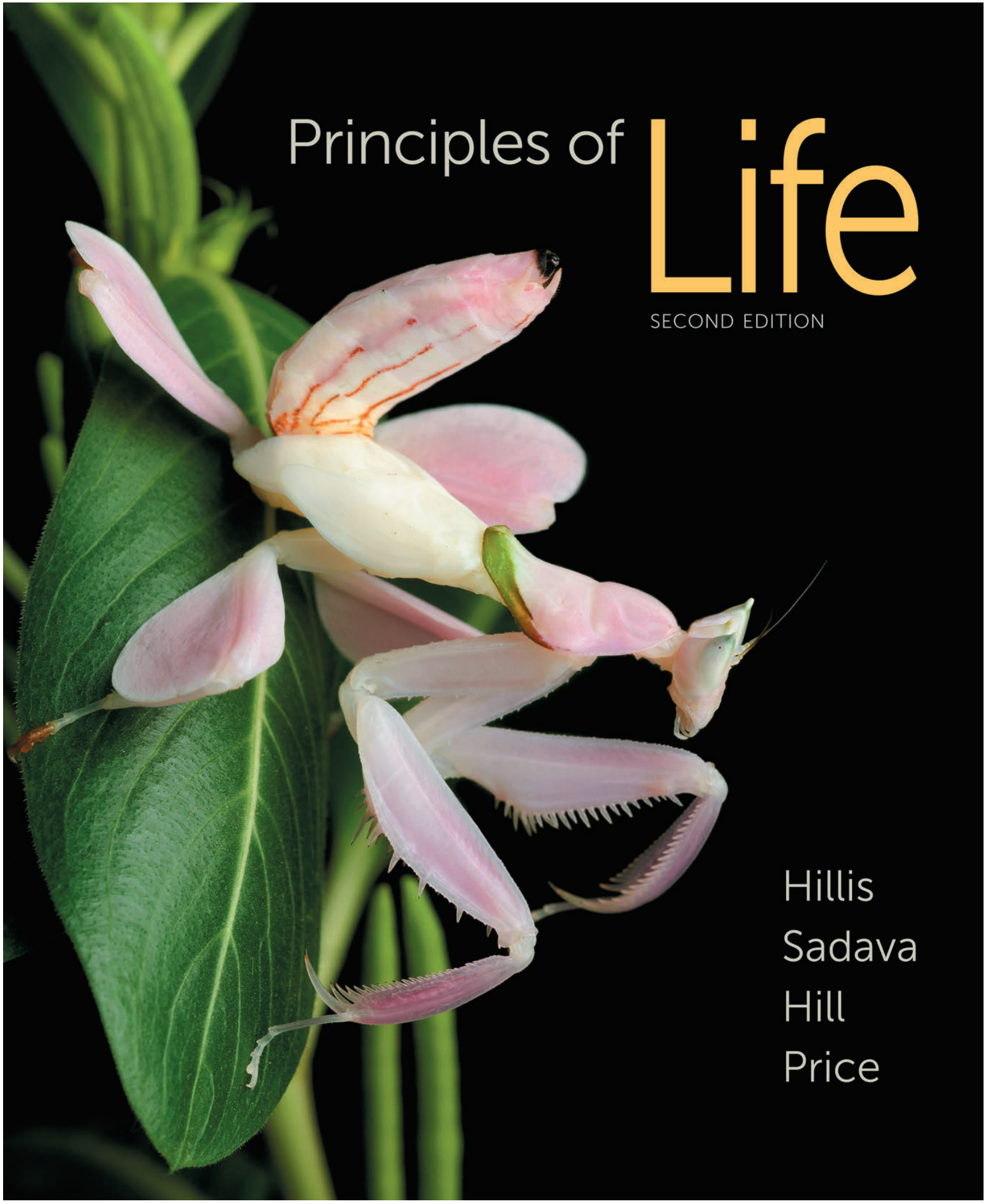


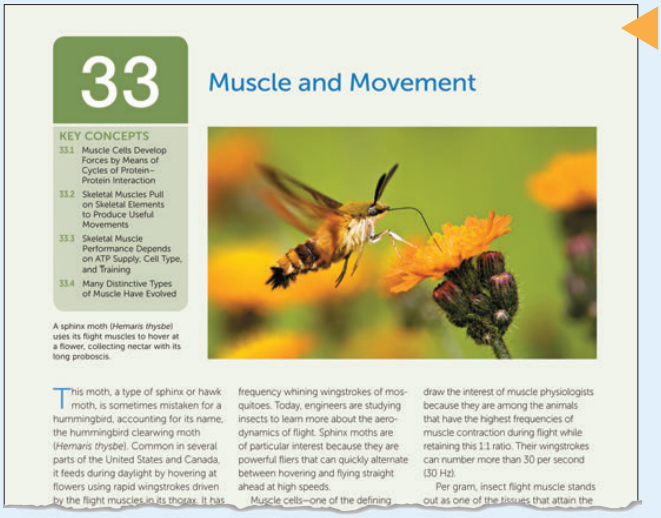
Principles of Life

SECOND EDITION

Hillis
Sadava
Hill
Price

How Principles of Life SECOND EDITION

Each chapter introduces essential biological concepts and the science that led to our understanding of them. Chapters are designed to help you focus on what's



33 Muscle and Movement

KEY CONCEPTS
33.1 Muscle Cells Develop Forces by Means of Cycles of Protein–Protein Interaction
33.2 Skeletal Muscles Pull on Skeletal Elements to Produce Useful Movements
33.3 Skeletal Muscle Performance Depends on ATP Supply, Cell Type, and Training
33.4 Many Distinctive Types of Muscle Have Evolved

A sphinx moth (*Hemaris thysbe*) uses its flight muscles to hover at a flower, collecting nectar with its long proboscis.

This moth, a type of sphinx or hawk moth, is sometimes mistaken for a hummingbird, accounting for its name, the hummingbird clearwing moth (*Hemaris thysbe*). Common in several parts of the United States and Canada, it feeds during daylight by hovering at flowers using rapid wingstrokes driven by the flight muscles in its thorax. It has frequency whining wingstrokes of mosquitoes. Today, engineers are studying insects to learn more about the aerodynamics of flight. Sphinx moths are of particular interest because they are powerful fliers that can quickly alternate between hovering and flying straight ahead at high speeds.

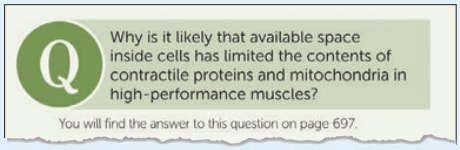
Muscle cells—one of the defining draw the interest of muscle physiologists because they are among the animals that have the highest frequencies of muscle contraction during flight while retaining this 1:1 ratio. Their wingstrokes can number more than 30 per second (30 Hz).

Per gram, insect flight muscle stands out as one of the tissues that attain the

OPENING STORY & QUESTION

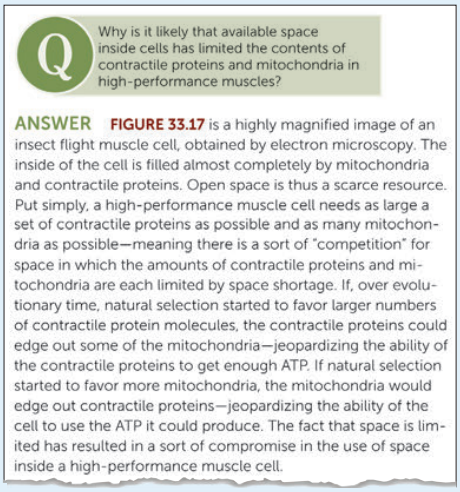
Chapters begin with an **OPENING STORY** designed to show you how the biology relates to historical, medical, or social issues. Each story ends with an intriguing question.

The **ANSWER** comes at the chapter's conclusion, with references to relevant information and illustrations in the chapter.



Q Why is it likely that available space inside cells has limited the contents of contractile proteins and mitochondria in high-performance muscles?

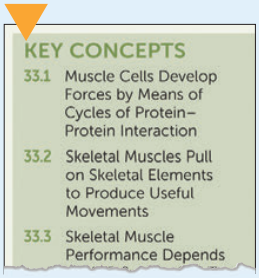
You will find the answer to this question on page 697.



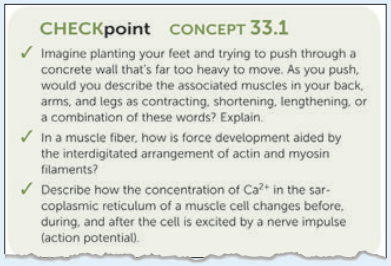
Q Why is it likely that available space inside cells has limited the contents of contractile proteins and mitochondria in high-performance muscles?

ANSWER **FIGURE 33.17** is a highly magnified image of an insect flight muscle cell, obtained by electron microscopy. The inside of the cell is filled almost completely by mitochondria and contractile proteins. Open space is thus a scarce resource. Put simply, a high-performance muscle cell needs as large a set of contractile proteins as possible and as many mitochondria as possible—meaning there is a sort of "competition" for space in which the amounts of contractile proteins and mitochondria are each limited by space shortage. If, over evolutionary time, natural selection started to favor larger numbers of contractile protein molecules, the contractile proteins could edge out some of the mitochondria—jeopardizing the ability of the contractile proteins to get enough ATP. If natural selection started to favor more mitochondria, the mitochondria would edge out contractile proteins—jeopardizing the ability of the cell to use the ATP it could produce. The fact that space is limited has resulted in a sort of compromise in the use of space inside a high-performance muscle cell.

KEY CONCEPTS & CHECKPOINTS



KEY CONCEPTS
33.1 Muscle Cells Develop Forces by Means of Cycles of Protein–Protein Interaction
33.2 Skeletal Muscles Pull on Skeletal Elements to Produce Useful Movements
33.3 Skeletal Muscle Performance Depends



CHECKpoint CONCEPT 33.1

✓ Imagine planting your feet and trying to push through a concrete wall that's far too heavy to move. As you push, would you describe the associated muscles in your back, arms, and legs as contracting, shortening, lengthening, or a combination of these words? Explain.

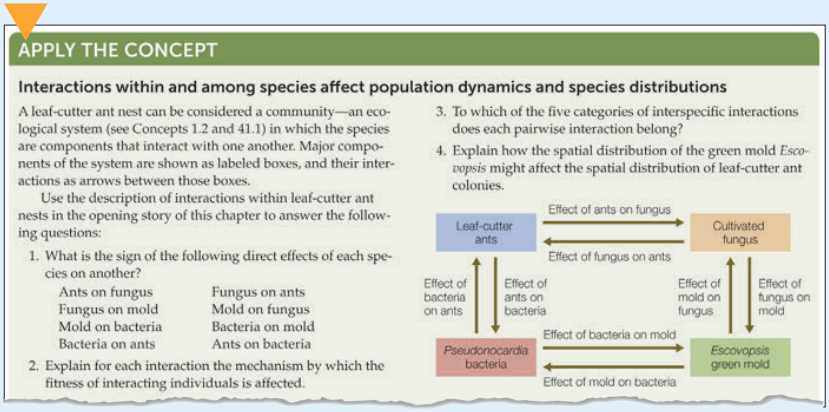
✓ In a muscle fiber, how is force development aided by the interdigitated arrangement of actin and myosin filaments?

✓ Describe how the concentration of Ca^{2+} in the sarcoplasmic reticulum of a muscle cell changes before, during, and after the cell is excited by a nerve impulse (action potential).

KEY CONCEPTS begin each chapter.

CHECKpoints revisit the Key Concepts at the end of each section.

APPLY THE CONCEPT



APPLY THE CONCEPT

Interactions within and among species affect population dynamics and species distributions

A leaf-cutter ant nest can be considered a community—an ecological system (see Concepts 1.2 and 41.1) in which the species are components that interact with one another. Major components of the system are shown as labeled boxes, and their interactions as arrows between those boxes.

Use the description of interactions within leaf-cutter ant nests in the opening story of this chapter to answer the following questions:

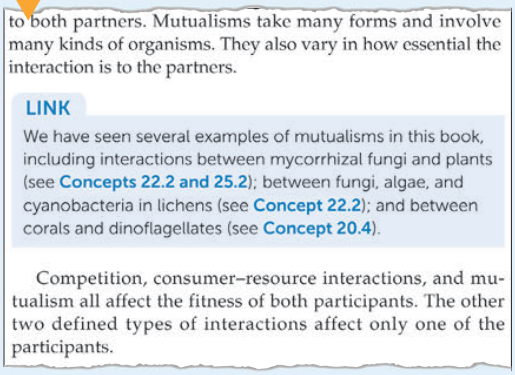
1. What is the sign of the following direct effects of each species on another?

Ants on fungus	Fungus on ants
Fungus on mold	Mold on fungus
Mold on bacteria	Bacteria on mold
Bacteria on ants	Ants on bacteria

2. Explain for each interaction the mechanism by which the fitness of interacting individuals is affected.
3. To which of the five categories of interspecific interactions does each pairwise interaction belong?
4. Explain how the spatial distribution of the green mold *Escovopsis* might affect the spatial distribution of leaf-cutter ant colonies.

APPLY THE CONCEPT exercises ask you to use a concept in a real-world setting to interpret actual research data and draw your own conclusions.

LINKS



to both partners. Mutualisms take many forms and involve many kinds of organisms. They also vary in how essential the interaction is to the partners.

LINK

We have seen several examples of mutualisms in this book, including interactions between mycorrhizal fungi and plants (see **Concepts 22.2 and 25.2**); between fungi, algae, and cyanobacteria in lichens (see **Concept 22.2**); and between corals and dinoflagellates (see **Concept 20.4**).

Competition, consumer–resource interactions, and mutualism all affect the fitness of both participants. The other two defined types of interactions affect only one of the participants.

In-text **LINKS** point you to additional discussion of a concept or key term elsewhere in the book.

Works for You...

important, and they offer a number of ways to analyze and review what you've read as you prepare for class or exams.

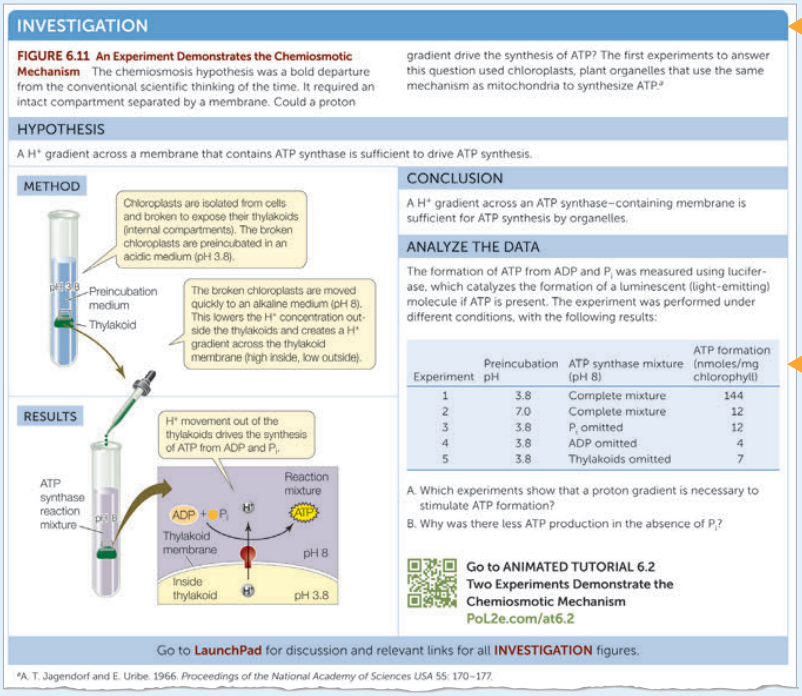

INVESTIGATION

FIGURE 6.11 An Experiment Demonstrates the Chemiosmotic Mechanism The chemiosmosis hypothesis was a bold departure from the conventional scientific thinking of the time. It required an intact compartment separated by a membrane. Could a proton gradient drive the synthesis of ATP? The first experiments to answer this question used chloroplasts, plant organelles that use the same mechanism as mitochondria to synthesize ATP.[a]

HYPOTHESIS

A H^+ gradient across a membrane that contains ATP synthase is sufficient to drive ATP synthesis.

METHOD

Chloroplasts are isolated from cells and broken to expose their thylakoids (internal compartments). The broken chloroplasts are preincubated in an acidic medium (pH 3.8).

Preincubation medium

Thylakoid

The broken chloroplasts are moved quickly to an alkaline medium (pH 8). This lowers the H^+ concentration outside the thylakoids and creates a H^+ gradient across the thylakoid membrane (high inside, low outside).

RESULTS

H^+ movement out of the thylakoids drives the synthesis of ATP from ADP and P_i.

ATP synthase reaction mixture

Reaction mixture

ADP + P_i

ATP

Thylakoid membrane

pH 8

Inside thylakoid

pH 3.8

CONCLUSION

A H^+ gradient across an ATP synthase–containing membrane is sufficient for ATP synthesis by organelles.

ANALYZE THE DATA

The formation of ATP from ADP and P_i was measured using luciferase, which catalyzes the formation of a luminescent (light-emitting) molecule if ATP is present. The experiment was performed under different conditions, with the following results:

Experiment	Preincubation pH	ATP synthase mixture (pH 8)	ATP formation (nmoles/mg chlorophyll)
1	3.8	Complete mixture	144
2	7.0	Complete mixture	12
3	3.8	P_i omitted	12
4	3.8	ADP omitted	4
5	3.8	Thylakoids omitted	7

A. Which experiments show that a proton gradient is necessary to stimulate ATP formation?

B. Why was there less ATP production in the absence of P_i?

Go to ANIMATED TUTORIAL 6.2
Two Experiments Demonstrate the Chemiosmotic Mechanism
PoL2e.com/at6.2

Go to LaunchPad for discussion and relevant links for all INVESTIGATION figures.

[a]A. T. Jagendorf and E. Uribe. 1966. *Proceedings of the National Academy of Sciences USA* 55: 170–177.

INVESTIGATION

INVESTIGATION figures emphasize the process of scientific inquiry to give you a realistic sense of how science is done. Each Investigation figure is organized in order of **hypothesis**, **method**, **results,** and **conclusion** and cites the original research paper(s).

ANALYZE THE DATA

Most **INVESTIGATION** figures are followed by **ANALYZE THE DATA** problems, which ask you to work with data from published biological research and make your own connections between observations, analyses, hypotheses, and conclusions.

INSTANT ACCESS CODES

QUICK RESPONSE (QR) CODES and **DIRECT WEB ADDRESSES** integrated into the text link you immediately to engaging animations, media clips, and activities. Just scan the code with your smartphone or tablet, or type the short Web address into any browser.
(Free QR reader apps are available from your mobile device's app store.)

Go to MEDIA CLIP 23.7
Octopuses Can Pass through Small Openings
PoL2e.com/mc23.7

HELPFUL ART WITH BALLOON CAPTIONS

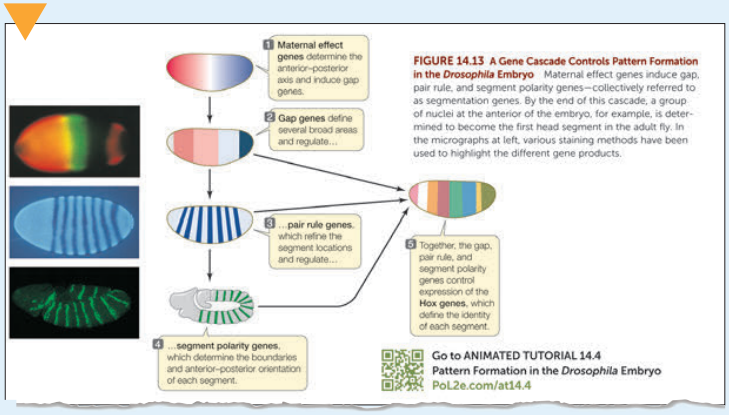


FIGURE 14.13 A Gene Cascade Controls Pattern Formation in the *Drosophila* Embryo Maternal effect genes induce gap, pair rule, and segment polarity genes—collectively referred to as segmentation genes. By the end of this cascade, a group of nuclei at the anterior of the embryo, for example, is determined to become the first head segment in the adult fly. In the micrographs at left, various staining methods have been used to highlight the different gene products.

Go to ANIMATED TUTORIAL 14.4
Pattern Formation in the *Drosophila* Embryo
PoL2e.com/at14.4

Numbered **BALLOON CAPTIONS** in the illustrations make it easy to follow key processes step by step.

CHAPTER SUMMARY

SUMMARY

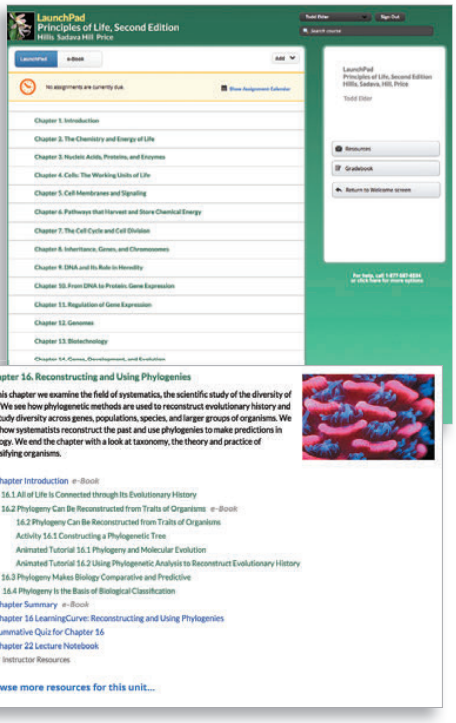
CONCEPT 19.1 **Life Consists of Three Domains That Share a Common Ancestor**

- Two of life's three **domains**, Bacteria and Archaea, are **prokaryotic**. They are distinguished from Eukarya in several ways, including their lack of a nucleus and of membrane-enclosed organelles. **Review Table 19.1**
- Eukaryotes are related to both Archaea and Bacteria and appear to have formed through endosymbiosis between members of these two lineages. The last common ancestor of all three domains probably lived about 3 billion years ago. **Review Figure 19.1 and ANIMATED TUTORIAL 19.1**

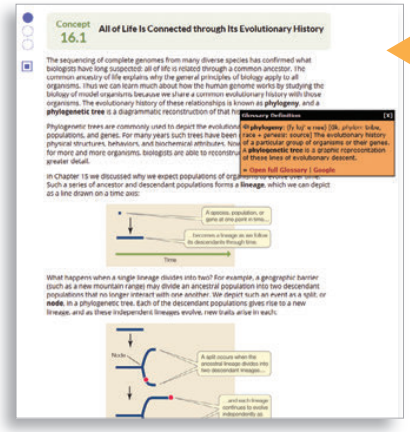
CHAPTER SUMMARIES provide a thorough review of chapter content, including key figures, and references to supporting online resources, including Animated Tutorials and Activities.

LaunchPad

MACMILLANHIGHERED.COM/LAUNCHPAD



LaunchPad is the easy-to-use online course space that offers rich content organized in a breakthrough user interface. LaunchPad's learning units combine animations, activities, and exercises with the e-Book version of the textbook to give you a complete set of resources for each chapter.

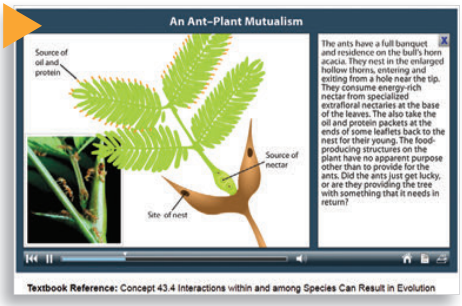


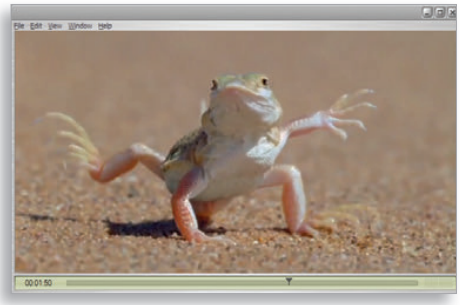
INTERACTIVE e-BOOK

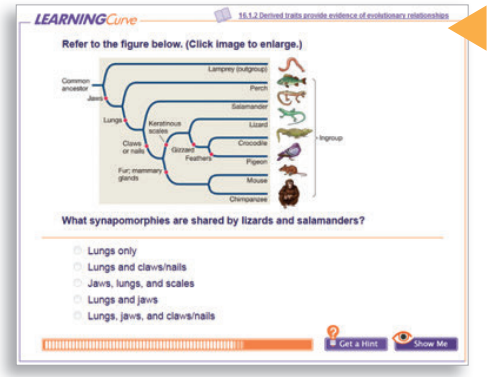
LaunchPad's interactive e-Book incorporates powerful study tools and integrated media resources. You can highlight text and add notes, click terms to get instant definitions and pronunciations (shown here as a pop-up), and review key animations and activities as you read.

MEDIA RESOURCES

LaunchPad features a wealth of study and review resources for each chapter of the textbook. These include Animated Tutorials, Activities, Media Clips, Flashcards and Key Terms, Investigation Links, Analyze the Data exercises, and more, all designed to help you learn and apply the material.







LEARNINGCURVE SELF-ASSESSMENT

LEARNINGCurve is a dynamic quizzing system that helps you learn the material in each chapter. By adapting to your responses, it presents questions at appropriate difficulty levels within each topic area and across the chapter. With its game-like format, direct links to the e-Book, and instant feedback, *LEARNINGCurve* is a great way to understand the concepts and master the content.

This textbook may have come packaged with a LaunchPad access code. If not, you can purchase access online at macmillanhighered.com/launchpad

ONLINE RESOURCES

The following online resources are referenced throughout the textbook, on the pages indicated. Each is available via the instant access code (QR code) and/or direct Web address that appears with the in-text reference. All of these resources (along with many additional resources) are also available within **LaunchPad** at macmillanhighered.com/launchpad.

MEDIA CLIPS

ANIMATED TUTORIALS

ONLINE RESOURCES continued

The following online resources are referenced throughout the textbook, on the pages indicated. Each is available via the instant access code (QR code) and/or direct Web address that appears with the in-text reference. All of these resources (along with many additional resources) are also available within **LaunchPad** at macmillanhighered.com/launchpad.

ACTIVITIES

ONLINE RESOURCES continued

The following online resources are referenced throughout the textbook, on the pages indicated. Each is available via the instant access code (QR code) and/or direct Web address that appears with the in-text reference. All of these resources (along with many additional resources) are also available within **LaunchPad** at macmillanhighered.com/launchpad.

Principles of Life

SECOND EDITION

Principles of Life

SECOND EDITION

David M. Hillis
University of Texas at Austin

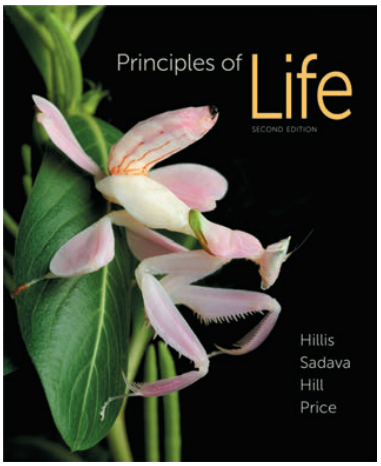
David Sadava
Emeritus, The Claremont Colleges

Richard W. Hill
Michigan State University

Mary V. Price
Emerita, University of California, Riverside

SINAUER ASSOCIATES

MACMILLAN



About the Cover

A juvenile pink orchid mantis (*Hymenopus coronatus*) looks, at first glance, like an orchid flower. Its abdomen, head, and four walking legs look like the petals of the flower, and the small black dot at the posterior tip of the abdomen resembles a small fly investigating the flower. This mimicry is advantageous to the mantid for two reasons. The mantis is concealed from potential predators, which mistake the mantis for a flower. At the same time, insects looking for nectar become prey for the mantis, which captures visiting insects with its front pair of toothed, grasping legs. As a result of these advantages, natural selection favored the evolution of this spectacular example of an insect that resembles an orchid flower. © Ch'ien Lee/Minden Pictures.

Principles of Life, Second Edition

Address editorial correspondence to:
Sinauer Associates, Inc., P.O. Box 407, Sunderland, MA 01375 U.S.A.

www.sinauer.com
publish@sinauer.com

Address orders to:
MPS / W.H. Freeman & Co., Order Dept., 16365 James Madison Highway,
U.S. Route 15, Gordonsville, VA 22942 U.S.A. or call 1-888-330-8477
Examination copy information: www.whfreeman.com/request or 1-800-446-8923

This SFI label applies to text and cover stocks.

Courier Corporation, the manufacturer
of this book, owns the *Green Edition* Trademark

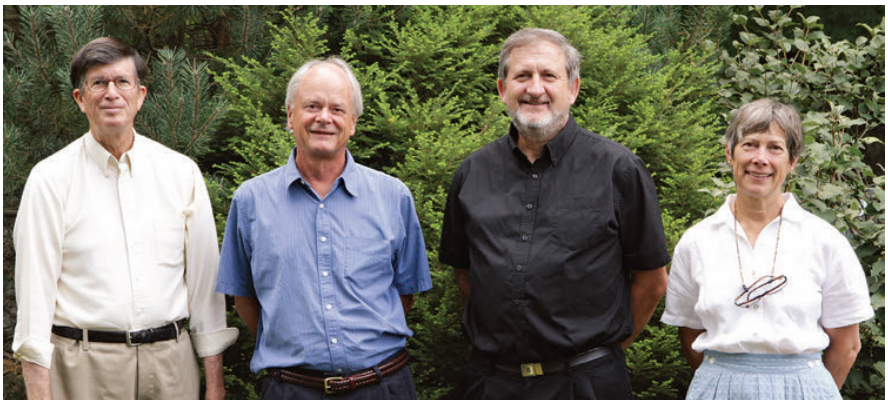
Library of Congress Cataloging-in-Publication Data
Hillis, David M.
Principles of life / David M. Hillis, University of Texas at Austin, David Sadava, Emeritus, The Claremont Colleges, Richard W. Hill, Michigan State University, Mary V. Price, Emerita, University of California, Riverside. -- Second edition.
pages cm
Includes index.
ISBN 978-1-4641-0947-8
1. Biology. I. Title.
QH308.2.P75 2013
570--dc23

2013036105

Printed in U.S.A.
First Printing November 2013
The Courier Companies, Inc.

To all our students. You have taught us, too, and inspired us to write this book.

The Authors



David Sadava, Richard W. Hill, David M. Hillis, and Mary V. Price

David M. Hillis is the Alfred W. Roark Centennial Professor in Integrative Biology at the University of Texas at Austin, where he also has directed the Center for Computational Biology and Bioinformatics and the School of Biological Sciences. Dr. Hillis has taught courses in introductory biology, genetics, evolution, systematics, and biodiversity. He has been elected to the National Academy of Sciences and the American Academy of Arts and Sciences, awarded a John D. and Catherine T. MacArthur Fellowship, and has served as President of the Society for the Study of Evolution and of the Society of Systematic Biologists. He served on the National Research Council committee that wrote the report BIO 2010: Transforming Undergraduate Biology Education for Research Biologists, and currently serves on the Executive Committee of the National Academies Scientific Teaching Alliance. His research interests span much of evolutionary biology, including experimental studies of evolving viruses, empirical studies of natural molecular evolution, applications of phylogenetics, analyses of biodiversity, and evolutionary modeling. He is particularly interested in teaching and research about the practical applications of evolutionary biology.

David Sadava is the Pritzker Family Foundation Professor of Biology, Emeritus, at the Keck Science Center of Claremont McKenna, Pitzer, and Scripps, three of The Claremont Colleges. In addition, he is Adjunct Professor of Cancer Cell Biology at the City of Hope Medical Center. Twice winner of the Huntoon Award for superior teaching, Dr. Sadava has taught courses in introductory biology, biotechnology, biochemistry, cell biology, molecular biology, plant biology, and cancer biology. In addition to *Life: The Science of Biology*, he is the author or coauthor of books on cell biology and on plants, genes, and crop biotechnology. His research has resulted in many papers coauthored with his students, on topics ranging from plant biochemistry to pharmacology of narcotic analgesics to human genetic diseases. For the past 15 years, he has investigated multi-drug resistance in human small-cell lung carcinoma cells with a view to understanding and overcoming this clinical challenge. At the City of Hope, his current work focuses on new anti-cancer agents from plants.

Richard W. Hill is Professor in the Department of Zoology at Michigan State University and a frequent Guest Investigator at Woods Hole Oceanographic Institution. He received his Ph.D. in Zoology from the University of Michigan. Apart from Sinauer Associates' editions of *Animal Physiology*, Dr. Hill has authored two other books on the subject (the second with Gordon Wyse), as well as numerous articles for scientific journals, encyclopedias, and edited volumes. Among the awards he has received are the Outstanding Faculty Award (Michigan State University Senior Class Council) and election as Fellow of the American Association for the Advancement of Science. He was a U.S. Senior Fulbright Scholar in 2000–2001. He has taught courses in environmental physiology and marine biology, and for the last five years has led student trips to the Galápagos Islands. His research interests include: temperature regulation and energetics in birds and mammals, especially neonates; and environmental physiology of marine tertiary sulfonium and quaternary ammonium compounds, especially in the contexts of biogeochemistry and animal–algal symbioses.

Mary V. Price is Professor of Biology, Emerita, at the University of California, Riverside and Adjunct Professor in the School of Natural Resources and the Environment at the University of Arizona. In "retirement," she continues to teach and study, having learned the joy and art of scientific discovery as an undergraduate student at Vassar College and doctoral student at the University of Arizona. Dr. Price has mentored and published with independent-research students and has developed and taught general biology and ecology courses from introductory (majors and nonmajors) to graduate levels. She has particularly enjoyed leading field classes in the arid regions of North America and Australia, and the tropical forests of Central America, Africa, and Madagascar. Dr. Price's research focuses on understanding the ecology of North American deserts and mountains. She has asked why so many desert rodents can coexist, how best to conserve endangered kangaroo rat species, how pollinators and herbivores influence floral evolution and plant population dynamics, and how climate change affects ecological systems.

Brief Table of Contents

Preface

Since the First Edition of *Principles of Life* was published, the surge in biology education research and the availability of resources for teachers continues to create excitement among the teaching community. Just as the First Edition appeared, the American Association for the Advancement of Science (supported by the National Science Foundation) published *Vision and Change in Undergraduate Biology Education: A Call to Action*. This report endorsed teaching core concepts and core competencies, and promoted the active involvement of students in problem-solving activities. When the First Edition of *Principles of Life* was published, it offered a radically new approach to teaching introductory biology that aligns with the goals put forward in the *Vision and Change* report (see below).

Principles of Life emphasizes mastering major concepts in biology through active learning, problem solving in realistic scenarios, and understanding rather than memorization. Now other textbook authors are beginning to follow our lead. We are proud that *Principles of Life* has helped to usher in this change in the way biology courses are taught, and we intend to continue to develop our book as the leading vehicle for this new approach to biology education.

Leading the Change in Undergraduate Biology Education

We are standing at an important crossroads in biology education, and many recent efforts have converged to produce an opportunity for lasting change in the way that instructors teach introductory biology. The validity of our approach in *Principles of Life* is supported by numerous reports and studies published by education agencies and national study groups since the turn of the millennium. In particular, two major reports have encouraged this change: the *Vision and Change* report mentioned above and *BIO2010: Transforming Undergraduate Education for Future Research Biologists*, sponsored by the National Institutes of Health and the Howard Hughes Medical Institute. These reports recommended focusing on core concepts and competencies, teaching students through active learning rather than memorization, and improving the integration of statistical and computational approaches. At about the same time, the College Board was redesigning the Advanced Placement Biology course with the same objectives. In *Principles of Life*, we have used our experience as authors and educators to implement these recommendations in a new approach to teaching introductory biology.

The *Vision and Change* report (2011) identified five "core concepts for biological literacy" that should be integrated throughout the curriculum. These core concepts center around the themes of:

- evolution
- the relationship between structure and function
- information flow, exchange, and storage
- pathways and transformations of energy and matter
- biological systems

In the Second Edition of *Principles of Life*, we have worked to ensure that these five core concepts are stressed and reinforced throughout the text, problems, media links, and other activities. To help students build bridges between different portions of the course and areas of knowledge, we have provided **Links** throughout the book. Using these *Links*, students can see, for example, that information they learn about molecular or cell biology is connected to topics in evolution, diversity, physiology, and ecology.

In addition to urging a focus on core concepts, the *Vision and Change* report argued that students need to cultivate certain core competencies to become successful scientists. Students should be able to:

- apply the process of science
- use quantitative reasoning
- use modeling and simulation
- tap into the interdisciplinary nature of science
- communicate and collaborate with other disciplines
- understand the relationship between science and society

Students are encouraged to practice these core competencies throughout the Second Edition of *Principles of Life*. Every chapter contains *Apply the Concept* exercises, which give students opportunities to practice working with data. These problems tie in with our *Making Sense of Data: A Statistics Primer* (Appendix B), which helps students understand why and how biologists draw conclusions from biological data, and thus helps them develop quantitative reasoning skills. We have also added more online *Animated Tutorials* and *Activities*, which include opportunities for students to use modeling and simulation modules to further reinforce their understanding of concepts. By engaging in these activities, students also learn about the importance of biological concepts and analyses for addressing societal issues and challenges.

The *Principles of Life* Story

Prior to our launch of the First Edition of *Principles of Life*, introductory biology textbooks for science majors presented encyclopedic summaries of biological knowledge. We believe that students who spend their time diligently memorizing myriad details and vast terminology actually retain fewer of the concepts that are the foundation for further study in advanced courses. In *Principles of Life*, we take the opposite approach: we promote understanding over memorization. Details are important, but no modern biology textbook can

begin to cover all the information biologists have learned to date, and students today have many other ways to access the details as they need them.

To help us create this new breed of biology textbook, in 2009 our publishers Sinauer Associates and W. H. Freeman brought together an Advisory Board of 20 leading biology educators and instructors in introductory biology from throughout North America. During an intensive meeting of the authors and the Board, dynamic discussions led to the solidification of the core concepts we believe are essential for teaching introductory biology. The book took shape, and the Advisory Board members then reviewed the emerging chapters and provided considerable feedback at every stage of the book's development. The result was a book that showcased the logical structure of scientific investigation, including lab, field, and computer modeling approaches. *Principles of Life* helped students apply the concepts they learn by providing opportunities for them to analyze original data in every chapter. In this and many other ways *Principles of Life* incorporated inquiry-based approaches that encourage active learning.

The First Edition of *Principles of Life* was widely adopted and well received. Adopters and reviewers praised the approach, and encouraged us to expand the effort to include even more problem-solving opportunities for students and more examples of the experiments that have formed the basis of our understanding. For the Second Edition, all chapters underwent extensive between-edition review by experts in each respective subdiscipline, and the chapters were revised accordingly. We now provide more references to original data and analyses so that students and instructors can easily explore the original experiments in greater depth. Moreover, we have expanded opportunities for students to apply what they have learned by using real data and examples, and have better integrated and explained the concepts of statistical analysis of data. We have included links to online videos (the new Media Clips) that help students to appreciate the relevance of what they have learned and to enjoy the excitement of biology.

How Is *Principles of Life* Different?

Each chapter of *Principles of Life* is organized into a series of **Concepts** that are important for mastering introductory biology. We have carefully chosen these concepts in light of feedback from our colleagues, from the Advisory Board, and from the numerous reports examining introductory biology. Concepts are elaborated upon, but not with the extensive detail found in most introductory texts. *Principles of Life* is focused; it is not meant to be encyclopedic.

Students learn concepts best when they apply them to practical problems. Each chapter of *Principles of Life* contains exercises, called **Apply the Concept**, that present data for students to analyze. Each of these exercises reinforces a concept that is central to that chapter. Science students need to understand basic methods for data presentation and analysis, so many of these problems ask students about statistical significance of the results. To help students understand issues in data presentation and interpretation, we have provided a short introduction to the reasoning behind biological statistics in Appendix B. Although this Appendix is not meant to replace a more formal introduction to statistics, we believe that statistical thinking is an important skill that should be developed in all introductory science courses. We have kept the problems and examples straightforward to emphasize the concepts of statistical analysis rather than the details of any particular statistical test. Some of the *Apply the Concept* exercises are simple enough that they can be presented, analyzed, and discussed in class; others are better suited for homework assignments.

Our **Investigation** figures let students see *how* we know *what* we know. These figures present a Hypothesis, Method, Results, and Conclusion. Most of these *Investigation* figures now include a section titled **Analyze the Data**, in which we have extracted a subset of data from the published experiment. Students are asked to analyze these data and to make connections between observations, analyses, hypotheses, and conclusions. As with *Apply the Concept* problems, students are asked to apply basic statistical approaches to understand the results and draw conclusions. We have also provided original references and extensive online resources for each *Investigation* figure. The online resources are available in LaunchPad, *Principles of Life*'s new online platform. These resources include expanded discussions of the original research, links to the original publications, and discussion and links for any follow-up investigations that have been published.

Each chapter begins with an application of a major concept—a story that illustrates and provides a motivation for understanding the chapter's content, and provides a social, medical, scientific, or historical context for the material. Each of these vignettes ends with an open-ended question that students can keep in mind as they read and study the rest of the chapter. We return to this opening question at the close of the chapter to show how information presented throughout the chapter illuminates the question and helps provide an answer. By pondering these questions as they read and study, students can begin to think like scientists.

At the end of each conceptual discussion we provide **Checkpoints** designed to help students self-evaluate their understanding of the material. These *Checkpoints* span the incremental levels of Bloom's Taxonomy of Cognitive Domains: factual knowledge, comprehension, application, analysis, synthesis, and evaluation.

Another important element for student success is reinforcement and application of concepts through online **Animated Tutorials**, **Activities**, and **Media Clips**. Each chapter contains instant access codes (in the form of both a direct URL and a Quick Response, or QR, code, a barcode students can scan with a smartphone or tablet) that allow students to quickly access these online resources while reading. For many concepts, students can conduct their own simulations, explore a concept in greater depth, and understand concepts through active discovery. Using the *Media Clips*, they can also watch videos that help explain concepts or introduce students to the wonders of biological diversity.

Students need to learn about some of the major **Research Tools** that are used in biology, including major laboratory, computational, and field methods. Our *Research Tools* figures

explain these tools and provide a context for how they are used by biologists.

Our art program for *Principles of Life* continues to build on our success from *Life: The Science of Biology*. We pioneered the use of balloon captions to help students understand and interpret the biological processes illustrated in figures without repeatedly going back and forth between a figure, its legend, and the text. These guides help students connect critical points of figures to the concepts that are developed in the text.

Media and Supplements

The Second Edition of *Principles of Life* features an expanded collection of online resources to support and reinforce the material covered in the textbook. In an effort to more closely link the printed book to the online resources, you will find references with instant access (QR) codes and direct Web addresses for all of the new Media Clips, Animated Tutorials, Activities, and Interactive Summaries throughout the book. These allow students to link instantly to these resources from any device—computer, smartphone, or tablet—while reading the book.

The new **LaunchPad** online platform integrates all of the student resources, instructor resources, the complete eBook, and all assessment tools within a streamlined interface that groups essential content into easily assignable learning units. LaunchPad features a range of assessment tools including the new **LearningCurve** adaptive quizzing engine, and pre-built summative quizzes for each chapter. To support course preparation, classroom sessions, and assessment programs, there is a wide range of instructor resources available, including multiple versions of all textbook figures, a wealth of PowerPoint resources, multiple banks of assessment questions, a large collection of videos, and in-class active learning exercises.

For a complete list of all the media and supplements available for *Principles of Life*, please refer to "Media and Supplements to accompany *Principles of Life*" following this Preface. Also, please refer to the inside front cover for a full list of the student media resources referenced in the text.

Special Contributions

Many people contributed to the creation of the Second Edition of *Principles of Life* (see below). However, two individuals deserve special mention for their contributions. Susan D. Hill did a masterful job in writing Chapter 38 on Animal Development. Nickolas Waser worked extensively with Mary Price on the Ecology section (Part 7), and was otherwise intimately involved in discussions of the book's planning and execution.

Many People to Thank

In addition to the many biologists listed on the next page who provided formal reviews, each of us benefitted enormously from personal contacts with colleagues who helped us resolve issues and made critical suggestions for new material. They are: Walter Arnold, University of Veterinary Medicine (Vienna); Harry Greene, Cornell University; Will Petry, University of California, Irvine; David Sleboda, Brown University; Thomas Ruf, University of Veterinary Medicine (Vienna); Andrew Zanella, The Claremont Colleges; Edward McCabe, University of Colorado and the March of Dimes Foundation; and Frank Price, Utica College.

Our editor and publisher, Andy Sinauer, embraced the need for change in introductory biology textbooks and has helped make our vision into a reality. Bill Purves, Gordon Orians, and Craig Heller, our co-authors on earlier editions of *Life: The Science of Biology* and/or *Principles of Life*, were instrumental in articulating the concepts developed in this Second Edition of *Principles of Life*, and many aspects of this book can be traced back to their critical contributions.

For this new Edition, Sinauer Associates assembled a talented duo, Laura Green and Danna Niedzwiecki, who coordinated the editorial team and did much of the developmental editing. Annie Reid and Carol Pritchard-Martinez worked to ensure that the level and terminology are appropriate for beginning undergraduate students. Jane Murfett also contributed to developmental editing. Laura and Danna worked closely with a top-notch copyeditor, Liz Pierson. Carol Wigg was the principle production editor on previous editions of *Principles of Life* and *Life: The Science of Biology* and her mark endures. Elizabeth Morales, our artist, again worked with each of us to create effective and beautiful line art. She also revised many figures to make them more effective for people with common forms of color blindness. David McIntyre again rose to the challenge of finding new, even better photographs. Designer Joan Gemme brought a fresh look to the book and did a fine job of assembling all of the book's elements into clear and attractive pages. Chris Small coordinated production and imposed his exacting standards on keeping the myriad components consistent. Johannah Walkowicz organized and commissioned the many expert academic reviews. Jason Dirks coordinated the team that created the vast array of online media and supplements. Dean Scudder, Director of Sales and Marketing, and Azelie Fortier, Biology Acquisitions Editor, participated in every stage of the book's development.

At W. H. Freeman, we continue to benefit from the long-term input of Biology Publisher Susan Winslow. John Britch, Director of Marketing, in collaboration with the Regional Specialists, Regional Sales Managers, and the Market Development team, coordinated all the stages of informing Freeman's skilled sales force of our book's story. We also wish to thank the Freeman media group for their expertise in producing LaunchPad.

David M. Hillis

David Sadava

Richard W. Hill

Mary V. Price

Advisors and Reviewers

Between-editions Reviewers

Rebecca Achterman, Western Washington University
William Anderson, Harvard University
Felicitas Avendano, Grand View University
Dave Bailey, St. Norbert College
Annalisa Berta, San Diego State University
Lori Boies, St. Mary's University
Nicole Bournias-Vardiabasis, California State University, Santa Bernardino
Elizabeth Braker, Occidental College
Cheryl Burrell, Wake Forest University
Robert Cabin, Brevard College
David Carroll, Florida Institute of Technology
Karen Champ, College of Central Florida
Shelton Charles, Forsyth Technical Community College
Sixue Chen, University of Florida
Doug Darnowski, Indiana University Southeast
Maria Davis, Olivet College
Weston Dulaney, Nashville State Community College
Laurie Eberhardt, Valparaiso University
Ingeborg Eley, Hudson Valley Community College
Lewis Feldman, University of California, Berkeley
Austin Francis, Jr., Armstrong Atlantic State University
Jason Garvon, Lake Superior State University
Patricia Geppert, The University of Texas at San Antonio
Nicole Gerlanc, Frederick Community College
Arundhati Ghosh, University of Pittsburgh
Suzanne Gollery, Sierra Nevada College
Elizabeth Good, University of Illinois, Urbana-Champaign
Eileen Gregory, Rollins College
Susan Hester, The University of Arizona
Samantha Hilber, University of Florida
Jill Holliday, University of Florida
Brian Hurdle, Frederick Community College
Sanjeeda Jafar, Morgan State University
Jerald Johnson, Brigham Young University
Susan Jorstad, The University of Arizona
Erin Keen-Rhineheart, Susquehanna University
Stephen Kilpatrick, University of Pittsburgh
Michael Koban, Morgan State University
Daniel Kueh, Emory University
Kristen Lennon, Frostburg State University
Jun Liu, Cornell University
James Malcolm, The University of Redlands
Steve Marcus, The University of Alabama
Nilo Marin, Broward College
Erin Martin, University of Portland
Robert Minckley, University of Rochester
Alexey Nikitin, Grand Valley State University
David Oppenheimer, University of Florida
Aswini Pai, St. Lawrence University
Crima Pogge, City College of San Francisco
Mary Preuss, Webster University
Frank Price, Hamilton College
David Puthoff, Frostburg State University
Sabiha Rahman, University of Ottawa
Nancy Rice, Western Kentucky University
Lori Rose, Hill College
Paul Schulte, University of Nevada, Las Vegas
Jon Shenker, Florida Institute of Technology
Cara Shillington, Eastern Michigan
Miles Silman, Wake Forest University
Thomas Terleph, Sacred Heart University
Mark Thogerson, Grand Valley State University
Elethia Tillman, Spelman College
Mike Troyan, Pennsylvania State University
Ximena Valderrama, Ramapo College of New Jersey
Sebastian Vélez, Worcester State University
Alexander Wait, Missouri State University
Katherine Warpeha, University of Illinois, Chicago
Ted Weinert, The University of Arizona
Erika Whitney, Morgan State University
Michelle Wien, Bryn Mawr College

Accuracy Reviewers

Brian Antonsen, Marshall University
Michael Baltzley, Western Oregon University
Alexa Bely, University of Maryland
Smriti Bhotika, University of Florida
Greta Bolin, Tarrant County College
Katherine Boss-Williams, Emory University
Michelle Boone, Miami University of Ohio
Jackie Brittingham, Simpson College
W. Randy Brooks, Florida Atlantic University
Justin Brown, Simpson College
Deborah Cardenas, Collin College
Trevor Caughlin, University of Florida
Shelton Charles, Forsyth Technical Community College
Donna Charley-Johnson, University of Wisconsin, Oshkosh
Kelly Jean Craig, Colorado Mesa University
Kathryn Craven, Armstrong Atlantic State University
Noelle Cutter, Molloy College
Nancy Dalman, University of North Georgia
Doug Darnowski, Indiana University Southeast
Sandra L. Davis, University of Indianapolis
Ann Marie Davison, Kwantlen Polytechnic University
James Demastes, University of Northern Iowa
Jesse Dillon, California State University, Long Beach
Wes Dulaney, Nashville State Community College

Michele DuRand, Memorial University
Todd Egan, Elmira College
Gordon Fain, The University of California, Los Angeles
Victor Fet, Marshall University
Jeffrey Firestone, University of Maryland
Michael Fultz, Morehead State University
Mark Garcia, Collin College
Deborah Gelman, Pace University
Emily Gillespie, Marshall University
Christopher Gilson, Emory University
Richard Gonzalez-Diaz, Seminole State College of Florida
Harry Greene, Cornell University
Shala Hankison, Ohio Wesleyan University
Amy Hark, Muhlenberg College
Phillip Harris, The University of Alabama
Julie Havens, Armstrong Atlantic State University
Marshal Hedin, San Diego State University
Anna Hicks, Memorial University
Nicole Huber, University of Colorado, Colorado Springs
Amanda Hyde, Greenfield Community College
Erika Iyengar, Muhlenberg College
Laura Jackson, University of Northern Iowa
Jerald Johnson, Brigham Young University
Carly Jordan, The George Washington University
Matthew Kayser, University of Pennsylvania Medical School
Kim Lackey, University of Alabama, Tuscaloosa
Marc Lajeunesse, University of South Florida
Evan Lampert, University of North Georgia
Charlie Lawrence, Viterbo University
Eric Liebgold, Salisbury University
Clark Lindgren, Grinnell College
Jessica Lucas, Santa Clara University
Lisa Lyons, Florida State University
C. Smoot Major, University of Southern Alabama
Marlee Marsh, Columbia College
Megan McNulty, The University of Chicago
Kyle McQuade, Colorado Mesa University
Susan T. Meiers, Western Illinois University
Tamra Mendelson, University of Maryland, Baltimore County
Anthony Metcalf, California State University, San Bernadino
Clara Moore, Franklin and Marshall College
Jeanelle Morgan, University of North Georgia
Jacqueline Bee Nesbit, University of New Orleans
Marcia Newcomer, Louisiana State University
Richard Niesenbaum, Muhlenberg College
John Ophus, University of Northern Iowa
Rebecca Orr, Collin College
Cassandra Osborne, Western State Colorado University
Andrew Palmer, Florida Institute of Technology
Karla Passalacqua, Emory University
Lisa Petrella, Marquette University
Eric Ribbens, Western Illinois University
George Robinson, University at Albany, State University of New York
Nicole M. Roy, Sacred Heart University
Aaron Schrey, Armstrong Atlantic State University
James Schulte, Clarkson University
Sarah Shannon, Indiana University
Frederick Sheldon, Louisiana State University
Bin Shuai, Wichita State University
Nancy Solomon, Miami University of Ohio
Sarah Spinette, Rhode Island College
Robert Sterner, University of Minnesota, Twin Cities
Oliver Sturm, Rhodes College
Timothy Sullivan, Armstrong Atlantic State University
Thomas Terleph, Sacred Heart University
Dustin Vale-Cruz, Springfield College
Kenneth van Golen, University of Delaware
Christine Weilhoefer, University of Portland
Robert Wise, University of Wisconsin, Oshkosh
MaryJo A. Witz, Monroe Community College
John Yoder, University of Alabama, Tuscaloosa
Ximena Valderrrama, Ramapo College of New Jersey

Media and Supplements Authors and Reviewers

Juliann Aukema, University of California, Santa Barbara
Edward Awad, Vanier College
Roberta Batorsky, Rowan University
Alexa Bely, University of Maryland
Joe Bruseo, Holyoke Community College
Shannon Compton, University of Massachusetts, Amherst
Douglas Darnowski, Indiana University Southeast
Ann Marie Davison, Kwantlen Polytechnic University
Donna Francis, University of Massachusetts, Amherst
Emily Gillespie, Marshall University
Richard Gonzalez-Diaz, Seminole State University
Carol Hand, science writer
Phillip Harris, University of Alabama
Laurel Hester, Keuka College
Peggy Hill, science writer
Jerry Johnson, Brigham Young University
Norman Johnson, University of Massachusetts, Amherst
Stephen Kilpatrick, University of Pittsburgh
Kim Lackey, University of Alabama
Marlee Marsh, Columbia College
Betty McGuire, Cornell University
Stephanie Moeckel-Cole, University of Massachusetts, Amherst
Paul Nolan, The Citadel
Kristine Nowak, Kennesaw State University
Don O'Malley, Northeastern University
Lesley Evans Ogden, science writer
Sherry Ogg, Johns Hopkins University
Sarah Shannon, Indiana University
Leo Shapiro, University of Maryland
Molly Solanki, science writer
Nancy Solomon, Miami University
Meredith Stafford, Johns Hopkins University
John Townsend-Mehler, Montana State University
Mary Tyler, University of Maine
Robert Wise, University of Wisconsin, Oshkosh

LaunchPad

for *Principles of Life*, SECOND EDITION

LaunchPad (**macmillanhighered.com/launchpad**) is the easy-to-use online platform that integrates the e-Book, all of the student and instructor media resources, and assessment functions into a unified interface. In addition to a wealth of course management, communication, organization, and gradebook features, LaunchPad includes the following resources.

e-Book

A complete online version of the textbook, the e-Book is fully integrated into LaunchPad and includes media resources, in-text links to all glossary entries (with audio pronunciations), and flexible notes and highlighting features. In addition, instructors can easily hide chapters or sections that they don't cover in their course, rearrange the order of chapters and sections, and add their own content directly into the e-Book.

LEARNINGCurve

LearningCurve is a powerful adaptive quizzing system with a game-like format to engage students. Rather than simply answering a fixed set of questions, students answer dynamically selected questions to progress toward a target level of understanding. At any point, students can view a report (with links to e-Book sections and media resources) of how well they are performing in each topic, to help them focus on problem areas.

Student Resources

INTERACTIVE SUMMARIES. For each chapter, these dynamic summaries combine a review of important concepts with links to all of the key figures, Activities, and Animated Tutorials.

ANIMATED TUTORIALS. In-depth animations and simulations present complex topics in a clear, easy-to-follow format that includes a brief quiz.

MEDIA CLIPS. New for the Second Edition, these short, engaging video clips depict fascinating examples of some of the many organisms, processes, and phenomena discussed in the textbook.

ACTIVITIES. A range of interactive activities helps students learn and review key facts and concepts through labeling diagrams, identifying steps in processes, and matching concepts.

LECTURE NOTEBOOK. Available online as PDF files, the Lecture Notebook includes all of the textbook's figures and tables, with space for note-taking.

BIONEWS FROM *SCIENTIFIC AMERICAN*. BioNews makes it easy for instructors to bring the dynamic nature of the biological sciences and up-to-the-minute currency into their course via an automatically updated news feed.

BIONAVIGATOR. A unique visual way to explore all of the Animated Tutorials and Activities across the various levels of biological inquiry—from the global scale down to the molecular scale.

ANALYZE THE DATA. Online versions of the Analyze the Data exercises included in many of the Investigation figures in the textbook.

FLASHCARDS AND KEY TERMS. An ideal way for students to learn and review the extensive terminology of introductory biology, featuring a review mode and a quiz mode.

INVESTIGATION LINKS. An overview of the experiments featured in each of the textbook's Investigation figures, with links to the original paper(s), related research or applications that followed, and additional information related to the experiment.

GLOSSARY. The full glossary, with audio pronunciations for all terms.

TREE OF LIFE. An interactive version of the Tree of Life from Appendix A, with links to a wealth of information on each group listed.

MATH FOR LIFE. A collection of mathematical shortcuts and references to help students with the quantitative skills they need in the biology laboratory.

SURVIVAL SKILLS. A guide to more effective study habits, including time management, note-taking, highlighting, and exam preparation.

Assessment Resources

SUMMATIVE QUIZZES. The pre-built summative quizzes assess overall student understanding of each chapter, and provide instructors with data on class and individual-student comprehension of chapter material.

LEARNINGCURVE. Reports provide instructors with instant information on student performance, broken down by individual section.

TEST BANK. The *Principles of Life* Test Bank includes over 5,000 questions, all referenced to specific textbook sections and categorized according to Bloom's taxonomy. Each chapter includes a wide range of multiple-choice, fill-in-the-blank, and short-answer questions, including diagram questions.

COMPREHENSIVE QUESTION BANKS include questions from the Test Bank, LearningCurve, Summative Quizzes, and the Study Guide. Question filtering allows instructors to select questions based on Bloom's taxonomy, textbook section, and question type, in order to easily select the desired mix of question types.

CUSTOM ASSESSMENT TOOLS allow instructors to create quizzes and many other types of assignments using any combination of publisher-provided questions and those created by the instructor.

Instructor Media and Course Planning Resources

(*Also available on the stand-alone Instructor's Media Library disc, ISBN 978-1-4641-8468-0.*)

TEXTBOOK FIGURES AND TABLES. Every figure and table from the textbook (including all photos and all unnumbered figures) is provided in both JPEG (high- and low-resolution) and PDF formats, in multiple versions, including whole, reformatted, and unlabeled.

POWERPOINT RESOURCES. For each chapter of the textbook, many different PowerPoint presentations are available, providing instructors the flexibility to build presentations in the manner that best suits their needs, including the following:

- Textbook Figures and Tables
- Lecture Presentations
- Figures with Editable Labels
- Layered Art Figures
- Supplemental Photos
- Active Learning Exercises

SUPPLEMENTAL PHOTOS. The supplemental photograph collection contains over 1,500 photographs, giving instructors a wealth of additional imagery to draw upon.

ANIMATIONS. An extensive collection of detailed animations and simulations.

VIDEOS. Featuring many new videos for the Second Edition, the wide-ranging collection of video segments helps demonstrate the complexity and beauty of life.

ACTIVE LEARNING EXERCISES. Set up for easy integration into lectures, each exercise poses a question or problem for the class to discuss or solve. Each also includes a multiple-choice element, for easy use with clicker systems.

ANSWERS TO TEXTBOOK QUESTIONS. Complete answers to all of the textbook's CheckPoint, Apply the Concept, and Analyze the Data questions.

INSTRUCTOR'S MANUAL. A wealth of course planning material, the Instructor's Manual includes the following sections for each chapter of the textbook: *Chapter Overview, What's New, Key Concepts & Learning Objectives, Chapter Outline, Lecture Outline*, and *Key Terms*.

MEDIA GUIDE. A visual guide to the extensive media resources available with *Principles of Life*.

Additional Instructor Resources

Computerized Test Bank

(CD, ISBN 978-1-4641-8467-3)

The entire Test Bank, plus the Summative Quizzes, LearningCurve questions, and Study Guide questions are all included in Blackboard's easy-to-use Diploma program (software included). Designed for both novices and advanced users, Diploma allows instructors to quickly and easily create or edit questions, create quizzes or exams with a "drag-and-drop" feature (using any combination of publisher-provided and instructor-added questions), publish to online courses, and print paper-based assessments.

Course Management System Support

As a service for *Principles of Life* adopters using Blackboard, Moodle, Desire2Learn, or other course management systems, full electronic course packs are available.

i>clicker

Developed for educators by educators, i>clicker is a hassle-free radio-frequency classroom response system that makes it easy for instructors to ask questions, record responses, take attendance, and direct students through lectures as active participants. For more information, visit **www.iclicker.com**.

LabPartner is a site designed to facilitate the creation of customized lab manuals. Its database contains a wide selection of experiments published by W. H. Freeman and Hayden-McNeil Publishing. Instructors can preview, choose, and reorder labs, interleave their own original experiments, add carbonless graph paper and a pocket folder, customize the cover both inside and out, and select a binding type. Manuals are printed on demand.
www.whfreeman.com/labpartner

The Scientific Teaching Book Series is a collection of practical guides, intended for all science, technology, engineering and mathematics (STEM) faculty who teach undergraduate and graduate students in these disciplines. The purpose of these books is to help faculty become more successful in all aspects of teaching and learning science, including classroom instruction, mentoring students, and professional development. Authored by well-known science educators, the Series provides concise descriptions of best practices and how to implement them in the classroom, the laboratory, or the department. For readers interested in the research results on which these best practices are based, the books also provide a gateway to the key educational literature. The Series includes:

Scientific Teaching
Jo Handelsman, Sarah Miller, and Christine Pfund
(ISBN 978-1-4292-0188-9)

Transformations: Approaches to College Science Teaching
Deborah Allen and Kimberly Tanner
(ISBN 978-1-4292-5335-2)

Entering Research: A Facilitator's Manual
Janet L. Branchaw, Christine Pfund, and Raelyn Rediske
(ISBN 978-1-4292-5857-9)

Discipline-Based Science Education Research: A Scientist's Guide
Stephanie J. Slater, Timothy F. Slater, and Janelle M. Bailey
(ISBN 978-1-4292-6586-7)

Assessment in the College Classroom
Clarissa Dirks, Mary Pat Wenderoth, and Michelle Withers
(ISBN 978-1-4292-8197-3)

Additional Student Resources

Principles of Life *Study Guide*
(Paper, ISBN 978-1-4641-8475-8)

The *Principles of Life* Study Guide offers a variety of study and review resources to accompany each chapter of the textbook. The opening Big Picture section gives students a concise overview of the main concepts covered in the chapter. The Study Strategies section points out common problem areas that students may find more challenging, and suggests strategies for learning the material most effectively. The Key Concept Review section combines a detailed review of each section with questions that help students synthesize and apply what they have learned, including diagram questions, short-answer questions, and more open-ended questions. Each chapter concludes with a Test Yourself section that allows students to test their comprehension. All questions include answers, explanations, and references to textbook sections.

Principles of Life *Flashcards App*
Available for iPhone/iPad and Android, the Flashcards App is a great way for students to learn and review all the key terminology from the textbook, whenever and wherever they want to study, in an intuitive flashcard interface.

CatchUp Math & Stats for the Life Sciences
Michael Harris, Gordon Taylor, and Jacquelyn Taylor
(ISBN 978-1-4292-0557-3)

Presented in brief, accessible units, this primer will help students quickly brush up on the quantitative skills they need to succeed in biology.

A Student Handbook for Writing in Biology, Fourth Edition
Karen Knisely (ISBN 978-1-4641-5076-0)

This book provides practical advice to students who are learning to write according to the conventions in biology, using the standards of journal publication as a model.

Bioethics and the New Embryology: Springboards for Debate
Scott F. Gilbert, Anna L. Tyler, and Emily J. Zackin
(ISBN 978-0-7167-7345-0)

Our ability to alter the course of human development ranks among the most significant changes in modern science and has brought embryology into the public domain. The question that must be asked is: Even if we can do such things, should we?

BioStats Basics: A Student Handbook
James L. Gould and Grant F. Gould
(ISBN 978-0-7167-3416-1)

Engaging and informal, *BioStats Basics* provides introductory-level biology students with a practical, accessible introduction to statistical methodology.

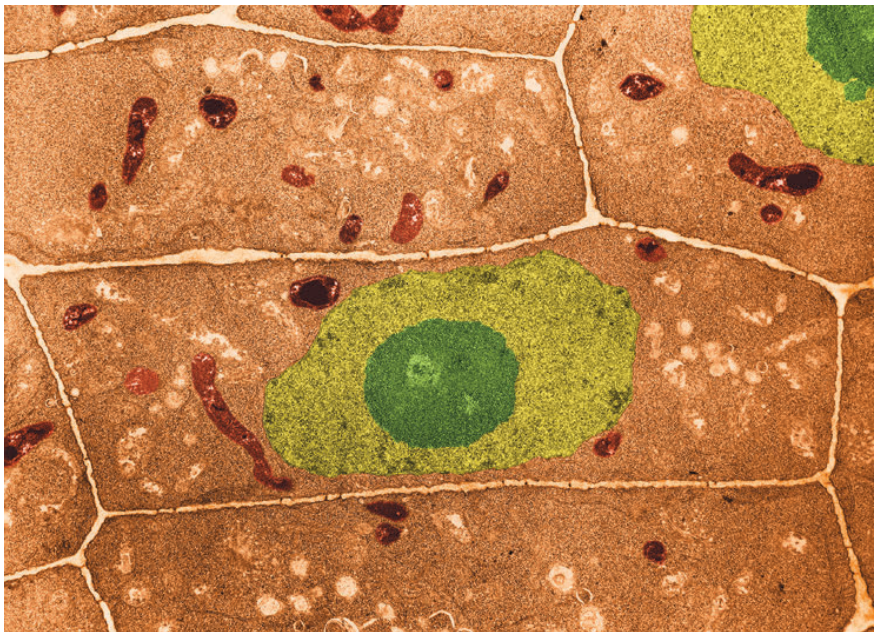
Inquiry Biology: A Laboratory Manual, Volumes 1 and 2
Mary Tyler, Ryan W. Cowan, and Jennifer L. Lockhart
(Volume 1 ISBN 978-1-4292-9288-7; Volume 2 ISBN 978-1-4292-9289-4)

This introductory biology laboratory manual is inquiry-based—instructing in the process of science by allowing students to ask their own questions, gather background information, formulate hypotheses, design and carry out experiments, collect and analyze data, and formulate conclusions.

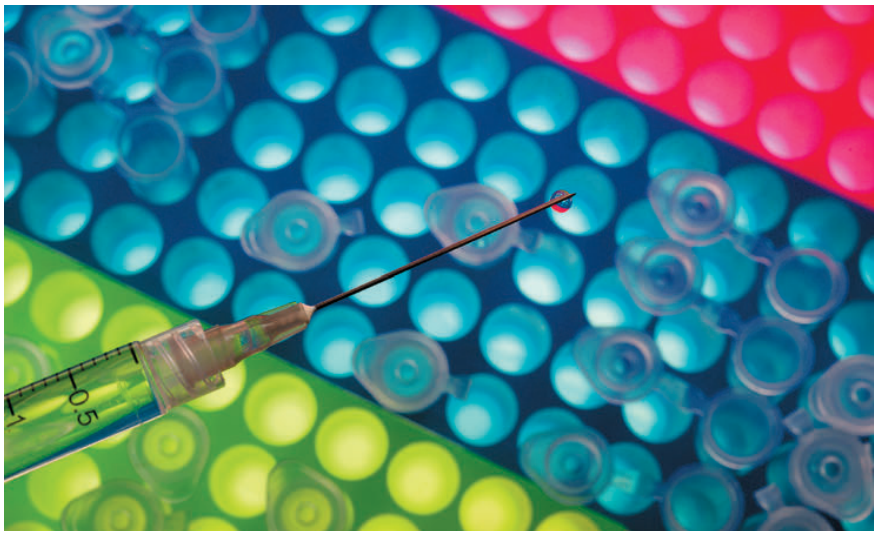
Hayden-McNeil Life Sciences Lab Notebook
(ISBN 978-1-4292-3055-1)

This durable, high-quality carbonless laboratory notebook allows students to hand in originals *or* copies of their lab work, not entire composition books. The Lab Notebook contains Hayden-McNeil's unique white paper carbonless copies and biology-specific reference materials.

Table of Contents

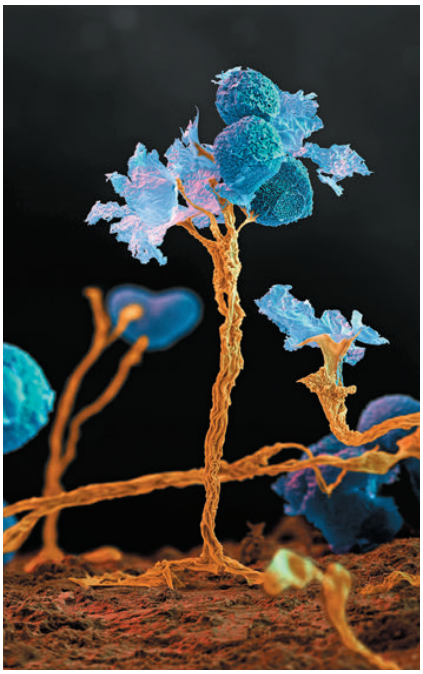


PART 2 Genetics



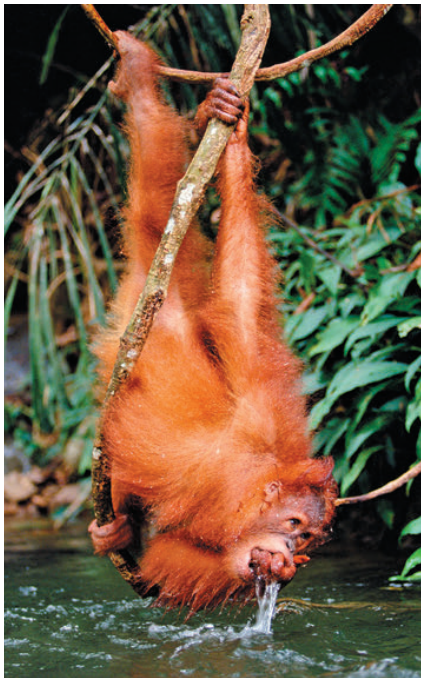
PART 3 Evolution

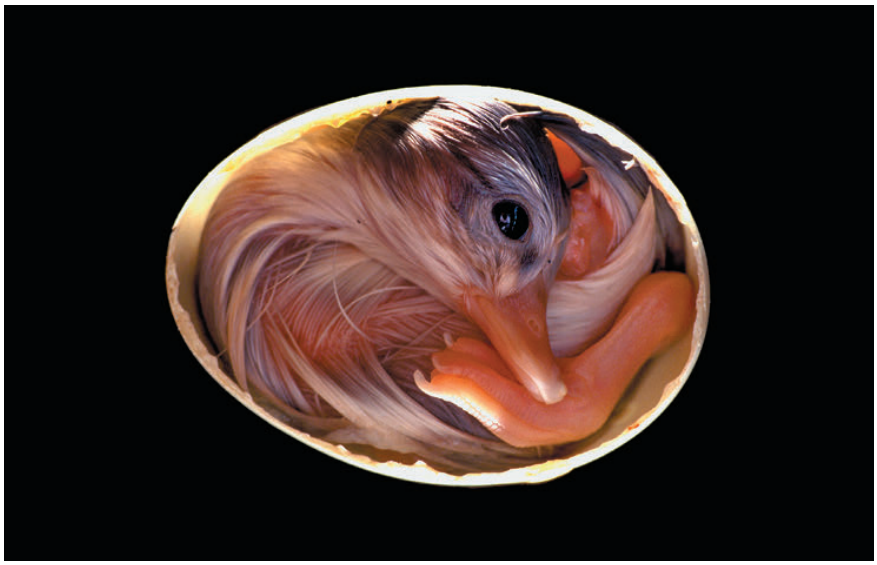
PART 4 Diversity



PART 5 Plant Form and Function

PART 6 Animal Form and Function





1 Principles of Life

KEY CONCEPTS

What principles of life are illustrated in this scene?

When you take a walk through the woods and fields or a park near your home, what do you see? Like most people, you probably notice the trees, colorful flowers, and some animals. But do you spend more than a little time thinking about how these living things survive, reproduce, interact with one another, or affect their environment? With the introduction to biology in this book, we would like to inspire you to ask questions about what life is, how living systems work, and how the living world came to be as we observe it today.

Biologists have amassed a huge amount of information about the living world, and some introductory biology classes focus on memorizing these details. In this book we take a different approach, focusing on the major principles of life that underlie everything in biology.

What do we mean by "principles of life"? Look at the photograph. Why is the view so overwhelmingly green? A fundamental principle of life, namely that all living organisms require energy to grow, move, reproduce, and maintain their bodies, can explain the color. Ultimately, most of that energy comes from the sun. The leaves of plants contain chlorophyll, a green pigment that captures energy from the sun and uses it to transform water and carbon dioxide into sugar and oxygen (in the process called photosynthesis). That sugar stores some of the energy from the sun in its chemical bonds. The plant, or other organisms that eat the plant, can then obtain energy by breaking down the sugar. The frog in the photo used energy to climb up the tree. That energy came from molecules in the bodies of insects eaten by the frog. The insects, in turn, had built up their bodies by ingesting tissues of plant leaves, which grew by using some of the sun's energy captured through photosynthesis. The frog, like the plants, is ultimately solar-powered, as is the human observer who took this photograph.

The photograph also illustrates other principles of biology. One is that living organisms often survive and thrive by interacting with one another in complex ways. You probably noticed the frog and the trees. But did you notice the patches of growth on the trunk of the tree? Most of those are lichens, a complex interaction between a fungus and a photosynthetic organism (in this case, a kind of alga). In lichens, the fungus and the alga depend on each other for survival. Many other organisms in this scene are too small to be seen, but they are critical components for keeping this living system functioning over time.

After reading this book, you should understand the main principles of life. You'll be able to describe how organisms capture and transform energy; pass genetic information to their offspring in reproduction; grow, develop, and behave; and interact with other organisms and with their physical environment. You will also have learned how this system of life on Earth evolved, and how it continues to change. May a walk in the park never be the same for you again!

CONCEPT 1.1 Living Organisms Share Common Aspects of Structure, Function, and Energy Flow

Biology is the scientific study of life, which encompasses all living things, or **organisms**. The living things we know about are all descended from a single-celled ancestor that lived on Earth almost 4 billion years ago. We can imagine that something with some similarities to life as we know it might have originated differently, perhaps on other planets. But the evidence suggests that all of life on Earth today has a single origin—a single common ancestor—and we consider all the organisms that descended from that common ancestor to be a part of life.

Life as we know it had a single origin

The overwhelming evidence for the common ancestry of life lies in the many characteristics that are shared among living organisms. Typically, living organisms

- are composed of a common set of chemical parts, such as nucleic acids (one example is DNA, which is the important molecule that carries our genetic information) and amino acids (the chemical building blocks that make up proteins), and similar structures, such as cells enclosed within membranes
- depend on intricate interactions among structurally complex parts to maintain the living state
- contain genetic information that uses a nearly universal code to specify how proteins are assembled
- convert molecules obtained from their environment into new biological molecules
- extract energy from the environment and use it to carry out life functions
- replicate their genetic information in the same manner when reproducing themselves
- share structural similarities among a fundamental set of genes
- evolve through gradual changes in their genetic information

Taken together, these shared characteristics logically lead to the conclusion that all life has a common ancestry, and that the diverse organisms that exist today originated from one life form. If life had multiple origins, there would be little reason to expect a nearly universal genetic code, or the similarities among many genes, or a common set of amino acids. If we were to discover something similar to life, such as a self-replicating system that originated independently on another planet, we would expect it to be fundamentally different in these aspects. It might be similar in some ways to life on Earth, such as using genetic information to reproduce. But we would not expect the details of its genetic code, for example, to be like ours.

The simple list of shared characteristics above, however, does not describe the incredible complexity and diversity of life. Some forms of life may not even display all of these characteristics all of the time. For example, the seed of a desert plant may exist for many years in a dormant state in which it doesn't extract energy from the environment, convert molecules, or reproduce. Yet the seed is alive.

And then there are viruses, which are not composed of cells and cannot carry out physiological functions on their own. Instead they use the cells they invade to perform these functions for them. Yet viruses contain genetic information, and they mutate and evolve. So even though viruses are not independent cellular organisms, their existence depends on cells, and there is strong evidence that viruses evolved from cellular life forms. For these reasons, most biologists consider viruses to be a part of life. But as viruses illustrate, the boundaries between "living" and "nonliving" are not always clear, and all biologists do not agree exactly on where we should draw the lines.

Major steps in the history of life are compatible with known physical and chemical processes

Geologists estimate that Earth formed between 4.6 and 4.5 billion years ago. At first the planet was not a very hospitable place. It was some 600 million years or more before the earliest life evolved. If we picture the history of Earth as a 30-day month, with each day representing about 150 million years, life first appeared somewhere toward the end of the first week (FIGURE 1.1).

How might life have arisen from nonliving matter? In thinking about this question, we must take into account that the young Earth's atmosphere, oceans, and climate all were very different than they are today. Biologists have conducted many experiments that simulate the conditions on early Earth. These experiments have confirmed that the formation of complex organic molecules under such conditions is possible, even probable.

The critical step for the evolution of life, however, was the appearance of **nucleic acids**—molecules that could reproduce themselves and also contain the information for the synthesis, or manufacture, of large molecules with complex but stable shapes. These large, complex molecules were proteins. Their shapes varied enough to enable them to participate in increasing numbers and kinds of chemical reactions with other molecules.

CELLULAR STRUCTURE EVOLVED IN THE COMMON ANCESTOR OF LIFE In the next big step in the origin of life, a membrane surrounded and enclosed complex proteins and other biological molecules, forming a tiny **cell**. This membrane kept the enclosed components separate from the surrounding external environment. Molecules called fatty acids played a critical role because these molecules form membrane-like films instead of dissolving in water. When agitated, these films can form hollow spheres, which could have enveloped assemblages of biological molecules. The creation of a cell interior, separate from the external environment, allowed the reactants and products of chemical reactions to be concentrated, opening up the possibility that those reactions could be integrated and controlled. This natural process of membrane formation likely resulted in the first cells with the ability to reproduce—the evolution of the first cellular organisms.

For more than 2 billion years after cells originated, every organism consisted of only one cell. These first organisms were **prokaryotes**, which are made up of single cells containing genetic material and other biochemical structures enclosed in a membrane (FIGURE 1.2). Vast numbers of their descendants, such as bacteria, exist in similar form today. Early prokaryotes were confined to the oceans, which had an abundance of complex molecules they could use as raw materials and sources of energy. The oceans also shielded them from the damaging effects of ultraviolet (UV) light, which was intense at that time because there was little or no oxygen (O_2) in the atmosphere, and for that reason, no protective ozone (O_3) layer in the upper atmosphere.

PHOTOSYNTHESIS ALLOWED LIVING ORGANISMS TO CAPTURE THE SUN'S ENERGY To fuel the chemical reactions inside them, the earliest prokaryotes took in molecules directly from their environment and broke down these small molecules to release and use the energy contained in their chemical bonds. Many modern prokaryotes still function this way, and very successfully.

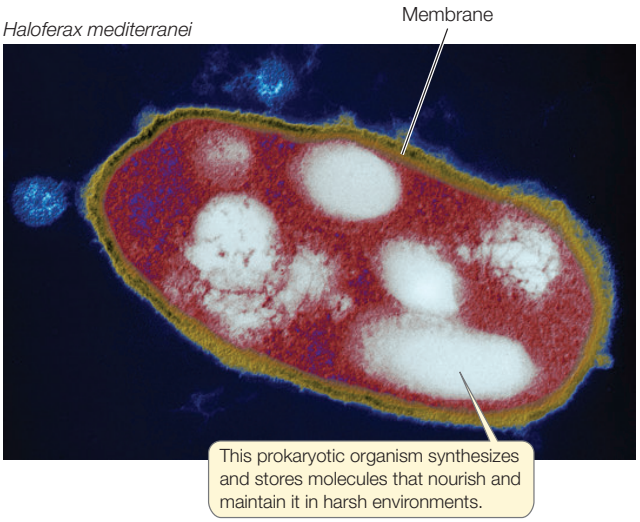


FIGURE 1.2 The Basic Unit of Life Is the Cell The concentration of reactions within the enclosing membrane of a cell allowed the evolution of integrated organisms. Today all organisms, even the largest and most complex, are made up of cells. Single-celled organisms such as this one, however, remain the most abundant living organisms (in absolute numbers) on Earth.

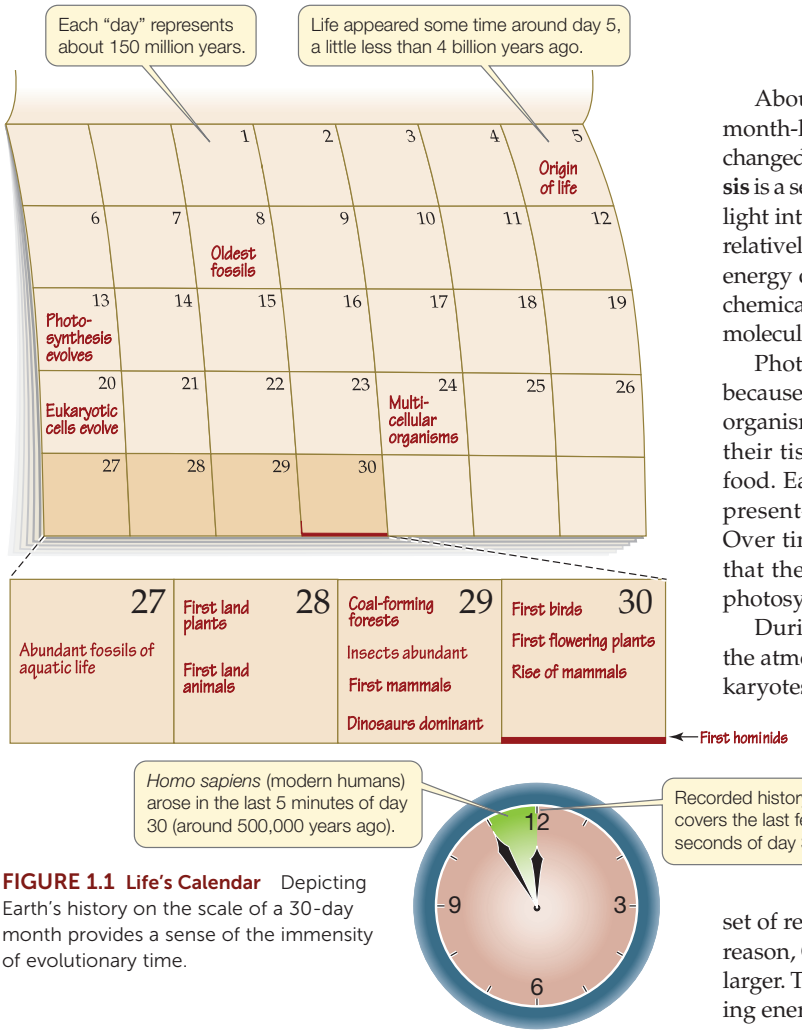


FIGURE 1.1 Life's Calendar Depicting Earth's history on the scale of a 30-day month provides a sense of the immensity of evolutionary time.

About 2.7 billion years ago, or on day 13 of our imaginary month-long calendar of life, the emergence of photosynthesis changed the nature of life on Earth (see Figure 1.1). **Photosynthesis** is a set of chemical reactions that transforms the energy of sunlight into chemical-bond energy of the sugar glucose and other relatively small biological molecules. In turn, the chemical-bond energy of these small molecules can be tapped to power other chemical reactions inside cells, including the synthesis of large molecules, such as proteins, that are the building blocks of cells.

Photosynthesis is the basis of much of life on Earth today because its energy-capturing processes provide food for other organisms. Photosynthetic organisms use solar energy to build their tissues, and then other organisms use those tissues as food. Early photosynthetic cells were probably similar to the present-day prokaryotes called cyanobacteria (FIGURE 1.3). Over time, photosynthetic prokaryotes became so abundant that they produced vast quantities of O_2 as a by-product of photosynthesis.

During the early eons of life on Earth, there was no O_2 in the atmosphere. In fact, O_2 was poisonous to many of the prokaryotes that lived at that time. But those organisms that tolerated O_2 were able to proliferate as O_2 slowly began to accumulate in the atmosphere. The presence of O_2 opened up vast new avenues of evolution. **Aerobic metabolism**, a set of chemical reactions that releases energy from life's molecules by using O_2, proved to be more efficient than **anaerobic metabolism**, a set of reactions that extracts energy without using O_2. For this reason, O_2 allowed organisms to live more intensely and grow larger. The majority of living organisms today use O_2 in extracting energy from molecules.

(A)

(B)

FIGURE 1.3 Photosynthetic Organisms Changed Earth's Atmosphere Cyanobacteria were the first photosynthetic organisms on Earth. (A) Colonies of cyanobacteria called stromatolites are known from the ancient fossil record. (B) Living stromatolites are still found in suitable environments on Earth today.

Oxygen in the atmosphere also made it possible for life to move onto land. For most of life's history, UV radiation falling on Earth's surface was so intense that it destroyed any living cell that was not well shielded by water. But as a result of photosynthesis, O_2 accumulated in the atmosphere for more than 2 billion years and gradually resulted in a layer of ozone in the upper atmosphere. By about 500 million years ago, or about day 28 on our imaginary calendar of life, the ozone layer was sufficiently dense and absorbed enough of the sun's UV radiation to make it possible for organisms to leave the protection of the water and live on land (see Figure 1.1).

EUKARYOTIC CELLS AROSE THROUGH ENDOSYMBIOSIS Another important, earlier step in the history of life was the evolution of cells with membrane-enclosed compartments called **organelles**. Organelles were—and are—important because specialized cellular functions could be performed inside them, separated from the rest of the cell. The first organelles probably appeared about 2.5 billion years after life first appeared on Earth, or about day 20 on Figure 1.1.

One of these organelles, the **nucleus**, came to contain the cell's genetic information. The nucleus (Latin *nux*, "nut" or "core") gives these cells their name: **eukaryotes** (Greek *eu*, "true"; *karyon*, "kernel" or "core"). The eukaryotic cell is distinct from the cells of prokaryotes (*pro*, "before"), which lack nuclei and other internal compartments.

Some organelles are hypothesized to have originated by **endosymbiosis**, which means "living inside another" and may have occurred when larger cells ingested smaller ones. The **mitochondria** that release energy for use by a eukaryotic cell probably evolved from engulfed prokaryotic organisms. And **chloroplasts**—the organelles specialized to conduct photosynthesis in eukaryotic photosynthetic organisms—could have originated when larger eukaryotes ingested photosynthetic prokaryotes. If the larger cell failed to break down this intended food object, a partnership could have evolved in which the ingested prokaryote provided sugars from photosynthesis and the host cell provided a good environment for its smaller partner.

MULTICELLULARITY ALLOWED SPECIALIZATION OF TISSUES AND FUNCTIONS For the first few billion years of life, all organisms—whether prokaryotic or eukaryotic—were single-celled. At some point, the cells of some eukaryotes failed to separate after cell division and remained attached to each other. These groupings of cells made it possible for some cells in the group to specialize in certain functions, such as reproduction, while other cells specialized in other functions, such as absorbing nutrients. **Cellular specialization** enabled multicellular eukaryotes to increase in size and become more efficient at gathering resources and living in specific environments.

Biologists can trace the evolutionary tree of life

If all the organisms on Earth today are the descendants of a single kind of unicellular organism that lived almost 4 billion years ago, how have they become so different? An organism reproduces by replicating its **genome**, which is the sum total of its genetic material, as we will discuss shortly. This replication process is not perfect, however, and changes, called **mutations**, are introduced almost every time a genome is replicated. Some mutations give rise to structural and functional changes in organisms. As individuals mate with one another, these changes can spread within a population, but the population continues to be made up of one kind, or species, of organism. However, if something happens to isolate some members of a population from the others, structural and functional differences between the two groups will accumulate over time. The two groups may eventually differ enough that their members no longer regularly reproduce with one another. In this way the two populations become two different species.

Tens of millions of species exist on Earth today. Many times that number lived in the past but are now extinct. As biologists discover species, they give each one a scientific name called a **binomial** (because it is made up of two Latinized words). The first word identifies the species' genus—a group of species that share a recent common ancestor. The second word indicates the species. For example, the scientific name for the

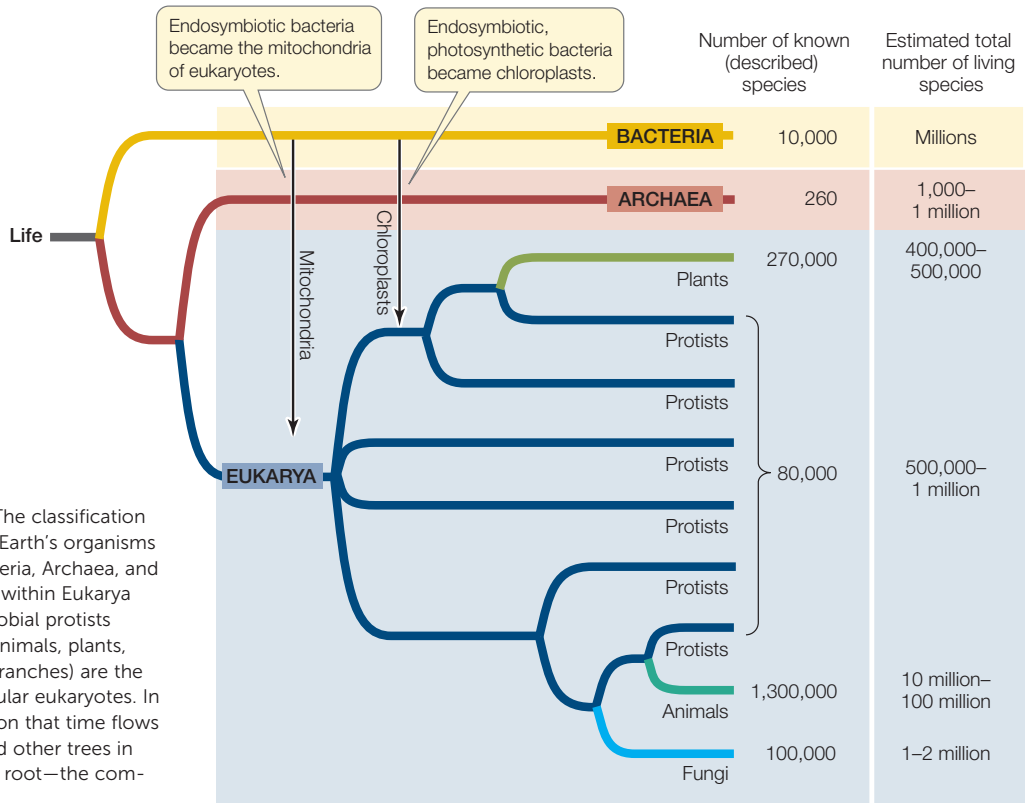


FIGURE 1.4 The Tree of Life The classification system used in this book divides Earth's organisms into three primary domains: Bacteria, Archaea, and Eukarya. The dark blue branches within Eukarya represent various groups of microbial protists (mostly unicellular eukaryotes). Animals, plants, and fungi (green and turquoise branches) are the most familiar groups of multicellular eukaryotes. In this book we adopt the convention that time flows from left to right, so this tree (and other trees in this book) lies on its side, with its root—the common ancestor—at the left.

human species is *Homo sapiens*: *Homo* is our genus and *sapiens* our species. *Homo* is Latin for "man," and *sapiens* is from the Latin word for "wise" or "rational." Our closest relatives in the genus *Homo* are the Neanderthals (*Homo neanderthalensis*), which are now extinct and are known only from fossil remains.

Much of biology is based on comparisons among species. Our ability to make relevant comparisons has improved greatly in recent decades as a result of our relatively newfound ability to study and compare the genomes of different species. We do this by sequencing a genome (in whole or in part), which means we can determine the order of the nucleotides that serve as the building blocks of the organism's DNA. Genome sequencing and other molecular techniques have allowed biologists to add a vast array of molecular evidence to existing evolutionary knowledge based on the fossil record. The result is the ongoing compilation of **phylogenetic trees** that document and diagram evolutionary relationships as part of an overarching **tree of life**. The broadest categories of this tree are shown in FIGURE 1.4. (The tree is expanded in Appendix A, and you can also explore the tree interactively online.)

Many details remain to be clarified, but the broad outlines of the tree of life have been determined. Its branching patterns are based on a rich array of evidence from fossils, structures, chemical processes, behavior, and molecular analyses of genomes. Molecular data in particular have been used to separate the tree into three major branches called **domains**: Archaea, Bacteria, and Eukarya. The organisms of each domain have been evolving separately from those in the other domains for more than a billion years. Note that all organisms that are alive today descended from common ancestors in the past. In other words, living species did not evolve from other species living today. Rather, all living organisms evolved from now-extinct common ancestors. For example, humans did not evolve from our close relatives, the chimpanzees, but humans and chimpanzees both evolved from a common (now extinct) ancestral species.

Organisms in the domains **Archaea** and **Bacteria** are single-celled prokaryotes. However, members of these two groups differ so fundamentally that they are thought to have separated into distinct evolutionary lineages very early. Species belonging to the third domain—**Eukarya**—have eukaryotic cells whose mitochondria and chloroplasts originated from endosymbioses with bacteria, as we have described.

Plants, fungi, and animals are examples of familiar multicellular eukaryotes. We know that multicellularity arose independently in each of these three multicellular groups because they are each most closely related to different groups of unicellular eukaryotes (commonly called protists), as you can see from the branching pattern of Figure 1.4.

Life's unity allows discoveries in biology to be generalized

Knowledge gained from investigations of one kind of organism can, with care, be generalized to other organisms because all life is related by descent from a common ancestor, shares a genetic

code, and consists of similar molecular building blocks. Biologists use certain **model organisms** for research, knowing they can extend their findings to other organisms, including humans.

Our basic understanding of the chemical reactions in cells came from research on bacteria but is applicable to all cells, including those of humans. Similarly, the biochemistry of photosynthesis—the process by which plants use sunlight to produce sugars—was largely worked out from experiments on *Chlorella*, a unicellular green alga. Much of what we know about the genes that control plant development is the result of work on *Arabidopsis thaliana*, a member of the mustard family. Knowledge about how animals develop has come from work on sea urchins, frogs, chickens, roundworms, and fruit flies. And recently, the discovery of a major gene controlling human skin color came from work on zebrafish. Being able to generalize from model systems is a powerful tool in biology.

CONCEPT 1.2 Life Depends on Organization and Energy

All of life depends on organization. Physics gives us the second law of thermodynamics, which states that, left to themselves, organized entities tend to become more random. Any loss of organization threatens the well-being of organisms. Cells, for example, must combat the thermodynamic tendency for their molecules, structures, and systems to lose organization—to become disorganized. Energy is required to maintain organization. For this reason, cells require energy throughout their lives.

Organization is apparent in a hierarchy of levels from molecules to ecosystems

Cells synthesize, or manufacture, proteins and other complex molecules by assembling atoms into new, highly organized configurations. Such complex molecules give cells their structure and enable them to function. For example, a fatty acid molecule that the cell synthesizes may become part of a membrane that structures the inside of the cell by dividing it into compartments. Or a protein made by a cell may enable a specific chemical reaction to take place in the cell by helping start or speed up the reaction—that is, by acting as a catalyst for the reaction.

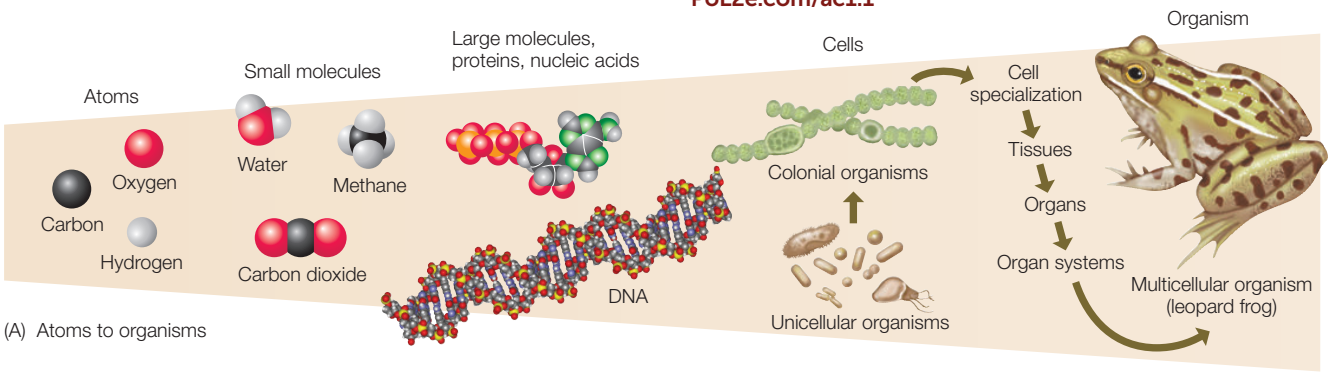
Organization is also essential for many cells to function together in a multicellular organism. As we have seen, multicellularity allows individual cells to specialize and depend on other cells for functions they themselves do not perform. But the different specialized cells also work together. For example, division of labor in a multicellular organism usually requires a circulatory system so that the functions of specialized cells in one part of the body are of use to cells in other, distant parts of the body.

Overall, a multicellular organism exhibits many hierarchical levels of organization (**FIGURE 1.5A**). Small molecules are organized into larger ones, such as DNA and proteins. Large molecules are organized into cells, and assemblages of differentiated cells are organized into **tissues**. For example, a single muscle cell cannot generate much force, but when many cells combine to form the tissue of a working muscle, considerable force and movement can be generated. Different tissue types are organized to form **organs** that accomplish specific functions. The heart, brain, and stomach are each constructed of several types of tissues, as are the roots, stems, and leaves of plants. Organs whose functions are interrelated can be grouped into **organ systems**; the esophagus, stomach, and intestines, for example, are all part of the digestive system. Because all these levels of organization are subject to the second law of thermodynamics, they all tend to degrade unless energy is applied to the system. This is why an organism must use energy to maintain its functions.

Matching the internal hierarchy of an individual organism is an external hierarchy in the larger biological world where organisms interact with their physical environment—an **ecological system**, often shortened to **ecosystem** (**FIGURE 1.5B**). Individual organisms interacting with their immediate

FIGURE 1.5 Life Consists of Organized Systems at a Hierarchy of Scales (A) The hierarchy of systems within a multicellular organism. DNA—a molecule—encodes the information for cells—a higher level of organization. Cells, in turn, are the components of still higher levels of organization: tissues, organs, and the organism itself. (B) Organisms interacting with their external environment form ecological systems on a hierarchy of scales. Individual organisms form the smallest ecological system. Individuals of a species form populations, which interact with other populations to form communities. Multiple communities in turn interact within landscapes at progressively larger scales until they include all the landscapes and organisms of Earth: the entire biosphere.

Go to ACTIVITY 1.1 The Hierarchy of Life
PoL2e.com/ac1.1



environment form the smallest ecological system. Groups of individuals of any one species live together and interact in **populations**, and populations of different species that live and interact in a single area form ecological **communities**. Multiple communities interact within **landscapes**. The landscape of the entire Earth and all its life is known as the **biosphere**.

But there are some important differences between biological systems at the organismal level and these larger scales. All the hierarchical levels of organization within an individual organism are encoded by its single genome, so that these levels generally interact harmoniously. By contrast, the external hierarchy of populations, communities, and landscapes involves interactions among multiple species with multiple genomes, so that interactions are not always harmonious. For example, individuals often prevent others of their own species from exploiting a necessary resource such as food, or they exploit members of their own or different species as food.

Each level of biological organization consists of systems

We have already discussed organ systems and ecological systems. More generally, a **system** is a set of interacting parts in which neither the parts nor the whole can be understood without taking into account the interactions. A simple biological system might consist of a few **components** (e.g., proteins, pools of nutrients, or organisms) and the **processes** by which the components interact (e.g., protein synthesis, nutrient metabolism, or grazing) (**FIGURE 1.6**).

Consider, for example, the system within a cell that synthesizes and controls the quantity of a particular protein, which we'll call Protein T (**FIGURE 1.7A**). The components of the system are the amino acids from which Protein T is made, Protein T, and the breakdown products of Protein T. The processes are the biochemical pathways that synthesize and break down Protein T. To understand how the cell controls the amount of Protein T, we must understand how all the other components and processes in this system function.

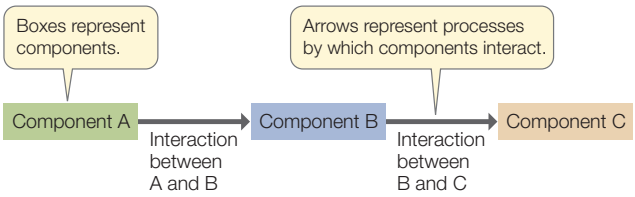


FIGURE 1.6 A Generalized System Systems in cells, whole organisms, and ecosystems can be represented with boxes and arrows.

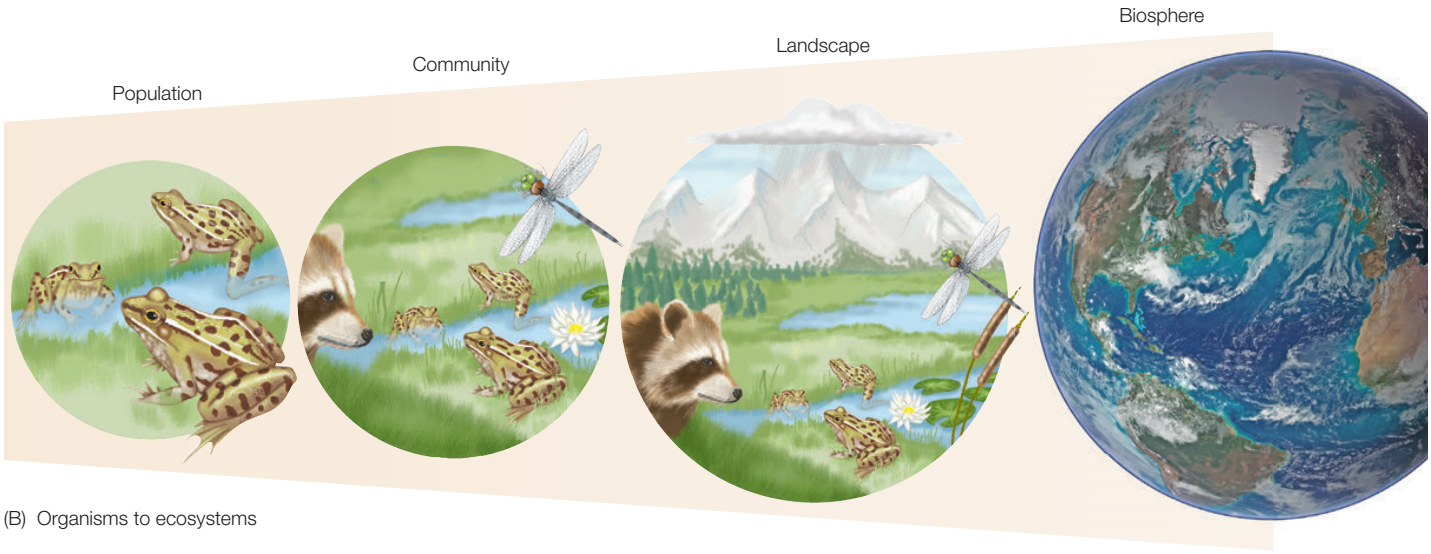
Go to ANIMATED TUTORIAL 1.1
System Simulation
PoL2e.com/at1.1

Systems are found at every level of biological organization. For example, our bodies have a physiological system that controls the amount of sodium (Na^+) in our body fluids (**FIGURE 1.7B**). Grass, voles, and predators (foxes and owls) are components of a community-level system (**FIGURE 1.7C**).

Biological systems are highly dynamic even as they maintain their essential organization

Given the central importance of organization, you might think that biological systems are inflexible and static. Actually, they are often incredibly dynamic—characterized by rapid flows of matter and energy. On average, for example, a cell in your body breaks down and rebuilds 2–3 percent of its protein molecules per day. Each day it also makes and uses more than 100,000 trillion (10^{14}) molecules of adenosine triphosphate (ATP), the molecule responsible for shuttling energy from sources to uses. Collectively, all the cells in your body liberate more than 90 grams of hydrogen every day from the foods they break down to obtain energy. Your cells also combine that hydrogen with oxygen (O_2) to make almost a liter of water every day.

This dynamic aspect of biological systems means that they constantly exchange energy and matter with their surroundings. For example, even after a single-celled or multicellular



(B) Organisms to ecosystems

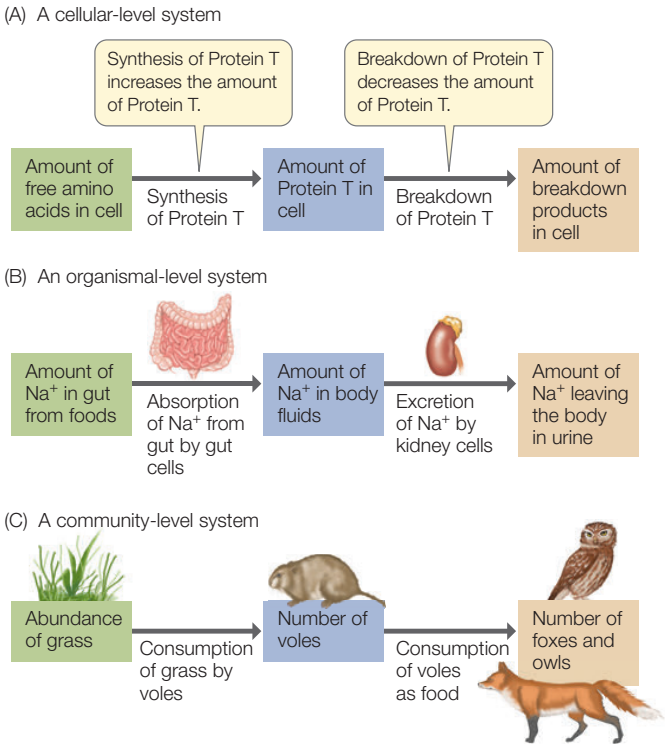


FIGURE 1.7 Organized Systems Exist at Many Levels (A) This cellular-level system synthesizes and breaks down a cell protein called Protein T. (B) This organismal-level system determines the amount (and thus the concentration) of sodium (Na^+) in the blood plasma and other extracellular body fluids of a human. (C) This community-level system helps determine the number of meadow voles (*Microtus pennsylvanicus*) in a field in the spring.

organism has reached maturity, most of its molecules are steadily replaced. In this ceaseless, dynamic process, atoms are lost from the cells in the organism to the surrounding soil, air, or water, and they are replaced with atoms from the soil, air, or water. Yet as the atomic building blocks of any particular cell come and go, the organization of the molecules, structures, and systems in the cell persists. This fact emphasizes the central importance of organization.

Positive and negative feedback are common in biological systems

Often, the amount of one of the components of a system, such as component C in **FIGURE 1.8**, affects the rate of one of the earlier processes in the system. This effect is called **feedback** and may be described as positive or negative. Feedback is often diagrammed simply with a line and symbol, but its actual mechanism may be complex.

Positive feedback occurs in a system when a product of the system *speeds up* an earlier process. The effect of positive feedback is to cause the product to be produced faster and faster. To return to one of our earlier examples, if the breakdown products of Protein T sped up synthesis of Protein T, this would lead to more breakdown products, then even more Protein T, then even more breakdown products, and so on. Positive feedback tends to destabilize a system, but destabilization can sometimes be advantageous, provided it is ultimately brought under control.

Negative feedback occurs when a product of a system *slows down* an earlier process in the system. Often, as the product increases in amount or concentration, it exerts more and more of a slowing effect. Negative feedback stabilizes the amount of the product in this way: if a high amount of the product accumulates, that accumulation tends to reduce further production of the product. For example, if an increase in the amount of breakdown products of Protein T slowed down synthesis of Protein T, this would lead to a decreased amount of breakdown products and a return to the previous rate of Protein T synthesis. Negative feedback is very common in **regulatory systems**, which are systems that tend to stabilize amounts or concentrations.

Systems analysis is a conceptual tool for understanding all levels of organization

Biologists today employ an approach known as **systems analysis** to understand how biological systems function. In systems analysis, we identify the parts or components of a biological system and specify the processes by which the components interact (see Figure 1.6). We may also be able to specify the *rates* of these interactions and how the rates are affected by feedback. What we can do then is analyze how the system will change through time. Will the amounts of different components increase or decrease, and how quickly, and how will this depend on the rates of the interactions? Will there be any stable balance, or equilibrium, that the system eventually reaches?

To do the analysis we write out mathematical equations that express the amounts of the different components and that include the processes and their rates. Expressed in words, such an equation for component B in Figure 1.6 has the following form:

The amount of B present at some time in the future = the amount of B now + the amount of A converted into B – the amount of B converted into C

We write out a similar equation for each component in the system.

We can analyze the relatively simple biological systems in Figure 1.7 by hand, but the analysis of larger systems quickly

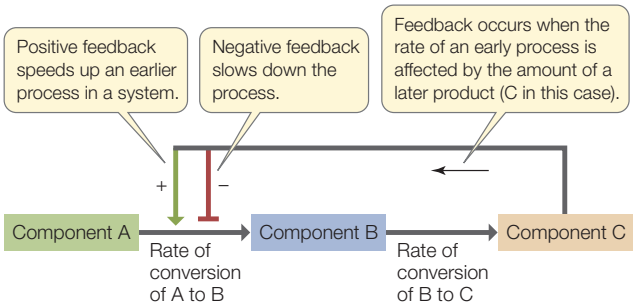
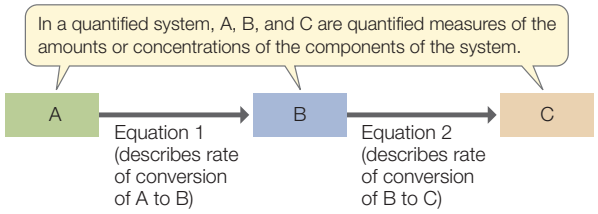


FIGURE 1.8 Feedback Can Be Positive or Negative Positive feedback tends to destabilize a system, whereas negative feedback typically stabilizes a system.

becomes very complicated and is typically carried out using computers. The approach, however, is the same: We express the rates of all processes as mathematical equations.



After this analysis is done, we have a **computational model** of the biological system. If the computational model is well grounded in factual knowledge of the biological system, the model will mimic the biological system.

An important use of computational models is prediction. For instance, if atmospheric temperature affects a biological system, we can use a computational model to develop a hypothetical prediction of the future behavior of the system in a warming world by adjusting the model to take into account the expected increases in temperature.

CONCEPT 1.3 Genetic Systems Control the Flow, Exchange, Storage, and Use of Information

The information required for an organism to function—the "blueprint" for its existence—is contained in the organism's genome, which as we noted earlier is the sum total of all the information encoded by its genes. The presence of genetic information and the processes by which organisms "decode" and use it to build the proteins that underlie a body's structure and function involve fundamental principles that we will discuss and expand on throughout the book, especially in Chapters 10–14.

Genomes encode the proteins that govern an organism's structure

Early in the chapter we noted the importance of self-replicating nucleic acids in the origin of life. Nucleic acid molecules contain long sequences of four subunits called **nucleotides**. The sequence of these nucleotides in **deoxyribonucleic acid**, or **DNA**, allows the organism to assemble **proteins**. Each **gene** is a specific segment of DNA whose sequence carries the information for building, or controlling the building of, one or more proteins (FIGURE 1.9). Proteins, in turn, are the molecules that govern the chemical reactions within cells and form much of an organism's structure. For these reasons, in biology we often say that genes "encode" proteins.

By analogy with a book, the nucleotides of DNA are like the letters of an alphabet. The sentences in the book are genes that encode proteins, which means that the genes provide instructions for making the proteins at a particular time or place. If you were to write out your own genome using four letters to represent the four DNA nucleotides, you would write more

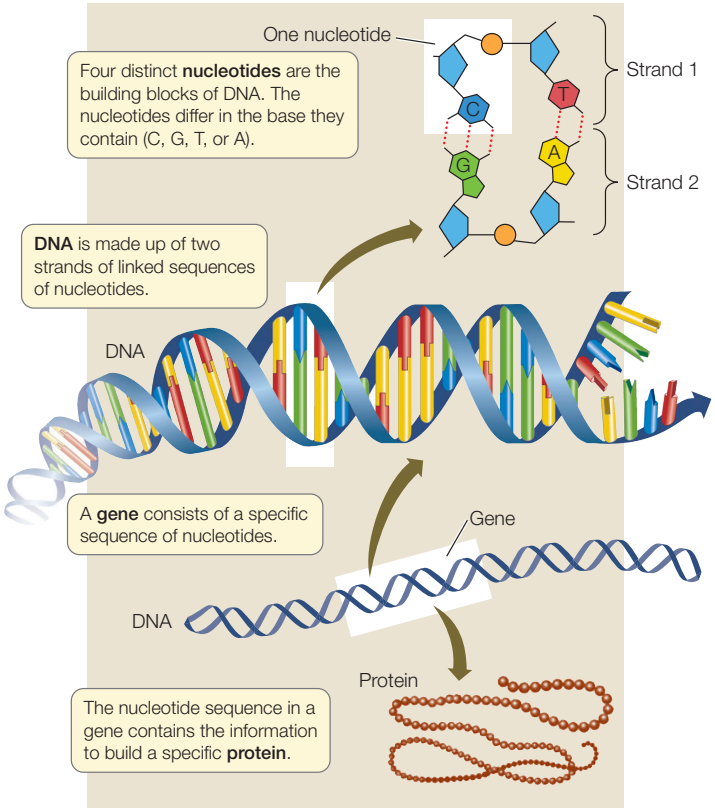


FIGURE 1.9 DNA Is Life's Blueprint The instructions for life are contained in the sequences of nucleotides in DNA molecules. Specific DNA nucleotide sequences comprise genes. The average length of a single human gene is 27,000 nucleotides. The information in each gene provides the cell with the information it needs to manufacture molecules of a specific protein.

than 3 billion letters. Using the size type you are reading now, your genome would fill more than 1,000 books the size of this one.

All the cells of a given multicellular organism contain the same genome, yet the different cells have different functions and form different proteins. For example, oxygen-carrying hemoglobin occurs in red blood cells, gut cells produce digestive proteins, and so on. Therefore different types of cells in an organism must express, or use, different parts of the genome. How any given cell controls which genes it expresses, or uses (and which genes it suppresses, or doesn't use), is a major focus of current biological research.

The genome of an organism contains thousands of genes. If mutations alter the nucleotide sequence of a gene, the protein that the gene encodes is often altered as well. Mutations may occur spontaneously, as happens when mistakes take place during replication of DNA. Mutations can also be caused by certain chemicals (such as those in cigarette smoke) and radiation (including UV radiation from the sun). Most mutations either are harmful or have no effect. Occasionally a mutation improves the functioning of the organism under the environmental conditions the individual encounters. Mutations are the raw material of evolution.

Genomes provide insights into all aspects of an organism's biology

Scientists determined the first complete DNA sequence of an organism's genome in 1976. This first sequence belonged to a virus, and viral genomes are very small compared with those of most cellular organisms. It was another two decades before the first bacterial genome was sequenced, in 1995. The first animal genome to be sequenced was a relatively small one—that of a roundworm—and was determined in 1998. A massive effort to sequence the complete human genome began in 1990 and finished 13 years later.

Since then, scientists have used the methods developed in these pioneering projects, as well as new DNA sequencing technologies that appear each year, to sequence genomes of hundreds of species. As methods have improved, the cost and time for sequencing a complete genome have dropped dramatically. The day is rapidly approaching when the sequencing of genomes from individual organisms will be commonplace for many biological applications.

What are we learning from genome sequencing? One surprise came when some genomes turned out to contain many fewer genes than expected. For example, there are only about 21,000 different genes that encode proteins in a human genome, but most biologists had expected many times that number. Gene sequence information is a boon to many areas of biology, making it possible to study the genetic basis of everything from physical structures to inherited diseases. Biologists can also compare genomes from many species to learn how and why one species differs from another. Such comparative genomic studies allow biologists to trace the evolution of genes through time and to document how particular changes in gene sequences result in changes in structure and function.

The vast amount of information being collected from genome studies has led to rapid development of the field of bioinformatics, the study of biological information. In this emerging field, biologists and computer scientists work together closely to develop new computational tools to organize, process, and study databases used in comparing genomes.

CONCEPT 1.4 Evolution Explains the Diversity as Well as the Unity of Life

Evolution—change in the genetic makeup of biological populations through time—is a major unifying principle of biology. Any process that can lead to changes in the frequencies of genes in a population from generation to generation is an evolutionary process. A common set of evolutionary processes is at work in populations of all organisms. The constant change that occurs in these populations gives rise to all the diversity we see in life. These two themes—unity and diversity—provide a framework for organizing and thinking about the evolution of life. The similarities of life allow us to make comparisons and predictions from one species to another, as we have discussed. The differences are what make biology such a rich and exciting field for investigation and discovery.

Natural selection is an important process of evolution

Charles Darwin compiled factual evidence for evolution in his 1859 book *On the Origin of Species*. Since then, biologists have gathered massive amounts of data supporting Darwin's idea that all living organisms descended from a common ancestor. Darwin also proposed one of the most important processes that produce evolutionary change. He argued that the differing survival and reproduction among individuals in a population, which he termed **natural selection**, could account for much of the evolution of life.

When Darwin proposed that living organisms descended from a common ancestor and are therefore related to one another, he did not have the advantage we have today of understanding the processes of genetic inheritance. Those processes, which we will cover in depth in Chapters 7–9, were not widely understood until the early 1900s. But he knew that offspring differed from their parents, even though they showed strong similarities. And he knew that any population of a plant or animal species displays variation.

Darwin himself bred pigeons, and he knew that if you select breeding pairs on the basis of some particular trait, then that trait is more likely to be present in their offspring than in the general population. He was well aware of how pigeon fanciers selected breeding pairs to produce offspring with unusual feather patterns, beak shapes, or body sizes. He realized that if humans could select for specific traits in organisms such as pigeons, a similar process could operate in nature. Darwin emphasized that human-imposed selection, which he called "artificial selection," has been practiced on crop plants and domesticated animals since the dawn of human civilization. In coining the term "natural selection," he argued that a similar process occurs in nature. But in nature, the "selection" occurs not by human choice but by the fact that some individuals contribute more offspring to future generations than others.

How does natural selection work? Darwin thought that differing probabilities of survival and reproductive success could account for evolutionary change. He reasoned that the reproductive capacity of plants and animals, if unchecked, would result in unlimited growth of populations, but we do not observe such growth in nature. In most species, only a small percentage of offspring survive to reproduce. For this reason, any trait will spread in the population if that trait gives an individual organism even a small increase in the probability that the individual will survive and reproduce.

Because organisms with certain traits survive and reproduce best under specific sets of conditions, natural selection leads to **adaptations**: structural, physiological, or behavioral traits that increase an organism's chances of surviving and reproducing in its environment. For example, remember the frog in the opening photograph of this chapter? Look at the frog's feet and notice that the frog's toes appear to be greatly expanded. These expanded toes would be especially obvious if you could compare them with the toes of frog species that do not live in trees. Expanded toes increase the ability of tree frogs to climb trees, which allows them to hunt insects for food in the treetops and to escape predators on the ground. For this

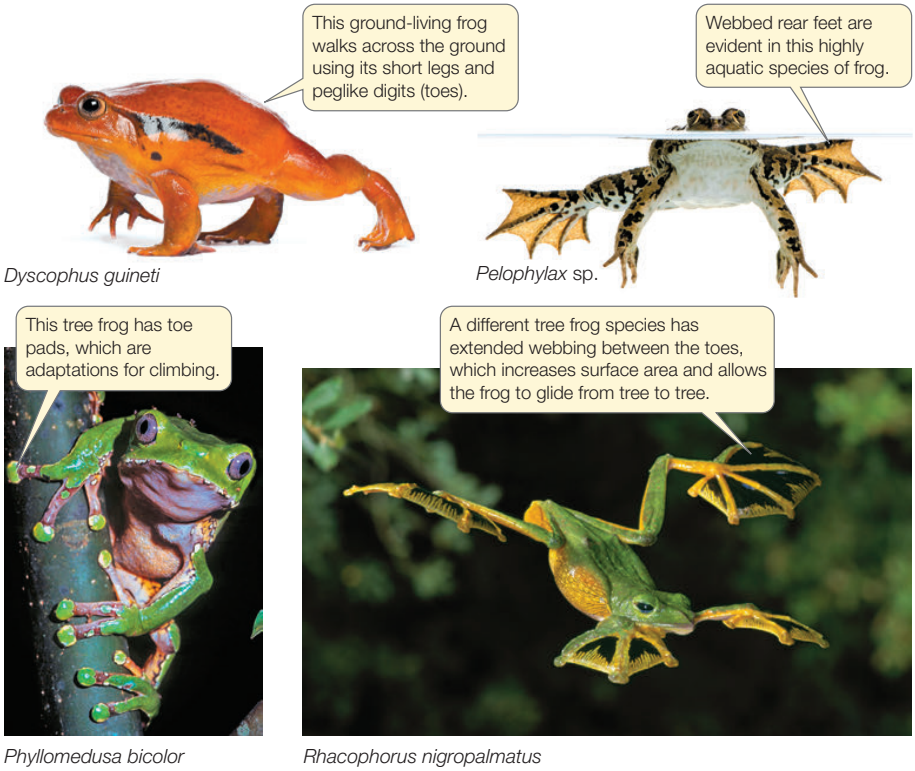


FIGURE 1.10 Adaptations to the Environment The limbs of frogs show adaptations to the different environments of each species.

Go to MEDIA CLIP 1.1
Wallace's Flying Frog
PoL2e.com/mc1.1

reason the expanded toe pads of tree frogs are an adaptation to life in trees. **FIGURE 1.10** shows other adaptations in the limbs of frogs to different environments.

Biologists often think about two different kinds of explanations for adaptations. On the one hand, we can consider the immediate genetic, physiological, neurological, and developmental processes that explain how an adaptation works. We call these **proximate explanations**. For example, a proximate explanation for the toes of tree frogs might examine the physical structure of the toe pads and explain how expansion of the toe leads to greater adhesion to a substrate. Such an explanation tells us how the adaptation works, but it does not explain how tree frogs came to possess such toe pads. An **ultimate explanation**, on the other hand, concerns the processes that led to the evolution of toe pads in various groups of climbing frogs. Ultimate explanations involve comparison of variation within and among species and describe how a given trait affects an organism's chances for survival and reproduction.

Natural selection has been demonstrated in countless biological investigations, but it is not the only process that results in evolution, as we will explore in Chapters 15–18. An example of another evolutionary process is genetic drift, which refers to random changes in gene frequencies in a population because of chance events. As a result of the various evolutionary processes, all biological populations evolve through time. All the evolutionary processes operating over the long history of Earth have led to the remarkable diversity of life on our planet.

Evolution is a fact, as well as the basis for broader theory

The famous biologist Theodosius Dobzhansky once wrote that "Nothing in biology makes sense except in the light of evolution." Dobzhansky was emphasizing the need to include an evolutionary perspective and approach in all aspects of biological study. Everything in biology is a product of evolution, and biologists need a perspective of change and adaptation to fully understand biological systems.

You may have heard someone say that evolution is "just a theory," implying that there is some question about whether or not biological populations evolve. This is a common misunderstanding that originates in part from the different meanings of the word "theory" in everyday language and in science. In everyday speech, some people use the word "theory" to mean "hypothesis" or even—disparagingly—"a guess." In science, however, a **theory** is *a body of scientific work in which rigorously tested and well-established facts and principles are used to make predictions about the natural world.* In short, evolutionary theory is both (1) a body of knowledge supported by facts and (2) the resulting understanding of the various processes by which biological populations have changed and diversified over time, and by which Earth's populations continue to evolve.

We can observe and measure evolution directly, and many biologists conduct experiments on evolving populations. We constantly observe changes in the genetic makeup of populations over relatively short time periods. For example, every year health agencies need to produce new flu vaccines, because populations of influenza viruses evolve so quickly that last year's vaccines may not be effective against this year's populations of viruses. In addition, we can directly observe a record of the history of evolution in the fossil record over the almost unimaginably long periods of geological time. Exactly *how* biological populations change through time is something that is subject to testing and experimentation. The fact that biological populations evolve, however, is not disputed among biologists.

CONCEPT 1.5 Science Is Based on Quantitative Observations, Experiments, and Reasoning

Regardless of the many different tools and methods used in research, all scientific investigations are based on quantitative observation, experimentation, and reasoning. In each of these areas, scientists are guided by an established set of scientific methodological principles.

Observing and quantifying are important skills

Many biologists are motivated by their observations of the living world. Learning *what to observe* in nature is a skill that develops with experience in biology. An intimate understanding of the **natural history** of a group of organisms—how the organisms get their food, reproduce, behave, regulate their functions, and interact with other organisms—leads to better observations and prompts biologists to ask questions about those observations. The more a biologist knows about general principles of life, the more he or she is likely to gain new insights from observing nature.

Biologists have always observed the world around them, but today our ability to observe is greatly extended by technologies such as electron microscopes, rapid genome sequencing, magnetic resonance imaging, and global positioning satellites. These technologies allow us to observe everything from the distribution of molecules in the body (by using electron microscopes) to the daily movement of animals across continents and oceans (by using global positioning satellites).

Observation is a basic tool of biology, but as scientists we must also be able to **quantify** our observations—turn the observations into explicit counts or measures. Whether we are testing a new drug or mapping the migrations of whales, mathematical and statistical calculations are essential. For example, biologists once classified organisms entirely on the basis of qualitative descriptions of the physical differences among them. There was no way of objectively determining evolutionary relationships of organisms, and biologists had to depend on the fossil record for insight. Today our ability to quantify the molecular and physical differences among species, combined with explicit mathematical models of the evolutionary process, enables quantitative analyses of evolutionary history. These mathematical calculations, in turn, make it easier to compare all other aspects of the biology of different organisms.

Scientific methods combine observation, experimentation, and logic

Often, science textbooks describe "*the* scientific method," as if there is a single flow chart that all scientists follow. This view is an oversimplification. Such flow charts include much of what scientists do, but you should not conclude that scientists necessarily go through these steps in one prescribed, linear order.

Observations lead to questions, and scientists make additional observations and often do experiments to answer those questions. This approach, called the hypothesis–prediction method, has five steps: (1) making observations; (2) asking questions; (3) forming hypotheses, or tentative answers to the questions; (4) making predictions based on the hypotheses; and (5) testing the predictions by making additional observations or conducting experiments. These are the steps in traditional flow charts such as the one shown in **FIGURE 1.11**.

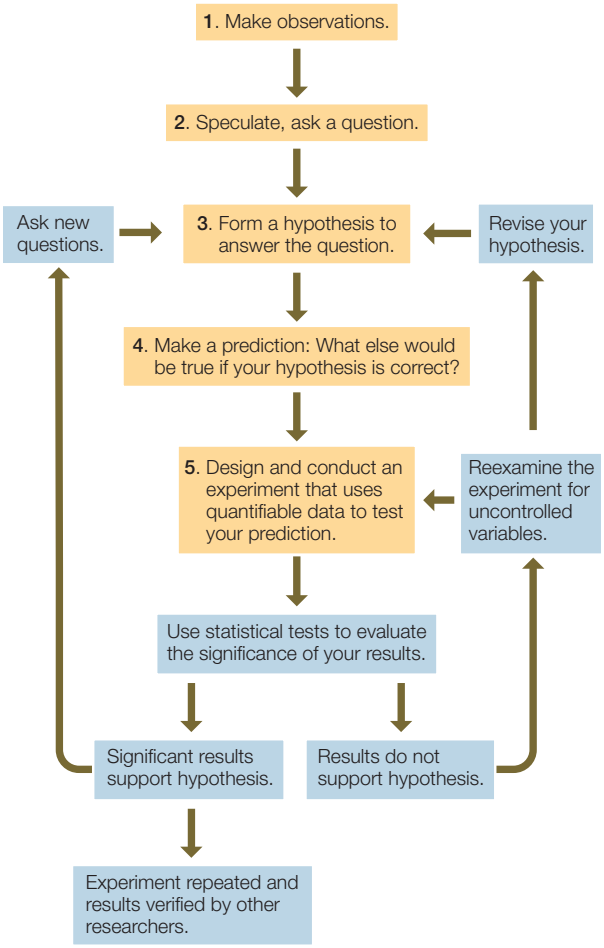


FIGURE 1.11 Scientific Methodology The process of observation, speculation and questioning, hypothesis formation, prediction, and experimentation is a cornerstone of modern science, although scientists may initiate their research at any of several different points.

Go to ANIMATED TUTORIAL 1.2
Using Scientific Methodology
PoL2e.com/at1.2

Getting from questions to answers

Let's consider an example of how scientists start with a general question and work to find answers. Amphibians—such as the frog in the opening photograph of this chapter—have been around for a long time. They watched the dinosaurs come and go. But today scientists have observed that amphibian populations around the world are in dramatic decline. In quantitative terms, more than a third of the world's amphibian species are threatened with extinction. Why is this happening?

To answer big questions like this, biologists begin by sifting through what is already known to arrive at possible answers, or **hypotheses**. In the case of amphibians, biologists know that instances of recent population declines have been associated

with various environmental changes such as loss of moist habitats, changing climate, pathogens, or increased quantities of environmental toxins. Tyrone Hayes, a biologist at the University of California at Berkeley, chose to test the hypothesis that frog populations have been adversely affected by agricultural insecticides and herbicides (weed killers). He did so even though several prior studies had shown that many of these chemicals tested at realistic concentrations do not kill amphibians.

Hayes focused on atrazine, the most widely used herbicide in the world and a common contaminant in fresh water. More than 70 million pounds of atrazine are applied to farmland in the United States every year, and it is used in at least 20 countries. Atrazine kills several types of weeds that can choke fields of important crops such as corn. The chemical is usually applied before weeds emerge in the spring—at the same time many amphibians are breeding and thousands of tadpoles swim in the ditches, ponds, and streams that receive runoff from farms.

In his laboratory, Hayes and his associates raised frog tadpoles in water containing no atrazine and in water with concentrations ranging from 0.01 parts per billion (ppb) up to 25 ppb. The U.S. Environmental Protection Agency considers environmental levels of atrazine of 10–20 ppb of no concern, and it considers 3 ppb a safe level in drinking water. Rainwater in Iowa has been measured to contain 40 ppb. In Switzerland, where the use of atrazine is illegal, the chemical has been measured at approximately 1 ppb in rainwater.

In the Hayes laboratory, an atrazine concentration as low as 0.1 ppb had a dramatic effect on tadpole development: it feminized the males. In some of the adult males that developed from these tadpoles, the vocal structures used in mating calls were smaller than normal, female sex organs developed, and eggs were found growing in the testes. In other studies, normal adult male frogs exposed to 25 ppb had a tenfold reduction in levels of the male sex hormone testosterone and did not produce sperm. You can imagine the disastrous effects these changes could have on the capacity of frogs to breed and reproduce.

But Hayes's experiments were performed in the laboratory, with a species of frog bred for laboratory use. Could atrazine be affecting frogs in nature? If so, then developmental abnormalities in natural frog populations should occur where atrazine is present in their environment. Hayes and his students traveled across the middle of North America, sampling water and collecting frogs to test this prediction. They analyzed the water for atrazine and examined the frogs. In the only site where atrazine was undetectable in the water, the frogs were normal. In all the other sites, male frogs had abnormalities of the sex organs.

Like other biologists, Hayes made observations. He then made predictions based on those observations, and designed and carried out experiments to test his predictions. Some of the conclusions from his experiments, described below, could have profound implications not only for amphibians but also for other animals, including humans.

Well-designed experiments have the potential to falsify hypotheses

Once predictions are made from a hypothesis, experiments can be designed to test those predictions. The most informative experiments have the ability to show that the prediction is wrong. If the prediction is wrong, the hypothesis must be modified or rejected.

There are two general types of experiments. Both compare data from different groups or samples. A **controlled experiment** changes, or manipulates, one or more of the factors being tested. A **comparative experiment** compares unmanipulated data gathered from different sources.

In a controlled experiment, we start with groups or samples that are as similar as possible. We predict on the basis of our hypothesis that some critical factor, or **variable**, has an effect on the phenomenon we are investigating. We devise some method to manipulate *only that variable* in an "experimental" group, and we compare the resulting data with data from an unmanipulated "control" group. If the predicted difference occurs, we then apply statistical tests to find the probability that the manipulation created the difference (as opposed to the difference being the result of random chance). **FIGURE 1.12** describes one of the many controlled experiments performed by the Hayes laboratory to quantify the effects of atrazine on male frogs.

The basis of controlled experiments is that one variable is manipulated while all others are held constant. The variable that is manipulated is called the independent variable (because an investigator can manipulate it independently of other considerations). The response that is measured is the dependent variable (because it is not manipulated directly by the investigator but is permitted to vary in ways that depend on the independent variable). A perfectly controlled experiment is not easy to design because biological variables are so interrelated that it is difficult to alter just one.

A comparative experiment starts with the prediction that there will be a difference among naturally existing samples or groups based on the hypothesis. In comparative experiments, we do not control any of the variables, and often we cannot even identify all the variables that are present. We simply gather and compare data from different naturally occurring sample groups.

When his controlled experiments indicated that atrazine indeed affects reproductive development in frogs, Hayes and his colleagues performed a comparative experiment. They collected frogs and water samples from eight widely separated sites across the United States and compared the percentages of abnormal frogs from environments with very different levels of atrazine (**FIGURE 1.13**). Of course, the sample sites differed in many ways besides the level of atrazine present.

The results of experiments frequently reveal that the situation is more complex than the hypothesis anticipated, thus raising new questions. There are no "final answers" in science. As a result, biologists often develop new questions, hypotheses, and experiments as they collect more data. The process of science is open-ended in this regard, and continued research leads to an ever-better understanding of the living world, with practical implications for agriculture, medicine, conservation of species, and other endeavors.

Statistical methods are essential scientific tools

Whether we do controlled or comparative experiments, at the end we have to decide whether there is a difference among the

INVESTIGATION

FIGURE 1.12 Controlled Experiments Manipulate a Variable
The Hayes laboratory created controlled environments that differed only in the concentrations of atrazine in the water. Eggs from leopard frogs (*Rana pipiens*) raised specifically for laboratory use were allowed to hatch and the tadpoles were separated into experimental tanks containing water with different concentrations of atrazine.[a]

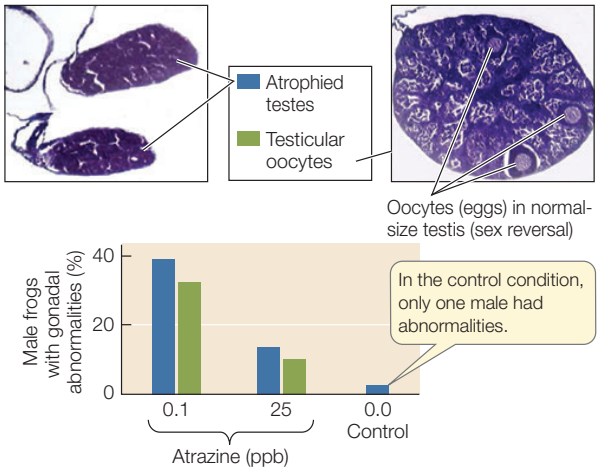
HYPOTHESIS

Exposure to atrazine during larval development causes abnormalities in the reproductive tissues of male frogs.

METHOD

1. Establish 9 tanks in which all attributes are held constant except the water's atrazine concentration. Establish 3 atrazine conditions (3 replicate tanks per condition): 0 ppb (control condition), 0.1 ppb, and 25 ppb.
2. Place *Rana pipiens* tadpoles from laboratory-reared eggs in the 9 tanks (30 tadpoles per tank).
3. When tadpoles have transitioned into adults, sacrifice the animals and evaluate their reproductive tissues.
4. Test for relationship between degree of atrazine exposure and the presence of abnormalities in the gonads (testes) of male frogs.

RESULTS



CONCLUSION

Exposure to atrazine at concentrations as low as 0.1 ppb induces abnormalities in the gonads of male frogs. The effect is not proportional to the level of exposure.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]T. Hayes et al. 2003. *Environmental Health Perspectives III*: 568–575.

samples, individuals, groups, or populations in the study. How do we decide whether a measured difference is enough to support or reject a hypothesis? In other words, how do we decide in an unbiased, objective way whether the measured difference is meaningful, or significant?

Statistical significance refers to the extent to which a result is unlikely to be due to chance alone. Scientists use statistics because they recognize that variation is always present in any set of measurements. Statistical tests calculate the probability that the differences observed in an experiment could be due to random variation. The results of statistical tests are therefore probabilities. Many statistical tests start with a **null hypothesis**—the premise that any observed differences are simply the result of random differences that arise from drawing two samples from the same population. When scientists collect quantified observations, or **data**, they apply statistical methods to those data to calculate the likelihood that the null hypothesis is correct.

More specifically, statistical methods tell us the probability of obtaining a particular result by chance alone, even if the samples being tested are drawn from the same population. As scientists, we need to eliminate, insofar as possible, the possibility that any differences seen are simply due to chance variation in the samples. Appendix B in this book is a short primer on statistical methods that you can refer to as you analyze data that will be presented throughout the text.

Not all forms of inquiry into nature are scientific

Science is a human endeavor that is bounded by certain standards of practice. Other areas of scholarship share with science the practice of making observations and asking questions, but scientists are distinguished by what they do with their observations and how they answer their questions. Data, subjected to appropriate statistical analysis, are critical in testing hypotheses. Science is the most powerful approach humans have devised for learning about the world and how it works.

Scientific explanations for natural processes are objective and reliable because the hypotheses proposed *must be testable* and *must have the potential of being rejected* by direct observations and experiments. Scientists must clearly describe the methods they use to test hypotheses so that other scientists can repeat their experiments to see if they get the same results. Not all experiments are repeated, but surprising or controversial results are always subjected to independent verification. Scientists worldwide share this process of testing and rejecting hypotheses, contributing to a common body of scientific knowledge.

If you understand the methods of science, you can distinguish science from non-science. Art, music, and literature all contribute to the quality of human life, but they are not science. They do not use scientific methods to establish what is fact. Religion is not science, although religions have historically attempted to explain natural events ranging from unusual weather patterns to crop failures to human diseases. Most such phenomena that at one time were mysterious can now be explained in terms of scientific principles. Fundamental tenets of religious faith, such as the existence of a supreme deity or deities, cannot be confirmed or refuted by experimentation and for this reason are outside the realm of science.

The power of science derives from the uncompromising objectivity and absolute dependence on evidence that comes from reproducible and quantifiable observations. A religious or spiritual explanation of a natural phenomenon may be

INVESTIGATION

FIGURE 1.13 Comparative Experiments Look for Differences among Groups To see whether the presence of atrazine correlates with testicular abnormalities in male frogs, the Hayes lab collected frogs and water samples from different locations around the U.S. The analysis that followed was "blind," meaning that the frogs and water samples were coded so that experimenters working with each specimen did not know which site the specimen came from.

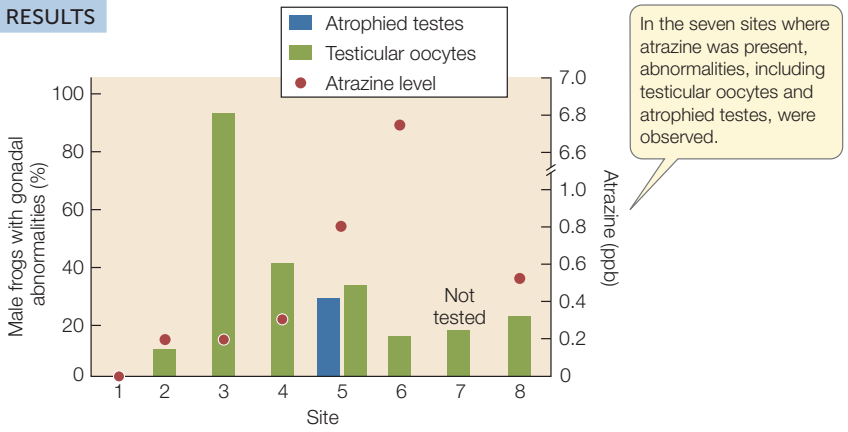
HYPOTHESIS

Presence of the herbicide atrazine in environmental water correlates with gonadal abnormalities in frog populations.

METHOD

1. Based on commercial sales of atrazine, select 4 sites (sites 1–4) less likely and 4 sites (sites 5–8) more likely to be contaminated with atrazine.
2. Visit all sites in the spring (i.e., when frogs have transitioned from tadpoles into adults); collect frogs and water samples.
3. In the laboratory, sacrifice frogs and examine their reproductive tissues, documenting abnormalities.
4. Analyze the water samples for atrazine concentration (the sample for site 7 was not tested).
5. Quantify and correlate the incidence of reproductive abnormalities with environmental atrazine concentrations.

RESULTS



CONCLUSION

Reproductive abnormalities exist in frogs from environments in which aqueous atrazine concentration is 0.2 ppb or above. The frequency of abnormalities does not appear to be proportional to atrazine concentration at the time of transition to adulthood.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]T. Hayes et al. 2002. *Nature* 419: 895–896.

coherent and satisfying for the person holding that view, but it is not testable and therefore it is not science. To invoke a supernatural explanation (such as a "creator" or "intelligent designer" with no known bounds) is to depart from the world of science. Science does not say that untestable religious beliefs are necessarily wrong, just that they are not something that we can address using scientific methods.

Science describes how the world works. It is silent on the question of how the world "ought to be." Many scientific advances that contribute to human welfare also raise major ethical issues. Recent developments in genetics and developmental biology may enable us to select the sex of our children, to use stem cells to repair our bodies, and to modify the human genome. Scientific knowledge allows us to do these things, but science cannot tell us whether or not we should do so, or if we choose to do them, how we should regulate them. Such questions are as crucial to human society as the science itself. A responsible scientist does not lose sight of these questions or neglect the contributions of the humanities in attempting to come to grips with them.

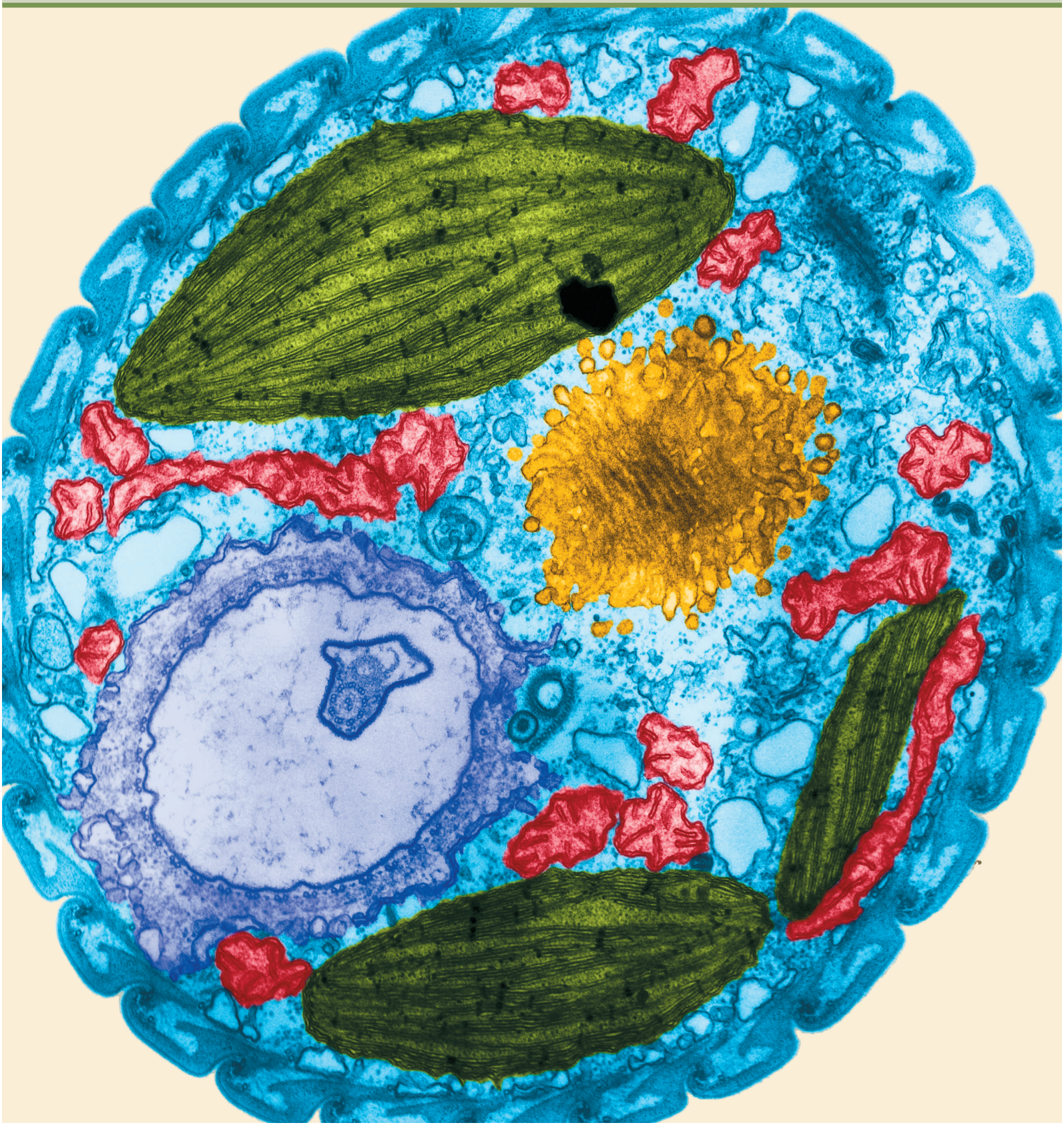
Consider the big themes of biology as you read this book

You will see evolution and the other fundamental principles of life introduced in this chapter at work in each part of this book. In Part 1 you will learn about the molecular organization of life. We will discuss the origin of life, the energy in atoms and molecules, and how proteins and nucleic acids became the self-replicating cellular systems of life. Part 2 will describe how these self-replicating systems work and the genetic principles that explain heredity and mutation, which are the basis of evolution. In Part 3 we will describe the processes of evolution and go into detail about how evolution works. Part 4 will examine the products of evolution: the vast diversity of life and the many different ways organisms solve common problems such as how to reproduce, defend themselves, and obtain nutrients. Parts 5 and 6 will explore the physiological adaptations that allow plants and animals to survive and function in a wide range of physical environments. Finally, in Part 7 we will discuss these environments and the integration of individual organisms, populations, and communities into the interrelated ecological systems of Earth.

You may enjoy returning to this chapter occasionally as the course progresses. The brief explanations we have given here should become more meaningful as you read about the facts and phenomena that underlie the principles. Our knowledge of the "facts" of biology, however, is not based just on reading, contemplation, or discussion, although all of these activities are important. Scientific knowledge is based on active and always-ongoing research.

PART 1

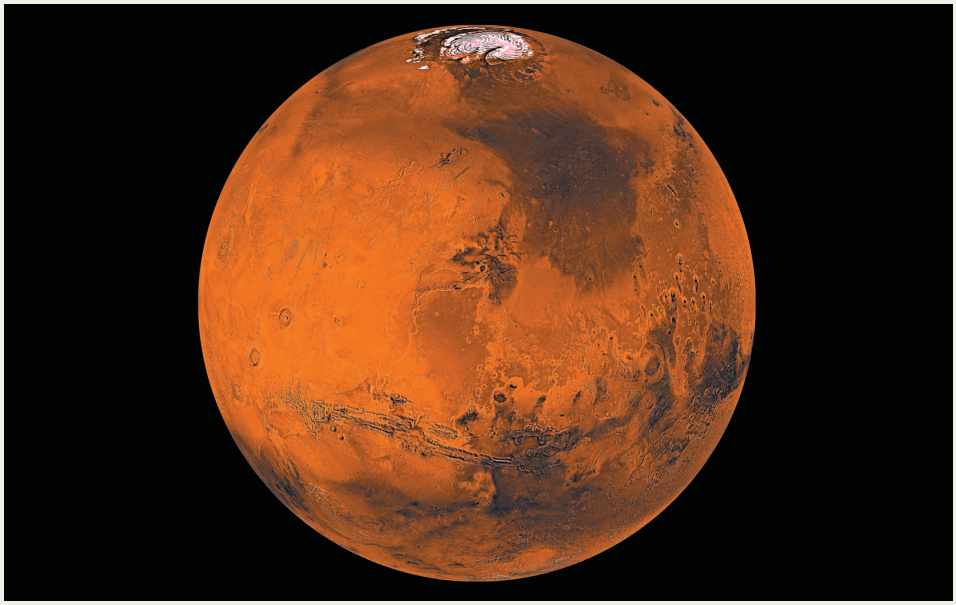
Cells



2

The Chemistry and Energy of Life

KEY CONCEPTS



Polar ice caps, as shown here, have been observed on Mars for a long time, but recent evidence also shows water at the milder mid-latitudes of Mars. Finding water on Mars may indicate that it is or was hospitable to life.

A major discovery of biology was that living things are composed of the same chemical elements as the vast nonliving portion of the universe. This mechanistic view—that life is chemically based and obeys the universal laws of chemistry and physics—is relatively new in human history. Until the nineteenth century, many scientists thought that a "vital force," distinct from the forces governing the inanimate world, was responsible for life. Many people still assume that such a vital force exists. However, the mechanistic view of life has led to great advances in biological science, and it underpins many of the applications of biology to medicine and agriculture. We use a mechanistic view throughout this book.

Among the most abundant chemical elements in the universe are hydrogen and oxygen, and life as we know it requires the presence of these elements as water (H_2O). Water makes up about 70 percent of the bodies of most organisms, and those that live on land have evolved elaborate ways to retain the water in their bodies. Aquatic organisms do not need these water-retention mechanisms; thus biologists think that life originated in a watery environment.

Life has been found in some surprising places, often in extreme conditions. There are organisms living in hot springs at temperatures above the boiling point of water, 5 kilometers below Earth's surface, at the bottom of the ocean, and in extremely acid or salty conditions.

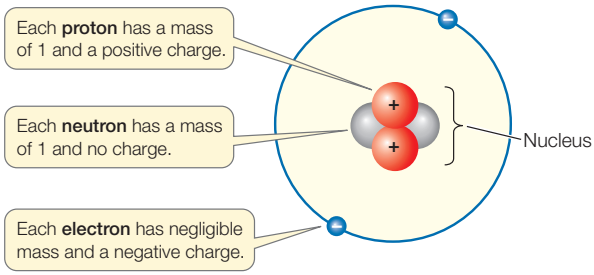
Life has been found even below the Antarctic ice, a finding especially relevant to the search for life beyond Earth. Some moons that orbit Jupiter and Saturn have ice on their surfaces, overlaying oceans of water much larger than the oceans on Earth. The subsurface water on these moons may be maintained as a liquid in large part because ice is a good insulator and traps heat generated by tidal forces between the moons and their planets. Europa is a moon of Jupiter that is slightly smaller than Earth's moon. It has an atmosphere composed primarily of oxygen, and scientists speculate that Europa's subsurface oceans could harbor life, just as the Antarctic ice lakes do. Mars also has ice on its surface, with some evidence of subsurface water. Two of NASA's highest priorities are to launch missions to Mars and Europa to bring back samples from these lakes to Earth.

Q Why is the search for water important in the search for life?

You will find the answer to this question on page 34.

CONCEPT 2.1 Atomic Structure Is the Basis for Life's Chemistry

Living and nonliving matter is composed of **atoms**. Each atom consists of a dense, positively charged **nucleus**, with one or more negatively charged **electrons** moving around it. The nucleus contains one or more positively charged **protons**, and may contain one or more **neutrons** with no electrical charge:



Charges that are different (+/–) attract each other, whereas charges that are alike (+/+, –/–) repel one another. Most atoms are electrically neutral because the number of electrons in an atom equals the number of protons.

The standard unit of measure for the mass of an atom (atomic mass) is the dalton—named after the English chemist John Dalton. A single proton or neutron has a mass of about 1 dalton (Da), which is 1.7×10^{-24} grams, but an electron is even tinier, at 0.0005 Da (9×10^{-28} g). Because the mass of an electron is only about 1/2,000th of the mass of a proton or neutron, the contribution of electrons to the mass of an atom can usually be ignored when chemical measurements and calculations are made.

An element consists of only one kind of atom

An **element** is a pure substance that contains only one kind of atom. The element hydrogen consists only of hydrogen atoms, the element gold only of gold atoms. The atoms of each element have characteristics and properties that distinguish them from the atoms of other elements.

There are 94 elements in nature, and at least another 24 have been made in physics laboratories. Most of the 94 natural elements have been detected in living organisms, but just a few predominate. About 98 percent of the mass of every living organism (bacterium, turnip, or human) is composed of just six elements:

Carbon (symbol C)	Hydrogen (H)	Nitrogen (N)
Oxygen (O)	Phosphorus (P)	Sulfur (S)

The chemistry of these six elements will be our primary concern in this chapter, but other elements found in living organisms are important as well. Sodium and potassium, for example, are essential for nerve function; calcium can act as a biological signal; iodine is a component of a human hormone; and magnesium is bound to chlorophyll in green plants.

The physical and chemical (reactive) properties of atoms depend on the numbers of protons, neutrons, and electrons they contain. The atoms of an element differ from those of other elements by the number of protons in their nuclei. The number of protons is called the **atomic number**, and it is unique to and characteristic of each element. A carbon atom has six protons and thus an atomic number of 6; the atomic number of oxygen is 8. For electrical neutrality, each atom has the same number of electrons as protons, so a carbon atom has six electrons and an oxygen atom has eight.

Along with a definitive number of protons, every element except hydrogen has one or more neutrons. The **mass number** of an atom is the total number of protons and neutrons in its nucleus. The number of neutrons may vary among atoms of a particular element. For example, carbon atoms with six, seven, and eight neutrons are found in nature. These variants are referred to as **isotopes**. The most common carbon isotope has six neutrons and a mass number of 12, and is referred to as carbon-12 (often written ^{12}C). The most common oxygen isotope (^{16}O) has eight protons and eight neutrons, and a mass number of 16.

Go to MEDIA CLIP 2.1
The Elements Song
PoL2e.com/mc2.1

Electrons determine how an atom will react

The **Bohr model** for atomic structure (see diagram in previous column) provides a concept of an atom that is largely empty space, with a central nucleus surrounded by electrons in orbits, or **electron shells**, at various distances from the nucleus. This model is much like our solar system, with planets orbiting around the sun. Although highly oversimplified (you will learn about the reality of atomic structure in physical chemistry courses), the Bohr model is useful for describing how atoms behave. Specifically, *the behaviors of electrons determine whether a chemical bond will form and what shape the bond will have*. These are two key properties for determining biological changes and structure.

In the Bohr model, each electron shell is a certain distance from the nucleus. Since electrons are negatively charged and protons are positive, an electron needs energy to escape from the attraction of the nucleus. The farther away an electron shell is from the nucleus, the more energy the electron must have. We will return to this topic when we discuss biological energetics in Chapter 6.

The electron shells, in order of their distance from the nucleus, can be filled with electrons as follows:

- First shell: up to 2 electrons
- Second shell: up to 8 electrons
- Third shell: up to 18 electrons
- Fourth and subsequent shells: up to 32 electrons

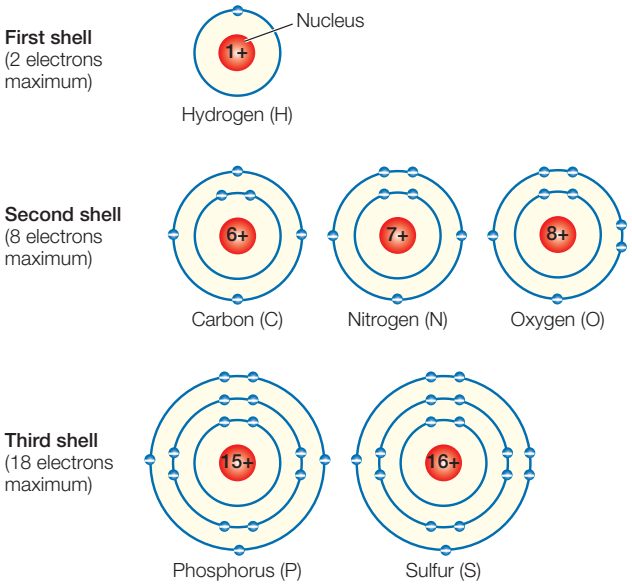


FIGURE 2.1 Electron Shells Each shell can hold a specific maximum number of electrons and must be filled before electrons can occupy the next shell. The energy level of an electron is higher in a shell farther from the nucleus. An atom with fewer than eight electrons in its outermost shell (or two in the case of hydrogen) can react (bond) with other atoms.

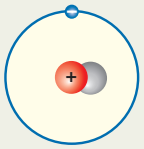
Electrons fill shells closest to the nucleus before occupying shells farther from the nucleus. FIGURE 2.1 illustrates the electron shell configurations for the six major elements found in living systems.

For elements with atomic numbers between 6 and 20 there is a chemical rule of thumb called the **octet rule**, which states that an atom will lose, gain, or share electrons in order to achieve a stable configuration of eight electrons in its outermost shell. Oxygen, for example, which has six electrons in its outermost shell, will undergo chemical reactions to gain two electrons. When atoms share electrons, they form stable associations called **molecules**. Most atoms in biologically important molecules—for example, carbon and nitrogen—follow the octet rule. However, very small atoms such as hydrogen (with one proton and one electron) tend to gain, lose, or share electrons such that their single shell contains its maximum number of two electrons.

CHECKpoint CONCEPT 2.1

- ✓ What is the arrangement of protons, neutrons, and electrons in an atom?
- ✓ Sketch the electron shell configuration of a sodium atom (symbol Na), which has 11 protons. According to the octet rule, what would be the simplest way for a sodium atom to achieve electron stability?
- ✓ Many elements have isotopes, which are rare variants of the element with additional neutrons in the nucleus. Deuterium is an isotope of hydrogen that has one neutron (normal hydrogen has no neutrons).

Does the neutron change the chemical reactivity of deuterium, compared with normal hydrogen? Explain why or why not.



Deuterium
1 proton
1 neutron

We have introduced the basic unit of matter that makes up all living organisms—the atom. We have discussed the tendency of atoms to attain stable configurations of electrons: a single shell of two electrons in the case of hydrogen, and an outer shell of eight electrons in the case of larger atoms. Next we will describe the different types of chemical bonds that can lead to stability, joining atoms together into molecular structures with different properties.

CONCEPT 2.2 Atoms Interact and Form Molecules

A **chemical bond** is an attractive force that links two atoms together in a molecule. There are several kinds of chemical bonds (TABLE 2.1; see p. 22). In this section we will begin with covalent bonds, the strong bonds that result from the sharing of electrons. Next we will consider weaker interactions, including hydrogen bonds, which are enormously important to biology. We will then examine ionic attractions, which form when an atom gains or loses electrons to achieve stability. Finally, we will see how atoms are bonded to make functional groups—groups of atoms that give important properties to biological molecules.

Go to ANIMATED TUTORIAL 2.1
Chemical Bond Formation
PoL2e.com/at2.1

Covalent bonds consist of shared pairs of electrons

A **covalent bond** forms when two atoms attain stable electron numbers in their outermost shells by *sharing* one or more pairs of electrons. In this case, each atom contributes one member of each electron pair. Consider two hydrogen atoms coming close together, each with just one electron in its single shell (FIGURE 2.2). When the electrons pair up, a stable association is formed, and this links the two hydrogen atoms in a covalent bond, forming the molecule H_2.

Let's see how covalent bonds are formed in the somewhat more complicated methane molecule (CH_4). The carbon atom has six electrons: two electrons fill its inner shell, and four electrons are in its outer shell. Because of the octet rule, carbon is most stable when it shares electrons with four other atoms—*it can form four covalent bonds* (FIGURE 2.3A). Methane forms when an atom of carbon reacts with four hydrogen atoms. As a result of electron sharing, the outer shell of the carbon atom is now filled with eight electrons—a stable configuration. The single shell of each hydrogen atom is also filled. Four covalent bonds—four shared electron pairs—hold methane together.

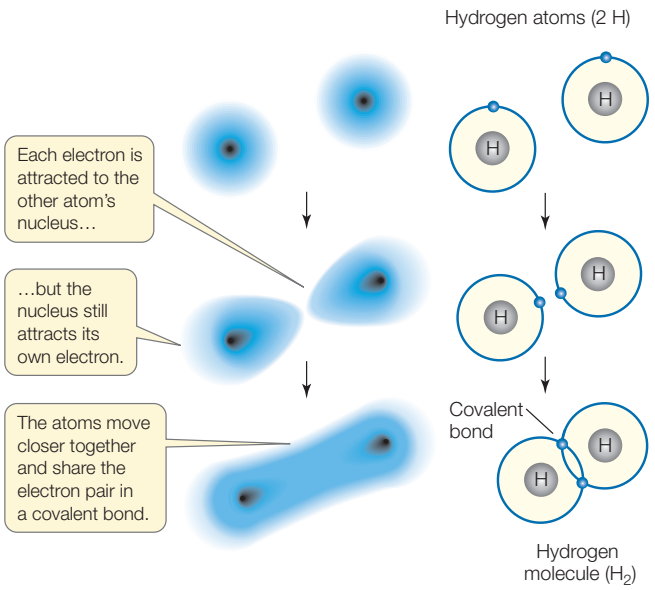


FIGURE 2.2 Electrons Are Shared in Covalent Bonds Two hydrogen atoms can combine to form a hydrogen molecule. A covalent bond forms when the electron shells of the two atoms overlap in an energetically stable manner.

FIGURE 2.3B shows several different ways to represent the molecular structure of methane. **TABLE 2.2** shows the covalent bonding capacities of some biologically important elements.

The properties of molecules are influenced by the characteristics of their covalent bonds. Four important aspects of covalent bonds are orientation, strength and stability, multiple covalent bonds, and the degree of sharing of electrons.

ORIENTATION For a given pair of elements, such as carbon bonded to hydrogen, the length of the covalent bond is always the same. And for a given atom within a molecule, the angle of each covalent bond with respect to the others is generally the same. This is true regardless of the type of larger molecule that contains the atom. For example, the four covalent bonds formed by the carbon atom in methane are always distributed in space so that the bonded hydrogens point to the corners of a regular tetrahedron, with the carbon in the center (see Figure 2.3B). Even when carbon is bonded to four atoms other than hydrogen, this three-dimensional orientation is more or less maintained. As you will see, the orientations of covalent bonds in space give molecules their three-dimensional geometry, and the shapes of molecules contribute to their biological functions.

STRENGTH AND STABILITY Covalent bonds are very strong (see Table 2.1), meaning it takes a lot of energy to break them. At the temperatures at which life exists, the covalent bonds of biological molecules are quite stable, as are their three-dimensional structures. However, this stability does not rule out change, as we will discover.

MULTIPLE COVALENT BONDS As shown in Figure 2.3B, covalent bonds can be represented by lines between the chemical symbols for the linked atoms:

- A *single bond* involves the sharing of a single pair of electrons (for example, H—H or C—H).

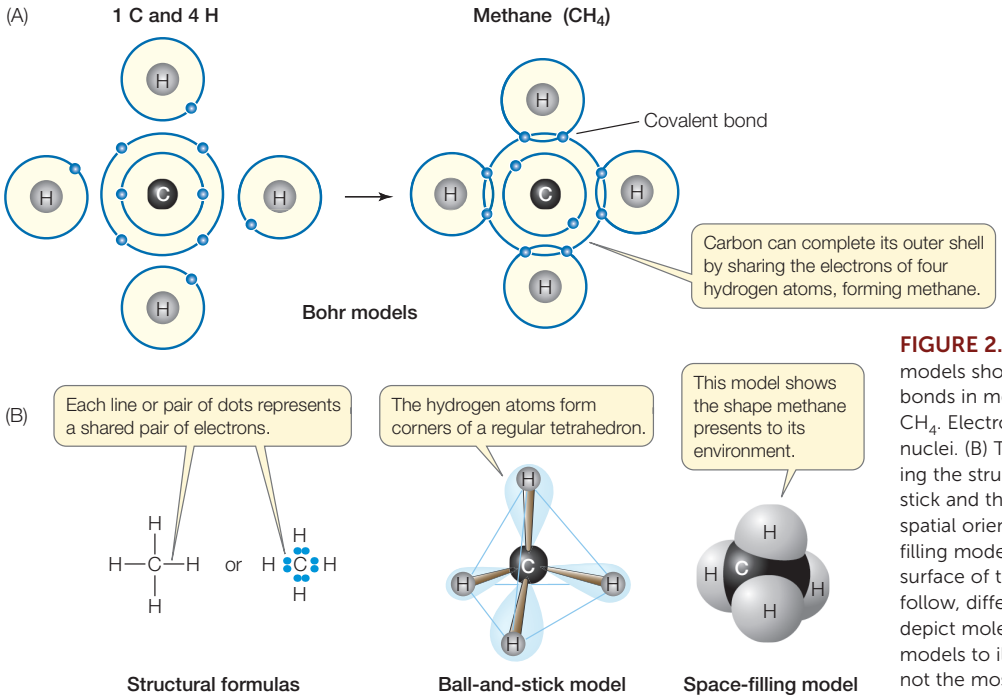
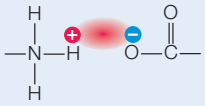
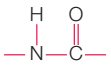
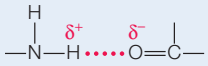
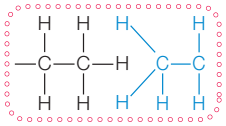
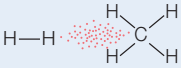


FIGURE 2.3 Covalent Bonding (A) Bohr models showing the formation of covalent bonds in methane, whose molecular formula is CH_4. Electrons are shown in shells around the nuclei. (B) Three additional ways of representing the structure of methane. The ball-and-stick and the space-filling models show the spatial orientations of the bonds. The space-filling model indicates the overall shape and surface of the molecule. In the chapters that follow, different conventions will be used to depict molecules. Bear in mind that these are models to illustrate certain properties and are not the most accurate portrayals of reality.

TABLE 2.1 Chemical Bonds and Interactions

Name	Basis of interaction	Structure	Bond energy[a]
Ionic attraction	Attraction of opposite charges		3–7
Covalent bond	Sharing of electron pairs		50–110
Hydrogen bond	Attraction between H (δ^+) and a strongly electronegative atom		3–7
Hydrophobic interaction	Interaction of nonpolar substances in the presence of polar substances (especially water)		1–2
van der Waals interaction	Interaction of electrons of nonpolar substances		1

[a]Bond energy is the amount of energy (Kcal/mol) needed to separate two bonded or interacting atoms under physiological conditions.

- A *double bond* involves the sharing of four electrons (two pairs; C═C).
- *Triple bonds*—six shared electrons—are rare, but there is one in nitrogen gas (N≡N), which is the major component of the air we breathe.

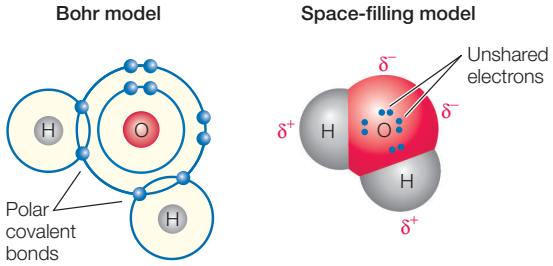
UNEQUAL SHARING OF ELECTRONS If two atoms of the same element are covalently bonded, there is an equal sharing of the pair(s) of electrons in their outermost shells. However, when the two atoms are different elements, the sharing is not necessarily equal. One nucleus may exert a greater attractive force on the electron pair than the other nucleus, so that the pair tends to be closer to that atom.

The attractive force that an atomic nucleus exerts on electrons in a covalent bond is called its **electronegativity**. The electronegativity of a nucleus depends on how many positive charges it has (nuclei with more protons are more positive and thus more attractive to electrons) and on the distance between the electrons in the bond and the nucleus (the closer the electrons, the greater the electronegative pull). **TABLE 2.3** shows the electronegativities of some elements important in biological systems. (Electronegativity is calculated to produce a dimensionless quantity, meaning that it has no unit of measurement such as for length, time, mass, etc.)

TABLE 2.2 Covalent Bonding Capacities of Some Biologically Important Elements

Element	Usual number of covalent bonds
Hydrogen (H)	1
Oxygen (O)	2
Sulfur (S)	2
Nitrogen (N)	3
Carbon (C)	4
Phosphorus (P)	5

If two atoms are 0.5 or less apart in electronegativity, they will share electrons equally in what is called a nonpolar covalent bond. Two oxygen atoms, for example, each with an electronegativity of 3.4, will share electrons equally. So will two hydrogen atoms (each with an electronegativity of 2.2). But when hydrogen bonds with oxygen to form water, the electrons involved are *unequally shared*: they tend to be nearer to the oxygen nucleus because it is more electronegative than hydrogen. When electrons are drawn to one nucleus more than to the other, the result is a **polar covalent bond**:



Because of this unequal sharing of electrons, the oxygen end of the bond has a slightly negative charge (symbolized by δ^- and spoken of as "delta negative," meaning a partial unit of charge), and the hydrogen end has a slightly positive charge (δ^+). The bond is *polar* because these opposite charges are separated at the two ends, or poles, of the bond. The partial charges that result from polar covalent bonds produce polar molecules or

TABLE 2.3 Some Electronegativities

Element	Electronegativity
Oxygen (O)	3.4
Chlorine (Cl)	3.2
Nitrogen (N)	3.0
Carbon (C)	2.6
Phosphorus (P)	2.2
Hydrogen (H)	2.2
Sodium (Na)	0.9
Potassium (K)	0.8

polar regions of large molecules. Polar bonds within molecules greatly influence the interactions they have with other polar molecules. The polarity of the water molecule has significant effects on its physical properties and chemical reactivity, as we will see in later chapters.

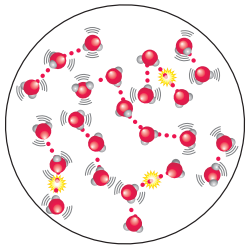
Hydrogen bonds may form within or between molecules with polar covalent bonds

In liquid water, the negatively charged oxygen (δ^-) atom of one water molecule is attracted to the positively charged hydrogen (δ^+) atoms of other water molecules (**FIGURE 2.4A**). The bond resulting from this attraction is called a **hydrogen bond** (see Table 2.1). These bonds are not restricted to water molecules. A hydrogen bond may also form between a strongly electronegative atom and a hydrogen atom that is covalently bonded to another electronegative atom (oxygen or nitrogen), as shown in **FIGURE 2.4B**.

A hydrogen bond is much weaker than a covalent bond (see Table 2.1). Although individual hydrogen bonds are weak, many of them can form within one molecule or between two molecules. In these cases, the hydrogen bonds together have considerable strength and can greatly influence the structure and properties of the substances. Hydrogen bonds play important roles in determining and maintaining the three-dimensional shapes of giant molecules such as DNA and proteins (see Chapter 3).

Hydrogen bonding also contributes to several properties of water that have great significance for life. As we will discuss later in this Concept, water plays a vital role in living systems as a **solvent**: a liquid in which other molecules dissolve. Other important properties of water are its heat capacity, cohesion, adhesion, and surface tension.

HEAT CAPACITY At any given time in liquid water, a water molecule forms an average of 3.4 hydrogen bonds (dotted red lines below) with other water molecules:



Liquid water

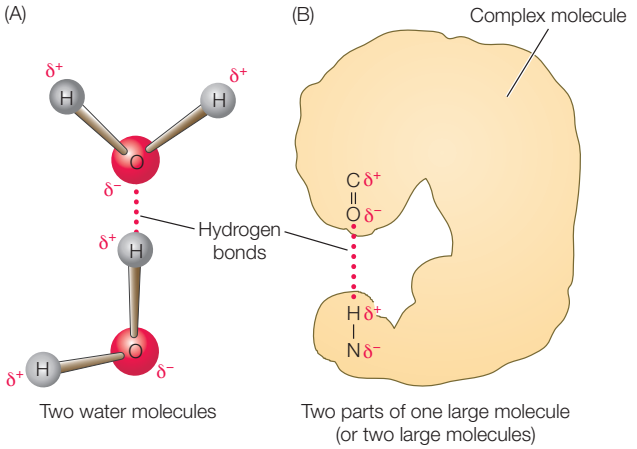


FIGURE 2.4 Hydrogen Bonds Can Form between or within Molecules (A) A hydrogen bond forms between two molecules because of the attraction between an atom with a partial negative charge on one molecule and a hydrogen with a partial positive charge on a second molecule. (B) Hydrogen bonds can form between different parts of the same large molecule.

These multiple hydrogen bonds contribute to the high **heat capacity** of water. Raising the temperature of liquid water takes a lot of heat, because much of the heat energy is used to break the hydrogen bonds that hold the liquid together (indicated by the yellow energy bursts in the liquid water diagram in the previous column). Think of what happens when you apply heat to a pan of water on the stove: it takes a while for the water to begin boiling. The same happens with a living organism—the large amount of water in living tissues shields the organism from fluctuations in environmental temperature.

Hydrogen bonding also gives water a high **heat of vaporization**. This means that a lot of heat is required to change water from its liquid to its gaseous state, in the process of evaporation. Once again, much of the heat energy is used to break the many hydrogen bonds between the water molecules. This heat must be absorbed from the environment in contact with the water. Evaporation thus has a cooling effect on the environment—whether a leaf, a forest, or an entire land mass. This effect explains why sweating cools the human body: as sweat evaporates from the skin, it absorbs some of the adjacent body heat.

LINK

Evaporation is important in the physiology of both plants and animals; see **Concepts 25.3 and 29.4**

COHESION, ADHESION, AND SURFACE TENSION The numerous hydrogen bonds that give water a high heat capacity and high heat of vaporization also explain the cohesive strength of liquid water. This cohesive strength, or **cohesion**, is defined as the capacity of water molecules to resist coming apart from one another when placed under tension. Hydrogen bonding between liquid water molecules and the solid surfaces surrounding them also allows for **adhesion** between the water and the solid surfaces. Together, cohesion and adhesion permit

narrow columns of liquid water to move from the roots to the leaves of tall trees (see Concept 25.3). When water evaporates from the leaves, the entire column moves upward in response to the pull of the molecules at the top.

The surface of liquid water exposed to air is difficult to puncture because the water molecules at the surface are hydrogen-bonded to other molecules below them. This **surface tension** permits a container to be filled with water above its rim without overflowing, and it permits spiders to walk on the surface of a pond.

Polar and nonpolar substances: Each interacts best with its own kind

Just as water molecules can interact with one another through hydrogen bonds, any polar molecule can interact with any other polar molecule through the weak (δ^+ to δ^-) attractions of hydrogen bonds. Polar molecules interact with water in this way and are called **hydrophilic** ("water-loving"). In aqueous (watery) solutions, these molecules become separated and surrounded by water molecules (FIGURE 2.5A). Thus, water functions as a solvent for polar molecules.

Because they do not have partial charges, nonpolar molecules do not interact with water in this way. Instead, these molecules tend to aggregate with one another. This allows more hydrogen bonds to form between water molecules. Therefore, nonpolar molecules are known as **hydrophobic** ("water-hating"), and the interactions between them are called hydrophobic interactions (FIGURE 2.5B). Hydrophobic substances do not really "hate" water—they can form weak interactions with it, since the electronegativities of the atoms in many nonpolar bonds (e.g., C—H bonds) are not exactly the same. But these interactions are far weaker than the hydrogen bonds between the water molecules, so the nonpolar substances tend to aggregate.

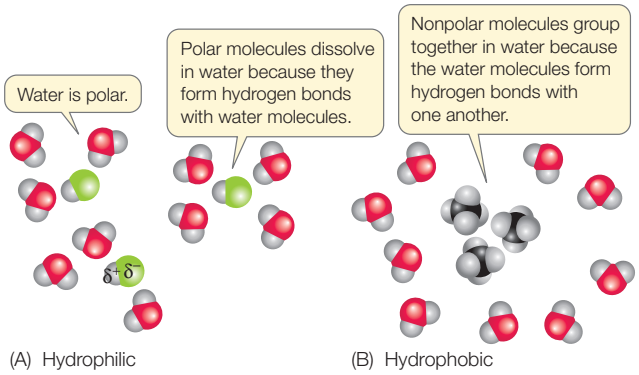


FIGURE 2.5 Hydrophilic and Hydrophobic (A) Molecules with polar covalent bonds are attracted to polar water (they are hydrophilic). (B) Molecules with nonpolar covalent bonds show greater attraction to one another than to water (they are hydrophobic). The color convention in the models shown here (gray, H; red, O; black, C; green, F) is often used.

APPLY THE CONCEPT

Atoms interact and form molecules

The concepts of chemical bonding and electronegativity (see Table 2.2) allow us to predict whether a molecule will be polar or nonpolar, and how it will interact with water. Typically, a difference in electronegativity greater than 0.5 will result in polarity. For each of the bonds below, indicate:

1. Whether the bond is polar or nonpolar
2. If polar, which is the δ^+ end
3. How a molecule with the bond will interact with water (hydrophilic or hydrophobic)

N—H	C—H	C═O	C—N
O—H	C—C	H—H	O—P

Ionic attractions form between anions and cations

When one interacting atom is much more electronegative than the other (see Table 2.3), a complete transfer of one or more electrons may occur. Consider sodium (11 protons) and chlorine (17 protons). A sodium atom has only one electron in its outermost shell; this condition is unstable. A chlorine atom has seven electrons in its outermost shell—another unstable condition. The most straightforward way for both atoms to achieve stability is to transfer an electron from sodium's outermost shell to that of chlorine (FIGURE 2.6). This reaction makes the two atoms more stable because they both have eight electrons in their outer shells. The result is two ions.

An **ion** is an electrically charged particle that forms when an atom gains or loses one or more electrons:

- The sodium ion (Na^+) in our example has a charge of +1 because it has one less electron than it has protons. The outermost electron shell of the sodium ion is stable because it has eight electrons. Positively charged ions are called **cations**.
- The chloride ion (Cl^-) has a charge of –1 because it has one more electron than it has protons. This additional electron gives Cl^- a stable outermost shell with eight electrons. Negatively charged ions are called **anions**.

Ionic attractions form as a result of the electrical attraction between ions bearing opposite charges. Ionic attractions result in stable crystalline structures that are often referred to as ionic compounds or salts. An example is sodium chloride (NaCl; table salt), where cations and anions are held together by ionic attractions. Ionic attractions in salt crystals may be stronger, but attractions between ions in solution, as occur in living systems, are typically weak (see Table 2.1).

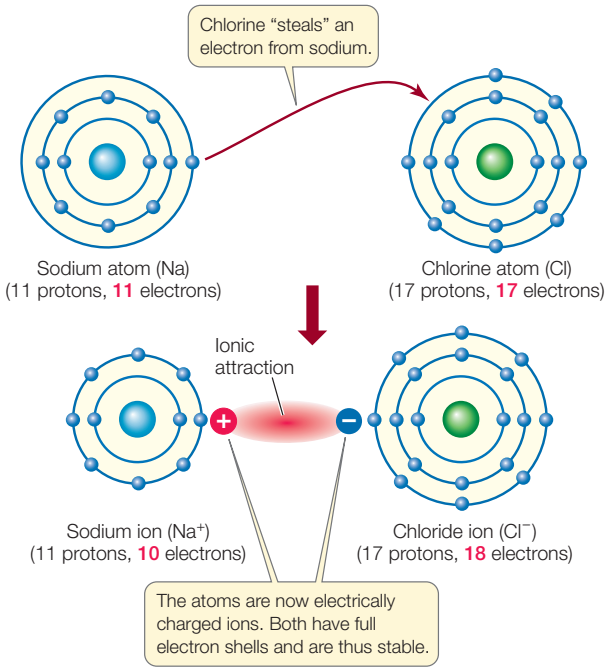
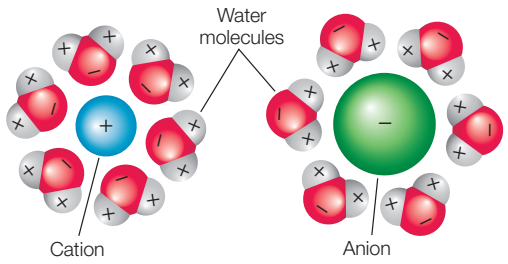


FIGURE 2.6 Ionic Attraction between Sodium and Chlorine
When a sodium atom reacts with a chlorine atom, the chlorine fills its outermost shell by "stealing" an electron from the sodium. In so doing, the chlorine atom becomes a negatively charged chloride ion (Cl^-). With one less electron, the sodium atom becomes a positively charged sodium ion (Na^+).

Given that living organisms consist of about 70 percent water, most biological processes occur in the presence of water. Because ionic attractions are weak, salts dissolve in water; the ions separate from one another and become surrounded by water molecules. The water molecules are oriented with their negative poles nearest to the cations and their positive poles nearest to the anions:



Functional groups confer specific properties to biological molecules

Certain small groups of atoms, called **functional groups**, are consistently found together in very different biological molecules. You will encounter several functional groups repeatedly in your study of biology (**FIGURE 2.7**). Each functional group has specific chemical properties (for example, polarity), and when attached to a larger molecule, it gives those properties to the larger molecule. The consistent chemical behavior of functional groups helps us understand the properties of the molecules that contain them.

Biological molecules often contain many different functional groups. A single large protein may contain hydrophobic, polar, and charged functional groups. Each group gives a different

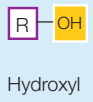
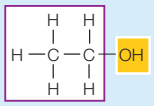
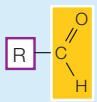
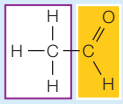
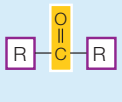
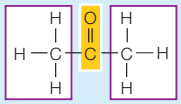
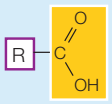
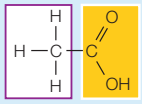
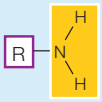
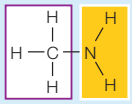
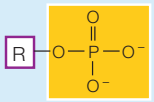
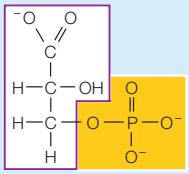
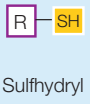
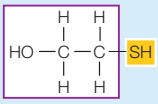
Functional group	Class of compounds and an example	Properties
 Hydroxyl	Alcohols  Ethanol	Polar. Hydrogen bonds with water to help dissolve molecules. Enables linkage to other molecules by condensation (see Figure 2.8).
 Aldehyde	Aldehydes  Acetaldehyde	C═O group is very reactive. Important in building molecules and in energy-releasing reactions.
 Keto	Ketones  Acetone	C═O group is important in carbohydrates and in energy reactions.
 Carboxyl	Carboxylic acids  Acetic acid	Acidic. Ionizes in living tissues to form $-COO^-$ and H^+. Enters into condensation reactions by giving up $-OH$. Some carboxylic acids important in energy-releasing reactions.
 Amino	Amines  Methylamine	Basic. Accepts H^+ in living tissues to form $-NH_3^+$. Enters into condensation reactions by giving up H.
 Phosphate	Organic phosphates  3-Phosphoglycerate	Acidic. Enters into condensation reactions by giving up $-OH$. When bonded to another phosphate, hydrolysis releases much energy.
 Sulfhydryl	Thiols  Mercaptoethanol	By giving up H, two $-SH$ groups can react to form a disulfide bridge, thus stabilizing protein structure.

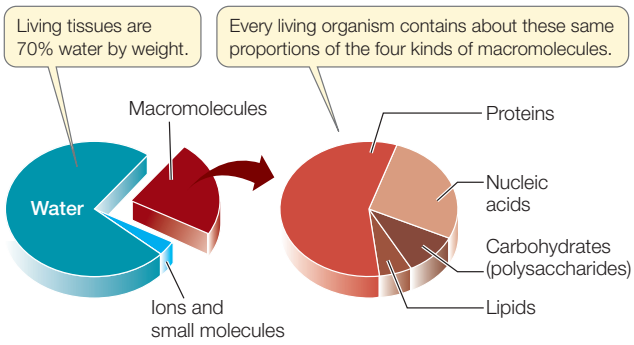
FIGURE 2.7 Functional Groups Important to Living Systems
Highlighted in yellow are the seven functional groups most commonly found in biological molecules. "R" represents the rest of the molecule.

specific property to its local site on the protein, and it may interact with another functional group on the same protein or with another molecule. The functional groups thus determine molecular shape and reactivity.

Go to ACTIVITY 2.1 Functional Groups
PoL2e.com/ac2.1

Macromolecules are formed by the polymerization of smaller molecules

Large molecules, called **macromolecules**, are formed by covalent linkages between smaller molecules. Four kinds of macromolecules are characteristic of living things: proteins, carbohydrates, nucleic acids, and lipids.



With the exception of small carbohydrates and lipids, these biological molecules are **polymers** (*poly*, "many"; *mer*, "unit") constructed by the covalent bonding of smaller molecules called **monomers**.

- *Proteins* are formed from different combinations of 20 amino acids, all of which share chemical similarities.
- *Carbohydrates* can be giant molecules, and are formed by linking together chemically similar sugar monomers (monosaccharides) to form polysaccharides.
- *Nucleic acids* are formed from four kinds of nucleotide monomers linked together in long chains.
- *Lipids* also form large structures from a limited set of smaller molecules, but in this case noncovalent forces maintain the interactions between the lipid monomers.

Polymers are formed and broken down by two types of reactions involving water (**FIGURE 2.8**):

- In **condensation**, the removal of water links monomers together.
- In **hydrolysis**, the addition of water breaks a polymer into monomers.

How the macromolecules function and interact with other molecules depends on the properties of the functional groups in their monomers.

Go to ANIMATED TUTORIAL 2.2
Macromolecules: Carbohydrates and Lipids
PoL2e.com/at2.2

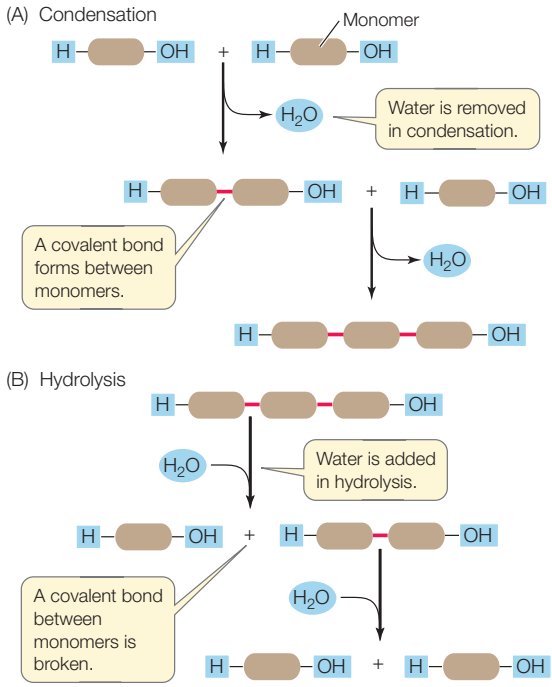
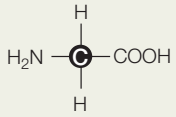


FIGURE 2.8 Condensation and Hydrolysis of Polymers
(A) Condensation reactions link monomers into polymers and produce water. (B) Hydrolysis reactions break polymers into individual monomers and consume water.

CHECKpoint CONCEPT 2.2

- ✓ How do differences in electronegativity result in the unequal sharing of electrons in polar molecules?
- ✓ Some functional groups (see Figure 2.7) can either donate or accept hydrogen bonds with other molecules, acting either as a donor (like the oxygen in a water molecule) or as an acceptor (like the hydrogens in a water molecule). For each of the following, is it an H-bond donor, acceptor, both, or neither?
  a. Aldehyde
  b. Amino
  c. Hydroxyl
- ✓ Here is the structure of the molecule glycine:



  a. What are the functional groups on this molecule? What is the R group to which they are attached? Is the R group hydrophilic or hydrophobic? Explain.
  b. Draw two glycine molecules and show how they can be linked by a condensation reaction.
- ✓ The boiling point (the temperature at which a liquid vaporizes) of water (H_2O) is 100°C, whereas the boiling point of methane (CH_4) is −161°C. Explain this difference in terms of hydrogen bonding between molecules.

We will begin our discussion of the molecules of life with carbohydrates, as they exemplify many of the chemical principles we have outlined so far.

CONCEPT 2.3 Carbohydrates Consist of Sugar Molecules

Carbohydrates are a large group of molecules that all have similar atomic compositions but differ greatly in size, chemical properties, and biological functions. Carbohydrates usually have the general formula $C_mH_{2n}O_n$, where m and n represent numbers. This makes them appear to be hydrates of carbon [associations between water molecules and carbon in the ratio $C_m(H_2O)_n$], hence their name. However, carbohydrates are not really "hydrates" because the water molecules are not intact. Rather, the linked carbon atoms are bonded with hydrogen atoms (—H) and hydroxyl groups (—OH), the components of water. Carbohydrates have four major biochemical roles:

- They are a source of stored energy that can be released in a form that organisms can use.
- They are used to transport stored energy within complex organisms.
- They function as structural molecules that give many organisms their shapes.
- They serve as recognition or signaling molecules that can trigger specific biological responses.

Some carbohydrates are relatively small, such as the simple sugars (for example, glucose) that are the primary energy source for many organisms. Others are large polymers of simple sugars; an example is starch, which is stored in seeds.

Monosaccharides are simple sugars

Monosaccharides (*mono*, "one") are relatively simple molecules with up to seven carbon atoms. They differ in their arrangements of carbon, hydrogen, and oxygen atoms (FIGURE 2.9).

Pentoses (*pente*, "five") are five-carbon sugars. Two pentoses are of particular biological importance: the backbones of the nucleic acids RNA and DNA contain ribose and deoxyribose, respectively.

LINK

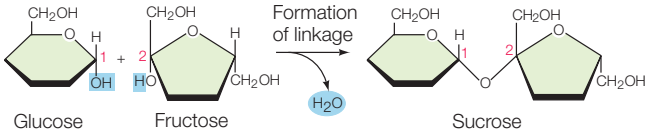
For a description of the nucleic acids RNA and DNA see **Concept 3.1**

The hexoses (*hex*, "six") all have the formula $C_6H_{12}O_6$. They include glucose, fructose (so named because it was first found in fruits), mannose, and galactose.

Go to ACTIVITY 2.2 Forms of Glucose
PoL2e.com/ac2.2

Glycosidic linkages bond monosaccharides

The disaccharides, oligosaccharides, and polysaccharides are all constructed from monosaccharides that are covalently bonded by condensation reactions that form **glycosidic linkages**. A single glycosidic linkage between two monosaccharides forms a **disaccharide**. For example, sucrose—common table sugar—is a major disaccharide formed in plants from a glucose molecule and a fructose molecule (see Figure 2.9 for complete structures):



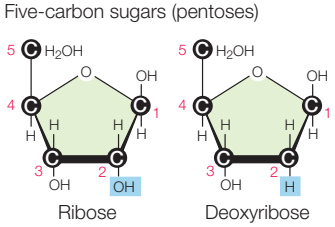
Another disaccharide is maltose, which is formed from two glucose units and is a product of starch digestion. Maltose is an important carbohydrate for making beer.

Oligosaccharides contain several monosaccharides bound together by glycosidic linkages. Many oligosaccharides have additional functional groups, which give them special properties. Oligosaccharides are often covalently bonded to proteins and lipids on the outer surfaces of cells, where they serve as recognition signals. For example, the different human blood groups (the ABO blood types) get their specificity from oligosaccharide chains.

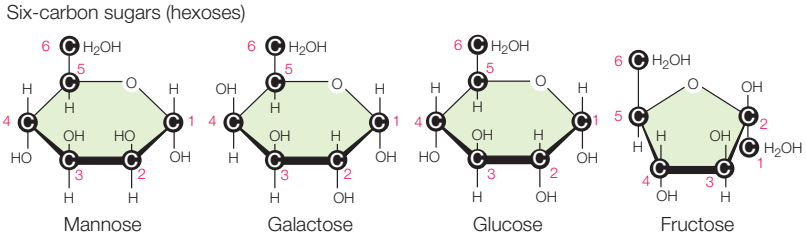
Polysaccharides store energy and provide structural materials

Polysaccharides are large polymers of monosaccharides connected by glycosidic linkages (FIGURE 2.10). Polysaccharides are not necessarily linear chains of monomers. Each monomer unit has several sites that are capable of forming glycosidic linkages, and thus branched molecules are possible.

FIGURE 2.9 Monosaccharides Monosaccharides are made up of varying numbers of carbons. Many have the same kind and number of atoms, but the atoms are arranged differently.



Ribose and deoxyribose each have five carbons, but very different chemical properties and biological roles.



These hexoses all have the formula $C_6H_{12}O_6$, but each has distinct biochemical properties.

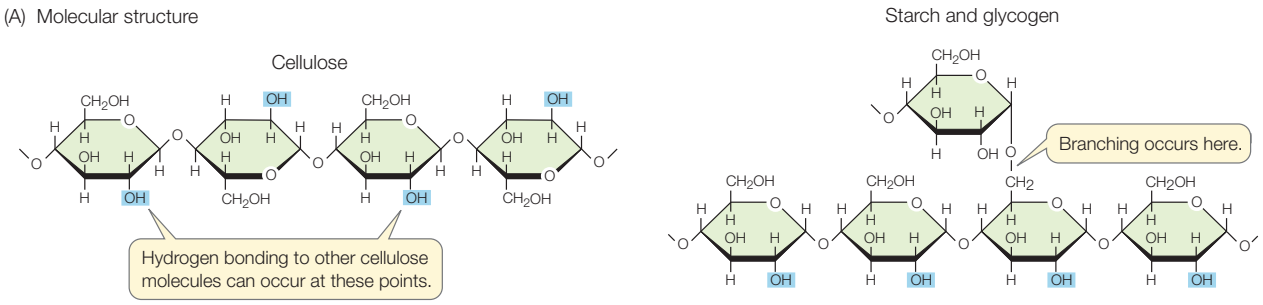


FIGURE 2.10 Polysaccharides Cellulose, starch, and glycogen are all composed of long chains of glucose but with different levels of branching and compaction.

Starches comprise a family of giant molecules that are all polysaccharides of glucose. The different starches can be distinguished by the amount of branching in their polymers. Starch is the principal energy storage compound of plants.

Glycogen is a water-insoluble, highly branched polymer of glucose that is the major energy storage molecule in mammals. It is produced in the liver and transported to the muscles. Both glycogen and starch are readily hydrolyzed into glucose monomers, which in turn can be broken down to liberate their stored energy.

If glucose is the major source of fuel, why store it in the form of starch or glycogen? The reason is that 1,000 glucose molecules would exert 1,000 times the osmotic pressure of a single glycogen molecule, causing water to enter the cells (see Concept 5.2). If it were not for polysaccharides, many organisms would expend a lot of energy expelling excess water from their cells.

As the predominant component of plant cell walls, cellulose is by far the most abundant carbon-containing (organic) compound on Earth. Like starch and glycogen, cellulose is a polysaccharide of glucose, but its glycosidic linkages are arranged in such a way that it is a much more stable molecule. Whereas starch is easily broken down by chemicals or enzymes to supply glucose for energy-producing reactions, cellulose is an excellent structural material that can withstand harsh environmental conditions without substantial change.

LINK

Most animals cannot digest (hydrolyze) cellulose; **Chapter 30** describes adaptations in some animals to use cellulose as an energy source

CHECKpoint CONCEPT 2.3

- ✓ Draw the chemical structure of a disaccharide formed from two glucose monosaccharides.
- ✓ Examine the glucose molecule shown in Figure 2.9. What are the functional groups on the molecule?
- ✓ Notice the large number of hydrogen bonding groups present in the linear structure of cellulose (see Figure 2.10A). Why is this structure so strong?
- ✓ Some sugars have other functional groups in addition to those typically present. Draw the structure of the amino sugar glucosamine, which has an amino group bonded at carbon 2 of glucose. Would this molecule be more or less polar than glucose? Explain why.

We have seen that carbohydrate structure is an example of the monomer–polymer theme in biology. Now we will turn to lipids, which are unusual among the four classes of biological macromolecules in that they are not, strictly speaking, polymers.

CONCEPT 2.4 Lipids Are Hydrophobic Molecules

Lipids—commonly called fats and oils—are hydrocarbons (composed of C and H atoms) that are insoluble in water because of their many nonpolar covalent bonds. As you have seen, nonpolar molecules are hydrophobic and preferentially aggregate together, away from polar water (see Figure 2.5). When nonpolar hydrocarbons are sufficiently close together, weak but additive **van der Waals interactions** (see Table 2.1) hold them together. The huge macromolecular aggregations that can form are not polymers in a strict chemical sense, because the individual lipid molecules are not covalently bonded.

Lipids play several roles in living organisms, including the following:

- They store energy in the C—C and C—H bonds.
- They play important structural roles in cell membranes and on body surfaces, largely because their nonpolar nature makes them essentially insoluble in water.
- Fat in animal bodies serves as thermal insulation.

Fats and oils are triglycerides

The most common units of lipids are **triglycerides**, also known as simple lipids. Triglycerides that are solid at room temperature (around 20°C) are called fats; those that are liquid at room temperature are called oils. A triglyceride contains three fatty acid molecules and one glycerol molecule. **Glycerol** is a small molecule with three hydroxyl (—OH) groups; thus it is an alcohol. A **fatty acid** consists of a long nonpolar hydrocarbon chain attached to the polar carboxyl (—COOH) group, and it is therefore a carboxylic acid. The long hydrocarbon chain is very hydrophobic because of its abundant C—H and C—C bonds.

Synthesis of a triglyceride involves three condensation reactions (FIGURE 2.11). The resulting molecule has very little polarity and is extremely hydrophobic. That is why fats and oils do not mix with water but float on top of it in separate globules or layers. The three fatty acids in a single triglyceride molecule need not all have the same hydrocarbon chain length or structure; some may be saturated fatty acids, while others may be unsaturated:

- In a **saturated fatty acid**, all the bonds between the carbon atoms in the hydrocarbon chain are single; there are no double bonds. That is, all the available bonds are saturated with hydrogen atoms (FIGURE 2.12A). These fatty acid molecules are relatively rigid and straight, and they pack together tightly, like pencils in a box.
- In an **unsaturated fatty acid**, the hydrocarbon chain contains one or more double bonds. Linoleic acid is an example of a polyunsaturated fatty acid that has two double bonds near the middle of the hydrocarbon chain, causing kinks in the chain (FIGURE 2.12B). Such kinks prevent the unsaturated molecules from packing together tightly.

The kinks in fatty acid molecules are important in determining the fluidity and melting point of the lipid. The triglycerides of animal fats tend to have many long-chain saturated fatty acids, which pack tightly together; these fats are usually solid at room temperature and have a high melting point. The triglycerides of plants, such as corn oil, tend to have short or

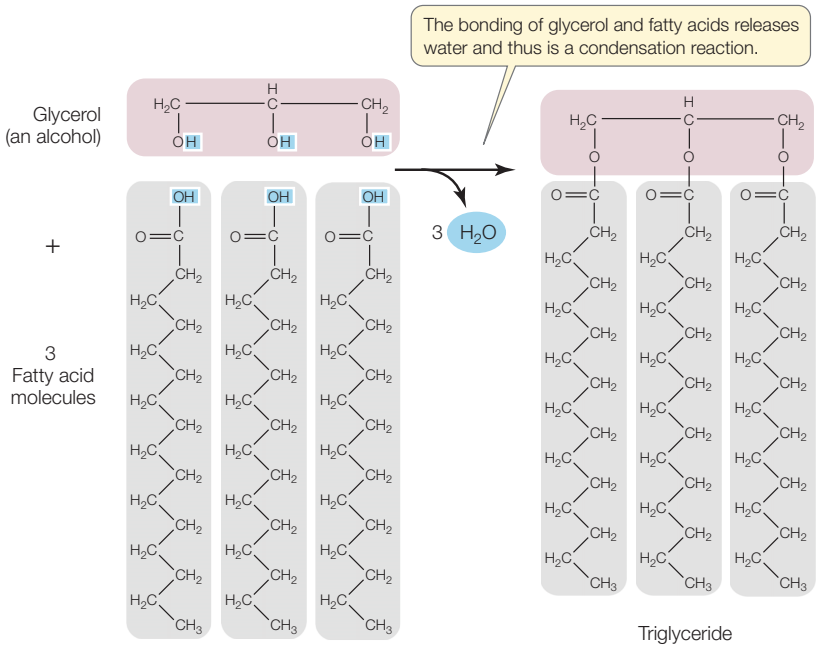


FIGURE 2.11 Synthesis of a Triglyceride In living things, the reaction that forms a triglyceride is more complex than the single step shown here.

unsaturated fatty acids. Because of their kinks, these fatty acids pack together poorly, have low melting points, and are usually liquid at room temperature.

Fats and oils are excellent storehouses for chemical energy. As you will see in Chapter 6, energy is released when these compounds are broken down into smaller molecules. The released energy can be used by an organism for other purposes, such as movement or the synthesis of complex molecules. On a per weight basis, broken-down lipids yield more than twice as much energy as degraded carbohydrates.

Phospholipids form biological membranes

We have mentioned the hydrophobic nature of the many C—C and C—H bonds in a fatty acid. But what about the carboxyl functional group at the end of the molecule? When it ionizes and forms COO^-, it is strongly hydrophilic. So a fatty acid has two opposing chemical properties: a hydrophilic end and a long hydrophobic tail. A molecule that is partly hydrophilic and partly hydrophobic is said to be **amphipathic**.

In triglycerides, a glycerol molecule is bonded to three fatty acid chains and the resulting molecule is entirely hydrophobic. **Phospholipids** are like triglycerides in that they contain fatty acids bound to glycerol. However, in phospholipids, a phosphate-containing compound replaces one of the fatty acids, giving these molecules amphipathic properties (FIGURE 2.13A). The phosphate functional group (there are several different kinds in different phospholipids) has a negative electric charge, so this portion of the molecule is hydrophilic, attracting polar water molecules. But the two fatty acids are hydrophobic, so they tend to avoid water and aggregate together or with other hydrophobic substances.

In an aqueous environment, phospholipids line up in such a way that the nonpolar, hydrophobic "tails" pack tightly together and the phosphate-containing "heads" face outward, where they interact with water. The phospholipids thus form a **bilayer**: a sheet two molecules thick, with water excluded from the core (FIGURE 2.13B). Although no covalent bonds link individual lipids in these large aggregations, such stable aggregations form readily in aqueous conditions. Biological membranes have this kind of **phospholipid bilayer** structure, and we will devote Chapter 5 to their biological functions.

CHECKpoint CONCEPT 2.4

- ✓ What is the difference between fats and oils?
- ✓ Why are phospholipids amphipathic, and how does this result in a lipid bilayer membrane?
- ✓ If fatty acids are carefully put onto the surface of water, they form a single molecular layer. If the mixture is then shaken vigorously, the fatty acids will form round structures called micelles. Explain these observations.

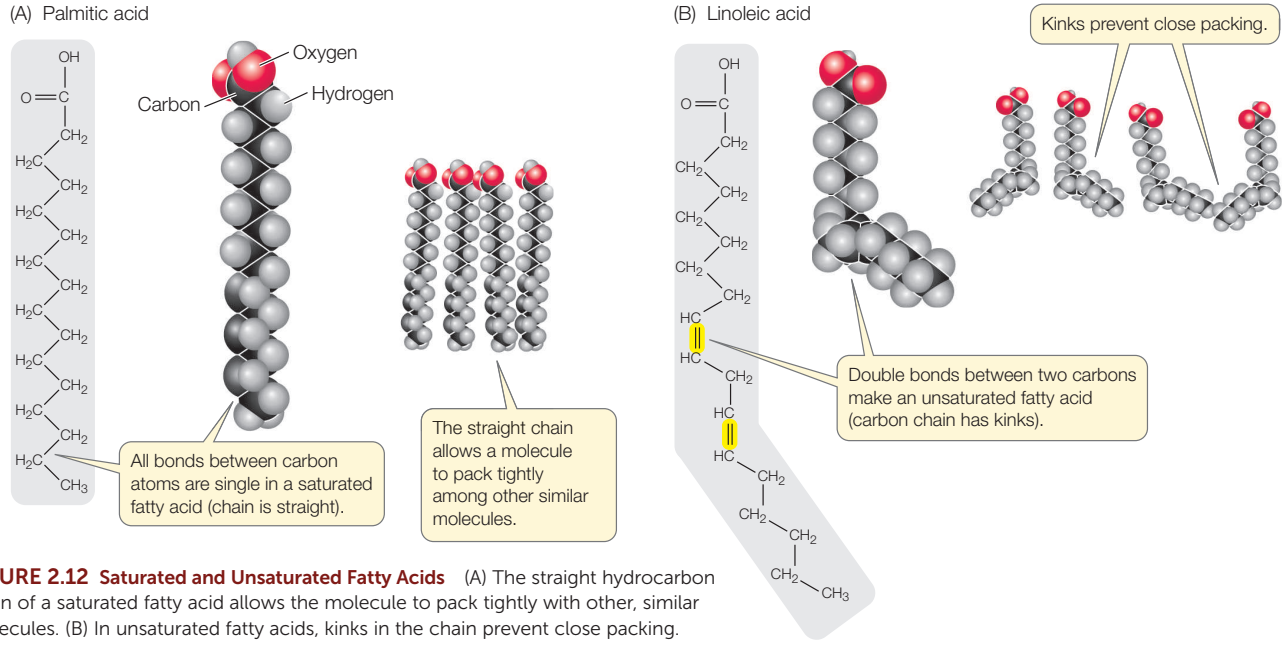


FIGURE 2.12 Saturated and Unsaturated Fatty Acids (A) The straight hydrocarbon chain of a saturated fatty acid allows the molecule to pack tightly with other, similar molecules. (B) In unsaturated fatty acids, kinks in the chain prevent close packing.

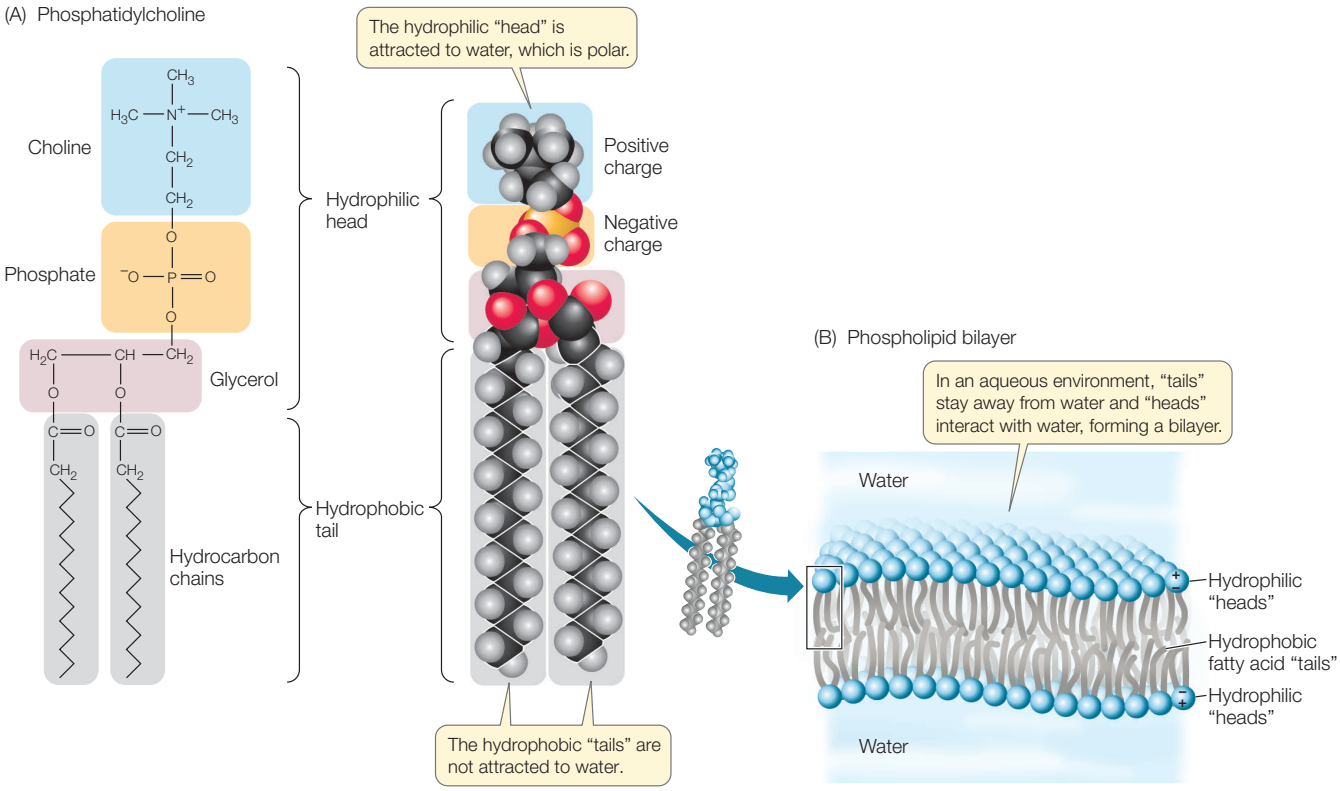


FIGURE 2.13 Phospholipids (A) Phosphatidylcholine (lecithin) is an example of a phospholipid molecule. In other phospholipids, the amino acid serine, the sugar alcohol inositol, or another compound replaces choline. (B) In an aqueous environment, hydrophobic interactions bring the "tails" of phospholipids together in the interior of a bilayer. The hydrophilic "heads" face outward on both sides of the bilayer, where they interact with the surrounding water molecules.

Molecules such as carbohydrates and lipids are not always stable in living systems. Rather, a hallmark of life is its *ability to transform molecules*. This involves making and breaking covalent bonds, as some atoms are removed and others are attached. As part of our introduction to biochemical concepts, we will now turn to these processes of chemical change.

CONCEPT 2.5 Biochemical Changes Involve Energy

A **chemical reaction** occurs when atoms have sufficient energy to combine, or to change their bonding partners. Consider the hydrolysis of the disaccharide sucrose to its component monomers, glucose and fructose (see p. 27 for the chemical structures). We can express this reaction using a chemical equation:

$$\underset{(C_{12}H_{22}O_{11})}{\text{Sucrose}} + H_2O \rightarrow \underset{(C_6H_{12}O_6)}{\text{glucose}} + \underset{(C_6H_{12}O_6)}{\text{fructose}}$$

In this equation, sucrose and water are the **reactants**, and glucose and fructose are the **products**. The reaction proceeds as some bonds in the reactants are broken and new bonds form to make the products. Electrons and protons are transferred from one reactant to the other to form the products. The products of this reaction have different properties from those of the reactants. Chemical reactions involve *changes in energy*; for example, the energy contained in the chemical bonds of sucrose and water (the reactants) is greater than the energy in the bonds of the two products, glucose and fructose.

What is energy? Physicists define it as the capacity to do work, which occurs when a force operates on an object over a distance. But in biochemistry, it is more useful to consider energy as *the capacity for change*. In biochemical reactions, energy changes are usually associated with changes in the chemical composition and properties of molecules. Energy comes in many forms: chemical, electrical, heat, light, and mechanical. But all forms of energy can be considered as one of two basic types:

- **Potential energy** is the energy of state or position—that is, stored energy. It can be stored in many forms: in chemical bonds, as a concentration gradient, or even as an electric charge imbalance.
- **Kinetic energy** is the energy of movement—that is, the type of energy that does work, that makes things change. For example, heat causes molecular motions and can even break chemical bonds.

Potential energy can be converted into kinetic energy and vice versa, and the form that the energy takes can be converted.

Think of reading this book: light energy is converted to chemical energy in your eyes, and then is converted to electrical energy in the nerve cells that carry messages to your brain. When you decide to turn a page, the electrical and chemical energy of nerves and muscles are converted to mechanical energy for movement of your hand and arm.

Metabolism involves reactions that store and release energy

The sum total of all the chemical reactions occurring in a biological system at a given time is called **metabolism**. Metabolic reactions involve energy changes, in which energy is either stored in, or released from, chemical bonds. In general, the formation of a bond releases energy, whereas the breaking of a bond requires an input of energy. More energy is released in the formation of a stronger, more stable (lower energy) bond than in the formation of a less stable (higher energy) one. Conversely, the breaking of a stronger bond consumes more energy than the breaking of a weaker bond. A chemical reaction will occur spontaneously if the total energy consumed by breaking bonds in the reactants is *less than* the total energy released by forming bonds in the products.

Anabolic reactions (collectively called anabolism) link simple molecules to form more complex molecules. Anabolic reactions require an input of energy because strong bonds within the smaller molecules must be broken to form the more complex molecules. For example, the formation of sucrose requires the breaking of strong O—H bonds in glucose and fructose. Reactions that require an input of energy are called endergonic or endothermic (FIGURE 2.14A). The energy used in anabolic reactions is stored in the newly formed (higher energy) chemical bonds.

Catabolic reactions (collectively called catabolism) break down complex molecules into simpler ones and release the energy that was used to make the complex molecules. Chemists call such reactions exergonic or exothermic (FIGURE 2.14B). For example, when sucrose is hydrolyzed, energy is released by the formation of more stable (lower energy) bonds within the monosaccharides.

Catabolic and anabolic reactions are often linked. The energy released in catabolic reactions is often used to drive anabolic reactions—that is, to do biological work. For example, the energy released by the breakdown of glucose (catabolism) is used to drive anabolic reactions such as the synthesis of triglycerides. That is why fat accumulates if you eat food in excess of your energy needs.

Biochemical changes obey physical laws

Recall from the opening of this chapter that we described the mechanistic view of life, whereby living systems obey the same rules that govern the nonliving world. The **laws of thermodynamics** (thermo, "energy"; dynamics, "change") were derived from studies of the fundamental properties of energy, and the ways energy interacts with matter. These laws apply to all

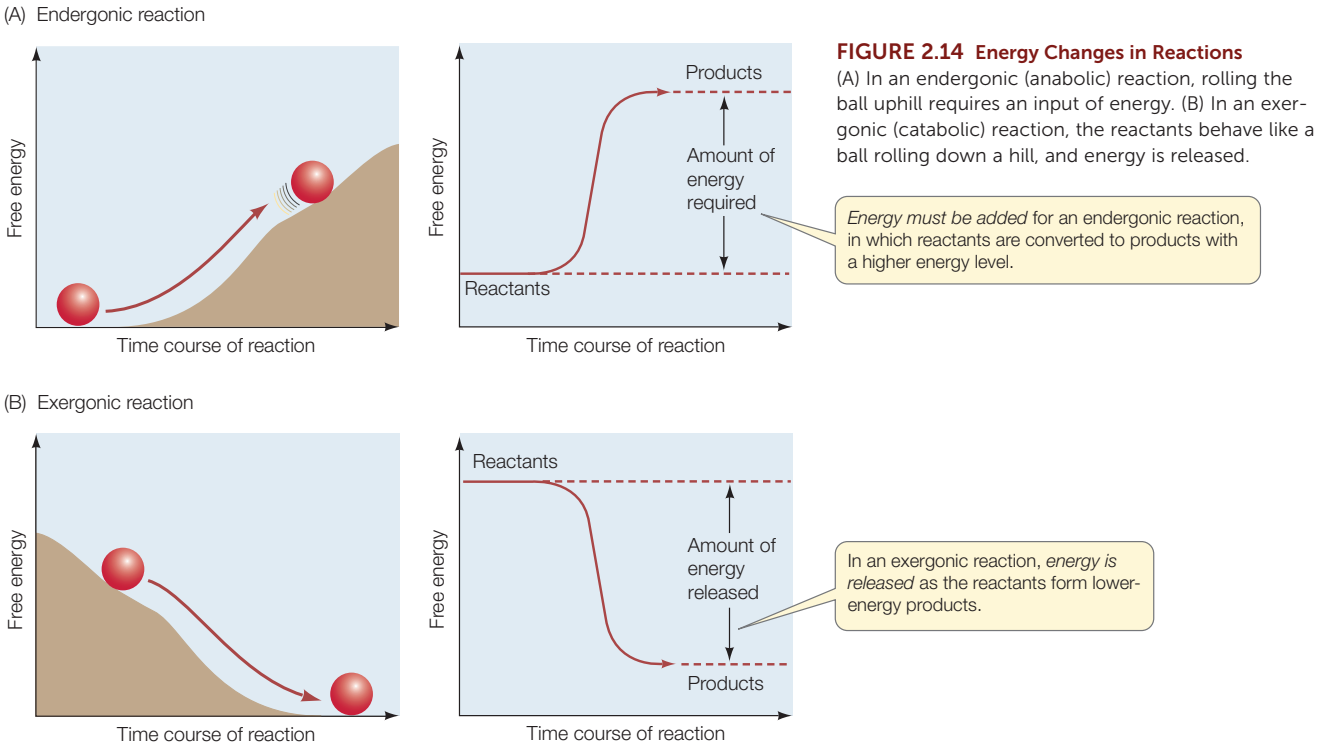
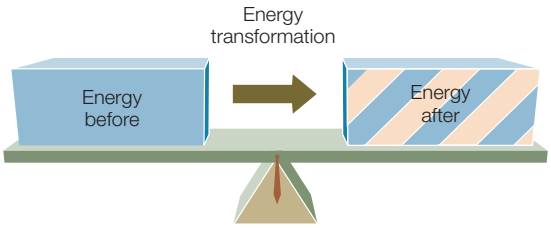


FIGURE 2.14 Energy Changes in Reactions
(A) In an endergonic (anabolic) reaction, rolling the ball uphill requires an input of energy. (B) In an exergonic (catabolic) reaction, the reactants behave like a ball rolling down a hill, and energy is released.

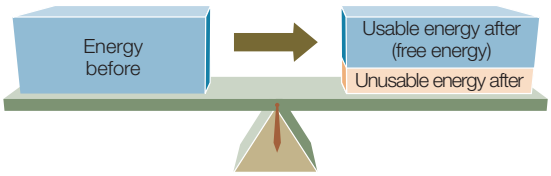
(A)

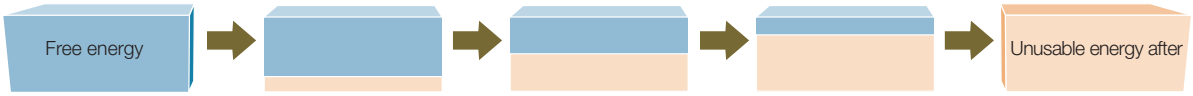
The First Law of Thermodynamics
The total amount of energy before a transformation equals the total amount after a transformation. No new energy is created, and no energy is lost.



(B)

The Second Law of Thermodynamics
A transformation does not change the total amount of energy within a closed system (one that is not exchanging matter or energy with the surroundings), but after any transformation the amount of energy available to do work (called free energy) is always less than the original amount of energy.





Another statement of the second law is that in a closed system, with repeated energy transformations, free energy decreases and unusable energy (disorder) increases—a phenomenon known as the increase in **entropy**.

FIGURE 2.15 The Laws of Thermodynamics (A) The first law states that energy cannot be created or destroyed. (B) The second law states that after energy transformations, some energy becomes unavailable to do work.

matter and all energy transformations in the universe. Their application to living systems helps us understand how organisms and cells harvest and transform energy to sustain life.

The first law of thermodynamics: Energy is neither created nor destroyed. The first law of thermodynamics states that in any conversion, energy is neither created nor destroyed. Another way of stating this is that the total energy before and after an energy conversion is the same (**FIGURE 2.15A**). [Similarly, matter is also conserved: in the hydrolysis of sucrose (see p. 31), there are 12 carbons, 24 hydrogens, and 12 oxygens on both sides of the equation.]

Although the total amount of energy is conserved, chemical reactions involve changes in the amount of (potential) energy stored in chemical bonds. If energy is released during the reaction, it is available to do work—for example, to drive another chemical reaction. In general, reactions that release energy (catabolic, or exergonic reactions) can occur spontaneously.

The second law of thermodynamics: Useful energy tends to decrease. Although energy cannot be created or destroyed, the second law of thermodynamics implies that when energy is converted from one form to another, some of that energy becomes unavailable for doing work (**FIGURE 2.15B**). In other words, no physical process or chemical reaction is 100 percent efficient; some of the released energy is lost in a form associated with disorder. Think of disorder as a kind of randomness caused by the thermal motion of particles; this energy is so dispersed that it is unusable. **Entropy** is a measure of the disorder in a system.

If a chemical reaction increases entropy, its products are more disordered or random than its reactants. The disorder in a solution of glucose and fructose is greater than that in a solution of sucrose, where the glycosidic bond between the two monosaccharides prevents free movement. Conversely, if there are fewer products and they are more restrained in their movements than the reactants, the disorder is reduced. But this requires an energy input to achieve.

The second law of thermodynamics predicts that, as a result of energy transformations, disorder tends to increase; some energy is always lost to random thermal motion (entropy). Chemical changes, physical changes, and biological processes all tend to increase entropy (see Figure 2.15B), and this tendency gives direction to these processes. Changes in entropy are mathematically related to changes in free energy (the energy available to do work), and thus the second law helps explain why some reactions proceed in one direction rather than another.

How does the second law of thermodynamics apply to organisms? Consider the human body, with its highly organized tissues and organs composed of large, complex molecules. This level of complexity appears to be in conflict with the second law, but for two reasons, it is not. First, the construction of complex molecules also generates disorder. The anabolic reactions needed to construct 1 kg of an animal body require the catabolism of about 10 kg of food. So metabolism creates far more disorder (more energy is lost to entropy) than the amount of order stored in flesh. Second, life requires a constant input of energy

to maintain order. Without this energy, the complex structures of living systems would break down. Because energy is used to generate and maintain order, and biological processes cause an overall increase in entropy, there is no conflict with the second law of thermodynamics.

CHECKpoint CONCEPT 2.5

- ✓ When you eat a candy bar and then decide to go for a walk, energy transformations take place. Beginning with the food energy in the candy bar, describe the forms of energy used and the changes in energy that occur as you decide to walk and as you do the walking.
- ✓ What is the difference between anabolism and catabolism? Between endergonic and exergonic reactions?
- ✓ Predict whether these situations are endergonic or exergonic, and explain your reasoning:
  a. The formation of a phospholipid bilayer membrane
  b. Turning on a TV set

APPLY THE CONCEPT

Biochemical changes involve energy

Chemical reactions in living systems involve changes in energy. These can be expressed as changes in available energy, called free energy (designated G, for Gibbs—the scientist who first described this parameter). The overall direction of a spontaneous chemical reaction is from higher to lower free energy. In other words, if the $G_{reactants}$ is greater than the $G_{products}$ (negative ΔG; the Greek letter delta stands for "change in" or "difference"), the reaction will be spontaneous; it will tend to go in the direction from reactants to product and release free energy in the process. Reactions where the $G_{reactants}$ is less than the $G_{products}$ (positive ΔG) will occur only if additional free energy is supplied.

The table shows some reactions and the absolute values of their associated free energy changes, $|\Delta G|$ (the vertical lines indicate absolute value).

REACTION	REACTANTS	PRODUCTS	$\lvert\Delta G\rvert$
Hydrolysis of sucrose	Sucrose + H_2O	Glucose + fructose	7.0
Triglyceride attachment	Glycerol + fatty acid	Monoglyceride	3.5

1. For each reaction, would you expect ΔG to be positive or negative?
2. Which reactions will be spontaneous? Explain your answer.

Q Why is the search for water important in the search for life?

ANSWER You have seen throughout this chapter that water is essential for the chemistry of life. Water is composed of two of the most abundant elements (Concept 2.1). Water is a polar molecule (Concept 2.2), which allows biologically important polar molecules such as monosaccharides (Concept 2.3) to dissolve in water. Because they are hydrophobic, lipids interact with water to form important biological structures (Concept 2.4). Water molecules participate directly in the formation and breakdown of polymers (Concept 2.2). In short, all of the processes of life as we know it require water.

In the opening essay of this chapter, we described recent evidence for the presence of water on other bodies in our solar system. Could this water harbor life, now or in the past? One way to investigate this possibility is to study how life on Earth may have originated in an aqueous environment. Geological evidence suggests that Earth was formed about 4.5 billion years ago and that life arose about 3.8 billion years ago. During the time when life originated, there was apparently little oxygen gas (O_2) in the atmosphere. In the 1950s, Stanley Miller and Harold Urey at the University of Chicago set up an experimental "atmosphere" containing various gases thought to be present in Earth's early atmosphere. Among them were ammonia (NH_3), hydrogen (H_2), methane (CH_4), and water vapor (H_2O). Miller and Urey passed an electric spark over the mixture to simulate lightning, providing a source of energy for covalent bond formation. Then they cooled the system so the gases would condense and collect in a watery solution, or "ocean" (**FIGURE 2.16**). Note that water was essential for this experiment as a source of oxygen atoms.

After several days of continuous operation, the system contained numerous complex molecules, including amino acids, the building blocks of proteins. In later experiments the researchers added other gases, such as carbon dioxide (CO_2), nitrogen (N_2), and sulfur dioxide (SO_2). This resulted in the formation of functional groups such as carboxylic acids, fatty acids, and many three- to six-carbon sugars. Taken together, these data suggest a plausible mechanism for the formation of life's chemicals in the aqueous environment of early Earth.

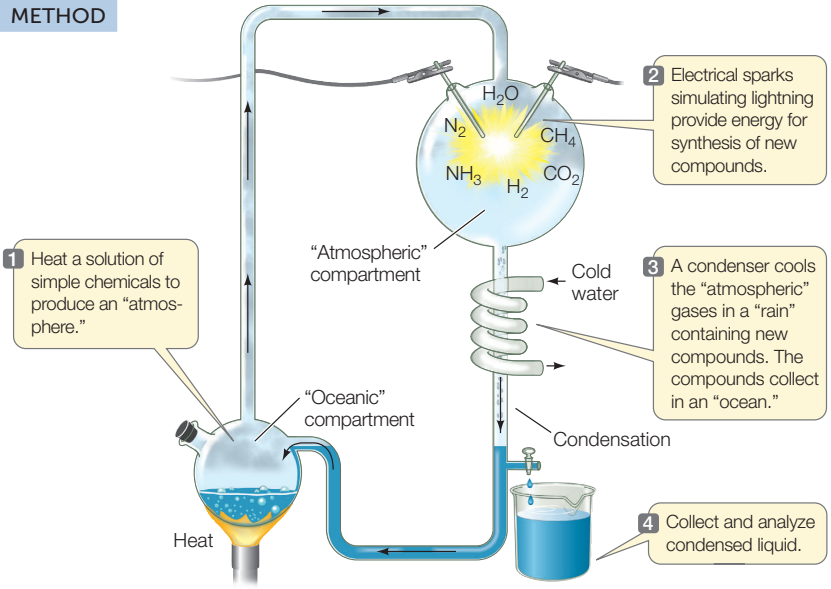
INVESTIGATION

FIGURE 2.16 Miller and Urey Synthesized Prebiotic Molecules in an Experimental Atmosphere With an increased understanding of the atmospheric conditions that existed on primitive Earth, the researchers devised an experiment to see if these conditions could lead to the formation of organic molecules.[a]

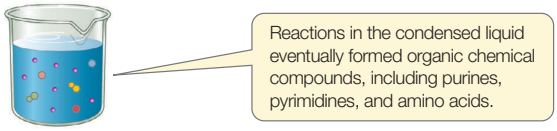
HYPOTHESIS

Organic chemical compounds can be generated under conditions similar to those that existed in the atmosphere of primitive Earth.

METHOD



RESULTS



CONCLUSION

The chemical building blocks of life could have been generated in the probable atmosphere of early Earth.

ANALYZE THE DATA

The following data show the amount of energy impinging on Earth in different forms.

Source	Energy (cal/[cm^2 × yr])
Total radiation from sun	260,000
Ultraviolet light	199
Wavelength <250 nm	570
Wavelength <200 nm	85
Wavelength <150 nm	3.5
Electric discharges	4
Cosmic rays	0.0015
Radioactivity	0.8
Volcanoes	0.13

A. Only a small fraction of the sun's energy is ultraviolet light (less than 250 nm). What is the rest of the solar energy?

B. The molecules CH_4, H_2O, NH_3, and CO_2 absorb light at wavelengths less than 200 nm. What fraction of total solar radiation is in this range?

C. Instead of electric discharges, what other sources of energy could be used in these experiments?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]S. L. Miller and H. C. Urey. 1959. *Science* 130: 245–251.

Go to ANIMATED TUTORIAL 2.3
Synthesis of Prebiotic Molecules
PoL2e.com/at2.3

SUMMARY

CONCEPT 2.1 Atomic Structure Is the Basis for Life's Chemistry

- Matter is composed of atoms. Each **atom** consists of a positively charged **nucleus** made up of **protons** and **neutrons**, surrounded by **electrons** bearing negative charges.
- The number of protons in the nucleus defines an **element**. There are many elements in the universe, but only a few of them (C, H, O, P, N, and S) make up the bulk of living organisms.
- Electrons are distributed in **electron shells** at varying energy levels away from the nucleus. The first shell can have a maximum of 2 electrons; the second shell, 8 electrons; the third shell, 18 electrons; and subsequent shells, 32 electrons. **Review Figure 2.1**

CONCEPT 2.2 Atoms Interact and Form Molecules

- A **chemical bond** is an attractive force that links two atoms together in a molecule. **Review ANIMATED TUTORIAL 2.1**
- A **covalent bond** is a strong bond formed when two atoms share one or more pairs of electrons. **Review Figures 2.2 and 2.3**
- When two atoms of unequal electronegativity bond with each other, a **polar covalent bond** is formed. The two ends, or poles, of the bond have partial charges (δ^+ or δ^-). A **hydrogen bond** is a weak electrical attraction that forms between a δ^+ hydrogen atom in one molecule and a δ^- atom in another molecule (or in another part of a large molecule). Hydrogen bonds are abundant in water. **Review Figure 2.4**
- **Ions** are electrically charged particles that form when atoms gain or lose one or more electrons in order to form more stable electron configurations. **Anions** and **cations** are negatively and positively charged ions, respectively. **Ionic attractions** form when ions with opposite charges attract. **Review Figure 2.6**
- **Functional groups** are covalently bonded groups of atoms that confer specific properties to biological molecules. **Review Figure 2.7**
- **Macromolecules** are formed by polymerization of smaller molecules call **monomers**. **Review Figure 2.8, ANIMATED TUTORIAL 2.2, and ACTIVITY 2.1**

CONCEPT 2.3 Carbohydrates Consist of Sugar Molecules

- **Carbohydrates** contain carbon bonded to hydrogen and oxygen.
- **Monosaccharides** include pentoses (with five carbons) and hexoses (with six carbons). **Review Figure 2.9 and ACTIVITY 2.2**
- Glycosidic linkages are covalent bonds between saccharides. **Disaccharides** such as sucrose each contain two monosaccharides, whereas **polysaccharides** such as starch and cellulose contain long chains of monomers. **Review Figure 2.10**

CONCEPT 2.4 Lipids Are Hydrophobic Molecules

- Fats and oils are **triglycerides**, composed of three **fatty acids** covalently linked to glycerol. **Review Figure 2.11**
- **Saturated fatty acids** have hydrocarbon chains with no double bonds. **Unsaturated fatty acids** contain double bonds in their hydrocarbon chains. **Review Figure 2.12**
- **Phospholipids** contain two fatty acids and a hydrophilic, phosphate-containing polar group attached to glycerol. They are **amphipathic**, with both polar and nonpolar ends. They form into a structural bilayer in water. **Review Figure 2.13**

CONCEPT 2.5 Biochemical Changes Involve Energy

- A **chemical reaction** occurs when atoms have sufficient energy to combine or to change their bonding partners.
- **Anabolic reactions** require energy and are endergonic. **Catabolic reactions** release energy and are exergonic. **Review Figure 2.14**
- The **laws of thermodynamics** govern biochemical reactions. The first law states that in any transformation, energy is neither created nor destroyed. The second law states that useful energy tends to decrease. In other words, **entropy** (disorder) tends to increase. **Review Figure 2.15**

See **ANIMATED TUTORIAL** 2.3

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is2

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

3 Nucleic Acids, Proteins, and Enzymes

KEY CONCEPTS

The bark of the willow tree (*Salix alba*) was the original source of salicylic acid, later modified to aspirin.

Despite suffering from the "ague," the Reverend Edward Stone went walking in the English countryside. Feverish, tired, with aching muscles and joints, he came across a willow tree. Although apparently unaware that many ancient healers used willow bark extracts to reduce fever, the clergyman knew of the tradition of natural remedies for various diseases. The willow reminded him of the bitter extracts from the bark of South American trees then being sold (at high prices) to treat fevers. Removing some willow bark, Stone sucked on it and found it did indeed taste bitter—and that it relieved his symptoms.

Later he gathered a pound of willow bark and ground it into a powder, which he gave to about 50 people who complained of pain; all said they felt better. Stone reported the results of this "clinical test" in a letter to the Royal Society, England's most respected scientific body. Stone had discovered the main source of salicylic acid, the basis of the most widely used drug in the world. The date of his letter (which still exists) was April 25, 1763.

The chemical structure of salicylic acid (named for *Salix*, the willow genus) was worked out about 70 years later, and soon chemists could synthesize it in the laboratory. Although the compound alleviated pain, its acidity irritated the digestive system. In the late 1890s, the German chemical company Bayer synthesized a milder yet equally effective form, acetylsalicylic acid, which it marketed as aspirin. The new medicine's success launched Bayer to world prominence as a pharmaceutical company, a position it maintains today.

In the 1960s and 1970s, aspirin use declined when two alternative medications, acetaminophen (Tylenol) and ibuprofen (Motrin and Advil), became widely available. But over this same time, clinical studies revealed a new use for aspirin: it is an effective anticoagulant, shown to prevent heart attacks and strokes caused by blood clots. Today many people take a daily low dose of aspirin as a preventive agent against clotting disorders.

Fever, joint pain, headache, blood clots. What do these symptoms have in common? They all are mediated by fatty acid products called prostaglandins and molecules derived from them. Salicylic acid blocks the synthesis of the primary prostaglandin. The biochemical mechanism by which aspirin works was described in 1971. As we will see, an understanding of this mechanism requires an understanding of protein and enzyme function—two subjects of this chapter.

Q How does an understanding of proteins and enzymes help explain how aspirin works?

You will find the answer to this question on page 57.

CONCEPT 3.1 Nucleic Acids Are Informational Macromolecules

Nucleic acids are polymers that store, transmit, and express hereditary (genetic) information. This information is encoded in the sequences of monomers that make up nucleic acids. There are two types of nucleic acids: **DNA** (*d*eoxyribo*n*ucleic *a*cid) and **RNA** (*r*ibo*n*ucleic *a*cid). DNA stores and transmits genetic information. Through RNA intermediates, the information encoded in DNA is used to specify the amino acid sequences of proteins. As you will see later in this chapter, proteins are essential for both metabolism and structure. Certain specialized RNA molecules also play roles in metabolism. Ultimately, *nucleic acids and the proteins encoded by them determine the metabolic functions of an organism*.

Nucleotides are the building blocks of nucleic acids

Nucleic acids are polymers composed of monomers called nucleotides. A **nucleotide** consists of three components: a nitrogen-containing **base**, a pentose sugar, and one to three phosphate groups (FIGURE 3.1). Molecules consisting of a pentose sugar and a base—but no phosphate group—are called nucleosides. The nucleotides that make up nucleic acids contain just one phosphate group—they are nucleoside monophosphates.

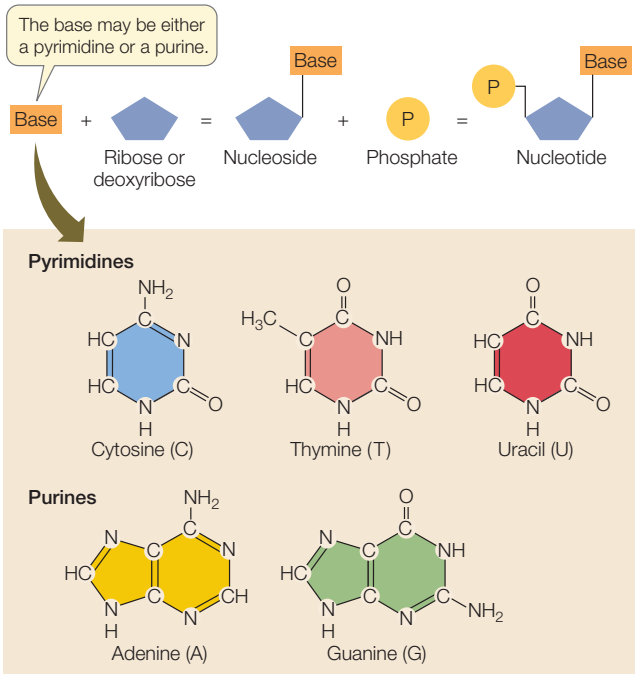


FIGURE 3.1 Nucleotides Have Three Components Nucleotide monomers are the building blocks of DNA and RNA polymers. Nucleotides may have one to three phosphate groups; those in DNA and RNA have one.

Go to ACTIVITY 3.1 Nucleic Acid Building Blocks
PoL2e.com/ac3.1

TABLE 3.1 Distinguishing RNA from DNA

Nucleic acid	Sugar	Bases	Strands
RNA	Ribose	Adenine Cytosine Guanine Uracil	Single
DNA	Deoxyribose	Adenine Cytosine Guanine Thymine	Double

The bases of the nucleic acids take one of two chemical forms: a six-membered single-ring structure called a **pyrimidine**, or a fused double-ring structure called a **purine** (see Figure 3.1). In DNA, the pentose sugar is **deoxyribose**, which differs from the **ribose** found in RNA by the absence of one oxygen atom (see Figure 2.9).

During the formation of a nucleic acid, new nucleotides are added to an existing chain one at a time. The pentose sugar in the last nucleotide of the existing chain and the phosphate on the new nucleotide undergo a condensation reaction (see Figure 2.8) and the resulting linkage is called a **phosphodiester bond**. The phosphate on the new nucleotide is attached to the 5′ (5 prime) carbon atom of its sugar, and the bond occurs between it and the 3′ (3 prime) carbon on the last sugar of the existing chain. Because each nucleotide is added to the 3′ carbon of the last sugar, nucleic acids are said to *grow in the 5′ to 3′ direction* (FIGURE 3.2).

Nucleic acids can be oligonucleotides, with a few to about 20 nucleotide monomers, or longer polynucleotides:

- *Oligonucleotides* include RNA molecules that function as "primers" to begin the duplication of DNA; RNA molecules that regulate the expression of genes; and synthetic DNA molecules used for amplifying and analyzing other, longer nucleotide sequences.
- *Polynucleotides*, more commonly referred to as nucleic acids, include DNA and most RNA. Polynucleotides can be very long, and indeed are the longest polymers in the living world. Some DNA molecules in humans contain hundreds of millions of nucleotides.

Base pairing occurs in both DNA and RNA

In addition to differing in their sugar groups, DNA and RNA also differ in their bases and general structures (TABLE 3.1). Four bases are found in DNA: **adenine** (**A**), **cytosine** (**C**), **guanine** (**G**), and **thymine** (**T**). RNA also contains adenine, cytosine, and guanine, but the fourth base in RNA is **uracil** (**U**) rather than thymine. The lack of a hydroxyl group at the 2′ position of the deoxyribose sugar in DNA makes the structure of DNA less flexible than that of RNA. As we describe below, DNA is composed of two polynucleotide strands whereas RNA is usually single-stranded. However, a long RNA can fold up on itself, forming a variety of structures.

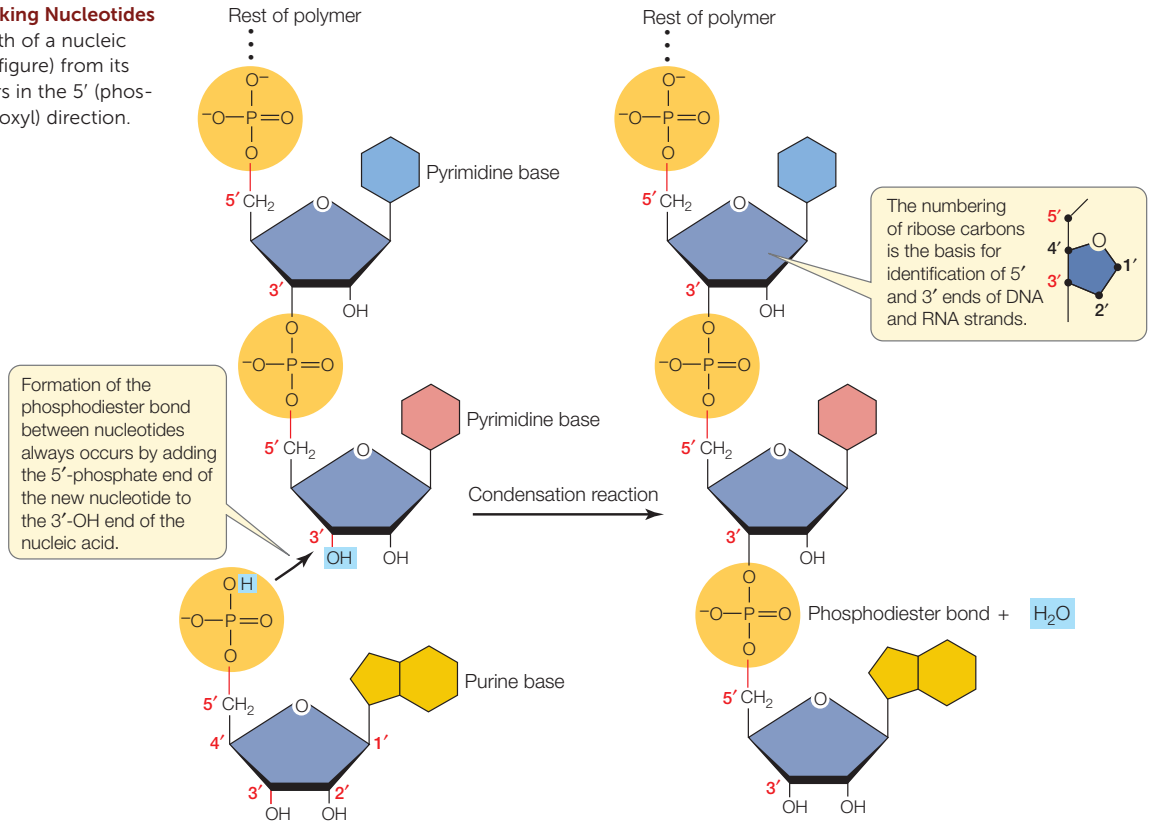
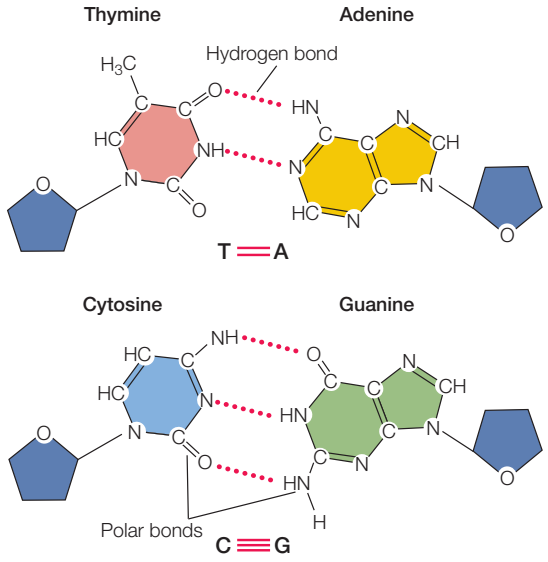


FIGURE 3.2 Linking Nucleotides Together Growth of a nucleic acid (RNA in this figure) from its monomers occurs in the 5′ (phosphate) to 3′ (hydroxyl) direction.

The key to understanding the structure and function of both DNA and RNA is the principle of **complementary base pairing**. In DNA, adenine and thymine always pair (A-T), and cytosine and guanine always pair (C-G):



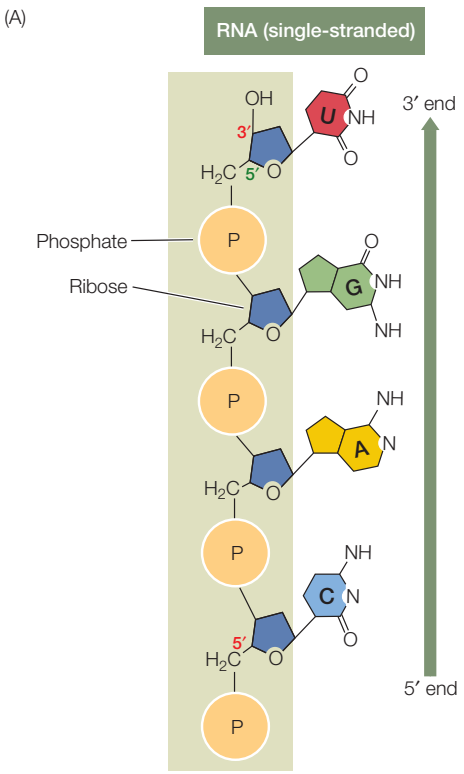
In RNA, the base pairs are A-U and C-G. Base pairs are held together primarily by hydrogen bonds. As you can see, there are polar C═O and N—H covalent bonds in the nucleotide bases (see Concept 2.2 for a discussion of polar covalent bonds).

Hydrogen bonds form between the partial negative charge (δ^-) on an oxygen or nitrogen atom of one base, and the partial positive charge (δ^+) on a hydrogen atom of another base. Complementary base pairing occurs because the arrangements of polar bonds in the nucleotide bases favor the pairing of bases as they occur (C with G, and A with U or T).

Individual hydrogen bonds are relatively weak, but there are so many of them in DNA and RNA that collectively they provide a considerable force of attraction. However, this attraction is not as strong as that provided by multiple covalent bonds. This means that base pairs are relatively easy to separate with a modest input of energy. As you will see in Chapters 9 and 10, the breaking and making of hydrogen bonds in nucleic acids is vital to their roles in living systems. Let's now look at a little more detail at the structures of RNA and DNA.

RNA Usually, RNA is single-stranded (**FIGURE 3.3A**). However, many single-stranded RNA molecules fold up into three-dimensional structures, because of hydrogen bonding between nucleotides in separate portions of the molecules (**FIGURE 3.3B**). An RNA strand can also fold back on itself to form a double-stranded helix. This results in a three-dimensional surface for the bonding and recognition of other molecules. It is important to realize that this folding occurs by complementary base pairing, and the structure is thus determined by the particular order of bases in the RNA molecule.

DNA Usually, DNA is double-stranded; that is, it consists of two separate polynucleotide strands of the same length



In RNA, the bases are attached to ribose. The bases in RNA are the purines adenine (A) and guanine (G) and the pyrimidines cytosine (C) and uracil (U).

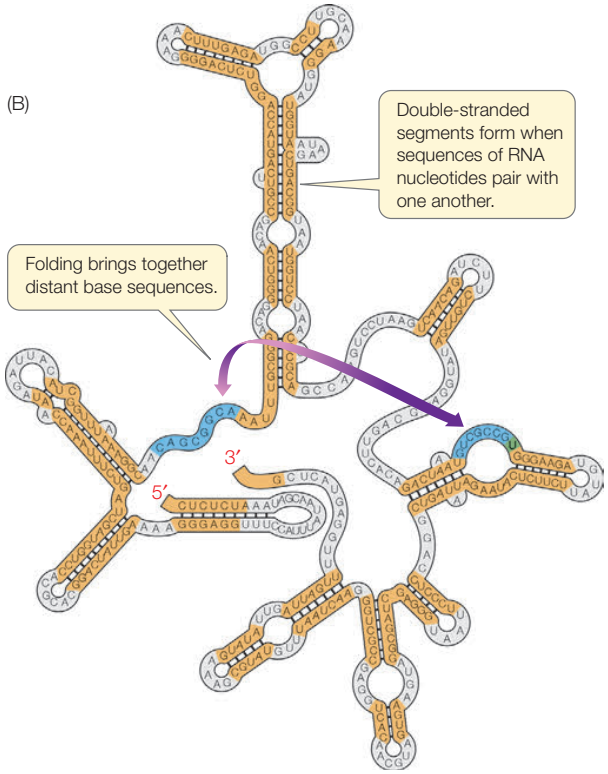
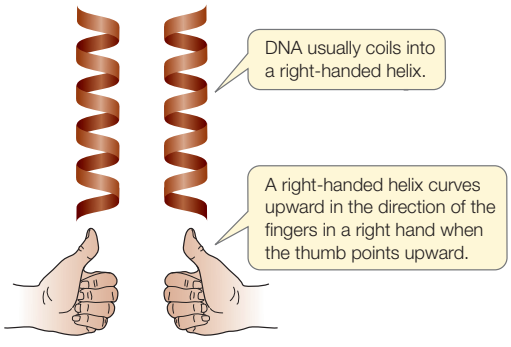


FIGURE 3.3 RNA (A) RNA is usually a single strand. (B) When a single-stranded RNA folds back on itself, hydrogen bonds between complementary sequences can stabilize it into a three-dimensional shape with distinct surface characteristics.

(**FIGURE 3.4A**). The two polynucleotide strands are antiparallel: they run in opposite directions so that their 5′ ends are at opposite ends of the double-stranded molecule. In contrast to RNA's diversity in three-dimensional structure, DNA is remarkably uniform. The A-T and G-C base pairs are about the same size (each is a purine paired with a pyrimidine), and the two polynucleotide strands form a "ladder" that twists into a double helix (**FIGURE 3.4B**). The sugar–phosphate groups form the sides of the ladder, and the bases with their hydrogen bonds form the rungs on the inside. The double helix is almost always right-handed:

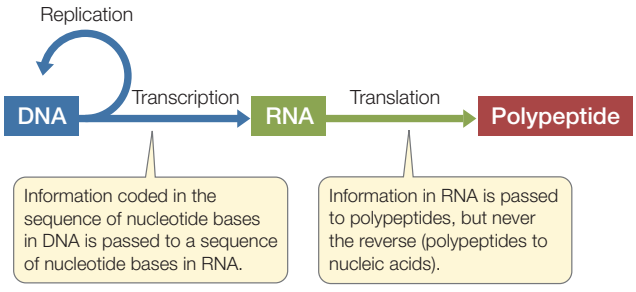


Go to ACTIVITY 3.2 DNA Structure
PoL2e.com/ac3.2

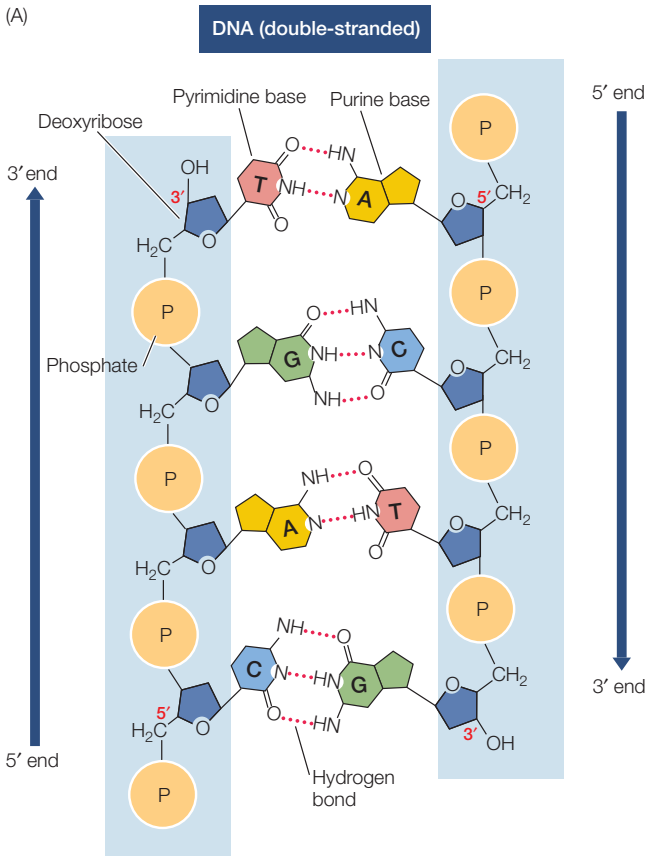
DNA carries information and is expressed through RNA

DNA is a purely informational molecule. The information is encoded in the sequence of bases carried in its strands. For example, the information encoded in the sequence TCAGCA is different from the information in the sequence CCAGCA. DNA has two functions in terms of information:

- DNA can be reproduced precisely by **DNA replication**. DNA is replicated by polymerization using an existing strand as a base-pairing template.
- Some DNA sequences can be copied into RNA, in a process called **transcription**. The nucleotide sequences in most RNA molecules can then be used to specify sequences of amino acids in proteins (polypeptides). This process is called **translation**.



The details of these important processes are described in Chapters 9 and 10, but it is important to realize several things at this point:



In DNA, the bases are attached to deoxyribose, and the base thymine (T) is found instead of uracil. Hydrogen bonds between purines and pyrimidines hold the two strands of DNA together.

FIGURE 3.4 DNA (A) DNA usually consists of two strands running in opposite directions that are held together by base pairing between purines and pyrimidines opposite one another on the two strands. (B) The two antiparallel strands in a DNA molecule are twisted into a double helix.

- *DNA replication and transcription depend on the base pairing properties of nucleic acids.* In both replication and transcription, the hydrogen bonds between two DNA strands are broken, so that complementary base pairing can occur between an existing DNA strand and a newly forming strand of DNA or RNA. The resulting new DNA or RNA strand is *complementary to* the existing DNA template strand. Recall that the hydrogen-bonded base pairs are A-T and G-C in DNA and A-U and G-C in RNA. Now, consider this double-stranded DNA region:

  5′-TCAGCA-3′

  3′-AGTCGT-5′

  Transcription of the lower strand will result in a single strand of RNA with the sequence 5′-UCAGCA-3′. Can you figure out what RNA sequence the top strand would produce?
- *DNA replication usually involves the entire DNA molecule.* Since DNA holds essential information, it must be replicated completely so that each new cell or new organism receives a complete set of DNA from its parent (**FIGURE 3.5A**).
- *Gene expression is the transcription and translation of specific DNA sequences.* Sequences of DNA that encode specific proteins and are transcribed into RNA are called **genes** (**FIGURE 3.5B**). The complete set of DNA in a living organism is called its **genome**. However, not all of the information in the genome is needed at all times and in all tissues. For example, in humans, the gene that encodes the major protein in hair (keratin) is expressed only in skin cells. The genetic information in the keratin-encoding gene is transcribed into RNA and then translated into the protein keratin. In other tissues such as the muscles, the keratin gene is not transcribed, but other genes are—for example, the genes that encode proteins present in muscles but not in skin.

The DNA base sequence reveals evolutionary relationships

Because DNA carries hereditary information from one generation to the next, a theoretical series of DNA molecules stretches back through the lineage of every organism to the beginning of biological evolution on Earth, about 3.8 billion years ago. The genomes of organisms gradually accumulate changes in their DNA base sequences over evolutionary time. Therefore closely related living species should have more similar base sequences than species that are more distantly related.

Over the past two decades there have been remarkable developments in technologies for determining the order of nucleotides in DNA molecules (DNA sequencing), and in computer technologies to analyze these sequences. These

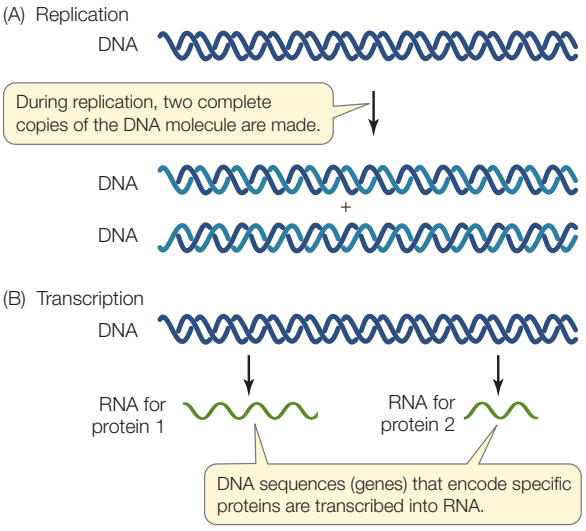


FIGURE 3.5 DNA Replication and Transcription DNA is completely replicated during cell reproduction (A), but it is only partially transcribed (B). In transcription, the DNA code is copied to RNA. The sequence of the latter determines the amino acid sequence of a protein. Transcription of the genes for many different proteins is activated at different times and, in multicellular organisms, in different cells of the body.

advances have enabled scientists to determine the entire DNA base sequences of whole organisms, including the human genome, which contains about 3 billion base pairs. These studies have confirmed many of the evolutionary relationships that were inferred from more traditional comparisons of body structure, biochemistry, and physiology. Traditional comparisons had indicated that the closest living relative of humans (*Homo sapiens*) is the chimpanzee (genus *Pan*). In fact, the chimpanzee genome shares nearly 99 percent of its DNA base sequence with the human genome. Increasingly, scientists turn to DNA analyses to figure out evolutionary relationships when other comparisons are not possible or are not conclusive. For example, DNA studies revealed a close relationship between starlings and mockingbirds that was not expected on the basis of their anatomy or behavior.

LINK

For more on the use of DNA sequences to reconstruct the evolutionary history of life, see **Concept 16.2**

CHECKpoint CONCEPT 3.1

- ✓ List the key differences between DNA and RNA and between purines and pyrimidines.
- ✓ What are the differences between DNA replication and transcription?
- ✓ If one strand of a DNA molecule has the sequence 5′-TTCCGGAT-3′, what is the sequence of the other strand of DNA? If RNA is transcribed from the 5′-TTCCGGAT-3′ strand, what would be its sequence? And if RNA is transcribed from the other DNA strand, what would be its sequence? (Note that it is conventional to write these sequences with the 5′ end on the left.)
- ✓ How can DNA molecules be so diverse when they appear to be structurally similar?

Nucleic acids are largely informational molecules that encode proteins. We will now turn to a discussion of proteins—the most structurally and functionally diverse class of macromolecules.

Go to ANIMATED TUTORIAL 3.1
Macromolecules: Nucleic Acids and Proteins
PoL2e.com/at3.1

CONCEPT 3.2 Proteins Are Polymers with Important Structural and Metabolic Roles

Proteins are the fourth and final type of biological macromolecule we will discuss, and in terms of structural diversity and function, they are at the top of the list. Here are some of the major functions of proteins in living organisms:

- *Enzymes* are catalytic molecules that speed up biochemical reactions. Most enzymes are proteins (some are RNA molecules).
- *Defensive proteins* such as antibodies recognize and respond to substances or particles that invade the organism from the environment.
- *Hormonal and regulatory proteins* such as insulin control physiological processes.
- *Receptor proteins* receive and respond to molecular signals from inside and outside the organism.
- *Storage proteins* store chemical building blocks—amino acids—for later use.
- *Structural proteins* such as collagen provide physical stability and enable movement.
- *Transport proteins* such as hemoglobin carry substances within the organism.
- *Genetic regulatory proteins* (transcription factors) regulate when, how, and to what extent a gene is expressed.

Clearly, the biochemistry of proteins warrants our attention!

Amino acids are the building blocks of proteins

As we noted in Chapter 2, **proteins** are polymers made up of monomers called **amino acids**. As their name suggests, the amino acids all contain two functional groups: the nitrogen-containing amino group and the (acidic) carboxyl group.

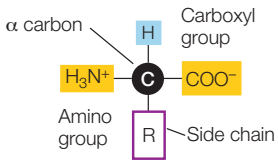
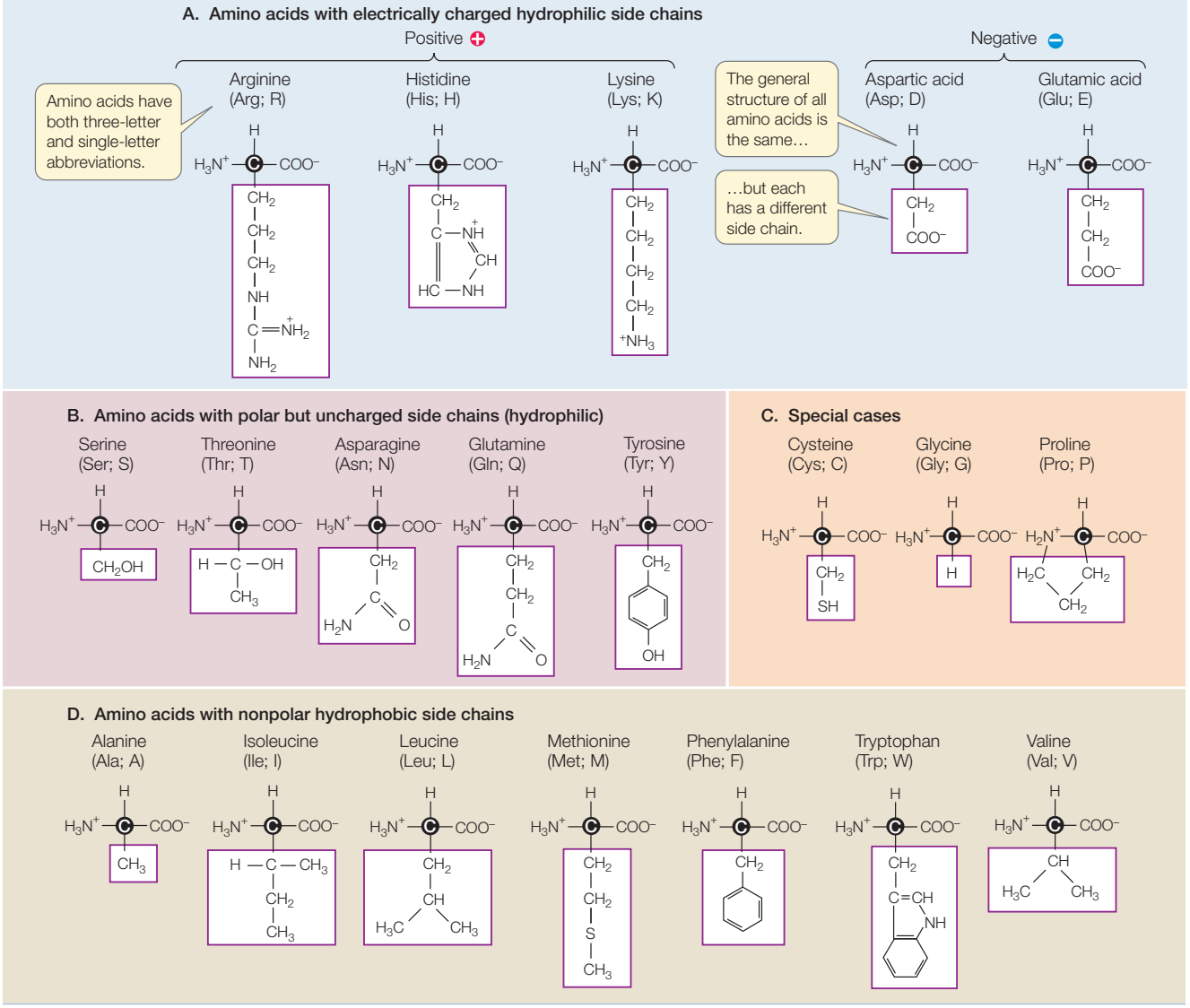


TABLE 3.2 The Twenty Amino Acids in Proteins



The amino and carboxyl groups shown in the diagram are charged. How does this happen? Under the conditions that exist in most living systems, the carboxyl group releases an H^+ (a cation), leaving the rest of the group as an anion:

$$—COOH \rightarrow —COO^- + H^+$$

From your studies of chemistry, you may recognize the carboxyl group as an acid. Conversely, under the same conditions the amino group tends to form a bond with H^+:

$$—NH_2 + H^+ \rightarrow —NH_3^+$$

Your chemistry knowledge should tell you that the amino group is a base.

The central carbon atom of an amino acid—the α (alpha) carbon—has four available electrons for covalent bonding. In all amino acids, two of the electrons are occupied by the two functional groups noted above, and a third is occupied by a hydrogen atom. The fourth bonding electron is shared with a group that differs in each amino acid. This is often referred to as the **R group**, or **side chain**, and is designated by the letter **R**. Each amino acid is identified by its R group.

Go to ACTIVITY 3.3 Features of Amino Acids
PoL2e.com/ac3.3

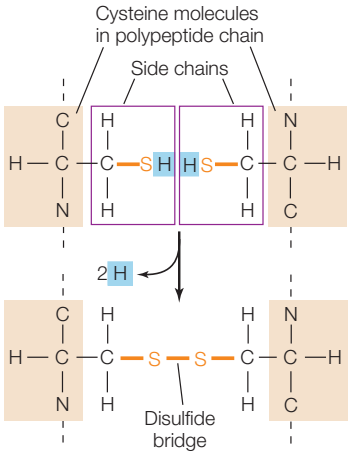
There are hundreds of amino acids known in nature, and many of these occur in plants. But *only 20 amino acids* (listed in **TABLE 3.2**) *occur extensively in the proteins of all organisms*. These 20 amino acids can be grouped according to the properties conferred by their side chains (R groups):

- Five amino acids have electrically charged side chains (+1 or –1), attract water (are hydrophilic), and attract oppositely charged ions of all sorts.

- Five amino acids have polar side chains (δ^+, δ^-) and tend to form hydrogen bonds with water and other polar or charged substances. These amino acids are also hydrophilic.
- Seven amino acids have side chains that are nonpolar hydrocarbons or very slightly modified hydrocarbons. In the watery environment of the cell, these hydrophobic side chains may cluster together in the interior of the protein.

Three amino acids—glycine, proline, and cysteine—are special cases, although the side chains of the former two generally are hydrophobic:

- The glycine side chain consists of a single hydrogen atom and is small enough to fit into tight corners in the interior of a protein molecule, where a larger side chain could not fit.
- Proline possesses a modified amino group that lacks a hydrogen atom and instead forms a covalent bond with the hydrocarbon side chain, resulting in a ring structure. This limits both its hydrogen-bonding ability and its ability to rotate. Thus proline often functions to stabilize bends or loops in proteins.
- The cysteine side chain, which has a terminal —SH group, can react with another cysteine side chain to form a covalent bond called a **disulfide bridge**, or disulfide bond (—S—S—). Disulfide bridges help determine how a protein molecule folds.



In the reaction shown above, the two —SH groups each lose a hydrogen atom (a proton and an electron) and become **oxidized**. The —SH group, which carries the extra electron and proton, is in its **reduced** state.

LINK

In addition to their role in protein structure, oxidation and reduction reactions are important in cellular metabolism; see **Concept 6.1**

Amino acids are linked together by peptide bonds

Amino acids can form short polymers of 20 or fewer amino acids, called **oligopeptides** or simply **peptides**. These include some hormones and other molecules involved in signaling from one part of an organism to another. Even with their relatively short chains of amino acids, oligopeptides have distinctive three-dimensional structures.

More common are the longer polymers called **polypeptides**, each with a unique sequence of amino acids. A functional protein may be made up of one or more polypeptides. Proteins range in size from small ones such as insulin, which has 51 amino acids, to huge molecules such as the muscle protein titin, with 34,350 amino acids.

Like nucleic acids, oligopeptides and polypeptides form via the sequential addition of new amino acids to the ends of existing chains. The amino group of the new amino acid reacts with the carboxyl group of the amino acid at the end of the chain. This condensation reaction forms a **peptide bond** (**FIGURE 3.6**). Note that there is directionality here, just as with the nucleic acids. In this case, *polymerization takes place in the amino to carboxyl direction*.

The precise sequence of amino acids in a polypeptide chain is the **primary structure** of a protein. Scientists have determined the primary structures of many proteins. The single-letter abbreviations for amino acids (see Table 3.2) are used to record the amino acid sequences of proteins. Here, for example, are the first 20 amino acids (out of a total of 1,827) in the human protein sucrase:

MARKKFSGLEISLIVLFVIV

The theoretical number of different proteins is enormous. Since there are 20 different amino acids, there could be $20 \times 20 = 400$ distinct dipeptides (two linked amino acids) and $20 \times 20 \times 20 = 8{,}000$ different tripeptides (three linked amino acids). So for even a small polypeptide of 100 amino acids there are 20^{100} possible sequences, each with its own distinctive primary structure. How large is the number 20^{100}? Physicists tell us there aren't that many electrons in the entire universe.

Higher-level protein structure is determined by primary structure

The primary structure of a protein is established by covalent bonds, but higher levels of structure are determined largely by weaker forces, including hydrogen bonds and hydrophobic and hydrophilic interactions. Follow **FIGURE 3.7** as we describe how a protein chain becomes a three-dimensional structure.

SECONDARY STRUCTURE A protein's **secondary structure** consists of regular, repeated spatial patterns in different regions of a polypeptide chain. There are two basic types of secondary structure, both determined by hydrogen bonding between the amino acids that make up the primary structure:

- The α (**alpha**) **helix** is a right-handed coil that turns in the same direction as a standard wood screw (see Figure 3.7B). The R groups extend outward from the peptide backbone of the helix. The coiling results from hydrogen bonds that form between the N—H group on one amino acid and the C═O group on another within the same turn of the helix.

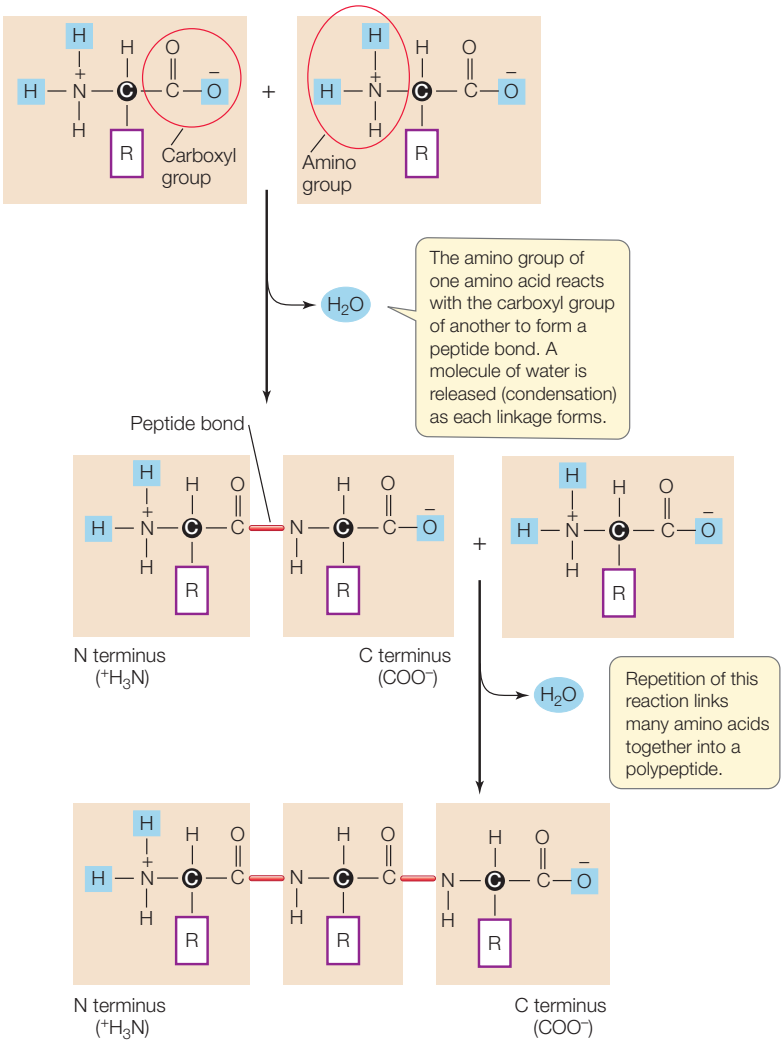


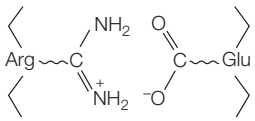
FIGURE 3.6 Formation of a Peptide Bond In living things, the reaction leading to a peptide bond has many intermediate steps, but the reactants and products are the same as those shown in this simplified diagram.

- The **β (beta) pleated sheet** is formed from two or more sequences of amino acids that are extended and aligned. The sheet is stabilized by hydrogen bonds between the N—H groups and the C═O groups on the two chains (see Figure 3.7C). A β pleated sheet may form between separate polypeptide chains or between different regions of a single polypeptide chain that is bent back on itself. Many proteins contain both α helices and β pleated sheets in different regions of the same polypeptide chain.

TERTIARY STRUCTURE In many proteins, the polypeptide chain is bent at specific sites and then folded back and forth, resulting in **tertiary structure** (see Figure 3.7D). Tertiary structure results in the polypeptide's definitive three-dimensional shape, including a buried interior as well as a surface that is exposed to the environment. The protein's exposed outer surfaces present functional groups capable of interacting with other molecules in the cell. These molecules might be other proteins or smaller chemical reactants (as in enzymes; see below).

Whereas hydrogen bonding between the N—H and C═O groups within and between chains is responsible for a protein's secondary structure, it is the interactions between R groups—the amino acid side chains—that determine tertiary structure (**FIGURE 3.8**):

- *Covalent disulfide bridges* can form between specific cysteine side chains, holding a folded polypeptide together.
- *Hydrogen bonds* between side chains also stabilize folds in proteins.
- *Hydrophobic side chains* can aggregate together in the interior of a protein, away from water, folding the polypeptide in the process.
- *van der Waals interactions* can stabilize close associations between hydrophobic side chains.
- *Ionic interactions* can form between positively and negatively charged side chains, forming "salt bridges" between amino acids. Ionic interactions can also be buried deep within a protein, away from water. These interactions occur between positively and negatively charged amino acids, for example arginine (which has a positively charged R group) and glutamic acid (which has a negatively charged R group):



LINK

To review the strong and weak interactions that can occur between atoms, see **Concept 2.2**

A complete description of a protein's tertiary structure would specify the location of every atom in the molecule in three-dimensional space, relative to all the other atoms. Many such descriptions are available, including one for the human protein sucrase (**FIGURE 3.9**).

Remember that both secondary and tertiary structure derive from primary structure. If a protein is heated slowly, the heat energy will disrupt only the weaker interactions, causing the secondary and tertiary structure to break down. The protein is then said to be **denatured**. Chemical treatments can also be used to denature proteins. In many cases a denatured protein can return to its normal tertiary structure when it cools or the denaturing

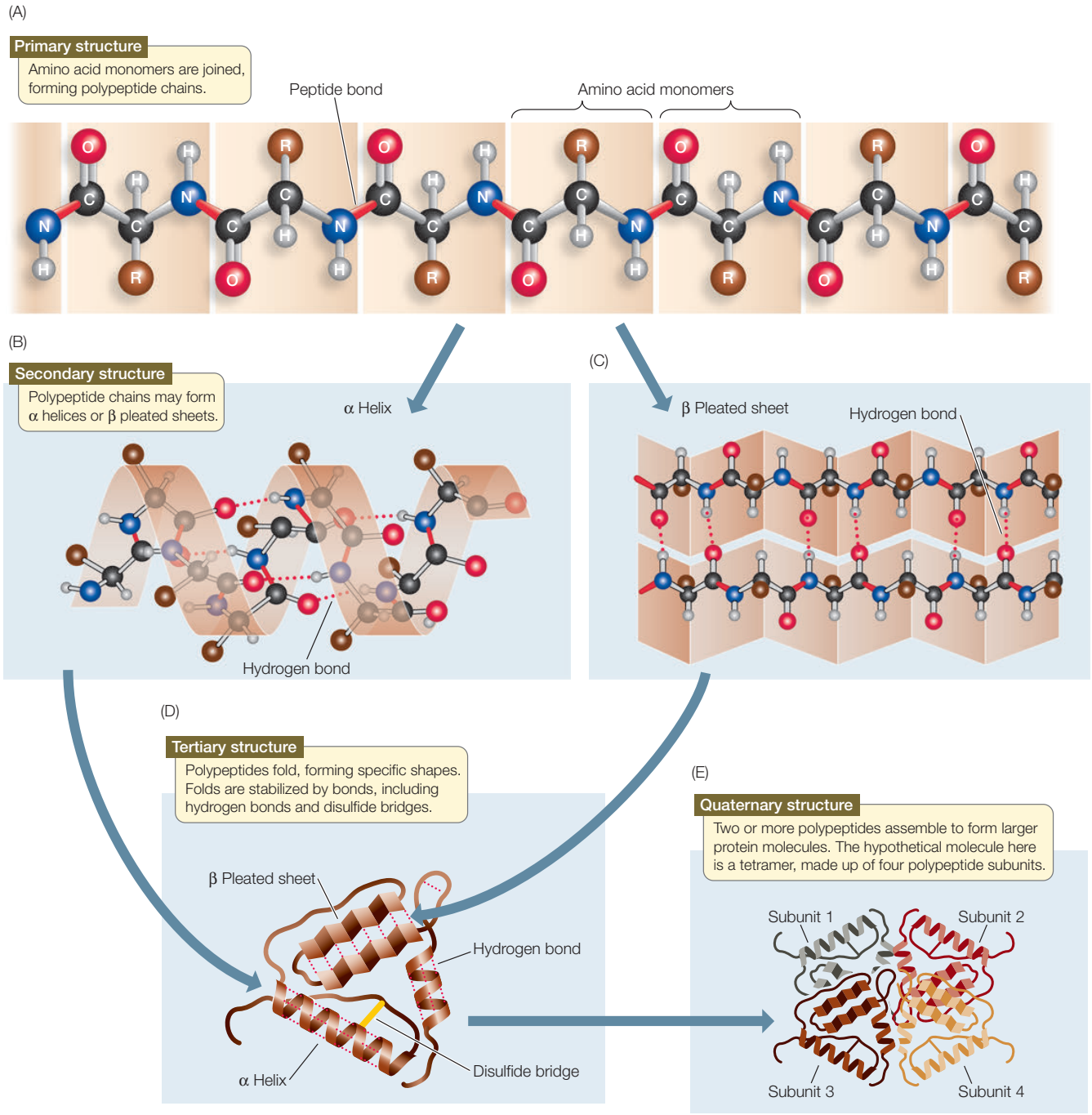


FIGURE 3.7 The Four Levels of Protein Structure The primary structure (A) of a protein determines what its secondary (B and C), tertiary (D), and quaternary (E) structures will be.

chemicals are removed, demonstrating that all the information needed to specify the protein's unique shape is contained in its primary structure. This fact was first shown by biochemist Christian Anfinsen for the protein ribonuclease (FIGURE 3.10).

QUATERNARY STRUCTURE Many functional proteins contain two or more polypeptide chains, called subunits, each folded into its own unique tertiary structure. The protein's **quaternary structure** results from the ways in which these

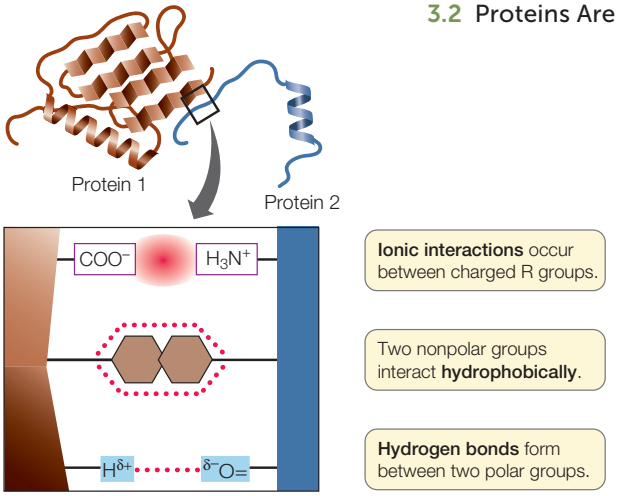
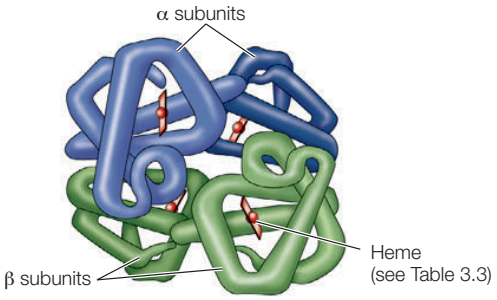


FIGURE 3.8 Noncovalent Interactions between Proteins and Other Molecules Noncovalent interactions allow a protein (brown) to bind tightly to another protein (blue) with specific properties. Noncovalent interactions also allow regions within a single protein to interact with one another.

subunits bind together and interact (see Figure 3.7E). Hemoglobin is an example of a protein with multiple subunits:



Hydrophobic interactions, hydrogen bonds, and ionic interactions all help hold the four subunits together to form a hemoglobin macromolecule. The weak nature of these forces permits small changes in the quaternary structure to aid the protein's function—which is to carry oxygen in red blood cells. As hemoglobin binds one O_2 molecule, the four subunits shift their relative positions slightly, changing the quaternary structure. Ionic interactions are broken, exposing buried side chains that enhance the binding of additional O_2 molecules. The quaternary structure changes again when hemoglobin releases its O_2 molecules to the cells of the body.

Protein structure can change

The environment that surrounds a protein, as well as interactions with other molecules, can change protein structure.

ENVIRONMENT Various conditions can alter the weak, noncovalent interactions that hold proteins together in their secondary, tertiary, and quaternary structures:

- *Increases in temperature* cause more rapid molecular movements and thus can break hydrogen bonds and hydrophobic interactions.

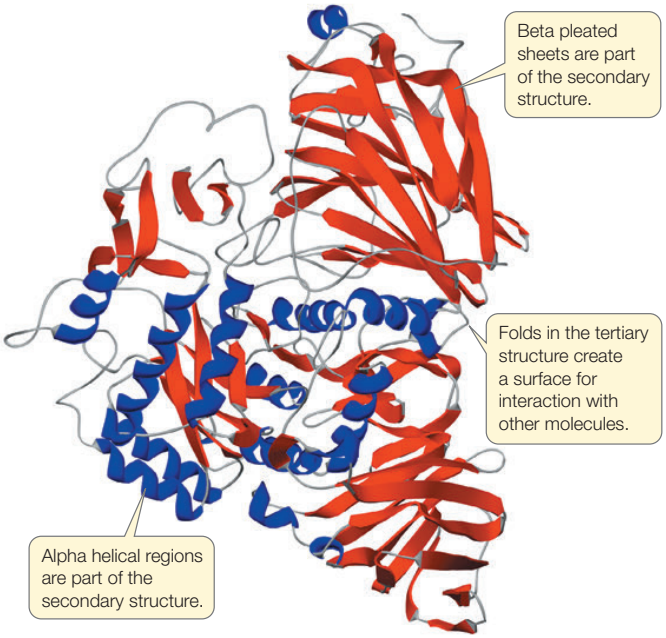


FIGURE 3.9 The Structure of a Protein Sucrase has a specific three-dimensional structure, determined by its primary structure. Sucrase plays a role in digestion in humans.

- *Alterations in the concentration of* H^+ (pH) in the solution surrounding the protein can change the patterns of ionization of the exposed carboxyl and amino groups. This can disrupt the patterns of ionic attractions and repulsions.
- *High concentrations of polar substances* such as urea can disrupt the hydrogen bonding that is crucial to protein structure.
- *Nonpolar substances* may also denature a protein in cases where hydrophobic groups are essential for maintaining the protein's structure.

Go to MEDIA CLIP 3.1
Protein Structures in 3D
PoL2e.com/mc3.1

Denaturation can be irreversible when amino acids that were buried in the interior of the protein become exposed at the surface, or vice versa, causing a new structure to form, or causing different molecules to bind to the protein. Boiling an egg denatures its proteins and is, as you know, not reversible.

MOLECULAR INTERACTIONS Proteins do not exist in isolation. Within a living organism, a protein may interact with other proteins, other kinds of macromolecules, or a variety of smaller molecules. These interactions are reminiscent of the interactions that make up quaternary structure (see above). If a polypeptide comes into contact with another molecule, R groups on its surface may form weak interactions (such as hydrogen bonds or ionic interactions) with groups on the surface

INVESTIGATION

FIGURE 3.10 Primary Structure Specifies Tertiary Structure
Using the protein ribonuclease, Christian Anfinsen showed that proteins spontaneously fold into functionally correct three-dimensional configurations.[a] As long as the primary structure is not disrupted, the information for correct folding (under the right conditions) is retained.

HYPOTHESIS

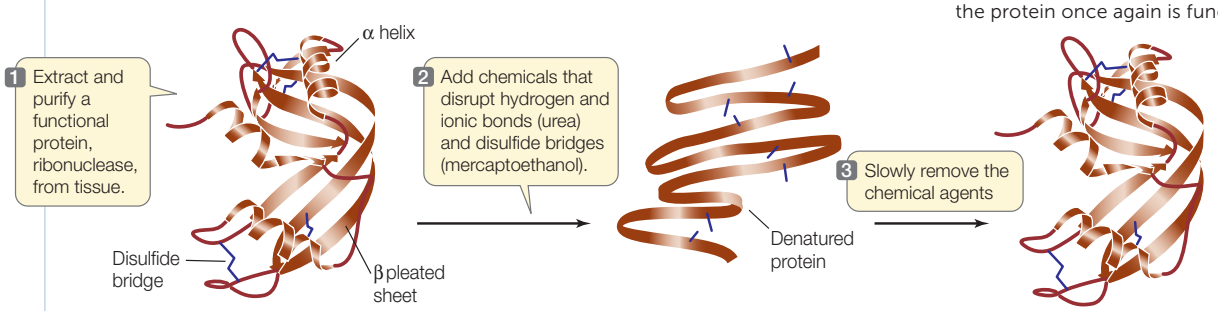
Under controlled conditions that simulate the normal cellular environment, a denatured protein can refold into a functional three-dimensional structure.

METHOD

Chemically denature a functional ribonuclease, so that only its primary structure (i.e., an unfolded polypeptide chain) remains.

RESULTS

When the disruptive agents are removed, three-dimensional structure is restored and the protein once again is functional.



CONCLUSION

In normal cellular conditions, the primary structure of a protein specifies how it folds into a functional, three-dimensional structure.

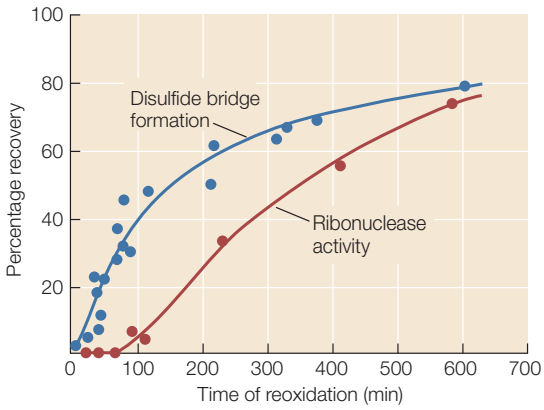
ANALYZE THE DATA

Initially, disulfide bridges (S—S) in ribonuclease were eliminated because the sulfur atoms in cysteine were reduced (—SH). At time 0, reoxidation began and at various times, the amount of disulfide bridge re-formation (blue circles) and the function of ribonuclease (enzyme activity; red circles) were measured by chemical methods. Here are the data:

A. At what time did disulfide bridges begin to form?

B. At what time did enzyme activity begin to appear?

C. Explain the difference between your answers for the times of (A) and (B).



Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a] C. B. Anfinsen et al. 1961. *Proceedings of the National Academy of Sciences USA*. 47: 1309–1314.

of the other molecule. This may disrupt some of the interactions between R groups within the polypeptide, causing it to undergo a change in shape (**FIGURE 3.11A**).

The structure of a protein can also be modified by the covalent bonding of a chemical group to the side chain of one or more of its amino acids. The chemical modification of just one amino acid can alter the shape and function of a protein. An example is the addition of a charged phosphate group to a relatively nonpolar R group. This can cause the amino acid to become more hydrophilic and to move to the outer surface of the protein, altering the shape of the protein in the region near the amino acid (**FIGURE 3.11B**).

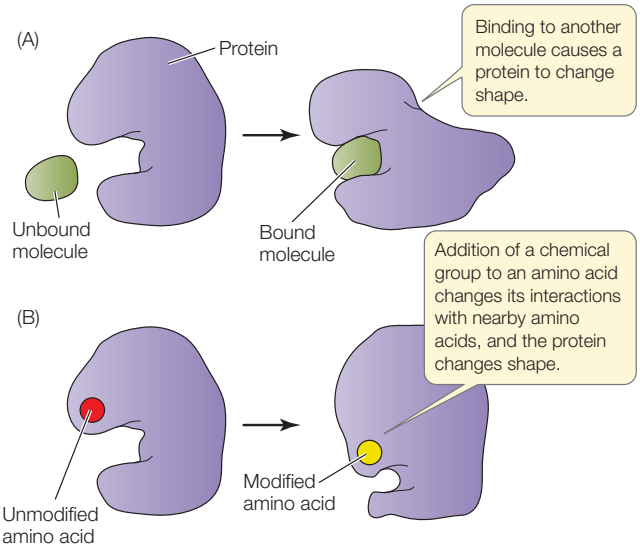


FIGURE 3.11 Protein Structure Can Change Proteins can change their tertiary structure when they bind to other molecules (A) or are modified chemically (B).

CHECKpoint CONCEPT 3.2

- ✓ Sketch the peptide bonding of the two amino acids glycine and leucine (in that order). Now add a third amino acid, alanine, in the position it would have if added within a biological system. What is the directionality of this process?
- ✓ Examine the structure of sucrase (see Figure 3.9). Where in the protein might you expect to find the following amino acids: valine, proline, glutamic acid, and threonine? Explain your answers.
- ✓ Detergents disrupt hydrophobic interactions by coating hydrophobic molecules with a molecule that has a hydrophilic surface. When hemoglobin is treated with a detergent, the four polypeptide chains separate and become random coils. Explain these observations.
- ✓ Several small molecules interact with a protein. The chemical groups on the small molecules interact with specific amino acids as shown in the table below. Fill in the table to show the types of noncovalent interactions that occur between the small molecules and the amino acids.

SMALL MOLECULE CHEMICAL GROUP	AMINO ACID IN PROTEIN	TYPE OF INTERACTION (HYDROGEN BOND; IONIC INTERACTION; HYDROPHOBIC INTERACTION)
$-NH_3^+$	Aspartic acid	
$-CH_3$	Isoleucine	
$-OH$	Glutamine	

We have discussed the remarkable diversity in protein structures. These structures carry functional groups (on exposed amino acid side chains) that can interact with other molecules. In the next section we will see how these interactions can result in catalysis, the speeding up of biochemical reactions.

APPLY THE CONCEPT

Proteins are polymers with important structural and metabolic roles

Biological systems contain "supermolecular complexes" (for example, the ribosome; see Chapter 4), which are composed of individual molecules of RNA and protein that fit together noncovalently. These complexes can be split apart with detergents that disrupt hydrophobic interactions. Based on the concepts discussed in this chapter, fill in the table below to indicate which of the observations are characteristic of RNA, which are characteristic of protein, and which are characteristic of both. Explain your answers.

OBSERVATION	CHARACTERISTIC OF: PROTEIN	RNA
Has three-dimensional (3-D) structure		
3-D structure destroyed by heat		
Monomers connected by N—C bonds		
Contains sulfur atoms		
Contains phosphorus atoms		

CONCEPT 3.3 Some Proteins Act as Enzymes to Speed up Biochemical Reactions

In Chapter 2 we introduced the concepts of biological energetics. We showed that some metabolic reactions are exergonic and some are endergonic, and that biochemistry obeys the laws of thermodynamics (see Figures 2.14 and 2.15). Knowing whether energy is supplied or released in a particular reaction tells us whether the reaction *can* occur in a living system. But it does not tell us *how fast* the reaction will occur.

Living systems depend on reactions that occur spontaneously. But without help, most of these reactions would proceed at such slow rates that an organism could not survive. The role of a **catalyst** is to speed up a reaction without itself being permanently altered. A catalyst does not cause a reaction to occur, but it increases the rate of the reaction. This is an important point: *No catalyst makes a reaction occur that would not proceed without it*.

Biological catalysts are called **enzymes**; for example, the synthesis of prostaglandin (see the opening story) is catalyzed by an enzyme (cyclooxygenase). Most enzymes are proteins, but a few important enzymes are RNA molecules called ribozymes. An enzyme can bind the reactants in a chemical reaction and participate in the reaction itself. However, this participation does not permanently change the enzyme. At the end of the reaction, the enzyme is unchanged and available to catalyze additional, similar reactions.

An energy barrier must be overcome to speed up a reaction

An exergonic reaction releases **free energy** (G), which is the amount of energy in a system that is available to do work. For

example, the free energy released in an exergonic reaction can be used by the cell to drive an endergonic reaction, or it can be converted to mechanical energy for movement (see Figure 6.1). But without a catalyst, a reaction will usually take place very slowly. This is because there is an energy barrier between the reactants and the products. Think about the hydrolysis of sucrose, which we described in Concept 2.5.

$$\text{Sucrose} + H_2O \rightarrow \text{glucose} + \text{fructose}$$

In humans, this reaction is part of the process of digestion. The reaction is exergonic, but even if water is abundant, the sucrose molecule will only rarely bind the H atom and the –OH group in the water molecule at the appropriate locations to break the covalent bond between the glucose and fructose—*unless there is an input of energy to initiate the reaction*. Such an input of energy will place the sucrose into a reactive mode called the **transition state**. The energy input required for sucrose to reach this state is called the **activation energy** (E_a). Once the transition state is reached, the reaction can proceed spontaneously with a release of free energy (ΔG is negative) (FIGURE 3.12A). The image of a ball rolling over a bump and then down a hill helps illustrate these concepts (FIGURE 3.12B).

Where does the activation energy come from? In any collection of reactants at room or body temperature, the molecules are moving around. Recall from Chapter 2 that the energy the molecules possess due to this motion is called kinetic energy. A few molecules are moving fast enough that their kinetic energy can overcome the energy barrier; they enter the transition state and react. So the reaction takes place—but very slowly. If the system is heated, all the reactant molecules have more kinetic energy, and the reaction speeds up. You have probably used this technique in the chemistry laboratory.

Adding enough heat to increase the average kinetic energy of the molecules would not work in a living system, however. Such a nonspecific approach would accelerate all reactions, including destructive ones such as the denaturation of proteins.

An enzyme lowers the activation energy for a reaction by enabling the reactants to come together and react more easily; the reactants need lower amounts of kinetic energy to enter their transition states (FIGURE 3.13). In this way, an enzyme can change the rate of a reaction substantially. For example, if a molecule of sucrose just sits in solution, hydrolysis may take hundreds of years. But with the enzyme sucrase present, the same reaction occurs in 1 second! Typically, an enzyme-catalyzed reaction proceeds 10^3 to 10^8 times faster than the uncatalyzed reaction, and the enzyme converts 100 to 1,000 substrate molecules into product per second.

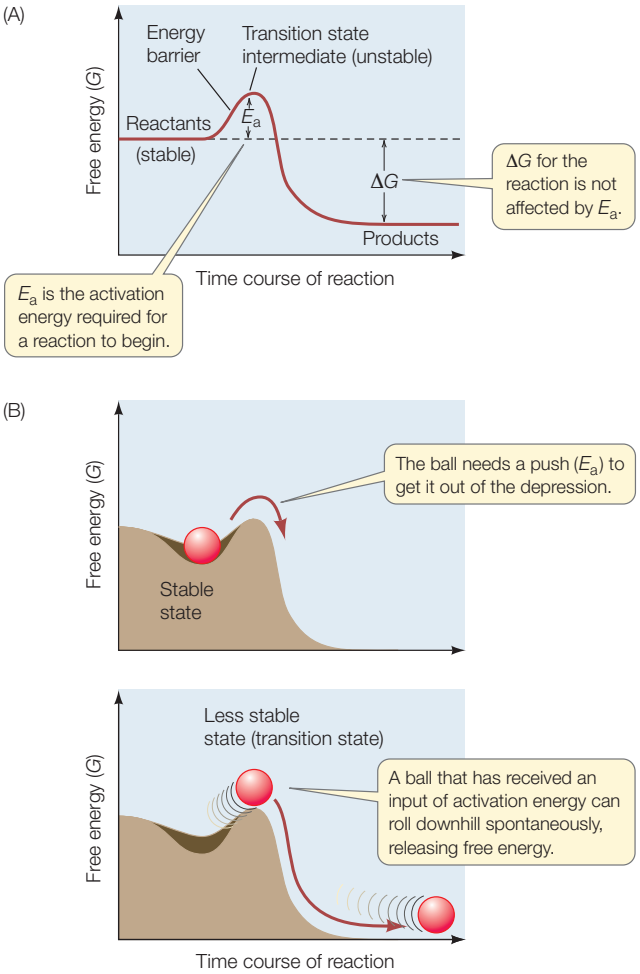


FIGURE 3.12 Activation Energy Initiates Reactions (A) In any chemical reaction, an initial stable state must become less stable before change is possible. (B) A ball on a hillside provides a physical analogy to the biochemical principle graphed in A. Although these graphs show an exergonic reaction, activation energy is needed for endergonic reactions as well.

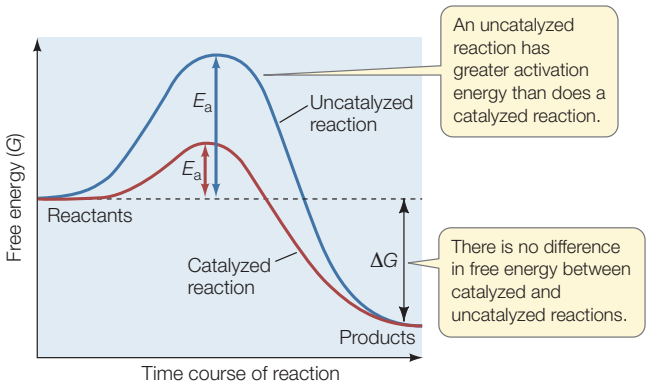


FIGURE 3.13 Enzymes Lower the Energy Barrier The activation energy (E_a) is lower in an enzyme-catalyzed reaction than in an uncatalyzed reaction, but the free energy released is the same with or without catalysis. A lower activation energy means the reaction will take place at a faster rate.

Go to ACTIVITY 3.4 Free Energy Changes
PoL2e.com/ac3.4

Enzymes bind specific reactants at their active sites

Catalysts increase the rates of chemical reactions. Most nonbiological catalysts are nonspecific. For example, powdered platinum catalyzes virtually any reaction in which molecular hydrogen (H_2) is a reactant. In contrast, most biological catalysts are highly specific. An enzyme usually recognizes and binds to only one or a few closely related reactants, and it catalyzes only a single chemical reaction.

In an enzyme-catalyzed reaction, the reactants are called **substrates**. Substrate molecules bind to a particular site on the enzyme, called the **active site**, where catalysis takes place (**FIGURE 3.14**). The specificity of an enzyme results from the exact three-dimensional shape (also called conformation) and chemical properties of its active site. Only a narrow range of substrates, with specific shapes, functional groups, and chemical properties, can fit properly and bind to the active site. The names of enzymes reflect their functions and often end with the suffix "ase." For example, the enzyme sucrase catalyzes the hydrolysis of sucrose, and we write the reaction as follows:

$$\text{Sucrose} + H_2O \xrightarrow{\text{Sucrase}} \text{glucose} + \text{fructose}$$

The binding of a substrate (S) to the active site of an enzyme (E) produces an **enzyme–substrate complex** (**ES**) that is held together by one or more means, such as hydrogen bonding, ionic attraction, or temporary covalent bonding. The enzyme–substrate complex gives rise to product (P) and free enzyme:

$$E + S \rightarrow ES \rightarrow E + P$$

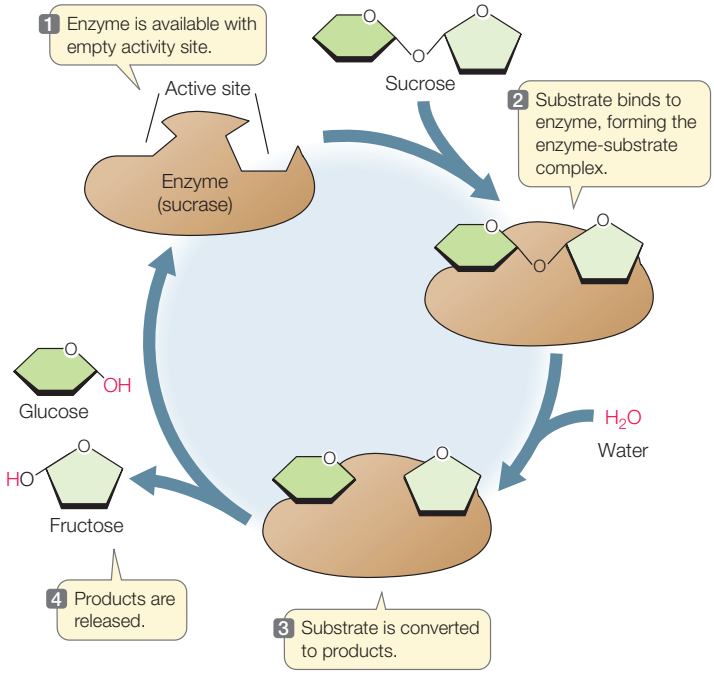
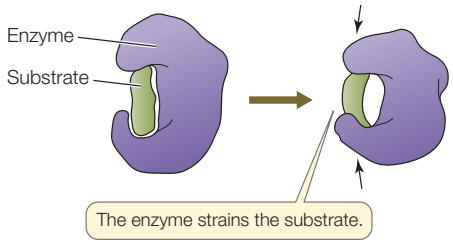


FIGURE 3.14 Enzyme Action Sucrase catalyzes the hydrolysis of sucrose. After the reaction, the enzyme is unchanged and is ready to accept another substrate molecule.

(As we have seen in the case of sucrase, a single enzyme-catalyzed reaction may involve multiple substrates and/or products.) The free enzyme (E) is in the same chemical form at the end of the reaction as at the beginning. While bound to the substrate(s), it may change chemically, but by the end of the reaction it has been restored to its initial form and is ready to catalyze the same reaction again (see Figure 3.14).

HOW ENZYMES WORK During and after the formation of the enzyme–substrate complex, chemical interactions occur. These interactions contribute directly to the breaking of old bonds and the formation of new ones. In catalyzing a reaction, an enzyme may use one or more mechanisms:

- *Inducing strain*: Once the substrate has bound to the active site, the enzyme causes bonds in the substrate to stretch, putting it in an unstable transition state:



- *Substrate orientation:* When free in solution, substrates are moving from place to place randomly while at the same time vibrating, rotating, and tumbling. They only rarely have the proper orientation to react when they collide. The enzyme lowers the activation energy needed to start the reaction, by bringing together specific atoms so that bonds can form.
- *Adding chemical groups:* The side chains (R groups) of an enzyme's amino acids may be directly involved in the reaction. For example, in acid–base catalysis, the acidic or basic side chains of the amino acids in the active site transfer H^+ ions to or from the substrate, destabilizing a covalent bond in the substrate and permitting the bond to break.

The active site is usually only a small part of the enzyme protein. But its three-dimensional structure is so specific that it binds only one or a few related substrates. The binding of the substrate to the active site depends on the same relatively weak forces that maintain the tertiary structure of the enzyme: hydrogen bonds, the attraction and repulsion of charged groups, and hydrophobic interactions. Scientists used to think of substrate binding as being similar to a lock and key fitting together. Actually, for most enzymes and substrates the relationship is more like a baseball and a catcher's mitt: the substrate first binds, and then the active site changes slightly to make the binding tight. **FIGURE 3.15** illustrates this "induced fit" phenomenon. (We introduced the concept of protein structure changes earlier; see Figure 3.11.)

Induced fit at least partly explains why enzymes are so large. The rest of the macromolecule has at least three roles:

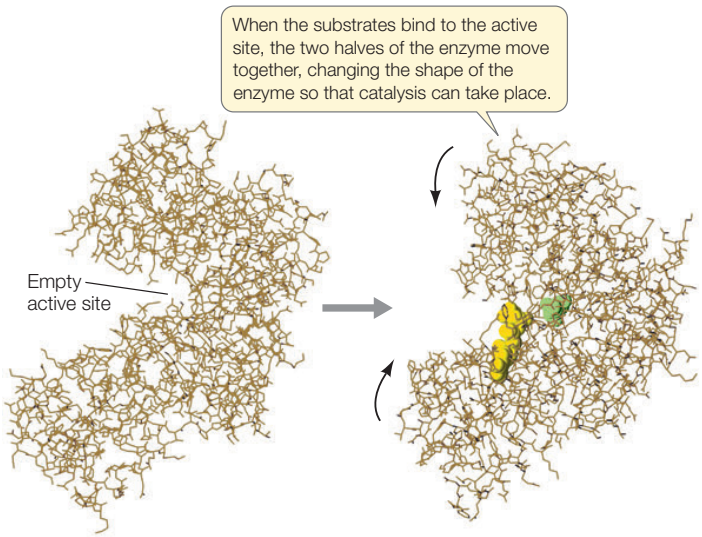


FIGURE 3.15 Some Enzymes Change Shape When Substrate Binds to Them Shape changes result in an induced fit between enzyme and substrate, improving the catalytic ability of the enzyme. Induced fit can be observed in the enzyme hexokinase, seen here with and without its substrates, glucose (green) and ATP (yellow).

- It provides a framework so the amino acids of the active site are properly positioned in relation to the substrate(s).
- It participates in the changes in protein shape and structure that result in induced fit.
- It provides binding sites for regulatory molecules (as we will discuss in Concept 3.4).

NONPROTEIN PARTNERS FOR ENZYMES Some enzymes require ions or other molecules in order to function. These molecules are referred to as **cofactors**, and they can be grouped into three categories (TABLE 3.3):

- *Metal ions* such as copper, zinc, and iron bind to certain enzymes and participate in the enzyme-catalyzed reactions. For example, the cofactor zinc binds to the enzyme alcohol dehydrogenase, which catalyzes the breakdown of toxic alcohol.
- A *coenzyme* is a relatively small, carbon-containing (organic) molecule that is required for the action of one or more enzymes. It binds to the active site of the enzyme, adds or removes a chemical group from the substrate, and then separates from the enzyme to participate in other reactions. A coenzyme differs from a substrate in that it can participate in many different reactions with different enzymes.
- *Prosthetic groups* are organic molecules that are permanently bound to their enzymes. An example is a flavin nucleotide, which binds to succinate dehydrogenase, an important enzyme in energy metabolism.

RATE OF REACTION The rate of an uncatalyzed reaction is directly proportional to the concentration of the substrate. The higher the concentration, the more reactions per unit of time.

As we have seen, the addition of the appropriate enzyme speeds up the reaction, but it also changes the shape of the plot of rate versus substrate concentration (FIGURE 3.16). For a given concentration of enzyme, the rate of the enzyme-catalyzed reaction initially increases as the substrate concentration increases from zero, but then it levels off.

Why does this happen? The concentration of an enzyme is usually much lower than that of its substrate and does not change as substrate concentration changes. When all the enzyme molecules are bound to substrate molecules, the enzyme is working at its maximum rate. Under these conditions the active sites are said to be saturated.

The maximum rate of a catalyzed reaction can be used to measure how efficient the enzyme is—that is, how many molecules of substrate are converted into product by an individual enzyme molecule per unit of time, when there is an excess of substrate present. This turnover number ranges from 1 molecule every second for sucrase to an amazing 40 million molecules per second for the liver enzyme catalase.

CHECKpoint CONCEPT 3.3

- ✓ Explain how the structure of an enzyme makes that enzyme specific.
- ✓ What is activation energy? How does an enzyme lower the activation energy needed to start a reaction?
- ✓ Compare coenzymes with substrates. How do they work together in enzyme catalysis?
- ✓ Compare the state of an enzyme active site at a low substrate concentration and at a high substrate concentration. How does this affect the rate of the reaction?

TABLE 3.3 Some Examples of Enzyme Cofactors

Type of cofactor	Role in catalyzed reactions
METAL IONS	
Iron (Fe^{2+} or Fe^{3+})	Oxidation/reduction
Copper (Cu^{+} or Cu^{2+})	Oxidation/reduction
Zinc (Zn^{2+})	Helps bind NAD
COENZYMES	
Biotin	Carries $-COO^{-}$
Coenzyme A	Carries $-CO-CH_3$
NAD	Carries electrons
FAD	Carries electrons
ATP	Provides/extracts energy
PROSTHETIC GROUPS	
Heme	Binds ions, O_2, and electrons; contains iron cofactor
Flavin	Binds electrons
Retinal	Converts light energy

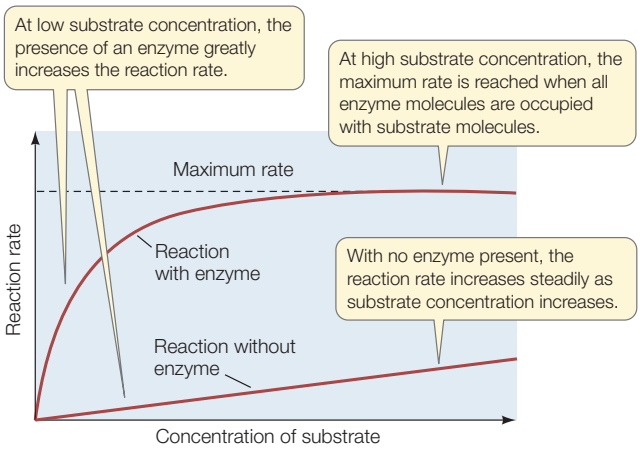


FIGURE 3.16 Catalyzed Reactions Reach a Maximum Rate Because there is usually less enzyme than substrate present, the reaction rate levels off when the enzyme becomes saturated.

Now that you understand more about how enzymes function, let's see how different enzymes work in the metabolism of living organisms.

CONCEPT 3.4 Regulation of Metabolism Occurs by Regulation of Enzymes

The enzyme-catalyzed reactions we have been discussing often operate within **metabolic pathways** in which the product of one reaction is a substrate for the next. For example, the pathway for the catabolism of sucrose begins with sucrase and ends many reactions later with the production of CO_2 and H_2O. Energy is released along the way. Each step of this catabolic pathway is catalyzed by a specific enzyme:

$$\text{Sucrose} + H_2O \xrightarrow{\text{Sucrase}} \text{glucose} + \text{fructose} \longrightarrow \xrightarrow[\text{Many enzymes}]{} CO_2 + H_2O$$

Other enzymes participate in anabolic pathways, which produce relatively complex molecules from simpler ones. A typical cell contains hundreds of enzymes that participate in many interconnecting metabolic pathways, forming a metabolic system (**FIGURE 3.17**). Consider a single molecule in the midst of this map:

- There may be two or more enzyme-catalyzed reactions affecting it: either making it or metabolizing it.
- Other pathways affect the concentrations of the substrates and products of these reactions.
- Each enzyme-catalyzed reaction has its own rate, depending on these concentrations.

Clearly, every component of this complex system is affected by every other component, making it difficult to predict what would happen if one or more components were altered. In the new field of **systems biology**, scientists describe mathematically the components of metabolic systems—the concentrations of all the reactants and the rates of the reactions—and use computer algorithms to make predictions about what would happen if a component of the system were altered (see Concept 1.2, pp. 8–9).

Cells need to maintain stable internal conditions, including constant levels of certain metabolites. In addition, cells need to regulate their metabolic pathways to respond to changes, either within the organism or in its environment. One way a cell can regulate its metabolism is to control the *amount* of an enzyme. For example, the product of a metabolic pathway may be available from the cell's environment in adequate amounts. In this case, it would be energetically wasteful for the cell to continue making large proteins (as most enzymes are) that it doesn't need. For this reason, cells often have the ability to turn off the synthesis of certain enzymes.

LINK

The amount of an enzyme is controlled by regulating the expression of gene(s), a topic covered in **Chapter 11**

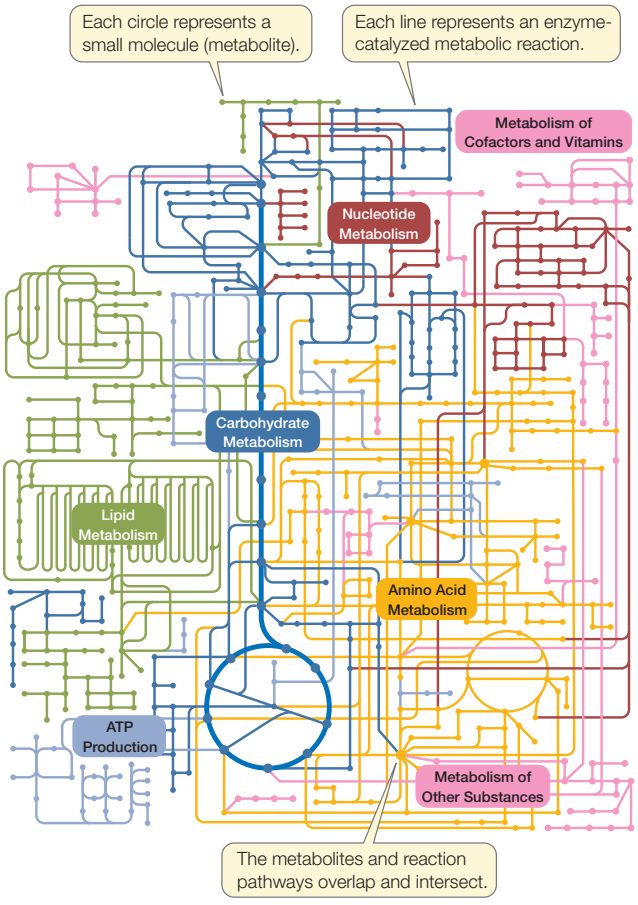


FIGURE 3.17 A Biochemical System The complex interactions of metabolic pathways can be studied using the tools of systems biology. Enzymes are a major element controlling these pathways.

The consequences of *too little* enzyme can be significant. For example, in humans sucrase is important in digestion. In rare cases, infants are born with a congenital sucrase deficiency and the pathway that begins with sucrose is essentially blocked. If these infants ingest foods containing sucrose, the sucrose accumulates rather than being catabolized, and the infant gets diarrhea and stomach cramps. In some cases this leads to slower growth. This deficiency can be treated by limiting sucrose consumption or taking tablets that contain sucrase at every meal.

Cells can also maintain stable internal conditions by regulating the *activity* of enzymes. An enzyme protein may be present continuously, but it may be active or inactive depending on the needs of the cell. Synthesizing and breaking down enzymes takes time, whereas regulating enzyme activity allows cells to fine-tune metabolism relatively quickly in response to changes in the environment. In this section, we will describe how enzyme regulation occurs.

Enzymes can be regulated by inhibitors

Various chemical inhibitors can bind to enzymes, slowing down the rates of the reactions they catalyze. Some inhibitors occur naturally in cells; others can be made in laboratories. Naturally occurring inhibitors regulate metabolism; artificial ones (such as the improved version of salicylic acid described in the opening story) can be used to treat disease, kill pests, or study how enzymes work. In some cases the inhibitor binds the enzyme irreversibly, and the enzyme becomes permanently inactivated. In other cases the inhibitor has reversible effects; it can separate from the enzyme, allowing the enzyme to function fully as before.

IRREVERSIBLE INHIBITION If an inhibitor covalently binds to an amino acid side chain at the active site of an enzyme, the enzyme is permanently inactivated because it cannot interact with its substrate. An example of an irreversible inhibitor is DIPF (diisopropyl phosphorofluoridate), which irreversibly inhibits acetylcholinesterase, an important enzyme that functions in the nervous system. DIPF does so by reacting with a hydroxyl group on a serine in the active site (**FIGURE 3.18**).

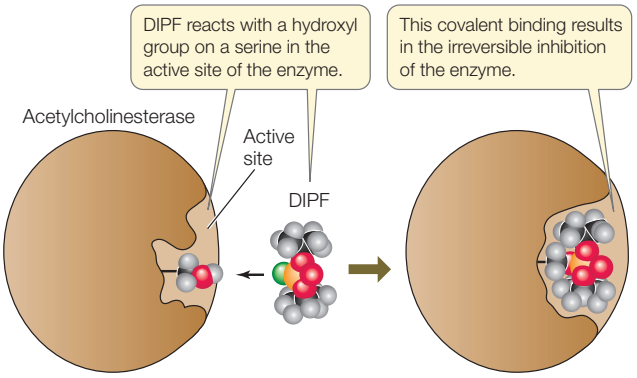


FIGURE 3.18 Irreversible Inhibition DIPF forms a stable covalent bond with the amino acid serine at the active site of the enzyme acetylcholinesterase, thus irreversibly disabling the enzyme.

The widely used insecticide malathion is a derivative of DIPF that inhibits only insect acetylcholinesterase, not the mammalian enzyme. The irreversible inhibition of enzymes is of practical use to humans, but this form of regulation is not common in the cell, because the enzyme is permanently inactivated and cannot be recycled. Instead, cells use reversible inhibition.

REVERSIBLE INHIBITION In some cases, an inhibitor is similar enough to a particular enzyme's natural substrate that it can bind noncovalently to the active site, yet different enough that no chemical reaction occurs. This is analogous to a key that inserts into a lock but does not turn it. When such a molecule is bound to the enzyme, the natural substrate cannot enter the active site and the enzyme is unable to function. Such a molecule is called a **competitive inhibitor** because it competes with the natural substrate for the active site (**FIGURE 3.19A**). Many drugs are competitive inhibitors of enzyme targets. For example, methotrexate is a drug designed with a structure similar to the metabolite dihydrofolate. The latter is converted by an enzyme to a substance essential to cell division. Acting as a competitive inhibitor of the enzyme, methotrexate blocks cell division and is used in cancer therapy. Competitive inhibition is reversible. When the concentration of the competitive inhibitor is reduced, the active site is less likely to be occupied by the inhibitor, and the enzyme regains activity.

A **noncompetitive inhibitor** binds to an enzyme at a site distinct from the active site. This binding causes a change in the shape (the conformation) of the enzyme, altering its activity (**FIGURE 3.19B**). The active site may no longer bind the substrate, or if it does, the rate of product formation may be reduced. Like competitive inhibitors, noncompetitive inhibitors can become unbound, so their effects are reversible.

An allosteric enzyme is regulated by changes in its shape

Noncompetitive inhibition is an example of allostery (*allo*, "different"; *stereos*, "shape"). **Allosteric regulation** occurs when a non-substrate molecule binds or modifies a site other than the active site of an enzyme. The site bound by the non-substrate molecule is called the allosteric site. This binding induces the enzyme to change its conformation, altering the chemical attraction (affinity) of the active site for the substrate. As a result, the rate of the reaction is changed.

An allosteric site may be modified by either noncovalent or covalent binding:

- *Noncovalent binding*: A regulatory molecule may bind noncovalently to an allosteric site, causing the enzyme to change shape. This noncovalent binding is reversible, and may result in the inactivation of an enzyme (see Figure 3.19B) or the activation of a formerly inactive enzyme (**FIGURE 3.20A**).
- *Covalent binding*: Some allosteric sites can be modified by the covalent binding of a molecule or chemical group. For example, an amino acid residue can be covalently modified by the addition of a phosphate group, in a process called phosphorylation (**FIGURE 3.20B**). If this occurs in a hydro-

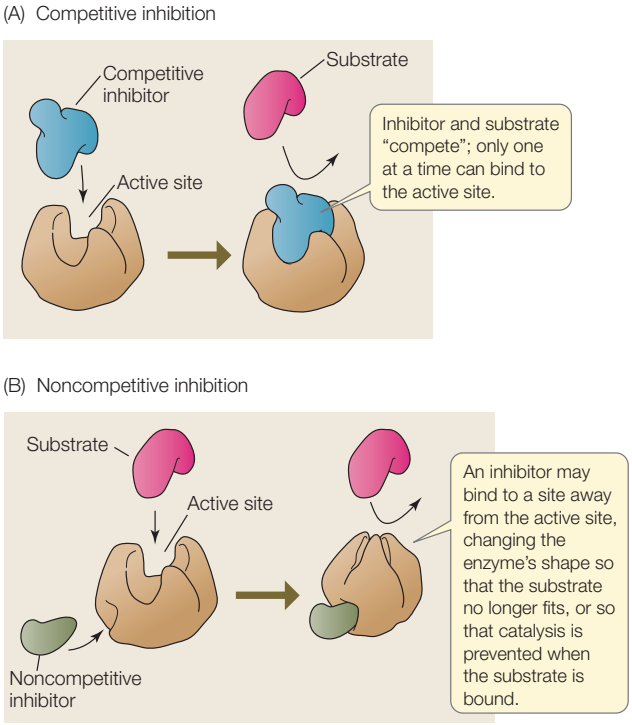


FIGURE 3.19 Reversible Inhibition (A) A competitive inhibitor binds temporarily to the active site of an enzyme. (B) A noncompetitive inhibitor binds temporarily to the enzyme at a site away from the active site. In both cases, the enzyme's function is disabled for only as long as the inhibitor remains bound.

Go to ANIMATED TUTORIAL 3.2
Enzyme Catalysis
PoL2e.com/at3.2

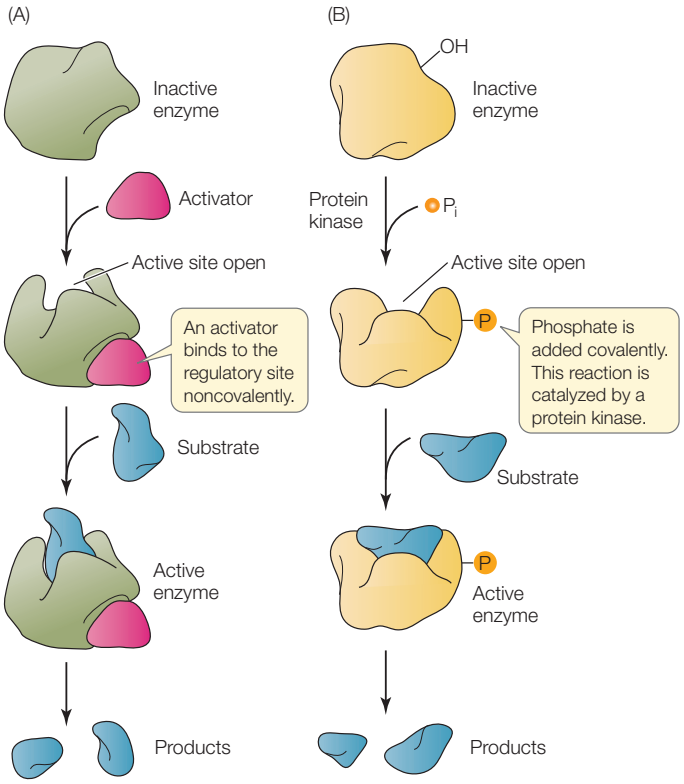


FIGURE 3.20 Allosteric Regulation of Enzyme Activity (A) Noncovalent binding of a regulator (in this case an activator) can cause an enzyme to change shape and expose an active site. B) Enzymes can also be activated by covalent modification, in this case phosphorylation. Note that allosteric regulation can be negative as well, with the active site becoming hidden.

Go to ANIMATED TUTORIAL 3.3
Allosteric Regulation of Enzymes
PoL2e.com/at3.3

phobic region of the enzyme, it makes that region hydrophilic, because phosphate carries a negative charge. The protein twists, and this can expose or hide the active site. Protein phosphorylation is an extremely important mechanism by which cells regulate many different enzymes and other proteins. It is a reversible process: a class of enzymes called protein kinases catalyze the addition of phosphate groups to proteins, whereas protein phosphatases remove phosphate groups from proteins. Humans have hundreds of different protein kinases and phosphatases. We will return to the exact functions of these proteins many times in this book.

LINK

Protein kinases are of particular importance in intracellular signaling pathways (see **Concepts 5.5 and 5.6**) and in the control of cell reproduction (see **Concept 7.3**)

Some metabolic pathways can be controlled by feedback inhibition

A metabolic pathway typically involves a starting material, various intermediate products, and an end product that is used for some purpose by the cell. In each pathway there are a number of reactions, each forming an intermediate product and each catalyzed by a different enzyme. In many pathways the first step is the commitment step, meaning that once this enzyme-catalyzed reaction occurs, the "ball is rolling," and the other reactions happen in sequence, leading to the end product. But as we pointed out earlier, it is energetically wasteful for the cell to make something it does not need.

One way to regulate a metabolic pathway is by having the final product inhibit the enzyme that catalyzes the commitment step (**FIGURE 3.21**). When the end product is present at a high concentration, some of it binds to a site on the commitment step enzyme, thereby causing it to become inactive. The end product may bind to the active site on the enzyme (as a competitive inhibitor) or an allosteric site (as a noncompetitive inhibitor). This mechanism is known as **feedback inhibition** or end-product inhibition. We will describe many other examples of such inhibition in later chapters.

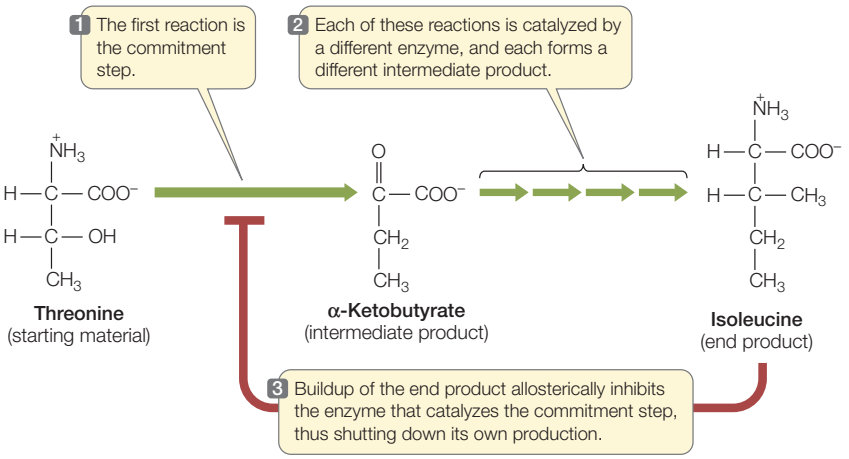


FIGURE 3.21 Feedback Inhibition of Metabolic Pathways The first reaction in a metabolic pathway is referred to as the commitment step. Often the end product of the pathway can inhibit the enzyme that catalyzes the commitment step. The specific pathway shown here is the synthesis of isoleucine from threonine in bacteria. It is typical of many enzyme-catalyzed biosynthetic pathways.

Enzymes are affected by their environment

As we have seen, the specificity and activity of an enzyme depend on its three-dimensional structure, and this in turn depends on weak forces such as hydrogen bonds (see Figure 3.7). In living systems, two environmental factors can change protein structure and thereby enzyme activity.

pH We introduced the concept of acids and bases when we discussed amino acids. Some amino acids have side chains that are acidic or basic (see Table 3.2). That is, they either generate H^+ and become anions, or attract H^+ and become cations. These reactions are often reversible. For example:

$$\text{Glutamic acid—COOH} \rightleftharpoons \text{glutamic acid—COO}^- + H^+$$

The ionic form of this amino acid (right) is far more hydrophilic than the nonionized form (left).

From your studies of chemistry, you may recall the law of mass action or Le Chatelier's principle. In this case the law implies that the higher the H^+ concentration in the solution, the more the reaction will be driven to the left (forming more of the nonionized form of glutamic acid). Therefore changes in the H^+ concentration can alter how hydrophobic some regions of a protein are and thus affect its shape. To generalize, protein tertiary structure, and therefore enzyme activity, is very sensitive to the concentration of H^+ in the aqueous environment. You may also recall that H^+ concentration is measured by pH (the negative logarithm of the H^+ concentration).

Although the water inside cells is generally at a neutral pH of 7, this can change, and different biological environments have different pH values. Each enzyme has a tertiary structure and amino acid sequence that make it optimally active at a particular pH. Its activity decreases as the solution is made more acidic or more basic than this ideal (optimal) pH (**FIGURE 3.22A**). As an example, consider the human digestive system (see Concept 30.4). The pH inside the human stomach is highly acidic, about pH 1.5. Pepsin, an enzyme that is active in the stomach, has a pH optimum near 2. Many enzymes that hydrolyze macromolecules in the intestine, such as proteases, have pH optima in the neutral range. So when food enters the small intestine, a buffer (bicarbonate) is secreted into the intestine to raise the pH to 6.5. This allows the hydrolytic enzymes to be active and digest the food.

TEMPERATURE In general, warming increases the rate of a chemical reaction because a greater proportion of the reactant molecules have enough kinetic energy to provide the activation energy for the reaction. Enzyme-catalyzed reactions are no different (**FIGURE 3.22B**). However, temperatures that are

APPLY THE CONCEPT

Regulation of metabolism occurs by regulation of enzymes

The concept of enzymes as biological catalysts has many applications. In a pile of clothes in your garage, you notice there are bacteria growing on some socks made of this synthetic polymer:

$$[\text{—CO—}(CH_2)_4\text{—CO—NH—}(CH_2)_4\text{—NH—}]_n$$

You make a protein extract from the bacteria and isolate what you think is an enzyme that can cleave the monomers from the polymer. You also synthesize the dipeptide glycine-glycine (see Table 3.2) to test as a possible inhibitor of the enzyme. The table shows the results from several of your experiments.

EXPERIMENT	CONDITION	RATE OF POLYMER CLEAVAGE
1	No enzyme	0.505
2	Enzyme	825.0
3	Enzyme pre-boiled at 100°C	0.520
4	Enzyme + dipeptide	0.495
5	Enzyme + RNA	799.0

1. Explain the results of each experiment.
2. How might the dipeptide work? How would you test your hypothesis?

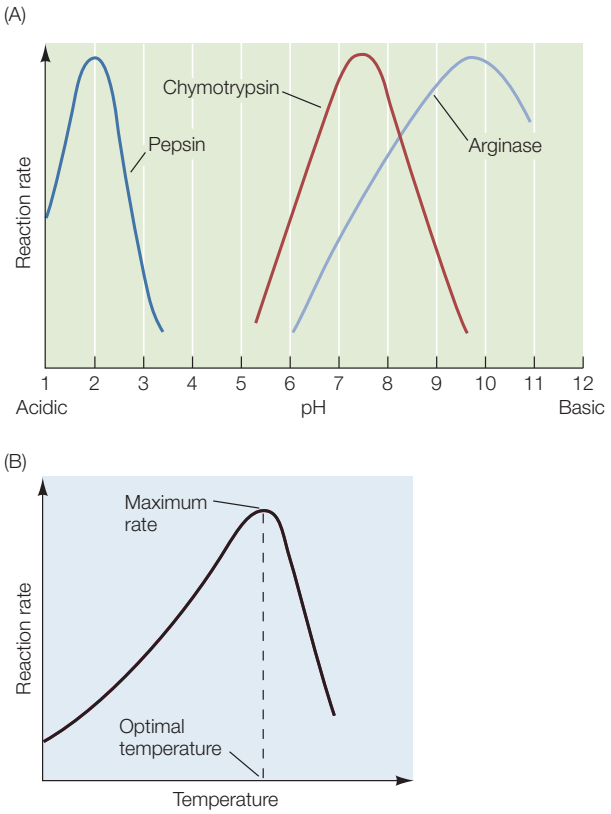


FIGURE 3.22 Enzyme Activity Is Affected by the Environment
(A) The activity curve for each enzyme peaks at its optimal pH. For example, pepsin is active in the acidic environment of the stomach, whereas chymotrypsin is active in the neutral environment of the small intestine, and arginase is active in a basic environment. (B) Similarly, there is an optimal temperature for each enzyme. At higher temperatures the enzyme becomes denatured and inactive; this explains why the activity curve falls off abruptly at temperatures that are above optimal.

too high inactivate enzymes, because at high temperatures the polypetides vibrate and twist so rapidly that some of their noncovalent bonds break. When an enzyme's tertiary structure is changed by heat, the enzyme can no longer function. Some enzymes denature at temperatures only slightly above that of the human body, but a few are stable even at the boiling point (or freezing point) of water. All enzymes have an optimal temperature for activity.

Individual organisms adapt to changes in the environment in many ways, one of which is based on groups of enzymes called isozymes, which catalyze the same reaction but have different chemical compositions and physical properties. Different isozymes within a given group may have different optimal temperatures. The rainbow trout, for example, has several isozymes of the enzyme acetylcholinesterase. If a rainbow trout is transferred from warm water to near-freezing water (2°C), the fish produces a different isozyme of acetylcholinesterase. The new isozyme has a lower optimal temperature, allowing the fish's nervous system to perform normally in the colder water.

In general, enzymes adapted to warm temperatures do not denature at those temperatures because their tertiary structures are held together largely by covalent bonds such as disulfide bridges, instead of the more heat-sensitive weak chemical interactions.

CHECKpoint CONCEPT 3.4

- ✓ Explain and give examples of irreversible and reversible enzyme inhibitors.
- ✓ The amino acid glutamic acid (see Table 3.2) is at the active site of an enzyme. Normally the enzyme is active at pH 7. At pH 4 (higher concentration of H^+), the enzyme is inactive. Explain these observations.
- ✓ An enzyme is subject to allosteric regulation. How would you design an inhibitor of the enzyme that was competitive? Noncompetitive? Irreversible?
- ✓ Some organisms thrive at pH 2; other organisms thrive at a temperature of 65°C. Yet mammals cannot tolerate either environment in their tissues. Explain.

Q How does an understanding of proteins and enzymes help explain how aspirin works?

ANSWER The mechanism by which aspirin works exemplifies many of the concepts introduced in this chapter. Robert Vane showed that aspirin binds to a protein with a specific three-dimensional structure (Concept 3.2). This protein is cyclooxygenase, an enzyme (Concept 3.3) that catalyzes the commitment step in a metabolic pathway (Concept 3.4). Aspirin acts as an irreversible inhibitor of cyclooxygenase (Concept 3.4). Follow the description below carefully, as it illustrates these important concepts.

Cyclooxygenase catalyzes the conversion of a fatty acid with 20 carbon atoms, arachidonic acid, to a structure with a ring (thus the "cyclo" in the name of the enzyme). O_2 is a substrate (thus the "oxygen"; **FIGURE 3.23**). The product of

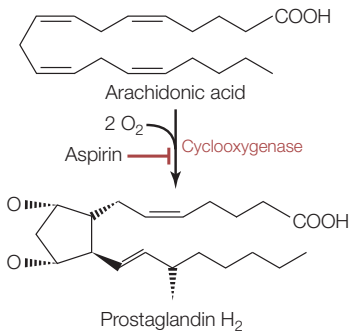


FIGURE 3.23 Aspirin: An Enzyme Inhibitor Aspirin inhibits a key enzyme in the metabolic pathways leading to inflammation and blood clotting.

this reaction (prostaglandin H_2) is the starting material for biochemical pathways that produce two types of molecules:

- prostaglandins, which are involved in inflammation and pain, and
- thromboxanes, which stimulate blood clotting and constriction of blood vessels.

Aspirin binds and reacts with a serine residue within the active site of cyclooxygenase. As a result of this binding, an acetyl group is transferred to the exposed hydroxyl group of the serine residue (**FIGURE 3.24**):

$$\text{(Cyclooxygenase)–serine–OH} \longrightarrow \text{(cyclooxygenase)–serine–O–CH}_2\text{–CH}_3$$

This covalent modification changes the exposed, polar serine to a less polar molecule, and it becomes slightly more hydrophobic. The conformation of the active site changes and becomes inaccessible to the substrate, arachidonic acid. The enzyme is inhibited, and the pathways leading to prostaglandins and thromboxanes are shut down. Less pain, inflammation, and blood clotting are the result. Small wonder that aspirin is taken as a pain reliever and a preventive medicine for heart attacks and strokes. It has come a long way from Edward Stone's walk in the woods

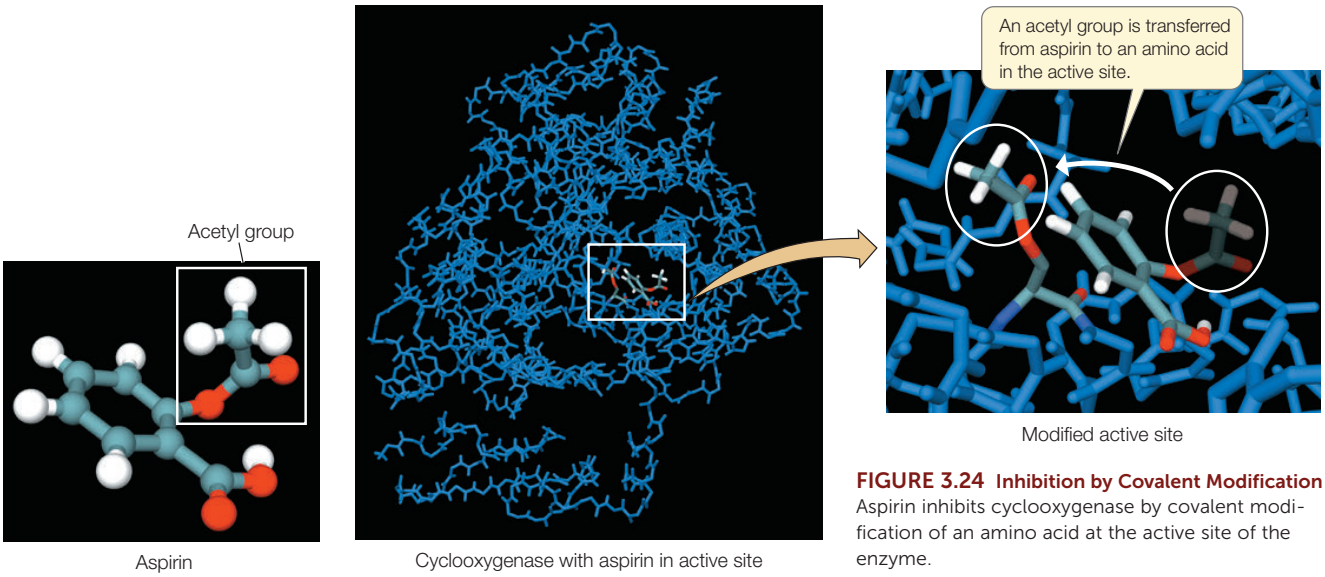


FIGURE 3.24 Inhibition by Covalent Modification
Aspirin inhibits cyclooxygenase by covalent modification of an amino acid at the active site of the enzyme.

SUMMARY

CONCEPT 3.1 Nucleic Acids Are Informational Macromolecules

- The **nucleic acids**—**DNA** and **RNA**—are used mainly to store, transmit, and express hereditary (genetic) information.
- Nucleic acids are polymers of nucleotides. A **nucleotide** consists of one to three phosphate groups, a pentose sugar (**ribose** in RNA and **deoxyribose** in DNA), and a nitrogen-containing **base.** Review Figure 3.1 and ACTIVITY 3.1
- In DNA, the nucleotide bases are **adenine** (**A**), **guanine** (**G**), **cytosine** (**C**), and **thymine** (**T**). **Uracil** (**U**) replaces thymine in RNA. The nucleotides are joined by **phosphodiester bonds** between the sugar of one and the phosphate of the next. RNA is usually single-stranded, whereas DNA is double-stranded. Review Figure 3.2
- **Complementary base pairing**, based on hydrogen bonds between A and T, A and U, and G and C, occurs in RNA and DNA. In RNA the hydrogen bonds result in a folded molecule; in DNA the hydrogen bonds connect two antiparallel strands into a double helix. Review Figures 3.3 and 3.4 and ACTIVITY 3.2
- DNA is expressed as RNA in the process of **transcription**. RNA can then specify the amino acid sequence of a protein in the process of **translation**.

See ANIMATED TUTORIAL 3.1

CONCEPT 3.2 Proteins Are Polymers with Important Structural and Metabolic Roles

- The functions of proteins include support, protection, catalysis, transport, defense, regulation, storage, and movement.
- **Amino acids** are the monomers from which polymeric proteins are made by **peptide bonds**. There are 20 different amino acids in proteins, each distinguished by a **side chain** (**R group**) that confers specific properties. Review Table 3.2 and ACTIVITY 3.3
- The **primary structure** of a protein is the sequence of amino acids in the polypeptide chain. This chain is folded into a **secondary structure**, which in different parts of the protein may take the form of an **α helix** or a **β pleated sheet**. Review Figure 3.7
- **Disulfide bridges** and noncovalent interactions between amino acids cause polypeptide chains to fold into three-dimensional **tertiary structures**. Multiple polypeptides can interact to form **quaternary structures**. A protein's unique shape and chemical structure allow it to bind specifically to other molecules.
- Heat and certain chemicals can result in a protein becoming **denatured**, which involves the loss of tertiary or secondary structure. Review Figure 3.10

CONCEPT 3.3 Some Proteins Act as Enzymes to Speed up Biochemical Reactions

- A chemical reaction must overcome an energy barrier to get started. An **enzyme** is a catalyst that affects the rate of a biological reaction by lowering the **activation energy** needed to initiate the reaction. Review Figure 3.13 and ACTIVITY 3.4
- A **substrate** binds to the enzyme's **active site**—the site of catalysis—forming an **enzyme–substrate complex**. Enzymes are highly specific for their substrates.
- At the active site, a substrate enters its **transition state**, and the reaction proceeds.
- Substrate binding causes many enzymes to change shape, exposing their active site(s) and allowing catalysis. Review Figure 3.15
- Some enzymes require nonprotein "partners" called **cofactors** to carry out catalysis. Review Table 3.3
- Substrate concentration affects the rate of an enzyme-catalyzed reaction. At the maximum rate, the enzyme is saturated with substrate. Review Figure 3.16

CONCEPT 3.4 Regulation of Metabolism Occurs by Regulation of Enzymes

- Metabolism is organized into pathways in which the product of one reaction is a substrate for the next reaction. A specific enzyme catalyzes each reaction in the pathway.
- Metabolic pathways are integrated into a biochemical system. **Systems biology** is a way to study how biochemical systems behave. Review Figure 3.17
- Enzyme activity is subject to regulation. Some inhibitors bind irreversibly to enzymes. Other inhibitors bind reversibly. Review Figures 3.18 and 3.19 and ANIMATED TUTORIAL 3.2
- In **allosteric regulation**, a molecule binds to a site on the enzyme other than the active site. This changes the overall structure of the enzyme (including that of its active site) and results in either activation or inhibition of the enzyme's catalytic activity. Review Figure 3.20 and ANIMATED TUTORIAL 3.3
- The end product of a metabolic pathway may inhibit an enzyme that catalyzes the "commitment step" of that pathway. This is called **feedback inhibition**. Review Figure 3.21
- Environmental pH and temperature affect enzyme activity. Review Figure 3.22

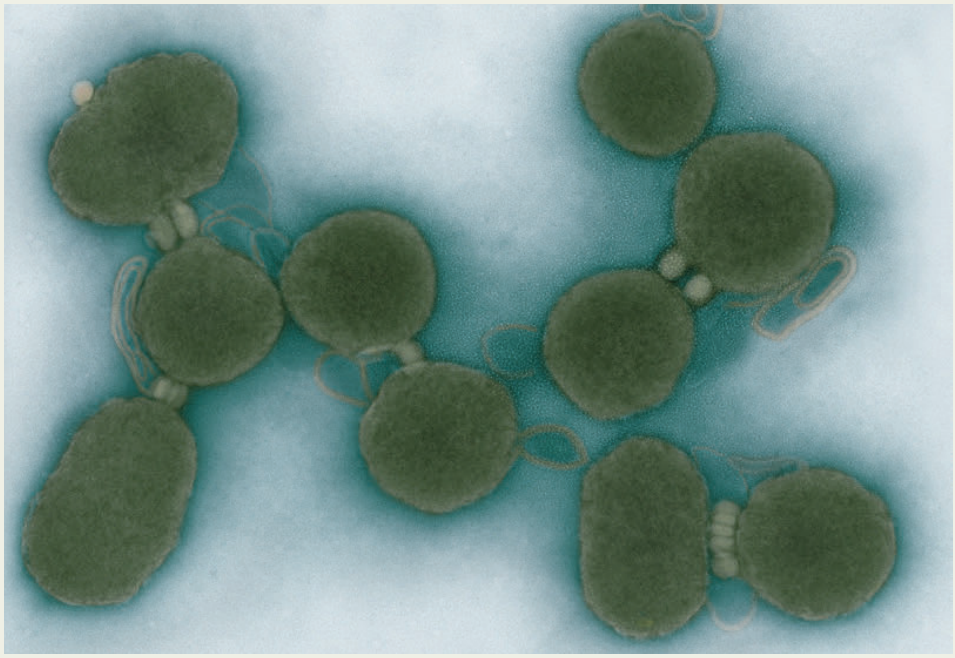
Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is3

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

4

Cells: The Working Units of Life

KEY CONCEPTS



Cells of *Mycoplasma mycoides JCVI-syn1.0*. These are the first synthetic cells.

In 1818, a 21-year-old London writer, Mary Shelley, published a novel that shocked a society in the midst of the Industrial Revolution. In Shelley's story, Dr. Victor Frankenstein discovers how to use electricity to reanimate dead creatures. Collecting body parts from graves and medical labs, the fictional doctor assembles them into a huge 8-foot-tall body and uses his secret method to bring it to life. The results are disastrous, and the novel became a cautionary tale about the limits of science.

Almost 200 years later, in 2010, biologists Craig Venter and Hamilton Smith also gave new life to an "empty shell." In this case, the "shell" was a cell of the tiny bacterium *Mycoplasma capricolum*, into which the scientists inserted a complete new set of genetic material, bringing a new organism to life. The scientists used a computer to design an artificial DNA sequence that had all the genes necessary for bacterial life, plus some unique sequences. Then they went into the chemistry lab and made the DNA from individual nucleotides. They inserted this synthetic genome (similar to the genome of the closely related bacterium *Mycoplasma mycoides*) into the host bacterium, where it replaced the host bacterium's normal DNA. The new DNA directed the cell to perform all the biochemical characteristics of life, including cell reproduction. Eventually, all of the cell's original proteins and RNAs were replaced with proteins and RNAs encoded by the new genome. Since the new genome had some distinctive DNA sequences devised by the scientists, these experiments resulted in an entirely new organism, called *Mycoplasma mycoides JCVI-syn1.0*.

Why did Venter and Smith need to start with a preexisting cell? The chemical reactions of life (metabolism, polymerization, and replication) cannot occur in a dilute aqueous environment; it would be too unlikely for reactants and enzymes to collide with one another. Life requires compartments that bring together and concentrate the molecules involved in these events, which ultimately are directed by the DNA genome.

After about 30 cell divisions, the cells of the new organism no longer had any of the original cell's proteins or small molecules. The cells had used substances in the environment to synthesize their own small and large molecules. They were truly individuals of a new organism, whose "parent" was a synthetic DNA molecule!

The practical aim of this research is to create cells with new capabilities, such as synthesizing clean-burning fuels. But it also puts cells into broader focus as the basic units of biological structure and function.

Q What do the characteristics of modern cells indicate about how the first cells originated?

You will find the answer to this question on page 80.

CONCEPT 4.1 Cells Provide Compartments for Biochemical Reactions

Cells contain water and other small and large molecules, which we examined in Chapters 2 and 3. Each cell contains at least 10,000 different types of molecules, most of them present in many copies. Cells use these molecules to transform matter and energy, to respond to their environments, and to reproduce. As we mentioned in the opening story, these biological processes would not be possible outside the enclosure of a cell.

The **cell theory**, developed in the nineteenth century, recognizes this basic fact about life. It was the first unifying principle of biology and has three critical components:

- Cells are the fundamental units of life.
- All living organisms are composed of cells.
- All cells come from preexisting cells.

Cell theory has two important conceptual implications:

- *Studying cell biology is in some sense the same as studying life.* The principles that underlie the functions of a single bacterial cell are similar to those governing the approximately 60 trillion cells in an adult human.
- *Life is continuous.* All those human cells came from a single cell, a zygote (or fertilized egg). The zygote was formed when two cells fused: a sperm from the father and an egg from the mother. The cells of the parents' bodies were all derived from their parents, and so on back through the generations—all the way back to the evolution of the first living cells.

Cell size can be limited by the surface area-to-volume ratio

Most cells are tiny. Their diameters range from about 1 to 100 micrometers (FIGURE 4.1). There are some exceptions: the eggs of birds are single cells that are, relatively speaking, enormous, and individual cells of several types of algae and bacteria are large enough to be viewed with the unaided eye.

Small cell size is a practical necessity arising from the decrease in the **surface area-to-volume ratio** of any object as it increases in size. As an object increases in volume, its surface area also increases, but not as quickly (FIGURE 4.2). This phenomenon has biological significance for two reasons:

- The *volume* of a cell determines the amount of metabolic activity it carries out per unit of time.
- The *surface area* of a cell determines the amount of substances that can enter it from the outside environment, and the amount of waste products that can exit to the environment.

As a living cell grows larger, its metabolic activity, and thus its need for resources and its rate of waste production, increases faster than its surface area. In addition, substances must move from one location to another within the cell; the smaller the cell, the more easily this is accomplished. The large surface area-to-volume ratio represented by the many small cells of a

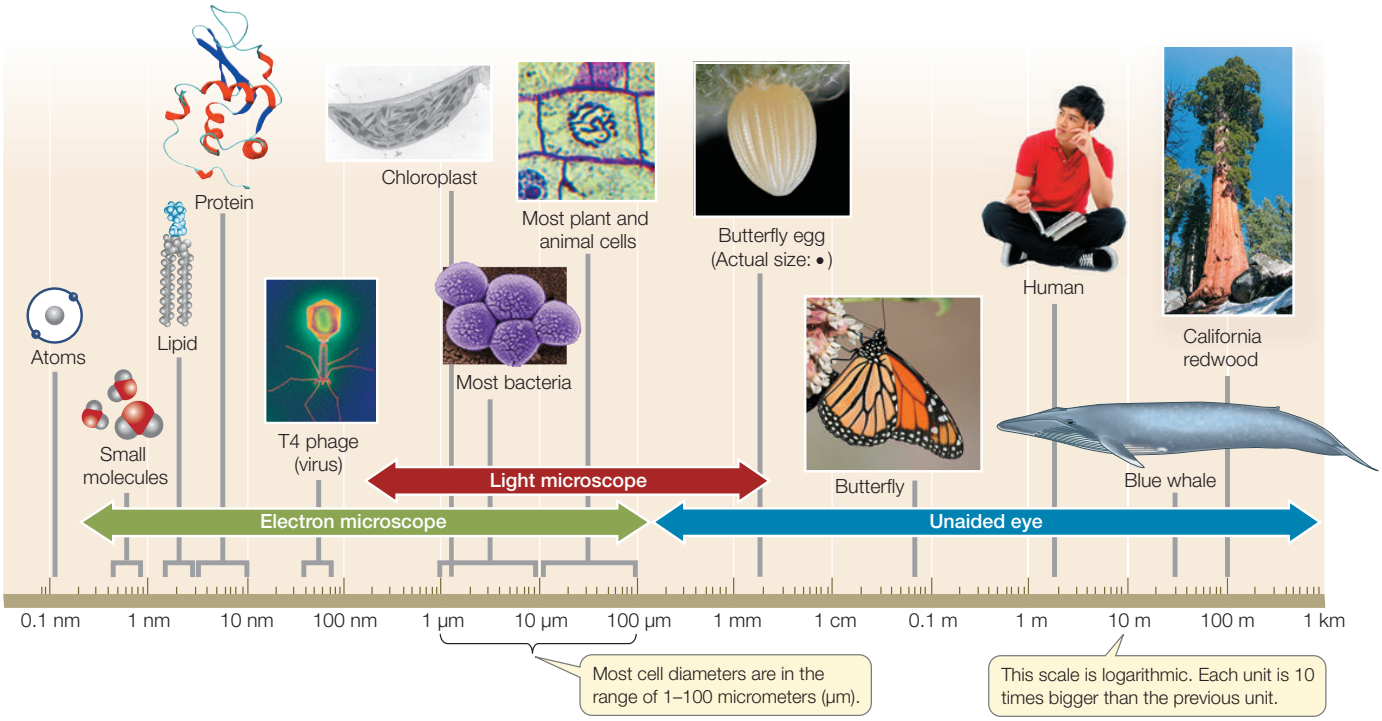
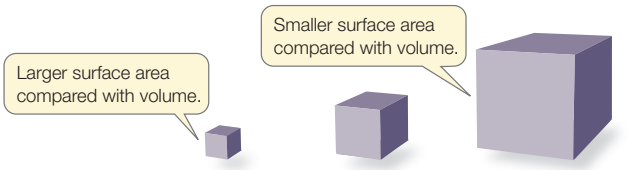


FIGURE 4.1 The Scale of Life This logarithmic scale shows the relative sizes of molecules, cells, and multicellular organisms.

Go to ACTIVITY 4.1 The Scale of Life
PoL2e.com/ac4.1



	1-mm cube	2-mm cube	4-mm cube
Surface area	6 sides × 1^2 = 6 mm^2	6 sides × 2^2 = 24 mm^2	6 sides × 4^2 = 96 mm^2
Volume	1^3 = 1 mm^3	2^3 = 8 mm^3	4^3 = 64 mm^3
Surface area-to-volume ratio	6 mm^{-1} (6:1)	3 mm^{-1} (3:1)	1.5 mm^{-1} (1.5:1)

FIGURE 4.2 Why Cells Are Small As an object grows larger, its volume increases more rapidly than its surface area. Cells must maintain a large surface area-to-volume ratio in order to function. This explains why multicellular organisms must be composed of many small cells rather than a few large ones.

multicellular organism enables it to carry out the many different functions required for survival.

However, for some cell types this general argument does not hold. Many cells are not shaped like cubes, but rather have an irregular shape; as they get larger, they can still have adequate exchange of materials with the environment by increasing their surface area by folds of their cell membrane. In other cases, cells can be quite large without greatly increased membrane surface. For instance, in a giraffe, nerve cells can be several meters long. In these cases, the rate of exchange of materials across the cell membrane must be increased, since the surface area is not adequately larger.

Cells can be studied structurally and chemically

The small sizes of most cells necessitate special instruments to study them and their constituents. For visualizing cells, there are two types of microscopes (**FIGURE 4.3**):

- *A light microscope* uses glass lenses and visible light to form images. Structures that absorb more light are seen as darker than regions that do not, because the absorbed light does not reach the eye. The smallest detail that can be seen with such a microscope is about 0.2 μm in diameter, which is about 1,000 times smaller than an object the human eye can see. Light microscopes are used to visualize living cells and general cell structure.
- *An electron microscope* uses an electron beam focused by magnets to illuminate a specimen and produce an image on a TV-like screen. Structures that absorb the electrons appear darker than regions that do not. The size limit is 0.1 nm, which is 2 million times smaller than something the human eye can see. For electron microscopy, specimens must be preserved and stained using toxic heavy metals, so living cells cannot be visualized this way.

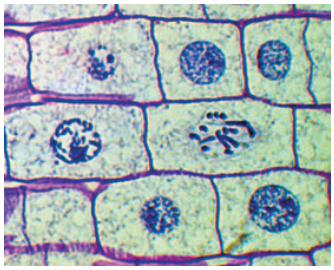
The chemical analysis of cells usually begins with breaking them open to make a cell-free extract. This can be done physically, using a blender or other homogenizing machinery, or by placing the cell in a chemical environment where it swells and bursts (see Figure 5.3). In either case, the resulting extract can be analyzed in terms of its composition and chemical reactions. For example, specific enzyme activities may be measured. If conditions are right in this test tube system, the *properties of the cell-free extract are the same as those inside the cell*. This last statement is of great importance, because it allows biologists to study the chemical processes that occur inside cells in the test tube, so that chemical changes can be easily measured.

A cell's internal structures and even some of its macromolecules can be separated according to their sizes in a centrifuge that spins the tubes at a high speed (**FIGURE 4.4**). Once the subcellular structures are separated from one another, they are much easier to study.

RESEARCH TOOLS

FIGURE 4.3 Microscopy Light and electron microscopes are used to examine cell structures.

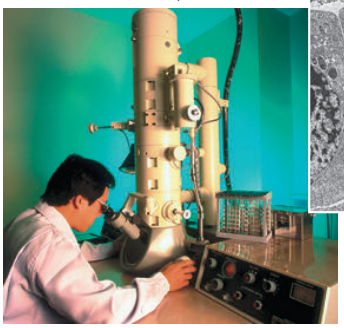
A light microscope

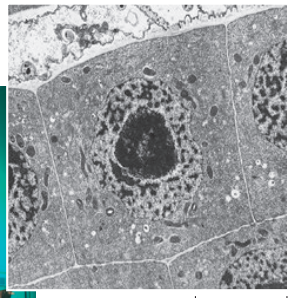


Onion (*Allium cepa*) root tip cells, as viewed through a light microscope.

10 μm

An electron microscope





10 μm

An electron microscope image of an onion root tip cell.

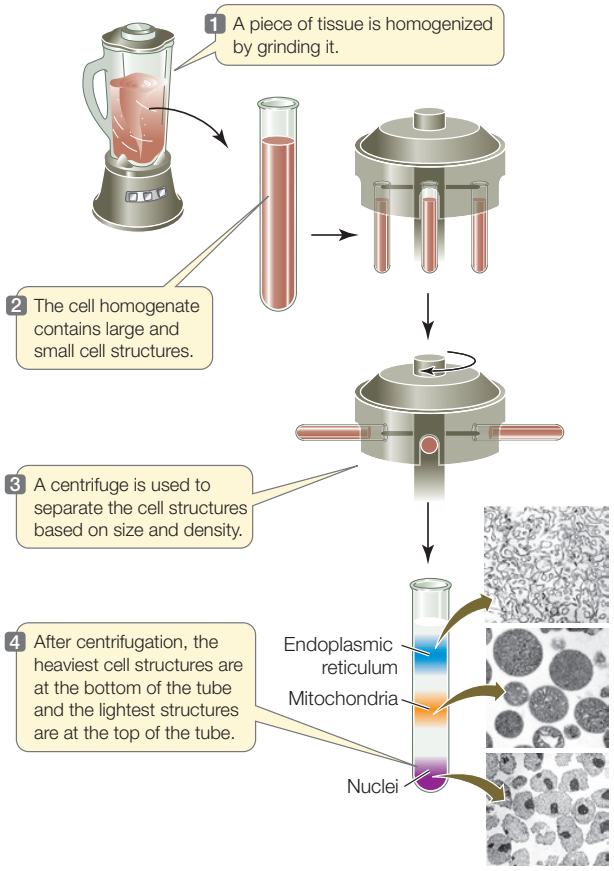
The cell membrane forms the outer surface of every cell

As we will discuss at the end of this chapter, a key to the origin of cells was the enclosure of biochemical functions within a membrane. We will describe the **cell membrane** in more detail in Chapter 5, but for now, note that it consists of a phospholipid bilayer with proteins (see Figure 2.13B). Unless stained for light microscopy, the very thin (7 nm) cell membrane is visible only with electron microscopy. It has several important roles:

- The cell membrane acts as a *selectively permeable barrier*, preventing some substances from crossing it while permitting other substances to enter and leave the cell. In doing

RESEARCH TOOLS

FIGURE 4.4 Centrifugation Structures within cells can be separated from one another on the basis of size and density, and the isolated structures can then be analyzed chemically.



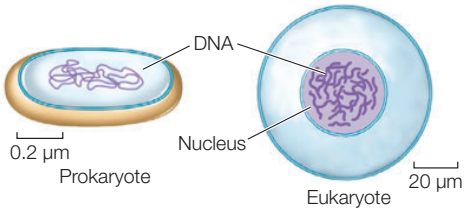
so, it allows the cell to maintain a stable internal environment that is distinct from the surrounding environment. This explains why a red blood cell contains the pigmented molecule hemoglobin but the surrounding blood plasma does not.

- As the cell's boundary with the outside environment, the cell membrane is important in *communicating* with adjacent cells and receiving signals from the environment.
- The cell membrane often has proteins protruding from it that are responsible for *binding and adhering* to adjacent cells or to a surface. Thus the cell membrane plays an important structural role and contributes to cell shape.

Cells are classified as either prokaryotic or eukaryotic

Biologists classify all living things into three domains: Archaea, Bacteria, and Eukarya. The organisms in Archaea and Bacteria are collectively called **prokaryotes** because they have in common a prokaryotic cellular organization. A prokaryotic cell typically does not have membrane-enclosed internal compartments; in particular, it does not have a nucleus.

Eukaryotic cell organization is found in members of the domain Eukarya—the **eukaryotes**—which includes the protists (a diverse group of microorganisms), plants, fungi, and animals. In contrast to the prokaryotes, eukaryotes contain membrane-enclosed compartments called **organelles** where specific metabolic functions occur. The most notable of these is the cell **nucleus**, where most of the cell's DNA is located and where gene expression begins:



Just as a cell is an enclosed compartment that separates its contents from the surrounding environment, each organelle provides a compartment that separates certain molecules and biochemical reactions from the rest of the cell. This impressive "division of labor" provides possibilities for regulation and efficiency that were important in the evolution of complex organisms.

LINK

Eukaryotes arose from prokaryotes by endosymbiosis; see **Concept 20.1**

CHECKpoint CONCEPT 4.1

- ✓ In considering the origin of cells, why do biologists focus on the origin of the cell membrane?
- ✓ If a cell has a cube shape that is 500 µm on a side, what is its surface area-to-volume ratio? If the surface area-to-volume ratio should be more than 0.1 μm^{-1} for optimal cell function, would dividing this cell into 1 million individual cells (also cubes) meet this standard?
- ✓ What is the surface area-to-volume ratio of a giraffe's nerve cell? [For simplicity, assume that it is tubular (a cylinder), with a length of 3 m and a diameter of 5 µm.]
- ✓ What evolutionary advantages does a eukaryotic cell have compared with a prokaryotic cell?

This section has introduced two structural themes in cell architecture: prokaryotic and eukaryotic. We'll turn now to the organization of prokaryotic cells.

CONCEPT 4.2 Prokaryotic Cells Do Not Have a Nucleus

In terms of sheer numbers and diversity, prokaryotes are the most successful organisms on Earth. As we generalize about the features of these cells, bear in mind that there are vast numbers

of prokaryotic species, and that the Bacteria and Archaea can be distinguished from one another in numerous ways. These differences, and the vast diversity of organisms in these two domains, are the subject of Chapter 19.

Prokaryotic cells, with diameters or lengths in the range of 1–10 µm, are generally smaller than eukaryotic cells, whose diameters or lengths are usually in the range of 10–100 µm. While some prokaryotes exist as single cells, other types form chains or small clusters of cells, and in some cases certain cells in a group perform specialized functions.

Prokaryotic cells share certain features

All prokaryotes have the same basic structure (**FIGURE 4.5**):

- The cell membrane encloses the cell, separating its interior from the external environment, and regulates the traffic of materials into and out of the cell.
- The **nucleoid** is a region in the cell where the DNA is located. DNA is the hereditary material that controls cell growth, maintenance, and reproduction (see Chapter 3).
- The rest of the material inside the cell is called the **cytoplasm**. The cytoplasm consists of a liquid component, the cytosol, and a variety of insoluble filaments and particles, the most abundant of which are ribosomes (see below).
- The **cytosol** consists mostly of water containing dissolved ions, small molecules, and soluble macromolecules such as proteins.
- **Ribosomes** are complexes of RNA and proteins that are about 25 nm in diameter. They can be visualized only with the electron microscope. They are the sites of protein synthesis, where the information encoded by nucleic acids directs the sequential linking of amino acids to form proteins.

The cytoplasm is not a static region. Rather, the substances in this environment are in constant motion. For example, a typical protein moves around the entire cell within a minute, and it collides with many other molecules along the way. This constant motion helps ensure that biochemical reactions proceed at sufficient rates to meet the needs of the cell. Prokaryotes may look simple, but in reality they are functionally complex, carrying out thousands of biochemical reactions.

Specialized features are found in some prokaryotes

As they evolved, some prokaryotes developed specialized structures that gave them a selective advantage in their particular environments. These cells were better able to survive and reproduce than cells lacking the specialized structures.

CELL WALLS Most prokaryotes have a **cell wall** located outside the cell membrane. The rigidity of the cell wall supports the cell and determines its shape. The cell walls of most bacteria, but not those of archaea, contain peptidoglycan, a polymer of carbohydrates that is linked at regular intervals to short peptides. Cross-linking among these peptides results in a single giant molecule that surrounds the entire cell. In some bacteria, another layer, the outer membrane (a polysaccharide-rich phospholipid membrane), encloses the peptidoglycan layer (see Figure 4.5). Unlike the cell membrane, this outer membrane is relatively permeable, allowing the movement of molecules across it.

Enclosing the cell wall of some bacteria is a slimy layer composed mostly of polysaccharides, referred to as the capsule. In some cases these capsules protect the bacteria from attack by white blood cells in the animals they infect. Capsules also help keep the cells from drying out, and sometimes they help bacteria attach to other cells. Many prokaryotes produce no capsule, and those that do have capsules can survive even if they lose

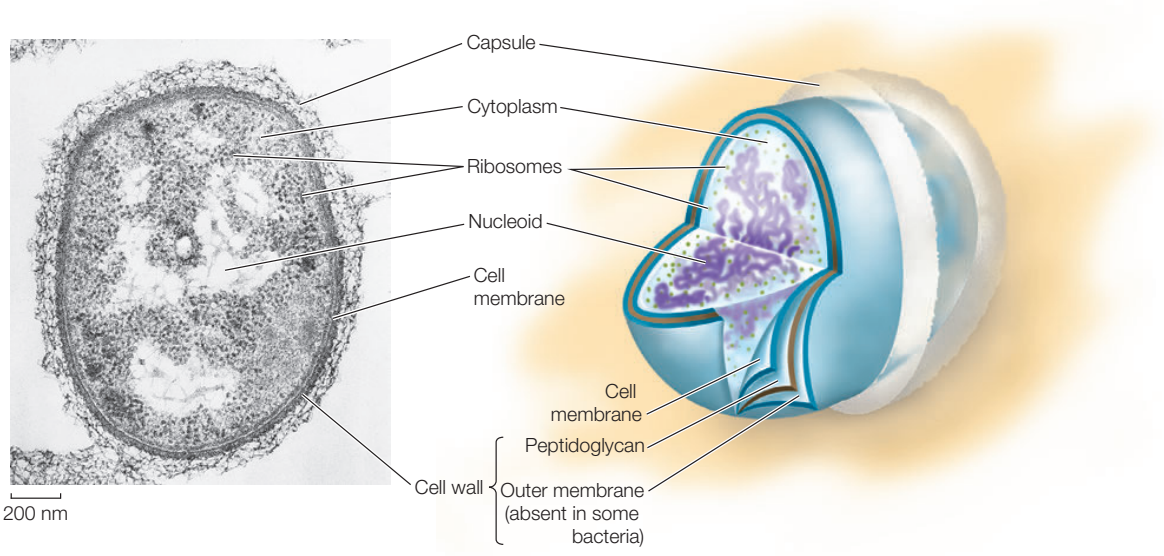


FIGURE 4.5 A Prokaryotic Cell The bacterium *Pseudomonas aeruginosa* illustrates the typical structures shared by all prokaryotic cells. This bacterium also has a protective outer membrane and a capsule, which are not present in all prokaryotes.

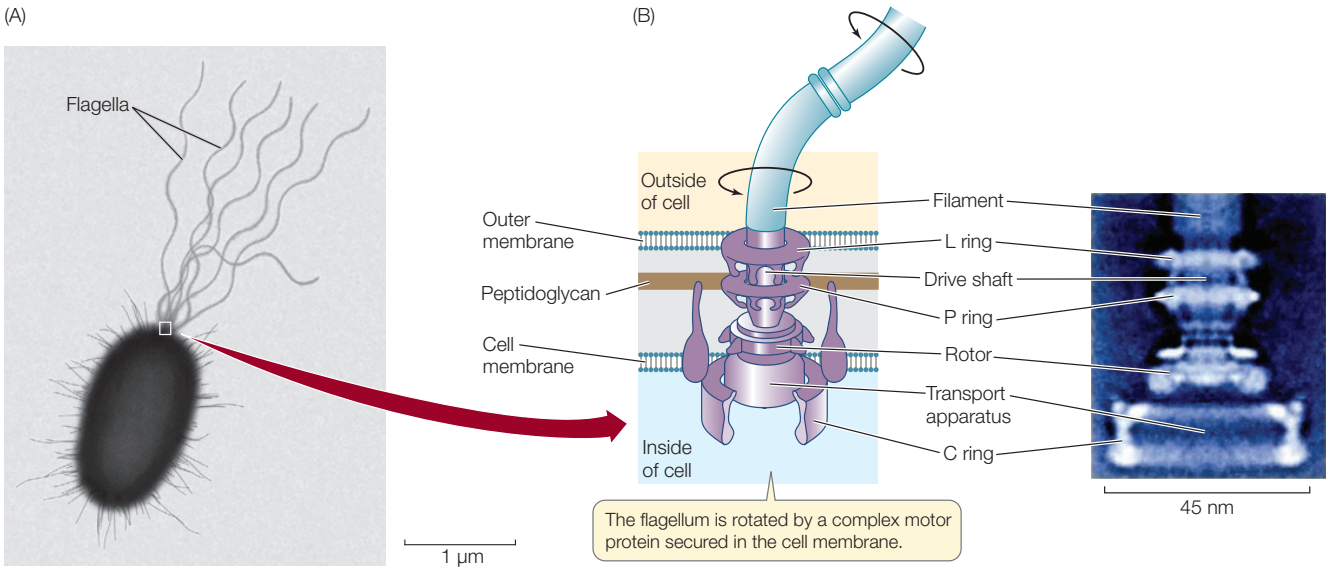


FIGURE 4.6 Prokaryotic Flagella (A) Flagella contribute to the movement and adhesion of prokaryotic cells. (B) Complex protein ring structures anchored in the cell membrane form a motor unit that rotates the flagellum and propels the cell.

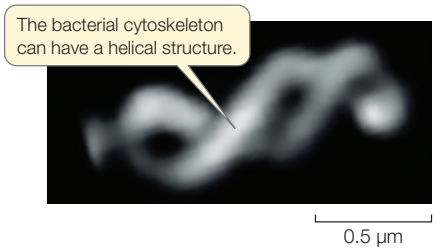
them, so the capsule is not essential to prokaryotic life. Some strains of the bacterium *Streptococcus pneumoniae* have a capsule and can cause pneumonia in humans and other mammals; however, non-encapsulated strains do not cause the disease.

INTERNAL MEMBRANES Some groups of bacteria—including the cyanobacteria—carry out photosynthesis: they use energy from the sun to convert carbon dioxide and water into carbohydrates. These bacteria have an **internal membrane** system that contains molecules needed for photosynthesis. The development of photosynthesis, which requires membranes, was an important event in the early evolution of life on Earth. Other prokaryotes have internal membrane folds that are attached to the cell membrane. These folds may function in cell division or in various energy-releasing reactions.

A bacterium with enclosed compartments would have several evolutionary advantages. Chemicals could be concentrated within particular regions of the cell, allowing chemical reactions to proceed more efficiently. Certain biochemical activities could be segregated within compartments with more favorable conditions for those reactions, such as a different pH from the rest of the cell.

FLAGELLA Some cells swim by using appendages called **flagella**, which sometimes look like tiny corkscrews (**FIGURE 4.6A**). These movements are important in allowing the cells to swim toward food, for example. In bacteria, the filament of the flagellum is made of a protein called flagellin. A complex motor protein (see Concept 4.4) spins each flagellum on its axis like a propeller, driving the cell along. This motor protein is anchored to the cell membrane and, in some bacteria, to the outer membrane of the cell wall (**FIGURE 4.6B**). We know that flagella cause the motion of cells because if they are removed, the cells do not move. Flagellar motion can result in impressive speeds. The prokaryote *Methanocaldococcus* has been clocked at over 400 μm (500 body lengths) per second. For a car, this would mean 4,000 miles per hour.

CYTOSKELETON The **cytoskeleton** is the collective name for filaments made up of polymers of monomer subunits that play roles in cell division or in maintaining the shapes of cells. One such protein forms helical structures that extend down the lengths of rod-shaped bacterial cells, helping maintain their shapes. This protein is similar to actin in eukaryotic cells, which we will discuss next.



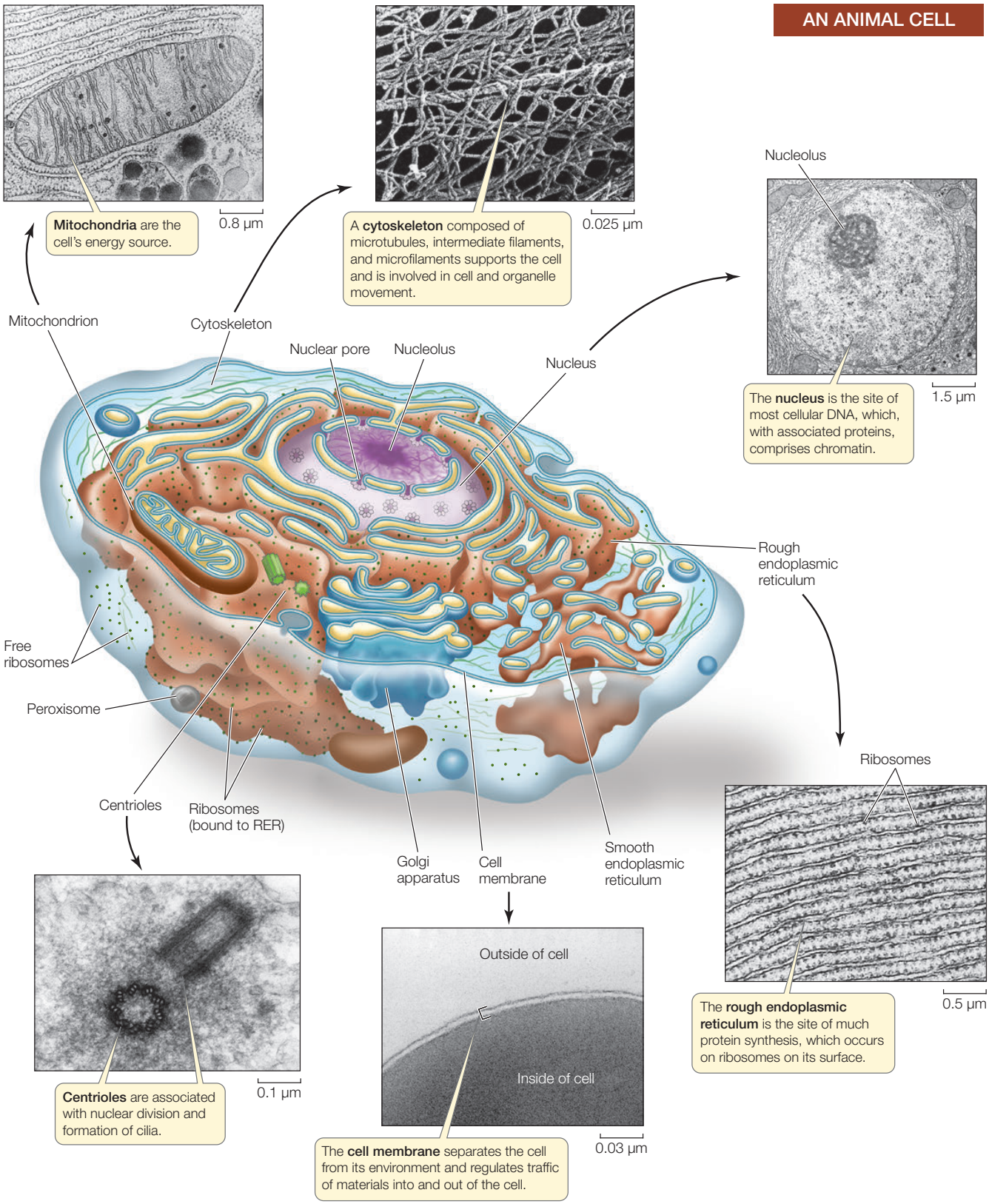
CHECKpoint CONCEPT 4.2

- ✓ Compare the structures and functions of bacterial cell walls with those of bacterial cytoskeletons.
- ✓ What is the evolutionary advantage of bacteria that have flagella over bacteria that do not?

The prokaryotic cell is one of two broad types of cells recognized in cell biology. The other is the eukaryotic cell. Eukaryotic cells, and multicellular eukaryotic organisms, are more structurally and functionally complex than prokaryotic cells.

CONCEPT 4.3 Eukaryotic Cells Have a Nucleus and Other Membrane-Bound Compartments

Like prokaryotic cells, eukaryotic cells have a cell membrane, cytoplasm, and ribosomes. But as we noted earlier in this chapter, eukaryotic cells also have organelles within the cytoplasm


AN ANIMAL CELL
Mitochondria are the cell's energy source.
0.8 µm
A cytoskeleton composed of microtubules, intermediate filaments, and microfilaments supports the cell and is involved in cell and organelle movement.
0.025 µm
Nucleolus
The nucleus is the site of most cellular DNA, which, with associated proteins, comprises chromatin.
1.5 µm
Mitochondrion
Cytoskeleton
Nuclear pore
Nucleolus
Nucleus
Rough endoplasmic reticulum
Free ribosomes
Peroxisome
Centrioles
Ribosomes (bound to RER)
Golgi apparatus
Cell membrane
Smooth endoplasmic reticulum
Ribosomes
The rough endoplasmic reticulum is the site of much protein synthesis, which occurs on ribosomes on its surface.
0.5 µm
Centrioles are associated with nuclear division and formation of cilia.
0.1 µm
Outside of cell
Inside of cell
The cell membrane separates the cell from its environment and regulates traffic of materials into and out of the cell.
0.03 µm

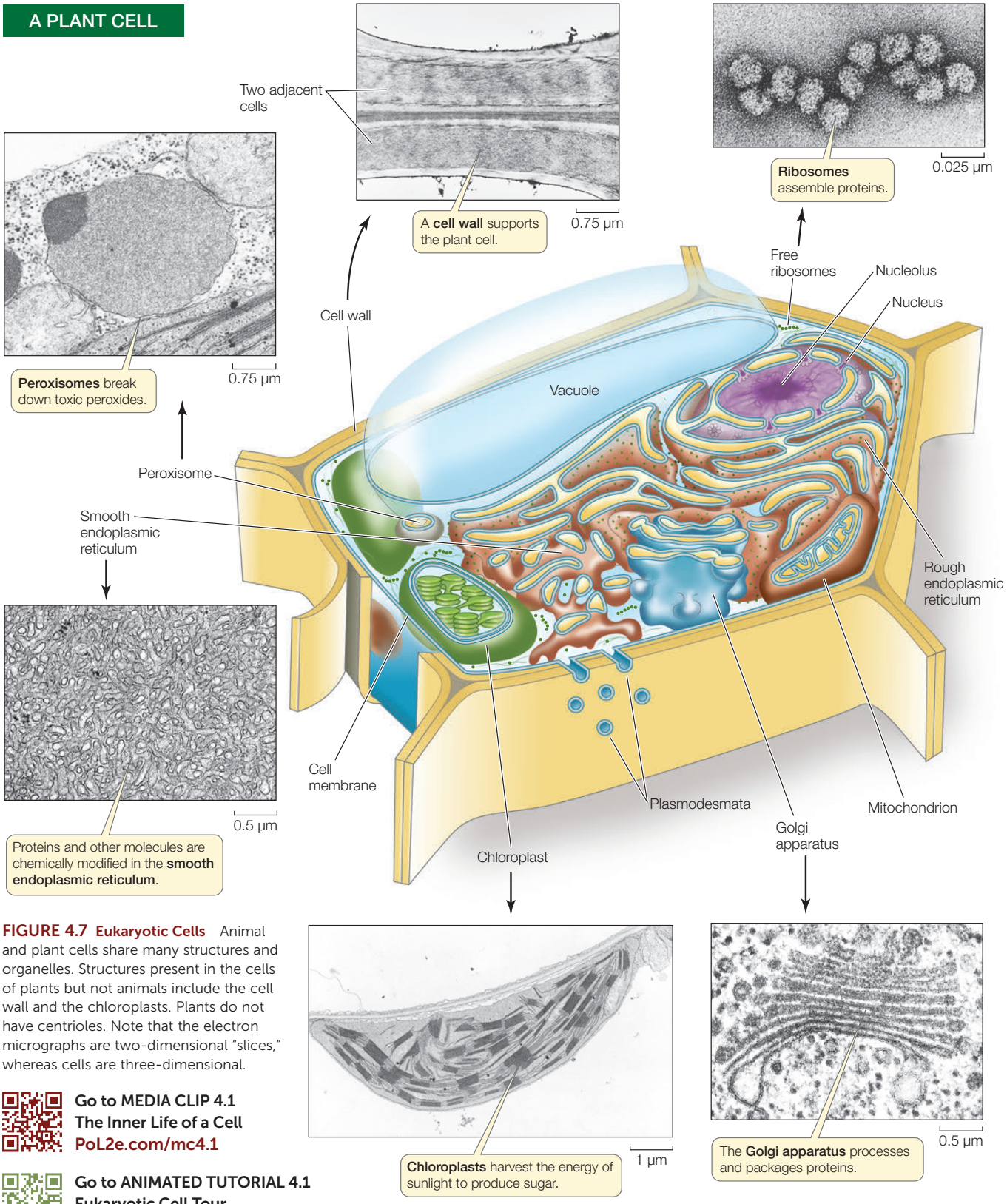


FIGURE 4.7 Eukaryotic Cells Animal and plant cells share many structures and organelles. Structures present in the cells of plants but not animals include the cell wall and the chloroplasts. Plants do not have centrioles. Note that the electron micrographs are two-dimensional "slices," whereas cells are three-dimensional.

Go to MEDIA CLIP 4.1
The Inner Life of a Cell
PoL2e.com/mc4.1

Go to ANIMATED TUTORIAL 4.1
Eukaryotic Cell Tour
PoL2e.com/at4.1

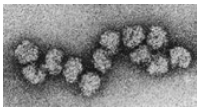
whose interiors are separated from the cytosol by membranes (FIGURE 4.7).

Compartmentalization is the key to eukaryotic cell function

Each type of organelle has a specific role in the cell. Some organelles have been characterized as factories that make specific products. Others are like power plants that take in energy in one form and convert it into a more useful form. In addition, eukaryotic cells have some structures that are analogous to those seen in prokaryotes. For example, they have a cytoskeleton composed of protein fibers, and outside the cell membrane, an extracellular matrix.

When animal and plant cells are examined using electron microscopy, they have many organelles and structures in common—the most obvious is the cell nucleus. But they also have some differences. For example, many plant cells have chloroplasts that are colored green by the pigment used in photosynthesis. Figure 4.7 shows diagrams of an animal and a plant cell, with electron micrographs of some of the subcellular structures.

Ribosomes are factories for protein synthesis



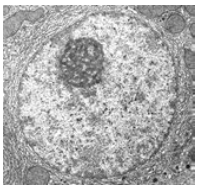
Ribosomes translate the nucleotide sequence of a messenger RNA molecule into a polypeptide chain. The ribosomes of both prokaryotes and eukaryotes consist of one larger and one smaller subunit; the sizes of the subunits differ between the two cell types. Each subunit consists of one to three large RNA molecules called ribosomal RNA (rRNA) and multiple smaller protein molecules that are bound noncovalently to one another and to the rRNA. The ribosome is an amazingly precise structure. If the individual macromolecules are separated by disruption of their hydrophobic interactions, they will spontaneously reassemble into a functional complex.

LINK

Protein synthesis is described in more detail in Concept 10.4

Ribosomes are not membrane-enclosed compartments. In prokaryotic cells, ribosomes float freely in the cytoplasm. In eukaryotic cells they are found in the cytoplasm, where they may be free or attached to the surface of the endoplasmic reticulum (a membrane-enclosed organelle; see below), and also inside certain organelles—namely the mitochondria and the chloroplasts.

The nucleus contains most of the cell's DNA



As we noted in Chapter 3, hereditary information is stored in the sequence of nucleotides in DNA molecules. In eukaryotic cells, most of the DNA is in the nucleus. Most cells have a single nucleus, and it is usually the largest organelle; at 5 μm in diameter, the nucleus is substantially larger than most prokaryotic cells. The nucleus has several functions:

- It is the location of the DNA and the site of DNA replication.
- It is where DNA is transcribed into RNA (see Concept 3.1).
- It contains the **nucleolus**, a region where ribosomes begin to assemble from RNA and proteins.

As you can see in Figure 4.7, the nucleus is enclosed by not one but two membranes: two lipid bilayers that together form the nuclear envelope. Functionally, this barrier separates DNA transcription (which occurs in the nucleus) from translation (which occurs in the cytoplasm). The two membranes of the nuclear envelope are perforated by thousands of nuclear pores, each measuring approximately 9 nm in diameter, which connect the interior of the nucleus to the cytoplasm. The pores regulate traffic between these two cellular compartments by allowing some molecules to enter or leave the nucleus and by blocking others. This allows the nucleus to regulate its information-processing functions.

Inside the nucleus, each DNA molecule is combined with proteins to form exceedingly long, thin threads called **chromosomes**. Different eukaryotic organisms have different numbers of chromosomes (ranging from two in one kind of Australian ant to hundreds in some plants). These DNA–protein complexes, which are also called **chromatin**, become much more compact during cell division, as you will see in Concept 7.2.

The outer membrane of the nuclear envelope folds outward into the cytoplasm and is continuous with the membrane of another organelle, the endoplasmic reticulum (see Figure 4.7).

The endomembrane system is a group of interrelated organelles

Much of the volume of many eukaryotic cells is taken up by an extensive **endomembrane system**. This interconnected system of membrane-enclosed compartments includes the nuclear envelope, endoplasmic reticulum, Golgi apparatus, and lysosomes, which are derived from the Golgi apparatus. Tiny, membrane-surrounded droplets called **vesicles** shuttle substances between the various components of the endomembrane system, as well as the cell membrane (FIGURE 4.8). In drawings and electron micrographs this system appears static, but in the living cell, membranes and the materials they contain are in constant motion. Membrane components have been observed to shift from one organelle to another within the endomembrane system. This suggests that all these membranes must be functionally related.

ENDOPLASMIC RETICULUM Electron micrographs of eukaryotic cells reveal networks of interconnected membranes branching throughout the cytoplasm, forming tubes and flattened sacs about 1 μm across. These membranes are collectively called the **endoplasmic reticulum**, or **ER**. The interior compartment (lumen) of the ER is separate and distinct from the surrounding cytoplasm (see Figure 4.8). The ER can enclose up to 10 percent of the interior volume of the cell, and its extensive folding results in a surface area many times greater

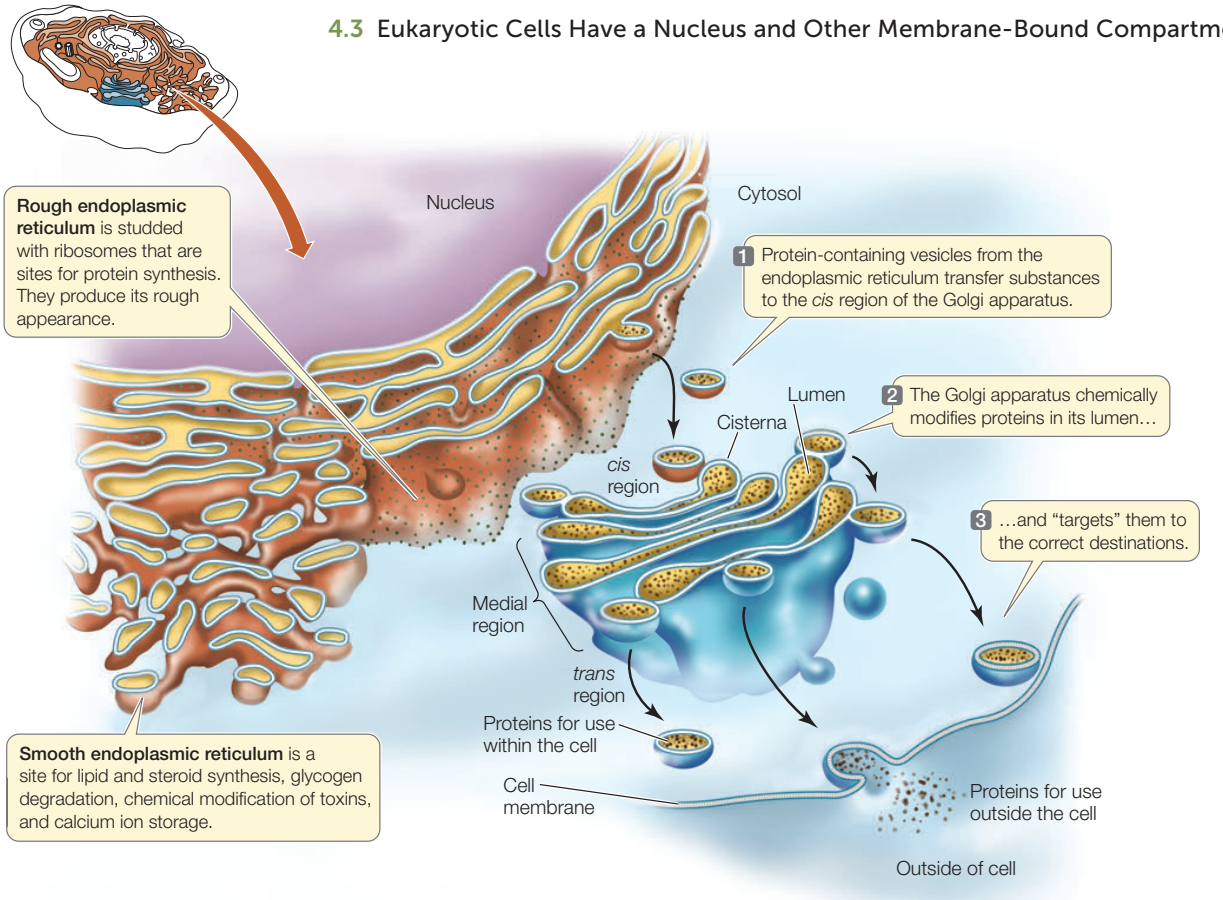


FIGURE 4.8 The Endomembrane System Membranes of the nucleus, endoplasmic reticulum (ER), and Golgi apparatus form a network that is connected by vesicles. Parts of the membrane move between these organelles. Membrane synthesized in the smooth ER becomes sequentially part of the rough ER, then the Golgi apparatus, then vesicles formed from the Golgi apparatus. These vesicles may eventually fuse with, and become part of, the cell membrane.

Go to ANIMATED TUTORIAL 4.2
The Golgi Apparatus
PoL2e.com/at4.2

APPLY THE CONCEPT

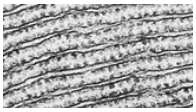
Eukaryotic cells have a nucleus and other membrane-bound compartments

Imagine that a group of scientists wanted to trace the path that the enzyme lipase follows between its site of synthesis and its final destination within a liver cell. First, the scientists exposed cultured liver cells to radioactive amino acids for 3 minutes. The amino acids entered the cells and became incorporated into all proteins synthesized during that time period. Then the radioactive amino acids were removed, so any proteins synthesized subsequently were *not* radioactive. At 5-minute intervals after the brief exposure to radioactive amino acids, some of the cells were broken open and fractionated to separate the organelles, as shown in Figure 4.4. An antibody that binds specifically to lipase was used to measure how much radioactive lipase was in each organelle at each time point (see Concept 39.4 for information on antibodies). The table shows the results.

	PERCENTAGE OF RADIOACTIVE LIPASE			
TIME (MIN)	ER LUMEN	GOLGI APPARATUS	LYSOSOMES	RIBOSOMES
5	5	0	0	95
10	25	10	0	65
15	75	20	5	0
20	25	55	20	0
25	0	65	35	0
30	0	25	75	0
35	0	0	100	0

1. Plot percentage radioactive lipase versus time for each organelle. What can you conclude about the pathway of lipase in the cell after it is synthesized?
2. Lipase breaks down lipids. Why is its organelle destination appropriate (see Figure 4.9)?

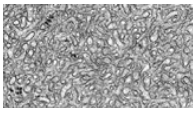
than that of the cell membrane. There are two types of ER: rough and smooth.



The **rough endoplasmic reticulum** (**RER**) is called "rough" because of the many ribosomes attached to the outer surface of the membrane, giving it a rough appearance in electron micrographs. These ribosomes are not permanently attached to the ER but become attached when they begin synthesizing proteins destined for modification within the RER:

- A protein enters the RER only if it contains a specific short sequence of amino acids that signals the ribosome to attach to the RER (see Concept 10.5).
- Once inside the RER, proteins are chemically modified to induce their three-dimensional functional shape and to chemically "tag" them for delivery to specific cellular destinations.
- The RER participates in transporting these proteins to other locations in the cell. The proteins are transported in vesicles that pinch off from the ER. All secreted proteins pass through the RER.
- Most membrane-bound proteins are made on the RER.

A polypetide that is synthesized on the RER surface is transported across the membrane and into the lumen while it is being translated. Once inside, it undergoes several changes, including the formation of disulfide bridges and folding into its tertiary structure. Many proteins are covalently linked to carbohydrate groups, thus becoming **glycoproteins**. These carbohydrate groups often have roles in recognition. For example, the carbohydrate groups on some secreted glycoproteins play roles in recognition and interactions between cells. Other carbohydrate groups "tag" their proteins for transfer to specific cellular locations. This "addressing" system is very important for ensuring that proteins arrive at their correct destinations. For example, the enzymes within the lysosomes (see below) are highly destructive and could destroy the cell if they were released into the cytosol.



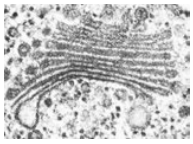
The **smooth endoplasmic reticulum** (**SER**) is connected to portions of the RER but lacks ribosomes and is more tubular (less like flattened sacs) than the RER. The SER has four important roles:

- It is responsible for the chemical modification of small molecules taken in by the cell that may be toxic to the cell. These modifications make the targeted molecules more polar, so they are more water-soluble and easily removed.
- It is the site for glycogen degradation in animal cells.
- It is the site where lipids and steroids are synthesized.
- It stores calcium ions, which when released trigger a number of cell responses, including muscle contraction.

LINK

The role of the SER in muscle contraction is described in **Concept 33.1**

Cells that synthesize a lot of protein for export are usually packed with RER. Examples include glandular cells that secrete digestive enzymes and white blood cells that secrete antibodies. In contrast, cells that carry out less protein synthesis (such as storage cells) contain less RER. Liver cells, which modify molecules (including toxins) that enter the body from the digestive system, have abundant SER.



GOLGI APPARATUS The **Golgi apparatus** is named after its discoverer, Camillo Golgi. It has two components: flattened membranous sacs called cisternae (singular cisterna), which are piled up like saucers in a stack about 1 μm thick, and small membrane-enclosed vesicles (see Figure 4.8). There can be many of these stacks in a cell.

When protein-containing vesicles from the RER fuse with the Golgi apparatus membrane, the proteins are released into the lumen of a Golgi apparatus cisterna, where they may be further modified. The Golgi apparatus has several roles:

- It concentrates, packages, and sorts proteins before they are sent to their cellular or extracellular destinations.
- It adds some carbohydrates to proteins.
- It is where some polysaccharides for the plant cell wall are synthesized.

The cisternae of the Golgi apparatus have three functionally distinct regions: the *cis* region lies nearest to the nucleus or a patch of RER, the *trans* region lies closest to the cell membrane, and the medial region lies in between (see Figure 4.8). (The terms *cis*, *trans*, and medial derive from Latin words meaning "on the same side," "on the opposite side," and "in the middle," respectively.) These three parts of the Golgi apparatus contain different enzymes and perform different functions.

Protein-containing vesicles from the ER fuse with the *cis* membrane of the Golgi apparatus. Other vesicles may transport proteins from one cisterna to the next, although it appears that some proteins move between cisterna through tiny channels. Vesicles budding off from the *trans* region carry their contents away from the Golgi apparatus. These vesicles go to the cell membrane or to the lysosome.

LYSOSOMES The **primary lysosomes** originate from the Golgi apparatus. They contain hydrolases (digestive enzymes), and they are the sites where macromolecules—proteins, polysaccharides, nucleic acids, and lipids—are hydrolyzed into their monomers (see Chapter 2):

$$R_1\text{-}R_2 \text{ (linked monomers)} + H_2O \rightarrow R_1\text{-}OH + R_2\text{-}H$$

For example,

$$\text{Polypeptide} + H_2O \rightarrow \text{amino acids}$$

A lysosome is about 0.5 μm in diameter; it is surrounded by a single membrane and has a densely staining, featureless interior (**FIGURE 4.9**). There may be dozens of lysosomes in a cell.

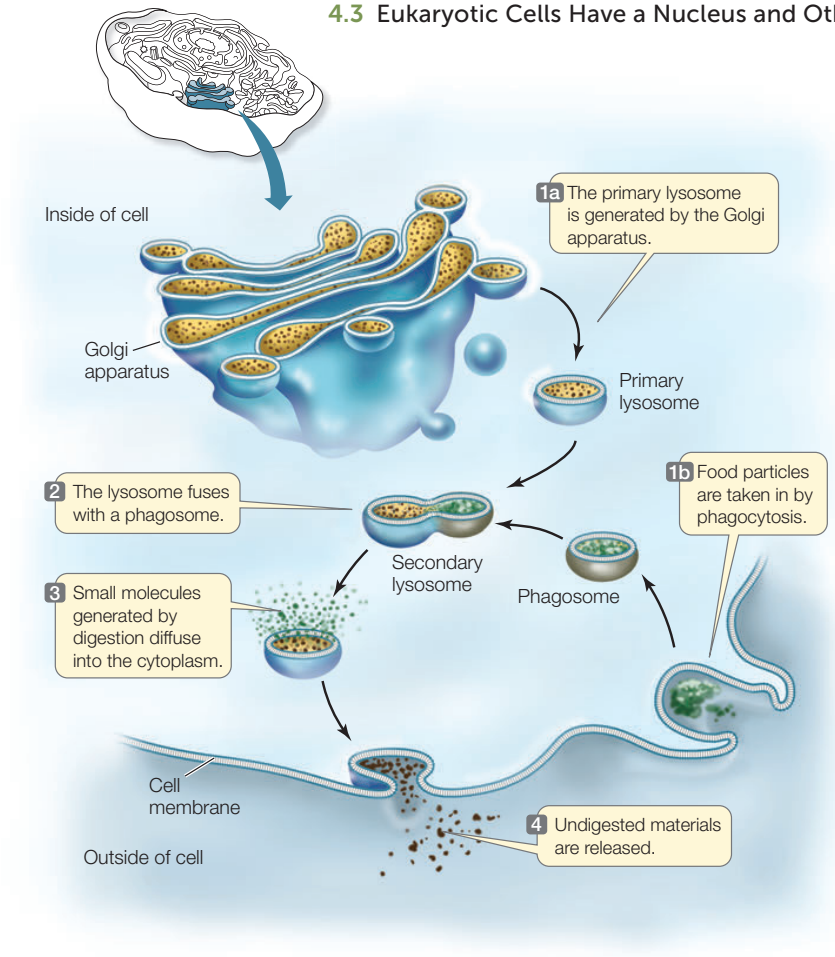
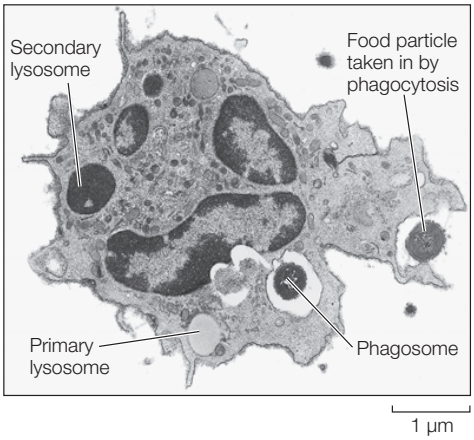


FIGURE 4.9 Lysosomes Isolate Digestive Enzymes from the Cytoplasm Lysosomes are sites for the hydrolysis of material taken into the cell by phagocytosis.

Go to ACTIVITY 4.2 Lysosomal Digestion
PoL2e.com/ac4.2



Some macromolecules that are hydrolyzed in lysosomes enter from the environment outside the cell by a process called **phagocytosis** (*phago*, "eat"; *cytosis*, "cellular"). In this process, a pocket forms in the cell membrane and then deepens and encloses material from outside the cell. The pocket becomes a small vesicle containing macromolecules (e.g., proteins), called a **phagosome**, which breaks free of the cell membrane to move into the cytoplasm. The phagosome fuses with a primary lysosome to form a **secondary lysosome**, in which hydrolysis occurs. The products of digestion (e.g., amino acids) pass through the membrane of the lysosome, providing monomers for other cellular processes. The "used" secondary lysosome, now containing undigested particles, then moves to the cell membrane, fuses with it, and releases the undigested contents to the environment.

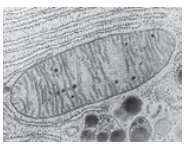
Phagocytes are specialized cells whose major role is to take in and break down materials; they are found in nearly all animals and many protists. However, lysosomes are active even in cells that do not perform phagocytosis. All cells continually break down some of their components and replace them with new ones. The programmed destruction of cell components is called **autophagy**, and lysosomes are where the cell breaks down its own materials, even entire organelles, hydrolyzing their constituents.

How important is autophagy? An entire class of human diseases called lysosomal storage diseases occur when lysosomes fail to digest internal components; these diseases are often very harmful or fatal. For example, Tay-Sachs disease occurs when a particular lipid called a ganglioside is not broken down in the lysosomes and instead accumulates in brain cells and damages them. In the most common form of this disease, a baby starts exhibiting neurological symptoms and becomes blind, deaf, and unable to swallow after six months of age. Death occurs before age four.

Plant cells do not appear to contain lysosomes, but the central vacuole of a plant cell (which we will describe below) may function in an equivalent capacity because it, like lysosomes, contains many digestive enzymes.

Some organelles transform energy

A cell requires energy to make the molecules it needs for activities such as growth, reproduction, responsiveness, and movement. Mitochondria (found in all eukaryotic cells) harvest chemical energy, whereas chloroplasts (found in plants and other photosynthetic cells) harvest energy from sunlight.



MITOCHONDRIA In eukaryotic cells, the breakdown of energy-rich molecules such as the monosaccharide glucose begins in the cytosol. The molecules that result from this partial degradation enter the **mitochondrion** (plural *mitochondria*), whose primary function is to harvest the chemical energy of those molecules in a form the cell

can use, namely the energy-rich nucleotide ATP (adenosine triphosphate). We will discuss these energy-harvesting processes in Chapter 6.

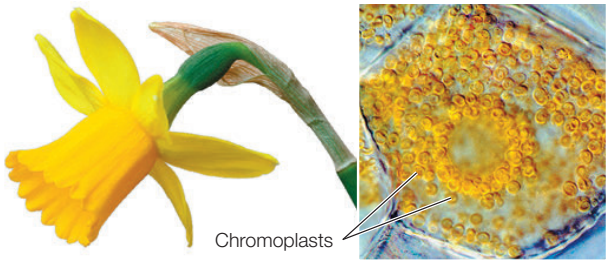
A typical mitochondrion is somewhat less than 1.5 µm in diameter and 2–8 µm in length—about the size of many bacteria. It contains some DNA and can divide independently of the central nucleus. The number of mitochondria per cell ranges from one gigantic organelle in some unicellular protists to a few hundred thousand in large egg cells. An average human liver cell contains more than 1,000 mitochondria. Cells that are active in movement and growth require the most chemical energy, and these tend to have the most mitochondria per unit of volume.

Mitochondria have two membranes. The outer membrane has large pores, and most substances can pass through it. The inner membrane separates the biochemical processes of the mitochondrion from the surrounding cytosol. The inner membrane is extensively folded into structures called cristae, and the fluid-filled region inside the inner membrane is referred to as the mitochondrial matrix. The mitochondrion contains many enzymes for energy metabolism, as well as DNA and ribosomes for the synthesis of a small proportion of the mitochondrial proteins.

PLASTIDS Plastids are present in the cells of plants and algae, and like mitochondria, they can divide autonomously. Plastids can differentiate into a variety of organelles, some of which are used for the storage of pigments (as in flowers), carbohydrates (as in potatoes), lipids, or proteins. An important type of plastid is the **chloroplast**, which contains the green pigment chlorophyll and is the site of photosynthesis (see Concepts 6.5 and 6.6). Photosynthesis is an anabolic process that converts light energy into the chemical energy contained in bonds between the atoms of carbohydrates.

A chloroplast is enclosed within two membranes. In addition, it contains a series of internal membranes that look like stacks of flat, hollow discs, called **thylakoids**. Each stack of thylakoids is called a granum (plural grana). Light energy is converted to chemical energy on the thylakoid membranes. The aqueous fluid surrounding the thylakoids is called the stroma, and it is there that carbohydrates are synthesized. Like the mitochondrial matrix, the chloroplast stroma contains ribosomes and DNA, which are used to synthesize some of the chloroplast proteins.

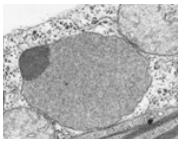
Other types of plastids have functions different from those of chloroplasts. Chromoplasts make and store red, yellow, and orange pigments, especially in flowers and fruits:



Leucoplasts are storage organelles for macromolecules such as starch:

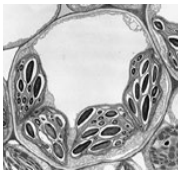
Several other membrane-enclosed organelles perform specialized functions

There are several other kinds of membrane-bound organelles with specialized functions: peroxisomes, glyoxysomes, and vacuoles, including contractile vacuoles.



PEROXISOMES **Peroxisomes** are small (0.2–1 µm diameter) organelles that accumulate toxic peroxides, such as hydrogen peroxide (H_2O_2), which occur as the by-products of some biochemical reactions in all eukaryotes. These peroxides are safely broken down inside the peroxisomes without mixing with other components of the cell. A peroxisome has a single membrane and a granular interior containing specialized enzymes.

GLYOXYSOMES **Glyoxysomes** are found only in plants. They are most abundant in young plants, which have many in each cell, and are the locations where stored lipids are converted into carbohydrates for transport to growing cells.



VACUOLES **Vacuoles** occur in many eukaryotic cells, but particularly those of plants and fungi (see Figure 4.7). There can be one large vacuole or many small ones in a cell. Plant vacuoles have several functions:

- *Storage*: Like all cells, plant cells produce a variety of toxic by-products and waste products. Plants store many of these in vacuoles. Because they are poisonous or distasteful, these stored materials deter some animals from eating the plants, and may thus contribute to the plants' defenses and survival.
- *Structure*: In many plant cells, enormous vacuoles take up more than 90 percent of the cell volume and grow as the cell grows. The presence of dissolved substances in the vacuole causes water to enter it from the cytoplasm (which in turn takes up water from outside the cell), making the vacuole swell like a water-filled balloon. (This osmotic effect is illustrated in Figure 5.3.) The plant cell wall resists the swelling, causing the cell to stiffen from the increase in water pressure. This pressure is called turgor pressure, and it helps support the plant.

- *Reproduction*: Some pigments in the petals and fruits of flowering plants are contained in vacuoles. These pigments—the red, purple, and blue anthocyanins—are visual cues that help attract animals, which assist in pollination and seed dispersal.
- *Catabolism*: In the seeds of some plants, the vacuoles contain enzymes that hydrolyze stored seed proteins into monomers. The developing plant seedling uses these monomers as building blocks and sources of energy.

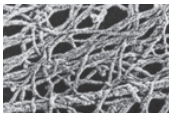
Many freshwater protists have contractile vacuoles. Their function is to get rid of the excess water that rushes into the cell because of the imbalance in solute concentration between the interior of the cell and its freshwater environment. The contractile vacuole enlarges as water enters, and then abruptly contracts, forcing the water out of the cell through a special pore structure.

CHECKpoint CONCEPT 4.3

- ✓ Make a table that summarizes eukaryotic cell organelles with regard to typical size, numbers per cell, and functions.
- ✓ What are some functions of the cell nucleus? What are the advantages of confining these functions within the nucleus, separated from the cytoplasm?
- ✓ Compare the structural and functional differences between rough and smooth endoplasmic reticulum.
- ✓ In I-cell disease, an enzyme in the endomembrane system that normally adds phosphorylated sugar groups to proteins is lacking, and the proteins are not targeted to the lysosomes as they would be in normal cells. The "I" stands for inclusion bodies that appear in the cells. What do you think these inclusions are, and why do they accumulate?

So far, we have discussed numerous membrane-enclosed organelles. Now we'll turn to a group of cytoplasmic structures that do not directly involve membranes.

CONCEPT 4.4 The Cytoskeleton Provides Strength and Movement



The interior of the cell has a meshwork of protein filaments. Each type of filament is a polymer, made up of monomers that are proteins (which in turn are polymers of amino acids). This cytoskeleton fills several important roles:

- It supports the cell and maintains its shape.
- It holds cell organelles and other particles in position within the cell.
- It moves organelles and other particles around within the cell.
- It is involved with movements of the cytoplasm called cytoplasmic streaming.
- It interacts with extracellular structures, helping anchor the cell in place.

There are three components of the eukaryotic cytoskeleton: microfilaments (smallest diameter), intermediate filaments, and microtubules (largest diameter). These filaments have very different functions.

Microfilaments are made of actin

Microfilaments (FIGURE 4.10A) are usually in bundles. Each filament is about 7 nm in diameter and up to several micrometers long. Microfilaments have two major roles:

- They help the entire cell or parts of the cell to move.
- They determine and stabilize cell shape.

Microfilaments are assembled from **actin** monomers that attach to the filament at one end (the "plus end") and detach at the other (the "minus end"). In an intact filament, assembly and detachment are in equilibrium. But sometimes the filaments can shorten (more detachment) or lengthen (more assembly):

$$\text{Actin polymer (filament)} \rightleftharpoons \text{actin monomers}$$

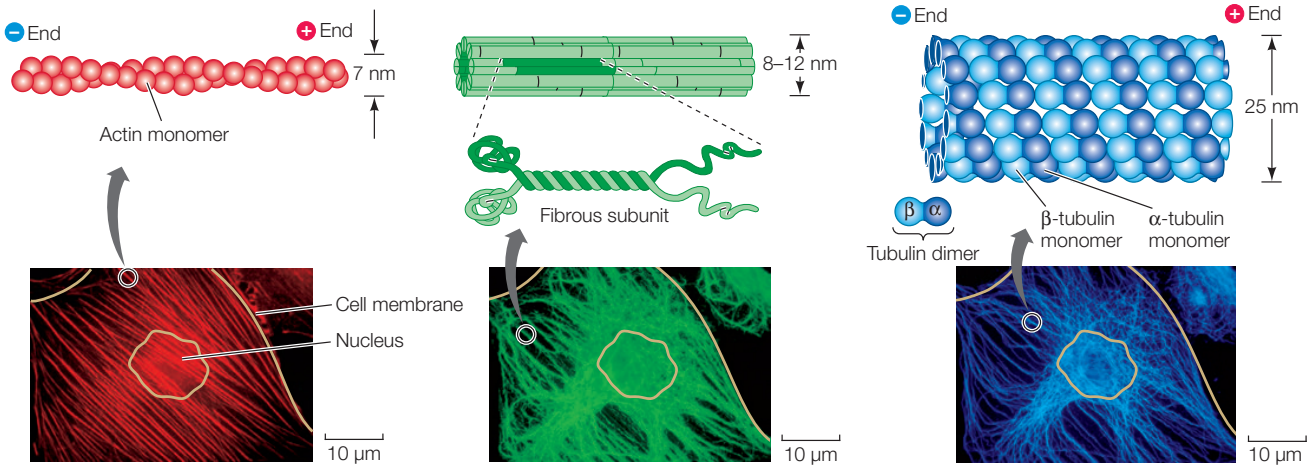
This property of dynamic instability is a hallmark of the cytoskeleton. Portions of it can be made and broken down rather quickly, depending on cell function. Actin-associated proteins work at both ends of the filament to catalyze assembly and disassembly.

In the muscle cells of animals, actin filaments are associated with the motor protein myosin, and the interactions of these two proteins account for the contraction of muscles. A **motor protein** (or molecular motor) is any protein that causes movement within a cell. In non-muscle cells, actin filaments are associated with localized changes in cell shape. For example, microfilaments are involved in the flowing movement of the cytoplasm called cytoplasmic streaming, and in the "pinching" contractions that divide an animal cell into two daughter cells. Microfilaments are also involved in the formation of cellular extensions called pseudopodia (*pseudo*, "false"; *podia*, "feet") that enable some cells (such as *Amoeba*, see Figure 4.14) to move.

Intermediate filaments are diverse and stable

There are at least 50 different kinds of **intermediate filaments** (FIGURE 4.10B), many of them specific to just a few cell types. They generally fall into six molecular classes based on amino acid sequence. One of these classes consists of the fibrous keratin proteins, which are also found in hair and fingernails. The intermediate filaments are tough, ropelike protein assemblages 8–12 nm in diameter. Intermediate filaments are more permanent than the other two types of filaments and do not show dynamic instability. Intermediate filaments have two major structural functions:

- They anchor cell structures in place. In some cells, intermediate filaments radiate from the nuclear envelope and help



(A) Microfilaments
Made up of strands of the protein actin; often interact with strands of other proteins.

(B) Intermediate filaments
Made up of fibrous proteins organized into tough, ropelike assemblies that stabilize a cell's structure and help maintain its shape.

(C) Microtubules
Long, hollow cylinders made up of many molecules of the protein tubulin. Tubulin consists of two subunits, α-tubulin and β-tubulin.

FIGURE 4.10 The Cytoskeleton Three highly visible and important structural components of the cytoskeleton are shown here in detail. Specific stains were used to visualize them in a single cell. These structures maintain and reinforce cell shape and contribute to cell movement.

maintain the positions of the nucleus and other organelles in the cell.

- They resist tension. For example, they maintain rigidity in body surface tissues by extending through the cytoplasm and connecting specialized membrane structures called desmosomes (see Figure 4.18).

Microtubules are the thickest elements of the cytoskeleton

Microtubules (FIGURE 4.10C) are long, hollow, unbranched cylinders about 25 nm in diameter and up to several micrometers long. Microtubules have two roles:

- They form a rigid internal skeleton for some cells or cell regions.
- They act as a framework along which motor proteins can move structures within the cell.

Microtubules are assembled from dimers of the protein **tubulin**. The dimers consist of one molecule each of α-tubulin and β-tubulin. Thirteen chains of tubulin dimers surround the hollow microtubule. Like microfilaments, microtubules show dynamic instability, with plus and minus ends and associated proteins.

$$\text{Microtubule} \rightleftharpoons \text{tubulin dimers}$$

Tubulin polymerization results in a rigid structure, and tubulin depolymerization leads to its collapse. Microtubules often form an interior skeleton for projections that come out of the cell membrane, such as cilia and flagella (see below).

LINK

Microtubules are important components of the spindle, which separates chromosomes during cell division; see Concept 7.2

Cilia and flagella provide mobility

Microtubules internally line movable cell appendages: the **cilia** (singular *cilium*; FIGURE 4.11) and the flagella (see Concept 4.2). Many eukaryotic cells have one or both of these appendages, which are projections of the cell membrane lined with microtubules and their associated proteins:

- Cilia are only 0.25 μm in length. They are present by the hundreds and move stiffly either to propel a cell (for example, in protists) or to move fluid over a stationary cell (as in the human respiratory system).
- Flagella are much longer—100 to 200 μm—and occur singly or in pairs. They can push or pull the cell through its aqueous environment.

The microtubules that line cilia and flagella do more than just make them rigid. Microtubules and their associated proteins are responsible for the movement of these organelles by bending.

In cross section, a typical cilium or eukaryotic flagellum is surrounded by the cell membrane and contains a "9 + 2" array of microtubules. As Figure 4.11B shows, nine fused pairs of microtubules—called doublets—form an outer cylinder, and one pair of unfused microtubules runs up the center. Each doublet is connected to the center of the structure by a radial spoke. This structure is essential to the bending motions of both cilia and flagella. How does this bending occur?

The motion of cilia and flagella results from the sliding of the microtubule doublets past one another. This sliding is driven

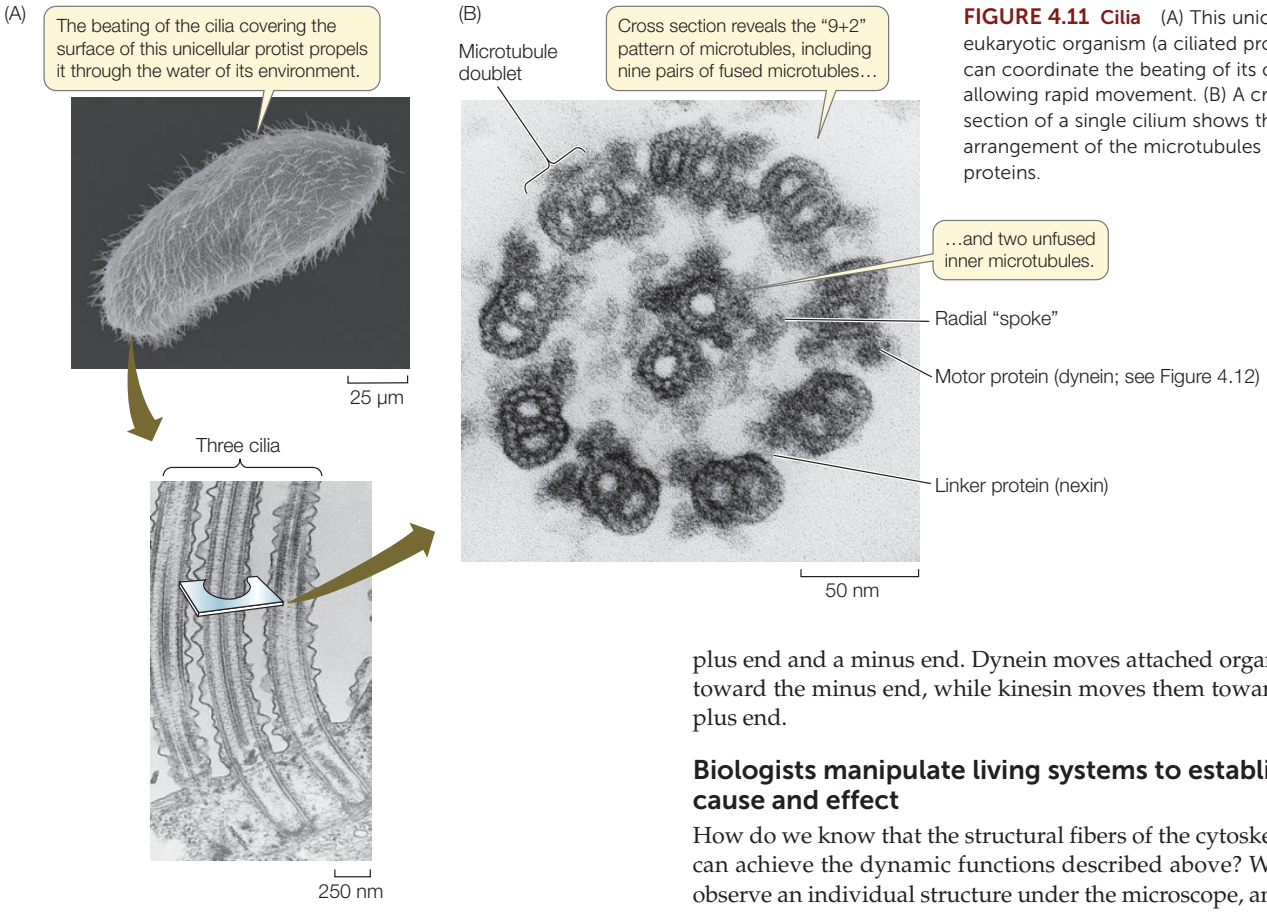


FIGURE 4.11 Cilia (A) This unicellular eukaryotic organism (a ciliated protist) can coordinate the beating of its cilia, allowing rapid movement. (B) A cross section of a single cilium shows the arrangement of the microtubules and proteins.

by a motor protein called dynein, which can change its three-dimensional shape. (All motor proteins work by undergoing reversible shape changes powered by energy from ATP hydrolysis.) Dynein molecules bind between two neighboring microtubule doublets. As the dynein molecules change shape, they move the doublets past one another (**FIGURE 4.12**). Another protein, nexin, can cross-link the doublets and prevent them from sliding past one another; in this case, the cilium bends.

Other motor proteins, including kinesin, carry protein-laden vesicles or other organelles from one part of the cell to another (**FIGURE 4.13**). These proteins bind to the organelle and "walk" it along a microtubule by a repeated series of shape changes. A slightly different form of dynein from the one that moves cilia also performs this function. Recall that microtubules are directional, with a plus end and a minus end. Dynein moves attached organelles toward the minus end, while kinesin moves them toward the plus end.

Biologists manipulate living systems to establish cause and effect

How do we know that the structural fibers of the cytoskeleton can achieve the dynamic functions described above? We can observe an individual structure under the microscope, and we can note that living cells containing that structure can perform a particular function. Such simultaneous observations suggest that the structure may carry out that function, but *mere correlation does not establish cause and effect*. For example, light microscopy of living cells reveals the movement of the cytoplasm within the cell. The observed presence of cytoskeletal components *suggests*, *but does not prove*, their role in this process. Scientists seek to understand the specific relationship between a

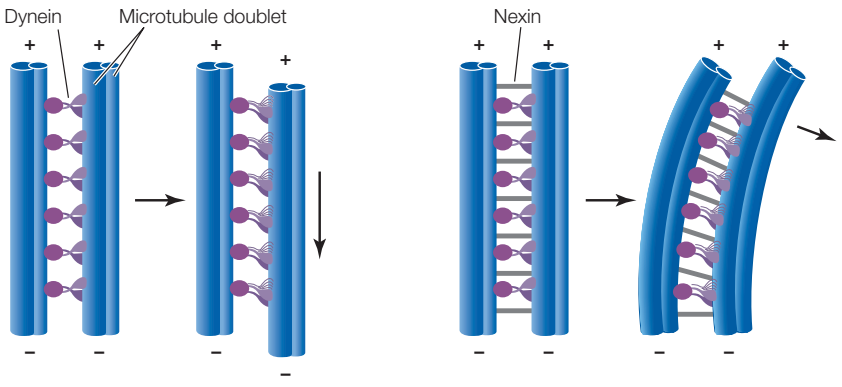


FIGURE 4.12 A Motor Protein Moves Microtubules in Cilia and Flagella The motor protein dynein causes microtubule doublets to slide past one another. If the protein nexin is present to anchor the microtubule doublets together, the flagellum or cilium bends.

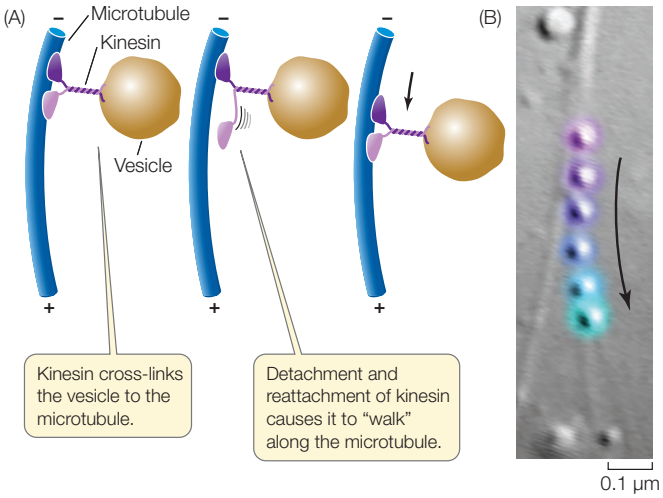


FIGURE 4.13 A Motor Protein Drives Vesicles along Microtubules (A) Kinesin delivers vesicles or organelles to various locations in the cell by moving along microtubule "railroad tracks." (B) The process is seen by time-lapse photography at half-second intervals in the protist *Dictyostelium*.

structure or molecule ("A") and a function ("B"). Two manipulative approaches are commonly used in cell biology:

- *Inhibition*: Use a drug that inhibits A, and see if B still occurs. If B does not occur, then A is probably a causative factor for B. **FIGURE 4.14** shows an experiment in which an inhibitor is used to demonstrate cause and effect in the case of microfilaments and cell movement.
- *Mutation*: Examine a cell or organism that lacks the gene (or genes) for A, and see if B still occurs. If it does not, then A is probably a causative factor for B. You will see many examples of this experimental approach later in this book.

CHECKpoint CONCEPT 4.4

- ✓ Make a table that compares the three major components of the cytoskeleton with regard to composition, structure, and function.
- ✓ The neuron (nerve cell) has a long extension called an axon. Molecules made in the cell's main body must travel a long distance to reach the end of the axon. The axon is lined with microtubules. Explain how motor proteins, vesicles, and microtubules move these molecules along the axon.
- ✓ In a dividing cell, the chromosomes become very compact, and then the duplicated sets of chromosomes move along microtubules to opposite ends of the cell. How would you use an inhibitor to show that microtubules are essential for this chromosomal separation? What control treatments would you suggest?

All cells interact with their environments, and many eukaryotic cells are part of multicellular organisms and must interact with

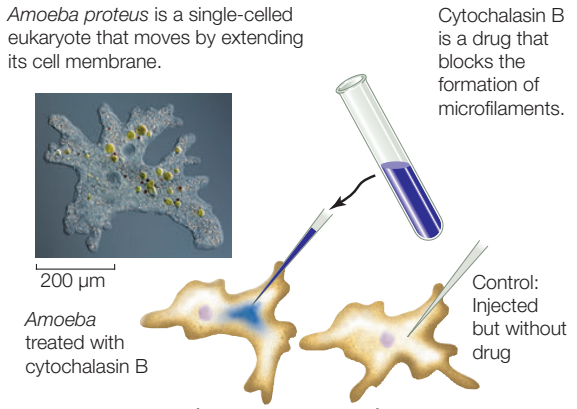
INVESTIGATION

FIGURE 4.14 The Role of Microfilaments in Cell Movement: Showing Cause and Effect in Biology In test tubes, the drug cytochalasin B prevents microfilament formation from monomeric precursors. This led to the question: Will the drug work like this in living cells and inhibit the movement of *Amoeba*[a]?

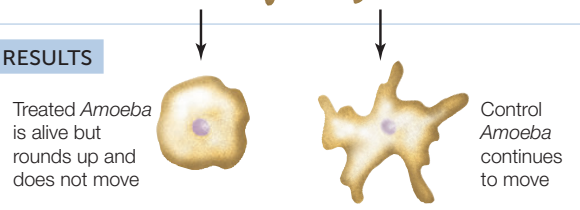
HYPOTHESIS

Amoeboid cell movements are caused by microfilaments.

METHOD



RESULTS



CONCLUSION

Microfilaments are essential for amoeboid cell movement.

ANALYZE THE DATA

Several important controls were done to validate the conclusions of this experiment. The experiment was repeated in the presence of cycloheximide, which inhibits new protein synthesis, and colchicine, which inhibits the polymerization of microtubules. Here are the results:

Condition	Rounded cells (%)
No drug	3
Cytochalasin B	95
Colchicine	4
Cycloheximide	3
Cycloheximide + cytochalasin B	94

A. Explain the reasoning behind each condition. Why were the controls important?

B. Interpret the results of this experiment. What can you conclude about movements in *Amoeba* and the cytoskeleton?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]T. D. Pollard and R. R Weihing. 1974. *CRC Critical Reviews of Biochemistry* 2: 1–65.

other cells. The cell membrane (which we will discuss in detail in Chapter 5) plays a crucial role in these interactions, but other structures outside that membrane are involved as well. We will now turn to these extracellular structures in animals and plants.

CONCEPT 4.5 Extracellular Structures Provide Support and Protection For Cells and Tissues

In Chapter 5 we will look at the role of the cell membrane in cell communication. Although the cell membrane is the functional barrier between the inside and the outside of a cell, cells produce molecules and secrete them to the outside of the cell membrane. There these molecules form structures that play essential roles in protecting, supporting, or attaching cells to each other. Because they are outside the cell membrane, these structures are said to be "extracellular." In eukaryotes, these structures are made up of two components:

- A fibrous macromolecule
- A gel-like medium in which the fibers are embedded

The plant cell wall is an extracellular structure

The plant cell wall is a semirigid structure outside the cell membrane (FIGURE 4.15). The fibrous component is the polysaccharide cellulose (see Figure 2.10), and the gel-like matrix contains extensively cross-linked polysaccharides and proteins. The wall has three major roles:

- It provides support for the cell and limits the volume of a mature cell by remaining rigid.
- It acts as a barrier to infection by fungi and other organisms that can cause plant diseases.
- It contributes to plant form by controlling the direction of cell expansion during growth and development.

Because of their thick cell walls, plant cells viewed under a light microscope appear to be entirely isolated from one another. But electron microscopy reveals that this is not the case. The cytoplasms of adjacent plant cells are connected by numerous cell membrane–lined channels called **plasmodesmata**. These are about 20–40 nm in diameter and extend through the cell walls (see Figure 4.7). Plasmodesmata allow water, ions, small molecules, hormones, and even some RNA and protein molecules to move between connected cells. In this way, energy-rich molecules such as sugars can be shared among cells, and plant hormones can affect growth at sites far from where they were synthesized. This intercellular communication integrates a plant organ composed of thousands of cells.

The extracellular matrix supports tissue functions in animals

Animal cells lack the semirigid wall that is characteristic of plant cells, but many animal cells are surrounded by, or in contact with, an **extracellular matrix** (FIGURE 4.16). The fibrous component of the extracellular matrix is the protein **collagen**,

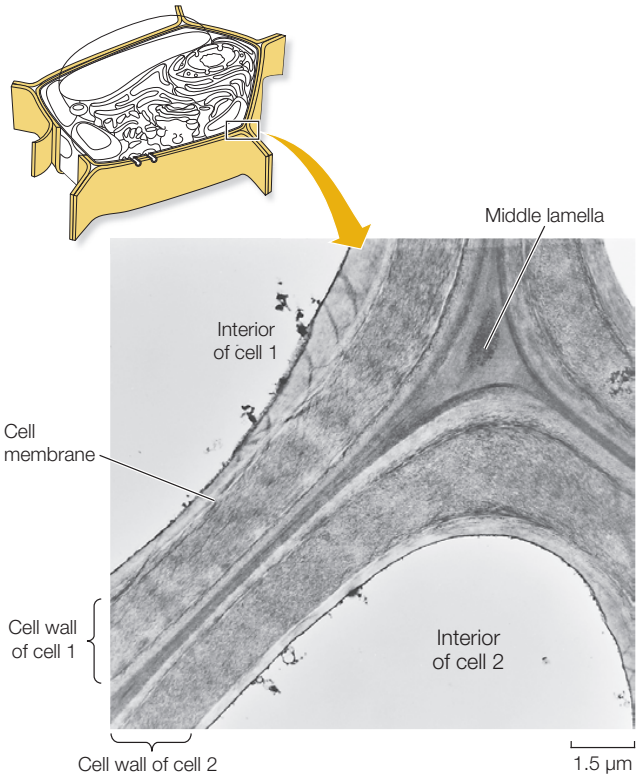


FIGURE 4.15 The Plant Cell Wall The semirigid cell wall provides support for plant cells. It is composed of cellulose fibers embedded in a matrix of polysaccharides and proteins.

and the gel-like medium consists of **proteoglycans**, which are glycoproteins with long carbohydrate side chains. A third group of proteins links the collagen and the proteoglycan matrix together.

The extracellular matrices of animal cells have several roles:

- They hold cells together in tissues.
- They contribute to the physical properties of cartilage, skin, and other tissues. For example, the mineral component of bone is laid down on an organized extracellular matrix.
- They help filter materials passing between different tissues. This is especially important in the kidney.
- They help orient cell movements during embryonic development and during tissue repair.

Proteins connect the cell's cell membrane to the extracellular matrix. These proteins (for example, integrin) span the cell membrane and have two binding sites: one on the interior of the cell, usually to microfilaments in the cytoplasm just below the cell surface, and the other to collagen in the extracellular matrix. These binding sites are noncovalent and reversible. When a cell moves its location in an organism, the first step is

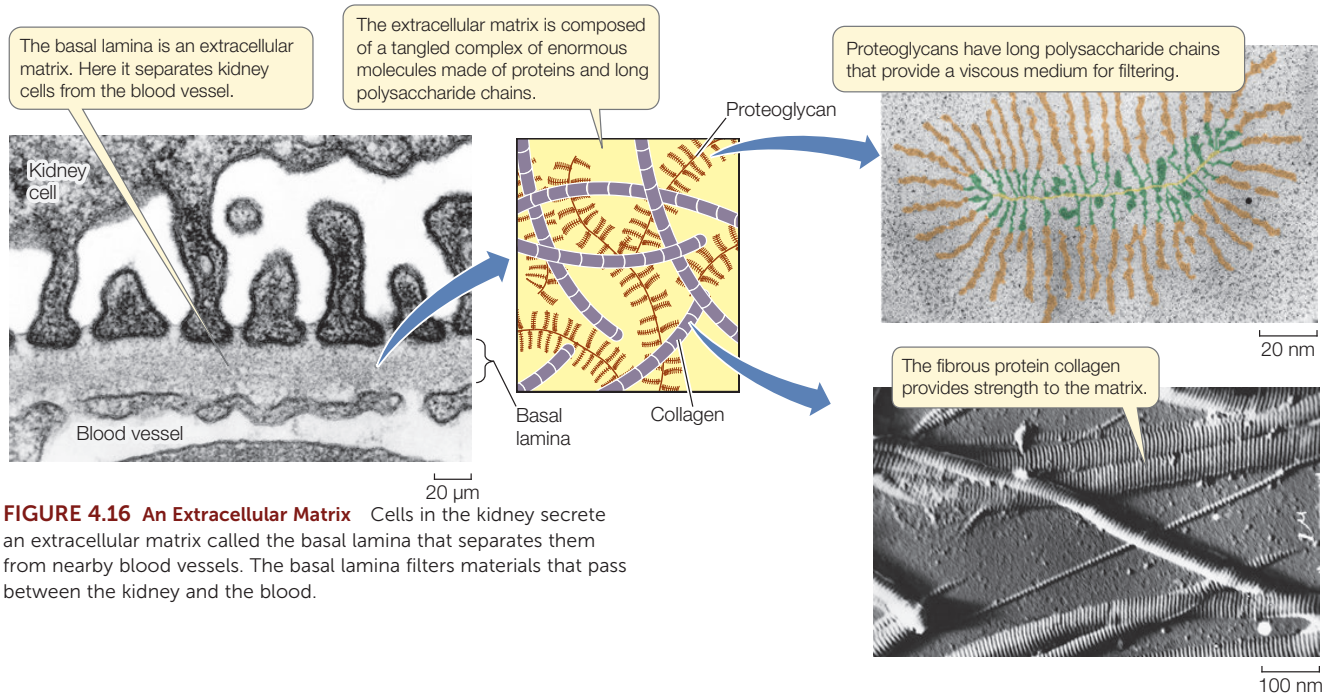


FIGURE 4.16 An Extracellular Matrix Cells in the kidney secrete an extracellular matrix called the basal lamina that separates them from nearby blood vessels. The basal lamina filters materials that pass between the kidney and the blood.

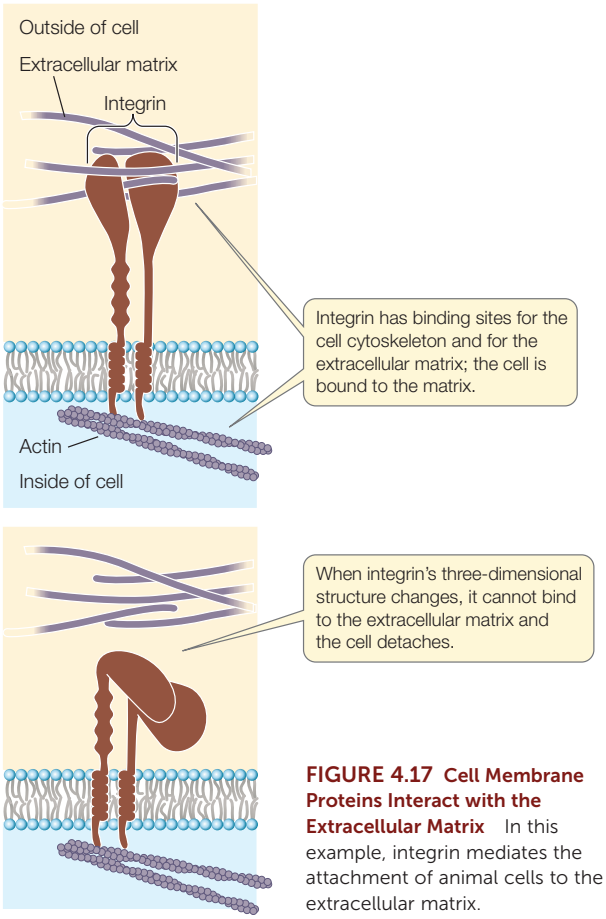


FIGURE 4.17 Cell Membrane Proteins Interact with the Extracellular Matrix In this example, integrin mediates the attachment of animal cells to the extracellular matrix.

for integrin to change its three-dimensional structure so that it detaches from the collagen (FIGURE 4.17).

Cell junctions connect adjacent cells

In a multicellular animal, specialized structures protrude from adjacent cells to "glue" them together. These **cell junctions** are most evident in electron micrographs of epithelial tissues, which are layers of cells that line body cavities or cover body surfaces (examples are skin and the lining of the windpipe leading to the lungs). These surfaces are often exposed to environmental factors that might disrupt the integrity of the tissues, so it is particularly important that their cells stick together tightly. There are three types of junctions (FIGURE 4.18):

- *Tight junctions* prevent substances from moving through spaces between cells. For example, the epithelium of the urinary bladder contains tight junctions to prevent urine from leaking out into the body.
- *Desmosomes* hold adjacent cells together with stable protein connections, but materials can still move around in the extracellular matrix. This provides mechanical stability for tissues such as skin that receive physical stress.
- *Gap junctions* are like plant plasmodesmata: they are channels that run between membrane pores in adjacent cells, allowing substances to pass between cells. In the heart, for example, gap junctions allow the rapid spread of electric current mediated by ions so the heart muscle cells can beat in unison.

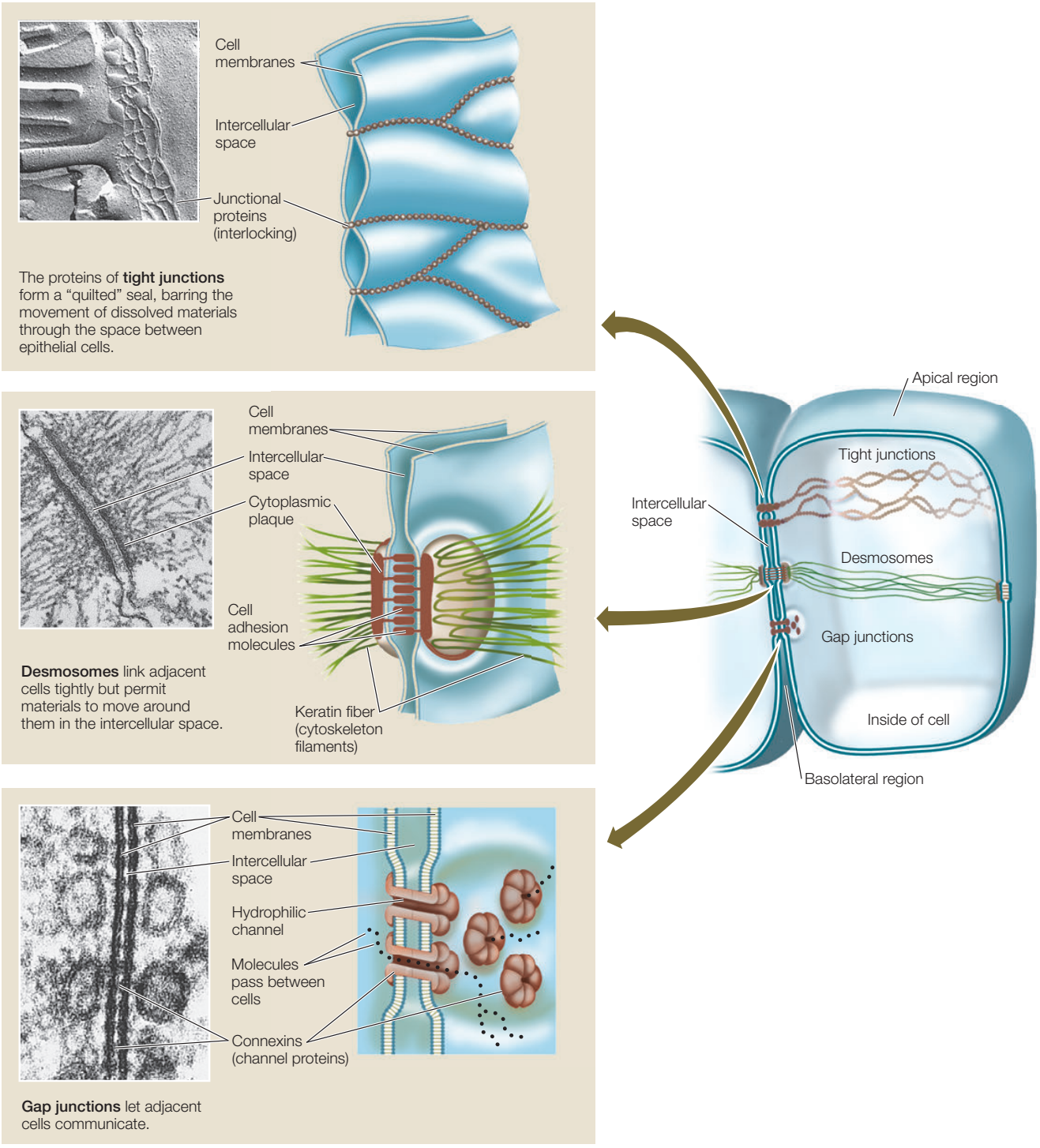


FIGURE 4.18 Junctions Link Animal Cells Although all three types of junctions are shown in the cell at right, they don't necessarily all occur in the same cell.

Go to ACTIVITY 4.3 Animal Cell Junctions
PoL2e.com/ac4.3

CHECKpoint CONCEPT 4.5

- ✓ Compare the fibrous and gel-like components of the extracellular matrices of plant and animal cells.
- ✓ What kinds of cell junctions would you expect to find, and why, in the following situations?
  - a. In the digestive system, where material must pass through cells and not go through the extracellular material, to get from the intestine to the blood vessels.
  - b. In a small animal, where a chemical signal passes rapidly through cells to go from the head to the tail.
  - c. In the lining of the intestine, where cells in the lining are constantly jostled by the churning of the underlying muscle.
- ✓ When cancer spreads from its primary location to other parts of the body (a process called metastasis), tumor cells detach from their original location and then reattach at a different location. How would the integrin–collagen system be involved in this process?

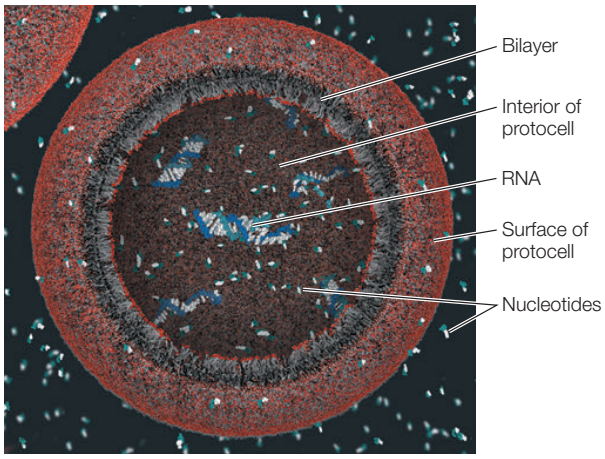


FIGURE 4.19 A Protocell A protocell can be made in the lab and can carry out some functions of modern cells—in particular, it provides a compartment for biochemical reactions.

Q What do the characteristics of modern cells indicate about how the first cells originated?

ANSWER Ideas about how the first cells may have formed focus on two questions: how and when. As to how cells could arise from a chemical-rich environment, most biologists assume that a cell membrane formed first and was necessary to provide a compartment for the chemical transformations of life to occur, separated from the environment (Concept 4.1). Biologists also assume that the first cells were relatively simple prokaryotes (Concept 4.2), without the organelles that define eukaryotic cells (Concept 4.3).

Jack Szostak, a Nobel laureate at Harvard University, builds synthetic cell models that give insights into the origin of cells. He and his colleagues make small membrane-lined droplets by putting fatty acids into water and then shaking the mixture. The lipids form water-filled droplets, each surrounded by a lipid bilayer "membrane" (see Figure 2.13). With water (and other molecules of the scientists' choosing) trapped inside, these spheres have many properties characteristic of modern cells—so many that they have been called **protocells** (**FIGURE 4.19**). For example, the membrane barrier determines what goes in and out of a protocell, by excluding macromolecules like RNA but allowing smaller molecules such as nucleotides to pass through. Moreover, RNA inside the protocell can act as a catalyst, replicating itself from nucleotides that enter the protocell. The spheres are somewhat unstable, and under the microscope they can be seen to grow, elongate, and break, a possible precursor of more precise cell division.

Is this really a cell, possibly like the one where life started? Certainly not: it cannot fully reproduce itself, and its capacities for metabolism are limited. But by providing a compartment for biochemical reactions with a boundary that separates it from the environment, the protocell is a model for the first cell.

When did the first cells on Earth appear? According to geologists, Earth is about 4.5 billion years old. Heat and atmospheric conditions precluded life for at least a half-billion years after Earth formed. The oldest fossils of multicellular organisms date from about 1.2 billion years ago.

In all probability, life began with single-celled organisms resembling modern bacteria. Unfortunately, such cells lack the structures that are typically preserved in fossils, and so they die without a trace. Recently, however, geochemist and paleontologist William Schopf at the University of California, Los Angeles used a new method of microscopy called confocal laser scanning microscopy, combined with chemical analyses, to identify fossil cells that are about 800 million years old. Some of these look like Szostak's protocells. These were probably not the first cells, as there is chemical evidence in some rocks that life was present about 3.8 billion years ago. But so far, Schopf's fossilized cells are the oldest cells that anyone has been able to find.

SUMMARY

CONCEPT 4.1 Cells Provide Compartments for Biochemical Reactions

See ACTIVITY 4.1

- Cell theory states that the cell is the fundamental unit of biological structure and function.
- Cells are small because a cell's surface area must be large compared with its volume to accommodate exchanges between the cell and its environment. **Review Figure 4.2**
- All cells are enclosed by a selectively permeable **cell membrane** that separates their contents from the external environment.

CONCEPT 4.2 Prokaryotic Cells Do Not Have a Nucleus

- Prokaryotic cells usually have no internal compartments, but have a **nucleoid** containing DNA, and a **cytoplasm** containing **cytosol**, **ribosomes** (the sites of protein synthesis), proteins, and small molecules. Many have an extracellular **cell wall**. **Review Figure 4.5**
- Some prokaryotes have folded membranes, for example photosynthetic membranes, and some have **flagella** for motility. **Review Figure 4.6**

CONCEPT 4.3 Eukaryotic Cells Have a Nucleus and Other Membrane-Bound Compartments

- Eukaryotic cells contain many membrane-enclosed **organelles** that compartmentalize their biochemical functions. **Review Figure 4.7 and ANIMATED TUTORIAL 4.1**
- The **nucleus** contains most of the cell's DNA.
- The **endomembrane system**—consisting of the nuclear envelope, **endoplasmic reticulum**, **Golgi apparatus**, and **lysosomes**—is a series of interrelated compartments enclosed by membranes. It segregates proteins and modifies them. Lysosomes contain many digestive enzymes. **Review Figures 4.8 and 4.9, ANIMATED TUTORIAL 4.2, and ACTIVITY 4.2**
- **Mitochondria** and **chloroplasts** are semiautonomous organelles that process energy.
- A **vacuole** is prominent in many plant cells. It is a membrane-enclosed compartment full of water and dissolved substances.

CONCEPT 4.4 The Cytoskeleton Provides Strength and Movement

- The **microfilaments**, **intermediate filaments**, and **microtubules** of the **cytoskeleton** provide the cell with shape, strength, and movement. **Review Figure 4.10**
- Microfilaments and microtubles have dynamic instability and can grow or shrink in length rapidly.
- **Cilia** and **flagella** are microtubule-lined extensions of the cell membrane that produce movements of cells or their surrounding fluid medium. **Review Figures 4.11 and 4.12**
- Motor proteins move cellular components, such as **vesicles**, around the cell by "walking" them along the microtubules. **Review Figure 4.13**
- Biologists establish cause-and-effect relationships by manipulating biological systems. **Review Figure 4.14**

CONCEPT 4.5 Extracellular Structures Provide Support and Protection for Cells and Tissues

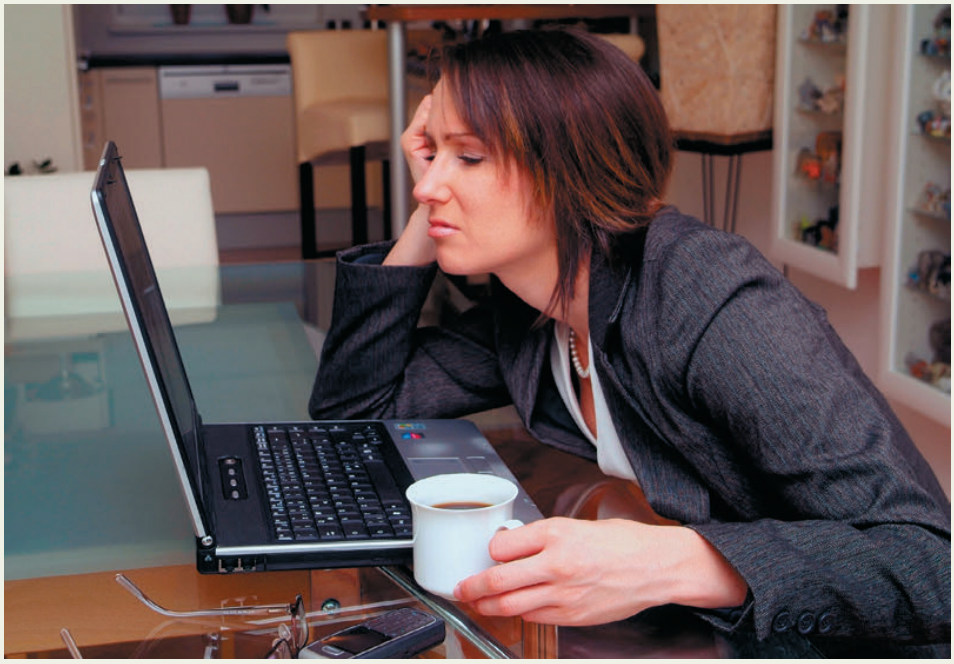
- The plant cell wall consists principally of cellulose. Cell walls are pierced by **plasmodesmata** that join the cytoplasms of adjacent cells. **Review Figure 4.15**
- In animals, the **extracellular matrix** consists of different kinds of proteins, including **collagen** and **proteoglycans**. Integrins connect the cell cytoplasm with the extracellular matrix. **Review Figures 4.16 and 4.17**
- Specialized **cell junctions** connect cells in animal tissues. These include tight junctions, desmosomes, and gap junctions. Gap junctions are involved in intercellular communication. **Review Figure 4.18 and ACTIVITY 4.3**

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is4

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

5 Cell Membranes and Signaling

KEY CONCEPTS



Many people rely on caffeine to wake themselves up and to keep their minds alert.

If you are like most people, you consume a significant amount of caffeine every day. In fact, more than 90 percent of North Americans and Europeans drink coffee or tea to get their "caffeine fix." Coffee and tea plants contain caffeine as a defense against the insects that eat them. Caffeine acts as an insecticide in plant parts that are particularly vulnerable to insect attacks, such as seeds, young seedlings, and leaves. But it is not toxic to humans.

Legend has it that about 5,000 years ago, a Chinese emperor found out by accident that a pleasant beverage could be made by boiling tea leaves. About 1,000 years ago, monks living in what is now Ethiopia found that roasting coffee seeds (also called "beans") gave a similarly pleasant effect and that the beverage kept them awake during long periods of prayer. Caffeine is now the most widely consumed psychoactive molecule in the world, but unlike other psychoactive drugs, it is not subject to government regulation.

Most people know from personal experience what caffeine does to the body: because it keeps us awake, it obviously affects the brain. In fact, it is often given to premature babies in the hospital nursery when they stop breathing. But it also affects other parts of the body—for example, it increases urination and speeds up the heart. How does this molecule work?

The key to understanding caffeine's action is to understand how it interacts with the cell membrane. In Chapter 4 we introduced the concept of the membrane as a structural boundary between the inside of a cell and the surrounding environment. The cell membrane physically separates the cell cytoplasm from its surroundings and helps maintain chemical differences between these two environments. The same can be said of the membranes that surround cell organelles, separating them from the cytoplasmic environment.

When caffeine arrives at a cell in the body, it first encounters the cell membrane. The properties of this membrane determine whether and how the cell will react to caffeine. Will it cross the membrane boundary and enter the cell? What determines whether it crosses the membrane? If it does not, how can caffeine's interactions with membrane components lead to changes in cell function?

Q What role does the cell membrane play in the body's response to caffeine?

You will find the answer to this question on page 102.

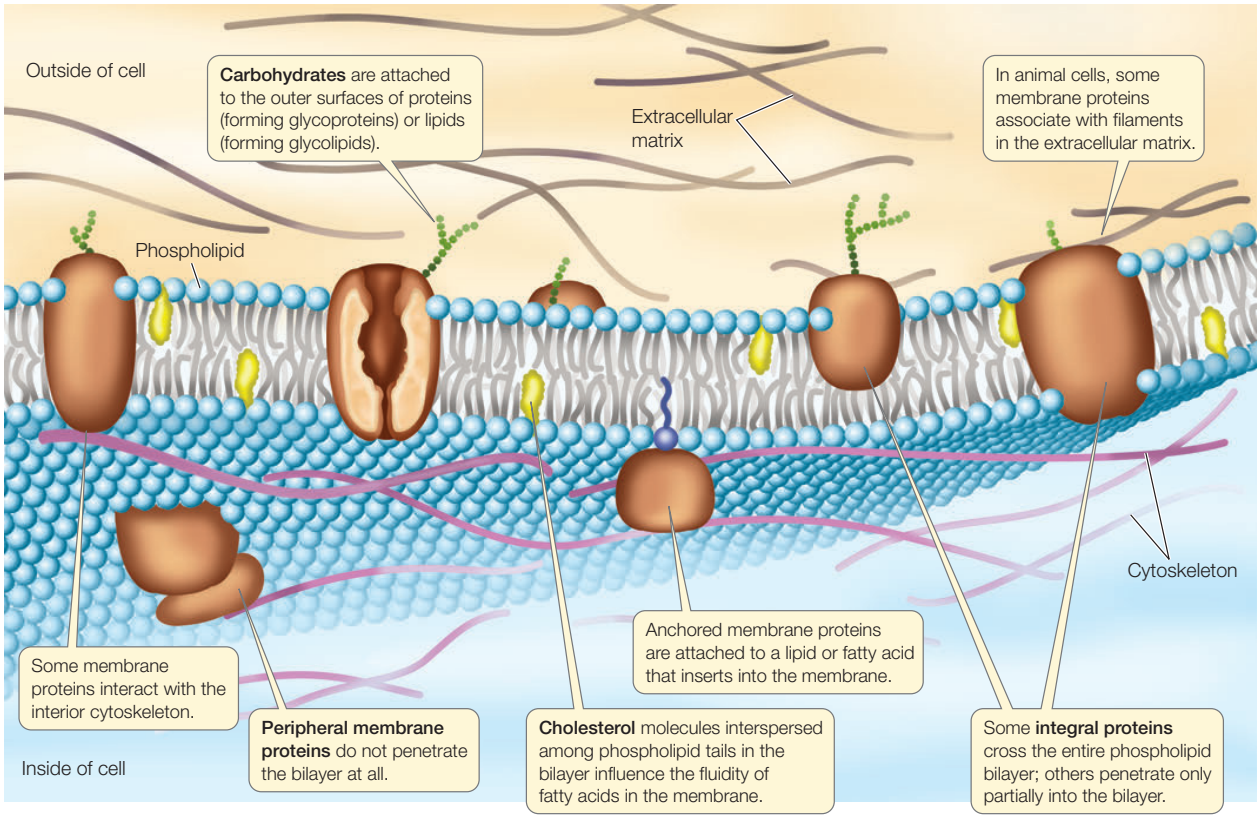


FIGURE 5.1 Membrane Structure The general molecular structure of biological membranes is a continuous phospholipid bilayer in which proteins are embedded. The phospholipid bilayer separates two aqueous regions, the external environment outside the cell and the cell cytoplasm.

Go to ACTIVITY 5.1 Membrane Molecular Structure
PoL2e.com/ac5.1

CONCEPT 5.1 Biological Membranes Have a Common Structure and Are Fluid

The evolution of cellular life required the presence of a boundary, a way to separate the inside of the cell from the surrounding environment. This need was—and still is—fulfilled by biological membranes. Like many of the basic processes of life at the cellular level, the functions of membranes are carried out by a molecular structure shared by most organisms.

A biological membrane's structure and functions are determined by the chemical properties of its constituents: lipids, proteins, and carbohydrates. An important concept that emerges from our consideration of these molecules is polarity and how polarity influences the way a molecule interacts with water. Recall from Concept 2.2 that some compounds are polar and hydrophilic ("water-loving"), whereas others are nonpolar and hydrophobic ("water-hating"), and that a phospholipid has both polar and nonpolar regions. The nonpolar regions of phospholipids and membrane proteins interact to form an insoluble barrier. The phospholipid bilayer serves as a lipid "lake" in which a variety of proteins "float" (**FIGURE 5.1**). This general design is known as the **fluid mosaic model**.

Membranes contain a wide array of proteins, most of which are noncovalently embedded in the phospholipid bilayer. These proteins are held within the membrane by their hydrophobic regions (also called their hydrophobic "domains"). The proteins' hydrophilic regions are exposed to the watery conditions on one or both sides of the bilayer. Membrane proteins have three major functions: some move materials through the membrane, others are involved in intercellular recognition and adhesion, while others receive chemical signals from the cell's external environment.

The carbohydrates associated with membranes are attached to either lipids or protein molecules. They are generally located on the outside of the cell, where they interact with substances in the external environment. Like some membrane proteins, carbohydrates are crucial for recognizing specific molecules, such as those on the surfaces of adjacent cells.

Each membrane has constituents that are suitable for the specialized functions of the cell or organelle it surrounds. As you read about the different molecules in membranes, keep in mind that some membranes contain many protein molecules, others are lipid-rich, others have significant amounts of cholesterol, and still others are rich in carbohydrates.

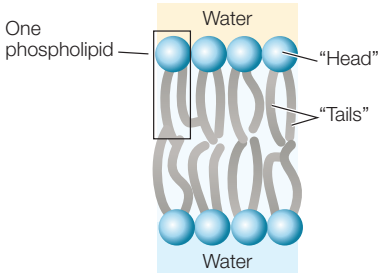
Lipids form the hydrophobic core of the membrane

The lipids in biological membranes are usually **phospholipids**, with hydrophilic and hydrophobic regions:

- *Hydrophilic regions.* The phosphorus-containing "head" of a phospholipid is electrically charged and therefore associates with polar water molecules.

- *Hydrophobic regions.* The long, nonpolar fatty acid "tails" of a phospholipid associate with other nonpolar materials, but they do not dissolve in water or associate with hydrophilic substances.

One way in which phospholipids can coexist with water is to form a bilayer, with the fatty acid "tails" of the two layers interacting with each other, and the polar "heads" facing the outside, aqueous environment:



LINK

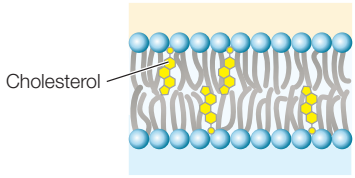
The properties of phospholipid bilayers are described in **Concept 2.4**

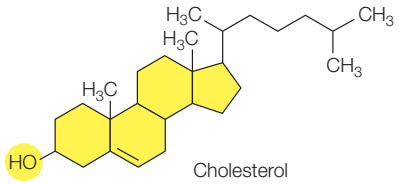
The thickness of a biological membrane is about 8 nm (0.008 μm), which is twice the length of a typical phospholipid. The page you are reading is about 3,000 times thicker than this.

As we noted in Chapter 4, it is possible in the laboratory to make artificial bilayers with the same organization as natural membranes. Small holes in such bilayers seal themselves spontaneously. The capacity of phospholipids to associate with one another and maintain a bilayer organization helps biological membranes fuse during vesicle formation, phagocytosis, and related processes (see Concept 4.3, especially Figure 4.9).

Although biological membranes all share a similar structure, there are many different kinds of phospholipids, and membranes from different cells or organelles may differ greatly in their lipid composition. Phospholipids can differ in terms of fatty acid chain length (number of carbon atoms), degree of unsaturation (number of double bonds) in the fatty acids, and the kinds of polar (phosphate-containing) groups present. The most common fatty acids in membranes have chains with 16–18 carbon atoms and 0–2 double bonds. Saturated fatty acid chains (those with no double bonds) allow close packing of phospholipids in the bilayer, whereas the "kinks" in the unsaturated fatty acids make for a less dense, more fluid packing (see Figure 2.12).

Up to 25 percent of the lipid content of an animal cell's cell membrane may be the steroid cholesterol. Steroids are a family of carbon compounds that have multiple linked rings. Cholesterol plays an important role in modulating membrane fluidity (see below); other steroids function as hormones (see Concept 35.2). The hydroxyl group (—OH) on the cholesterol molecule interacts with the polar heads of the phospholipids, while the nonpolar rings insert among the fatty acid chains of the membrane:





The fatty acids of the phospholipids make the membrane somewhat fluid—about as fluid as olive oil. This fluidity permits some molecules to move laterally within the plane of the membrane. A given phospholipid molecule in the cell membrane can travel from one end of the cell to the other in a little more than 1 second!

APPLY THE CONCEPT

Biological membranes have a common structure and are fluid

The membrane lipids of a cell can be labeled with a fluorescent tag so the entire surface of the cell will glow evenly under ultraviolet light. If a strong laser light is then shone on a tiny region of the cell, that region gets bleached (the strong light destroys the fluorescent tag) and there is a "hole" in the cell surface fluorescence (though not an actual hole in the cell's membrane). After the laser is turned off, the hole gradually fills in with fluorescent lipids that diffuse in from other parts of the membrane. The time it takes for the "hole" to disappear is a measure of membrane fluidity. The table shows some data for cells with altered membrane compositions.[a] Explain the effect of each alteration.

CONDITION	TIME (sec) FOR "HOLE" TO BECOME FLUORESCENT
No alteration	65
Decreased length of fatty acid chains	38
Increased cholesterol	88
Increased percentage of unsaturated fatty acids	42
Increased membrane protein content	90

[a] Adapted from E. Wu et al. 1977. *Biochemistry* 16: 3936–3941.

Membrane fluidity is affected by several factors, two of which are particularly important:

- *Lipid composition.* Cholesterol and long-chain, saturated fatty acids pack tightly together, resulting in less fluid membranes. Unsaturated fatty acids or those with shorter chains tend to increase membrane fluidity. Some anesthetics are nonpolar and act by inserting into cell membranes. They reduce the fluidity of nerve cell membranes and thereby decrease nerve activity.
- *Temperature.* Membrane fluidity declines under cold conditions because molecules move more slowly at lower temperatures. For example, when your fingers get numb after contact with ice, it is due to a reduction in membrane fluidity in the nerve cells. To address this problem, some organisms simply change the lipid composition of their membranes when their environment gets cold, replacing saturated with unsaturated fatty acids and using fatty acids with shorter chains. These changes play a role in the survival of plants, bacteria, and hibernating animals during the winter.

While phospholipid molecules can easily move laterally within a membrane, it is rare for a phospholipid in one half of the bilayer to spontaneously flip over to the other side. For that to happen, the polar part of the molecule would have to move through the hydrophobic interior of the membrane. Since spontaneous flip-flops are rare, the inner and outer halves of the bilayer may be quite different in the kinds of phospholipids they contain.

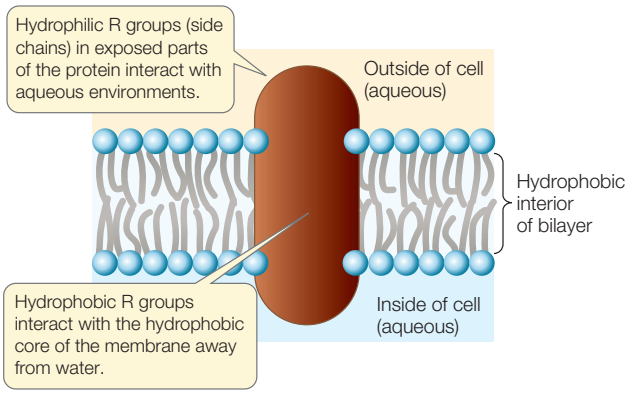
Go to ANIMATED TUTORIAL 5.1
Lipid Bilayer Composition
PoL2e.com/at5.1

Proteins are important components of membranes

All biological membranes contain proteins. Typically, cell membranes have about 1 protein molecule for every 25 phospholipid molecules. This ratio varies depending on membrane function. In the inner membrane of the mitochondrion, which is specialized for energy processing, there is 1 protein for every 5 lipids. By contrast, myelin—a membrane that encloses portions of some neurons (nerve cells) and acts as an electrical insulator—has only 1 protein for every 70 lipids.

Recall from Table 3.2 that some amino acids contain nonpolar, hydrophobic R groups, whereas others contain polar (or charged), hydrophilic R groups. The arrangement of these amino acids in a membrane protein determines whether the membrane protein will insert into the nonpolar lipid bilayer and how it will be positioned. There are two general types of membrane proteins:

- **Peripheral membrane proteins** lack exposed hydrophobic groups and are not embedded in the bilayer. Instead, they have polar or charged regions that interact with exposed parts of integral membrane proteins, or with the polar heads of phospholipid molecules (see Figure 5.1).
- **Integral membrane proteins** are at least partly embedded in the phospholipid bilayer. Like phospholipids, these proteins have both hydrophilic and hydrophobic regions:



Membrane proteins and lipids generally interact only noncovalently. The polar ends of proteins can interact with the polar ends of lipids, and the nonpolar regions of both molecules can interact hydrophobically. However, some membrane proteins have fatty acids or other lipid groups covalently attached to them. These are referred to as anchored membrane proteins, because it is their hydrophobic lipid components that anchor them in the phospholipid bilayer (see Figure 5.1).

Proteins are asymmetrically distributed on the inner and outer surfaces of membranes. An integral membrane protein that extends all the way through the phospholipid bilayer and protrudes on both sides is known as a **transmembrane protein**. In addition to one or more transmembrane domains (regions) that extend through the bilayer, such a protein may have domains with other specific functions on the inner and outer sides of the membrane. Transmembrane proteins are always oriented the same way—domains with specific functions inside or outside the cell are always found on the correct side of the membrane. Peripheral membrane proteins are located on one side of the membrane or the other. This asymmetrical arrangement gives the two surfaces of the membrane different properties. As we will soon see, these differences have great functional significance.

Like lipids, some membrane proteins move relatively freely within the phospholipid bilayer. Cell fusion experiments illustrate this migration dramatically. When two cells fuse, a single continuous membrane forms around both cells, and some proteins from each cell distribute themselves uniformly around this membrane (**FIGURE 5.2**).

Although some proteins are free to migrate throughout the membrane, others appear to be contained within specific regions. These membrane regions are like a corral of horses on a farm: the horses are free to move around within the fenced area but not outside it. For example, a muscle cell protein that recognizes a chemical signal from a neuron is normally found only within a specific region of the cell membrane, where the neuron meets the muscle cell.

How does this happen? Proteins inside the cell can restrict the movement of proteins within a membrane. Components

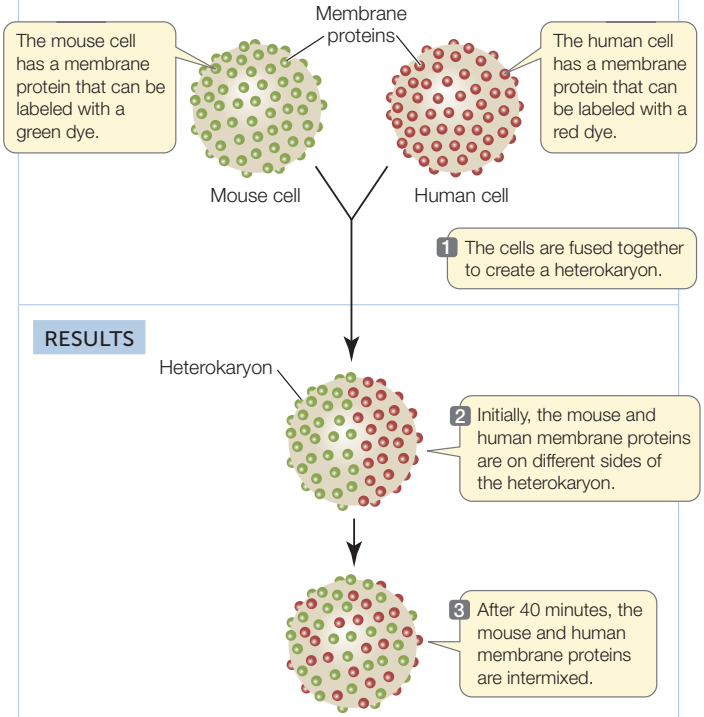
INVESTIGATION

FIGURE 5.2 Rapid Diffusion of Membrane Proteins A human cell can be fused to a mouse cell in the laboratory, forming a single large cell (heterokaryon). This phenomenon was used to test whether membrane proteins can diffuse independently in the plane of the cell membrane.[a]

HYPOTHESIS

Proteins embedded in a membrane can diffuse freely within the membrane.

METHOD



RESULTS

CONCLUSION

Membrane proteins can diffuse rapidly in the plane of the membrane.

ANALYZE THE DATA

The experiment was repeated at various temperatures with the following results:

Temperature (°C)	Cells with mixed proteins (%)
20	0
15	8
20	42
26	77

Plot these data on a graph of Percentage Mixed vs. Temperature. Explain these data, relating the results to the concepts of diffusion and membrane fluidity.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]L. Frye and M. Edidin. 1970. *Journal of Cell Science* 7: 319–335.

of the cytoskeleton may be attached to membrane proteins protruding into the cytoplasm (see Figure 5.1). The stability of the cytoskeleton may thus restrict the movement of attached membrane proteins.

Cell membrane carbohydrates are recognition sites

In addition to lipids and proteins, the cell membrane contains carbohydrates. The carbohydrates are located on the outer surface of the cell membrane and may be covalently bonded to lipids or to proteins:

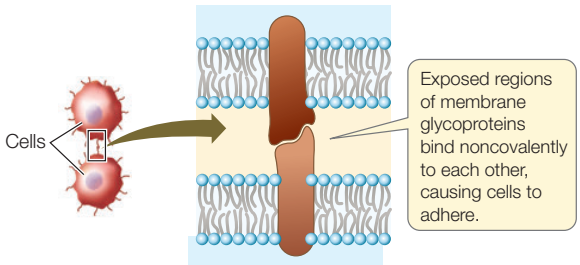
- A **glycolipid** consists of a carbohydrate covalently bonded to a lipid.
- A **glycoprotein** consists of one or more short carbohydrate chains covalently bonded to a protein. The bound carbohydrates are oligosaccharides, usually not exceeding 15 monosaccharide units in length (see Concept 2.3).
- A **proteoglycan** is a protein with even more carbohydrate molecules attached to it, and the carbohydrate chains are often longer than in glycoproteins.

These chains of monosaccharides can generate a large diversity of linear or branched structures. These diverse structures play roles in cell communication and cell adhesion. For example, the carbohydrates on some glycolipids change when cells become cancerous. This change may allow white blood cells to target cancer cells for destruction.

LINK

The activities of white blood cells are described in **Chapter 39**

Cells can stick together (adhere) due to interactions between similar carbohydrates on the outer surfaces of two cells, or between a carbohydrate on one cell and a membrane protein on another cell. Or, two proteins can interact directly:



Cell adhesion occurs in all kinds of multicellular organisms.

Membranes are constantly changing

Membranes in eukaryotic cells are constantly forming, transforming from one type to another, fusing with one another, and breaking down. As we discussed in Concept 4.3, fragments of membrane move (in the form of vesicles) from the endoplasmic reticulum to the Golgi apparatus, and from the Golgi apparatus to the cell membrane (see Figure 4.8). Secondary lysosomes form when primary lysosomes from the Golgi apparatus fuse with phagosomes from the cell membrane (see Figure 4.9).

Because membranes interconvert so readily, we might expect all subcellular membranes to be chemically identical. However, that is not the case: there are major chemical differences among the membranes of even a single cell. Membranes are changed chemically when they form parts of certain organelles. In the Golgi apparatus, for example, the membranes of the *cis* face closely resemble those of the endoplasmic reticulum in chemical composition, but those of the *trans* face are more similar to the cell membrane.

CHECKpoint CONCEPT 5.1

- ✓ What are the differences between peripheral and integral membrane proteins?
- ✓ A membrane protein has the following amino acid sequence (see Table 3.2 for abbreviations):

  EWDRHDFESGPTFIWLIWLVLAVLFLLLWAVLRPGKYKDKHE

  Considering the R groups on the amino acids, predict the region of the protein that will be embedded within the membrane.
- ✓ What is the evidence for membrane fluidity?
- ✓ If the cells of certain sponges are separated, they reaggregate because of binding between their membrane-associated proteoglycans. What would happen if the same experiment were conducted with cells treated to remove cell surface carbohydrates?

Now that you understand the structure of biological membranes, let's see how their components function. In the sections that follow, we will focus on the cell membrane. We'll look at how the cell membrane regulates the passage of substances that enter or leave a cell. Bear in mind that these principles also apply to the membranes that surround organelles.

CONCEPT 5.2 Passive Transport across Membranes Requires No Input of Energy

An important property of all life is the ability to regulate the internal composition of a cell, distinguishing it from the surrounding environment. Biological membranes allow some substances, but not others, to pass through them. This characteristic of membranes is called **selective permeability**. If a membrane is permeable to a particular substance, that substance can simply diffuse (as we describe below) across the membrane from a region of higher concentration to a region of lower concentration. However, some substances must be transported across membranes, and this process is facilitated by specialized membrane proteins.

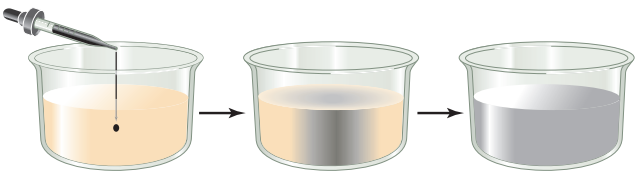
There are two fundamentally different processes by which substances cross biological membranes:

- The processes of **passive transport** do not require direct inputs of metabolic energy to drive them. In general, passive transport occurs when a substance moves from the side of the membrane where its concentration is higher to the side where its concentration is lower. In other words, the substance moves *down its concentration gradient*. Passive transport can also occur when an ion is transported across a membrane to reduce charge differences between the two sides of the membrane.
- The processes of **active transport** require the input of metabolic energy because they involve the movement of substances *against their concentration gradients*. Even if the membrane is permeable to a substance, that substance must be actively transported from the side where its concentration is lower to the side where it is higher.

This section focuses on passive transport across the membrane. Passive transport can occur by simple diffusion through the phospholipid bilayer, or it can be facilitated by channel proteins or carrier proteins.

Simple diffusion takes place through the phospholipid bilayer

In a solution, there is a tendency for all of the components to be evenly distributed. You can see this when a drop of ink is allowed to fall into a gelatin suspension (a "gel"). Initially the pigment molecules are very concentrated, but they will move about at random, slowly spreading until the intensity of color is exactly the same throughout the gel:



A solution in which the solute molecules are uniformly distributed is said to be at equilibrium. This does not mean the molecules have stopped moving; it just means they are moving in such a way that their overall distribution does not change.

Diffusion is the process of random movement toward a state of equilibrium. In effect, it is a net movement from regions of greater concentration to regions of lesser concentration. Diffusion is generally a *very slow process in living tissues* except over short distances, especially when we consider the gel-like consistency of the cell cytoplasm. For example, it would take about 3 years for a molecule of oxygen gas (O_2) to diffuse from the human lung to a cell at the fingertip! So it is not surprising that as plants and animals evolved and became larger and multicellular, those with circulatory systems to distribute vital molecules such as O_2 had a distinct advantage over organisms relying on simple diffusion.

How fast a substance diffuses depends on three factors:

- The *diameter* of the molecules or ions: smaller molecules diffuse faster.
- The *temperature* of the solution: higher temperatures lead to faster diffusion because the heat provides more energy for movement.
- The *concentration gradient* in the system—that is, the change in solute concentration with distance in a given direction. The greater the concentration gradient, the more rapidly a substance diffuses.

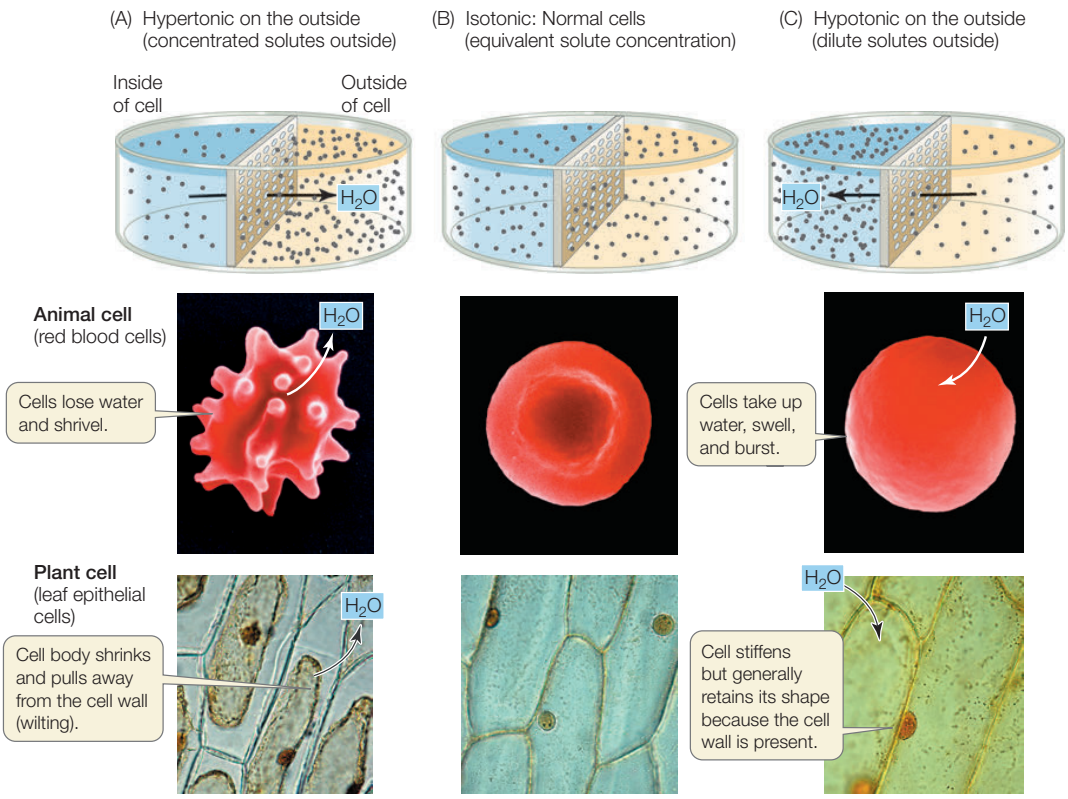


FIGURE 5.3 Osmosis Can Modify the Shapes of Cells
(A) In a solution that is hypertonic to the cytoplasm of a plant or animal cell, water flows out of the cell. (B) In a solution that is isotonic with the cytoplasm, the cell maintains a consistent, characteristic shape because there is no net movement of water into or out of the cell. (C) In a solution that is hypotonic to the cytoplasm, water enters the cell. An animal cell will swell and may burst under these conditions; a plant cell will not swell too much because of its rigid cell wall.

What does this mean for a cell surrounded by a membrane? The cytoplasm is largely a water-based (aqueous) solution, and so is the surrounding environment. In a complex solution (one with many different solutes), the diffusion of each solute depends only on its own concentration, not on the concentrations of other solutes. So one might expect a substance with a higher concentration inside the cell to diffuse out, and one with a higher concentration outside the cell to diffuse in. Indeed, some small molecules can pass through the phospholipid bilayer of the membrane by **simple diffusion**. Gases, including oxygen and carbon dioxide, can cross membranes this way. Small nonpolar and uncharged molecules can enter the membrane readily and pass through it. The more lipid-soluble the molecule is, the more rapidly it diffuses through the lipid bilayer.

In contrast, electrically charged or polar (hydrophilic) molecules, such as amino acids, sugars, ions, and water, do not pass readily through a membrane because they are not soluble in the hydrophobic interior of the lipid bilayer. However, as we discuss below, specialized proteins facilitate the transport of these molecules across membranes.

Osmosis is the diffusion of water across membranes

Water molecules pass through specialized channels in membranes (see below) by a diffusion process called **osmosis**. This process depends on the relative concentrations of water molecules on both sides of the membrane. In a particular solution, the higher the *total* solute concentration, the lower the concentration of water molecules. **Osmotic pressure** is defined as the pressure that needs to be applied to a solution to prevent the flow of water across a membrane by osmosis. This pressure is proportional to the total concentration of solutes in the solution—the more dissolved solutes there are, the fewer water molecules there are, and so water moves across the membrane and into the solution. The equation for osmotic pressure (symbolized by the Greek capital letter *pi*, Π) due to water is

$$\Pi = cRT$$

where c is total solute concentration, R is the gas constant, and T is the absolute temperature. In thermodynamic terms, the higher concentration of a substance in a compartment on one side of a membrane represents stored energy.

Consider a situation where a membrane separating two different solutions allows water, *but not solutes*, to pass through. The water molecules will move across the membrane toward the solution with the higher solute concentration and the lower concentration of water molecules.

Here we are referring to the *net* movement of water. Since it is so abundant, water is constantly moving (through channel proteins) across the cell membrane, into and out of cells. But if there is a concentration difference between the two sides of the membrane, the overall movement will be greater in one direction or the other.

Three terms are used to compare the solute concentrations of two solutions separated by a membrane:

- A **hypertonic** solution has a higher solute concentration than the other solution (**FIGURE 5.3A**).

- **Isotonic** solutions have equal solute concentrations (FIGURE 5.3B).
- A **hypotonic** solution has a lower solute concentration than the other solution (FIGURE 5.3C).

The concentration of solutes in the environment determines the direction of osmosis in all animal cells. A red blood cell takes up water from a solution that is hypotonic to the cell's contents. If this happens, the cell bursts because its cell membrane cannot withstand the pressure created by the water entry and the resultant swelling (see Figure 5.3C). Conversely, the cell shrinks if the solution surrounding it is hypertonic to its contents (see Figure 5.3A). The integrity of blood cells is absolutely dependent on the maintenance of a constant solute concentration in the surrounding blood plasma—the plasma must be isotonic to the blood cells. Regulation of the solute concentrations of body fluids is thus an important process for organisms without cell walls.

In contrast to animal cells, the cells of plants, archaea, bacteria, fungi, and some protists have cell walls that limit their volumes and keep them from bursting. Cells with sturdy walls take up a limited amount of water, and in so doing they build up internal pressure against the cell wall, which prevents further water from entering. This pressure within the cell is called **turgor pressure**; it keeps the green parts of plants upright and is the driving force for enlargement of plant cells (see Concept 25.3). It is a normal and essential component of plant growth. If enough water leaves the cells, turgor pressure drops and the plant wilts. Turgor pressure reaches about 100 pounds per square inch (0.7 kg/cm^2), which is greater than the pressure in auto tires (about 35 pounds per square inch).

LINK

The roles of osmosis in plant physiological processes are described in Concept 25.3. Excretion in animals also involves osmosis; see Concept 36.1

Diffusion may be aided by channel proteins

As we saw earlier, polar or charged substances such as water, amino acids, sugars, and ions do not readily diffuse across membranes. But they can cross the hydrophobic phospholipid bilayer passively (that is, without the input of energy) in one of two ways, depending on the substance:

- **Channel proteins** are integral membrane proteins that form channels across the membrane through which certain substances can pass.
- Some substances can bind to membrane proteins called **carrier proteins** that speed up their diffusion through the phospholipid bilayer.

Both of these processes are forms of **facilitated diffusion**. The substances diffuse according to their concentration gradients, but their diffusion is made easier by channel or carrier proteins. Particular channel or carrier proteins allow diffusion both into and out of a cell or organelle. In other words, they can operate in both directions.

We will focus here on two examples of channel proteins and discuss carrier proteins in the next section.

ION CHANNELS The best-studied channel proteins are the **ion channels**. As you will see in later chapters, the movement of ions across membranes is important in many biological processes, including ATP production within the mitochondria, the electrical activity of the nervous system, and the opening of pores in plant leaves to allow gas exchange with the environment. Several types of ion channels have been identified, each of them specific for a particular ion. All of them show the same basic structure of a hydrophilic pore that allows a particular ion to move through it.

Just as a fence may have a gate that can be opened or closed, most ion channels are "gated": they can be opened or closed to ion passage. A **gated channel** opens when a stimulus causes a change in the three-dimensional shape of the channel. In some cases, this stimulus is the binding of a chemical signal, or **ligand**. Channels controlled in this way are called ligand-gated channels (FIGURE 5.4). In contrast, a voltage-gated channel is stimulated to open or close by a change in the voltage (electrical charge difference) across the membrane (see Figure 34.5).

AQUAPORINS Water crosses membranes at a much faster rate than would be expected if it simply diffused through the phospholipid bilayer. One way water does this is by "hitchhiking" with some ions, such as Na^+, as they pass through ion channels. Up to 12 water molecules may coat an ion as it

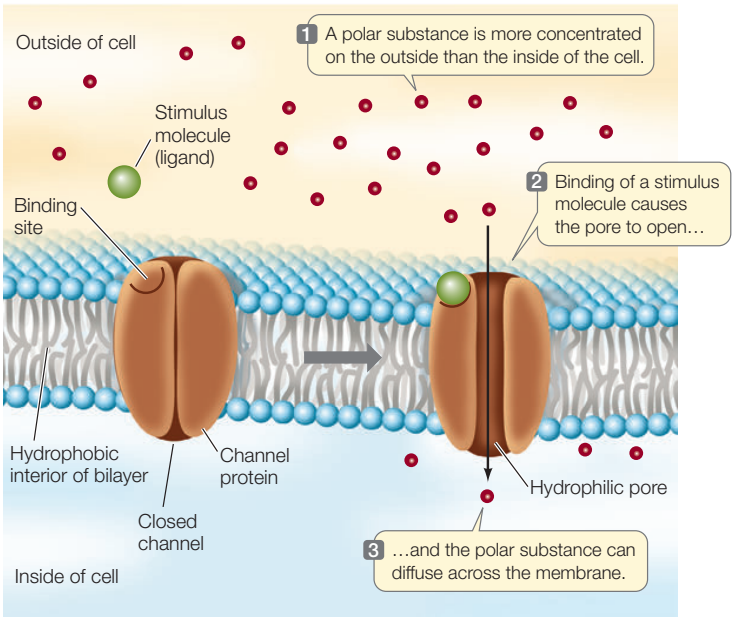


FIGURE 5.4 A Ligand-Gated Channel Protein Opens in Response to a Stimulus The channel protein is anchored in the lipid bilayer by the nonpolar (hydrophobic) amino acids exposed on the protein's surface. The protein changes its three-dimensional shape when a stimulus molecule (ligand) binds to it, opening a pore lined with polar amino acids. This allows hydrophilic, polar substances to pass through.

traverses a channel. But there is an even faster way for water to cross membranes. Plants and some animal cells (such as red blood cells and kidney cells) have membrane channels called **aquaporins**. These specific channels allow large amounts of water to move down its concentration gradient, as you will see when we discuss water relations in plants (see Chapter 25) and animals (see Chapter 36).

Aquaporins were first identified by Peter Agre at Duke University. He noticed a membrane protein that was present in red blood cells, kidney cells, and plant cells, all of which show rapid diffusion of water across their membranes. To test the idea that the membrane protein might be a water channel, Agre injected egg cells (oocytes) with the mRNA for the protein. The injected cells produced the protein and inserted it into their membranes. An oocyte membrane does not normally permit much diffusion of water. However, the injected oocytes began swelling immediately after being transferred to a hypotonic solution, indicating the rapid diffusion of water into the cells (FIGURE 5.5).

Carrier proteins aid diffusion by binding substances

Another kind of facilitated diffusion involves the actual binding of the transported substance to a membrane protein called a carrier protein. Carrier proteins transport polar molecules such as sugars and amino acids.

Glucose is the major energy source for most mammalian cells, and they require a great deal of it. Their membranes contain a carrier protein—the glucose transporter—that facilitates glucose uptake into the cell. Binding of glucose to a specific three-dimensional site on one side of the transport protein causes the protein to change its shape and release glucose on the other side of the membrane (FIGURE 5.6A). Since glucose is usually broken down as soon as it enters the cell, there is almost always a strong concentration gradient favoring glucose entry (that is, a higher concentration outside the cell than inside). The transporter allows glucose molecules to cross the membrane and enter the cell much faster than they would by simple diffusion through the bilayer. This rapid entry is necessary to ensure that the cell receives enough glucose for its energy needs.

Transport by carrier proteins is different from simple diffusion. In both processes, the rate of movement depends on the concentration gradient across the membrane. However, in carrier-mediated transport, a point is reached at which increases in the concentration gradient are not accompanied by an increased rate of diffusion. At this point, the facilitated diffusion system is said to be saturated (FIGURE 5.6B). Because there are only a limited number of carrier protein molecules per unit of membrane area, the rate of diffusion reaches a maximum when all the carrier molecules are fully loaded with solute molecules. This situation is similar to that of enzyme saturation (see Figure 3.16).

Go to ANIMATED TUTORIAL 5.2
Passive Transport
PoL2e.com/at5.2

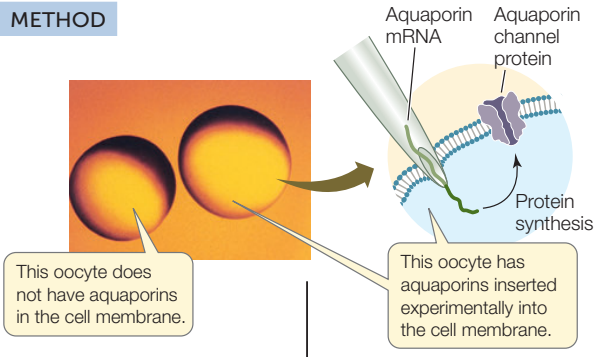
INVESTIGATION

FIGURE 5.5 Aquaporins Increase Membrane Permeability to Water A protein was isolated from the membranes of cells in which water diffuses rapidly across the membranes. When mRNA encoding the protein was inserted into and translated in oocytes, which do not normally have the protein, the water permeability of the oocytes was greatly increased.[a]

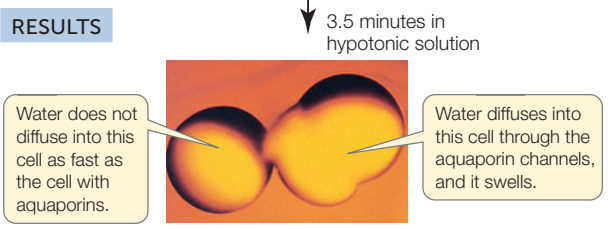
HYPOTHESIS

Aquaporin increases membrane permeability to water.

METHOD



RESULTS

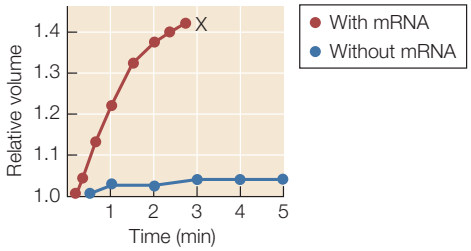


CONCLUSION

Aquaporin increases the rate of water diffusion across the cell membrane.

ANALYZE THE DATA

Oocytes were injected with aquaporin mRNA (red circles) or a solution without mRNA (blue circles). Water permeability was tested by incubating the oocytes in hypotonic solution and measuring cell volume. After time X in the upper curve, intact oocytes were not visible:



A. Why did the cells with aquaporin mRNA increase in volume?

B. What happened at time X?

C. Calculate the relative rates (volume increase per minute) of swelling in the control and experimental curves. What does this show about the effectiveness of mRNA injection?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]G. M. Preston et al. 1992. *Science* 256: 385–387.

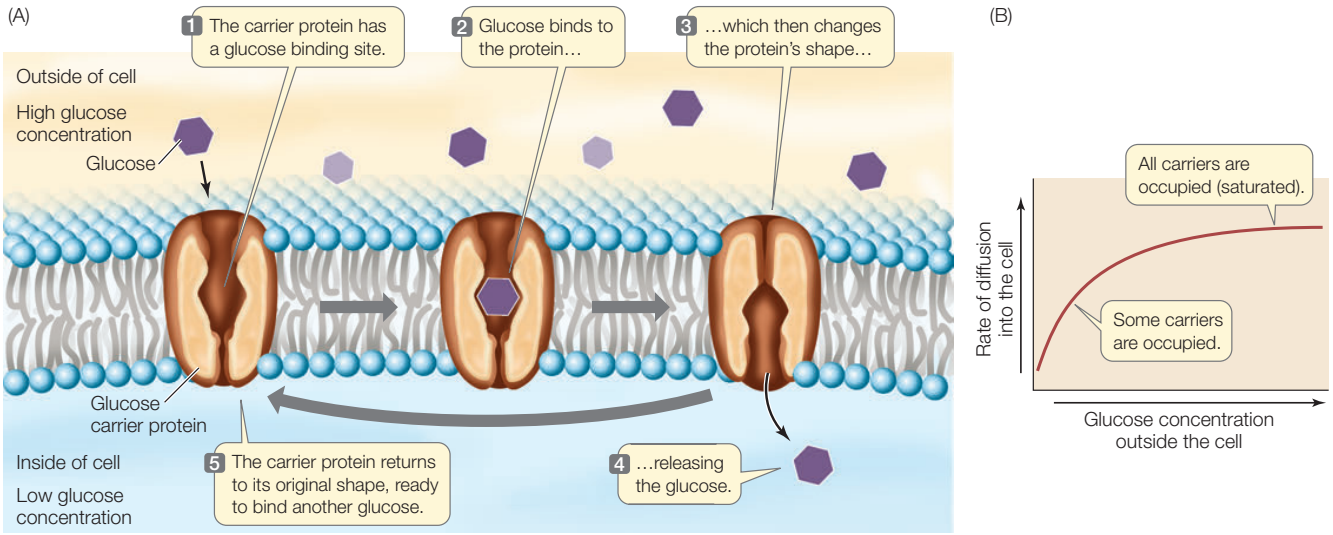


FIGURE 5.6 A Carrier Protein Facilitates Diffusion The glucose transporter is a carrier protein that allows glucose to enter the cell at a faster rate than would be possible by simple diffusion. (A) The transporter binds to glucose, and as it does so, it changes shape, releasing the glucose into the cell cytoplasm. (B) The graph shows the rate of glucose entry via a carrier versus the concentration of glucose outside the cell. As the glucose concentration increases, the rate of diffusion increases until the point at which all the available transporters are being used (the system is saturated).

CHECKpoint CONCEPT 5.2

- ✓ What properties of a substance determine whether, and how fast, it will diffuse across a membrane?
- ✓ Compare the process of facilitated diffusion through a channel and by a carrier protein. Which might be faster, and why?
- ✓ After celery is stored in an open, dry container in the refrigerator for two days, it is wilted. However, immersing the cut stalk in water for a few hours restores the integrity of the celery. How?

Diffusion tends to equalize the concentrations of substances between the outsides and insides of cells or organelles. However, one hallmark of a living thing is that it can have an internal composition quite different from that of its environment. To achieve this, a cell must sometimes move substances *against their concentration gradients*. This process requires work—the input of energy—and is known as active transport.

CONCEPT 5.3 Active Transport Moves Solutes against Their Concentration Gradients

In many biological situations, there is a different concentration of a particular ion or small molecule inside compared with outside a cell. In these cases, the concentration imbalance is maintained by a protein in the cell membrane that moves the substance against its concentration gradient. This is called active transport, and because it is acting "against the normal flow," it requires the expenditure of energy. Often the energy source is the nucleotide adenosine triphosphate (ATP). In eukaryotes, ATP is produced in the mitochondria and plastids, and it has chemical energy stored in its terminal phosphate bond. This energy is released when ATP is converted to adenosine diphosphate (ADP) in a hydrolysis reaction that breaks the bond between the terminal phosphate and the rest of the molecule.

LINK

You will find more details about how ATP functions as an energy shuttle in cells in **Concept 6.1**

The differences between diffusion and active transport are summarized in **TABLE 5.1**. In many cases of simple and facilitated diffusion, ions or molecules can move down their concentration gradients in either direction across the cell membrane. In contrast, *active transport is directional*, and moves a substance either into or out of a cell or organelle, depending on the transport protein's function. As in facilitated diffusion, there is usually a specific carrier protein for each substance that is transported.

Different energy sources distinguish different active transport systems

There are two basic types of active transport:

- **Primary active transport** involves the direct hydrolysis of ATP, which provides the energy required for transport.
- **Secondary active transport** does not use ATP directly. Instead, its energy is supplied by an ion concentration gradient or an electrical gradient, established by primary active transport. This transport system uses the energy of ATP indirectly to set up the gradient.

TABLE 5.1 Membrane Transport Mechanisms

	Simple diffusion	Facilitated diffusion (channel or carrier protein)	Active transport
Cellular energy required?	No	No	Yes
Driving force	Concentration gradient	Concentration gradient	ATP hydrolysis (against concentration gradient)
Membrane protein required?	No	Yes	Yes
Specificity	No	Yes	Yes

In primary active transport, energy released by the hydrolysis of ATP drives the movement of specific ions against their concentration gradients. For example, the concentration of potassium ions (K^+) inside a cell is often much higher than the concentration in the fluid bathing the cell. However, the concentration of sodium ions (Na^+) is often much higher outside the cell. A protein in the cell membrane pumps Na^+ out of the cell and K^+ into the cell against these concentration gradients, ensuring that the gradients are maintained. This **sodium–potassium (Na^+–K^+) pump** is an integral membrane glycoprotein that is found in all animal cells. It breaks down a molecule of ATP to ADP and a free phosphate ion (P_i) and uses the released energy to bring two K^+ ions into the cell, and export three Na^+ ions (FIGURE 5.7).

In secondary active transport, the movement of a substance against its concentration gradient is accomplished using energy "regained" by letting ions move across the membrane *down* their concentration gradients. For example, once the Na^+–K^+ pump establishes a concentration gradient of sodium ions, the passive diffusion of some Na^+ back into a cell can provide energy for the secondary active transport of glucose into the cell. This occurs when glucose is absorbed into the bloodstream from the digestive tract. Secondary active transport is usually accomplished by a single protein that moves both the ion and the actively transported molecule across the membrane. In some cases, the ion and the transported molecule move in opposite directions, whereas in others they move in the same direction (as for glucose and Na^+ in the digestive tract). Secondary active transport aids in the uptake of amino acids and sugars, which are essential raw materials for cell maintenance and growth.

Go to ANIMATED TUTORIAL 5.3
Active Transport
PoL2e.com/at5.3

CHECKpoint CONCEPT 5.3

- ✓ Why is energy required for active transport?
- ✓ The drug ouabain inhibits the activity of the Na^+–K^+ pump. A nerve cell is incubated in ouabain. Make a table in which you predict what would happen to the concentrations of Na^+ and K^+ inside and outside the cell, as a result of the action of ouabain.
- ✓ How would you use experiments to distinguish between the following two ways for glucose to enter a cell: (1) facilitated diffusion via a carrier protein and (2) secondary active transport?

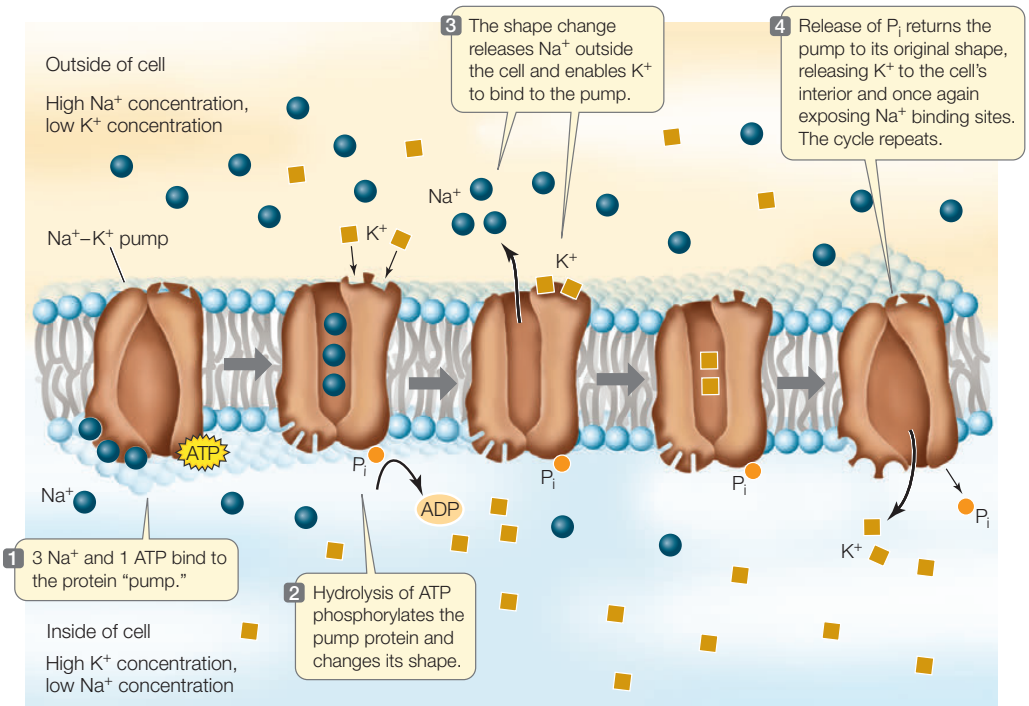


FIGURE 5.7 Primary Active Transport: The Sodium–Potassium Pump In active transport, energy is used to move a solute against its concentration gradient. Here, energy from ATP is used to move Na^+ and K^+ against their concentration gradients.

We have examined a number of passive and active ways by which ions and small molecules can enter and leave cells. But what about large molecules such as proteins? Many proteins are so large that they diffuse very slowly, and their bulk makes it difficult for them to pass through the phospholipid bilayer. It takes a completely different mechanism to move intact large molecules across membranes.

CONCEPT 5.4 Large Molecules Cross Membranes via Vesicles

Macromolecules such as proteins, polysaccharides, and nucleic acids are simply too large and too charged or polar to pass through biological membranes. This is a fortunate property—cellular integrity depends on containing these macromolecules in specific locations. However, cells must sometimes take up or **secrete** (release to the external environment) intact large molecules. This is done via vesicles, and the general terms for the mechanisms by which cells secrete and take up large molecules or particles are exocytosis and endocytosis (FIGURE 5.8).

Exocytosis moves materials out of the cell

Exocytosis is the process by which materials packaged in vesicles are secreted from the cell (see Figure 5.8B). When the vesicle membrane fuses with the cell membrane, an opening is made to the outside of the cell. The contents of the vesicle are released into the environment, and the vesicle membrane is smoothly incorporated into the cell membrane.

In Chapter 4 we encountered exocytosis as the last step in the processing of material engulfed by phagocytosis—the release of undigested materials back to the extracellular environment (see Figure 4.9). Secreted proteins are also transported out of the cell via exocytosis. The proteins are folded and modified in the endoplasmic reticulum and then transported in vesicles to the Golgi apparatus, where they may be further modified. Finally, the proteins are packaged in new vesicles for secretion (see Figure 4.8).

Exocytosis is important in the secretion of many types of substances, including digestive enzymes from the pancreas, neurotransmitters from neurons, and materials for the construction of the plant cell wall. You will encounter these processes in later chapters.

Macromolecules and particles enter the cell by endocytosis

Endocytosis is a general term for a group of processes that bring small molecules, macromolecules, large particles, and even small cells into eukaryotic cells (see Figure 5.8A). The cell membrane invaginates (folds inward), forming a small pocket around materials from the environment. The pocket deepens, forming a vesicle. This vesicle separates from the cell membrane and migrates with its contents to the cell's interior.

Endocytosis often depends on **receptors** (see Concept 5.5), which are proteins that bind to specific molecules (their ligands) and then set off specific cellular responses. In endocytosis, the receptors are integral membrane proteins located on

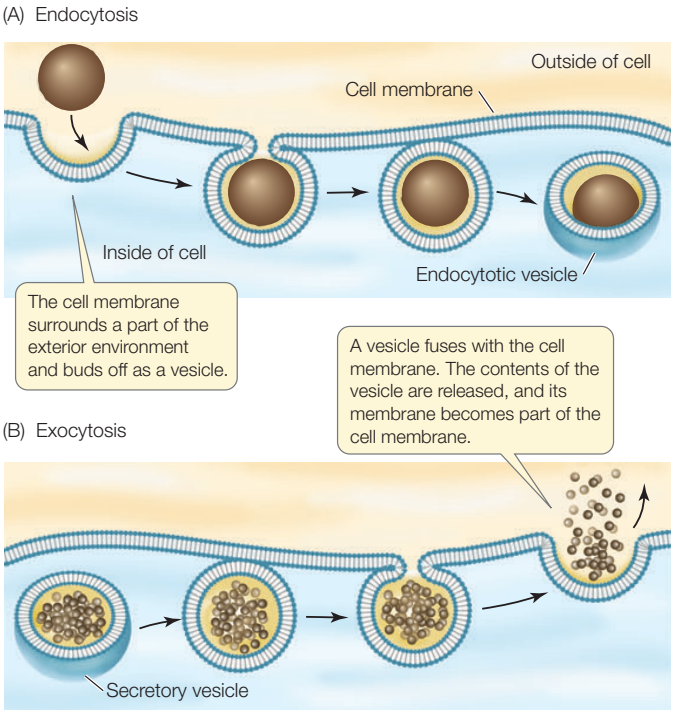


FIGURE 5.8 Endocytosis and Exocytosis Eukaryotic cells use endocytosis (A) and exocytosis (B) to take up and release large molecules and particles. Even small cells can be engulfed via endocytosis.

Go to MEDIA CLIP 5.1
An Amoeba Eats by Phagocytosis
PoL2e.com/mc5.1

the extracellular surface of the cell membrane. Vesicle formation results in the internalization of both the receptor and its ligand, along with other substances present near the site of invagination.

There are three broad types of endocytosis: phagocytosis, pinocytosis, and receptor endocytosis:

- In **phagocytosis** ("cellular eating"), receptors in the cell membrane recognize a specific ligand on the surface of a large particle or even an entire cell. The binding of the ligand to the receptor causes the phagocytic cell to engulf the particle or other cell. Phagocytosis is restricted to specialized cells; for example, unicellular protists use phagocytosis for feeding, and some white blood cells use phagocytosis to engulf foreign cells and substances. The food vesicle (phagosome) that forms usually fuses with a lysosome, where the vesicle's contents are digested.

LINK
Review the discussion of phagocytosis in **Concept 4.3**

- Vesicles also form in **pinocytosis** ("cellular drinking"). However, in this case the vesicles bring fluids and dissolved substances, including proteins, into the cell. Pinocytosis is relatively nonspecific regarding what it brings into the cell. For example, pinocytosis goes on constantly in

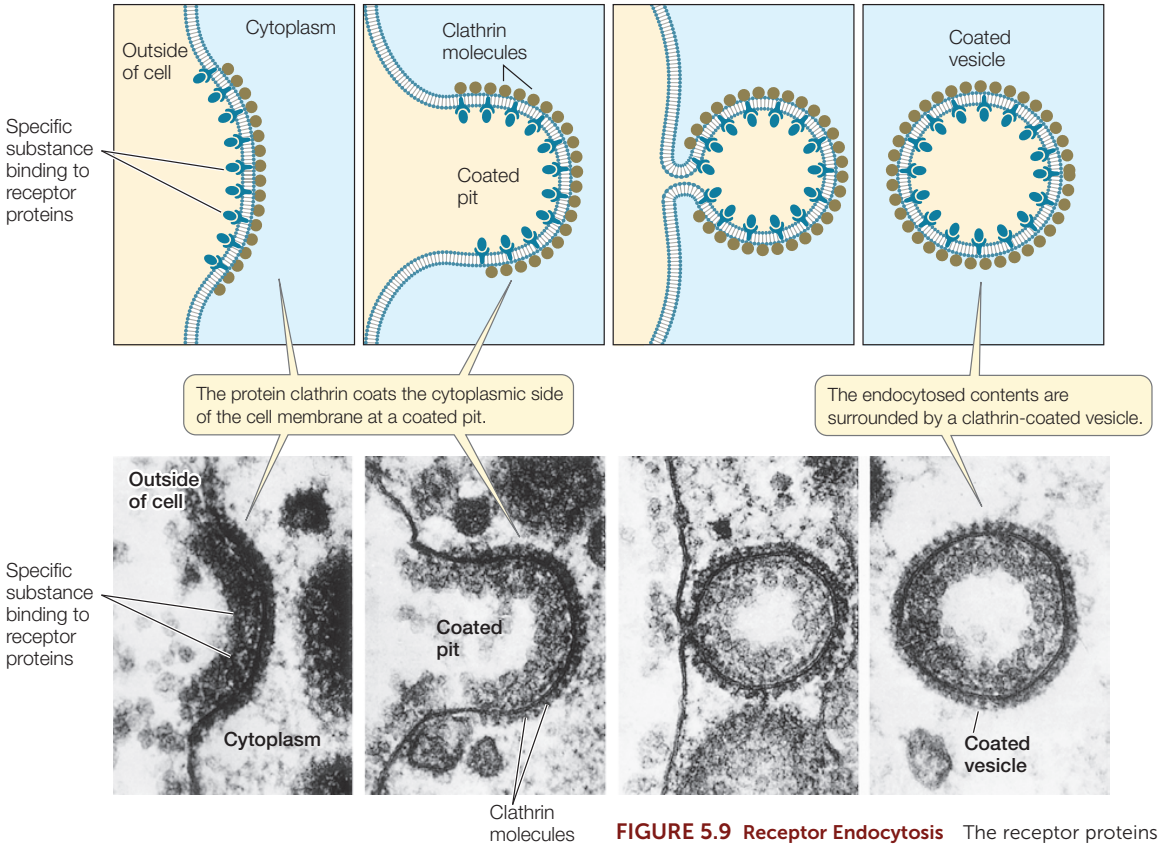


FIGURE 5.9 Receptor Endocytosis The receptor proteins in a coated pit bind specific macromolecules, which are then carried into the cell by a coated vesicle.

the endothelium—the single layer of cells that separates a blood capillary from the surrounding tissue. Pinocytosis allows cells of the endothelium to rapidly acquire fluids and dissolved solutes from the blood.

- **Receptor endocytosis** (also called receptor-mediated endocytosis) is a mechanism for bringing specific large molecules, recognized by specific receptors, into the cell. In recent years it has become clear that receptor endocytosis also plays an important role in cell signaling, which we will discuss in Concepts 5.5 and 5.6. Put simply, receptor endocytosis allows cells to control their internal processes by controlling the location and abundance of each type of receptor on the cell membrane.

Let's take a closer look at the process of receptor endocytosis.

Receptor endocytosis often involves coated vesicles

In receptor endocytosis, the receptors are often located at particular regions, called coated pits, on the extracellular surface of the cell membrane. These pits form slight depressions in the cell membrane, and their cytoplasmic surfaces are coated by another protein, often clathrin. The uptake process is similar to that in phagocytosis. The clathrin (or other protein) molecules strengthen and stabilize the vesicle (FIGURE 5.9).

Once inside the cell, the vesicle loses its clathrin coat and fuses with a membrane-enclosed compartment called an **endosome**, where the ligands, receptors, and other substances in the vesicle are separated and sorted. Some of these components are transferred to the lysosome for degradation, while others may be transferred back to the cell membrane. Thus a receptor may be recycled to the cell membrane or degraded in the lysosome, and as we mentioned above, this is an important mechanism by which the cell controls the abundance of each kind of receptor at its surface.

Receptor endocytosis is the way cholesterol is taken up by most mammalian cells. Cholesterol and triglycerides, which have low solubility in water, are packaged by liver cells into lipoprotein particles. Most of the cholesterol is packaged into low-density lipoproteins (LDLs) and circulated via the bloodstream. When a particular cell requires cholesterol, it produces LDL receptors, which are inserted into the cell membrane. The receptors diffuse laterally through the membrane until they become associated with clathrin-coated pits. LDLs bind to the receptors and are taken into the cell via receptor endocytosis. After separation from the receptors, the LDL particles are transferred to the lysosome, where they are broken down and the cholesterol made available for use by the cell.

APPLY THE CONCEPT

Some substances require energy to cross the membrane

The liver plays several vital metabolic roles, including protein synthesis, detoxification, and the production of substances necessary for digestion. Liver cells are in contact with the blood and exchange a variety of substances with the blood plasma (the noncellular part of blood). Below is a list of observations about the relative concentrations of various molecules in a liver cell cytoplasm and in the blood plasma. Explain each observation in terms of membrane permeability and transport mechanisms.

1. The concentration of serum albumin, a blood protein synthesized in the liver, is much higher in the plasma.
2. The concentration of RNA is much higher in the cytoplasm.
3. The concentration of Na^+ is lower in the cytoplasm.
4. The concentration of water is equal in the plasma and the cytoplasm.
5. The concentration of low-density lipoproteins is higher in the cytoplasm.
6. The concentration of glucose is equal in the plasma and the cytoplasm.
7. If K^+ enters the plasma, its concentration rapidly equalizes between the plasma and the cytoplasm.

In healthy individuals, the liver takes up unused LDLs for recycling. People with the inherited disease familial hypercholesterolemia have a defective LDL receptor in their livers. This prevents receptor endocytosis of LDLs in the liver, resulting in dangerously high levels of cholesterol in the blood. The cholesterol builds up in the arteries that nourish the heart and causes heart attacks. In extreme cases where only the defective receptor is present, children and teenagers can have severe cardiovascular disease.

Receptor endocytosis also plays an important role in cell signaling, which we will discuss in the following concepts.

Go to ANIMATED TUTORIAL 5.4
Endocytosis and Exocytosis
PoL2e.com/at5.4

CHECKpoint CONCEPT 5.4

- ✓ What is the difference between phagocytosis and pinocytosis?
- ✓ Would a small molecule such as an amino acid enter a cell by receptor endocytosis?

We have just introduced the concept of a membrane-bound receptor, which is a key factor in a cell's interaction with its environment. Let's look more closely at receptors and how they respond to signals.

CONCEPT 5.5 The Membrane Plays a Key Role in a Cell's Response to Environmental Signals

A hallmark of living cells is their ability to process information from their environments. We can think of this information in terms of **cell signaling**. In this context, the signal may be a physical stimulus such as light or heat, or a chemical such as a hormone. A chemical signal may also be referred to as a **ligand**: a molecule that binds to a receptor (see Concept 5.4). The mere presence of a signal, however, does not mean a particular cell will respond to it. In order to respond, the cell must have a specific receptor that can detect the signal. Once the signal activates its receptor, it sets off a **signal transduction pathway**, a sequence of molecular events and chemical reactions within a cell that lead to the cell's response to the signal. This ability of cells to sense and respond to signals in the environment is key to the maintenance of stable intracellular conditions, a theme that recurs throughout this book.

Cells are exposed to many signals and may have different responses

Inside a large multicellular animal, chemical signals made by the body itself reach a target cell by local diffusion or by circulation within the blood. These signals are usually in tiny concentrations (as low as 10^{-10} M) and differ in their sources and mode of delivery (**FIGURE 5.10**):

- **Autocrine signals** affect the same cell that releases them. For example, many tumor cells reproduce uncontrollably because they self-stimulate cell division by making their own division signals.
- **Paracrine signals** diffuse to and affect nearby cells. An example is a neurotransmitter made by a nerve cell that diffuses to an adjacent cell and stimulates it.
- **Juxtacrine signaling** requires direct contact between the signaling and the responding cell, and usually involves interaction between signaling molecules bound to the surfaces of the two cells.
- Signals that travel through the circulatory systems of animals or the vascular systems of plants to reach receptors on distant cells are generally called **hormones**.

Chemical signals do not always come from within the multicellular organism—some come from the external environment. For example, specific molecules produced by pathogenic (disease-causing) organisms trigger signal transduction pathways in plants, leading to defense responses.

For the information from a signal to be transmitted to a cell, the target cell must be able to sense the signal and respond to it. In a multicellular animal, all the cells may receive chemical signals that are circulated in the blood, but most body cells are not capable of responding to every signal. *Only the cells with the necessary receptors can respond*.

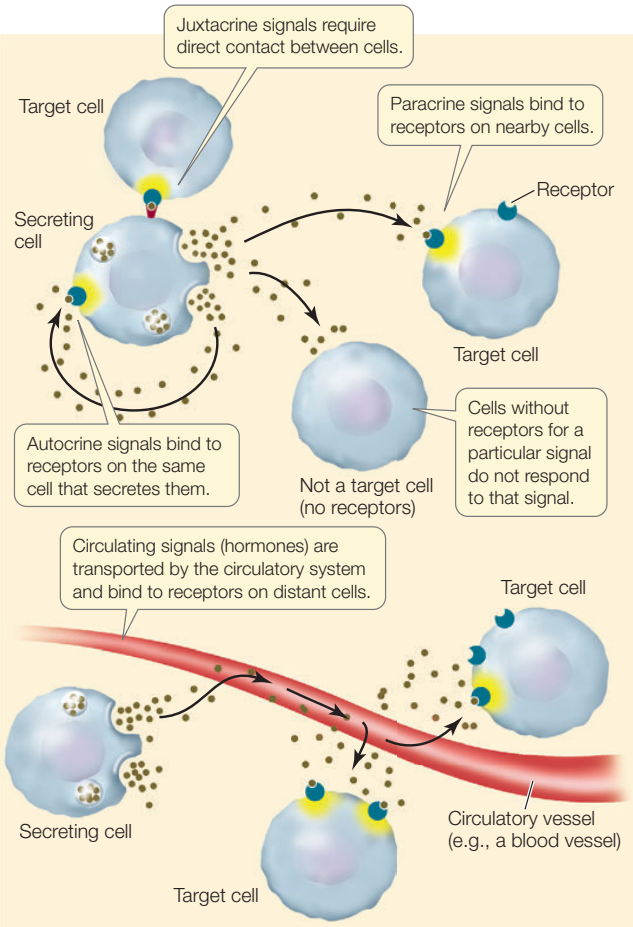


FIGURE 5.10 Chemical Signaling Concepts A signal molecule can act on the cell that produces it, on a nearby cell, or be transported by the organism's circulatory system to a distant target cell.

Typically, a signal transduction pathway involves a signal, a receptor, and a response (**FIGURE 5.11**). These pathways vary in their details, but they commonly include allosteric regulation. Recall that **allosteric regulation** involves an alteration in the three-dimensional shape of a protein as a result of the binding of another molecule at a site other than the protein's active site (see Figure 3.20). You saw an example of allosteric regulation earlier in this chapter when we considered a ligand-gated channel, which opens (changes shape) after binding to another molecule (see Figure 5.4).

A signal transduction pathway may end in a response that is short-term, such as the activation of an enzyme, or long-term, such as an alteration in gene expression.

LINK

Gene expression—the transcription of specific DNA sequences and the translation of these sequences into proteins—is described in **Chapter 10**

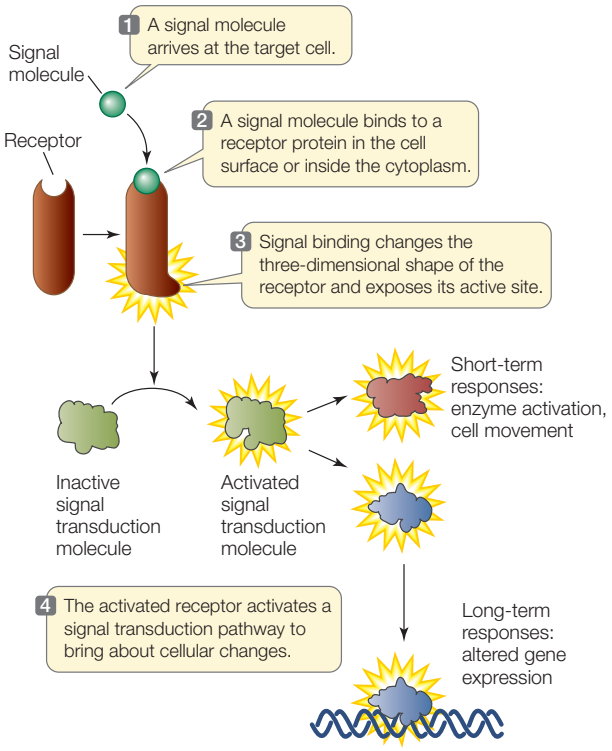


FIGURE 5.11 Signal Transduction Concepts This general pathway is common to many cells and situations. The ultimate cellular responses are either short-term or long-term.

Receptors can be classified by location and function

Chemical signals (ligands) are quite variable, but they can be divided into two groups based on whether or not they can diffuse through membranes. Physical signals such as light and sound also vary in their ability to penetrate particular cells and tissues. Accordingly, we can classify a receptor by its location in the cell, which largely depends on the nature of its ligand:

- **Intracellular receptors** are located inside the cell. Small or nonpolar ligands can diffuse across the phospholipid bilayer of the cell membrane and enter the cell. Estrogen, for example, is a lipid-soluble steroid hormone that can easily diffuse across the cell membrane; it binds to a receptor inside the cell.
- **Membrane receptors** are located on the cell surface. Large or polar ligands cannot cross the lipid bilayer. Insulin, for example, is a protein hormone that cannot diffuse through the cell membrane. Instead, it binds to a transmembrane receptor with an extracellular ligand-binding domain.

Many receptors are associated with the cell membrane

Membrane receptors are found on the surfaces of cells, and they respond to signals from outside the cell. If the signal is a

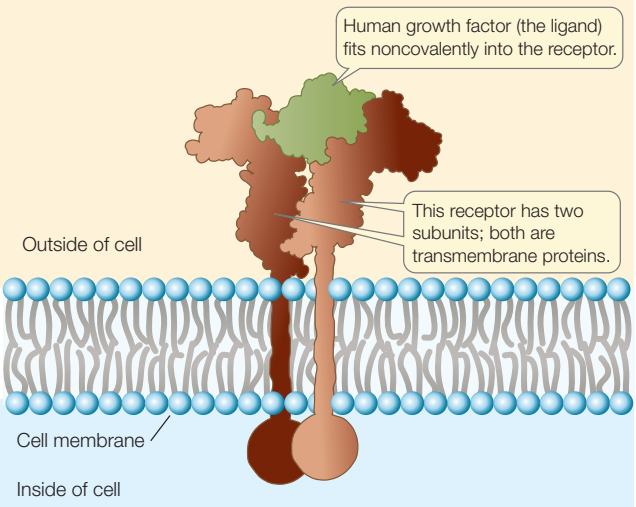
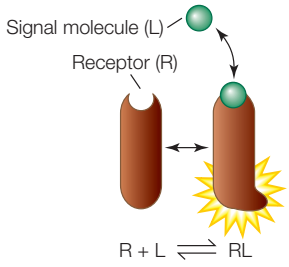


FIGURE 5.12 A Signal Binds to Its Receptor Human growth factor fits into its membrane-bound receptor (a protein with two subunits) and binds to it noncovalently.

chemical ligand, it fits into a three-dimensional site on its corresponding receptor protein (**FIGURE 5.12**). In many cases the receptor has a catalytic domain and functions as an enzyme, with its active site on the cytoplasmic side of the membrane. The ligand acts as an allosteric regulator, exposing the active site of the catalytic domain. The ligand does not contribute further to the cellular response; its role is purely to "knock on the door." (This is in sharp contrast to the enzyme–substrate interactions we described in Concept 3.3. The whole purpose of those interactions is to change substrates into useful products.)

Ligands (L) bind to their receptors (R) noncovalently and reversibly, according to chemistry's law of mass action:



Reversibility is important because if the ligand were never released, the receptor would be continuously stimulated. In most cases, the cell needs to *stop* responding to a signal after the appropriate response has occurred. For example, if a signal transduction pathway results in the production of a particular protein, the cell needs to stop producing the protein when enough of it has been made. Nevertheless, for most ligand–receptor complexes, the equilibrium point is far to the right in the above reaction—that is, binding is favored, even at low ligand concentrations.

As noted above, receptors are removed from the cell membrane by endocytosis and either degraded inside the cell or recycled back to the membrane.

An inhibitor (or antagonist) can also bind to a receptor protein, preventing the binding of the normal ligand. This is analogous to the competitive inhibition of enzymes (see Concept 3.4). There are both natural and artificial antagonists of receptor binding. For example, many substances that alter human behavior (such as caffeine; see the opening story) bind to specific receptors in the brain and prevent the binding of the receptors' specific ligands.

In complex eukaryotes such as mammals and higher plants, there are three well-studied categories of cell membrane receptors, which are grouped according to their activities: ion channels, protein kinase receptors, and G protein–linked receptors. Because you will see these receptors several times later in this book, we describe them in some detail here.

ION CHANNEL RECEPTORS As described in Concept 5.2, the cell membranes of many cells contain ligand-gated channels for ions such as Na^+, K^+, Ca^{2+}, or Cl^- (see Figure 5.4). These proteins are receptors because their functioning depends on ligand binding. An example is the acetylcholine receptor, a ligand-gated sodium channel located in the cell membranes of skeletal muscle cells. Acetylcholine is a neurotransmitter—a chemical signal released from nerve cells. Opening of the channel allows Na^+, which is more concentrated outside the cell than inside, to diffuse into the cell. This initiates a series of events that result in muscle contraction (see Figure 34.9).

PROTEIN KINASE RECEPTORS Like ligand-gated channel receptors, protein kinase receptors change shape upon ligand binding. But in this case, the new conformation exposes or activates a catalytic domain on the cytoplasmic side of the transmembrane protein that has **protein kinase** activity—it modifies specific target proteins in the cell by adding phosphate groups to them. In general, protein kinases catalyze the following reaction:

$$\text{ATP} + \text{protein} \xrightarrow{\text{Protein kinase}} \text{ADP} + \text{phosphorylated protein}$$

This reaction results in the covalent modification (phosphorylation) of the target protein, thereby changing its activity (see Figure 3.20B). Protein kinases are extraordinarily important in biological signaling: about 1 human gene in 50 is a protein kinase gene, and there is an even higher proportion of such genes in some plants.

An example of a protein kinase receptor is the receptor for the hormone insulin (**FIGURE 5.13**). The activation of this receptor results in the phosphorylation of target proteins, which then bring about the cell's response, which includes the insertion of glucose transport proteins into the cell membrane.

It should be noted that not all protein kinases are receptors—many function in later steps of signal transduction pathways, as we will discuss in Concept 5.6.

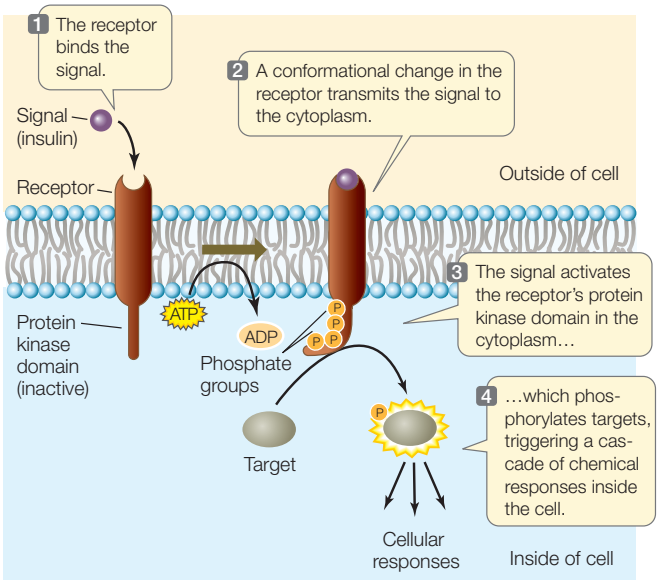


FIGURE 5.13 A Protein Kinase Receptor The mammalian hormone insulin binds to a protein kinase receptor on the outside surface of the cell and initiates a response.

G PROTEIN–LINKED RECEPTORS A third category of eukaryotic cell membrane receptors is the family of **G protein–linked receptors**. In this case, ligand binding on the extracellular domain of the receptor exposes a site on the cytoplasmic side that can bind to a mobile membrane protein called a **G protein**. The G protein is partially inserted in the lipid bilayer and partially exposed on the cytoplasmic surface of the membrane.

Many G proteins have three polypeptide subunits and can bind three different molecules:

- The G protein–linked receptor
- GDP and GTP (guanosine diphosphate and guanosine triphosphate; these are nucleotides, like ADP and ATP)
- An effector protein (a protein that causes an effect in the cell)

The activated G protein–linked receptor functions as a guanine nucleotide exchange factor. It exchanges a GDP nucleotide bound to the G protein for a GTP, inducing a shape change in the G protein. The G protein then activates the effector protein, leading to downstream signal amplification (**FIGURE 5.14**). G protein–linked receptors are especially important in the sensory systems of animals (see Concept 34.4).

Go to ANIMATED TUTORIAL 5.5
G Protein–Linked Signal Transduction and Cancer
PoL2e.com/at5.5

CHECKpoint CONCEPT 5.5

- ✓ Name three major steps in cell signaling that were discussed in this concept.
- ✓ What are the differences and similarities between ion channel receptors and G protein–linked receptors?
- ✓ If an intact cell is treated to remove cell surface proteins, will the cell be able to receive any environmental signals? Explain.

When a signal activates a receptor, a signal transduction pathway ensues. This often involves multiple steps, and it leads to one or more specific cellular responses. We will discuss these processes next.

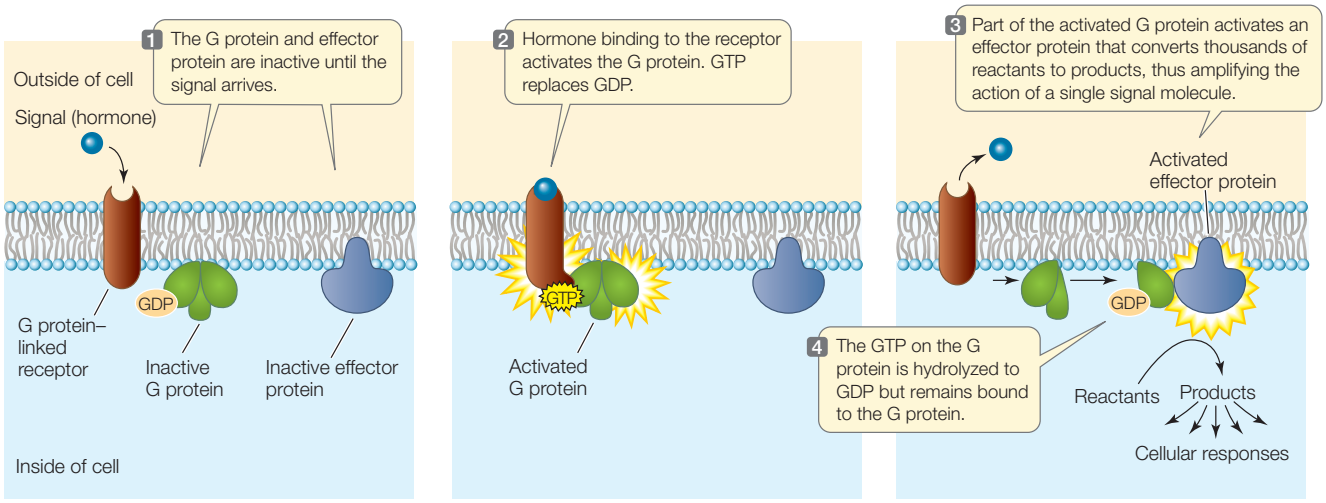


FIGURE 5.14 A G Protein–Linked Receptor The G protein is an intermediary between the receptor and an effector protein.

CONCEPT 5.6 Signal Transduction Allows the Cell to Respond to Its Environment

As we mentioned in Concept 5.5, a signal may be a chemical ligand or a physical stimulus such as light or heat. Its effect is to activate a specific receptor, leading to a cellular response that is brought about by a signal transduction pathway. Typically, signaling at the cell membrane initiates a cascade (or series) of events in the cell. Proteins interact with other proteins until the final responses are achieved. Through such a cascade, an initial signal can be both *amplified* and *distributed* to cause several different responses.

Before we discuss how signals are amplified and distributed by signal transduction pathways, let's look at some of the cellular responses that can result from cell signaling.

Cell functions change in response to environmental signals

The activation of a receptor by a signal, and the subsequent transduction and amplification of the signal, ultimately leads to changes in cell function. There are many ways in which a cell might respond, some of which we mention here:

- *Opening of ion channels* changes the balance of ion concentrations between the outside of the cell membrane and its interior (see Figure 5.4). As you will see in Chapter 34, this results in a change in the electrical potential across the membrane, with important consequences in nerve and muscle cells.
- Many signal transduction pathways lead to *alterations in gene expression*. The expression of some genes may be switched on (upregulated), whereas others may be switched off (downregulated). This alters the abundance of the proteins (often enzymes) encoded by the genes, thus changing cell function. You will see many examples that highlight the importance of gene regulation throughout this book.
- A third kind of response involves the *alteration of enzyme activities*. An example is the activation of specific enzymes in liver cells exposed to the hormone epinephrine, which we discuss below. An alteration in enzyme activity is a much more rapid response than one that involves a change in gene expression.

LINK

The different types of enzyme regulation are discussed in **Concept 3.4**

The same signal can lead to different responses in different types of cells. For example, in heart muscle cells, the hormone epinephrine *activates* a signal transduction cascade that results in glucose mobilization for energy and muscle contraction. However, in the smooth muscle cells that line the digestive tract, epinephrine stimulates a pathway that *inhibits* a target enzyme, allowing the muscle cells to relax. This increases the diameter of the blood vessels, allowing more nutrients to be carried from the digestive system to the rest of the body. Heart and digestive tract muscle cells respond differently to the same signal—epinephrine—because the signal transduction pathways stimulated by epinephrine are different in the different cell types. Let's take a closer look at the mechanism by which cells amplify and transduce signals to bring about these responses.

Second messengers can stimulate signal transduction

Often there is a small molecule intermediary between the activated receptor and the cascade of events that ensues. In a series of clever experiments, Earl Sutherland and his colleagues at Case Western Reserve University discovered that a small, water-soluble chemical could mediate cytoplasmic events initiated by a cell membrane receptor. The researchers were investigating the activation of the liver enzyme glycogen phosphorylase by the hormone epinephrine (also called adrenaline)—the "fight-or-flight" hormone (see Concept 35.2). The enzyme is activated when an animal faces life-threatening conditions and needs energy fast for the fight-or-flight response. Glycogen phosphorylase catalyzes the breakdown of glycogen stored in the liver so that the resulting glucose molecules can be released to the blood (see Figure 30.16). The enzyme is present in the liver cell cytoplasm but is inactive in the absence of epinephrine.

The researchers found that epinephrine could activate glycogen phosphorylase in liver cells that had been broken open, but only if the entire cell contents, including cell membrane fragments, were present. Under these conditions epinephrine was bound to the cell membrane fragments, but the active phosphorylase was in the solution. The researchers hypothesized that there must be a second "messenger" that transmits the epinephrine signal (the "first messenger") from the cell membrane to the phosphorylase in the cytoplasm. They investigated this by separating cell membrane fragments from the cytoplasmic fractions of broken liver cells and following the sequence of steps described in **FIGURE 5.15**. This experiment confirmed the existence of a second messenger, later identified as **cyclic AMP** (**cAMP**; **FIGURE 5.16**).

A second messenger is a small molecule that brings about later steps in a signal transduction pathway. Second messengers do not have enzymatic activity themselves; rather, they act to regulate target enzymes by binding to them noncovalently. Whereas receptor binding is highly specific, second messengers allow a cell to respond to a single event at the cell membrane with *many events inside the cell*—in other words, the second messenger *distributes* the initial signal. Second messengers also serve to *amplify* the signal—for example, the binding of a single epinephrine molecule leads to the production of many molecules of cAMP. In turn, cAMP activates many enzyme targets by binding to them noncovalently. In the case of epinephrine and the liver cell, glycogen phosphorylase is just one of several enzymes that are activated.

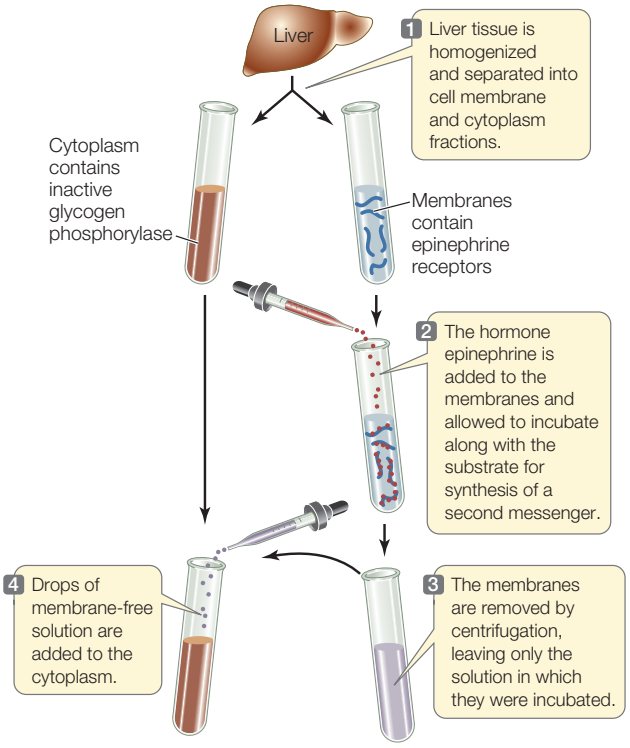
INVESTIGATION

FIGURE 5.15 The Discovery of a Second Messenger Glycogen phosphorylase is activated in liver cells after epinephrine binds to a membrane receptor. Sutherland and his colleagues observed that this activation could occur in a test tube only if fragments of the cell membrane were present. They designed experiments to show that a second messenger caused the activation of glycogen phosphorylase.[a]

HYPOTHESIS

A second messenger mediates between receptor activation at the cell membrane and enzyme activation in the cytoplasm.

METHOD



RESULTS

Active glycogen phosphorylase is present in the cytoplasm.

CONCLUSION

A soluble second messenger, produced by hormone-activated membranes, is present in the solution and activates enzymes in the cytoplasm.

ANALYZE THE DATA

The experiment was repeated under various conditions with the following results:

Condition	Enzyme activity (units)
Homogenate	0.4
Homogenate + epinephrine	2.5
Cytoplasm fraction	0.2
Cytoplasm + epinephrine	0.4
Membranes + epinephrine	0.4
Cytoplasm + membranes + epinephrine	2.0

A. What do these data show?

B. Propose an experiment to show that the factor that activates the enzyme is stable on heating (and therefore probably not a protein) and give predicted data.

C. Propose an experiment to show that cAMP can replace the membrane fraction and hormone treatment and give predicted data.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

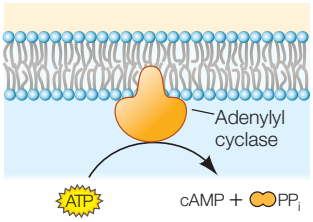
[a]T. W. Rall et al. 1957. *Journal of Biological Chemistry* 224: 463–470.

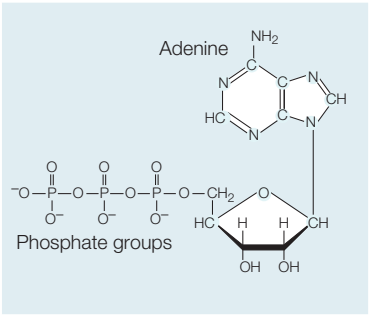
A signaling cascade involves enzyme regulation and signal amplification

Signal transduction pathways often involve multiple sequential steps, in which particular enzymes are either activated or inhibited by other enzymes in the pathway. For example, a protein kinase adds a phosphate group to a target protein, and this covalent change alters the protein's conformation and activates or inhibits its function. Cyclic AMP binds noncovalently to a target protein, and this changes the protein's shape, activating or inhibiting its function. In the case of activation, a previously inaccessible active site is exposed, and the target protein goes on to perform a new cellular role.

A good example of a signaling cascade is the G protein–mediated protein kinase pathway stimulated by epinephrine in liver cells (**FIGURE 5.17**). Binding of epinephrine to the membrane receptor results in the activation of a G protein, followed by the production of cAMP, which activates a key signaling molecule, the enzyme protein kinase A. In turn, protein kinase A phosphorylates two other enzymes, with opposite effects:

- *Inhibition*. Glycogen synthase, which catalyzes the joining of glucose molecules to form the energy-storing molecule glycogen, is inactivated when a phosphate group is added





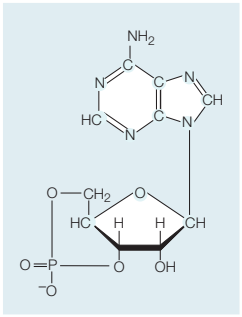


FIGURE 5.16 The Formation of Cyclic AMP The formation of cAMP from ATP is catalyzed by adenylyl cyclase, an enzyme that is activated by G proteins.

to it by protein kinase A. Thus the epinephrine signal *prevents glucose from being stored* in glycogen (see Figure 5.17, step 1).

- *Activation*. Phosphorylase kinase is activated when a phosphate group is added to it. It is part of a cascade of reactions that ultimately leads to the activation of glycogen phosphorylase, another key enzyme in glucose metabolism. This enzyme results in the liberation of glucose molecules from glycogen (see Figure 5.17, steps 2 and 3).

An important consequence of having multiple steps in a signal transduction cascade is that the signal is amplified with each step. The amplification of the signal in the pathway illustrated in Figure 5.17 is impressive. Each molecule of epinephrine that arrives at the cell membrane ultimately results in 10,000 molecules of blood glucose:

1	molecule of epinephrine bound to the membrane activates
1	molecule of adenylyl cyclase, which produces
20	molecules of cAMP, which activate
20	molecules of protein kinase A, which activate
100	molecules of phosphorylase kinase, which activate
1,000	molecules of glycogen phosphorylase, which produce
10,000	molecules of glucose 1-phosphate, which produce
10,000	molecules of blood glucose

Signal transduction is highly regulated

Signal transduction is a temporary event in the cell, and gets "turned off" once the cell has responded. We have already discussed the turnover of cell surface receptors by endocytosis. In addition, there are enzymes that convert

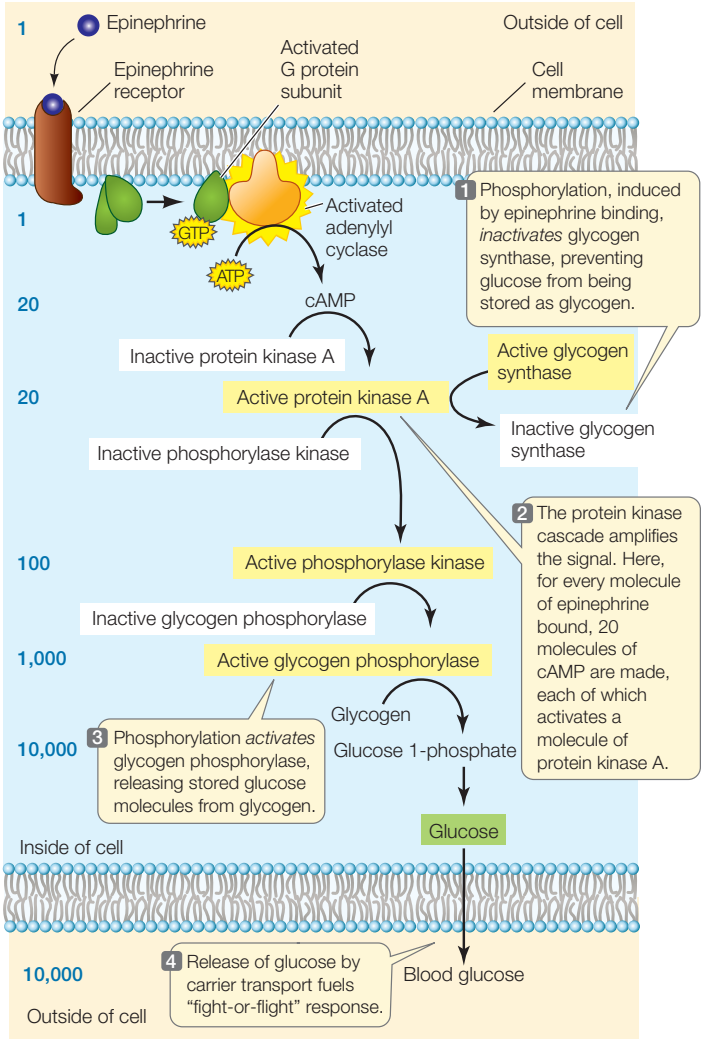


FIGURE 5.17 A Cascade of Reactions Leads to Altered Enzyme Activity Liver cells respond to epinephrine by activating G proteins, which in turn activate the synthesis of the second messenger cAMP. Cyclic AMP initiates a protein kinase cascade, greatly amplifying the epinephrine signal, as indicated by the blue numbers. The cascade both inhibits the conversion of glucose to glycogen and stimulates the release of previously stored glucose.

Go to ANIMATED TUTORIAL 5.6
Signal Transduction Pathway
PoL2e.com/at5.6

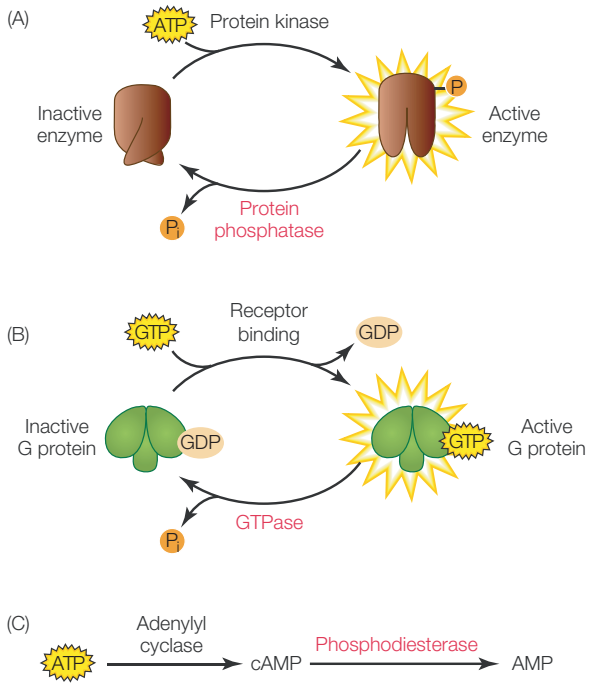


FIGURE 5.18 Signal Transduction Regulatory Mechanisms
Some signals lead to the production of active signal transduction molecules such as (A) protein kinases, (B) G proteins, and (C) cAMP. Other enzymes (red type) inactivate or remove these active molecules.

CHECKpoint CONCEPT 5.6

- ✓ Compare "first messengers" (e.g., hormones) with "second messengers" (e.g., cAMP) with regard to their chemical nature, where and when they are made, and their activity.
- ✓ Outline the steps in the amplification of signaling by epinephrine, resulting in the release of glucose to the bloodstream. At each step, is the amplification due to a covalent or noncovalent interaction?
- ✓ What would happen to a liver cell exposed to epinephrine and at the same time to a drug that inhibits protein kinase A? To epinephrine and to a drug that inhibits the hydrolysis of GTP? (Assume that both these drugs are able to cross the cell membrane.)
- ✓ The disease cholera is caused by a toxin released from the bacterium *Vibrio cholerae*. Cholera toxin causes continuous activation of a G protein at the cell membrane of cells lining the intestine. This in turn results in continuous activation of adenylyl cyclase. As a result, there is continuous release of Na^+ from the intestine, followed by massive outflow of water, resulting in severe diarrhea, dehydration, and if untreated, death. How does cholera toxin work on the second messenger system, and what is the normal role of that second messenger in the intestine cell membrane?

signal transduction molecules back to their inactive precursors. For example, protein phosphatases remove phosphate groups from target proteins, thus reversing the effects of protein kinases (**FIGURE 5.18A**). G proteins have GTPase activity, which removes a phosphate group from GTP, converting it to GDP (**FIGURE 5.18B**). Cyclic AMP is converted back to AMP by the enzyme phosphodiesterase (**FIGURE 5.18C**). The balance between the activities of these regulating enzymes and the signaling enzymes themselves is what determines the ultimate cellular response to a signal. Cells can alter this balance in several ways, including:

- Synthesis or breakdown of the enzymes involved
- Activation or inhibition of the enzymes by other molecules (see Concept 3.4)

A great deal has been learned about signal transduction pathways and cellular responses in the past two decades, and there is still much to learn. As biologists tease apart specific pathways, they find that many of them are interconnected: one pathway may be switched on by a particular signal or molecule, and another may be switched off. In this chapter we have concentrated on signaling pathways that occur in animal cells. However, signal transduction pathways are important in the functioning of all living organisms.

Q What role does the cell membrane play in the body's response to caffeine?

ANSWER Caffeine has many effects on the body, but the most noticeable is that it keeps us awake. The caffeine molecule is somewhat large and polar, and it is unlikely to diffuse through the nonpolar lipids of the cell membrane (Concept 5.2). Instead, it binds to receptors on the surfaces of nerve cells in the brain (Concept 5.5).

The nucleoside adenosine (adenine attached to a five-carbon sugar) accumulates in the brain when a person is under stress or has prolonged mental activity. When it binds to a specific receptor in the brain, adenosine sets in motion a signal transduction pathway (Concept 5.6) that results in reduced brain activity, which usually means drowsiness. This membrane-associated signaling by adenosine has evolved as a protective mechanism against the adverse effects of stress.

Caffeine has a three-dimensional structure similar to that of adenosine and is able to bind to the adenosine receptor (**FIGURE 5.19**). Because its binding does not activate the receptor, caffeine functions as an antagonist of adenosine signaling, with the result that the brain stays active and the person remains alert.

When we discussed the interaction between a ligand and its receptor, we noted that this is a reversible, noncovalent interaction. In time, after drinking coffee or tea, the caffeine molecules come off the adenosine receptors in the brain,

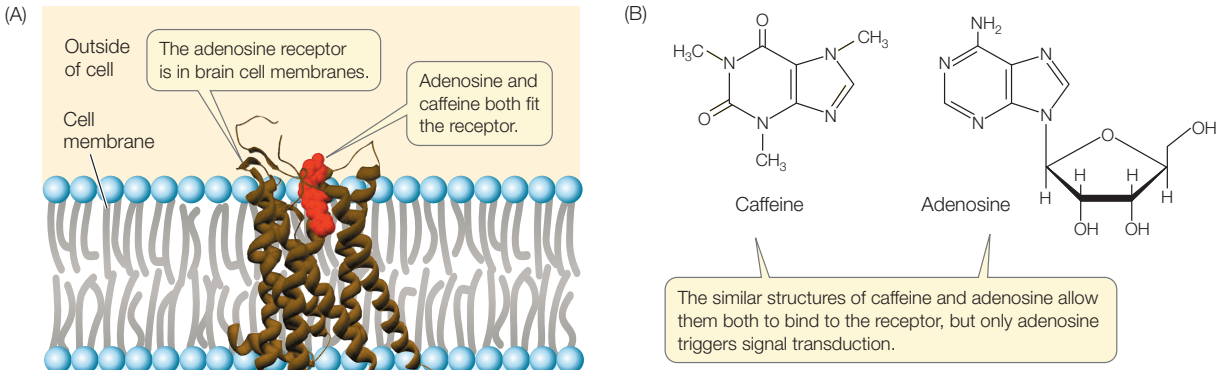


FIGURE 5.19 Caffeine and the Cell Membrane (A) The adenosine 2A receptor is present in the human brain, where it is involved in inhibiting arousal. (B) Adenosine is the normal ligand for the receptor. Caffeine has a structure similar to that of adenosine and can act as an antagonist that binds the receptor and prevents its normal functioning.

allowing adenosine to bind once again. Otherwise, coffee drinkers might never get to sleep!

In addition to competing with adenosine for a membrane receptor, caffeine blocks the enzyme cAMP phosphodiesterase. This enzyme acts in signal transduction (Concept 5.6) to break down the second messenger cAMP. Looking at the signal transduction pathway in Figure 5.17, can you explain how caffeine augments the fight-or-flight response, which includes an increase in blood sugar and increased heartbeat?

SUMMARY

CONCEPT 5.1 Biological Membranes Have a Common Structure and Are Fluid

- Biological membranes consist of lipids, proteins, and carbohydrates. The **fluid mosaic model** of membrane structure describes a **phospholipid** bilayer in which proteins can move about within the plane of the membrane.
- The two layers of a membrane may have different properties because of their different phospholipid compositions, exposed domains of **integral membrane proteins**, and **peripheral membrane proteins**. **Transmembrane proteins** span the membrane. **Review Figure 5.1, ACTIVITY 5.1, and ANIMATED TUTORIAL 5.1**

CONCEPT 5.2 Passive Transport across Membranes Requires No Input of Energy

- Membranes exhibit **selective permeability** that regulates which substances can pass through them.
- A substance can diffuse passively across a membrane by one of two processes: **simple diffusion** through the phospholipid bilayer or **facilitated diffusion**, either through a channel created by a **channel protein** or by means of a **carrier protein**. In both cases, molecules diffuse down their concentration gradients. **Review Figure 5.4 and ANIMATED TUTORIAL 5.2**
- In **osmosis**, water diffuses from a region of higher water concentration to a region of lower water concentration, largely through membrane channels called **aquaporins**. Ions diffuse across membranes through **ion channels**. **Review Figures 5.3 and 5.5**
- Carrier proteins bind to polar molecules such as sugars and amino acids and transport them across the membrane. **Review Figure 5.6**

CONCEPT 5.3 Active Transport Moves Solutes against Their Concentration Gradients

- **Active transport** requires the use of chemical energy to move substances across membranes against their concentration gradients. The **sodium–potassium** (Na^+–K^+) **pump** uses energy released from the hydrolysis of ATP. **Review Figure 5.7 and ANIMATED TUTORIAL 5.3**

CONCEPT 5.4 Large Molecules Cross Membranes via Vesicles

- **Endocytosis** is the transport of molecules, large particles, and small cells into eukaryotic cells via the invagination of the cell membrane and the formation of vesicles. **Review Figure 5.8A**
- In **receptor endocytosis**, a specific receptor on the cell membrane binds to a particular macromolecule that is to be transported into the cell. **Review Figure 5.9 and ANIMATED TUTORIAL 5.4**
- In **exocytosis**, materials in vesicles are secreted from the cell when the vesicles fuse with the cell membrane. **Review Figure 5.8B**

CONCEPT 5.5 The Membrane Plays a Key Role in a Cell's Response to Environmental Signals

- Cells receive many signals from the physical environment and from other cells. Chemical signals are often at very low concentrations. **Review Figure 5.10**
- A **signal transduction pathway** involves the interaction of a signal (often a chemical **ligand**) with a receptor; the transduction and amplification of the signal via a series of steps within the cell; and a cellular response. The response may be short-term or long-term. **Review Figure 5.11**

(continued)

SUMMARY *(continued)*

- Cells respond to signals only if they have specific receptor proteins that can be activated by those signals. Many receptors are located at the cell membrane. They include ion channels, **protein kinases**, and **G protein–linked receptors**. **Review Figures 5.13 and 5.14 and ANIMATED TUTORIAL 5.5**

CONCEPT 5.6 Signal Transduction Allows the Cell to Respond to Its Environment

- A cascade of events, one following another, occurs after a receptor is activated by a signal.
- Often, a soluble second messenger conveys signaling information from the primary messenger (ligand) at the membrane to downstream signaling molecules in the cytoplasm. **Cyclic AMP** (**cAMP**) is an important second messenger. **Review Figure 5.16**
- Activated enzymes may in turn activate other enzymes in a signal transduction pathway, leading to impressive amplification of a signal. **Review Figure 5.17 and ANIMATED TUTORIAL 5.6**
- Protein kinases covalently add phosphate groups to target proteins; cAMP binds target proteins noncovalently. Both kinds of binding change the target protein's conformation to expose or hide its active site.
- Signal transduction can be regulated in several ways. The balance between the activation and inactivation of the molecules involved determines the ultimate cellular response to a signal. **Review Figure 5.18**
- The cellular responses to signals may include the opening of ion channels, changes in gene expression, or the alteration of enzyme activities.

See **ACTIVITY 5.2** for a concept review of this chapter.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is5

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

6

Pathways that Harvest and Store Chemical Energy

KEY CONCEPTS

An Old Brew Carvings from more than 4,000 years ago in ancient Egypt show barley being crushed and mixed with water (left), then put into closed vessels (center) where airless conditions are suitable for the production of alcohol by yeast cells residing on the vessels' walls. The beer is then ready for consumption (right).

Agriculture was a key invention in the development of human civilizations. The planting and harvesting of seeds began about 10,000 years ago. One of the first plants to be turned into a reliable crop was barley, and one of the first uses of barley was to brew beer. Living in what is now Iraq, ancient Sumerians learned that partly germinated and then mashed-up barley seeds, stored under the right conditions, could produce a potent and pleasant drink. An ancient king, Hammurabi, laid down the oldest known laws regarding an alcoholic beverage: the daily beer ration was 2 liters for a normal worker, 3 liters for a civil servant, and 5 liters for a high priest. Alcoholic beverages were not just a diversion to these people; their health depended on them. Drinking water from rivers and ponds caused diseases, and whatever caused these diseases was not present in liquids containing alcohol.

Early chemists and biologists were interested in how mashed barley seeds (or grapes, in the case of wine) were transformed into alcoholic beverages. By the nineteenth century there were two theories. Chemists claimed that these transformations were simply chemical reactions, not some special property of the plant material. Biologists, armed with their microscopes and cell theory (see Chapter 4), said that the barley and grape extracts were converted to beer and wine by living cells.

The great French scientist Louis Pasteur tackled the question in the 1860s, responding to a challenge posed by a group of distillers who wanted to use sugar beets to produce alcohol. Pasteur found that (1) nothing happened to beet mash unless microscopic yeast cells were present; (2) in the presence of fresh air, yeast cells grew vigorously on the mash, and bubbles of CO_2 were formed; and (3) without fresh air, the yeast grew slowly, less CO_2 was produced, and alcohol was formed. So the biologists were right: living cells produced alcohol from ground-up, sugary extracts. Later, biochemists broke open yeast cells and unraveled the sequence of chemical transformations from sugar to alcohol. It turned out that the chemists were right too: the production of alcohol involves a series of chemical reactions that require energy transfers. The flow of energy in living systems (such as yeast cells) involves the same chemical principles as energy flow in the inanimate world.

Q Why does fresh air inhibit the formation of alcohol by yeast cells?

You will find the answer to this question on page 126.

CONCEPT 6.1 ATP and Reduced Coenzymes Play Important Roles in Biological Energy Metabolism

In Chapters 2 and 3 we introduced the general concepts of energy, enzymes, and metabolism. Energy is stored in the chemical bonds of molecules, and it can be released and transformed by the metabolic pathways of living cells. There are five general principles governing metabolic pathways:

- A complex chemical transformation occurs in a series of separate, intermediate reactions that form a metabolic pathway.
- Each reaction is catalyzed by a specific enzyme.
- Most metabolic pathways are similar in all organisms, from bacteria to plants to humans.
- In eukaryotes, many metabolic pathways are compartmentalized, with certain reactions occurring inside specific organelles.
- Each metabolic pathway is controlled by key enzymes that can be inhibited or activated, thereby determining how fast the reactions will go.

Chemical energy available to do work is termed free energy (*G*). According to the laws of thermodynamics, a biochemical reaction may change the *form* of energy but not the net *amount*. A biochemical reaction is exergonic if it releases energy from the reactants, or endergonic if energy must be added to the reactants.

LINK

You can review the principles of energy transformations in **Concept 2.5**

In the chemistry lab, energy can be released or added in the form of heat. But in cells, energy-transforming reactions are often coupled; that is, an energy-releasing (exergonic) reaction is coupled in time and location to an energy-requiring (endergonic) reaction. Two widely used coupling molecules are the coenzymes ATP and NADH.

ATP hydrolysis releases energy

Cells use adenosine triphosphate (ATP) as a kind of "energy currency." Just as it is more effective, efficient, and convenient for you to trade money for a lunch than to trade your actual labor, it is useful for cells to have a currency for transferring energy between different reactions and cell processes. Some of the energy that is released in exergonic reactions is captured in chemical bonds when ATP is formed from adenosine diphosphate (ADP) and inorganic phosphate (hydrogen phosphate; commonly abbreviated as P_i). The ATP can then be hydrolyzed at other sites in the cell, releasing free energy to drive endergonic reactions (**FIGURE 6.1**).

An active cell requires the production of millions of molecules of ATP per second to drive its biochemical machinery.

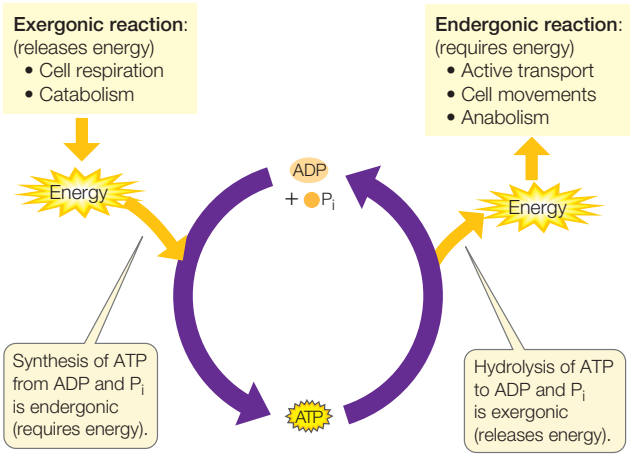


FIGURE 6.1 The Concept of Coupling Reactions Some exergonic cellular reactions are coupled with the formation of ATP from ADP and P_i (an endergonic reaction). The cell can later couple the (exergonic) hydrolysis of ATP with endergonic cellular processes.

Go to ACTIVITY 6.1 ATP and Coupled Reactions
PoL2e.com/ac6.1

You are already familiar with some of the activities in the cell that require free energy derived from the hydrolysis of ATP:

- Active transport across a membrane (Concept 5.3)
- Condensation reactions that use enzymes to form polymers (Concept 2.2)
- Motor proteins that move vesicles along microtubules (Concept 4.4)

An ATP molecule consists of the nitrogen-containing base adenine bonded to ribose (a sugar), which is attached to a sequence of three phosphate groups (**FIGURE 6.2**). The hydrolysis

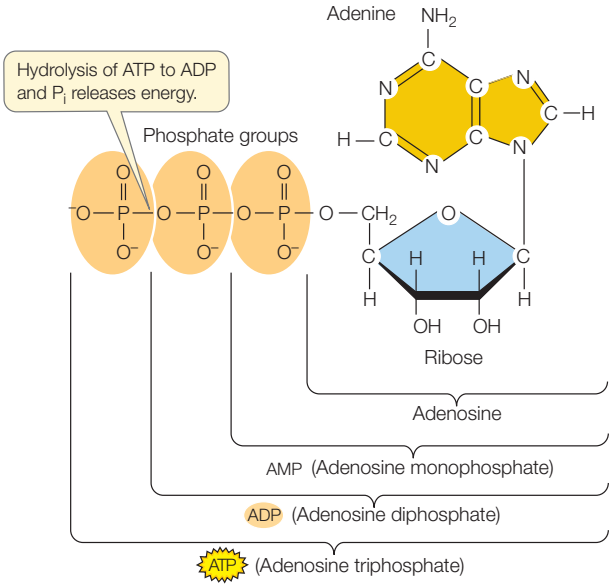


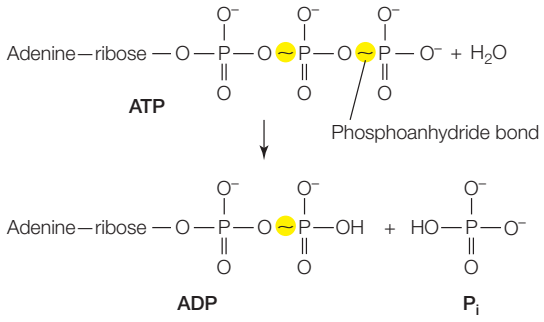
FIGURE 6.2 ATP ATP is built by the addition of terminal phosphate groups onto the nucleoside adenosine.

of a molecule of ATP yields free energy, ADP, and the inorganic phosphate ion (P_i):

$$ATP + H_2O \rightarrow ADP + P_i + \text{free energy}$$

The important property of this reaction is that it is exergonic, releasing free energy. Under standard laboratory conditions, the change in free energy for this reaction (ΔG) is about –7.3 kcal/mol (–30 kJ/mol). Recall that a negative change in free energy means that the product molecules (in this case, ADP and P_i) have less free energy than the reactants (ATP and H_2O), so the change is negative. A molecule of ATP can also be hydrolyzed to adenosine monophosphate (AMP) and a pyrophosphate ion ($P_2O_7^{4-}$; commonly abbreviated as PP_i). In this case, additional energy may be released by the subsequent conversion of PP_i to two molecules of P_i.

Energy is released as a result of ATP hydrolysis because the P—O bonds in a free hydrogen phosphate (P_i) molecule are stronger and more stable than the relatively weak P—O bonds (called phosphoanhydride bonds) between the phosphate groups in ATP. (Phosphoanhydride bonds are often denoted by wavy lines in chemical structures, as highlighted below). Recall that in general, stable bond formation is an exergonic process, whereas breaking bonds requires an input of energy. In this case, the amount of energy released by the formation of a new P—O bond in the P_i molecule is greater than the energy needed to break the phosphoanhydride bond.



In some reactions, ATP is formed by substrate-level phosphorylation—the enzyme-mediated direct transfer of phosphate from another molecule (the substrate) to ADP. This is the case for some reactions of glycolysis, as we will see in Concept 6.2. But most of the ATP in living cells is formed by oxidative phosphorylation, which we will discuss shortly.

Redox reactions transfer electrons and energy

Another way of transferring energy in chemical reactions is to transfer electrons. A reaction in which one substance transfers one or more electrons to another substance is called a reduction–oxidation reaction, or **redox** reaction.

- **Reduction** is the gain of one or more electrons by an atom, ion, or molecule.
- **Oxidation** is the loss of one or more electrons.

Oxidation and reduction *always occur together*: as one chemical is oxidized, the electrons it loses are transferred to another chemical, reducing it. Thus some molecules are called oxidizing agents and others are reducing agents:

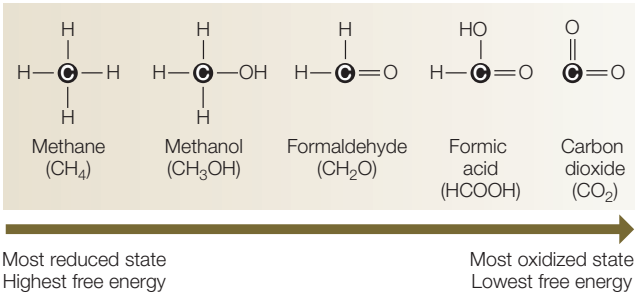
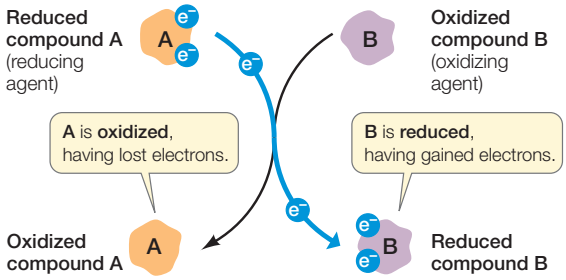


FIGURE 6.3 Oxidation, Reduction, and Energy The more oxidized a carbon atom is, the less free energy it has.



Although oxidation and reduction are defined in terms of traffic in electrons, it is often helpful to think in terms of the gain or loss of hydrogen atoms. Transfers of hydrogen atoms involve transfers of electrons ($H = H^+ + e^-$). So when a molecule loses a hydrogen atom, it becomes oxidized.

In general, the more reduced a molecule is, the more energy is stored in its covalent bonds (**FIGURE 6.3**). Indeed, highly reduced molecules can be used as energy sources; for example, methane and methanol can be burned as fuel. However, oxidized molecules such as CO_2 cannot be used as sources of energy. In a redox reaction, some energy is transferred from the reducing agent to the reduced product. Some energy remains in the reducing agent (now oxidized), and some is lost to entropy.

Cells use the coenzyme nicotinamide adenine dinucleotide (NAD) as an electron carrier in redox reactions (**FIGURE 6.4**). NAD exists in two chemically distinct forms, one oxidized (NAD^+) and the other reduced (NADH). The reduction reaction

$$NAD^+ + H^+ + 2\,e^- \rightarrow NADH$$

involves the transfer of a proton (the hydrogen ion, H^+) and two electrons, which are released by an accompanying oxidation reaction.

The reduction of NAD^+ is highly endergonic, and within the cell, the electrons do not remain with NADH. Oxygen is highly electronegative and readily accepts electrons from the reduced NADH molecule. The oxidation of NADH by O_2 (which occurs in several steps):

$$NADH + H^+ + \tfrac{1}{2}\,O_2 \rightarrow NAD^+ + H_2O$$

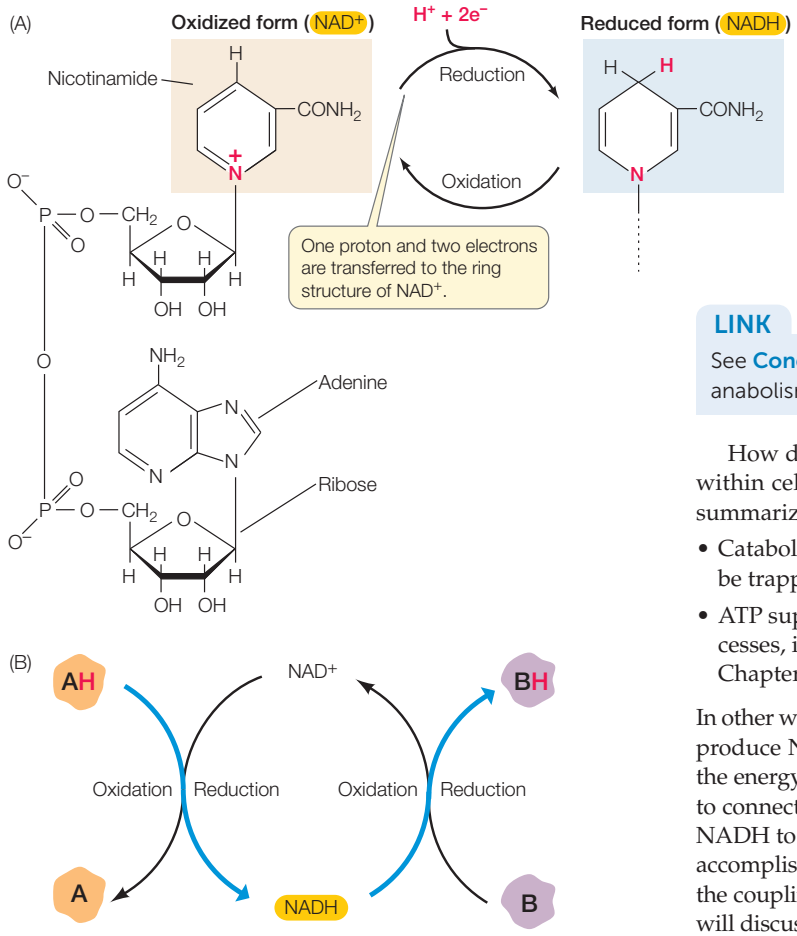


FIGURE 6.4 NAD⁺/NADH Is an Electron Carrier in Redox Reactions (A) NAD^+ is an important electron acceptor in redox reactions, and its reduced form, NADH, is an important energy intermediary in cells. The unshaded portion of the molecule (left) remains unchanged by the redox reaction. (B) Coupling of redox reactions using NAD^+/NADH.

is highly exergonic, releasing energy with a ΔG of –52.4 kcal/mol (–219 kJ/mol). Note that the oxidizing agent appears here as "$^1/_2\ O_2$" instead of "O." This notation emphasizes that it is molecular oxygen (O_2) that acts as the oxidizing agent. This is clearer if the molecules of the reaction above are doubled:

$$2\ NADH + 2\ H^+ + O_2 \rightarrow 2\ NAD^+ + 2\ H_2O$$

Because the oxidation of NADH releases more energy than the hydrolysis of ATP, NADH can be thought of as a larger package of free energy than ATP. NAD^+ is a common electron carrier in cells, but not the only one. Others include flavin adenine dinucleotide (FAD), which also transfers electrons during glucose metabolism (see Concept 6.2), and nicotinamide adenine dinucleotide phosphate ($NADP^+$), which is used in photosynthesis (see Concept 6.5).

The processes of NADH oxidation and ATP production are coupled

In order to carry out the many metabolic processes needed to sustain life, cells release and reuse the energy contained in chemical bonds. The energy-coupling coenzymes (in particular, ATP and NADH) play vital roles in the transfer of energy between cellular reactions that release energy (catabolism) and those that require energy (including anabolism).

LINK

See **Concept 2.5** to review the principles of catabolism and anabolism

How do these coenzymes participate in the flow of energy within cells? The release and reuse of cellular energy can be summarized as follows:

- Catabolism releases energy by oxidation; this energy can be trapped by the reduction of coenzymes such as NAD^+.
- ATP supplies the energy for many energy-requiring processes, including anabolism. For example, as we noted in Chapter 5, active transport requires ATP.

In other words, most of the energy-releasing reactions in the cell produce NADH (or similar reduced coenzymes), but most of the energy-consuming reactions require ATP. Cells need a way to connect the two coenzymes; that is, to transfer energy from NADH to the phosphoanhydride bond of ATP. This transfer is accomplished in a process called oxidative phosphorylation—the coupling of NADH oxidation to the production of ATP. We will discuss the mechanisms of this process in Concept 6.2.

CHECKpoint CONCEPT 6.1

✓ For each of the reactions

a. $C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O$

b. $6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2$

which reactants get oxidized and which get reduced?

✓ What kinds of coenzymes might be involved in the following reactions? Explain your answer.

a. Glucose → glucose 6-phosphate

b. Fatty acid → $CO_2 + H_2O$

✓ A typical, active young man requires 2,800 kilocalories of food energy a day to fuel metabolism, movement, active transport, etc. The energy stored in the third phosphodiester bond of ATP is 0.0145 kcal/gram.

a. If the energy from the man's food were all stored as ATP, how much ATP would be produced each day from ADP and P_i?

b. The man actually has about 50 grams of ATP. What does this mean in terms of ATP hydrolysis and synthesis?

In this concept we saw that both ATP and NADH function as energy-coupling coenzymes, which are used by cells to store

and transfer energy. We will now look at how cells capture energy from the catabolism of glucose to produce NADH, and then transfer this energy from NADH to ATP.

CONCEPT 6.2 Carbohydrate Catabolism in the Presence of Oxygen Releases a Large Amount of Energy

Cellular respiration is the set of metabolic reactions used by cells to harvest energy from food. Energy is released when reduced organic molecules, with many C—C and C—H bonds, are oxidized to CO_2. We will consider in detail only the oxidation (catabolism) of carbohydrates, but bear in mind that cells also obtain energy from the catabolism of other molecules, such as lipids.

The chemical energy released from the complete oxidation of glucose to CO_2 is considerable:

$$\text{Glucose} + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O + \text{energy (686 kcal per mole of glucose)}$$

In a chemistry lab, where sugar is "burned" in the presence of O_2, this energy is all lost as heat. In the cell, some of the released energy (234 kcal/mol; 34% of the total) is trapped as ATP. The efficiency of this energy-trapping process is impressive, even when compared with motors that humans have devised. The cell can achieve this through the general principles that govern metabolism listed in Concept 6.1. Most notably, the oxidation occurs in a series of small steps (FIGURE 6.5).

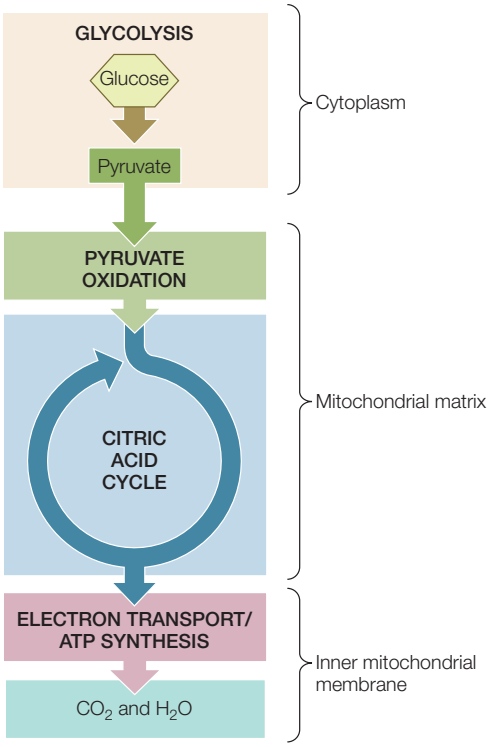


FIGURE 6.6 Energy-Releasing Metabolic Pathways The catabolism of glucose under aerobic conditions occurs in three sequential metabolic pathways: glycolysis, pyruvate oxidation, and the citric acid cycle. The reduced coenzymes are then oxidized by the respiratory chain, and ATP is made.

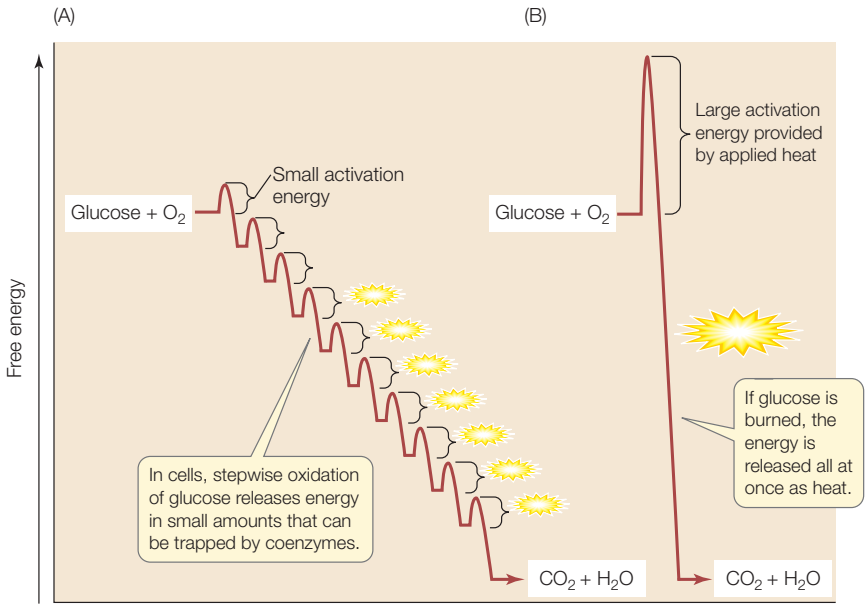


FIGURE 6.5 Energy Metabolism Occurs in Small Steps (A) In living systems, glucose is oxidized via a series of steps, releasing small amounts of energy that can be efficiently trapped by coenzymes. (B) Glucose that is burned releases its energy as heat in one big step.

In the catabolism of glucose under **aerobic** conditions (in the presence of O_2), the small steps can be grouped into three linked biochemical pathways (FIGURE 6.6):

- In **glycolysis**, the six-carbon monosaccharide glucose is converted into two three-carbon molecules of pyruvate.
- In **pyruvate oxidation**, two three-carbon molecules of pyruvate are oxidized to two two-carbon molecules of acetyl CoA and two molecules of CO_2.
- In the **citric acid cycle**, two two-carbon molecules of acetyl CoA are oxidized to four molecules of CO_2.

In glycolysis, glucose is partially oxidized and some energy is released

Glycolysis takes place in the cytosol and involves ten enzyme-catalyzed reactions. During glycolysis, some of the C—H bonds in the glucose molecule are oxidized, releasing some stored energy. The final products are two molecules of pyruvate (the anion of pyruvic acid), two molecules of ATP, and

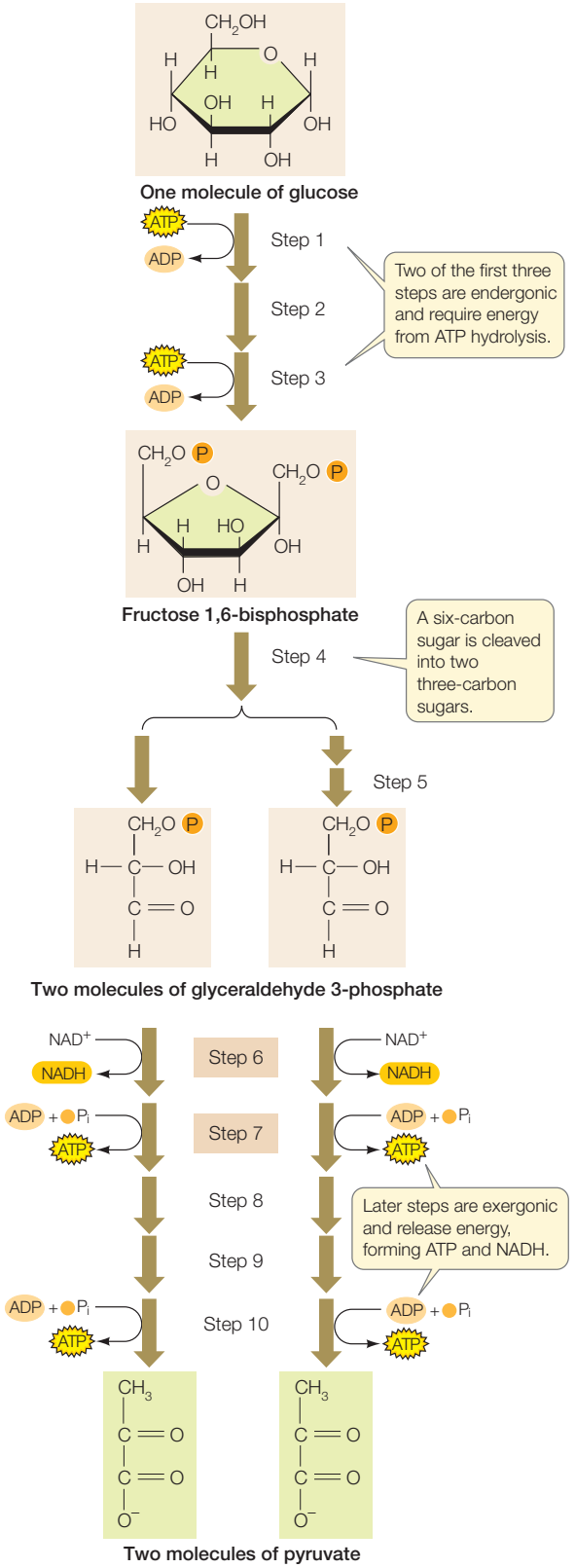
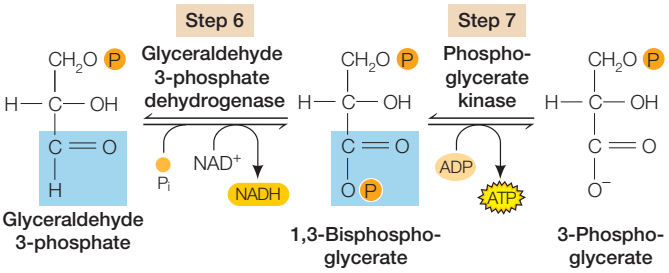


FIGURE 6.7 Glycolysis Converts Glucose into Pyruvate Glucose is converted to pyruvate in ten enzyme-catalyzed steps. Along the way, energy is released to form ATP and NADH.

two molecules of NADH. Glycolysis can be divided into two stages: the initial energy-investing reactions that consume chemical energy stored in ATP, and the energy-harvesting reactions that produce ATP and NADH (**FIGURE 6.7**).

To help you understand the process without getting into extensive detail, we will focus on two consecutive reactions in this pathway (steps 6 and 7 in Figure 6.7).



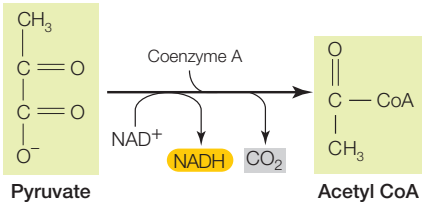
These are examples of two types of reactions that occur repeatedly in glycolysis and in many other metabolic pathways:

- *Oxidation–reduction*: In the exergonic reaction of step 6, more than 50 kcal/mol of energy are released in the oxidation of glyceraldehyde 3-phosphate. (Look at the first carbon atom, highlighted in blue, where an H is replaced by an O.) The energy is trapped via the reduction of NAD^+ to NADH.
- *Substrate-level phosphorylation*: The second reaction in this series is also exergonic, but in this case less energy is released. It is enough to transfer a phosphate from the substrate (1,3-bisphosphoglycerate) to ADP, forming ATP.

The end product of glycolysis, pyruvate, is somewhat more oxidized than glucose. In the presence of O_2, further oxidation can occur. In prokaryotes these subsequent reactions take place in the cytosol, but in eukaryotes they take place in the mitochondrial matrix.

Pyruvate oxidation links glycolysis and the citric acid cycle

The next step in the aerobic catabolism of glucose involves the oxidation of pyruvate to a two-carbon acetate molecule and CO_2. The acetate is then bound to **coenzyme A** (**CoA**), which is used in various biochemical reactions as a carrier of acetyl groups:



This is the link between glycolysis and further oxidative reactions (see Figure 6.6).

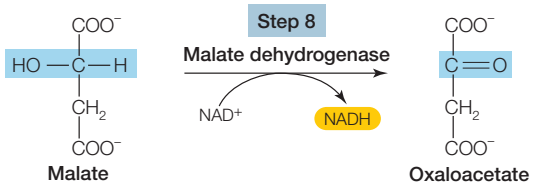
The formation of acetyl CoA is a multistep reaction catalyzed by the pyruvate dehydrogenase complex, which contains 60 individual proteins and 5 different coenzymes. The overall reaction is exergonic, and one molecule of NAD^+ is reduced.

The main role of acetyl CoA is to donate its acetyl group to the four-carbon compound oxaloacetate, forming the six-carbon molecule citrate (the anion of citric acid). This initiates the citric acid cycle, one of life's most important energy-harvesting pathways.

The citric acid cycle completes the oxidation of glucose to CO_2

Acetyl CoA is the starting point for the citric acid cycle. This pathway of eight reactions completely oxidizes the two-carbon acetyl group to two molecules of CO_2. The free energy released from these reactions is captured by ADP and the electron carriers NAD^+ and FAD (FIGURE 6.8). This is a cycle because the starting material, oxaloacetate, is regenerated in the last step and is ready to accept another acetate group from acetyl CoA. The citric acid cycle operates twice for each glucose molecule that enters glycolysis (once for each pyruvate that enters the mitochondrion).

Let's focus on the final reaction of the cycle (step 8 in Figure 6.8) as an example of the kind of reaction that occurs:



This oxidation reaction (see the blue-highlighted carbon atom) is exergonic, and the released energy is trapped by NAD^+, forming NADH. With four such reactions ($FADH_2$ is a reduced coenzyme similar to NADH), the citric acid cycle harvests a great deal of chemical energy from the oxidation of acetyl CoA.

Energy is transferred from NADH to ATP by oxidative phosphorylation

As we mentioned in Concept 6.1, energy-consuming processes in the cell use ATP as their source of energy. In order to fully use the energy harvested in catabolism, cells need to transfer energy from NADH (and $FADH_2$) to the phosphoanhydride bond of ATP. In eukaryotic mitochondria, this transfer is accomplished by **oxidative phosphorylation**: NADH oxidation is used to actively transport protons (H^+ ions) across the inner mitochondrial membrane, resulting in a proton gradient across the membrane. The diffusion of protons back across the membrane is then used to drive the synthesis of ATP. (In prokaryotes, oxidative phosphorylation takes place at the cell membrane.)

First let's examine how the oxidation of NADH and $FADH_2$ leads to the production of the proton gradient. For example, when NADH is reoxidized to NAD^+, O_2 is reduced to H_2O:

$$NADH + H^+ + \tfrac{1}{2}\,O_2 \rightarrow NAD^+ + H_2O$$

This does not happen in a single step. Rather, there is a series of redox electron carrier proteins called the **respiratory chain** embedded in the inner membrane of the mitochondrion (FIGURE 6.9). The electrons from the oxidation of NADH and $FADH_2$ pass from one carrier to the next in the chain in a process called **electron transport**. The oxidation reactions are exergonic, and they release energy that is used to actively transport H^+ ions across the membrane.

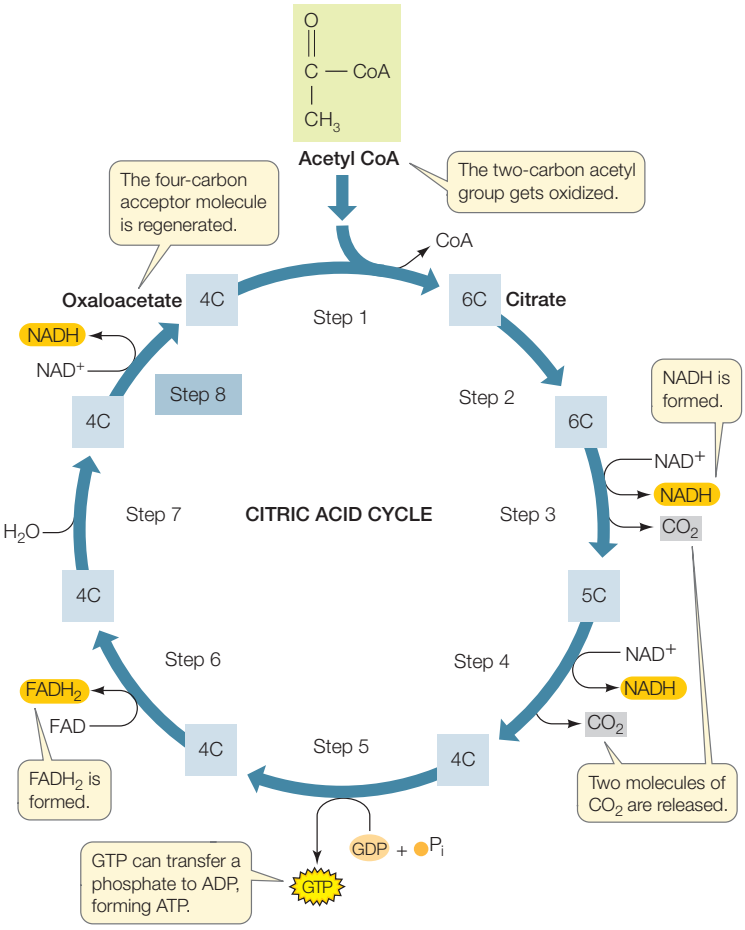


FIGURE 6.8 The Citric Acid Cycle Also called the Krebs cycle for its discoverer, Hans Krebs, the citric acid cycle involves eight steps and fully oxidizes acetyl CoA to CO_2.

Go to ACTIVITY 6.2 The Citric Acid Cycle
PoL2e.com/ac6.2

An important aspect of this process is that an oxidation reaction is always coupled with a reduction. When NADH is oxidized to NAD^+, the corresponding reduction reaction is the formation of water from O_2:

$$2\,H^+ + 2\,e^- + \tfrac{1}{2}\,O_2 \rightarrow H_2O$$

So the *key role of O_2 in cells*—the reason we breathe and have a blood system to deliver O_2 to tissues—*is to act as an electron acceptor and become reduced*.

Chemiosmosis uses the proton gradient to generate ATP

In addition to the electron transport carriers, the inner mitochondrial membrane contains an enzyme called **ATP synthase** (FIGURE 6.10A). This enzyme uses the H^+ gradient to drive the

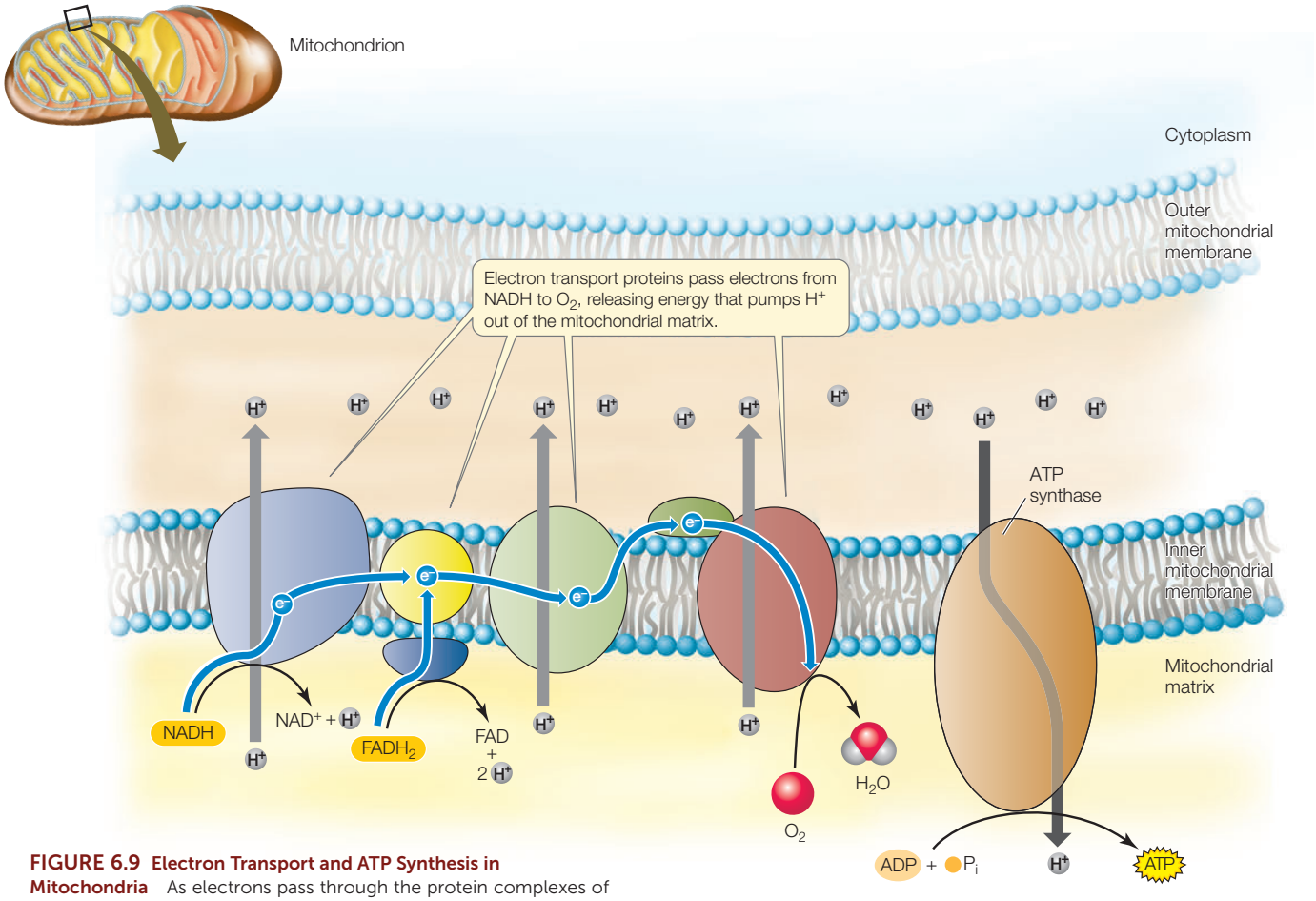


FIGURE 6.9 Electron Transport and ATP Synthesis in Mitochondria As electrons pass through the protein complexes of the respiratory chain, protons are pumped from the mitochondrial matrix into the intermembrane space. As the protons return to the matrix through ATP synthase, ATP is formed.

Go to ANIMATED TUTORIAL 6.1
Electron Transport and ATP Synthesis
PoL2e.com/at6.1

Go to ACTIVITY 6.3 Respiratory Chain
PoL2e.com/ac6.3

synthesis of ATP via a mechanism called **chemiosmosis**—the movement of ions across a semipermeable barrier from a region of higher concentration to a region of lower concentration. Chemiosmosis relies on concepts covered in earlier chapters:

- If the concentration of a substance is greater on one side of a membrane than the other, the substance will tend to diffuse across the membrane to its region of lower concentration (see Concept 5.2).
- If a membrane blocks this diffusion, the substance at the higher concentration has potential energy, which can be converted to other forms of energy (see Concept 2.5).
- Because the interior of a membrane is nonpolar, protons (H^+) cannot readily diffuse across the membrane, but they can cross the membrane through the ATP synthase enzyme. ATP synthase converts the potential energy of the proton gradient (called the proton motive force) into the chemical energy in ATP.

ATP synthase is a molecular motor composed of two parts: the F_o unit, which is a transmembrane domain that functions as the H^+ channel; and the F_1 unit, which contains the active sites for ATP synthesis (**FIGURE 6.10B**). The F_1 unit consists of six subunits (three each of two polypeptide chains), arranged like the segments of an orange around a central polypeptide. The potential energy set up by the proton gradient drives the passage of protons through the ring of polypeptides that make up the F_o component. This ring rotates as the protons pass through the membrane, causing part of the F_1 unit to rotate as well. ADP and P_i bind to active sites that become exposed on the F_1 unit as it rotates, and ATP is made.

Go to MEDIA CLIP 6.1
ATP Synthase in Motion
PoL2e.com/mc6.1

The structure and function of ATP synthase enzymes are shared by living organisms as diverse as bacteria and humans. These enzymes make ATP at rates of up to 100 molecules per second. In all organisms, these molecular motors rely on protein gradients across membranes:

- In prokaryotes, the gradient is set up across the cell membrane, using energy from various sources.
- In eukaryotes, chemiosmosis occurs in the mitochondria and the chloroplasts.

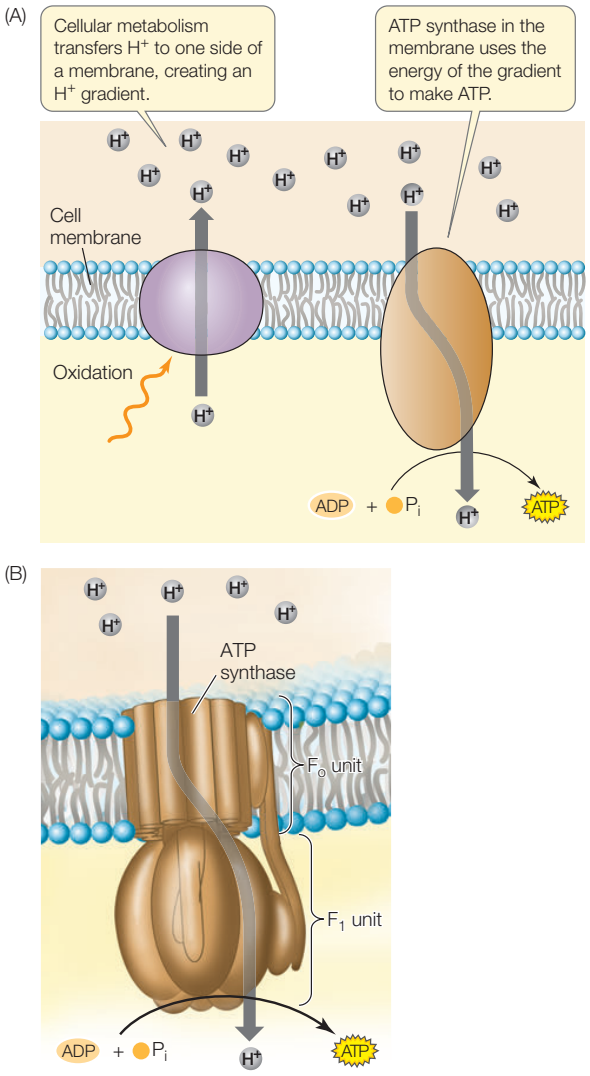


FIGURE 6.10 Chemiosmosis (A) If a cell can generate a proton (H^+) gradient across a membrane, the potential energy resulting from the concentration gradient can be used by a membrane-spanning enzyme to make ATP. (B) ATP synthase has a membrane-embedded channel for H^+ diffusion and a motor that turns, releasing some energy to produce ATP.

- As we have just seen, the H^+ gradient in mitochondria is set up across the inner mitochondrial membrane, using energy released by the oxidation of NADH and $FADH_2$.
- In chloroplasts, the H^+ gradient is set up across the thylakoid membrane using energy from light (see Concept 6.5). In this case, the reduced molecule is $NADP^+$, a relative of NAD^+.

Despite these differences in detail, the mechanism of chemiosmosis is similar in almost all forms of life.

Chemiosmosis can be demonstrated experimentally

If chloroplasts or mitochondria are isolated from cells and put in a test tube, a proton gradient can be introduced artificially. This artificial gradient drives ATP synthesis (**FIGURE 6.11**), but only if ATP synthase, ADP, inorganic phosphate, and the membrane are present.

What happens if the H^+ gradient is destroyed by the presence of a membrane channel that is always open to protons? ATP cannot be made, but the oxidation of NADH still occurs and O_2 is reduced, releasing considerable energy. The released energy forms heat instead of being used to make ATP. In newborn human infants, a membrane protein appropriately called uncoupling protein 1 disrupts the H^+ gradient in fat cell mitochondria, and this results in the release of heat. Because infants lack body hair, this process helps keep them warm.

A popular weight loss drug in the 1930s was the synthetic uncoupler molecule dinitrophenol. There were claims of dramatic weight loss when the drug was administered to obese patients. Unfortunately, the heat that was released caused fatally high fevers, and the effective dose and fatal dose were quite close. The use of this drug was discontinued in 1938, but the general strategy of using an uncoupler for weight loss remains a subject of research.

Oxidative phosphorylation and chemiosmosis yield a lot of ATP

For each NADH (or $FADH_2$) that begins the respiratory chain, two to three (let's say 2.5) ATP molecules are formed under the conditions in the cell. Thus the four molecules of reduced coenzyme produced by each turn of the citric acid cycle yield about ten (4 × 2.5) molecules of ATP. Two molecules of acetyl CoA are produced from each glucose, so the total is about 20 ATPs per molecule of glucose. Add to this the NADH produced by glycolysis and pyruvate oxidation, and the ATP formed by substrate-level phosphorylation during glycolysis and the citric acid cycle, and the total is about *32 molecules of ATP produced per fully oxidized glucose.*

The vital role of O_2 is now clear: most of the ATP produced in cellular respiration is formed by oxidative phosphorylation—the process of transferring electrons from NADH to O_2, resulting in the reoxidation of NADH to NAD^+. The accumulation of atmospheric O_2 as a result of photosynthesis by ancient microorganisms (see Concept 18.2) set the stage for the evolution of oxidative phosphorylation; organisms that could exploit the O_2 would have had a selective advantage.

Nevertheless, many microorganisms still thrive where O_2 is scarce. These anaerobic bacteria and archaea use alternative electron acceptors in their natural environments. For instance, the bacterium *Geobacter metallireducens* typically lives in sediments under streams or ponds, and uses metal ions as terminal electron acceptors. For example:

$$Fe^{3+} + e^- \text{ (from electron transport)} \rightarrow Fe^{2+}$$

This bacterium can also use radioactive uranium ions as electron acceptors. In the process, the uranium is converted from a soluble to an insoluble form, making *Geobacter* of potential use in environmental cleanup. The bacterium can convert uranium in contaminated water into a form that accumulates in the sediment instead and can be more readily removed.

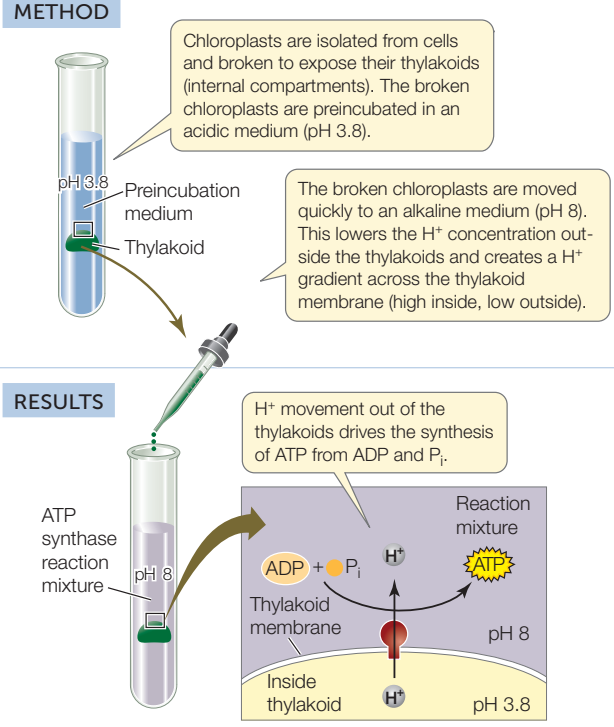
INVESTIGATION

FIGURE 6.11 An Experiment Demonstrates the Chemiosmotic Mechanism The chemiosmosis hypothesis was a bold departure from the conventional scientific thinking of the time. It required an intact compartment separated by a membrane. Could a proton gradient drive the synthesis of ATP? The first experiments to answer this question used chloroplasts, plant organelles that use the same mechanism as mitochondria to synthesize ATP.[a]

HYPOTHESIS

A H^+ gradient across a membrane that contains ATP synthase is sufficient to drive ATP synthesis.

METHOD



RESULTS

CONCLUSION

A H^+ gradient across an ATP synthase–containing membrane is sufficient for ATP synthesis by organelles.

ANALYZE THE DATA

The formation of ATP from ADP and P_i was measured using luciferase, which catalyzes the formation of a luminescent (light-emitting) molecule if ATP is present. The experiment was performed under different conditions, with the following results:

Experiment	Preincubation pH	ATP synthase mixture (pH 8)	ATP formation (nmoles/mg chlorophyll)
1	3.8	Complete mixture	144
2	7.0	Complete mixture	12
3	3.8	P_i omitted	12
4	3.8	ADP omitted	4
5	3.8	Thylakoids omitted	7

A. Which experiments show that a proton gradient is necessary to stimulate ATP formation?

B. Why was there less ATP production in the absence of P_i?

Go to ANIMATED TUTORIAL 6.2
Two Experiments Demonstrate the Chemiosmotic Mechanism
PoL2e.com/at6.2

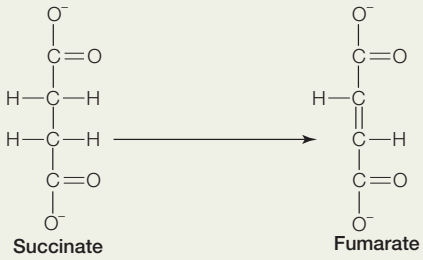
Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]A. T. Jagendorf and E. Uribe. 1966. *Proceedings of the National Academy of Sciences USA* 55: 170–177.

APPLY THE CONCEPT

Carbohydrate catabolism in the presence of oxygen releases a large amount of energy

The following reaction occurs in the citric acid cycle:



Answer each of the following questions, and explain your answers:

1. Is this reaction an oxidation or reduction?
2. Is the reaction exergonic or endergonic?
3. This reaction requires a coenzyme. What kind of coenzyme?
4. What happens to the fumarate after the reaction is completed?
5. What happens to the coenzyme after the reaction is completed?

CHECKpoint CONCEPT 6.2

- ✓ Draw the molecules of glucose and pyruvate and compare them with respect to their oxidation/reduction states. What does this mean in terms of available chemical bond energy?
- ✓ Compare the energy yields (in terms of ATP per glucose) from glycolysis and from the combination of pyruvate oxidation and the citric acid cycle.
- ✓ High levels of ammonia (NH_3) are toxic to mammals. One reason is that ammonia bonds with α-ketoglutarate (the substrate for step 4 in Figure 6.8), forming the amino acid glutamate. This removes α-ketoglutarate from the citric acid cycle. Explain the consequence of this for the cell.
- ✓ Hibernating animals usually have low rates of metabolism. Periodically, they have episodes of arousal, where they rapidly increase body temperature. During arousal, they synthesize a membrane protein not normally present in their cells. What might that protein do?

We have seen that a large amount of energy is released when carbohydrates are catabolized in the presence of O_2. If O_2 is absent, the yield of ATP is much lower. We will now turn to this situation.

CONCEPT 6.3 Carbohydrate Catabolism in the Absence of Oxygen Releases a Small Amount of Energy

In the absence of O_2—that is, when conditions are **anaerobic**—the respiratory chain cannot operate. (The exceptions, as we noted earlier, are the respiratory chains of anaerobic microbes adapted to use terminal electron acceptors other than oxygen.) Without an alternative, the NADH produced by glycolysis would not be reoxidized and glycolysis would stop, because there would be no NAD^+ for step 6 of glycolysis (see Figure 6.7). To solve this problem, organisms use **fermentation** to reoxidize the NADH, thus allowing glycolysis to continue (FIGURE 6.12).

Like glycolysis, fermentation pathways occur in the cytoplasm. There are many different types of fermentation used by different organisms, but all operate to regenerate NAD^+. The consequence of this is that the NADH made during glycolysis is not available for reoxidation by the respiratory chain to form ATP. Therefore the overall yield of ATP from fermentation is restricted to the ATP made in glycolysis (two ATP per glucose).

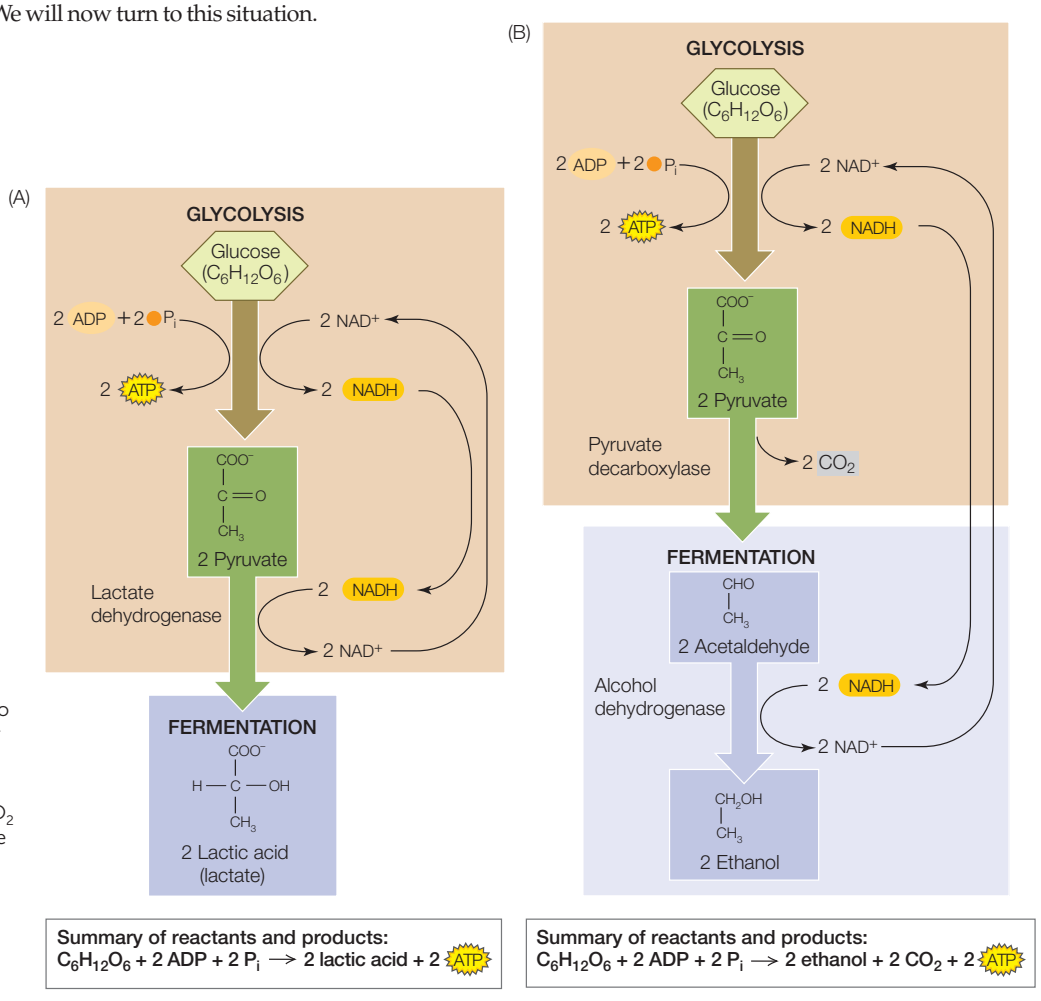


FIGURE 6.12 Fermentation
(A) In lactic acid fermentation, NADH is used to reduce pyruvate to lactic acid, thus regenerating NAD^+ to keep glycolysis operating. (B) In alcoholic fermentation, pyruvate is converted to acetaldehyde, and CO_2 is released. NADH is used to reduce acetaldehyde to ethanol, again regenerating NAD^+ for glycolysis.

Go to ACTIVITY 6.4
Glycolysis and Fermentation
PoL2e.com/ac6.4

Two fermentation pathways are found in a wide variety of organisms:

- Lactic acid fermentation, whose end product is lactic acid (lactate)
- Alcoholic fermentation, whose end product is ethyl alcohol (ethanol)

In **lactic acid fermentation**, pyruvate serves as the electron acceptor and lactate is the product (see Figure 6.12A). This process takes place in many microorganisms and complex organisms, including more complex plants and vertebrates. A notable example of lactic acid fermentation occurs in vertebrate muscle tissue. Usually, vertebrates get their energy for muscle contractions aerobically, with the circulatory system supplying O_2 to muscles. This is almost always adequate for small vertebrates, which explains why birds can fly long distances without resting. But in larger vertebrates such as humans, the circulatory system is not up to the task of delivering enough O_2 when the need is great, such as during a long sprint. At this point, the muscle cells break down glycogen (a stored polysaccharide) and undergo lactic acid fermentation. The process is reversible; lactate is converted back to pyruvate once O_2 is available again.

Alcoholic fermentation takes place in certain yeasts (eukaryotic microbes) and some plant cells under anaerobic conditions. In this process, pyruvate is converted to ethanol (see Figure 6.12B). We saw these reactions (as did Pasteur) in the opening story of this chapter. As with lactic acid fermentation, the reactions are essentially reversible.

By recycling NAD^+, fermentation allows glycolysis to continue, thus producing small amounts of ATP through substrate-level phosphorylation. The net yield of two ATPs per glucose molecule is much lower than the energy yield from oxidative phosphorylation. For this reason, most organisms that rely on fermentation instead of respiration are small microbes that grow relatively slowly.

Although its yield of ATP per glucose is generally low, in some circumstances cellular anaerobic metabolism can produce an adequate supply of ATP if the enzymatic reactions in the pathway are speeded up. Indeed, this occurs in vertebrate muscle cells (see Concept 33.3) and in some cancer cells where O_2 is in low supply.

Go to ACTIVITY 6.5 Energy Levels
PoL2e.com/ac6.5

CHECKpoint CONCEPT 6.3

- ✓ Why is replenishing NAD^+ crucial to cellular metabolism?
- ✓ Compare the sources and total energy yield in terms of ATP per glucose in human cells in the presence versus the absence of O_2.
- ✓ Conditions can become anaerobic in a heart muscle cell during a heart attack, because of the inadequate supply of blood. If O_2 is restored, what will happen to the lactate produced by the heart muscle?

You have seen how cells harvest chemical energy in cellular respiration. Now we will see how that energy moves through other metabolic pathways in the cell.

CONCEPT 6.4 Catabolic and Anabolic Pathways Are Integrated

Metabolic transformations are a hallmark of life. The pathways we have seen thus far in the chapter, including glycolysis and the citric acid cycle, do not operate in isolation. Rather, there is an interchange of molecules into and out of these pathways, to and from the metabolic pathways for the synthesis and breakdown of amino acids, nucleotides, fatty acids, and other building blocks of life. Carbon skeletons (a term describing molecules with covalently linked carbon atoms) can enter catabolic pathways and be oxidized to release their energy, or they can enter anabolic pathways to be used in the formation of the macromolecules that are the major constituents of the cell. These relationships are summarized in **FIGURE 6.13** and comprise a metabolic system (see Figure 3.17). The inputs and outputs of this system can be described, and predictions can be made regarding what would happen if the concentration of one component changes.

Catabolism and anabolism are linked

A hamburger or veggie burger on a bun contains three major sources of carbon skeletons: carbohydrates, mostly in the form of starch (a polysaccharide); lipids, mostly as triglycerides (three fatty acids attached to glycerol); and proteins (polymers of amino acids). Look at Figure 6.13 to see how each of these three types of macromolecules can be hydrolyzed and used in catabolism or anabolism.

CATABOLIC INTERCONVERSIONS Polysaccharides, lipids, and proteins can all be broken down to provide energy:

- *Polysaccharides* can be hydrolyzed to glucose. Glucose then passes through glycolysis, pyruvate oxidation, and the respiratory chain, where its energy is captured in ATP.
- *Lipids* are broken down into their constituents—glycerol and fatty acids. Glycerol is converted into dihydroxyacetone phosphate, an intermediate in glycolysis. Fatty acids are highly reduced molecules that are converted to acetyl CoA in a process called β-oxidation, catalyzed by a series of oxidation enzymes inside the mitochondrion. For example, the β-oxidation of a C_{16} (16-carbon) fatty acid occurs in several steps:

$$C_{16}\text{ fatty acid} + \text{CoA} \rightarrow C_{16}\text{ fatty acyl CoA}$$

$$C_{16}\text{ fatty acyl CoA} + \text{CoA} \rightarrow C_{14}\text{ fatty acyl CoA} + \text{acetyl CoA}$$

$$(\text{Repeat 6 times}) \rightarrow 8\text{ acetyl CoA}$$

  The acetyl CoA can then enter the citric acid cycle and be catabolized to CO_2.
- *Proteins* are hydrolyzed to their amino acid building blocks. The 20 different amino acids feed into glycolysis

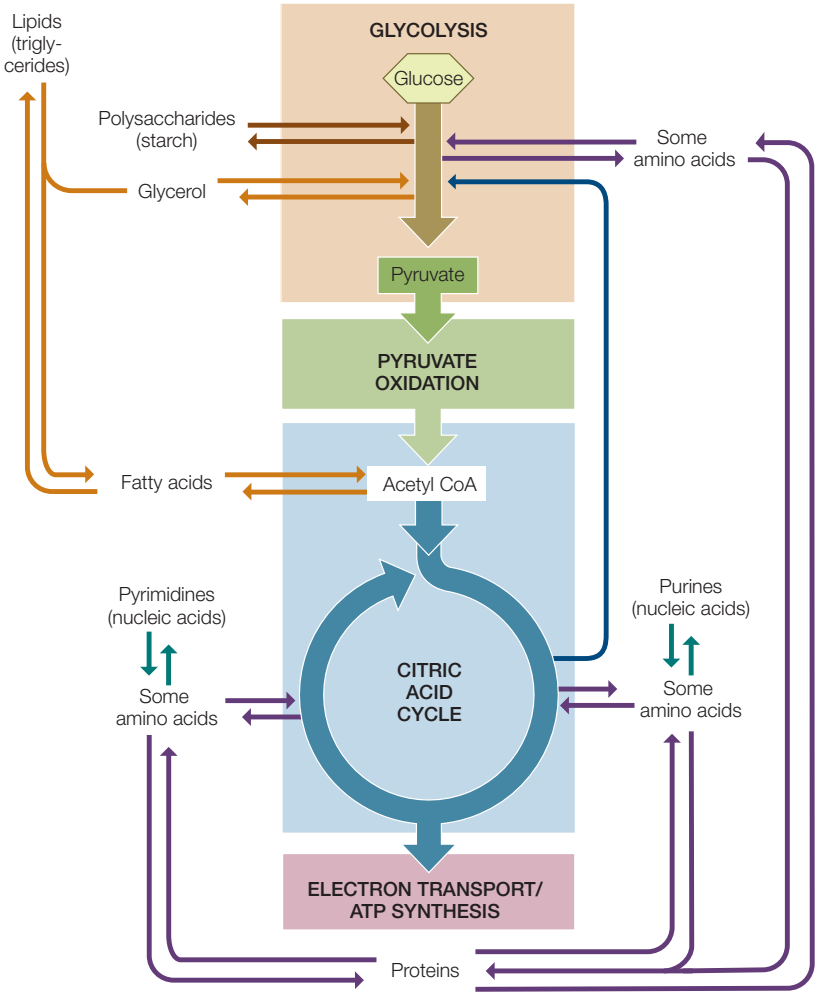


FIGURE 6.13 Relationships among the Major Metabolic Pathways of the Cell Note the central positions of glycolysis and the citric acid cycle in this system of metabolic pathways. Also note that many of the pathways can operate essentially in reverse.

or the citric acid cycle at different points. For example, the amino acid glutamate is converted into α-ketoglutarate, an intermediate in the citric acid cycle.

ANABOLIC INTERCONVERSIONS Many catabolic pathways can operate essentially in reverse, with some modifications. Glycolytic and citric acid cycle intermediates, instead of being oxidized to form CO_2, can be reduced and used to form glucose in a process called **gluconeogenesis** (which means "new formation of glucose"). Likewise, acetyl CoA can be used to form fatty acids. The most common fatty acids have even numbers of carbons: 14, 16, or 18. These are formed by the addition of two-carbon acetyl CoA "units" one at a time until the appropriate chain length is reached.

Some intermediates in the citric acid cycle are reactants in pathways that synthesize important components of nucleic acids. For example, α-ketoglutarate and oxaloacetate are starting points for purines and pyrimidines, respectively.

Catabolism and anabolism are integrated into a system

A carbon atom from a protein in your burger can end up in DNA, fat, or CO_2, among other fates. How does the organism "decide" which metabolic pathways to follow, in which cells? With hundreds of enzymes and all the possible interconversions, you might expect that the cellular concentrations of various biochemical molecules would fluctuate widely. Remarkably, the levels of these substances in what is called the metabolic pool—the sum total of all the biochemical molecules in a cell—are usually quite constant. Metabolic changes in the cell are a bit like the changes in traffic patterns in a city: if an accident blocks traffic on a major road, drivers take alternate routes, where the traffic volume consequently changes.

Consider what happens to the starch in your burger bun. In the digestive system, starch is hydrolyzed to glucose, which enters the blood. If it is needed, the glucose is distributed to the rest of the body. But if there is already enough glucose in the blood to supply the body's needs, the excess glucose is converted into glycogen and stored in the liver. If not enough glucose is supplied by food, glycogen is broken down, or other molecules are used to make glucose by gluconeogenesis. The end result is that the level of glucose in the blood is remarkably constant. How does the body accomplish this?

Metabolic enzymes (including those of glycolysis, the citric acid cycle, and the respiratory chain) are subject to regulation, and often the regulatory mechanisms involve allosteric effects. An example is feedback inhibition, illustrated in Figure 3.21. In a metabolic pathway, a high concentration of the final product can inhibit the enzyme that catalyzes the commitment step. For example, the product of glucose catabolism, ATP, feedback inhibits key enzymes in both glycolysis and the citric acid cycle. Sometimes feedback involves the product of one pathway speeding up reactions in another pathway. Feedback regulation generally occurs rapidly, affecting a pathway within minutes.

LINK

Review the discussion of enzyme regulation in **Concept 3.4**

The rate of a biochemical reaction can also be controlled by reducing or increasing the number of enzyme molecules present relative to substrate (see Figure 11.6). This can be done by altering the transcription of the genes that encode the enzymes. These events take time, and typically the effects on metabolism will take days to appear. For example, excess levels of glucose and other dietary factors can lead to increased transcription of the

gene for fatty acid synthase, a key enzyme in the synthesis of fatty acids. Excess citrate produced by the citric acid cycle is broken down to acetyl CoA, which in turn is used in fatty acid synthesis. This is one reason why people accumulate fat after eating too much. The fatty acids may be catabolized later to produce more acetyl CoA.

LINK

The regulation of gene transcription is described in **Chapter 11**

ATP and reduced coenzymes link catabolism, anabolism, and photosynthesis

Thus far in this chapter we have discussed the major catabolic pathway called cellular respiration, in which a reduced molecule, the carbohydrate glucose, is oxidized, often all the way to CO_2:

$$\text{Glucose} + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O + \text{chemical energy}$$

We have also seen how the energy derived from cellular respiration is transferred between catabolic and anabolic pathways (depending on a cell's needs), and that the coenzymes ATP and NADH play key roles in these energy transfers.

But where does the energy for all these processes come from? As we will see in the next two concepts, almost all living organisms on Earth ultimately derive their energy from the sun (exceptions include organisms that derive energy from geothermal vents deep within oceans). In addition to cellular respiration, green plants, algae, some protists, and some prokaryotes also carry out photosynthesis—a major anabolic pathway that converts light energy into chemical energy in the form of carbohydrates:

$$CO_2 + H_2O + \text{light energy} \rightarrow O_2 + \text{carbohydrates}$$

Just as catabolism and anabolism are linked, cellular respiration and photosynthesis are also linked—not only by their reactants and products (O_2, CO_2, and carbohydrates), but also by the energy "currency" of ATP and reduced coenzymes (**FIGURE 6.14**). The pathways of cellular respiration and photosynthesis can occur in the same cell. For example, cells in green plants often carry out both photosynthesis and cellular respiration simultaneously.

More often, however, cells do not carry out photosynthesis for themselves. Even in a green plant, a root cell is not photosynthetic and relies on carbohydrates transported from the leaf to carry out cellular respiration. Humans do not carry out photosynthesis anywhere in their bodies; they rely on carbohydrates obtained in their diet (ultimately derived from photosynthesis) to carry out cellular respiration, which provides the chemical energy for their bodies' activities, such as active transport and anabolism.

Go to ACTIVITY 6.6 Regulation of Energy Pathways
PoL2e.com/ac6.6

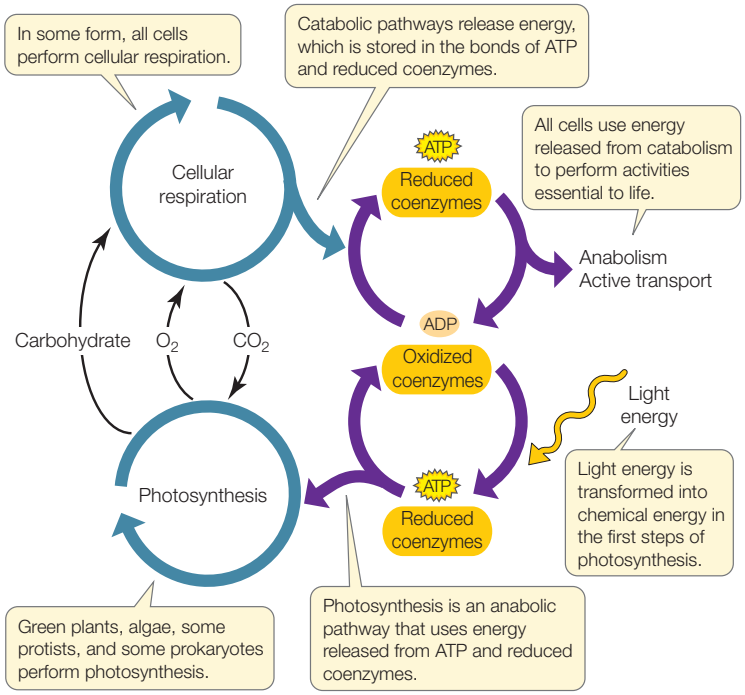


FIGURE 6.14 ATP, Reduced Coenzymes, and Metabolism The major pathways of energy metabolism, cellular respiration and photosynthesis, are related by their use of energy-transferring substrates, ATP and reduced coenzymes. Note that the net result of these pathways is to convert light energy into chemical energy to fuel the processes of life.

CHECKpoint CONCEPT 6.4

- ✓ Give examples of the catabolic conversion of a lipid and the anabolic conversion of a protein.
- ✓ Trace the biochemical pathway by which a carbon atom from a starch molecule in rice eaten today can end up in a muscle protein tomorrow.
- ✓ Describe what might happen if there were no mechanisms for modulating the level of acetyl CoA.

We have seen how cellular respiration allows organisms to harvest chemical energy from organic molecules. For the rest of the chapter, we'll look at how plants and other photosynthetic organisms produce these organic molecules using energy from light.

CONCEPT 6.5 During Photosynthesis, Light Energy Is Converted to Chemical Energy

The energy released by catabolic pathways in almost all organisms, including animals, plants, and prokaryotes, ultimately comes from the sun. **Photosynthesis** (literally, "synthesis from light") is an anabolic process by which the energy of sunlight is captured and used to convert carbon dioxide (CO_2) and water

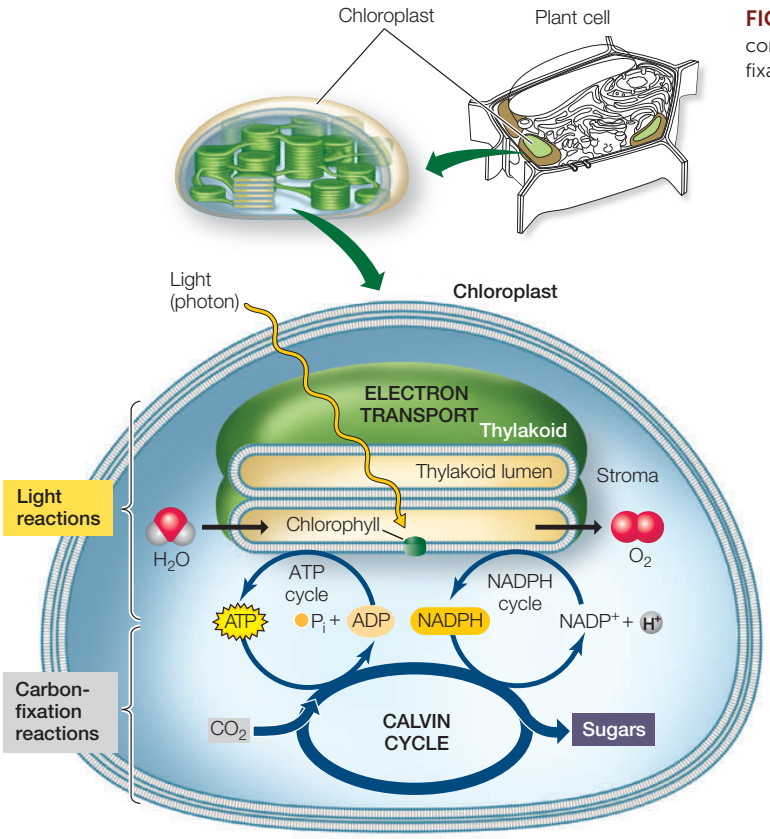


FIGURE 6.15 An Overview of Photosynthesis Photosynthesis consists of two pathways: the light reactions and the carbon-fixation reactions. In eukaryotes, these occur in the chloroplast.

(H_2O) into carbohydrates (which we represent as a six-carbon sugar, $C_6H_{12}O_6$) and oxygen gas (O_2):

$$6\,CO_2 + 6\,H_2O \rightarrow C_6H_{12}O_6 + 6\,O_2$$

This equation shows a highly endergonic reaction. The net outcome is the reverse of the general equation for glucose catabolism that we discussed in Concept 6.2. Many of the molecular processes of photosynthesis are similar to those for glucose catabolism. For example, both processes involve redox reactions, electron transport, and chemiosmosis. However, the details of photosynthesis are quite different.

Photosynthesis involves two pathways (**FIGURE 6.15**):

- The **light reactions** convert light energy into chemical energy in the form of ATP and the reduced electron carrier NADPH. This molecule is similar to NADH (see Figure 6.4A) but with an additional phosphate group attached to the sugar of its adenosine.
- The **carbon-fixation reactions** do not use light directly, but instead use the ATP and NADPH made by the light reactions, along with CO_2, to produce carbohydrates.

Both the light reactions and the carbon-fixation reactions stop in the dark because ATP synthesis and $NADP^+$ reduction require light. In photosynthetic prokaryotes (e.g., cyanobacteria), the light reactions take place on internal membranes and the carbon-fixation reactions occur in the cytosol. In plants, which will be our focus here, both pathways proceed within the chloroplast, but they occur in different parts of that organelle (see Figure 6.15).

Light energy is absorbed by chlorophyll and other pigments

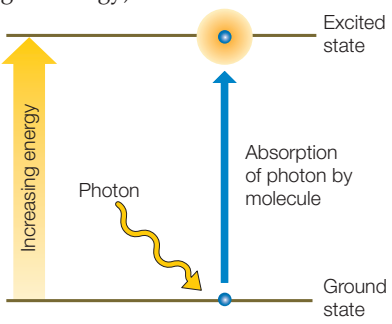
Light is a form of energy that can be converted to other forms, such as heat or chemical energy. It is helpful here to discuss light in terms of its photochemistry and photobiology.

PHOTOCHEMISTRY Light is a form of **electromagnetic radiation**. Electromagnetic radiation is propagated in waves, and the amount of energy in the radiation is inversely proportional to its **wavelength**—the shorter the wavelength, the greater the energy. The visible portion of the electromagnetic spectrum (**FIGURE 6.16**) encompasses a wide range of wavelengths and energy levels. In addition to traveling in waves, light also behaves as particles, packets of light energy called **photons**, which have no mass. In plants and other photosynthetic organisms, receptive molecules absorb photons in order to harvest their energy for biological processes. These receptive molecules absorb only specific wavelengths of light—photons with specific amounts of energy.

When a photon meets a molecule, one of three things can happen:

- The photon may bounce off the molecule—it may be *scattered* or *reflected*.
- The photon may pass through the molecule—it may be *transmitted*.
- The photon may be *absorbed* by the molecule, adding energy to the molecule.

Neither of the first two outcomes causes any change in the molecule. However, in the case of absorption, the photon disappears and its energy is absorbed by the molecule. The photon's *energy* cannot disappear, because according to the first law of thermodynamics, energy is neither created nor destroyed. When the molecule acquires the energy of the photon, it is raised from a ground state (with lower energy) to an excited state (with higher energy):



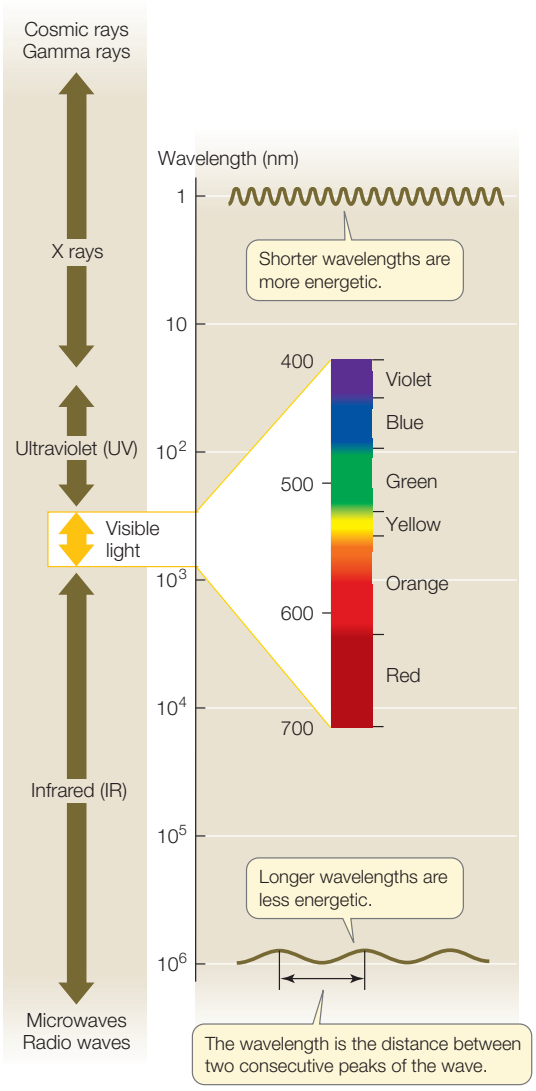


FIGURE 6.16 The Electromagnetic Spectrum The portion of the electromagnetic spectrum that is visible to humans as light is shown in detail at the right.

The difference in free energy between the molecule's excited state and its ground state is approximately equal to the free energy of the absorbed photon (a small amount of energy is lost to entropy). The increase in energy boosts one of the electrons within the molecule into a shell farther from its nucleus; this electron is now held less firmly, making the molecule unstable and more chemically reactive.

PHOTOBIOLOGY Each type of molecule absorbs light at specific, characteristic wavelengths. Molecules that absorb wavelengths in the visible spectrum are called **pigments**.

When a beam of white light (containing all the wavelengths of visible light) falls on a pigment, certain wavelengths are absorbed. The remaining wavelengths are scattered or transmitted and make the pigment appear to us as colored. For example, the pigment **chlorophyll** absorbs both blue and red light, and we see the remaining light, which is primarily green. If we

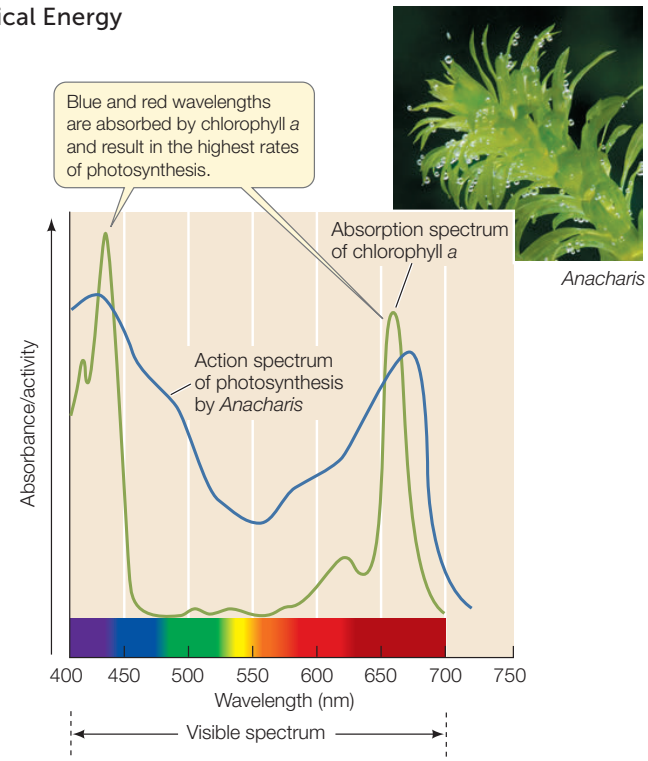


FIGURE 6.17 Absorption and Action Spectra The absorption spectrum of the purified pigment chlorophyll *a* from the aquatic plant *Anacharis* is similar to the action spectrum, obtained when different wavelengths of light are shone on the intact plant and the rate of photosynthesis is measured. In the thicker leaves of land plants, the action spectra show less of a dip in the green region (500–650 nm).

plot light absorbed by a purified pigment against wavelength, the result is an **absorption spectrum** for that pigment.

In contrast to the absorption spectrum, an **action spectrum** is a plot of the biological activity of an organism against the wavelengths of light to which it is exposed. An action spectrum can be determined as follows:

1. Place the organism (for example, a water plant with thin leaves) in a closed container.
2. Expose it to light of a certain wavelength for a period of time.
3. Measure the rate of photosynthesis by the amount of O_2 released.
4. Repeat with light of other wavelengths.

FIGURE 6.17 shows the absorption spectrum of the pigment chlorophyll *a*, which was isolated from the leaves of *Elodea* (also known as *Anacharis*), a common aquarium plant. Also shown is the action spectrum for photosynthetic activity by the same plant. A comparison of the two spectra shows that the wavelengths at which photosynthesis is highest are the same wavelengths at which chlorophyll *a* absorbs light.

In plants, two chlorophylls absorb light energy to drive the light reactions: chlorophyll *a* and chlorophyll *b*. These two molecules differ only slightly in their molecular structures. Both have a complex ring structure, similar to that of the heme group of hemoglobin, with a magnesium ion at the center (**FIGURE 6.18**).

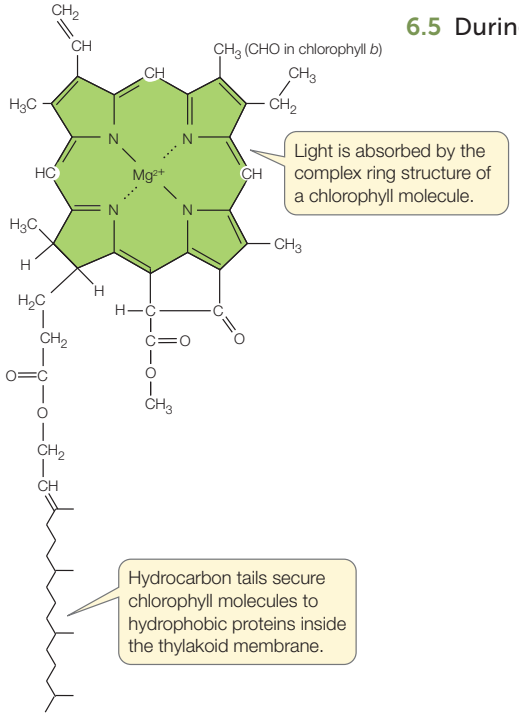


FIGURE 6.18 The Molecular Structure of Chlorophyll Chlorophyll consists of a complex ring structure (green) with a magnesium ion at its center, plus a hydrocarbon "tail." The tail anchors chlorophyll molecules to integral membrane proteins in the thylakoid membrane. Chlorophyll *a* and chlorophyll *b* are identical except for the replacement of a methyl group ($-CH_3$) with an aldehyde group ($-CHO$), shown on the upper right side of the ring structure.

A long hydrocarbon "tail" anchors the chlorophyll molecule to integral proteins in the thylakoid membrane of the chloroplast.

As mentioned above, the chlorophylls absorb blue and red light, which are near the two ends of the visible spectrum (see Figure 6.17). In addition, plants possess accessory pigments that absorb photons intermediate in energy between the red and the blue wavelengths, and then transfer a portion of that energy to the chlorophylls. Among these accessory pigments are carotenoids such as β-carotene, which absorb photons in the blue and blue-green wavelengths and appear deep yellow. The phycobilins, which are found in red algae and in cyanobacteria, absorb various yellow-green, yellow, and orange wavelengths.

LINK

Some plant pigments act as sensors that regulate growth and development; see **Concept 26.4**

Light absorption results in photochemical change

The pigments in photosynthetic organisms are arranged into energy-absorbing antenna systems, also called **light-harvesting complexes** (FIGURE 6.19A). These form part of a large multiprotein complex called a **photosystem**, where light energy is converted into chemical energy (FIGURE 6.19B). The photosystem spans the thylakoid membrane and consists of multiple antenna systems with their associated pigment molecules, all surrounding a **reaction center**.

When chlorophyll absorbs light, it enters an excited state. This is an unstable situation, and the chlorophyll rapidly returns to its ground state, releasing most of the absorbed energy.

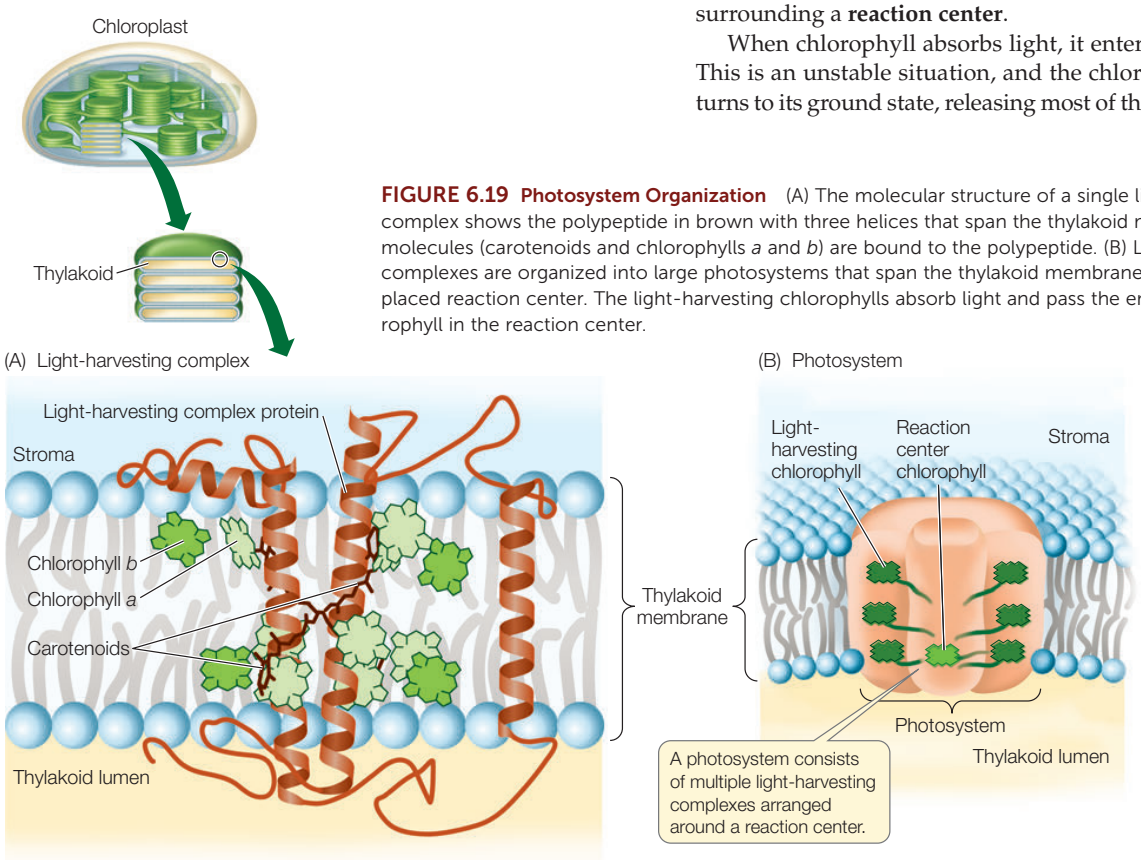


FIGURE 6.19 Photosystem Organization (A) The molecular structure of a single light-harvesting complex shows the polypeptide in brown with three helices that span the thylakoid membrane. Pigment molecules (carotenoids and chlorophylls *a* and *b*) are bound to the polypeptide. (B) Light-harvesting complexes are organized into large photosystems that span the thylakoid membrane and have a centrally placed reaction center. The light-harvesting chlorophylls absorb light and pass the energy on to a chlorophyll in the reaction center.

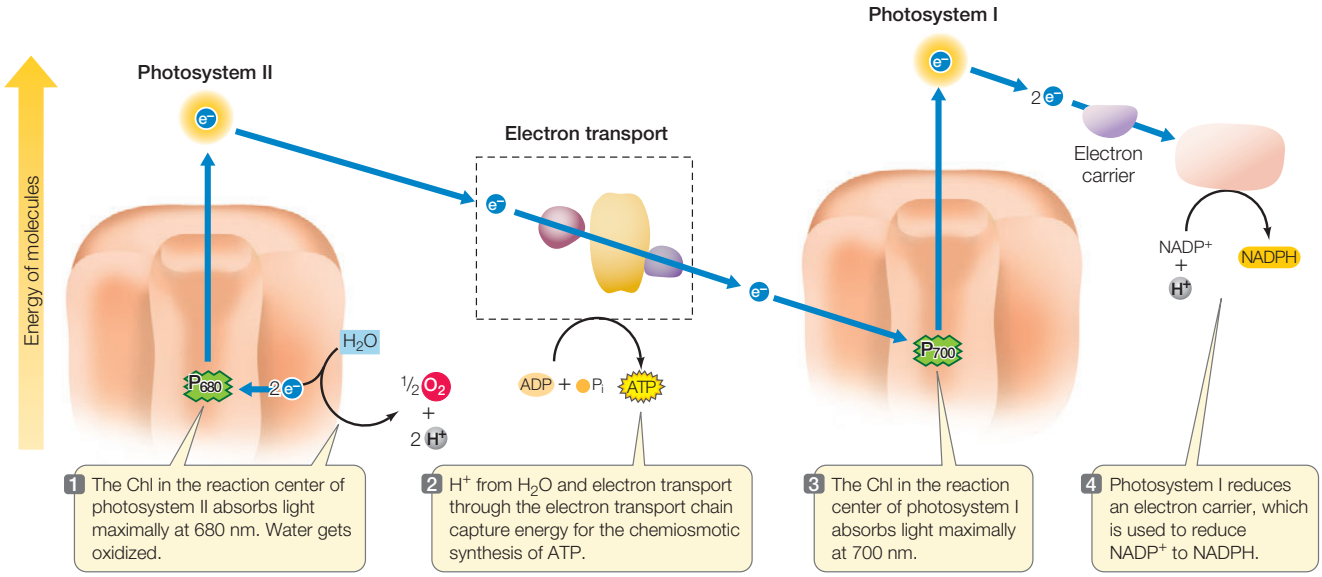
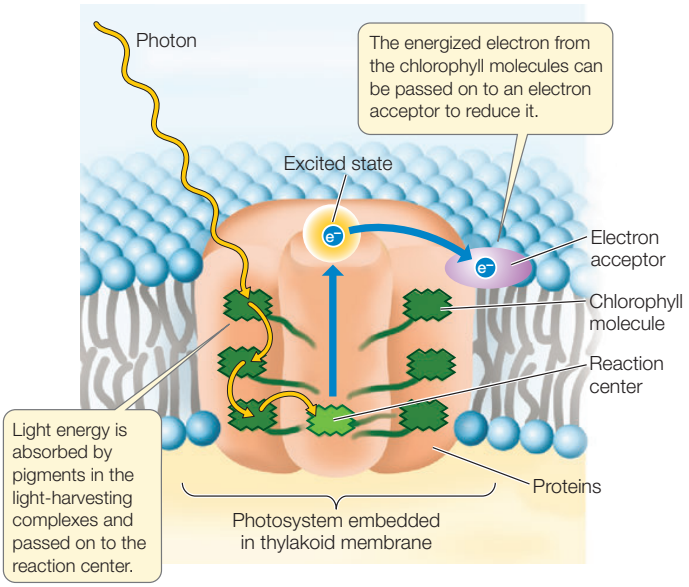


FIGURE 6.20 Noncyclic Electron Transport Uses Two Photosystems As chlorophyll molecules in the reaction centers of photosystems I and II absorb light energy, they pass electrons into a series of redox reactions, ultimately producing NADPH and ATP. The term "Z scheme" describes the path (blue arrows) of electrons as they travel through the two photosystems. In this scheme the vertical positions represent the energy levels of the molecules in the electron transport system.

Go to ANIMATED TUTORIAL 6.3
Photophosphorylation
PoL2e.com/at6.3

This is an extremely rapid process—measured in picoseconds (trillionths of a second)! For most chlorophyll molecules embedded in the thylakoid membrane, the released energy is absorbed by other, adjacent chlorophyll molecules. The energy eventually arrives at a ground-state chlorophyll molecule at the reaction center (symbolized by Chl), which absorbs the energy and becomes excited (Chl*). But when the reaction center chlorophyll returns to the ground state, something very different occurs. *The reaction center converts the absorbed light energy into chemical energy*. The chlorophyll molecule in the reaction center absorbs sufficient energy that it actually *gives up its excited electron to a chemical acceptor*:

$$\text{Chl}^* + \text{acceptor} \rightarrow \text{Chl}^+ + \text{acceptor}^-$$



This, then, is the first consequence of light absorption by chlorophyll: *the reaction center chlorophyll (Chl*) loses its excited electron in a redox reaction and becomes Chl⁺*. As a result of this transfer of an electron, the chlorophyll gets oxidized, while the acceptor molecule is reduced.

Reduction leads to ATP and NADPH formation

The electron acceptor that is reduced by Chl* is the first of a chain of electron carriers in the thylakoid membrane. Electrons are passed from one carrier to another in an energetically "downhill" series of reductions and oxidations. Thus the thylakoid membrane has an electron transport system similar to the respiratory chain of mitochondria (see Concept 6.2). The final electron acceptor is $NADP^+$, which gets reduced:

$$NADP^+ + H^+ + 2\ e^- \rightarrow NADPH$$

As in mitochondria, ATP is produced chemiosmotically during the process of electron transport (a process called photophosphorylation). **FIGURE 6.20** shows the series of electron transport reactions that use the energy from light to generate NADPH and ATP. There are two photosystems, each with its own reaction center:

- **Photosystem I** (containing the "P_{700}" chlorophylls at its reaction center) absorbs light energy at 700 nm and passes an excited electron to $NADP^+$, reducing it to NADPH.
- **Photosystem II** (with "P_{680}" chlorophylls at its reaction center) absorbs light energy at 680 nm, oxidizes water molecules, and initiates the electron transport chain that produces ATP.

Let's look in more detail at these photosystems, beginning with photosystem II.

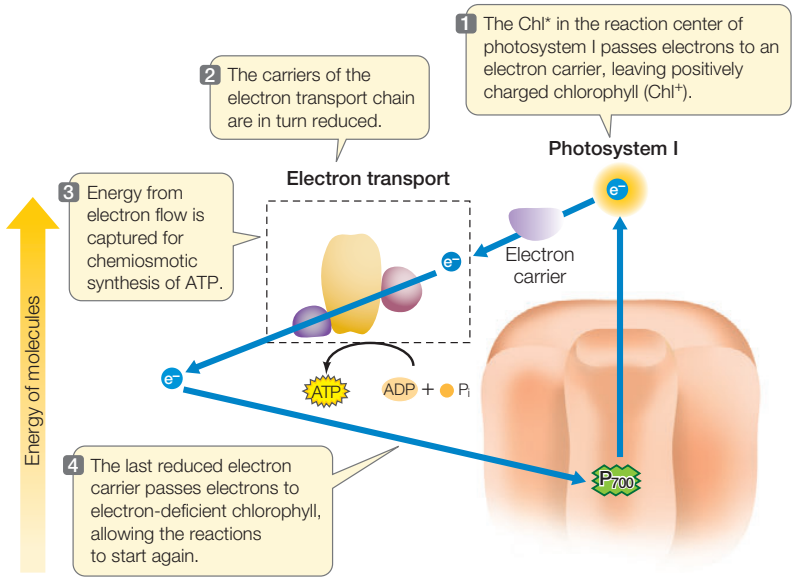


FIGURE 6.21 Cyclic Electron Transport Traps Light Energy as ATP Cyclic electron transport produces ATP but no NADPH.

PHOTOSYSTEM II After an excited chlorophyll in the reaction center (Chl*) gives up its energetic electron to reduce a chemical acceptor molecule, the chlorophyll lacks an electron and is very unstable. It has a strong tendency to "grab" an electron from another molecule to replace the one it lost—in chemical terms, it is a strong oxidizing agent. The replenishing electrons come from water, splitting the H—O—H bonds:

$$H_2O \rightarrow \tfrac{1}{2}\,O_2 + 2\,H^+ + 2\,e^-$$

$$2\,e^- + 2\,Chl^+ \rightarrow 2\,Chl$$

$$\text{Overall: } 2\,Chl^* + H_2O \rightarrow 2\,Chl + 2\,H^+ + \tfrac{1}{2}\,O_2$$

The source of the O_2 produced by photosynthesis is H_2O.

Go to ANIMATED TUTORIAL 6.4
The Source of the Oxygen Produced by Photosynthesis
PoL2e.com/at6.4

Back to the electron acceptor in the electron transport system: the energetic electrons are passed through a series of thylakoid membrane–bound carriers to a final acceptor at a lower energy level. As in the mitochondrion, a proton gradient is generated and is used by ATP synthase to store energy in the bonds of ATP.

PHOTOSYSTEM I In photosystem I, an excited electron from the Chl* at the reaction center reduces an acceptor. The oxidized chlorophyll (Chl^+) now "grabs" an electron, but in this case the electron comes from the last carrier in the electron transport system. This links the two photosystems chemically. They are also linked spatially, with the two photosystems in the thylakoid membrane. The energetic electrons from photosystem I pass through several molecules and end up reducing $NADP^+$ to NADPH.

Next in the process of harvesting light energy to produce carbohydrates is the series of carbon-fixation reactions. These reactions require more ATP than NADPH. If the pathway we just described—the linear, or noncyclic, pathway—were the only set of light reactions operating, there might not be sufficient ATP for carbon fixation. **Cyclic electron transport** makes up for this imbalance. This pathway uses only photosystem I and produces ATP but not NADPH; it is cyclic because the electrons flow from the reaction center of photosystem I, through the electron transport chain, and then back to photosystem I (**FIGURE 6.21**).

APPLY THE CONCEPT

During photosynthesis, light energy is converted to chemical energy

The key role of water in supplying electrons for reduction of light-activated chlorophyll in the light reactions and in the release of O_2 to the atmosphere in the process of energy conversion has been investigated using isotopes of oxygen. The ^{18}O isotope is heavier than normal oxygen (^{16}O), and a mass spectrometer can be used to detect the difference. Green plant cells were exposed to light, water, and CO_2. (The CO_2 was supplied as the bicarbonate ion HCO_3^-, which forms CO_2 when dissolved in water.) In the first experiment, some of the oxygen atoms in the water molecules were ^{18}O ($H_2{}^{18}O$), while CO_2 had the normal form of oxygen ($C^{16}O_2$). In the second experiment, the situation was reversed, with $H_2{}^{16}O$ and $C^{18}O_2$ being supplied to the cells. After 2 hours of photosynthesis, the ratio of ^{18}O to ^{16}O was measured in the O_2 produced by the cells.[a]

	ISOTOPE RATIO		
EXPERIMENT	H_2O	CO_2	O_2
1	0.85	0.31	0.84
2	0.20	0.50	0.20

1. In experiment 1, was the isotopic ratio of O_2 more similar to that of H_2O or CO_2?
2. What about experiment 2? What can you conclude?

[a] S. Ruben et al. 1941. *Journal of the American Chemical Society* 63: 877–879.

CHECKpoint CONCEPT 6.5

- ✓ What are the reactants and products of the light reactions of photosynthesis?
- ✓ Chlorophyll absorbs light of blue and red wavelengths, but leaves also absorb some light at other wavelengths. Explain why.
- ✓ Trace the flow of electrons in noncyclic electron transport in the chloroplast and compare it with that of cyclic electron transport.
- ✓ Write equations for the production of the following in photosynthesis, and indicate whether they are oxidations, reductions, or neither: Chl*; O_2; ATP; NADPH.

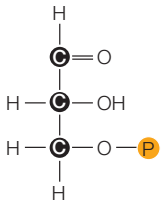
We have seen how photosystems I and II absorb light energy, which ultimately ends up as chemical energy in ATP and NADPH. Let's look now at how these two energy-rich molecules are used in the carbon-fixation reactions to reduce CO_2 and thereby form carbohydrates.

CONCEPT 6.6 Photosynthetic Organisms Use Chemical Energy to Convert CO_2 to Carbohydrates

The energy in ATP and NADPH is used in the carbon-fixation reactions to "fix" CO_2 into a reduced form and convert it to carbohydrates. Most CO_2 fixation occurs only in the light, when ATP and NADPH are being generated. The metabolic pathway occurs in the stroma, or central region, of the chloroplast (see Figure 6.15) and is called the **Calvin cycle** after one of its discoverers, Melvin Calvin.

Like all biochemical pathways, each reaction in the Calvin cycle is catalyzed by a specific enzyme. The cycle is composed of three distinct processes (**FIGURE 6.22**):

- *Fixation* of CO_2. The initial reaction of the Calvin cycle adds the one-carbon CO_2 to an acceptor molecule, the five-carbon ribulose 1,5-bisphosphate (RuBP). The immediate product is a six-carbon molecule, which quickly breaks down into two three-carbon molecules called 3-phosphoglycerate (3PG; **FIGURE 6.23**). The enzyme that catalyzes this reaction, **ribulose bisphosphate carboxylase/oxygenase** (**rubisco**), is rather sluggish as enzymes go. It typically catalyzes two to three fixation reactions per second. Because of this, plants need a lot of rubisco to perform enough photosynthesis to satisfy the needs of growth and metabolism. Rubisco constitutes about half of all the protein in a leaf, and it is probably the most abundant protein in the world!
- *Reduction* of *3PG to form glyceraldehyde 3-phosphate.* This series of reactions involves a phosphorylation (using the high-energy phosphate from an ATP made in the light reactions) and a reduction (using an NADPH made in the light reactions). The product is **glyceraldehyde 3-phosphate** (**G3P**), which is a three-carbon sugar phosphate, also called triose phosphate:



Glyceraldehyde 3-phosphate (G3P)

- *Regeneration of the CO_2 acceptor, RuBP.* Most of the G3P ends up as ribulose monophosphate (RuMP), and ATP is used to convert this compound into RuBP. So for every "turn" of the Calvin cycle, one CO_2 is fixed and the CO_2 acceptor is regenerated.

What happens to the extra G3P made by the Calvin cycle (see Figure 6.22)? It has two fates, depending on the time of day and the needs of different parts of the plant:

- Some of the extra G3P is exported out of the chloroplast to the cytosol, where it is converted to hexoses (glucose and fructose). This is the familiar $C_6H_{12}O_6$ from the general equation for photosynthesis. These molecules may be catabolized for energy in mitochondria as part of cellular respiration; used as carbon skeletons for the synthesis of amino acids and other molecules (see Figure 6.13); or converted to sucrose, which is transported out of the leaf to other organs in the plant.
- Late in the day when glucose has accumulated inside the chloroplast, the glucose units are linked to form the polysaccharide starch. This storage carbohydrate can then be drawn on during the night so that the photosynthetic tissues can continue to export sucrose to the rest of the plant, even when photosynthesis is not taking place. In addition, starch is abundant in nonphotosynthetic organs such as roots, underground stems, and seeds, where it provides a ready supply of glucose to fuel cellular activities, including plant growth.

The products of the Calvin cycle are of crucial importance to Earth's entire biosphere. The C—H covalent bonds generated by this cycle store almost all of the energy for life on Earth. Photosynthetic organisms, which are also called **autotrophs** ("self-feeders"), release most of this energy in cellular respiration, and use it to support their own growth, development, and reproduction. But plants are also the source of energy for other organisms. Much plant matter ends up being consumed by **heterotrophs** ("other-feeders"), including humans and other animals, which cannot photosynthesize. Heterotrophs depend on autotrophs for chemical energy, which they harvest via cellular respiration. In addition, many heterotrophs rely on plants to produce other molecules (such as vitamins) that the heterotrophs cannot synthesize for themselves.

LINK

Concept 44.3 describes how the energy captured by autotrophs flows through ecological communities

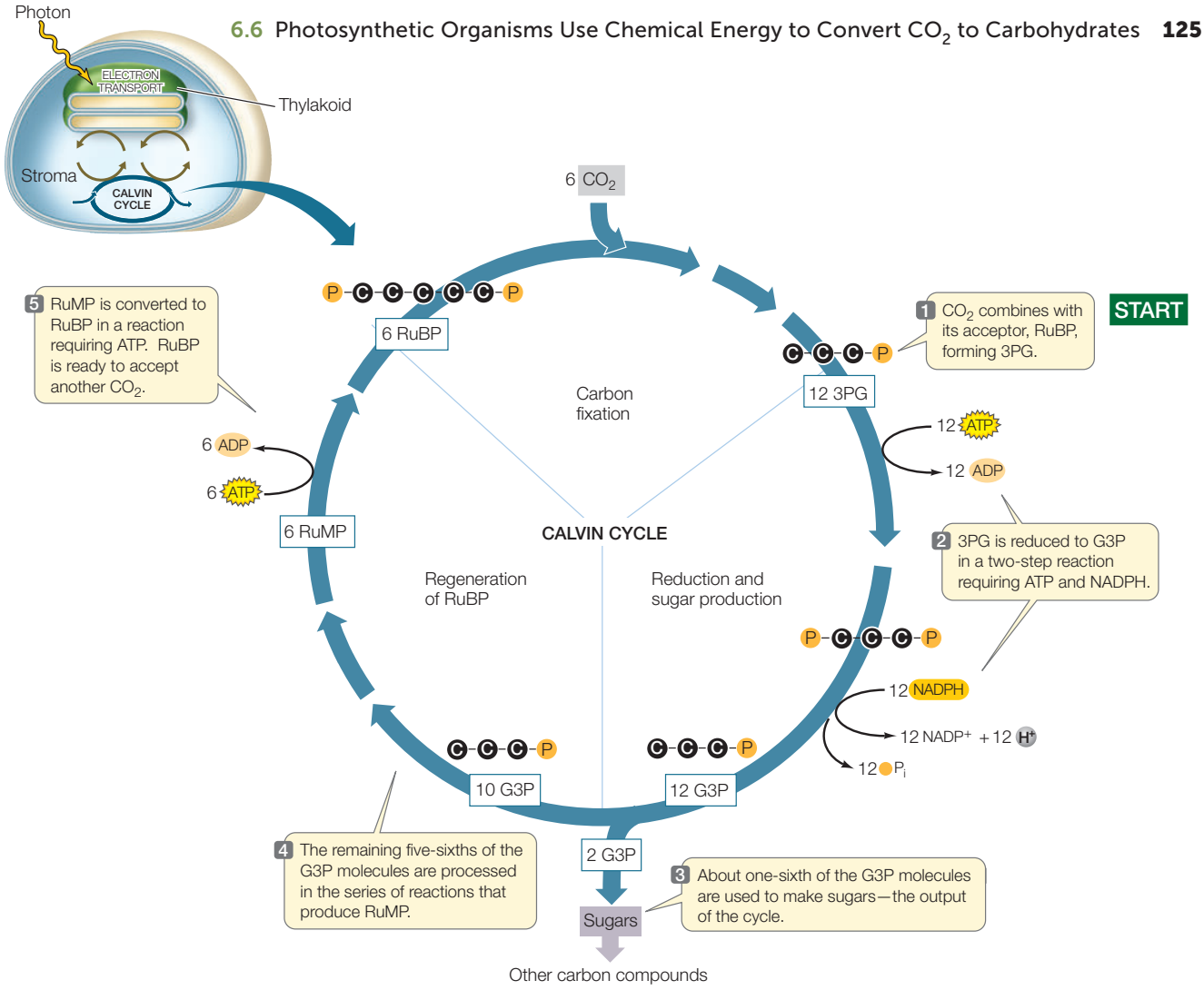


FIGURE 6.22 The Calvin Cycle The Calvin cycle uses the ATP and NADPH generated in the light reactions to produce G3P from CO_2. The G3P is used as a starting material for the production of glucose and other carbohydrates. Six turns of the cycle are needed to produce one molecule of the hexose glucose.

Go to ACTIVITY 6.7 The Calvin Cycle
PoL2e.com/ac6.7

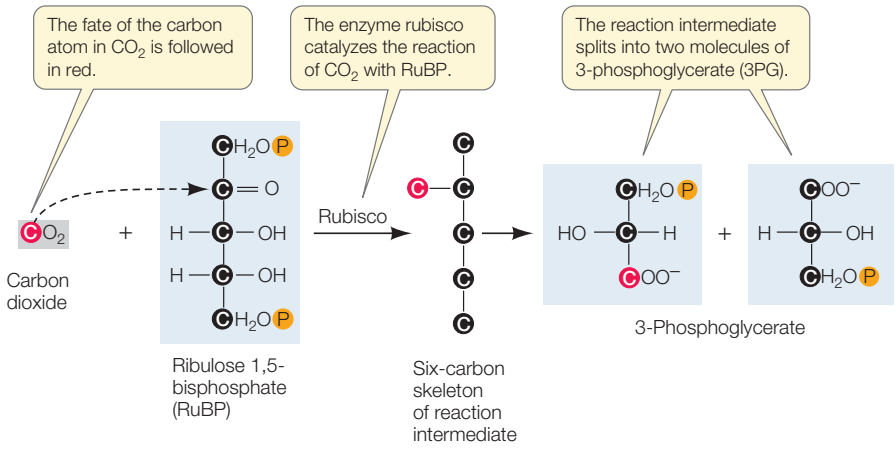


FIGURE 6.23 RuBP Is the Carbon Dioxide Acceptor The enzyme rubisco adds CO_2 to the five-carbon compound RuBP. The resulting six-carbon compound immediately splits into two molecules of 3PG.

Go to ANIMATED TUTORIAL 6.5
Tracing the Pathway of CO_2
PoL2e.com/at6.5

CHECKpoint CONCEPT 6.6

- ✓ What are the three processes of the Calvin cycle?
- ✓ If green plant cells are incubated in the presence of CO_2 molecules containing radioactive carbon atoms, the fate of the carbon atoms can be followed. In an experiment, radioactive CO_2 was given for 1 minute to plant cells, and then the cells were examined after 1, 5, 10, 20, and 30 minutes. The following molecules were labeled with radioactive carbon at some point(s): glucose, glyceraldehyde 3-phosphate, glycine (an amino acid), 3-phosphoglycerate, ribulose 1,5-bisphosphate, and sucrose. List these molecules in the order in which they first become labeled.

Q Why does fresh air inhibit the formation of alcohol by yeast cells?

ANSWER Armed with our knowledge of metabolism, we can now explain Pasteur's observations of beet sugar and alcohol:

1. *Nothing happened to beet mash unless microscopic yeast cells were present.* Beet sugar is a product of photosynthetic CO_2 fixation (Concept 6.6). Catabolism of carbohydrates is a cellular activity, characteristic of virtually all living things (Concepts 6.2 and 6.3).
2. *In the presence of fresh air, yeast cells grew vigorously on the mash, and bubbles of CO_2 were formed.* Carbohydrate catabolism under aerobic conditions yields a large amount of ATP (Concepts 6.1 and 6.2), which can be used to fuel anabolic pathways for growth (Concept 6.4). Under aerobic conditions, glucose is fully oxidized to CO_2 (Concept 6.2).
3. *Without fresh air, the yeast grew slowly, less CO_2 was produced, and alcohol was formed.* Alcoholic fermentation occurs in anaerobic conditions, and yields far less ATP (for growth) than aerobic metabolism; a small amount of CO_2 is produced (Concept 6.3).

As we noted in the opening story, the use by humans of yeasts for fermentation has a long history (**FIGURE 6.24**).

Beer comes from the fermentation of barley seeds. These are soaked to begin the process of germination, which includes the induction of an enzyme that hydrolyzes the starch stored in the seed. The resulting disaccharide, maltose, is what the yeast cells use for energy—they first break the maltose down to its glucose monomers. At a cool temperature and under anaerobic conditions, yeast produces alcohol. An herb called hops is added to impart a distinctive bitter flavor.

Wine grapes are crushed, and the resulting juice contains sugars that are used by yeast for fermentation. There are yeasts growing on the skins of grapes naturally, so the original winemakers did not add yeast. More recently, special yeast strains have been used to control the fermentation process. The longer this process is allowed to proceed, the less sugar remains (resulting in a less sweet taste) and the higher the alcohol content. After fermentation the wine is stored in wooden casks, typically at cool temperatures. During this time hundreds of molecular transformations occur, giving different wines distinctive properties.

Bread is made from ground-up plant seeds (flour), which contain abundant starch. Moisture activates enzymes inside the seed that catabolize the starch to monosaccharides. These monosaccharides are used by bread yeast in fermentation reactions that result in CO_2 production, and the resulting gas bubbles cause the complex flour mixture to rise.

FIGURE 6.24 Products of Glucose Metabolism Beer, wine, and bread are all made using fermentation reactions in yeast cells.

SUMMARY

CONCEPT 6.1 ATP and Reduced Coenzymes Play Important Roles in Biological Energy Metabolism

- Metabolism is carried out in small steps and involves coenzymes as carriers of chemical energy.
- Adenosine triphosphate (ATP) serves as "energy currency" in the cell. Hydrolysis of ATP releases a large amount of free energy. **Review Figure 6.1 and ACTIVITY 6.1**
- In **oxidation**, a material loses electrons by transfer to another material, which thereby undergoes **reduction**. Such **redox** reactions transfer large amounts of energy.
- The coenzyme nicotinamide adenine dinucleotide (NAD) is a key electron carrier in biological redox reactions. It exists in two forms, one oxidized (NAD^+) and the other reduced (NADH). **Review Figure 6.4**

CONCEPT 6.2 Carbohydrate Catabolism in the Presence of Oxygen Releases a Large Amount of Energy

- The sequential pathways of aerobic glucose catabolism are **glycolysis**, **pyruvate oxidation**, and the **citric acid cycle**. **Review Figure 6.6**
- In glycolysis, a series of ten enzyme-catalyzed reactions in the cell cytoplasm converts glucose to two molecules of pyruvate. Energy is released and captured as ATP and NADH. **Review Figure 6.7**
- The next pathway, pyruvate oxidation, links glycolysis to the citric acid cycle. Pyruvate oxidation converts pyruvate into the two-carbon molecule acetyl CoA.
- In the citric acid cycle, a series of eight enzyme-catalyzed reactions fully oxidizes acetyl CoA to CO_2. Much energy is released, and most is used to form NADH. **Review Figure 6.8 and ACTIVITY 6.2**
- The energy in NADH is used to make ATP via a series of **electron transport** carriers and chemiosmosis. **Review Figure 6.9, ANIMATED TUTORIAL 6.1, and ACTIVITY 6.3**
- In **oxidative phosphorylation**, ATP is formed with the energy derived from the reoxidation of reduced coenzymes. This depends on the process of **chemiosmosis**, in which a proton gradient across a membrane powers ATP formation. This occurs at the cell membrane in prokaryotes, and in the mitochondria and chloroplasts in eukaryotes. **Review Figures 6.10 and 6.11 and ANIMATED TUTORIAL 6.2**

CONCEPT 6.3 Carbohydrate Catabolism in the Absence of Oxygen Releases a Small Amount of Energy

- In the absence of O_2, glycolysis is followed by **fermentation**. Together, these pathways partially oxidize pyruvate and generate the end products lactic acid or ethanol. In the process, NAD^+ is regenerated from NADH so that glycolysis can continue, thus generating a small amount of ATP. **Review Figure 6.12 and ACTIVITY 6.4**
- For each molecule of glucose used, fermentation yields 2 molecules of ATP. In contrast, glycolysis, pyruvate oxidation, the citric acid cycle, and oxidative phosphorylation yield up to 32 molecules of ATP per molecule of glucose. **Review ACTIVITY 6.5**

CONCEPT 6.4 Catabolic and Anabolic Pathways Are Integrated

- The catabolic pathways for the breakdown of carbohydrates, lipids, and proteins feed into the energy-harvesting metabolic pathways. **Review Figure 6.13**
- Anabolic pathways use intermediate components of the energy-harvesting pathways to synthesize fatty acids, amino acids, and other essential building blocks.
- The formation of glucose from intermediates of glycolysis and the citric acid cycle is called **gluconeogenesis**.
- The enzymes of glycolysis and the citric acid cycle are regulated by various mechanisms, including allosteric regulation. Excess acetyl CoA is diverted into fatty acid synthesis. **Review ACTIVITY 6.6**

CONCEPT 6.5 During Photosynthesis, Light Energy Is Converted to Chemical Energy

- The **light reactions** of photosynthesis convert light energy into chemical energy. They produce ATP and reduce $NADP^+$ to NADPH. **Review Figure 6.15**
- Light is a form of **electromagnetic radiation**. It is emitted in particle-like packets called **photons** but has wavelike properties. Molecules that absorb light in the visible spectrum are called **pigments**. Photosynthetic organisms have several pigments, most notably **chlorophylls**. **Review Figures 6.16–6.18**
- The absorption of a photon puts a chlorophyll molecule into an excited state that has more energy than its ground state. This energy can be transferred via other chlorophylls to one in the **reaction center** of a photosystem. **Review Figure 6.19**
- An excited chlorophyll can act as a reducing agent, transferring excited electrons to other molecules. Oxidized chlorophyll regains electrons by the splitting of H_2O.
- In the thylakoid membrane of the chloroplast, **photosystems I** and **II** and a noncyclic electron transport system produce ATP via oxidative phosphorylation. NADPH and O_2 are also produced. **Review Figure 6.20 and ANIMATED TUTORIALS 6.3 and 6.4**
- **Cyclic electron transport** uses only photosystem I and produces only ATP. **Review Figure 6.21**

CONCEPT 6.6 Photosynthetic Organisms Use Chemical Energy to Convert CO_2 to Carbohydrates

- The **Calvin cycle** makes carbohydrates from CO_2. The cycle consists of three processes: fixation of CO_2, reduction and sugar production, and regeneration of RuBP. **Review Figure 6.22 and ACTIVITY 6.7**
- RuBP is the initial CO_2 acceptor, and 3PG is the first stable product of CO_2 fixation. The enzyme **rubisco** catalyzes the reaction of CO_2 and RuBP to form 3PG. **Review Figure 6.23 and ANIMATED TUTORIAL 6.5**
- ATP and NADPH formed by the light reactions are used to fuel the reduction of 3PG to form **glyceraldehyde 3-phosphate** (**G3P**)—a starting material for the synthesis of glucose and other carbohydrates.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is6

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

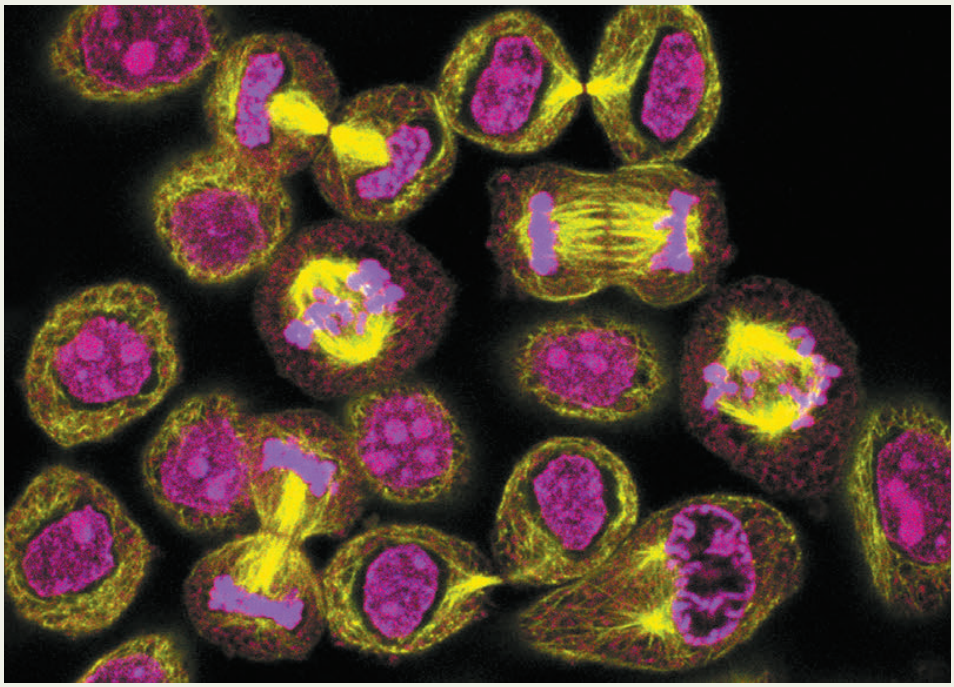
PART 2

Genetics

7

The Cell Cycle and Cell Division

KEY CONCEPTS



These cervical cancer cells are actively dividing; many of the cells are in various stages of mitosis and cytokinesis.

Ruth felt healthy and was surprised when she was called back to her physician's office a week after her annual checkup. "Your lab report indicates you have early cervical cancer," said the doctor. "I ordered a follow-up test, and it came back positive. At some point, you were infected with HPV."

Ruth felt numb as soon as she heard the word "cancer." Her mother had died of breast cancer in the previous year. The doctor's statement about HPV—human papillomavirus—did not register in her consciousness. Sensing Ruth's discomfort, the doctor quickly reassured her that the cancer had been caught at an early stage and that a simple surgical procedure would remove it. Two weeks later, the cancer was removed and Ruth remains cancer-free. She was fortunate that her annual medical exam included a Papanicolau (Pap) test, in which the cells lining the cervix are examined for abnormalities. Since they were begun almost 50 years ago in Europe, Pap tests have resulted in the detection and removal of millions of early cervical cancers, and the death rate from this potentially lethal disease has plummeted.

The role of HPV in causing most cervical cancers was discovered only recently. The German physician Harald Zur-Hausen was awarded the Nobel Prize in 2008 for this discovery, and it has led to a vaccine to prevent future infections. There are many different types of HPV, and many of the ones that infect humans cause warts, which are small, rough growths on the skin. The types of HPV that infect tissues at the cervix get there by sexual transmission, and this is a common infection.

When HPV arrives at the tissues lining the cervix, it has one of two fates. Most of the time it enters the cells and turns them into HPV factories, releasing a lot of HPV particles into the mucus outside the uterus. These viruses can infect another person during a sexual encounter. But in some cases the virus follows a different, more sinister path. The viral DNA becomes incorporated into the DNA of the cervical cells, and the cells are stimulated to reproduce.

Cell reproduction in healthy humans is tightly controlled by a variety of mechanisms, but the virus-infected cells lose these controls. Understanding how cell division is controlled is clearly an important subject for the development of cancer treatments. But cell division is not just important in medicine. It underlies the growth, development, and reproduction of all organisms.

Q How does infection with HPV result in uncontrolled cell reproduction?

You will find the answer to this question on page 148.

CONCEPT 7.1 Different Life Cycles Use Different Modes of Cell Reproduction

In Chapter 4 we described cells as the basic compartments of life, where biological processes are separated from the external environment. Cells are also essential for biological reproduction.

The life span of an organism from birth to death is intimately linked to cell reproduction, which is commonly referred to as **cell division**: a process by which a parent cell duplicates its genetic material and then divides into two similar cells. Cell division plays important roles in the growth and repair of tissues in multicellular organisms, as well as in the reproduction of all organisms (FIGURE 7.1). Although the details vary widely, organisms have two basic strategies for reproducing themselves: asexual reproduction and sexual reproduction. These two strategies make use of different types of cell division.

Asexual reproduction by binary fission or mitosis results in genetic constancy

Asexual reproduction is a rapid and effective means of making new individuals, and it is common in nature. The offspring resulting from asexual reproduction are **clones** of the parent organism—they are genetically identical (or virtually identical) to each other and the parent. Any genetic variations among the parent and offspring are due to changes called mutations, which are alterations in DNA sequence caused by environmental factors or errors in DNA replication (see Concept 9.3). This small amount of variation contrasts with the extensive variation possible in sexually reproducing organisms, as we will see later in this chapter.

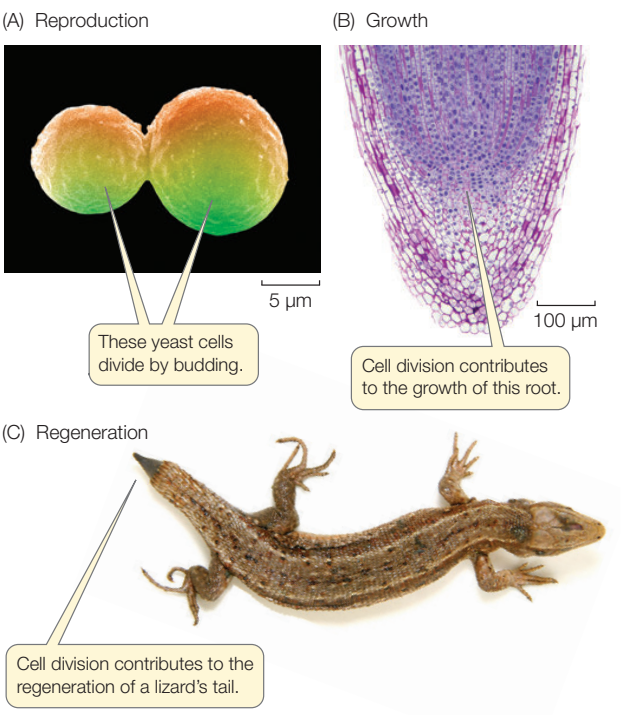


FIGURE 7.1 The Importance of Cell Division Cell division is the basis for (A) reproduction, (B) growth, and (C) repair and regeneration of tissues.

FIGURE 7.2 Asexual Reproduction on a Large Scale This forest of aspens in Utah's Wasatch Mountains arose via asexual reproduction. Genetically, all these trees are virtually identical.

In most cases, single-celled prokaryotes reproduce by binary fission, an asexual process that we will discuss in Concept 7.2. A cell of the bacterium *Escherichia coli* is the whole organism, so when it divides to form two new cells, it is reproducing. Similarly, single-celled eukaryotes (such as fission yeast) can reproduce asexually through mitosis followed by cytokinesis, processes that also produce two genetically identical cells (see Concept 7.2).

Many multicellular eukaryotes, including fungi and plants, can also reproduce asexually. Perhaps the most dramatic example of this is a forest containing thousands of aspen trees (*Populus tremuloides*) in the Wasatch Mountains of Utah (FIGURE 7.2). DNA analyses have shown that these trees are clones—they are virtually identical genetically. Aspen can reproduce sexually, with male and female plants, but in many aspen stands all the trees are the same sex and reproduction is asexual. An extensive root system spreads through the soil, and at intervals stems form and grow into new trees.

Sexual reproduction by meiosis results in genetic diversity

Sexual reproduction involves the fusion of two specialized cells called **gametes**, and can result in offspring with considerable genetic variation. In many diploid organisms, the gametes form by **meiosis**—a process of cell division (described in Concept 7.4) resulting in daughter cells with only half the genetic material of the original cell. During meiosis, the genetic material is randomly separated and reorganized so that the daughter cells differ genetically from one another. Because of this genetic variation, some offspring of sexual reproduction may be better adapted than others to survive and reproduce in a particular environment. Meiosis thus increases genetic diversity, which is the raw material for natural selection and evolution.

As we described in Chapter 4, the DNA in eukaryotic cells is organized into multiple structures called chromosomes. Each

chromosome consists of a double-stranded molecule of DNA and associated proteins. In multicellular organisms, the body cells that are *not* specialized for reproduction are called **somatic cells**. In many familiar organisms, including most vascular plants and animals, the somatic cells each contain two sets of chromosomes, and the chromosomes occur in pairs called **homologous pairs**. One chromosome of each pair comes from the organism's female parent, and the other comes from its male parent. For example, in humans with 46 chromosomes, 23 come from the mother and 23 from the father, with, for example, a chromosome 1 from each parent, and so on.

The two chromosomes in a homologous pair (called homologs of one another) bear corresponding, though not identical, genetic information. For example, a homologous pair of chromosomes in a plant may carry different versions of a gene that controls seed shape. One homolog may carry the version for wrinkled seeds, while the other may carry the version for smooth seeds.

LINK

The inheritance of characteristics such as seed shape is discussed in **Concept 8.1**

Gametes contain only a single set of chromosomes—that is, one homolog from each pair. The number of chromosomes in a gamete is denoted by n, and the cell is said to be **haploid**. During sexual reproduction, two haploid gametes fuse to form a **zygote** in a process called **fertilization**. The zygote thus has two sets of chromosomes, just as the somatic cells do. The chromosome number in the zygote is denoted by $2n$, and the cells are said to be **diploid**.

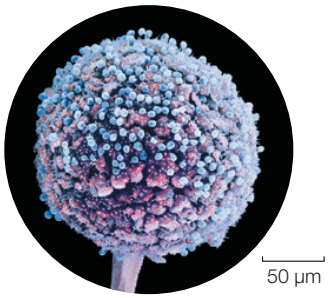
In many familiar organisms the zygote divides by mitosis, producing a new, mature organism with diploid somatic cells (as we have just described). But some other organisms have haploid stages in their life cycles.

Sexual life cycles are diverse

All sexual life cycles involve meiosis to produce haploid cells. In some cases, gametes develop immediately after meiosis. In others, each haploid cell divides and develops into a haploid organism—the haploid stage of the life cycle—that eventually produces gametes by mitosis. The fusion of gametes—fertilization—results in a zygote and begins the diploid stage of the life cycle. Since the origin of sexual reproduction, evolution has generated many different versions of the sexual life cycle. **FIGURE 7.3**

FIGURE 7.3 Sexual Life Cycles Involve Fertilization and Meiosis
In sexual reproduction, haploid (n) cells or organisms alternate with diploid ($2n$) cells or organisms.

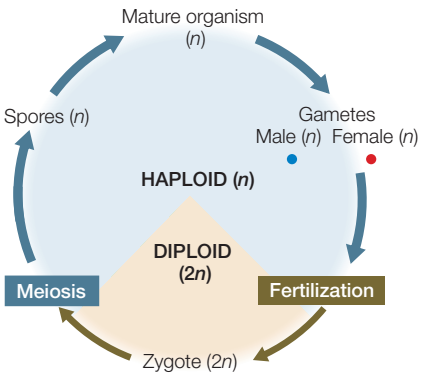
Go to ACTIVITY 7.1 Sexual Life Cycle
PoL2e.com/ac7.1



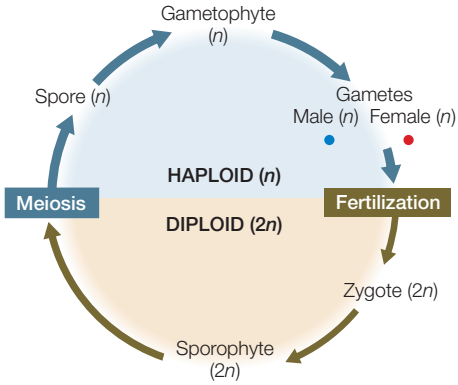
Bread mold (*Rhizopus stolonifer*)
(haploid organism)

Lady fern (*Athyrium felix-femina*)
(diploid sporophyte)

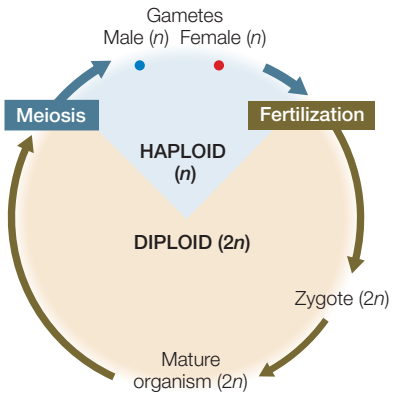
Nyala (*Tragelaphus angasii*)
(diploid organism)



In the **haplontic life cycle**, the mature organism is haploid and the zygote is the only diploid stage. Found in most protists, fungi, and some green algae.



In **alternation of generations**, the organism passes through haploid and diploid stages that are both multicellular. Found in most plants and some fungi.



In the **diplontic life cycle**, the organism is diploid and the gametes are the only haploid stage. Found in animals, brown algae, and some fungi.

presents three examples. The life cycles of a variety of organisms will be described in detail in Part 4. For now we will focus on the role of sexual reproduction in generating diversity among individuals.

The essence of sexual reproduction is the *random selection of half of the diploid chromosome set* to make a haploid gamete, followed by fusion of two haploid gametes from separate parents to produce a diploid cell. As we will see later in this chapter, further diversity is introduced by events that take place during meiosis. All of these steps contribute to a shuffling of genetic information in the population, so that no two individuals have exactly the same genetic constitution. The diversity provided by sexual reproduction opens up enormous opportunities for evolution.

CHECKpoint CONCEPT 7.1

- ✓ In terms of the genetic composition of offspring, what is the difference between sexual and asexual reproduction?
- ✓ Discuss the advantages of sexual versus asexual reproduction in terms of evolution. Could evolution proceed without sexual reproduction? Explain your answer.

We have briefly mentioned the different types of cell division and the roles they play in the life cycles of organisms. Now let's look in more detail at the processes of cell division, starting with binary fission and mitosis.

CONCEPT 7.2 Both Binary Fission and Mitosis Produce Genetically Identical Cells

Cell division by either binary fission or mitosis produces two genetically identical cells. This is the basis of asexual reproduction in single-celled organisms: prokaryotes reproduce by binary fission, and single-celled eukaryotes reproduce by mitosis. In multicellular organisms, mitosis is a way to build tissues and organs during development and to repair damaged tissues once development is complete.

In order for any cell to divide, the following events must occur:

- There must be one or more **reproductive signals**. These signals initiate cell division and may originate from either inside or outside the cell.
- **DNA replication** (i.e., replication of the genetic material) must occur so that each of the two new cells will have a full complement of genes to complete cell functions.
- The cell must distribute the replicated DNA to each of the two new cells. This process is called **DNA segregation**.
- The cytoplasm must divide to form the two new cells, each surrounded by a cell membrane and a cell wall in organisms that have one. This process is called **cytokinesis**.

Let's see how these events occur during the processes of binary fission in prokaryotes and mitosis in eukaryotes.

Prokaryotes divide by binary fission

In prokaryotes, cell division results in the reproduction of the entire single-celled organism. The cell grows in size, replicates its DNA, and then separates the cytoplasm and DNA into two new cells by a process called **binary fission**.

REPRODUCTIVE SIGNALS External factors such as environmental conditions and nutrient concentrations are common reproductive signals for prokaryotes. For example, the bacterium *Bacillus subtilis* can divide every 30 minutes under ideal conditions. But when nutrients in its environment are low, it stops dividing. It then resumes dividing when conditions improve.

DNA REPLICATION In most prokaryotic cells, almost all of the genetic information is carried on one single chromosome. In many cases the ends of the single DNA molecule are covalently joined, making the chromosome circular. Two regions of the prokaryotic chromosome play functional roles in cell reproduction:

- *ori*: the site where replication of the circular chromosome starts (the *ori*gin of replication)
- *ter*: the site where replication ends (the *ter*minus of replication)

Chromosome replication takes place as the DNA is threaded through a "replication complex" of proteins near the center of the cell. Replication begins at the *ori* site and moves toward the *ter* site. When replication is complete, the two daughter DNA molecules separate and segregate from one another at opposite ends of the cell. In rapidly dividing prokaryotes, DNA replication occupies the entire time between cell divisions.

DNA SEGREGATION Replication begins near the center of the cell, and as it proceeds, the *ori* regions move toward opposite ends of the cell (**FIGURE 7.4**). DNA sequences adjacent to the *ori* region bind proteins that are essential for this segregation. This is an active process, since the binding proteins hydrolyze ATP. Components of the prokaryotic cytoskeleton are involved in the segregation process. In particular, a bacterial protein that is structurally related to actin but functionally related to tubulin provides a filament along which the *ori* regions and their associated proteins move.

CYTOKINESIS The actual division of a single cell and its contents into two cells begins immediately after chromosome segregation. Initially, there is a pinching in of the cell membrane caused by the contraction of a ring of fibers on the inside surface of the membrane (similar to a drawstring on shorts being tightened). In this case, the major component of these fibers is structurally similar to eukaryotic tubulin (which makes up microtubules), but its function is analogous to that of actin in the contractile ring of an animal cell (see below). As the membrane pinches in, new cell wall materials are deposited, which finally separate the two cells.

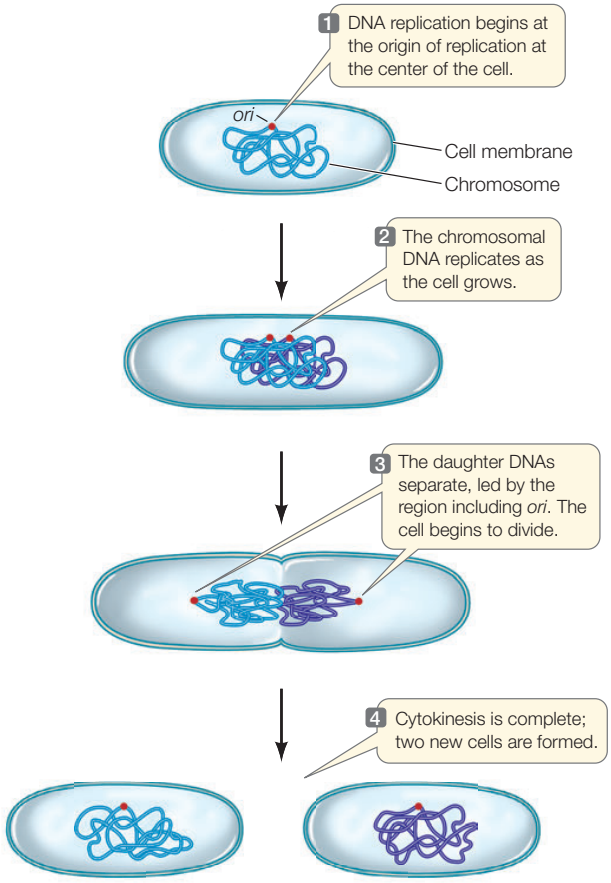


FIGURE 7.4 Prokaryotic Cell Division: Binary Fission The process of cell division in a bacterium involves DNA replication, DNA segregation, and cytokinesis.

Eukaryotic cells divide by mitosis followed by cytokinesis

As in prokaryotes, cell division in eukaryotes entails reproductive signals, DNA replication, DNA segregation, and cytokinesis. Some of the details, however, are quite different:

- *Reproductive signals*. Unlike prokaryotes, eukaryotic cells do not constantly divide whenever environmental conditions are adequate. In fact, most cells in a multicellular organism are specialized and do not divide. In a eukaryotic organism, the signals for cell division are usually not related to the environment of a single cell, but to the function of the entire organism. We will discuss the signals that control eukaryotic cell division in Concept 7.3.
- *DNA replication*. Unlike prokaryotes, eukaryotes have more than one chromosome. But the replication of each eukaryotic DNA molecule is similar to replication in prokaryotes, in that it is achieved by threading the long strands through replication complexes (see Concept 9.2). DNA replication occurs only during a specific stage of the cell cycle.
- *DNA segregation*. This is much more complicated than in prokaryotes, because first, there is a nuclear envelope, and second, there are multiple chromosomes. When a cell divides, one copy of each chromosome must end up in each of the two new cells—for example, each new somatic cell in a human will have all 46 chromosomes. In eukaryotes, the pairs of newly replicated chromosomes are initially attached to one another. They become highly condensed, and then the pairs are pulled apart before segregating into two new nuclei. The cytoskeleton is involved in this process.
- *Cytokinesis*. The process of cytokinesis in plant cells (which have cell walls) is different than in animal cells (which do not have cell walls). We describe both processes below.

LINK

The cytoskeleton is crucial for cell division. Review the description of the cytoskeleton and its molecular components in **Concept 4.4**

These events occur within the context of the **cell cycle**: the period from one cell division to the next. In eukaryotes, the cell cycle can be divided into several stages (**FIGURE 7.5**):

- **Mitosis** is the set of processes in which the chromosomes become condensed and then segregate into two new nuclei.

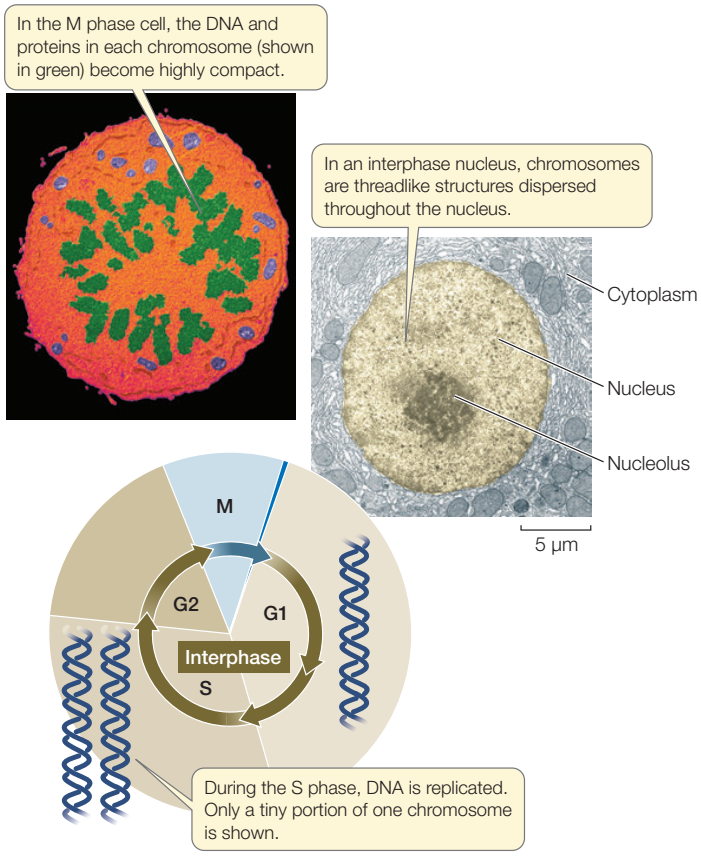


FIGURE 7.5 The Phases of the Eukaryotic Cell Cycle The eukaryotic cell cycle has several phases. DNA in the interphase nucleus is diffuse and becomes compacted as mitosis begins.

- Cytokinesis usually follows immediately after mitosis, and these two stages—mitosis and cytokinesis—are referred to as **M phase**.
- M phase is followed by a much longer period called **interphase**, when the cell nucleus is visible and typical cell functions occur—including DNA replication in cells that are preparing to divide.

Interphase has three subphases called G1, S, and G2 (the *G* stands for gap). **G1** is quite variable, and a cell may spend a long time in this phase carrying out its specialized functions. The cell's DNA is replicated during **S** phase (*S* for synthesis). During **G2**, the cell makes preparations for mitosis—for example, by synthesizing components of the microtubules that will move the segregating chromosomes to opposite ends of the dividing cell.

In mitosis, *a single nucleus gives rise to two daughter nuclei that each contain the same number of chromosomes as the parent nucleus*. Although mitosis is a continuous process in which each event flows smoothly into the next, it is convenient to subdivide it into a series of stages: prophase, prometaphase, metaphase, anaphase, and telophase. Next we will look at these stages in more detail.

Go to ANIMATED TUTORIAL 7.1
Mitosis
PoL2e.com/at7.1

Prophase sets the stage for DNA segregation

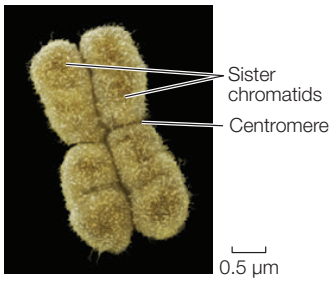
During interphase, only the nuclear envelope and the nucleolus (the region of the nucleus where ribosomes are formed; see Concept 4.3) are visible under the light microscope. The chromatin (the DNA with its associated proteins) is not yet condensed, and individual chromosomes cannot be discerned. The appearance of the nucleus changes as the cell enters **prophase**—the beginning of mitosis. Here we describe three structures that appear during prophase and contribute to the orderly segregation of the replicated DNA: the condensed chromosomes, the reoriented centrosomes, and the spindle.

CONDENSED CHROMOSOMES Before S phase of interphase, each chromosome contains one very long double-stranded DNA molecule. If all of the DNA in a typical human cell were put end to end, it would be nearly 2 meters long. Yet the nucleus is only 5 μm (0.000005 m) in diameter. So even during interphase, eukaryotic DNA is packaged in a highly organized way. The DNA is wound around specific proteins, and other proteins coat the DNA coils. During prophase the chromosomes become much more tightly coiled and condensed.

LINK

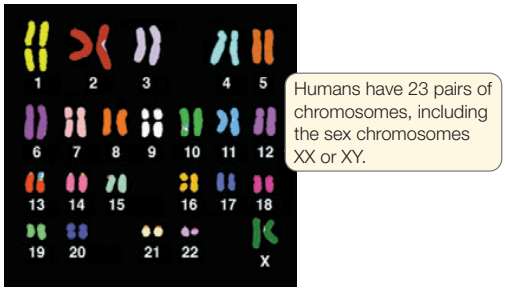
DNA "packaging" and the specialized proteins that accomplish it are described in detail in **Concept 11.3**

After DNA replication, each chromosome has *two* DNA molecules, known as **sister chromatids**. Until they are separated during anaphase (see below), the chromatids are held together at a region called the **centromere**. During prophase the chromosomes become so compact that they can be seen clearly with a light microscope after staining with special dyes. They are even more clearly visualized with an electron microscope:



Specialized protein structures called **kinetochores** assemble on the centromeres, one on each chromatid. These structures are important for chromosome movement.

For a given organism, the number and sizes of the condensed chromosomes constitute the **karyotype**. Each chromosome has a particular length, and the centromere is located at a particular position along its length. For example, humans have 46 chromosomes (23 homologous pairs) that can be distinguished from one another by their sizes and centromere positions. (In the image below, each chromosome is composed of two chromatids, but the individual chromatids cannot be distinguished. The colors are from dyes that were used to help identify the different chromosomes.)



In the past, karyotype analysis was used as a way to identify organisms, and this method is still used to detect chromosomal abnormalities in humans. However, DNA sequence analysis is now much more commonly used to identify individuals and to classify related organisms (see Concept 16.2).

REORIENTED CENTROSOMES Before the spindle apparatus forms (see below), its orientation is determined. In many cells this is accomplished by the **centrosome** ("central body"), an organelle in the cytoplasm near the nucleus. The centrosome consists of a pair of **centrioles**, each one a hollow tube formed by nine triplets of microtubules. During S phase the centrosome becomes duplicated, and at the G2–M transition, the two centrosomes separate from one another, moving to opposite sides of the nucleus. Eventually these identify "poles" toward which chromosomes move during segregation.

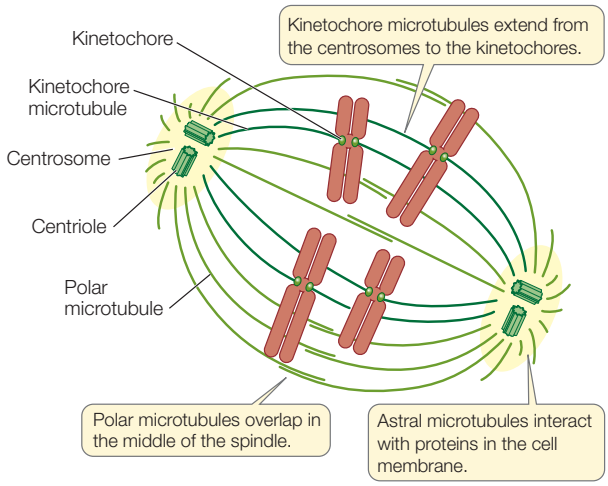
The positions of the centrosomes determine the plane at which the cell divides; therefore they determine the spatial relationship between the two new cells. This relationship may be

of little consequence to single free-living cells such as yeasts, but it is important for development in a multicellular organism. For example, during the development of an embryo, the daughter cells from some divisions must be positioned correctly to receive signals to form new tissues. Plant cells lack centrosomes, but distinct microtubule organizing centers at each end of the cell play the same role.

SPINDLE Each of the two centrosomes, when positioned on opposite sides of the nucleus, serves as a pole toward which the chromosomes move. Tubulin dimers from around the centrosomes aggregate into microtubules that extend from the poles into the middle region of the cell. Together these microtubules make up a **spindle**. The spindle forms during prophase and prometaphase, when the nuclear envelope breaks down. The microtubules are initially unstable, constantly forming and falling apart, until they contact kinetochores or microtubules from the other half-spindle and become more stable.

There are three types of microtubules in the spindle:

- *Polar microtubules* overlap in the middle region of the cell and keep the two poles apart.
- *Astral microtubules* interact with proteins attached to the cell membrane, and also assist in keeping the poles apart.
- *Kinetochore microtubules* attach to the kinetochores on the chromosomes. The two sister chromatids in each chromosome become attached to kinetochore microtubules from opposite sides of the cell. This ensures that the two chromatids will move to opposite poles.



Separation of the chromatids and movement of the **daughter chromosomes** (which the sister chromatids become after separation) is the central feature of mitosis. It accomplishes the DNA segregation that is needed for cell division and completion of the cell cycle. *Note the difference between chromatids and chromosomes:*

- Chromatids share a centromere.
- Chromosomes have their own centromere.

Go to ACTIVITY 7.2 The Mitotic Spindle
PoL2e.com/ac7.2

Chromosome separation and movement are highly organized

During the next three phases of mitosis—prometaphase, metaphase, and anaphase—dramatic changes take place in the cell and the chromosomes (FIGURE 7.6):

- In **prometaphase** the nuclear envelope breaks down and the compacted chromosomes, each consisting of two chromatids, attach to the kinetochore microtubules.
- In **metaphase** the chromosomes line up at the midline of the cell (the equatorial position).
- In **anaphase** the chromatids separate, and the daughter chromosomes move away from each other toward the poles.

The separation of chromatids into daughter chromosomes occurs at the beginning of anaphase. The migration of the daughter chromosomes to the poles of the cell is a highly organized, active process. Two mechanisms operate to move the chromosomes along. First, the kinetochores contain molecular motor proteins, including kinesin and dynein (see Concept 4.4), which use energy from ATP hydrolysis to move the chromosomes along the microtubules. Second, the kinetochore microtubules shorten from the poles, drawing the chromosomes toward the poles.

Telophase occurs after the chromosomes have separated and is the last phase of mitosis. During this period, a nuclear envelope forms around each set of new chromosomes, nucleoli appear, and the chromosomes become less compact. The spindle also disappears at this stage. As a result, there are two new nuclei in a single cell.

Cytokinesis is the division of the cytoplasm

Mitosis refers only to the division of the nucleus. Cytokinesis, the division of the cell's cytoplasm, is the final stage of cell reproduction. This process occurs differently in plants and animals.

ANIMAL CELLS Cytokinesis usually begins with a furrowing of the cell membrane, as if an invisible thread were cinching the cytoplasm between the two nuclei (FIGURE 7.7A). This contractile ring is composed of microfilaments of actin and myosin, which form a ring on the cytoplasmic surface of the cell membrane. These two proteins interact to produce a contraction (just as they do in muscles; described in Concept 33.1), pinching the cell in two. The microfilaments assemble rapidly from actin monomers that are present in the interphase cytoskeleton. Their assembly is controlled by calcium ions (commonly used in cellular signaling) that are released from storage sites in the center of the cell.

PLANT CELLS In plant cells, the cytoplasm divides differently because plants have cell walls. As the spindle breaks down after mitosis, vesicles derived from the Golgi apparatus appear along the plane of cell division, roughly midway between the two daughter nuclei. The vesicles are propelled

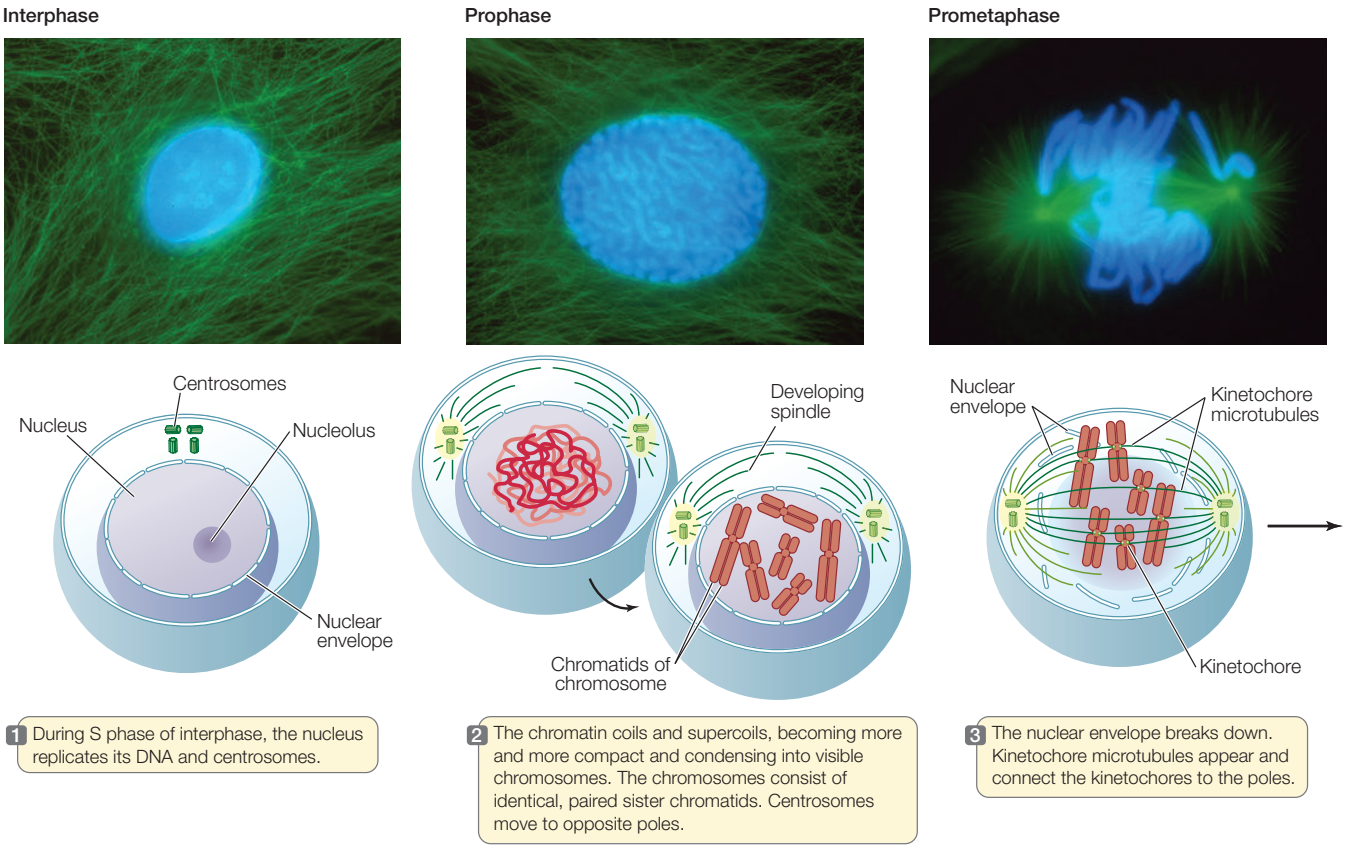


FIGURE 7.6 The Phases of Mitosis Mitosis results in two new nuclei which are genetically identical to each other and to the nucleus from which they were formed. In the micrographs, the green dye stains microtubules (and thus the spindle); the blue dye stains the chromosomes. The chromosomes in the diagrams are stylized to emphasize the fates of the individual chromatids.

Go to ACTIVITY 7.3 Images of Mitosis
PoL2e.com/ac7.3

Go to MEDIA CLIP 7.1
Mitosis: Live and Up Close
PoL2e.com/mc7.1

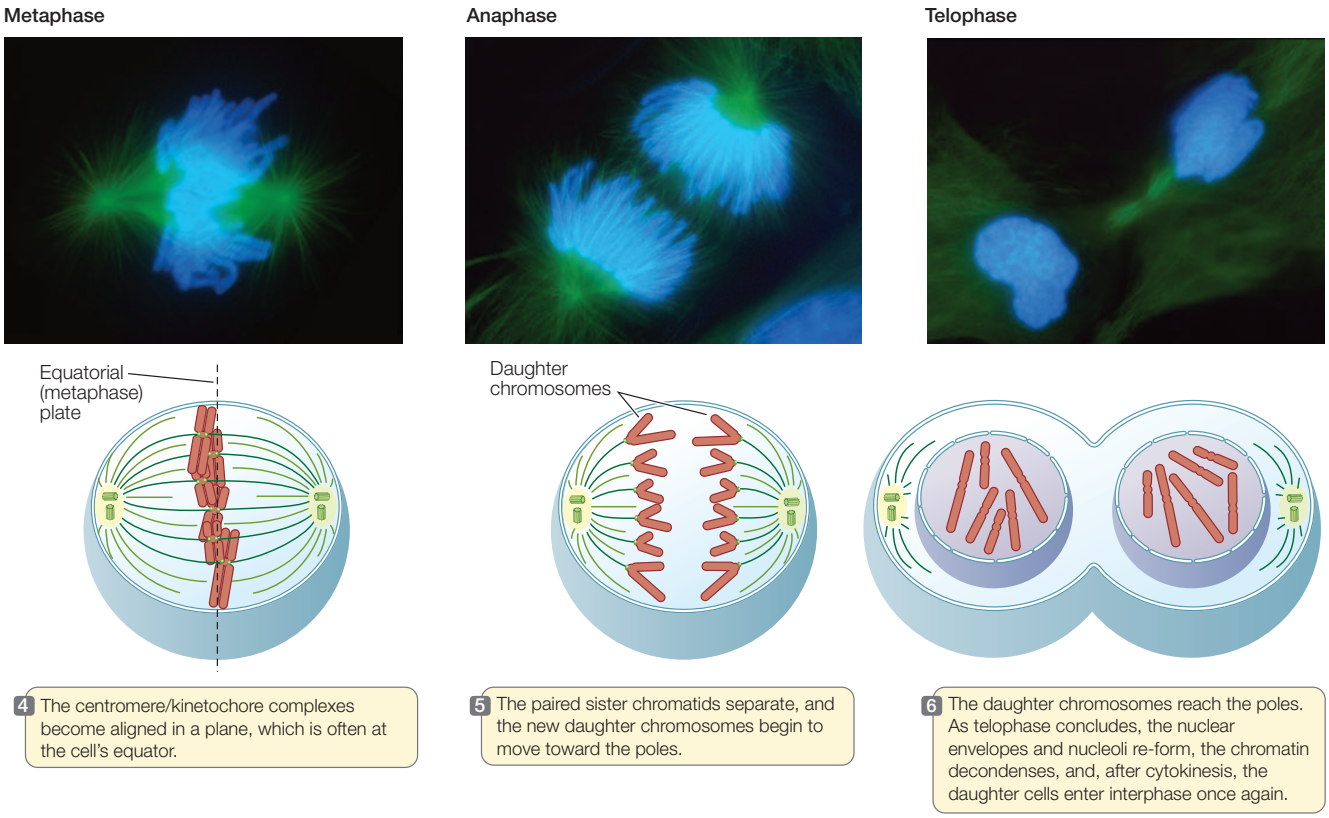
along microtubules by the motor protein kinesin and fuse to form a new cell membrane. At the same time they contribute their contents to a cell plate, which is the beginning of a new cell wall between the two daughter cells (**FIGURE 7.7B**).

Following cytokinesis, each daughter cell contains all the components of a complete cell. A precise distribution of chromosomes is ensured by mitosis. In contrast, organelles such as mitochondria and chloroplasts are not necessarily distributed equally, although at least one of each must be present in each daughter cell. The *orientation* of cell division is important in development (see above), but there does not appear to be a precise mechanism for the distribution of the cytoplasmic contents.

TABLE 7.1 summarizes the major events of the eukaryotic cell cycle.

TABLE 7.1 Summary of Eukaryotic Cell Cycle Events

Phase	Events
INTERPHASE	
G1	Growth; specialized cell functions
S	DNA replication
G2	Spindle synthesis begins; preparation for mitosis
MITOSIS	
Prophase	Condensation of chromosomes; spindle assembly
Prometaphase	Nuclear envelope breakdown; chromosome attachment to spindle
Metaphase	Alignment of chromosomes at equatorial plate
Anaphase	Separation of chromatids; migration to poles
Telophase	Chromosomes decondense; nuclear envelope re-forms
Cytokinesis	Cell separation; cell membrane and/or cell wall formation



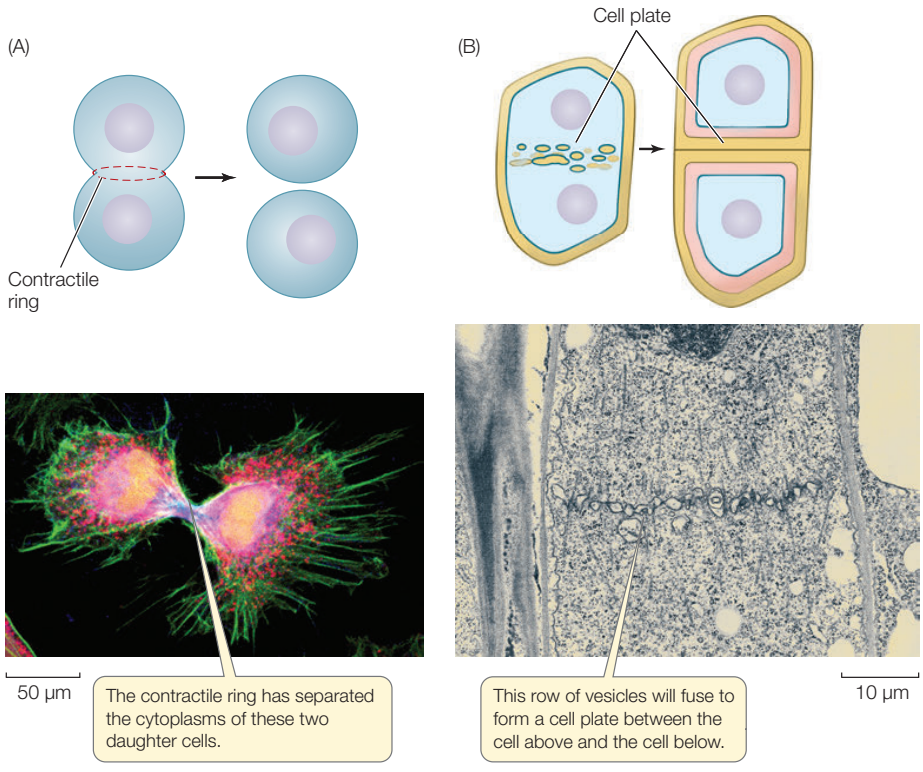


FIGURE 7.7 Cytokinesis Differs in Animal and Plant Cells (A) A HeLa cell (a type of human cancer cell) undergoing cytokinesis. In this flurorescence micrograph, nuclei are yellow, mitochondria are red, and actin filaments are green. (B) An electron micrograph of a plant cell in late telophase. Plant cells divide differently than animal cells because they have cell walls.

CHECKpoint CONCEPT 7.2

- ✓ How does the mitotic spindle ensure that each daughter cell receives a full complement of the genetic material in the cell nucleus?
- ✓ Compare the cell cycles of prokaryotes and eukaryotes with regard to reproductive signals for initiation, how many chromosomes are present, and how the replicated DNA segregates.
- ✓ Sketch the five stages of mitosis for a diploid organism with four chromosomes (two pairs). Clearly label chromosomes and chromatids and note the number of double-stranded DNA molecules in each structure at each stage.
- ✓ The drug cytochalasin B blocks the assembly and function of microfilaments. What would you expect to happen in dividing animal cells treated with this drug after telophase but before cytokinesis?

Light microscopists have studied the dramatic events of mitosis and cell division since the 1880s, and we now have detailed descriptions of these processes. More recently, biologists have focused on the mechanisms controlling cell reproduction, which will be our next topic.

CONCEPT 7.3 Cell Reproduction Is Under Precise Control

Cell reproduction cannot go on continuously and indefinitely. If a single-celled species had no control over its reproduction, it would soon overrun its environment and starve to death. In a multicellular organism, cell reproduction must be controlled to maintain the forms and functions of different parts of the body.

Unlike prokaryotes, eukaryotic cells do not constantly divide whenever environmental conditions are adequate. In fact, the specialized cells of a multicellular eukaryotic organism may seldom or never divide. The signals for eukaryotic cell division are related to the needs of the entire organism. Mammals produce a variety of substances called **growth factors** that stimulate cell division and differentiation. For example, if you cut yourself and bleed, a blood clot eventually forms. Cell fragments called platelets in the blood vessels surrounding the clot secrete various growth factors that stimulate nearby cells to divide and heal the wound.

The eukaryotic cell division cycle is regulated internally

As we discussed in Concept 7.2, the eukaryotic cell cycle can be divided into four stages: G1, S, G2, and M. Progression through these phases is tightly regulated. For example, the **G1–S transition** marks a key decision point for the cell: passing this point (called R, the restriction point) usually means the cell will proceed with the rest of the cell cycle and divide (FIGURE 7.8).

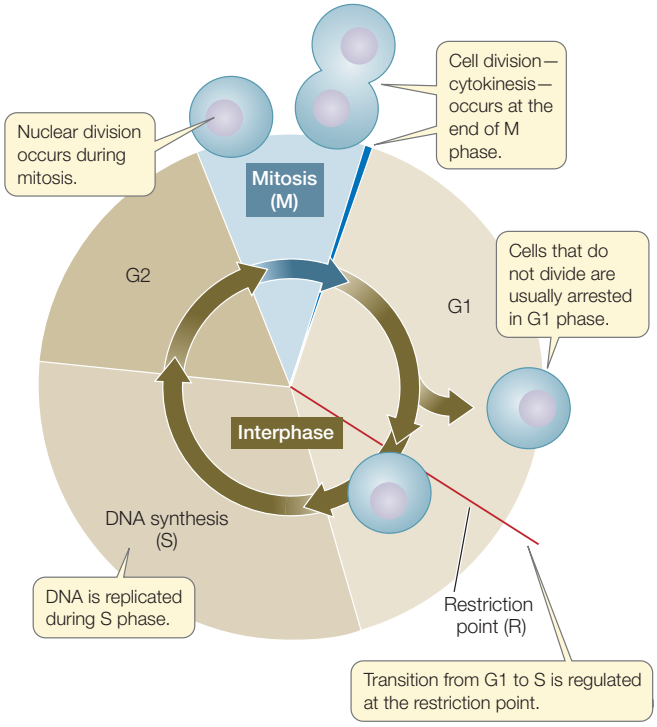


FIGURE 7.8 The Eukaryotic Cell Cycle The cell cycle consists of a mitotic (M) phase, during which mitosis and cytokinesis take place, and a long period of growth known as interphase. Interphase has three subphases (G1, S, and G2) in cells that divide.

What events cause a cell to enter S phase or M phase? A first indication that there were substances controlling these transitions came from cell fusion experiments. For example, an experiment involving the fusion of mammalian cells at G1 phase and S phase showed that a cell in S phase produces a substance that activates DNA replication (FIGURE 7.9). Similar experiments pointed to a molecular activator for entry into M phase.

The cell cycle is controlled by cyclin-dependent kinases

The molecular activators revealed by the cell fusion experiments turned out to be protein kinases, a class of enzymes that are common in cell signaling (see Concept 5.5). The kinases involved in cell cycle regulation are called **cyclin-dependent kinases** (**CDKs**). They catalyze the phosphorylation of target proteins that regulate the cell cycle:

$$\text{Cell cycle regulator} + \text{ATP} \xrightarrow{\text{CDK}} \text{cell cycle regulator-P} + \text{ADP}$$

As their name implies, CDKs are activated by binding to the protein **cyclin**. This binding changes the shape of a CDK such

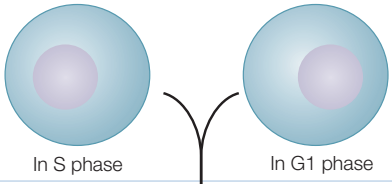
INVESTIGATION

FIGURE 7.9 Regulation of the Cell Cycle Nuclei in G1 do not undergo DNA replication, but nuclei in S phase do. To determine if there is some signal in the S cells that stimulates G1 cells to replicate their DNA, cells in G1 and S phases were induced to fuse, creating cells with both G1 and S properties.[a]

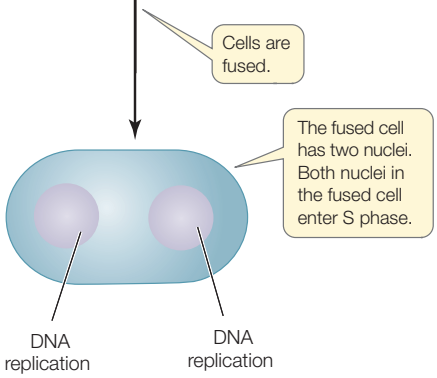
HYPOTHESIS

A cell in S phase contains an activator of DNA replication.

METHOD



RESULTS



CONCLUSION

The S phase cell contains a substance that diffuses to the G1 nucleus and activates DNA replication.

ANALYZE THE DATA

The researchers used mammalian cells undergoing the cell cycle synchronously. Radioactive labeling and microscopy were used to determine which nuclei were synthesizing DNA; only nuclei that were synthesizing DNA became labeled:

Type of cells	Cells with labeled nuclei/total cells
Unfused G1	6/300
Unfused S	435/500
Fused G1 and S cells	17*/19

*Both nuclei labeled

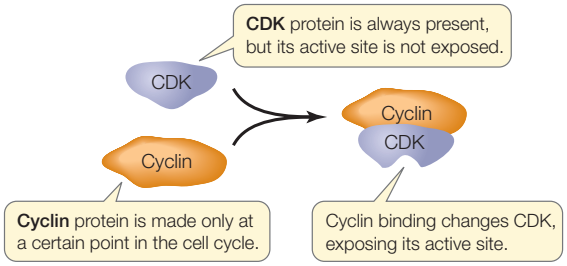
A. What were the percentages of cells in S phase in each of the three experiments?

B. What does this mean in terms of control of the cell cycle?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]P. N. Rao and R. T. Johnson. 1970. *Nature* 225: 159–164.

that its active site is exposed, and is an example of allosteric regulation (see Concept 3.4):



Several different CDKs function at specific stages of the cell cycle, called **cell cycle checkpoints**. At these points, signaling pathways regulate the progress of the cell cycle. For example, if DNA is substantially damaged by radiation or toxic chemicals, the cell may be prevented from successfully completing the cell cycle. So the damage to DNA is repaired before the cycle proceeds. There are three checkpoints during interphase and one during mitosis:

- G1 checkpoint is triggered by DNA damage.
- S checkpoint is triggered by incomplete replication or DNA damage.
- G2 checkpoint is triggered by DNA damage.
- M checkpoint is triggered by a chromosome that fails to attach to the spindle.

Each CDK has its own cyclin to activate it, and the cyclin is made only at the right time. After the CDK acts, the cyclin is broken down by a protease (**FIGURE 7.10**). So a key event controlling the transition from one cell cycle phase to the next is the

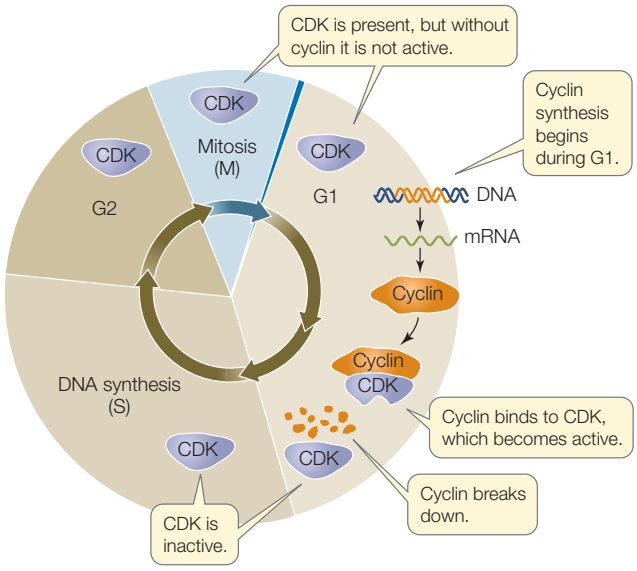


FIGURE 7.10 Cyclins Are Transient in the Cell Cycle Cyclins are made at a particular time and then break down. In this case, the cyclin is present during G1 and activates a CDK at that time.

synthesis and subsequent breakdown of a cyclin. Cyclins are synthesized in response to various molecular signals, including growth factors. This starts a chain reaction:

Growth factor → cyclin synthesis → CDK activation → cell cycle events

To illustrate the concept of cell cycle control by a particular cyclin–CDK complex, let's take a look at the complex that controls the R point at the G1–S transition (see Figure 7.8).

The G1–S cyclin–CDK catalyzes the phosphorylation of a protein called retinoblastoma protein (RB). In many cells, RB or a protein like it acts as an *inhibitor of the cell cycle* at the R point. To begin S phase, a cell must overcome the RB block. Here is where the G1–S cyclin–CDK comes in: it catalyzes the addition of a phosphate to RB. This causes a change in the three-dimensional structure of RB, thereby inactivating it. With RB out of the way, the cell cycle can proceed. To summarize:

$$\text{RB} + \text{ATP} \xrightarrow{\text{G1–S cyclin–CDK}} \text{RB-P} + \text{ADP}$$

(active: blocks cell cycle) (inactive: allows cell cycle)

Now we can be more specific about the chain of events involved in growth factor stimulation of cell division: the specific cyclin whose synthesis is activated is the one that allosterically activates the CDK that phosphorylates RB, and this allows the cell cycle to exit G1 and begin DNA replication in S phase. This example illustrates how regulation of the cell cycle involves a number of cellular processes that we have examined in this and other chapters: signal transduction (see Chapter 5), gene expression and protein synthesis (see Chapter 3), and cell division.

CHECKpoint CONCEPT 7.3

- ✓ Draw a diagram and describe the events that occur in the four stages of the eukaryotic cell cycle (M phase, G1, S, and G2).
- ✓ Cultures of eukaryotic cells can be synchronized, so they are all at the same phase of the cell cycle at the same time. If you examined a culture at the beginning of G1, would the CDK that acts at the R point be present? Would it be active? Would its cyclin be present? What would your answers be if the culture were at the R point?

Binary fission and mitosis result in daughter cells with the same number of chromosomes as their parent cells. Sexual reproduction, however, requires a process of cell division in which the number of chromosomes is halved. We'll look at this process next.

CONCEPT 7.4 Meiosis Halves the Nuclear Chromosome Content and Generates Diversity

In Concept 7.1 we described the role and importance of meiosis in sexual reproduction. Now we will see how the orderly and precise generation of haploid cells is accomplished.

Meiosis consists of *two* nuclear divisions that reduce the number of chromosomes to the haploid number. Although the nucleus divides twice during meiosis, the DNA is replicated only once. Unlike the products of mitosis, the haploid cells produced by meiosis are genetically different from one another and from the parent cell. **FIGURE 7.11** compares the two processes.

To understand the process of meiosis and its specific details, it is useful to keep in mind the overall functions that meiosis has evolved to serve:

- To reduce the chromosome number from diploid to haploid
- To ensure that each of the haploid products has a complete set of chromosomes
- To generate genetic diversity among the products (gametes)

The events of meiosis are illustrated in **FIGURE 7.12**. In the rest of this section we will discuss some of the key features that distinguish meiosis from mitosis.

Go to ANIMATED TUTORIAL 7.2
Meiosis
PoL2e.com/at7.2

Meiotic division reduces the chromosome number

As noted above, meiosis consists of two nuclear divisions, meiosis I and meiosis II. Two unique features characterize **meiosis I**:

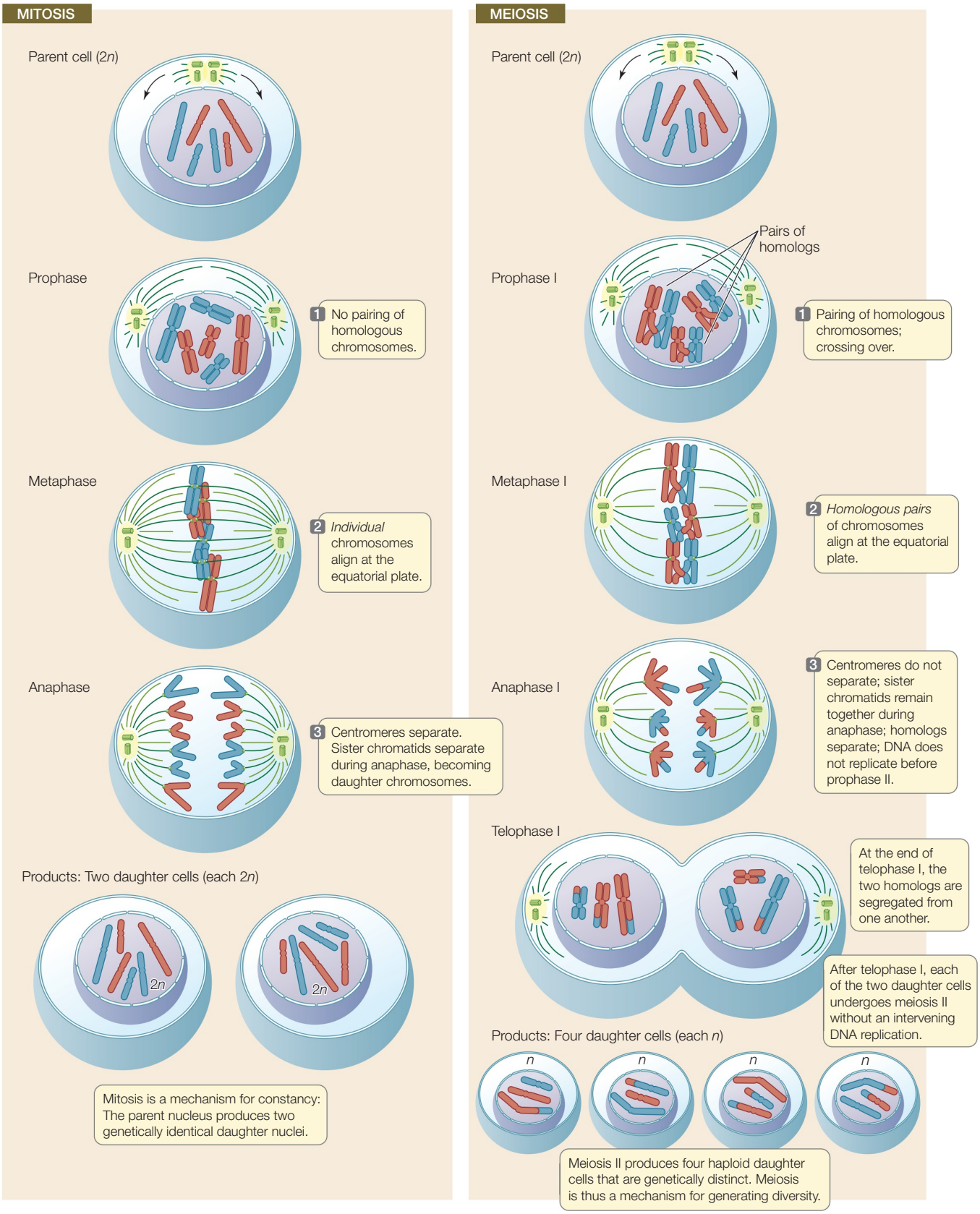
- *Homologous chromosomes come together and line up* along their entire lengths. No such pairing occurs in mitosis.
- *The homologous chromosome pairs separate*, but the individual chromosomes, each consisting of two sister chromatids, remain intact. (The chromatids will separate during meiosis II.)

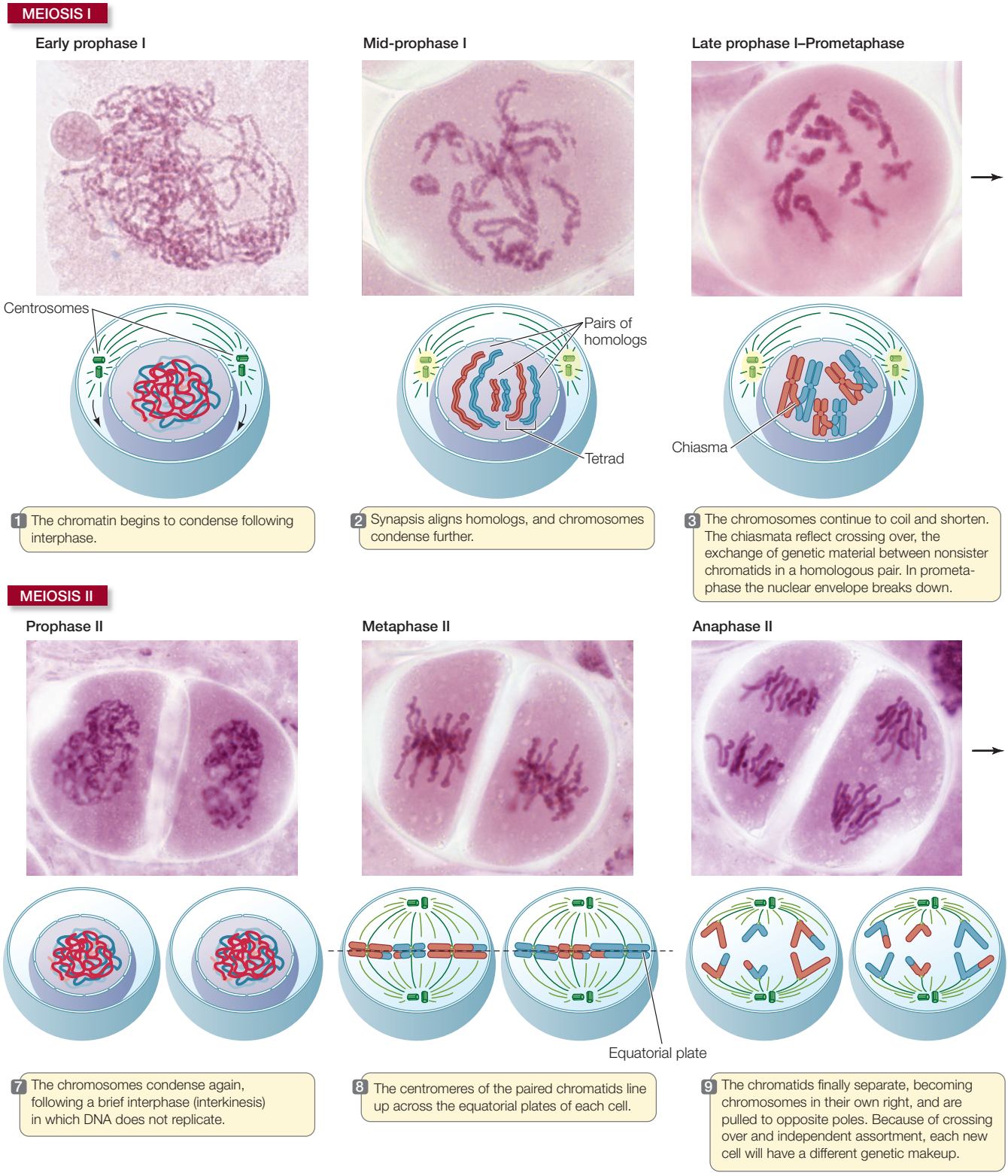
Like mitosis, meiosis I is preceded by an interphase with an S phase, during which each chromosome is replicated. As a result, each chromosome consists of two sister chromatids. At the end of meiosis I two nuclei form, each with half of the original chromosomes (one member of each homologous pair). Since the centromeres did not separate, these chromosomes are still double—composed of two sister chromatids. The sister chromatids are separated during **meiosis II**, which is *not* preceded by DNA replication. As a result, the products of meiosis I and II are four cells, each containing the haploid number of chromosomes. But *these four cells are not genetically identical*.

Crossing over and independent assortment generate diversity

A diploid organism has two sets of chromosomes ($2n$): one set derived from its male parent, the other from its female parent. As the organism grows and develops, its cells undergo mitotic

FIGURE 7.11 Mitosis and Meiosis: A Comparison Meiosis involves two cell divisions, the first of which is very different from the single division of mitosis. Meiosis II is similar to mitosis, in that the centromeres separate during anaphase, allowing the chromatids of the two homologous pairs to separate into four daughter chromosomes that are genetically distinct from the parental chromosomes. ▶


MITOSIS
Parent cell (2n)
Prophase
1 No pairing of homologous chromosomes.
Metaphase
2 Individual chromosomes align at the equatorial plate.
Anaphase
3 Centromeres separate. Sister chromatids separate during anaphase, becoming daughter chromosomes.
Products: Two daughter cells (each 2n)
2n
2n
Mitosis is a mechanism for constancy: The parent nucleus produces two genetically identical daughter nuclei.
MEIOSIS
Parent cell (2n)
Pairs of homologs
Prophase I
1 Pairing of homologous chromosomes; crossing over.
Metaphase I
2 Homologous pairs of chromosomes align at the equatorial plate.
Anaphase I
3 Centromeres do not separate; sister chromatids remain together during anaphase; homologs separate; DNA does not replicate before prophase II.
Telophase I
At the end of telophase I, the two homologs are segregated from one another.
After telophase I, each of the two daughter cells undergoes meiosis II without an intervening DNA replication.
Products: Four daughter cells (each n)
n
n
n
n
Meiosis II produces four haploid daughter cells that are genetically distinct. Meiosis is thus a mechanism for generating diversity.



divisions. In mitosis, each chromosome behaves independently of its homolog, and its two chromatids are sent to opposite poles during anaphase. Each daughter nucleus ends up with an identical set of $2n$ chromosomes. In meiosis, things are very different (see Figure 7.11).

An important consequence of meiosis is that the four resulting cells differ from one another genetically. The shuffling of genetic material occurs by two processes: crossing over and independent assortment.

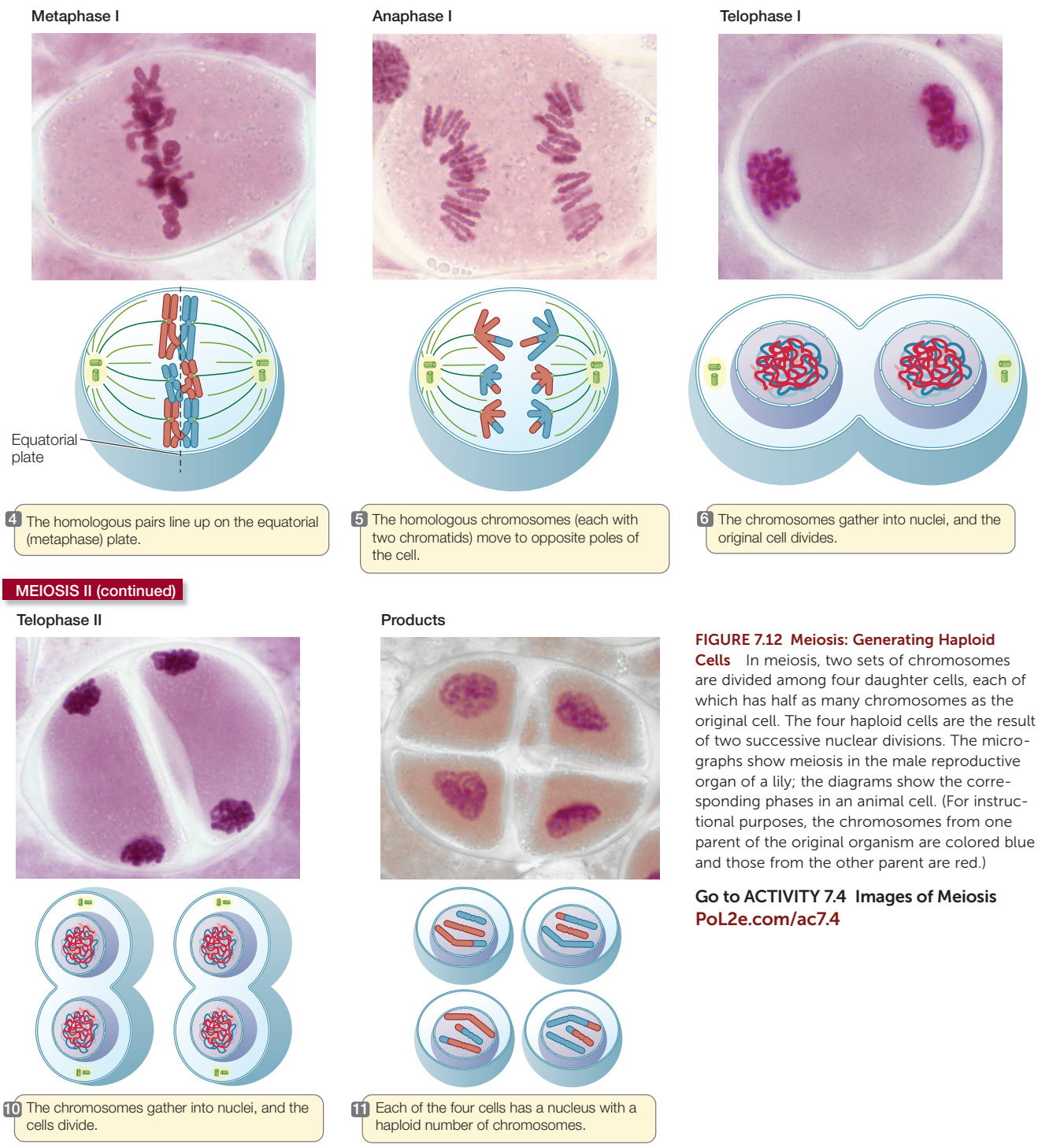


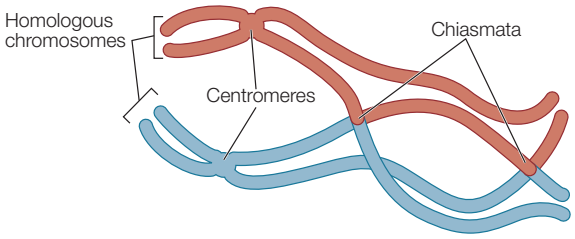
FIGURE 7.12 Meiosis: Generating Haploid Cells In meiosis, two sets of chromosomes are divided among four daughter cells, each of which has half as many chromosomes as the original cell. The four haploid cells are the result of two successive nuclear divisions. The micrographs show meiosis in the male reproductive organ of a lily; the diagrams show the corresponding phases in an animal cell. (For instructional purposes, the chromosomes from one parent of the original organism are colored blue and those from the other parent are red.)

Go to ACTIVITY 7.4 Images of Meiosis
PoL2e.com/ac7.4

CROSSING OVER Meiosis I begins with a long prophase I (the first three panels of Figure 7.12), during which the chromosomes change markedly. The homologous chromosomes pair by adhering along their lengths in a process called synapsis. (This does not happen in mitosis.) This pairing process lasts from prophase I to the end of metaphase I. The four chromatids of each pair of homologous chromosomes form a **tetrad**, or bivalent. For example, in a human cell at the end of prophase I there are 23 tetrads, each consisting of four chromatids. The four chromatids come from the two partners in each homologous pair of chromosomes.

Throughout prophase I and metaphase I, the chromatin continues to coil and compact and the chromosomes become more condensed. At a certain point, the homologous chromosome

pairs appear to repel each other, especially near the centromeres, but they remain attached. The X-shaped attachment points are called **chiasmata** (singular chiasma, "cross"):



A chiasma is a point where genetic material is exchanged between nonsister chromatids on homologous chromosomes—a process called **crossing over** (FIGURE 7.13). Any of the four chromatids in the tetrad can participate in this exchange, and a single chromatid can exchange material at more than one point along its length. Crossing over occurs shortly after synapsis begins, but chiasmata do not become visible until later, when the homologs are repelling each other. Crossing over results in **recombinant** chromatids, and it increases genetic variation among the products of meiosis by reshuffling genetic information between homologous chromosome pairs. In Concept 8.3 we will explore further the genetic consequences of crossing over.

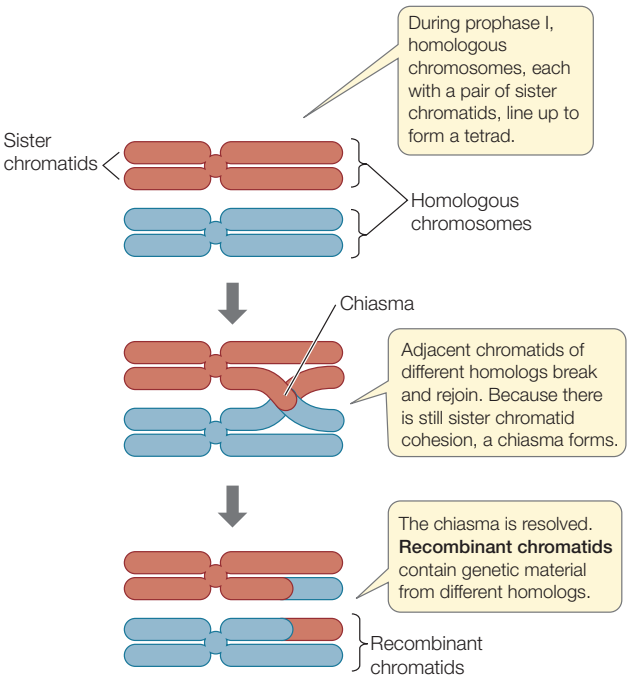
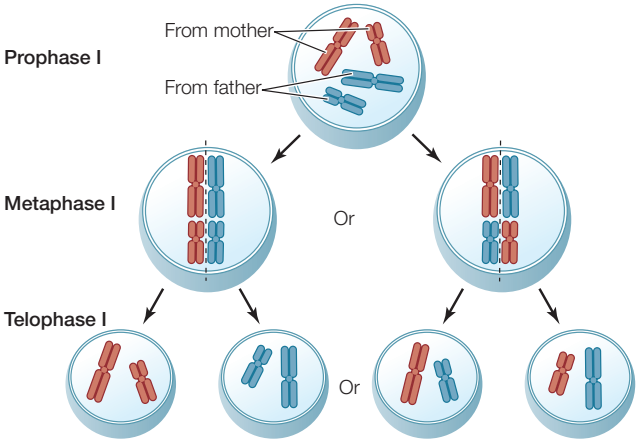


FIGURE 7.13 Crossing Over Forms Genetically Diverse Chromosomes The exchange of genetic material by crossing over results in new combinations of genetic information on the recombinant chromosomes. The two different colors distinguish the chromosomes contributed by the male and female parents of the organism whose cell is undergoing meiosis.

Mitosis seldom takes more than an hour or two, but meiosis can take *much* longer. In human males, the cells in the testis that undergo meiosis take about a week for prophase I and about a month for the entire meiotic cycle. In females, prophase I begins long before a woman's birth, during her early fetal development. Meiosis continues as much as decades later, during the monthly ovarian cycle, and is completed only after fertilization.

INDEPENDENT ASSORTMENT In addition to crossing over, meiosis provides a second source of genetic diversity. It is a matter of chance which member of a homologous pair goes to which daughter cell at anaphase I. For example, consider a diploid organism with two pairs of homologous chromosomes (pairs 1 and 2). One member of each pair came from the male parent of the organism (paternal 1 and 2), and the other came from the female parent (maternal 1 and 2). When cells in this organism undergo meiosis, a particular daughter nucleus could receive paternal 1 and maternal 2, paternal 2 and maternal 1, both maternal, or both paternal chromosomes. It all depends on how the homologous pairs line up at metaphase I. This phenomenon is called **independent assortment**.



Note that of the four possible outcomes in the figure above, only two daughter nuclei receive either all maternal or all paternal chromosomes (apart from material exchanged by crossing over). The greater the number of chromosomes, the lower the probability of reestablishing the original parental combinations, and thus the greater the potential for genetic diversity. Most species of diploid organisms have more than two pairs of chromosomes. In humans, with 23 chromosome pairs, 2^{23} (8,388,608) different combinations of maternal and paternal chromosomes can be produced just by the mechanism of independent assortment! Taking into account the extra genetic shuffling afforded by crossing over, the number of possible combinations is virtually infinite. Crossing over and independent assortment, along with the processes that result in mutations, provide the genetic diversity needed for evolution by natural selection.

We have seen how meiosis I is fundamentally different from mitosis. However, meiosis II is similar to mitosis in that it involves the separation of chromatids into daughter nuclei (see steps 7–11 in Figure 7.12). The final products of meiosis I and meiosis II are four haploid daughter cells, each with one set (n) of chromosomes.

Meiotic errors lead to abnormal chromosome structures and numbers

Meiosis is a complex process, and things occasionally go wrong. For example, chromosomes may break, homologs may fail to separate at anaphase I, or chromatids may fail to separate at anaphase II. The gametes formed from meiotic errors carry abnormal chromosomes, and when abnormal chromosomes take part in fertilization, the consequences for offspring can be significant.

NONDISJUNCTION Occasionally a homologous chromosome pair fails to separate (fails to "disjoin") at anaphase I, or a pair of chromatids fail to separate at anaphase II. This failure to separate is referred to as **nondisjunction**. If a chromosome pair fails to separate at anaphase I, two of the four daughter nuclei will each end up with both members of that homologous pair, and the other two will have neither member of the pair. If nondisjunction occurs at anaphase II, only two of the four daughter nuclei will be affected: one will have an extra chromosome and the other will have one less than the full complement of chromosomes.

Using humans as an example, if during anaphase I the two homologs of chromosome 10 fail to separate, half the gametes will have two copies of chromosome 10, with a total of 24 chromosomes instead of 23. If one of these gametes fuses with a normal gamete during fertilization, the zygote will have 47 (23 + 24) chromosomes, with three copies of chromosome 10. The condition of having an abnormal number of chromosomes is called **aneuploidy**; having one extra chromosome is called trisomy, and missing one chromosome is called monosomy (FIGURE 7.14).

For reasons that are unclear, aneuploidy is a common and harmful condition in humans. About 10–30 percent of all conceptions show aneuploidy, but most of the embryos that develop from such zygotes do not survive to birth, and those that do often die before the age of 1 year. At least one-fifth of all recognized human pregnancies are spontaneously terminated (miscarried) during the first 2 months, largely because of trisomies and monosomies. The actual proportion of spontaneously terminated pregnancies is certainly higher, because the earliest ones often go unrecognized. The most common form of aneuploidy in humans is trisomy 16 (three copies of chromosome 16), but almost none of these embryos survive to birth. Among the few aneuploidies that allow survival is Down syndrome—trisomy 21. Such individuals generally have intellectual disabilities but can lead long and productive lives.

POLYPLOIDY Most organisms are either diploid (for example, most animals) or haploid (for example, most fungi). Under some circumstances, triploid ($3n$), tetraploid ($4n$), or higher-order **polyploid** nuclei may form. This can occur in a variety of ways. For example, there could be an extra round of DNA replication preceding meiosis, or there could be no spindle formed in meiosis II. Polyploidy occurs naturally in some animal species and in many plants.

LINK

Polyploidy can lead to reproductive isolation (the inability of two individuals to produce fertile offspring) and has probably led to speciation—the evolution of new species—as described in **Concept 17.3**

A diploid nucleus can undergo normal meiosis because there are two sets of chromosomes to make up homologous pairs, which separate during anaphase I. Similarly, a tetraploid nucleus has an even number of each kind of chromosome, so each chromosome can pair with its homolog. However, a triploid nucleus cannot undergo normal meiosis because one-third of the chromosomes would lack partners. Polyploidy has implications for agriculture, particularly in the production of hybrid plants. For example, ploidy (the number of chromosomes in the nucleus) must be taken into account in wheat breeding because

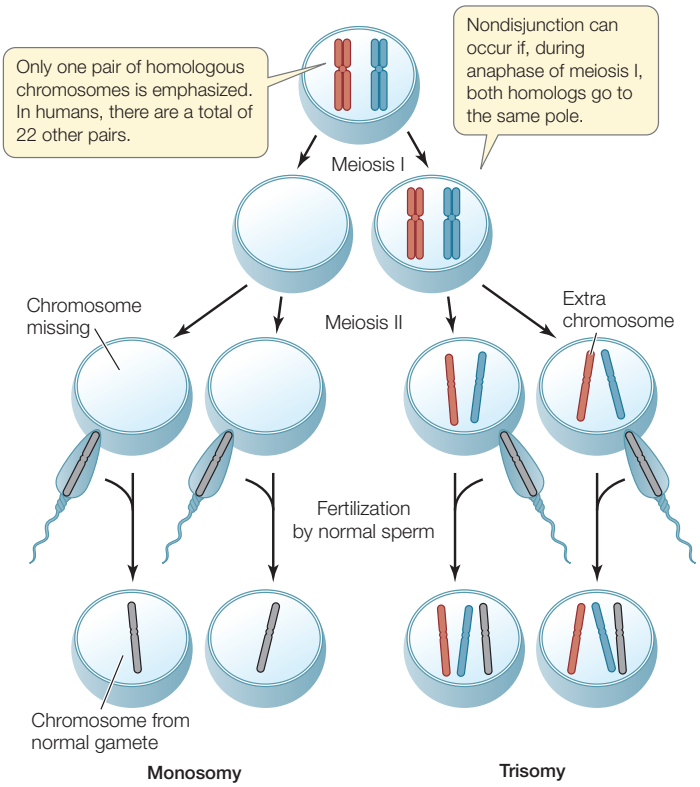
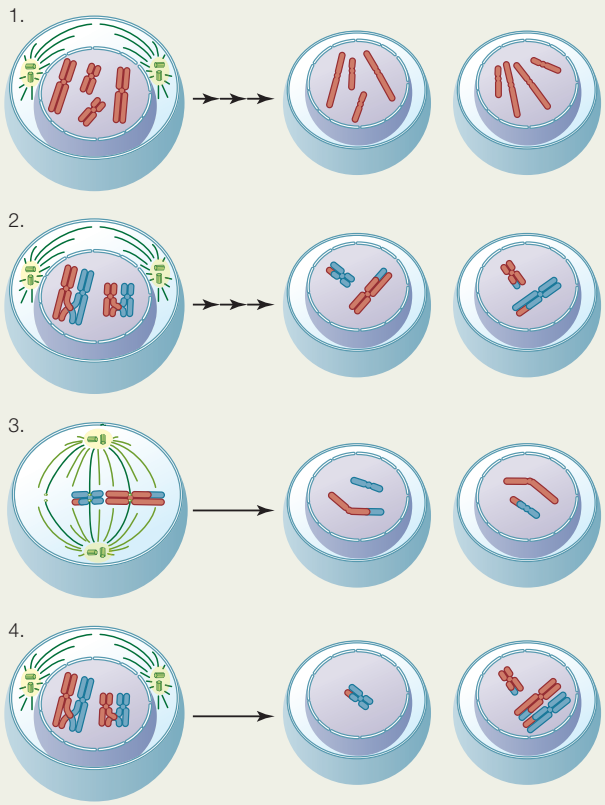


FIGURE 7.14 Nondisjunction Leads to Aneuploidy Nondisjunction, shown here occurring in meiosis I, results in aneuploidy: one or more chromosomes are either lacking or present in excess. Generally, aneuploidy is lethal to a developing embryo.

APPLY THE CONCEPT

Meiosis halves the nuclear chromosome content and generates diversity

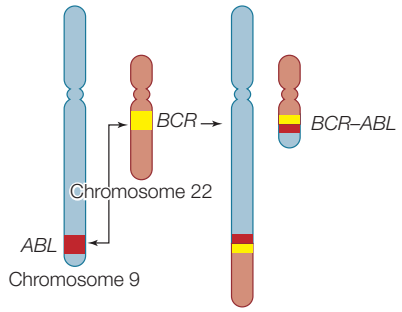
Cells from a diploid organism ($2n = 4$) are shown undergoing division in the diagrams. For each diagram, indicate the type of cell division (mitosis or meiosis), the phase of division, and any special condition that is depicted.



there are diploid, tetraploid, and hexaploid wheat varieties. Polyploidy can be a desirable trait in crops and ornamental plants because it often leads to more robust plants with larger flowers, fruits, and seeds. In addition, triploid fruit varieties are desirable because they are infertile and therefore seedless.

TRANSLOCATION During crossing over in meiosis I, chromatids from homologous chromosome pairs break and rejoin. Occasionally this can happen between *non-homologous chromosomes*. The result is a **translocation**, and these are quite common, even in mitotic cells. As we will point out in our discussion of gene expression and its regulation in Chapters 10 and 11, the location of genes relative to other DNA sequences is important, and translocations can have profound effects on gene expression.

An example of a translocation known to occur in humans is a swap of material between chromosomes 9 and 22:



In this case, part of the *BCR* gene sequence on chromosome 22 comes to lie adjacent to part of the *ABL* gene sequence, which was translocated from chromosome 9. If the translocation occurs in a mitotic cell forming white blood cells, the result of this combination is a form of leukemia, a cancer of white blood cells.

A translocation that occurs during meiosis may be carried on the gametes that result and passed on to offspring at fertilization.

CHECKpoint CONCEPT 7.4

- ✓ How do crossing over and independent assortment during meiosis result in daughter nuclei that differ genetically?
- ✓ What are the differences between meiosis and mitosis?
- ✓ A vertebrate animal has a diploid number of 6. How many chromosomes are present in the following cells: A gamete? A gamete with monosomy of chromosome 2? A skin cell? A sperm cell at meiotic anaphase II?

An essential role of cell division in complex eukaryotes is to replace cells that die. What causes cells to die?

CONCEPT 7.5 Programmed Cell Death Is a Necessary Process in Living Organisms

Cells die in one of two ways. The first type of cell death, **necrosis**, occurs when cells are damaged by mechanical means or toxins, or are starved of oxygen or nutrients. These cells often swell up and burst, releasing their contents into the extracellular environment. This process often results in inflammation (see Concept 39.1).

More typically, cell death is due to **apoptosis** (Greek, "falling apart"). Apoptosis is a genetically programmed series of events that result in cell death. Why would a cell initiate apoptosis, which is essentially cell suicide? In animals, there are two possible reasons:

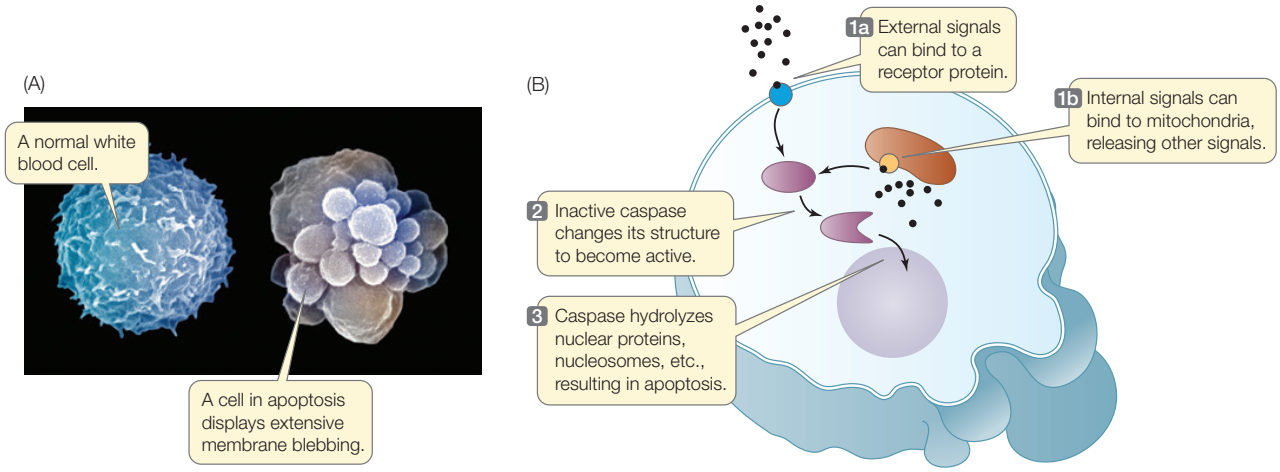


FIGURE 7.15 Apoptosis: Programmed Cell Death (A) Many cells are programmed to "self-destruct" when they are no longer needed, or when they have lived long enough to accumulate a burden of DNA damage that might harm the organism. (B) Both external and internal signals stimulate caspases (or similar enzymes in plants), which break down specific cell constituents, resulting in apoptosis.

- *The cell is no longer needed by the organism*. For example, before birth, a human fetus has weblike hands, with connective tissue between the fingers. As development proceeds, this unneeded tissue disappears as the cells undergo apoptosis in response to specific signals.
- *The longer cells live, the more prone they are to genetic damage that could lead to cancer*. This is especially true of epithelial cells on the surface of an organism, which may be exposed to radiation or toxic substances. Such cells normally die after only days or weeks and are replaced by new cells.

The events of apoptosis are similar in many organisms. The cell becomes detached from its neighbors, it hydrolyzes its DNA into small fragments, and forms membranous lobes, or "blebs," that break up into cell fragments (**FIGURE 7.15A**). In a remarkable example of the economy of nature, the surrounding living cells usually ingest the remains of the dead cell by phagocytosis. The remains are digested in the lysosomes, and the digestion products are recycled.

Apoptosis is also used by plant cells in an important defense mechanism called the hypersensitive response. Plants can protect themselves from disease by undergoing apoptosis at the site of infection by a fungus or bacterium. With no living tissue to grow in, the invading organism is not able to spread to other parts of the plant. Because of their rigid cell walls, plant cells do not form blebs the way animal cells do. Instead, they digest their own cell contents in the vacuole and then release the digested components into the vascular system.

Despite these differences between plant and animal cells, they share many of the signal transduction pathways that lead to apoptosis. Like the cell division cycle, programmed cell death is controlled by signals, which may come from inside

APPLY THE CONCEPT

Programmed cell death is a necessary process in living organisms

The DNA content of an individual cell can be measured by applying a DNA-specific dye to the cell and then passing it through an instrument that measures the staining intensity. A new drug was tested on a population of rapidly dividing tumor cells, and the DNA contents of the treated cells were analyzed and compared with those of untreated cells:[a]

DYE INTENSITY	% UNTREATED CELLS	% TREATED CELLS
<10	0	20
10	10	5
20	55	60
30	5	5
40	30	10

1. Plot percentage of cells versus dye intensity for the untreated and treated cells.
2. Explain the data for the untreated cells. Which cells are in G1? What do the data indicate about how much time cells spend in G1 relative to other phases?
3. Explain the data for treated cells and compare them with untreated cells. At what stage of the cell cycle do you think the new drug acts?

[a] Author's own, unpublished data.

or outside the cell. Internal signals may be linked to the age of the cell or the recognition of damaged DNA. External signals can be detected by receptors in the cell membrane, and in turn they activate signal transduction pathways. Both internal and external signals lead to the activation of a class of enzymes called **caspases** in animals or of a functionally similar class of enzymes in plants. These enzymes hydrolyze target proteins in a cascade of events. The cell dies as the caspases hydrolyze proteins of the nuclear envelope, nucleosomes, and cell membrane (**FIGURE 7.15B**).

CHECKpoint CONCEPT 7.5

- ✓ What are some differences between apoptosis and necrosis?
- ✓ Give examples of situations in which apoptosis occurs in animals and in plants.
- ✓ In the worm *Caenorhabditis elegans* the fertilized egg divides by mitosis to produce 1,090 somatic cells. But the adult worm has only 959 cells. What happens to the 131 other cells formed during worm embryo development? What might happen if the 131 cells did not undergo this process?

Q How does infection with HPV result in uncontrolled cell reproduction?

ANSWER Human papillomavirus (HPV) stimulates the cell cycle when it infects tissues lining the cervix. It does this by "hijacking" the regulatory mechanisms that control the cell cycle (Concept 7.3). There are two types of proteins that regulate the cell cycle:

- **Oncogene** proteins are positive regulators of the cell cycle in cancer cells. They are derived from normal positive regulators that have become mutated to be overly active, or that are present in excess, and they stimulate the cancer cells to divide more often than normal cells. An example of an oncogene protein is the growth factor receptor in a breast cancer cell (**FIGURE 7.16A**). Normal breast cells have relatively low numbers of the growth factor receptor *h*uman epidermal growth factor *r*eceptor 2 (HER2). So the growth factor does not normally find many receptors with which to bind and initiate cell division. In about 25 percent of breast cancers, a DNA change results in the increased production of HER2. This results in positive stimulation of the cell cycle, and a rapid proliferation of cells with the altered DNA.
- **Tumor suppressors** are negative regulators of the cell cycle in normal cells, but in cancer cells they are inactive. An example is the retinoblastoma (RB) protein that acts at R (the restriction point) in G1 (see Figure 7.8). When RB is active the cell cycle does not proceed, but it is inactive in cancer cells, allowing the cell cycle to occur (**FIGURE 7.16B**). *This is where HPV hijacks the system*. When it infects cells lining the cervix, HPV causes the synthesis of a protein called E7, which has a three-dimensional shape that just fits into the protein-binding site of RB, thereby inactivating it. With no active RB to prevent it, cell division proceeds. Uncontrolled cell reproduction is a hallmark of cancer—and so cervical cancer begins.

Most tumors are treated by surgery. But when a tumor has spread from its original site (a common occurrence, unfortunately), surgery does not cure it. Instead, drugs—chemotherapy—are used. Generally, these drugs stop cell division by targeting specific cell cycle events (Concepts 7.2 and 7.3). For example, some drugs block DNA replication (e.g., 5-fluorouracil); others damage DNA, stopping the cells at G2 (e.g., etoposide); and still others prevent the normal functioning of the mitotic spindle (e.g., paclitaxel). Many of these drugs do not kill the cell, but they cause the cell cycle to stop, and the damaged cell is stimulated to undergo apoptosis (Concept 7.5).

A major problem with these treatments is that they target normal cells as well as the tumor cells. They are toxic to

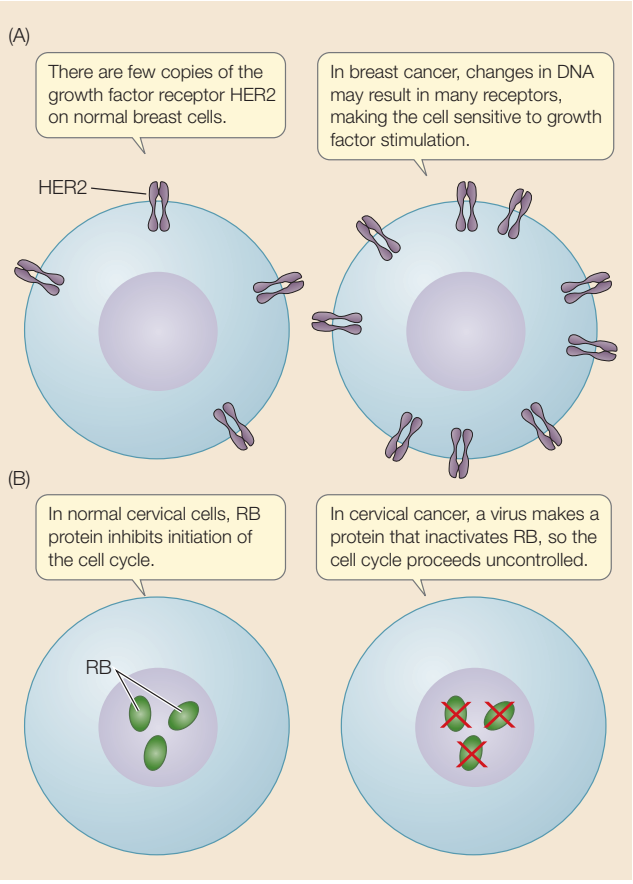


FIGURE 7.16 Molecular Changes Regulate the Cell Cycle in Cancer Cells In cancer cells, oncogene proteins become active (A) and tumor suppressor proteins become inactive (B).

tissues with large populations of normally dividing cells such as those in the intestine, skin, and bone marrow (producing blood cells). There is an ongoing search for better and more specific drugs. For example, a drug has been identified that affects the protein produced as a result of the translocation between chromosomes 9 and 22 (Concept 7.4). The drug is rather specific and has been very successful at treating leukemia caused by this translocation.

In this chapter we examined the cell cycle and cell division by binary fission and mitosis. We have seen how the normal cell cycle is disrupted in cancer. We also examined meiosis and the production of haploid cells in sexual life cycles. In the coming chapters we will examine heredity, genes, and DNA. In Concept 8.1 we will discuss Gregor Mendel's studies of heredity and how the enormous power of his discoveries founded the science of genetics.

SUMMARY

CONCEPT 7.1 Different Life Cycles Use Different Modes of Cell Reproduction

- **Cell division** is necessary for the reproduction, growth, and repair of organisms. **Review Figure 7.1**
- **Asexual reproduction** produces **clones**, new organisms that are virtually identical genetically to the parent. Any genetic variation is the result of mutations.
- In **sexual reproduction**, two **haploid** gametes—usually one from each parent—unite in **fertilization** to form a genetically unique **diploid zygote**. There are many different sexual life cycles that can be haplontic, diplontic, or involve alternation of generations. **Review Figure 7.3 and ACTIVITY 7.1**
- Diploid cells contain **homologous pairs** of chromosomes. In sexually reproducing organisms, certain cells undergo **meiosis**, a process of cell division in which the chromosome number is halved. Each of the haploid daughter cells contains one member of each homologous pair of chromosomes.

CONCEPT 7.2 Both Binary Fission and Mitosis Produce Genetically Identical Cells

- Cell division must be initiated by a reproductive signal. Before a cell can divide, the genetic material (DNA) must undergo **replication** and **segregation** to separate portions of the cell. **Cytokinesis** then divides the cytoplasm into two cells.
- In prokaryotes, most cellular DNA is a single molecule, usually in the form of a circular chromosome. Prokaryotes reproduce by **binary fission**. **Review Figure 7.4**
- During most of the eukaryotic cell cycle, the cell is in **interphase**, which is divided into three subphases: **G1**, **S**, and **G2**. DNA is replicated during S phase. **Mitosis** (**M phase**) and cytokinesis follow. **Review Figure 7.5 and ANIMATED TUTORIAL 7.1**
- In mitosis, a single nucleus gives rise to two nuclei that are genetically identical to each other and to the parent nucleus.
- At mitosis, the replicated chromosomes, called **sister chromatids**, are held together at the **centromere**. Each chromatid contains one double-stranded DNA molecule. During mitosis, sister chromatids line up at the equatorial plate and attach to the **spindle**. **Review ACTIVITY 7.2**
- Mitosis can be divided into several phases called **prophase**, **prometaphase**, **metaphase**, **anaphase**, and **telophase**. **Review Figure 7.6 and ACTIVITY 7.3**
- Nuclear division is usually followed by cytokinesis. Animal cell cytoplasms divide via a contractile ring made up of actin microfilaments. In plant cells, cytokinesis is accomplished by vesicles that fuse to form a cell plate. **Review Figure 7.7**

CONCEPT 7.3 Cell Reproduction Is Under Precise Control

- Interactions between **cyclins** and **CDKs** regulate the passage of cells through checkpoints in the cell cycle. External controls such as **growth factors** can stimulate the cell to begin a division cycle. **Review Figure 7.10**

CONCEPT 7.4 Meiosis Halves the Nuclear Chromosome Content and Generates Diversity

- Meiosis consists of two nuclear divisions, **meiosis I** and **meiosis II**, which collectively reduce the chromosome number from diploid to haploid. Meiosis results in four genetically diverse haploid cells, often gametes. **Review ANIMATED TUTORIAL 7.2**
- In meiosis I, entire chromosomes, each with two chromatids, migrate to the poles. In meiosis II, the sister chromatids separate. **Review Figures 7.11 and 7.12 and ACTIVITY 7.4**
- During prophase I, homologous chromosomes undergo synapsis to form pairs in a **tetrad**. Chromatids can form junctions called chiasmata, and genetic material may be exchanged between the two homologs by **crossing over**. **Review Figure 7.13**
- Both crossing over during prophase I and **independent assortment** of the homologs as they separate during anaphase I ensure that gametes are genetically diverse.
- Meiotic errors can result in abnormal numbers of chromosomes in the resulting gametes and offspring. **Review Figure 7.14**

CONCEPT 7.5 Programmed Cell Death Is a Necessary Process in Living Organisms

- A cell may die by **necrosis**, or it may self-destruct by **apoptosis**, a genetically programmed series of events that includes the fragmentation of the cell's nuclear DNA.
- Apoptosis is regulated by both external and internal signals. These signals result in activation of a class of enzymes called **caspases** that hydrolyze proteins in the cell. **Review Figure 7.15**

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is7

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

8

Inheritance, Genes, and Chromosomes

KEY CONCEPTS

A male infant undergoes ritual circumcision in accordance with Jewish laws. Sons of Jewish mothers who carry the gene for hemophilia may be exempted from this ritual.

In the Middle Eastern desert 1,800 years ago, a rabbi faced a dilemma. A Jewish woman had given birth to a son. As required by Jewish custom, the mother brought her 8-day-old son to the rabbi for ritual penile circumcision. The rabbi knew that the woman's two previous sons had bled to death when their foreskins were cut. Yet the biblical requirement remained: unless he was circumcised, the boy could not be counted among those with whom their God had made a solemn covenant. After consultation with other rabbis, the religious leaders decided to exempt this third son.

Almost 1,000 years later, in the twelfth century, the physician and biblical commentator Moses Maimonides reviewed this and other cases in the rabbinical literature and stated that in such instances the third son should not be circumcised. Furthermore, the exemption should apply whether the mother's son was "from her first husband or from her second husband." The bleeding disorder, he reasoned, was clearly carried by the mother and passed on to her sons. In all cases, the parents did not show any evidence of having the disease.

Without any knowledge of modern concepts of genes and genetics, the rabbis had linked a human disease with a pattern of inheritance. We now have a name for the disease: it is hemophilia, which affects about 18,000 people in the United States, almost all of them males. The bleeding disorder is due to the absence of a specific protein that is crucial for the formation of blood clots. When a person who does not have hemophilia gets a cut there is usually some bleeding, but soon a clot forms and prevents further bleeding. In the case of hemophilia, no clot forms and the bleeding can continue until the person dies. Indeed, well into the twentieth century the slightest accident could be lethal for such a person. Internal bleeding is also an extremely serious problem for people with this disease, and permanent joint damage due to bleeding in the joints is a common problem for untreated patients.

Treatment of hemophilia by injection of clotting factor into the bloodstream is now possible because the proteins can be isolated from donated blood or made in the laboratory using biotechnological techniques. An issue has been whether people who suffer from hemophilia should receive injections of clotting factor all the time as a preventive measure (an expensive proposition), or only when the factor is needed. Based on reductions in joint damage in children treated by the former (preventive) approach, recent studies have concluded that this approach is best.

Q How is hemophilia inherited, and why is it most frequent in males?

You will find the answer to this question on page 169.

CONCEPT 8.1 Genes Are Particulate and Are Inherited According to Mendel's Laws

Genetics, the field of biology concerned with inheritance, has a long history. There is good evidence that people were deliberately breeding animals (horses) and plants (the date palm tree) for desirable characteristics as long as 5,000 years ago. The general idea was to examine the natural variation among the individuals of a species and "breed the best to the best and hope for the best." This was a hit-or-miss method—sometimes the resulting offspring had all the good characteristics of the parents, but often they did not.

By the mid-nineteenth century, two hypotheses had emerged to explain the results of breeding experiments:

- The hypothesis of *blending inheritance* proposed that gametes contained hereditary determinants (what we now call genes) that blended when the gametes fused during fertilization. Like inks of different colors, the two different determinants lost their individuality after blending and could never be separated. For example, if a plant that made smooth, round seeds was mated (crossed) with a plant that made wrinkled seeds, the offspring would be intermediate between the two and the determinants for the two parental characteristics would be lost.
- The hypothesis of *particulate inheritance* proposed that each determinant had a physically distinct nature; when gametes fused in fertilization, the determinants remained intact. According to this theory, if a plant that made round seeds was crossed with a plant that made wrinkled seeds, the offspring (no matter the shape of their seeds) would still contain the determinants for the two characteristics.

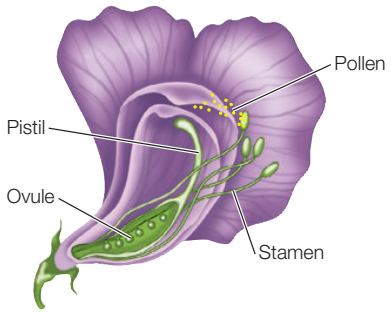
The story of how these competing hypotheses were tested provides a great example of how the scientific method can be used to support one theory and reject another. In the following sections we will look in detail at experiments performed in the 1860s by an Austrian monk and scientist, Gregor Mendel, whose work clearly supported the particulate hypothesis.

Mendel used the scientific method to test his hypotheses

After entering the priesthood at a monastery in Brno, in what is now the Czech Republic, Gregor Mendel was sent to the University of Vienna, where he studied biology, physics, and mathematics. He returned to the monastery in 1853 to teach. The abbot in charge had set up a small plot of land to do experiments with plants and encouraged Mendel to continue with them. Over seven years, Mendel made crosses with many thousands of plants. Analysis of his meticulously gathered data suggested to him that inheritance was due to particulate factors.

Mendel presented his theories in two public lectures in 1865 and a detailed written publication in 1866, but his work was ignored by mainstream scientists until 1900. By that time, the discovery of chromosomes had suggested to biologists that genes might be carried on chromosomes. When they read Mendel's work on particulate inheritance, the biologists connected the dots between genes and chromosomes.

Mendel chose to study the common garden pea because it is easily cultivated and is amenable to manipulation and controlled crosses. Pea flowers have both male and female sex organs—stamens and pistils, which produce gametes that are contained within the pollen and ovules, respectively:



Pea flowers normally self-fertilize. However, the male organs can be removed from a flower so that it can be manually fertilized with pollen from a different flower.

There are many varieties of pea plants with easily recognizable characteristics. A **character** is an observable physical feature, such as seed shape. A **trait** is a particular form of a character, such as round or wrinkled seeds. Mendel worked with seven pairs of varieties with contrasting traits for characters such as seed shape, seed color, and flower color. These varieties were true-breeding: that is, when he crossed a plant that produced wrinkled seeds with another of the same variety, all of the offspring plants produced wrinkled seeds.

As we will see, Mendel developed a set of hypotheses to explain the inheritance of particular pea traits, and then designed crossing experiments to test his hypotheses. He performed his crosses in the following manner:

- He removed the stamens to emasculate flowers of one parental variety so that it couldn't self-fertilize. Then he collected pollen from another parental variety and placed it on the pistils of the emasculated flowers. The plants providing and receiving the pollen were the **parental generation**, designated **P**. The pollen provider was the male parent, and the pollen receiver was the female parent.
- In due course, seeds formed and were planted. The seeds and the resulting new plants constituted the **first filial generation**, or $\mathbf{F_1}$. (The word "filial," from the Latin *filius*, "son," refers to the relationship between offspring and parents.) Mendel examined each F_1 plant to see which traits it bore and then recorded the number of F_1 plants expressing each trait.
- In some experiments the F_1 plants were allowed to self-pollinate and produce a **second filial generation**, the $\mathbf{F_2}$. Again, each F_2 plant was characterized and counted.

Go to MEDIA CLIP 8.1
Mendel's Discoveries
PoL2e.com/mc8.1

Mendel's first experiments involved monohybrid crosses

The term "hybrid" refers to the offspring of crosses between organisms differing in one or more characters. In Mendel's first experiments, he crossed parental (P) varieties with contrasting traits for a single character, producing monohybrids (from the Greek *monos*, "single") in the F_1 generation. He subsequently planted the F_1 seeds and allowed the resulting plants to self-pollinate to produce the F_2 generation. This technique is referred to as a **monohybrid cross**.

Mendel performed the same experiment for seven pea characters. His method is illustrated in **FIGURE 8.1**, using seed shape as an example. When he crossed a strain that made round seeds with one that made wrinkled seeds, the F_1 seeds were round—it was as if the wrinkled seed trait had disappeared completely. However, when F_1 plants were allowed to self-pollinate to produce F_2 seeds, about one-fourth of the seeds were wrinkled. These observations were key to distinguishing the two theories noted above:

- The F_1 offspring were *not a blend* of the two traits of the parents. Only one of the traits was present (in this case, round seeds).
- Some F_2 offspring had wrinkled seeds. The trait *did not disappear*.

These observations led to a rejection of the blending theory of inheritance and provided support for the particulate theory. We now know that hereditary determinants are not actually "particulate," but they are physically distinct entities: sequences of DNA carried on chromosomes (see Concept 8.3).

All seven crosses between varieties with contrasting traits gave the same kind of data (see Figure 8.1). In the F_1 generation only one of the two traits was seen, but the other one reappeared in about one-fourth of the offspring in the F_2 generation. Mendel called the trait that appeared in the F_1 and was more abundant in the F_2 the **dominant** trait, and the other trait **recessive**.

Mendel went on to expand the particulate hypothesis. He proposed that hereditary determinants—what we know today as genes, though Mendel did not use that term—occur in pairs and segregate (separate) from one another during the formation of gametes. He concluded that each pea plant has two copies of the gene for each character (such as seed shape), one inherited from each parent. We now use the term **diploid** to describe the state of having two copies of each gene; **haploids** have just one copy (see Concept 7.1).

Mendel concluded that while each gamete contains one copy of each gene (i.e., is haploid), the resulting zygote contains two copies (is diploid), because it is produced by the fusion of two gametes. Furthermore, different traits arise because there can be *different forms of a gene*—now called **alleles**—for a particular character. For example, Mendel studied two alleles for seed shape: one that caused round seeds and the other causing wrinkled seeds.

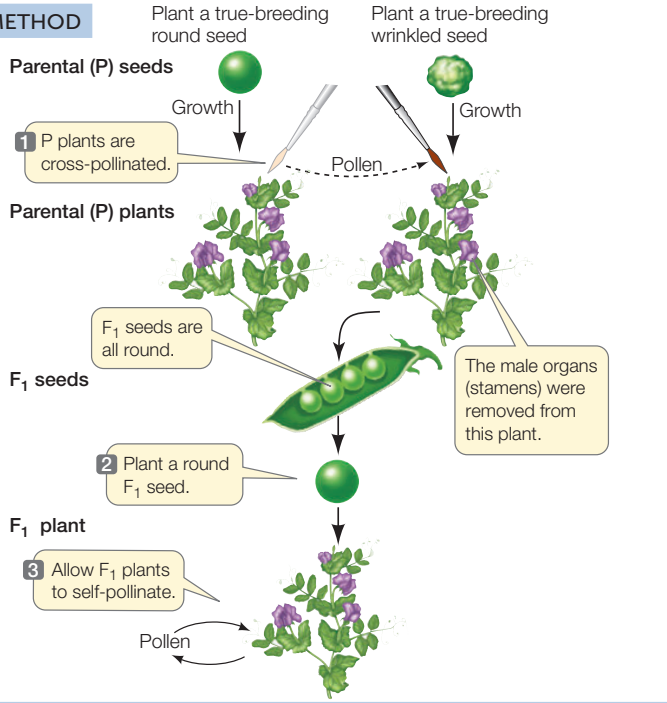
INVESTIGATION

FIGURE 8.1 Mendel's Monohybrid Experiments Mendel performed crosses with pea plants and carefully analyzed the outcomes to show that genetic determinants are particulate.[a]

HYPOTHESIS

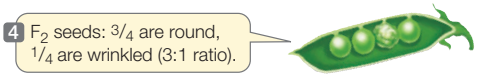
When two strains of peas with contrasting traits are bred, their characteristics are irreversibly blended in succeeding generations.

METHOD



RESULTS

F_2 seeds from F_1 plant



CONCLUSION

The hypothesis is rejected. There is no irreversible blending of characteristics, and a recessive trait can reappear in succeeding generations.

ANALYZE THE DATA

The table gives Mendel's data—the number of offspring showing each trait—for the F_2 from crosses between P generation plants with contrasting traits:

Characteristic	Dominant	Recessive
Seed shape	5,474 round	1,850 wrinkled
Seed color	6,022 yellow	2,001 green
Flower color	705 purple	224 white
Pod color	428 green	152 yellow
Stem height	787 tall	277 short

A. Calculate the phenotypic ratio of dominant:recessive in the F_2 offspring.

B. What can you conclude about the behavior of alleles during gamete formation in a plant that is heterozygous for a trait?

C. Perform a chi-square test to evaluate the statistical significance of these data (refer to Appendix B).

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]See www.mendelweb.org/Mendel.plain.html

- An organism that is **homozygous** for a gene has two alleles that are the same (for example, two copies of the allele for round seeds).
- An organism that is **heterozygous** for a gene has two different alleles (for example, one allele for round seeds and one allele for wrinkled seeds).

In a heterozygote, one of the two alleles may be dominant (such as round, *R*) and the other recessive (wrinkled, *r*). By convention, we designate dominant alleles with uppercase letters and recessive alleles with lowercase letters.

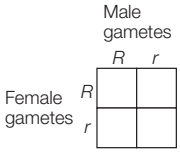
- The physical appearance of an organism is its **phenotype**.
- The phenotype is the result of the **genotype**, or genetic constitution, of the organism.

Round seeds and wrinkled seeds are two phenotypes resulting from three possible genotypes: the wrinkled seed phenotype is produced by the genotype *rr*, whereas the round seed phenotype is produced by either of the genotypes *RR* or *Rr* (because the *R* allele is dominant to the *r* allele).

Mendel's first law states that the two copies of a gene segregate

How do Mendel's theories explain the proportions of traits seen in the F_1 and F_2 generations of his monohybrid crosses? Mendel's first law—the **law of segregation**—states that *when any individual produces gametes, the two copies of a gene separate, so each gamete receives only one copy*. Thus gametes from a parent with the *RR* genotype will all be *R*; gametes from an *rr* parent will all be *r*; and the progeny derived from a cross between these parents will all be *Rr*, producing seeds with a round phenotype (FIGURE 8.2).

Now let's consider the composition of the F_2 generation. Because the alleles segregate, half of the gametes produced by the F_1 generation will have the *R* allele and the other half will have the *r* allele. What genotypes are produced when these gametes fuse to form the next (F_2) generation? We can predict the allele combinations that will result from a cross using a **Punnett square**, a method devised in 1905 by the British geneticist Reginald Punnett. This device ensures that we consider all possible random combinations of gametes when calculating expected genotype frequencies. A Punnett square looks like this:



All possible male gamete (haploid sperm) genotypes are shown along the top and all possible female gamete (haploid egg) genotypes along the left side. The grid is completed by filling in each square with the diploid genotype that can be generated from each combination of gametes. In this example, to fill in the top right square, we put in the *R* from the female gamete (the egg cell) and the *r* from the male gamete (the sperm cell in the pollen tube), yielding *Rr*.

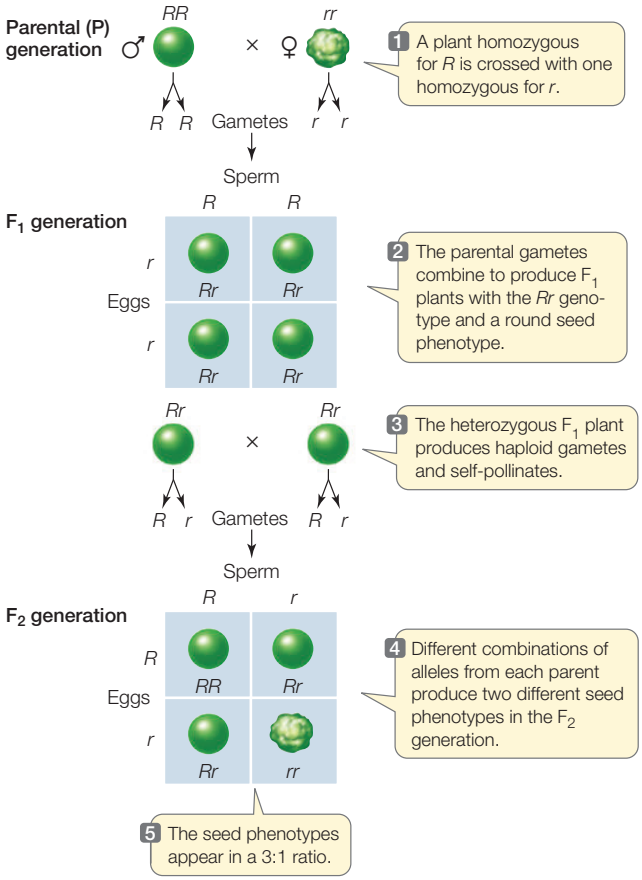


FIGURE 8.2 Mendel's Explanation of Inheritance Mendel concluded that inheritance depends on discrete factors from each parent that do not blend in the offspring.

Once the Punnett square is filled in, we can readily see that there are four possible combinations of alleles in the F_2 generation: *RR*, *Rr*, *rR*, and *rr* (see Figure 8.2). Since *R* is dominant, there are three ways to get round-seeded plants in the F_2 generation (*RR*, *Rr*, or *rR*), but only one way to get a plant with wrinkled seeds (*rr*). Therefore we predict a 3:1 ratio of these phenotypes in the F_2 generation, remarkably close to the values Mendel found experimentally for all seven of the traits he compared.

Mendel did not live to see his theories placed on a sound physical footing with the discoveries of chromosomes and DNA. Genes are now known to be relatively short sequences of DNA (usually a few thousand base pairs in length) found on the much longer DNA molecules that make up chromosomes (which are often millions of base pairs long). Today we can picture the different alleles of a gene segregating as chromosomes separate during meiosis I (FIGURE 8.3).

LINK

You can review the process of meiosis and the separation of chromosomes in Concept 7.4

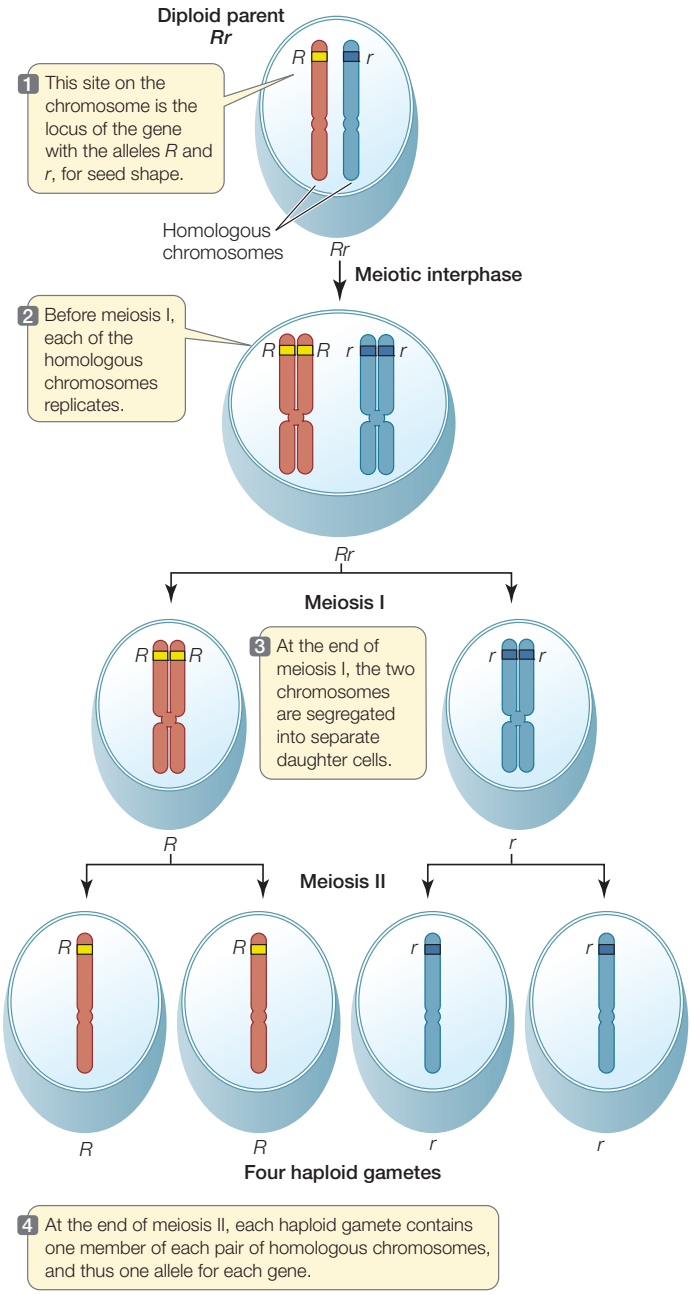


FIGURE 8.3 Meiosis Accounts for the Segregation of Alleles
Although Mendel had no knowledge of chromosomes or meiosis, we now know that a pair of alleles resides on homologous chromosomes, and that those alleles segregate during meiosis.

Genes determine phenotypes mostly by producing proteins with particular functions, such as enzymes. So in many cases a dominant gene is expressed (transcribed and translated; see Concept 3.1) to produce a functional protein, whereas a recessive gene is mutated so that it is no longer expressed, or it encodes a mutant protein that is nonfunctional. For example, the molecular nature of the wrinkled pea seed phenotype is the absence of an enzyme called starch branching enzyme 1 (SBE1), which is essential for starch synthesis. With less starch, the developing seed has more sucrose, which causes an inflow of water by osmosis. When the seed matures, this water is lost, leaving a shrunken, wrinkled seed.

Mendel verified his hypotheses by performing test crosses

As mentioned above, Mendel arrived at his laws of inheritance by developing a series of hypotheses and then designing experiments to test them. One such hypothesis was that there are two possible allele combinations (*RR* or *Rr*) for seeds with the round phenotype. Mendel verified this hypothesis by performing test crosses with F_1 seeds derived from a variety of other crosses. A **test cross** is used to determine whether an individual showing a dominant trait is homozygous or heterozygous. The individual in question is crossed with an individual that is homozygous for the recessive trait—an easy individual to identify, because all individuals with the recessive phenotype are homozygous for that trait.

The recessive homozygote for the seed shape gene has wrinkled seeds and the genotype *rr*. The individual being tested may be described initially as *R*_ because we do not yet know the identity of the second allele. We can predict two possible results:

- If the individual being tested is homozygous dominant (*RR*), all offspring of the test cross will be *Rr* and show the dominant trait (round seeds) (**FIGURE 8.4, LEFT**).
- If the individual being tested is heterozygous (*Rr*), then approximately half of the offspring of the test cross will be heterozygous and show the dominant trait (*Rr*), and the other half will be homozygous for the recessive trait (*rr*) (**FIGURE 8.4, RIGHT**).

Mendel obtained results consistent with both of these predictions; thus his hypothesis accurately predicted the results of his test crosses.

Go to ACTIVITY 8.1 Homozygous or Heterozygous?
PoL2e.com/ac8.1

Mendel's second law states that copies of different genes assort independently

Consider an organism that is heterozygous for two genes (*RrYy*). In this example, the dominant *R* and *Y* alleles came from one true-breeding parent, and the recessive *r* and *y* alleles came from the other true-breeding parent. When this organism produces gametes, do the *R* and *Y* alleles always go together in one gamete, and *r* and *y* alleles in another? Or can a single gamete receive one recessive and one dominant allele (*R* and *y* or *r* and *Y*)?

Mendel performed another series of experiments to answer these questions. He began with peas that differed in *two* characters: seed shape and seed color. One parental variety produced only round, yellow seeds (*RRYY*), and the other produced only wrinkled, green ones (*rryy*). A cross between these two varieties produced an F_1 generation in which all the plants were *RrYy*.

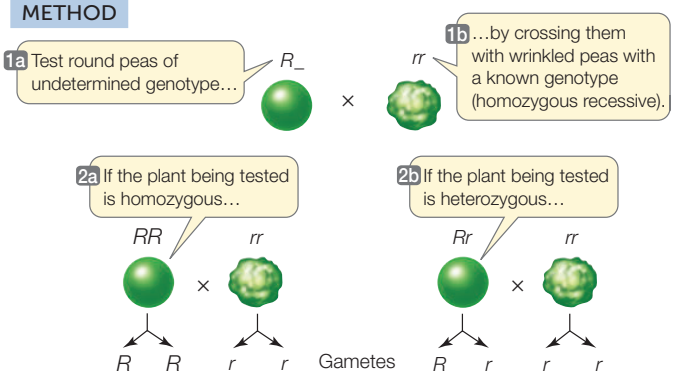
INVESTIGATION

FIGURE 8.4 Homozygous or Heterozygous? An individual with a dominant phenotype may have either a homozygous or a heterozygous genotype. The test cross determines which.[a]

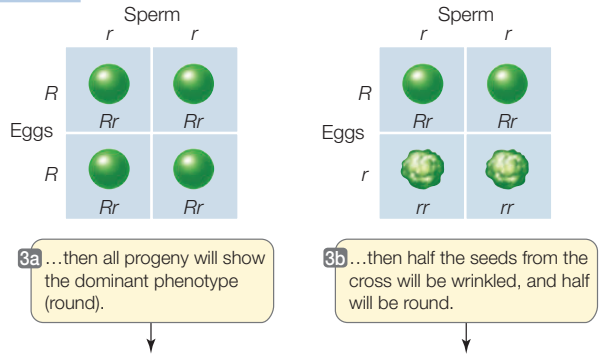
HYPOTHESIS

The progeny of a test cross can reveal whether a parental organism is homozygous or heterozygous.

METHOD



RESULTS



CONCLUSION

The plant being tested is homozygous.

CONCLUSION

The plant being tested is heterozygous.

ANALYZE THE DATA

A tall plant was crossed with a short plant to produce three F_1 plants (A, B, and C). Each of the F_1 plants was test crossed, with the following results:

F_1	Tall progeny	Short progeny
A	34	0
B	18	15
C	21	0

A. For each test cross, what were the genotypes and phenotypes of both parents?

B. Choose and perform a statistical test from Appendix B to evaluate the significance of these data.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]See www.mendelweb.org/Mendel.plain.html

Because the *R* and *Y* alleles were dominant, the F_1 seeds were all round and yellow.

Mendel continued this experiment into the F_2 generation by performing a **dihybrid cross**, a cross between individuals that are identical double heterozygotes. In this case he simply allowed the F_1 plants, which were all double heterozygotes, to self-pollinate (**FIGURE 8.5**). Depending on whether the alleles of the two genes are inherited together or separately, there are two possible outcomes, as Mendel saw:

- The alleles could maintain the associations they had in the parental generation—they could be near each other (linked) on the same chromosome. If this were the case, the F_1 plants would produce two types of gametes (*RY* and *ry*). The F_2 progeny resulting from self-pollination of these F_1 plants would consist of two phenotypes: round yellow and wrinkled green in the ratio of 3:1, just as in the monohybrid cross.
- The segregation of *R* from *r* could be independent of the segregation of *Y* from *y*—the two genes could be unlinked, on different chromosomes. In this case, four

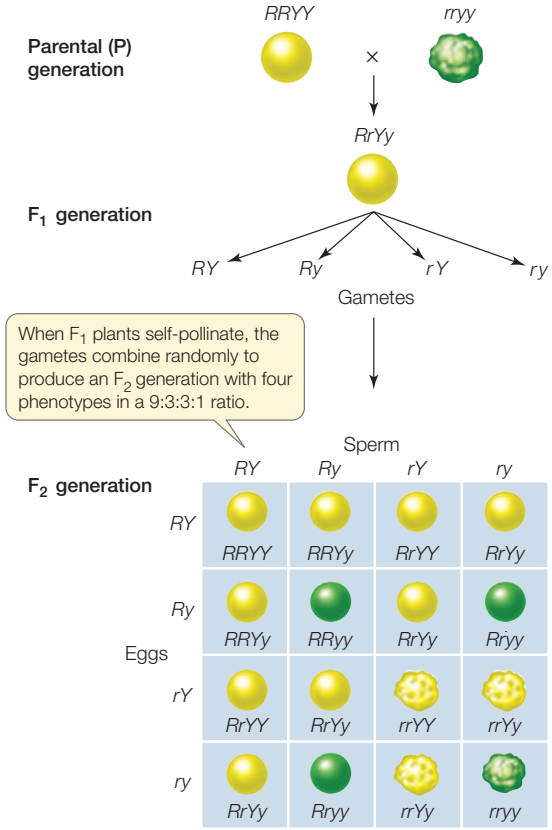


FIGURE 8.5 Independent Assortment The 16 possible combinations of gametes in this dihybrid cross result in nine different genotypes. Because *R* and *Y* are dominant over *r* and *y*, respectively, the nine genotypes result in four phenotypes in a ratio of 9:3:3:1. These results show that the two genes segregate independently.

kinds of gametes would be produced in equal numbers: *RY*, *Ry*, *rY*, and *ry*. When these gametes combine at random, they should produce an F_2 having nine different genotypes. The nine genotypes would produce four phenotypes (round yellow, round green, wrinkled yellow, wrinkled green). Putting these possibilities into a Punnett square, we can predict that these four phenotypes just mentioned would occur in a ratio of 9:3:3:1.

Mendel's dihybrid crosses supported the second prediction: four different phenotypes appeared in the F_2 generation in a ratio of about 9:3:3:1 (see Figure 8.5). On the basis of such experiments, Mendel proposed his second law, the **law of independent assortment**: *alleles of different genes assort independently of one another during gamete formation*. In the example above, the segregation of the *R* and *r* alleles is independent of the segregation of the *Y* and *y* alleles. As you will see in Concept 8.3, this is not as universal as the law of segregation because it does not apply to genes located near one another on the same chromosome. However, it is correct to say that *chromosomes segregate independently* during the formation of gametes, and so do any two genes located on separate chromosome pairs (FIGURE 8.6).

Probability is used to predict inheritance

One key to Mendel's success was his use of large sample sizes. By counting many progeny from each cross, he observed clear patterns that allowed him to formulate his theories. After his work became widely recognized, geneticists began using simple probability calculations to predict the ratios of genotypes and phenotypes in the progeny of a given cross or mating. They use statistics to determine whether the actual results match the prediction.

You can think of probabilities by considering a coin toss. The basic conventions of probability are simple:

- If an event is absolutely certain to happen, its probability is 1.
- If it cannot possibly happen, its probability is 0.
- All other events have a probability between 0 and 1.

There are two possible outcomes of a coin toss, and both are equally likely, so the probability of heads is ½—as is the probability of tails.

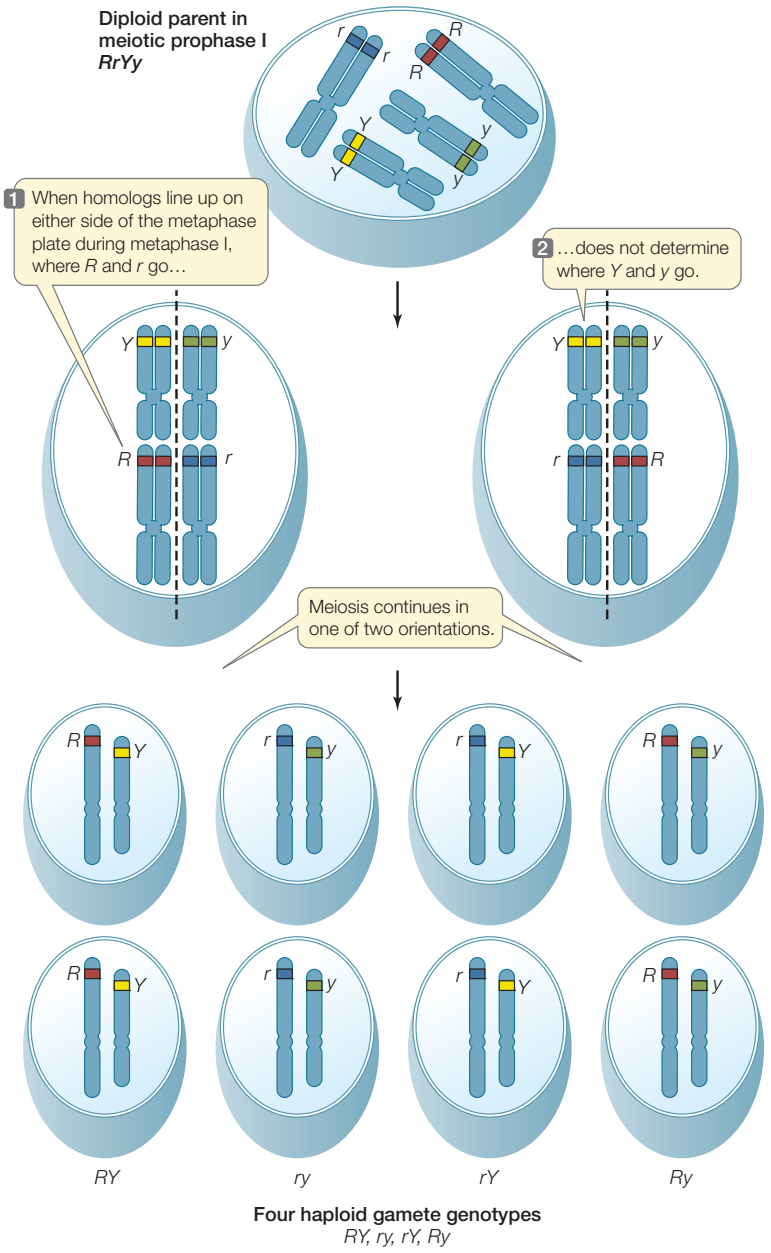
If two coins (say a penny and a dime) are tossed, each acts independently of the other (FIGURE 8.7).

FIGURE 8.6 Meiosis Accounts for Independent Assortment of Alleles We now know that copies of genes on different chromosomes are segregated independently during metaphase I of meiosis. Thus a parent of genotype *RrYy* can form gametes with four different genotypes.

Go to ANIMATED TUTORIAL 8.1
Independent Assortment of Alleles
PoL2e.com/at8.1

What is the probability of both coins coming up heads? In half of the tosses, the penny comes up heads, and in half of that fraction, the dime comes up heads. The probability of both coins coming up heads is ½ × ½ = ¼. In general, *the probability of two independent outcomes occurring together is found by multiplying the two individual probabilities*. This can be applied to a monohybrid cross (see Figure 8.2). After the self-pollination of an *Rr* F_1 plant, the probability that an F_2 plant will have the genotype *RR* is ½ × ½ = ¼, because the chance that the sperm will have the genotype *R* is ½, and the chance that the egg will have the genotype *R* is also ½. Similarly, the probability of *rr* offspring is also ¼.

What about the probability of getting a heterozygote? As you can see in Figures 8.2 and 8.7, there are *two* ways to get an



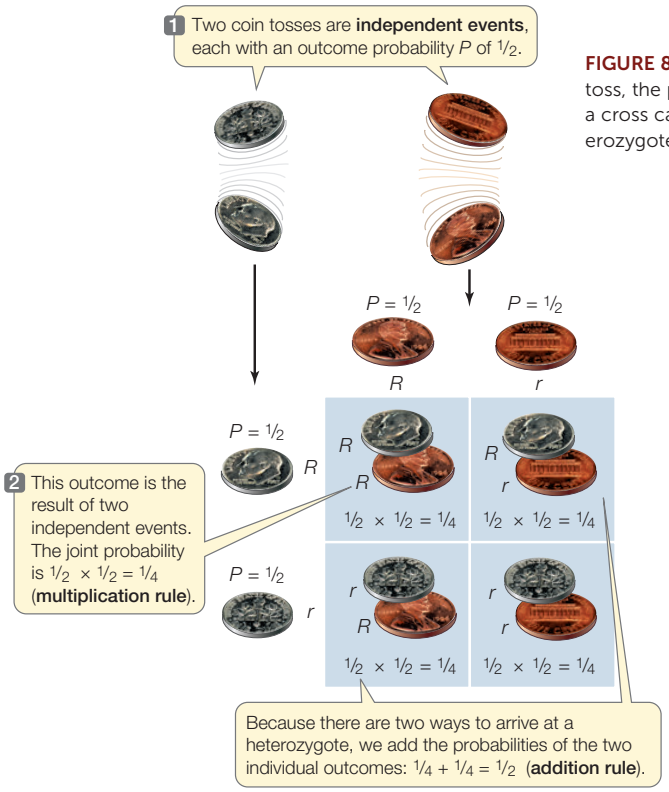


FIGURE 8.7 Using Probability Calculations in Genetics Like the results of a coin toss, the probability of any given combination of alleles appearing in the offspring of a cross can be obtained by multiplying the probabilities of each event. Since a heterozygote can be formed in two ways, these two probabilities are added together.

Rr plant or a head and a tail in a coin toss. In the case of the seed shape gene, the *R* allele can come from a sperm and the *r* from an egg (probability ¼). Or the *R* allele could come from the egg and the *r* from the sperm (probability ¼). *The probability of an event that can occur in two or more different ways is the sum of the individual probabilities of those ways*. Thus the probability that an F_2 plant will be a heterozygote is equal to the sum of the probabilities of the two ways of forming a heterozygote: ¼ + ¼ = ½.

Mendel's laws can be observed in human pedigrees

Mendel developed his theories by performing many planned crosses and counting many offspring. This approach is not possible with humans, so human geneticists rely on **pedigrees**: family trees that show the occurrence of inherited phenotypes in several generations of related individuals (**FIGURE 8.8**).

Because humans have relatively few offspring, human pedigrees do not show the clear proportions of phenotypes that Mendel saw in his pea plants. For example, when a man and a woman who are both heterozygous for a recessive allele (say, *Aa*) have children together, each child has a ¼ probability of

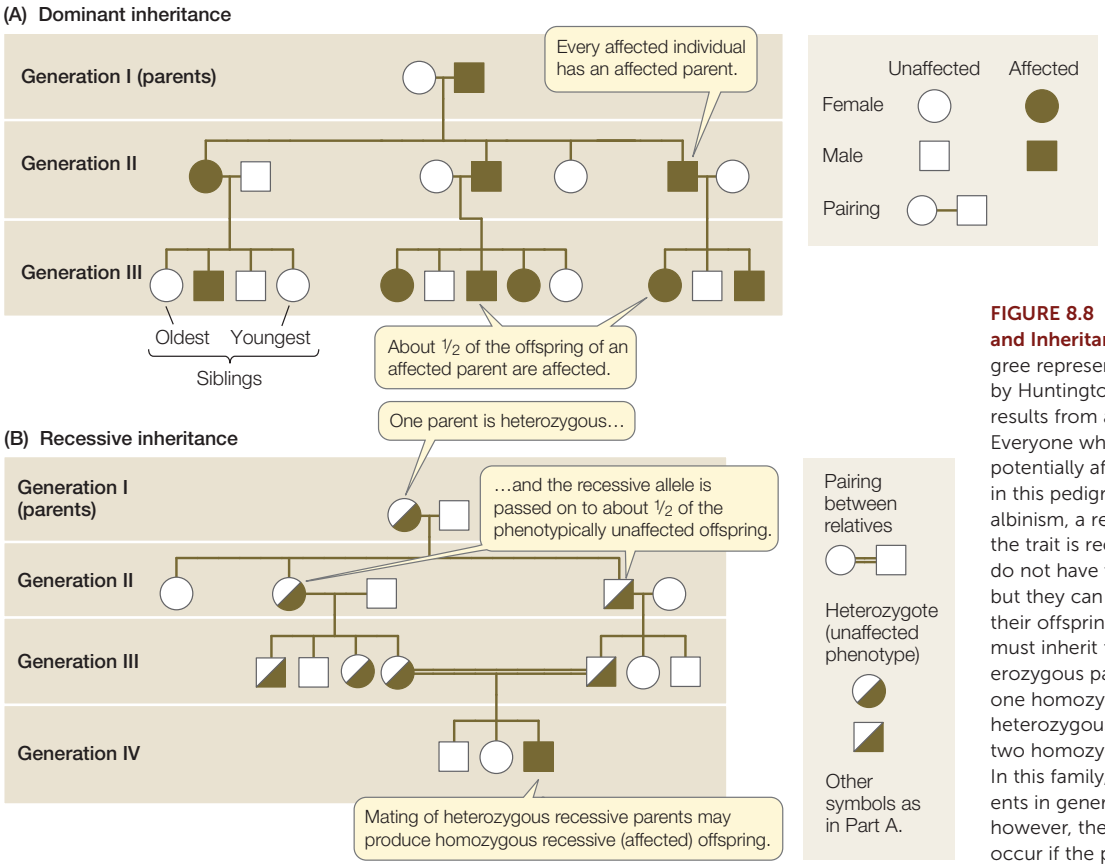


FIGURE 8.8 Pedigree Analysis and Inheritance (A) This pedigree represents a family affected by Huntington's disease, which results from a rare dominant allele. Everyone who inherits this allele is potentially affected. (B) The family in this pedigree carries the allele for albinism, a recessive trait. Because the trait is recessive, heterozygotes do not have the albino phenotype, but they can pass the allele on to their offspring. Affected persons must inherit the allele from two heterozygous parents, or (rarely) from one homozygous recessive and one heterozygous parent, or (very rarely) two homozygous recessive parents. In this family, the heterozygous parents in generation III are cousins; however, the same result would occur if the parents were unrelated.

being a recessive homozygote (*aa*). But the offspring of a single couple are likely to be too few to reliably show the one-fourth proportion. In a family with only two children, for example, both could easily be *aa* (or *Aa*, or *AA*).

Despite this limitation, pedigrees do show inheritance patterns that can provide information about the allele(s) controlling a particular phenotype. For example, it is useful to know whether a particular rare allele that causes an abnormal phenotype is dominant or recessive. Figure 8.8A is a pedigree showing the pattern of inheritance of a rare dominant allele. The following are the key features to look for in such a pedigree:

- Every person with the abnormal phenotype (affected) has an affected parent.
- Either all (in the case of a homozygous parent) or about half (in the case of a heterozygous parent) of the offspring are affected. In Figure 8.8A, each of the affected individuals in generations II and III is heterozygous for the rare dominant allele. This is clear because each person has inherited a recessive allele from his or her unaffected parent.

Compare this pattern with the one shown in Figure 8.8B, which is typical for the inheritance of a rare recessive allele:

- Affected people most often have two unaffected parents.
- Only a small proportion of people are affected: about one-fourth of children whose parents are both heterozygotes. Such parents are **heterozygous carriers** of the recessive allele. Unfortunately, it is not possible to use the pedigree alone to determine who else in the family may be a heterozygous carrier (that is, unless they parent a homozygous child), because the allele is recessive and shows no phenotype in the heterozygote. However, as we will see in Chapter 12, DNA sequencing or other techniques can sometimes be used to identify carriers of specific recessive alleles. This is especially important in cases where a rare allele causes serious disease.

Other patterns of inheritance can arise in special cases, such as sex-linked characters (see Concept 8.3).

Go to ANIMATED TUTORIAL 8.2
Pedigree Analysis Simulation
PoL2e.com/at8.2

CHECKpoint CONCEPT 8.1

- ✓ What are the differences between genes and alleles? Between homozygous and heterozygous conditions? Between genotype and phenotype? Between Mendel's definition of a gene and the current definition?
- ✓ In a monohybrid cross, how do the events of meiosis explain Mendel's first law? In a dihybrid cross, how does meiosis explain Mendel's second law?
- ✓ Using the cross shown in Figure 8.5, calculate the probability that an F_2 seed will be round and yellow.

The laws of inheritance as articulated by Mendel remain valid today, and his discoveries laid the groundwork for all future studies of genetics. However, as we will see next, the relationship of one gene to one phenotype can be complicated by interactions among alleles and among genes.

CONCEPT 8.2 Alleles and Genes Interact to Produce Phenotypes

Phenotypes do not always follow the simple patterns of inheritance shown by the pairs of alleles for seed color and seed shape in peas, described in Concept 8.1. Existing alleles are subject to change by mutation and can give rise to new alleles—in fact, a single gene can have many alleles. In addition, alleles do not always show simple dominant–recessive relationships. A single allele may have multiple phenotypic effects, and a single character may be controlled by multiple genes. The expression of a gene is generally affected by interactions with other genes and with the environment.

New alleles arise by mutation

Genes are subject to **mutations**, which are rare, stable, and inherited changes in the genetic material. In other words, an allele can mutate (change) to become a different allele (this can happen in a number of different ways, as will be detailed in Concept 9.3). For example, we can envision that at one time all pea plants made round seeds and had the seed shape allele *R*. At some point, a mutation in *R* resulted in a new allele, *r* (wrinkled seeds). If this mutation was present in a cell that underwent meiosis, some of the resulting gametes would carry the *r* allele, and some offspring of this pea plant would carry the *r* allele in all of their cells.

Geneticists usually define one allele of a gene as the **wild type**; this is the allele that is present in most individuals in nature ("the wild"). Other alleles of that gene are usually called mutant alleles, and they may produce different phenotypes. The wild-type and mutant alleles are inherited according to Mendelian laws. A gene with a wild-type allele that is present less than 99 percent of the time (the rest of the alleles being mutant) is said to be **polymorphic** (Greek *poly*, "many"; *morph*, "form").

LINK

By producing phenotypic variety, mutations provide raw material for evolution. An allele may become more or less prevalent in a population, depending on its effect on the fitness of the individuals carrying it; see **Concept 15.3**

Mendel developed his theories by studying just two alleles of each gene. But often a gene has multiple alleles (although any diploid individual will carry only two of them). The alleles may show a hierarchy of dominance when present in heterozygous individuals. An example is coat color in rabbits, determined by multiple alleles of the *C* gene (**FIGURE 8.9**):

- *C* determines dark gray.
- c^{chd} determines chinchilla, a lighter gray.

Possible genotypes	CC, Cc^{chd}, Cc^{h}, Cc	$c^{chd}c^{chd}$, $c^{chd}c$	$c^{h}c^{h}$, $c^{h}c$	cc
Phenotype	Dark gray	Chinchilla	Point restricted	Albino

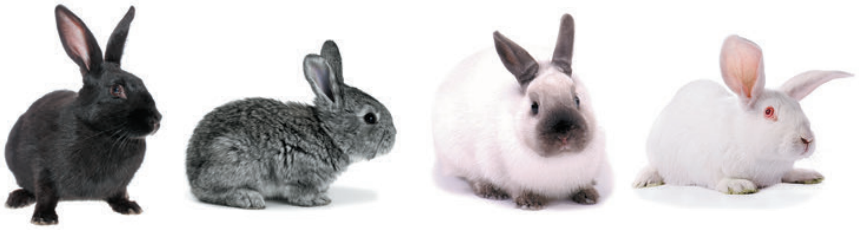


FIGURE 8.9 Multiple Alleles for Coat Color in Rabbits These photographs show the phenotypes conferred by four alleles of the *C* gene for coat color in rabbits. Different combinations of two alleles give different coat colors and pigment distributions.

- c^{h} determines Himalayan (point restricted).
- c determines albino, no pigment.

The hierarchy of dominance for these alleles is $C > c^{chd} > c^{h} > c$. Any rabbit with the *C* allele (paired with itself or another allele) is dark gray, and a *cc* rabbit is albino. Intermediate colors result from different allele combinations, as shown in Figure 8.9. As this example illustrates, multiple alleles can increase the number of possible phenotypes.

Dominance is not always complete

Many genes have alleles that are neither dominant nor recessive to one another. Instead, the heterozygotes have an intermediate phenotype, in a situation called **incomplete dominance**. For example, if a true-breeding red snapdragon is crossed with a white one, all the F_1 flowers are an intermediate pink. Such cases appear to support the old blending theory of inheritance. However, further crosses indicate that this apparent blending can still be explained in terms of Mendelian genetics (**FIGURE 8.10**). The red and white snapdragon alleles have not disappeared, as those colors reappear in the F_2 generation.

Sometimes two alleles of a gene both produce their phenotypes when present in a heterozygote—a phenomenon called **codominance**. An example is the ABO blood group in humans. The gene *I* encodes an enzyme involved in the attachment of sugars to a glycoprotein present on the surfaces of red blood cells. There are three alleles of the gene: I^{A}, I^{B}, and I^{O}. The I^{A} and I^{B} alleles both encode active enzymes, but the enzymes attach different sugars to the glycoprotein. The I^{O} allele does not encode an active enzyme, so no sugar is attached at that position on the glycoprotein. When two different alleles (e.g., I^{A} and I^{B}) are present, both alleles are expressed (both enzymes are

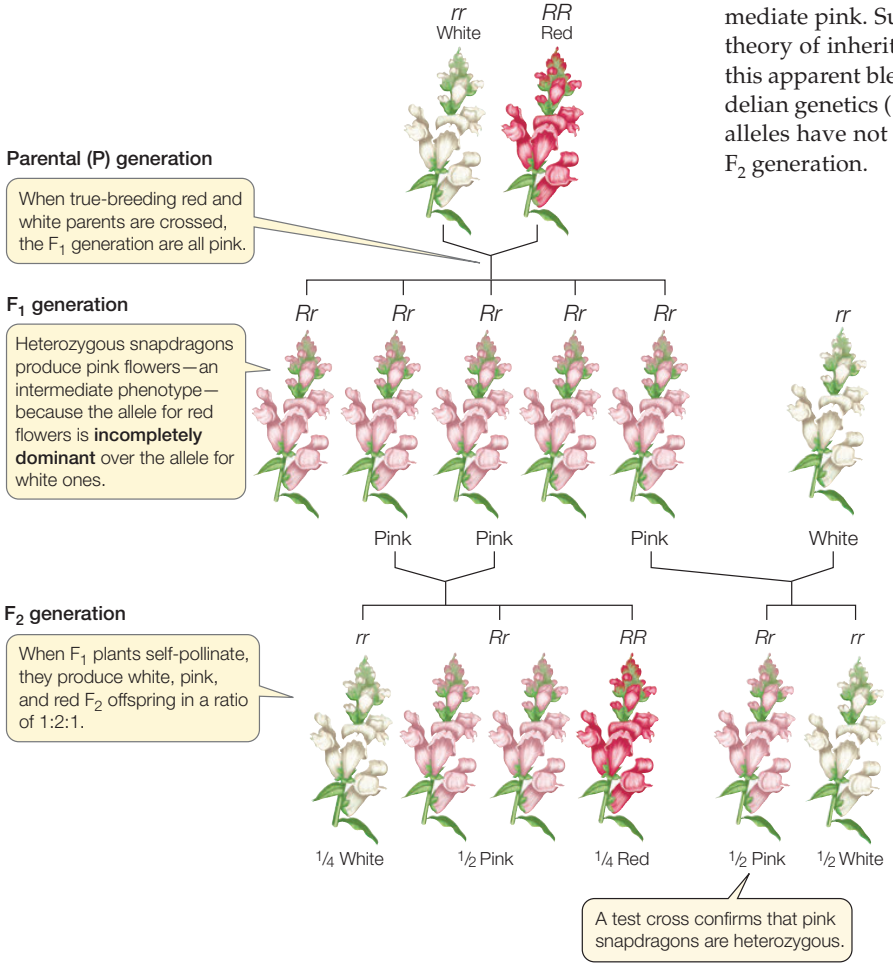


FIGURE 8.10 Incomplete Dominance Follows Mendel's Laws An intermediate phenotype can occur in heterozygotes when neither allele is dominant. The heterozygous phenotype (here, pink flowers) may give the appearance of a blended trait, but the traits of the parental generation reappear in their original forms in succeeding generations, as predicted by Mendel's laws of inheritance.

made, so both types of glycoproteins are made). The A and B glycoproteins are antigenic: if a red blood cell with the A glycoprotein on its surface gets into the bloodstream of a person who lacks the I^A allele, the recipient produces antibodies against the "nonself" cells (**FIGURE 8.11**). While the A and B glycoproteins are antigenic in people who do not have the I^A or I^B alleles, respectively, the O glycoprotein does not provoke antibody production. This makes people who are I^OI^O good blood donors.

LINK

Reactions against A or B glycoproteins are an adaptive immune response, as described in **Concept 39.3**

Genes interact when they are expressed

Epistasis occurs when the phenotypic expression of one gene is affected by another gene. For example, two genes (*B* and *E*) determine coat color in Labrador retrievers:

- Allele *B* (black pigment) is dominant to *b* (brown).
- Allele *E* (pigment deposition in hair) is dominant to *e* (no deposition, so hair is yellow).

An *EE* or *Ee* dog with *BB* or *Bb* is black; one with *bb* is brown. A dog with *ee* is yellow regardless of whether *B* or *b* alleles are present. Clearly, gene *E* determines the phenotypic expression of gene *B*, and is therefore epistatic to *B* (**FIGURE 8.12**).

Perhaps the most dramatic example of interacting genes or alleles is **hybrid vigor** (or heterosis). In 1876, Charles Darwin reported that when he crossed two different genetic varieties of corn, the offspring were 25 percent taller than either of the parent strains. Darwin's observation was largely ignored for the next 30 years. In 1908, George Shull "rediscovered" this idea, reporting that not just plant height but the weight of the corn grain produced was dramatically higher in the offspring:

Parent: *B73* Hybrid Parent: *Mo17*

Agricultural scientists took note, and Shull's paper had a lasting impact on the field of applied genetics. The cultivation of hybrid corn spread rapidly, and the practice of hybridization is now used for many other agricultural crops and animals. For example, beef cattle strains that are crossbred are larger and live longer than cattle bred within their own genetic strains.

What determines the "vigor" in hybrid vigor? A phenotype such as the amount of grain that a variety of corn produces in a given environment is determined by many genes and their alleles. Put more generally, *most complex phenotypes are determined by multiple genes*. Traits conferred by multiple genes are usually **quantitative traits** that need to be *measured* rather than assessed qualitatively. For example, grain yields must be measured, whereas a character such as Mendel's pea seed color is an either-or quality that can be assessed by eye. The genetic basis of hybrid vigor is not well understood, but presumably the combinations of alleles and their products from two different varieties can interact to produce more vigorous offspring.

The environment affects gene action

The phenotype of an individual does not always result from its genotype alone. *Genotype and environment often interact to determine the phenotype of an organism.* It is especially important to remember this fact in the era of genome sequencing (see Chapter 12). When the sequence of the human genome was completed in 2003, it was hailed as the "book of life," and public expectations of the benefits gained from this knowledge were (and are) high. But this kind of "genetic determinism"—the idea that an organism's genome sequence determines all of its phenotype—is wrong. Common knowledge tells us that environmental variables such as light, temperature, and nutrition can affect the phenotypic expression of a genotype. For example, in humans body weight is determined not only by multiple genes but also by nutrition and activity.

A common example of the interaction of genes and environment involves "point restriction" coat patterns found in Siamese cats and certain rabbit breeds. These animals carry the c^h allele (see Figure 8.9), a mutant version of the *C* gene that controls coat

Red blood cell type	Genotype	Blood group that body rejects	Reaction to added antibodies: Anti-A	Reaction to added antibodies: Anti-B
A	I^AI^A or I^AI^O	B and AB		
B	I^BI^B or I^BI^O	A and AB		
AB	I^AI^B	None		
O	I^OI^O	A, B, and AB		

Red blood cells that do not react with antibody remain evenly dispersed.

Red blood cells that react with antibody clump together (speckled appearance).

FIGURE 8.11 ABO Blood Groups Are Important in Transfusions The table shows the results of mixing red blood cells of types A, B, AB, and O with serum containing anti-A or anti-B antibodies. As you look down the two columns on the right, note that each of the types, when mixed separately with anti-A and with anti-B, gives a unique pair of results. This is the basic method by which blood is typed.

APPLY THE CONCEPT

Alleles and genes interact to produce phenotypes

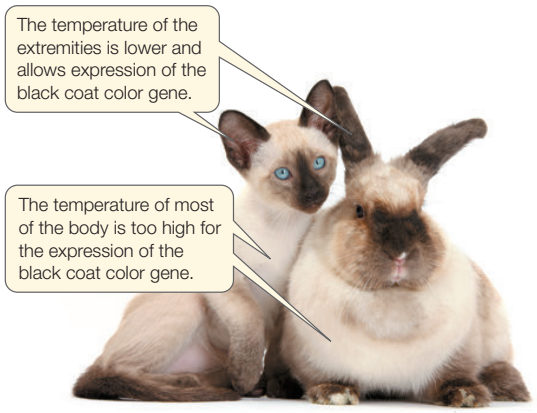
1. In the genetic cross *AaBbCcDdEE* × *AaBBCcDdEe* where all the genes are on separate chromosomes, what fraction of the offspring will be heterozygous for all of these genes?
2. In squash plants, fruit color is determined by one gene (*W*: white, dominant to *w*: yellow) and fruit shape by another gene on a different chromosome (*D*: disk shape, dominant to *d*: round). What were the genotypes of the parents for each cross in the table at right?
3. In chickens, when the dominant alleles of the genes for rose comb (*R*) and pea comb (*A*) are present together, the bird has a walnut comb. Birds that are homozygous recessive for both genes (i.e., genotype *ra*) have a single comb. A rose-combed bird mated with a walnut-combed bird, and the phenotypes of the 8 offspring were:

   ⅜ walnut : ⅜ rose : ⅛ pea : ⅛ single.

   What were the genotypes of the parents?

	NUMBER OF OFFSPRING			
PARENTS	WHITE, DISK	WHITE, ROUND	YELLOW, DISK	YELLOW, ROUND
White, round × White, round	0	52	0	16
White, disk × White, round	62	58	18	20
White, disk × White, disk	176	60	54	21
White, disk × Yellow, round	22	24	21	22

color. As a result of this mutation, the enzyme encoded by the gene is inactive at temperatures above a certain point—usually 35°C. The animals maintain a body temperature above this, so their body fur is light. However, their extremities—feet, ears, nose, and tail—are cooler (around 25°C), and the fur on these regions is dark:



A simple experiment shows that the dark fur is temperature-dependent. If a patch of white fur on a point-restricted animal is shaved off and an ice pack is placed on the skin where the patch was, the fur that grows in is dark. This indicates that although the allele for dark fur was expressed all

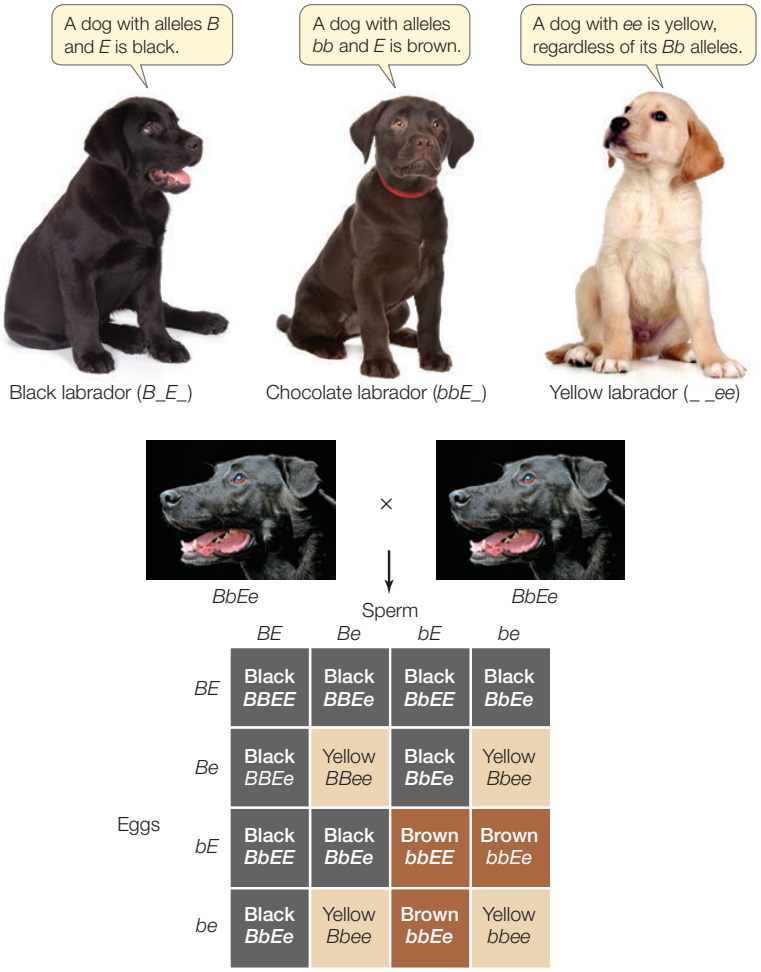


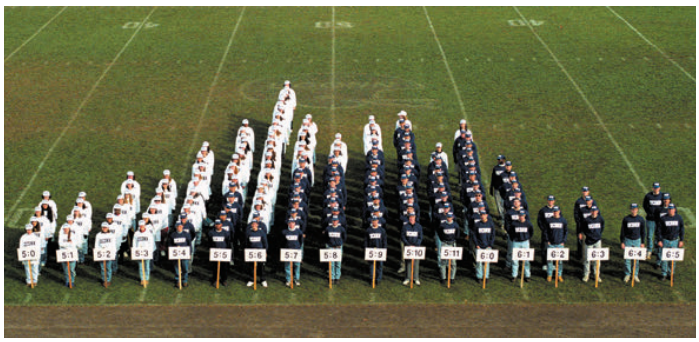
FIGURE 8.12 Genes Interact Epistatically Epistasis occurs when one gene alters the phenotypic effect of another gene. In Labrador retrievers, the *E* gene determines the expression of the *B* gene.

along, the environment inhibited the activity of the mutant enzyme. These animals are all white at birth because the extremities are kept warm in the mother's womb.

Two parameters describe the effects of genes and environment on phenotype:

- **Penetrance** is the proportion of individuals in a group with a given genotype that actually show the expected phenotype. For example, many people who inherit a mutant allele of the gene *BRCA1* develop breast cancer in their lifetimes. But for reasons that are not yet clear and must involve other genes and/or the environment, some people with the mutation do not develop breast cancer. So the *BRCA1* mutation is said to be incompletely penetrant.
- **Expressivity** is the degree to which a genotype is expressed in an individual. For example, a woman with the *BRCA1* allele may develop both breast and ovarian cancer as part of the phenotype, but another woman with the same mutation may only get breast cancer. So the mutation is said to have variable expressivity.

In a population of an organism, there is often wide variability with regard to a particular phenotype. Consider, for example, height in humans:



Human height is a quantitative trait. Many genes contribute to height, and the interactions between these genes are complex. But the environment also contributes to variation in height; for example, some people have better nutrition than others, and this can affect their growth. The **heritability** of a character is the relative contribution of genetic versus environmental factors to the variation in that character in a particular population. Typically, heritability varies from 0 to 1. For example, the heritability of human height varies from about 0.65 to 0.8, depending on the population studied. In a population where heritability is 0.65, 65 percent of the variation in height is due to genetic factors, while the remaining 35 percent is due to environmental effects. It is important to note that heritability estimates apply to variations within populations. They cannot be used to estimate the contribution of genetics to particular characters in an individual.

Heritability estimates are important for breeders of plants and animals because they provide information about whether it is more worthwhile to modify the environment or do genetic crosses to improve a phenotype.

CHECKpoint CONCEPT 8.2

- ✓ What is the difference between incomplete dominance and codominance?
- ✓ A point-restricted rabbit (see Figure 8.9) was mated with a chinchilla rabbit. The two offspring were albino and chinchilla. What were the genotypes of the parents?
- ✓ If a dominant allele of one gene, *A*, is necessary for hearing in humans, and the dominant allele of another gene, *B*, results in deafness regardless of the presence of other genes, what fraction of offspring in a pairing of *AaBb* and *Aabb* individuals will be deaf?
- ✓ Give an example from your own experience of a genotype whose expression is affected by the environment.

So far we have considered genes that obey Mendel's law of independent assortment. But many genes are inherited together, with multiple genes transmitted as one unit. This apparent anomaly can be explained by the presence of multiple genes on a single chromosome.

CONCEPT 8.3 Genes Are Carried on Chromosomes

Genes are parts of chromosomes. More specifically, a gene is a sequence of DNA that resides at a particular site on a chromosome called a **locus** (plural *loci*). You have seen how the behavior of chromosomes during meiosis can explain Mendel's laws of segregation (see Figure 8.3) and independent assortment (see Figure 8.6). However, the **genetic linkage** of genes on a single chromosome alters their pattern of inheritance.

Genetic linkage was first discovered in the fruit fly *Drosophila melanogaster*. This animal is an attractive experimental subject because it is small, easily bred, and has a short generation time (from fertilized egg to reproducing adult). In fact, the fruit fly has been a model organism for experimental genetics for more than a century. In Concept 8.1 we saw how Mendel successfully applied the scientific method to arrive at his laws of inheritance. Now we will examine the work of Thomas Hunt Morgan, who worked at Columbia University early in the twentieth century and used a similar approach to discover genetic linkage.

Genes on the same chromosome are linked, but can be separated by crossing over in meiosis

Some of the crosses Morgan performed with fruit flies yielded phenotypic ratios that were not in accordance with those predicted by Mendel's law of independent assortment. Morgan did a test cross between flies with two known genotypes: *BbVgvg* and *bbvgvg*. The *B* and *Vg* genes control two characters, body color and wing shape:

- *B* (wild-type gray body) is dominant over *b* (black body).
- *Vg* (wild-type wing) is dominant over *vg* (vestigial, or very small, wing).

INVESTIGATION

FIGURE 8.13 Some Alleles Do Not Assort Independently
Morgan's studies showed that the genes for body color and wing size in *Drosophila* are linked, so that their alleles do not assort independently.[a]

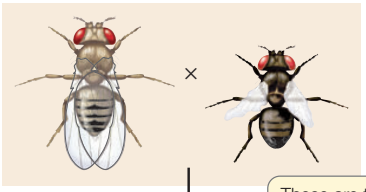
HYPOTHESIS

Alleles for different characteristics always assort independently.

METHOD

Parent (P)

BbVgvg
Wild type (gray body, normal wings)
♀



bbvgvg
(black body, vestigial wings)
♂

RESULTS

F_1

Genotypes	*BbVgvg* Gray normal	*bbvgvg* Black vestigial	*Bbvgvg* Gray vestigial	*bbVgvg* Black normal
Expected frequencies	575	575	575	575
Observed frequencies (number of individuals)	965	944	206	185
	Parental phenotypes		Recombinant phenotypes	

These are the results expected from Mendel's second law (independent assortment)…

…but the actual results were inconsistent with the law.

CONCLUSION

The hypothesis is rejected. These two genes do not assort independently, but are linked (on the same chromosome).

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]T. H. Morgan. 1912. *Science* 36: 719–720

Go to ANIMATED TUTORIAL 8.3
Alleles That Do Not Sort Independently
PoL2e.com/at8.3

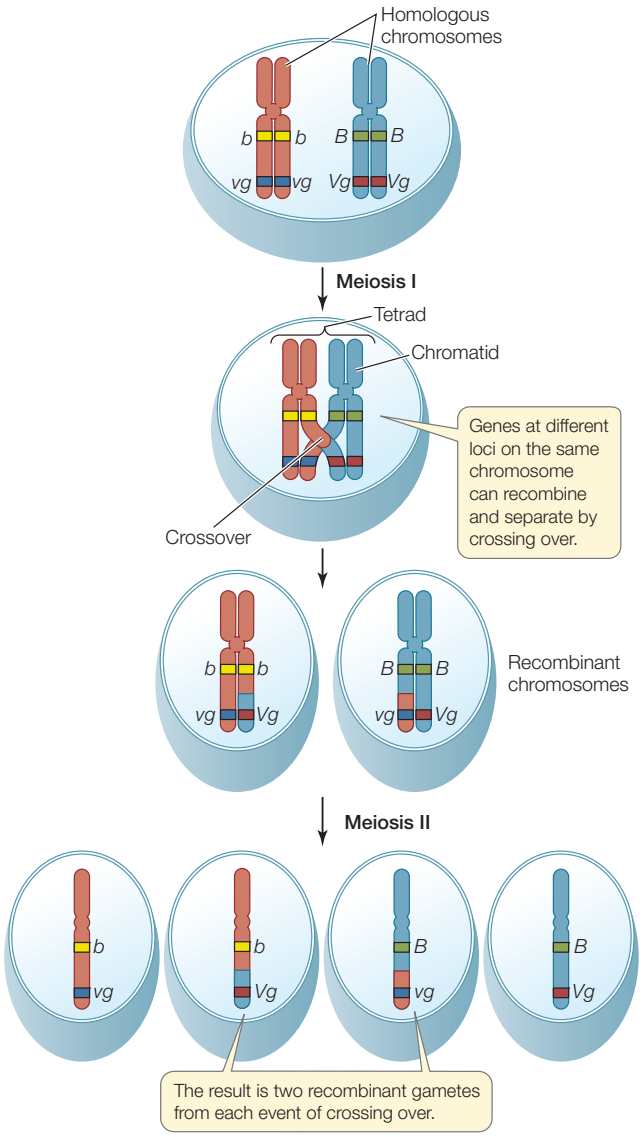


FIGURE 8.14 Crossing Over Results in Genetic Recombination
Recombination accounts for why linked alleles are not always inherited together. Alleles at different loci on the same chromosome can be recombined by crossing over and then being separated from one another. Such recombination occurs during prophase I of meiosis.

Morgan expected to see four phenotypes in a ratio of 1:1:1:1, but that is not what he observed. The body color gene and the wing size gene did not assort independently; instead, they were frequently inherited together, and most of the progeny showed one or the other of the parental phenotypes (**FIGURE 8.13**).

These results became understandable when Morgan considered the possibility that the two loci were *linked* on the same chromosome. Such genes would not be able to assort independently as predicted by Mendel's second law. In this case, the test cross offspring might be expected to have only the parental phenotypes (gray flies with normal wings or black flies with vestigial wings) in a 1:1 ratio. If linkage were absolute, we would see *only* these two types of progeny. However, this did not happen. Why did some of Morgan's flies show phenotypes different from their parents?

Some of Morgan's flies displayed **recombinant** phenotypes because two homologous chromosomes can physically exchange corresponding segments during prophase I of meiosis—that is, by crossing over (**FIGURE 8.14**; see also Figure 7.13). Each exchange event involves two of the four chromatids in a tetrad—one from each member of the homologous pair—and can occur at any point along the length of the chromosome. The chromosome segments are exchanged reciprocally, so both chromatids become recombinant (that is, each chromatid ends up with genes from both of the organism's parents).

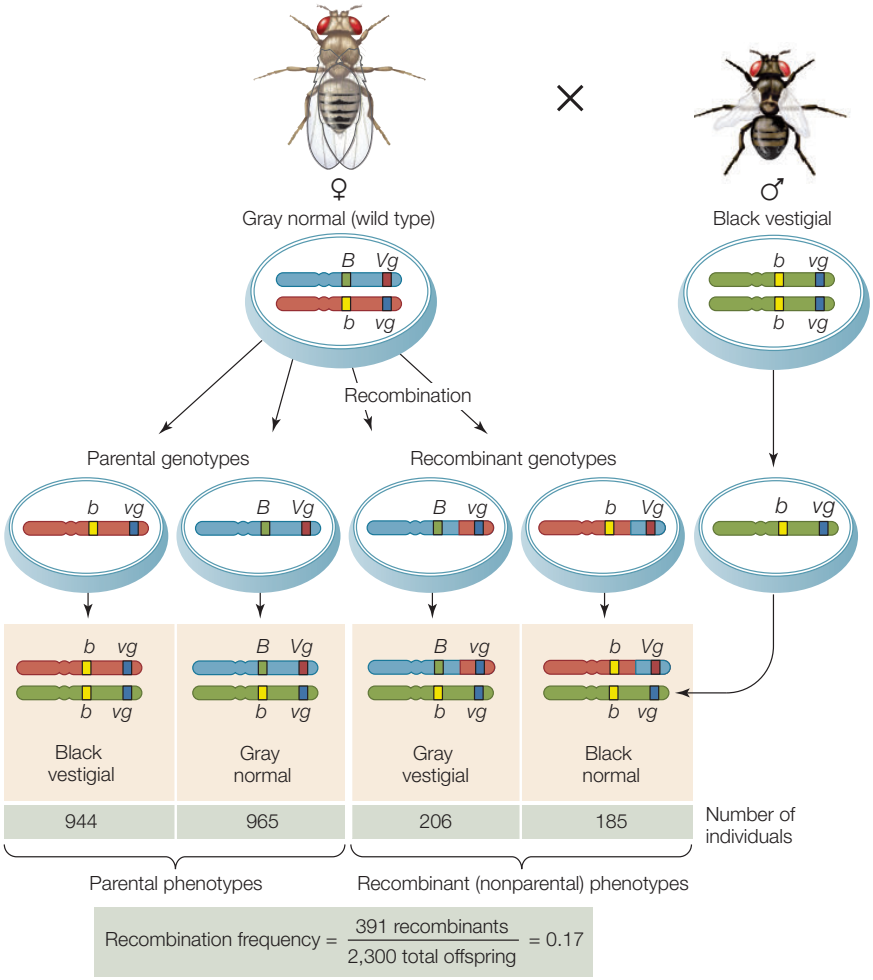
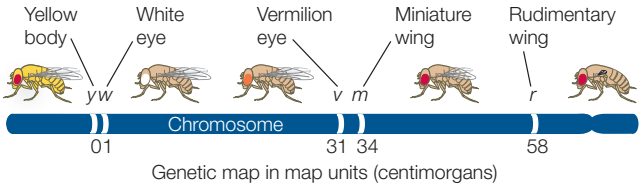


FIGURE 8.15 Recombination Frequencies The frequency of recombinant offspring (those with a phenotype different from either parent) can be calculated.

As a result of a crossing over event between two linked genes, not all the progeny of a cross have the parental phenotypes. Recombinant offspring appear as well, generally in proportions related to the **recombination frequency** between the two genes, which is calculated by dividing the number of recombinant progeny by the total number of progeny (**FIGURE 8.15**). *Recombination frequencies are greater for loci that are farther apart on the chromosome* than for loci that are closer together because crossing over is more likely to occur between genes that are far apart. By calculating recombination frequencies, geneticists can infer the locations of genes along a chromosome and generate a genetic map. Below is a map showing five genes on a fruit fly chromosome. It was constructed using the recombination frequencies generated by test crosses involving various pairs of the genes:

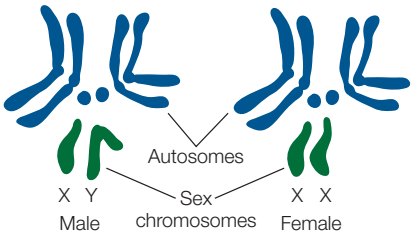


The recombination frequency between *y* and *w* is low, so they are close together on the map. Recombination between *y* and *v* is more frequent, so they are farther apart. The recombination frequencies are converted to map units (also called centimorgans, cM); one map unit is equivalent to an average recombination frequency of 0.01 (one percent).

The era of gene sequencing has made mapping less important in some areas of genetics research. However, mapping is still one way to verify that a particular DNA sequence corresponds with a particular phenotype. The phenomenon of linkage has allowed biologists to isolate genes and to create genetic markers that are linked to important genes, making it easy to identify individuals carrying particular alleles. This is particularly important in breeding new crops and animals for agriculture, and for identifying humans carrying medically significant mutations.

Linkage is also revealed by studies of the X and Y chromosomes

The fruit fly genome has four pairs of chromosomes: in three pairs, the chromosomes are similar in size to one another and are called **autosomes**. The fourth pair has two chromosomes of different sizes. These determine the sex of the fly and are called the **sex chromosomes**, shown here in green:



Note that the female fly has two X chromosomes and that the male has only one, the other being the Y chromosome: females are XX and males are XY. It turns out that in addition to being different sizes, *many genes on the X chromosome are not present on the Y*. This means that males have only one copy of these genes. The X chromosome was one of the first to have specific genes assigned to it.

Morgan identified a gene that controls eye color in *Drosophila*. The wild-type allele of the gene confers red eyes, whereas a recessive mutant allele confers white eyes. Morgan's

experimental crosses with flies carrying the mutant allele demonstrated that this eye color locus is on the X chromosome. If we abbreviate the eye color alleles as *R* (red eyes) and *r* (white eyes), the presence of the alleles on the X chromosome is designated by X^R and X^r.

Morgan crossed a homozygous red-eyed female (X^RX^R) with a white-eyed male. The male is designated X^rY because the Y does not carry any allele for this gene. (Any gene that is present as a single copy in a diploid organism is called **hemizygous**. A male will express all the alleles of his one X chromosome, whether or not they are dominant.) All the sons and daughters from this cross had red eyes, because red (*R*) is dominant over white (*r*) and all the progeny had inherited a wild-type X chromosome (X^R) from their mother (**FIGURE 8.16A**). Note that this phenotypic outcome would have occurred even if the *R* gene had been present on an autosome rather than a sex chromosome. In that case, the male would have been homozygous recessive—*rr*.

When Morgan performed the reciprocal cross, in which a white-eyed female (X^rX^r) was mated with a red-eyed male (X^RY), the results were unexpected: *all the sons were white-eyed and all the daughters were red-eyed* (**FIGURE 8.16B**). The sons from the reciprocal cross inherited their only X chromosome from their white-eyed mother and were therefore hemizygous for the white allele. The daughters, however, got an X chromosome bearing the white allele from their mother and an X chromosome bearing the red allele from their father; therefore they were red-eyed heterozygotes. When these heterozygous females were mated with red-eyed males, half their sons had white eyes but all their daughters had red eyes. Together, these results showed that eye color was carried on the X chromosome and not on the Y.

These and other experiments led to the term **sex-linked inheritance**: inheritance of a gene that is carried on a sex chromosome. (This term is misleading because "sex-linked" inheritance is not really linked to the sex of an organism—after all,

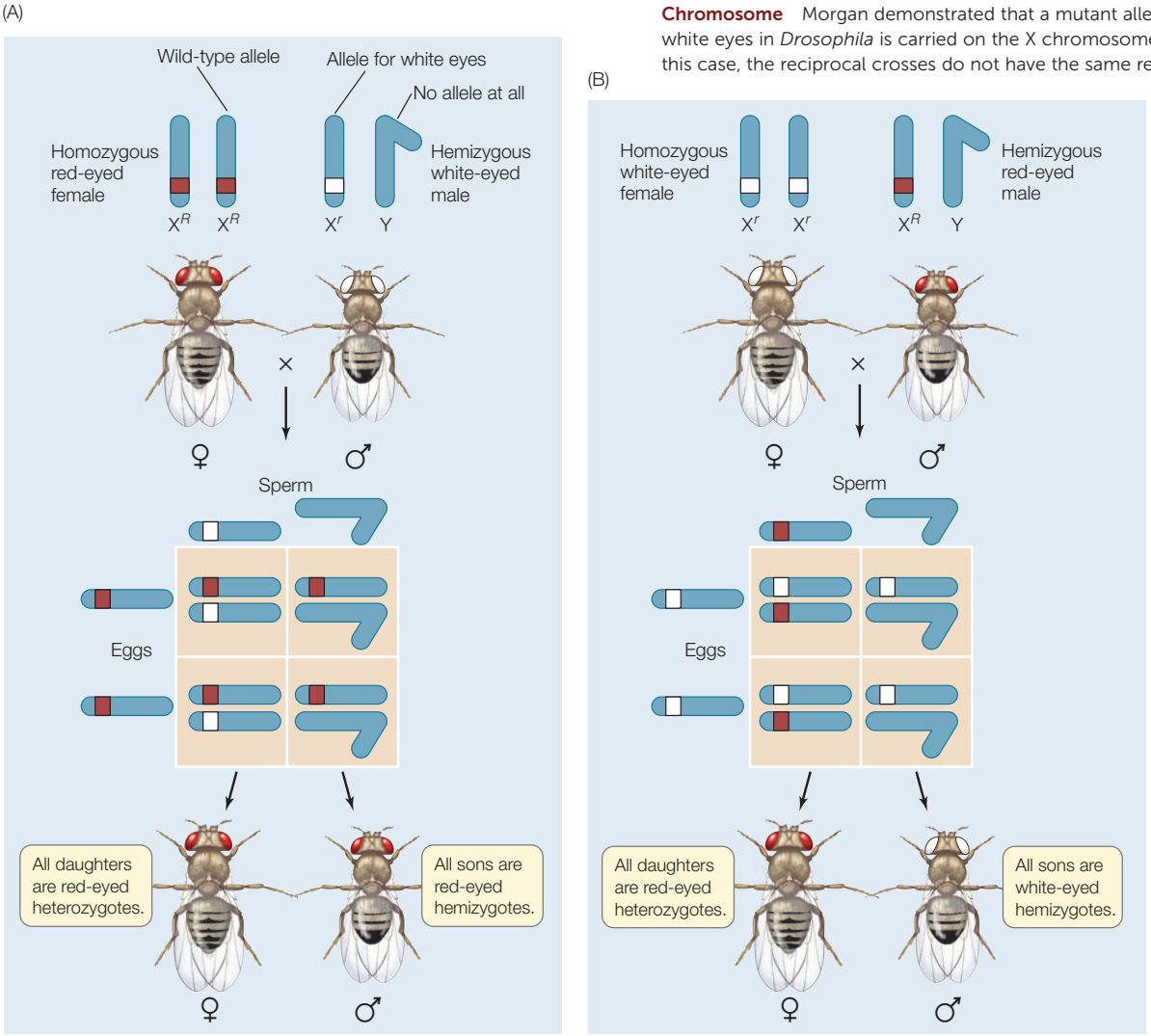


FIGURE 8.16 A Gene for Eye Color Is Carried on the *Drosophila* X Chromosome Morgan demonstrated that a mutant allele that causes white eyes in *Drosophila* is carried on the X chromosome. Note that in this case, the reciprocal crosses do not have the same results.

both males and females carry X chromosomes.) In mammals, the X chromosome is larger and carries more genes than the Y. For this reason, most examples of sex-linked inheritance involve genes that are carried on the X chromosome.

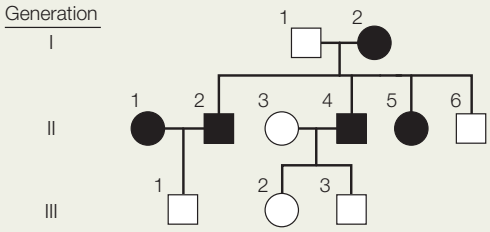
Many sexually reproducing species, including humans, have sex chromosomes. As in fruit flies, human males are XY, females are XX, and relatively few of the genes that are present on the X chromosome are present on the Y. Pedigree analyses of X-linked recessive phenotypes like the one in **FIGURE 8.17** reveal the following patterns (compare with the pedigrees of non-X-linked phenotypes in Figure 8.8):

- The phenotype appears much more often in males than in females, because only one copy of the allele is needed for its expression in males, whereas two copies must be present in females.
- A male with the mutation can pass it on only to his daughters; all his sons get his Y chromosome.
- Daughters who receive one X-linked mutation are heterozygous carriers. They are phenotypically normal, but they can pass the mutant allele to their sons or daughters. On average, half their children will inherit the mutant allele since half of their X chromosomes carry the normal allele.
- The mutant phenotype can skip a generation if the mutation passes from a male to his daughter (who will be phenotypically normal) and then to her son.

APPLY THE CONCEPT

Genes are carried on chromosomes

The pedigree shows the inheritance pattern of a rare mutant phenotype in humans, congenital cataract (filled-in symbols).



1. Are cataracts inherited as autosomal dominant? Autosomal recessive? Sex-linked dominant? Sex-linked recessive?
2. Person #5 in the second generation marries a man who does not have cataracts. Two of their four children, a boy and a girl, develop cataracts. What is the probability that their next child will be a girl with cataracts?

Some genes are carried on chromosomes in organelles

The nucleus is not the only organelle in a eukaryotic cell that carries genetic material. Mitochondria and plastids (including chloroplasts) each contain several copies of a small chromosome that carries a small number of genes. For example, in humans there are about 21,000 genes in the nuclear genome and 37 in the mitochondrial genome. Plastids have five times as many genes as mitochondria. But note that these organelle genomes do not encode all of the molecules that make up the organelle. Most of the proteins and some of the RNAs in organelles are encoded by the nuclear genome and imported from the cytoplasm.

The inheritance of organelle genes differs from that of nuclear genes because in most organisms, *mitochondria and plastids are inherited only from the mother*. Most egg cells in plants and animals contain abundant cytoplasm and organelles, but the only part of the sperm that survives to take part in the union of haploid gametes is the nucleus. You have inherited your mother's mitochondria (with their genes) but not your father's. The inheritance of organelles and their genes is therefore non-Mendelian and is described as maternal or cytoplasmic inheritance, since the inherited organelles come from the maternal cytoplasm.

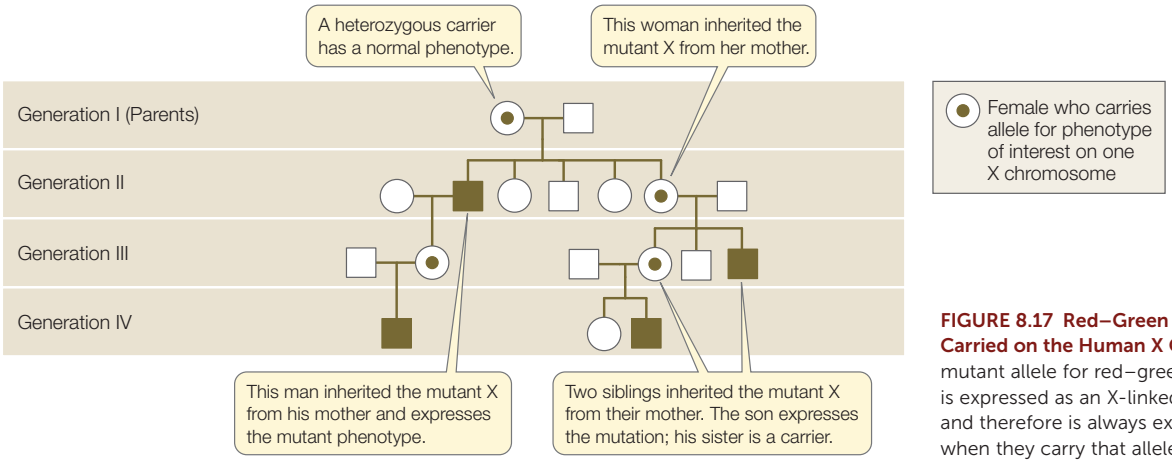


FIGURE 8.17 Red–Green Color Blindness Is Carried on the Human X Chromosome The mutant allele for red–green color blindness is expressed as an X-linked recessive trait, and therefore is always expressed in males when they carry that allele.

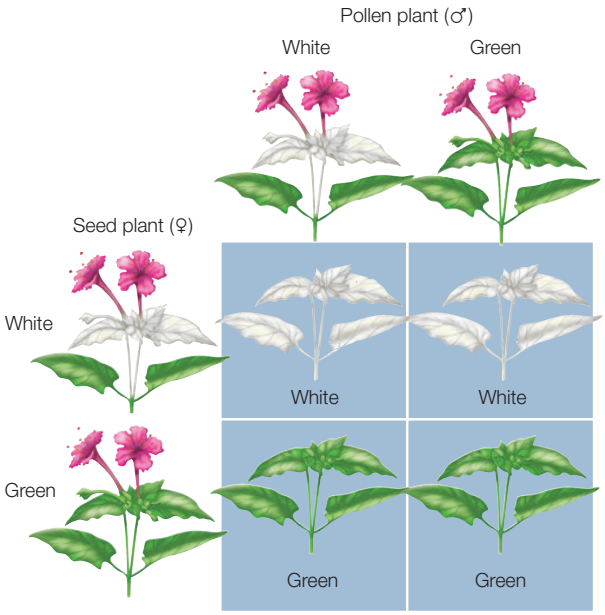


FIGURE 8.18 Cytoplasmic Inheritance In four o'clock plants, leaf color is inherited through the female plant only. In the parent plant with some white leaves, the white leaf color is caused by a chloroplast mutation that occurred during the life of the plant; the leaves that formed before the mutation occurred are green. When this plant is used as a pollen donor in a cross with an all green plant, the offspring are all green. But when the same plant is used as an egg donor, the offspring inherit the mutation cytoplasmically and are entirely white.

Some of the genes carried by cytoplasmic organelles are important for organelle assembly and function, and mutations of these genes can have profound effects on the organism. For example, in plants, certain plastid gene mutations affect the proteins that assemble chlorophyll molecules into photosystems (see Figure 6.19). These mutations result in a phenotype that is essentially white instead of green. **FIGURE 8.18** illustrates the cytoplasmic inheritance of a mutant plastid gene that confers white color to leaves.

CHECKpoint CONCEPT 8.3

- ✓ Describe the differences in patterns of inheritance between a gene present in the nucleus and a gene present in the mitochondria.
- ✓ Explain the concept of linkage. If you performed a test cross with a fruit fly that is heterozygous for two genes, how would you conclude that the two genes are linked?
- ✓ Red–green color blindness is inherited as a sex-linked recessive. Two parents with normal color vision have a child who is red–green color-blind. Is the child a boy or girl? Draw a pedigree of the family. Draw a Punnett square to show gametes and offspring with regard to the X and Y chromosomes and the normal and color-blind alleles.

Like eukaryotes, prokaryotes contain genes that determine their phenotypes. Sexual reproduction in eukaryotes involves two sets of chromosomes and meiosis, giving rise to haploid gametes. Let's look next at how reproduction and inheritance differ in prokaryotes, which are haploid.

CONCEPT 8.4 Prokaryotes Can Exchange Genetic Material

As described in Concept 4.2, prokaryotic cells lack nuclei; they contain their genetic material mostly as single chromosomes in central regions of their cells. Prokaryotes reproduce asexually by binary fission, a process that gives rise to progeny that are virtually identical genetically (see Concept 7.2). That is, the offspring of cell reproduction in prokaryotes constitute a clone.

How, then, do prokaryotes evolve? Mutations occur in prokaryotes just as they do in eukaryotes, and the resulting new alleles increase genetic diversity. You might expect, therefore, that there is no way for individuals of these organisms to exchange genes as happens in sexual reproduction. It turns out, however, that prokaryotes *do* have a way of transferring genes between cells. This transfer of genes from one individual organism to another without sexual reproduction is called **horizontal** or **lateral gene transfer** to distinguish it from vertical gene transfer (gene transfer from parent to offspring). Along with mutation, this process generates genetic diversity among prokaryotes.

LINK

The evolutionary consequences of lateral gene transfer and its role in identifying and classifying bacterial species are discussed in **Concepts 15.6** and **19.2**

Bacteria exchange genes by conjugation

To illustrate genetic exchange in bacteria, let's consider two strains of the bacterium *E. coli* with different alleles for each of six genes (which code for enzymes in a biochemical pathway that synthesizes a certain small molecule). The two strains have the following genotypes (remember that bacteria are haploid):

ABCdef and *abcDEF*

where capital letters stand for wild-type alleles that encode functional gene products and lowercase letters stand for mutant alleles that encode defective gene products. Neither strain is able to synthesize the small molecule because neither has a full complement of wild-type genes.

When the two strains are grown together in the laboratory, most of the cells produce clones. That is, almost all of the cells that grow have either genotype *ABCdef* or genotype *abcDEF*. However, out of millions of bacteria, a few occur that have the genotype *ABCDEF*.

Unlike the original strains, cells of genotype *ABCDEF* are able to synthesize the small molecule. How could these completely wild-type bacteria arise? One possibility is mutation: in

the *abcDEF* bacteria, the *a* allele could have mutated to *A*, the *b* allele to *B*, and the *c* allele to *C*. The problem with this explanation is that a mutation at any particular point in an organism's DNA sequence is a very rare event (about 1 in a million). The probability of all three events occurring in the same cell is extremely low—much lower than the observed rate of appearance of cells with genotype *ABCDEF*. So the mutant cells must have acquired wild-type genes some other way—and this turns out to be the transfer of DNA between cells.

Electron microscopy shows that gene transfers between bacteria can happen via physical contact between the cells (**FIGURE 8.19A**). Contact is initiated by a thin projection called a **sex pilus** (plural *pili*) that extends from one cell (the donor), attaches to another cell (the recipient), and draws the two cells together. Genetic material can then pass from the donor to the recipient through a thin cytoplasmic bridge called a conjugation tube. There is no reciprocal transfer of DNA from the recipient to the donor. This process is referred to as **bacterial conjugation**.

Once the donor DNA is inside the recipient cell, it can recombine with the recipient cell's genome. In much the same way that chromosomes pair up in prophase I of meiosis, the donor DNA can line up beside its homologous genes in the recipient, and crossing over can occur. Gene(s) from the donor can become integrated into the genome of the recipient, thus changing the recipient's genetic constitution (**FIGURE 8.19B**). In general, about half the transferred genes become integrated in this way. When the recipient cells proliferate, the integrated donor genes are passed on to all progeny cells, and the other transferred genes are lost. If the new combination of alleles is advantageous, the progeny of the recipient cell may be able to proliferate faster than the original strains.

Plasmids transfer genes between bacteria

In addition to their main chromosome, many bacteria harbor additional smaller, circular DNA molecules called **plasmids** that replicate independently inside the cell. Plasmids typically contain at most a few dozen genes, which may fall into one of several categories, including:

- Genes for unusual metabolic capacities, such as the ability to break down hydrocarbons. Bacteria carrying these plasmids can be used to clean up oil spills.
- Genes for antibiotic resistance. Plasmids carrying such genes are called R factors, and since they can be transferred between bacteria via conjugation, they are a major threat to human health.

Plasmids can move between cells during conjugation, thereby transferring new genes to the recipient bacterium (**FIGURE 8.20**). A single strand of the donor plasmid is transferred to the recipient; synthesis of complementary DNA strands results in two complete copies of the plasmid, one in the donor and one in the recipient. Because plasmids can replicate independently of the main chromosome, they do not need to recombine with the main chromosome to add their genes to the recipient cell's genome.

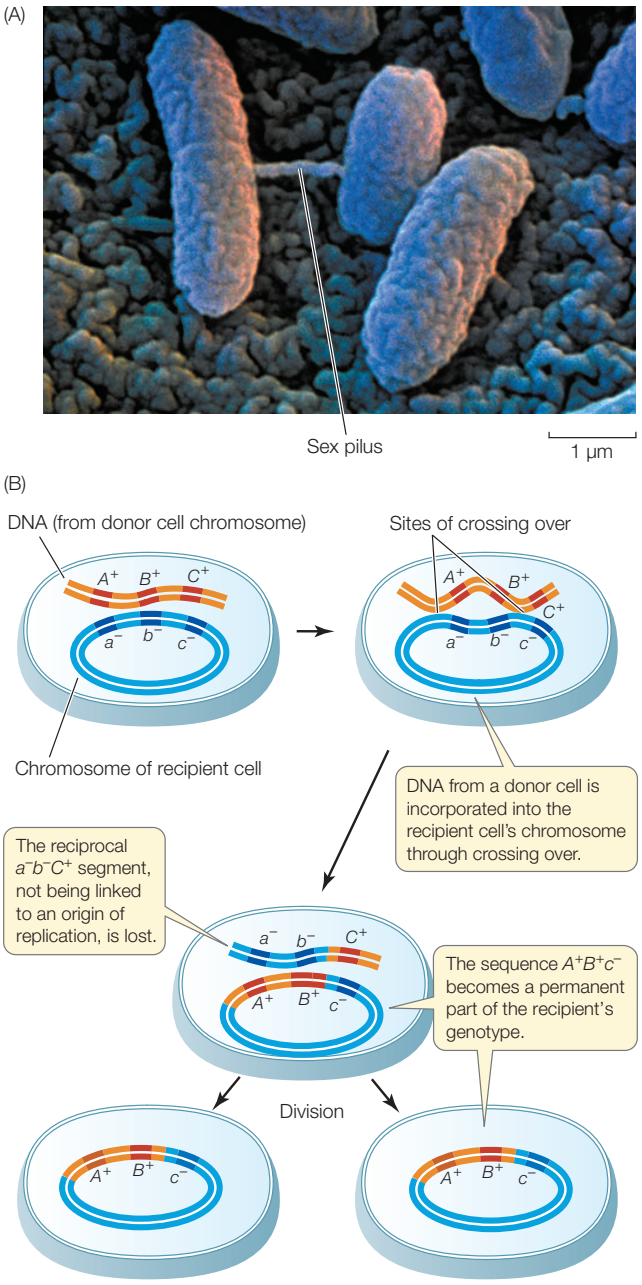


FIGURE 8.19 Bacterial Conjugation and Recombination (A) A sex pilus draws two bacteria into close contact, so that a cytoplasmic bridge (conjugation tube) can form. DNA is transferred from the donor cell to the recipient cell via the conjugation tube. (B) DNA from a donor cell can become incorporated into a recipient cell's chromosome through crossing over.

The evolution of drug-resistant bacteria is a major public health problem

Until the twentieth century, bacterial infections were a major scourge of humanity. With the discovery of antibiotics

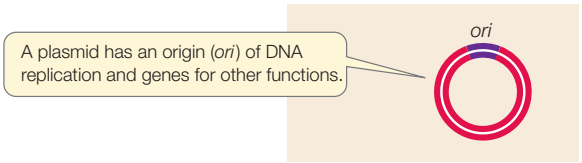


FIGURE 8.20 Gene Transfer by Plasmids When plasmids enter a cell via conjugation, their genes can be expressed in the recipient cell.

(particularly penicillin, which prevents the assembly of the bacterial call wall), many lethal infections were kept at bay. But over time some bacteria acquired mutations that rendered them resistant to penicillin. These bacteria had a selective advantage when faced with penicillin, and as penicillin use increased, the resistant bacteria became widespread. So scientists designed chemical variants, such as methicillin and vancomycin, to attack penicillin-resistant bacteria. With the development of strains of bacteria resistant to these antibiotics as well, the "arms race" has continued. The latest weapon against resistant bacteria is a group of antibiotics called carbapenems (for example, colistin), but bacteria resistant to these antibiotics are beginning to appear. Some resistance genes are carried on plasmids and are transferred between bacterial species by conjugation. This poses a major public health problem worldwide, since the new antibiotics are currently the last line of defense against lethal infections.

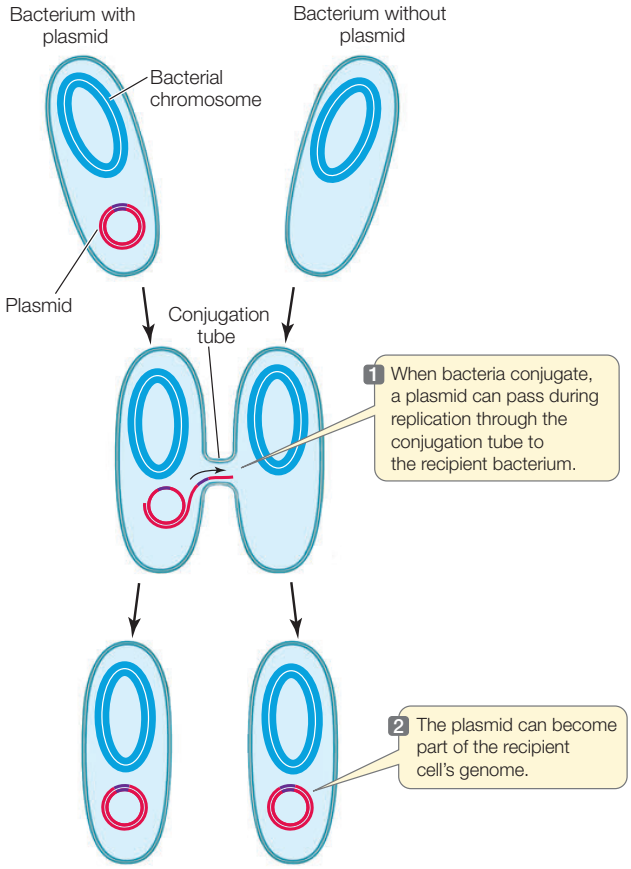
CHECKpoint CONCEPT 8.4

- ✓ How does recombination occur in prokaryotes?
- ✓ What is the evolutionary advantage of recombination in prokaryotes?
- ✓ What are the differences between recombination after conjugation in prokaryotes and recombination during meiosis in eukaryotes?

Q How is hemophilia inherited, and why is it more frequent in males?

ANSWER The ancient rabbis in the opening story were dealing with male babies that bled to death when they were cut. We know this as the blood-clotting disease hemophilia. The mutant allele of the gene coding for a blood clotting factor missing in hemophilia must be recessive, because the babies' parents did not suffer from the disease (Concept 8.1). The rabbis noted that the disease occurred in boys, and any relatives with the disease were males on the mother's side of the family. This is because the gene for the clotting factor is carried on the human X chromosome and its inheritance is sex-linked. Females have two X chromosomes, so even if they receive a mutant X chromosome from their mother, the second X chromosome, inherited from their father, will usually provide sufficient functional clotting factor. Males, however, have a single X chromosome, always inherited from their mother (Concept 8.3). If they receive their mother's recessive mutant chromosome (a 50–50 probability), they cannot produce clotting factor and thus suffer from hemophilia.

Hemophilia played a role in the history of modern Europe. England's Queen Victoria, who ruled for much of the nineteenth century, had nine children, one of whom, Leopold, had hemophilia and died after a minor accident at age 31. The queen did not have the disease, nor was it present in any of her forbears, so it is probable that a mutation arose spontaneously on one of Victoria's X chromosomes or that of her father (Concept 8.2). Three of Victoria's grandchildren had hemophilia, indicating that their mothers were carriers of the mutant allele. The mutation was thus introduced into the royal families of Spain, Germany, and Russia (**FIGURE 8.21**). While there are many descendants of Queen Victoria still living (including Queen Elizabeth II), none have hemophilia. It remains possible, however, that some of Victoria's female descendants are carriers.



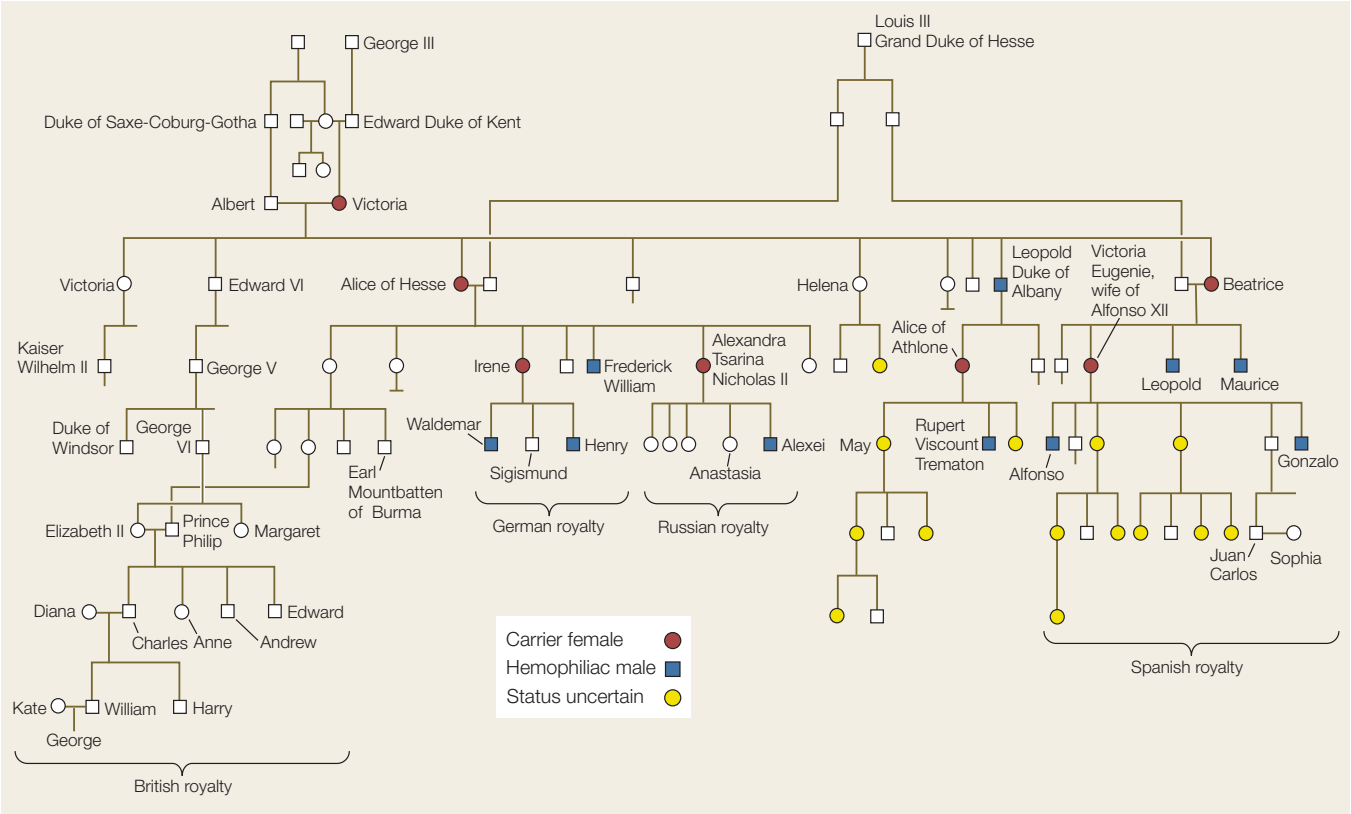


FIGURE 8.21 Sex Linkage in Royal Families of Europe England's Queen Victoria passed an X chromosome carrying the mutant allele for hemophilia to three of her children.

SUMMARY

CONCEPT 8.1 Genes Are Particulate and Are Inherited According to Mendel's Laws

- Mendel's experiments on pea plants supported the particulate theory of inheritance stating that discrete units (now called genes) are responsible for the inheritance of specific traits. **Review Figure 8.1**
- Mendel's first law, the **law of segregation**, states that when any individual produces gametes, the two copies of a gene separate, so that each gamete receives only one member of the pair. **Review Figures 8.2 and 8.3**
- Mendel used a **test cross** to find out if an individual with a dominant phenotype was homozygous or heterozygous for that phenotype. **Review Figure 8.4 and ACTIVITY 8.1**
- Mendel's use of **dihybrid crosses** to study the inheritance of two characters led to his second law: the **law of independent assortment**. The independent assortment of genes in meiosis leads to novel combinations of phenotypes. **Review Figures 8.5 and 8.6 and ANIMATED TUTORIAL 8.1**
- **Pedigree** analysis can determine whether an allele is **dominant** or **recessive**. **Review Figure 8.8 and ANIMATED TUTORIAL 8.2**

CONCEPT 8.2 Alleles and Genes Interact to Produce Phenotypes

- New alleles arise by random **mutation**. Many genes have multiple alleles. A **wild-type** allele gives rise to the predominant form of a trait. When the wild-type allele is present at a locus less than 99 percent of the time, the locus is said to be **polymorphic**.
- In **incomplete dominance**, neither of two alleles is dominant. The heterozygous phenotype is intermediate between the homozygous phenotypes. **Review Figure 8.10**
- **Codominance** exists when two alleles at a locus produce two different phenotypes that both appear in heterozygotes. **Review Figure 8.11**
- In **epistasis**, one gene affects the expression of another. **Review Figure 8.12**
- Environmental conditions can affect the expression of a genotype.
- **Heritability** is the relative contribution of genetic versus environmental factors to the variation in a character in a population.

CONCEPT 8.3 Genes Are Carried on Chromosomes

- Each chromosome carries many genes, and the genes on a single chromosome show **genetic linkage**. **Review Figure 8.13 and ANIMATED TUTORIAL 8.3**
- Genes on the same chromosome can recombine by crossing over. The resulting **recombinant** chromosomes have new combinations of alleles. **Review Figure 8.14**
- **Recombination frequencies** can be used to generate a genetic map of a chromosome. **Review Figure 8.15**
- In fruit flies and mammals, the X chromosome carries many genes, but the Y chromosome has only a few. Males have only one allele (are **hemizygous**) for X-linked genes, so recessive sex-linked mutations are expressed phenotypically more often in males than in females. Females may be unaffected **heterozygous carriers** of such alleles. **Review Figure 8.16**
- Some genes are present on the chromosomes of organelles such as plastids and mitochondria. In many organisms, cytoplasmic genes are inherited only from the mother because the male gamete contributes only its nucleus (i.e., no cytoplasm) to the zygote at fertilization. **Review Figure 8.18**

CONCEPT 8.4 Prokaryotes Can Exchange Genetic Material

- Prokaryotes reproduce asexually but can transfer genes from one cell to another in a process called **bacterial conjugation**. **Review Figure 8.19**
- **Plasmids** are small, extra DNA molecules in bacteria that carry genes involved in important metabolic processes. Plasmids can be transmitted from one cell to another. **Review Figure 8.20**

See **ACTIVITIES 8.2** and **8.3** for a concept review of this chapter.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is8

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

9

DNA and Its Role in Heredity

KEY CONCEPTS

Michael Crichton's novel *Jurassic Park* was based on the fictional premise that DNA retrieved from fossils could produce living dinosaurs, such as this *Triceratops*.

Jurassic Park, in both its literary and film incarnations, featured a fictional theme park populated with live dinosaurs. In the story, scientists isolated DNA from dinosaur blood extracted from the digestive tracts of fossil insects. The insects supposedly sucked the reptiles' blood immediately before being preserved in amber (fossilized tree resin). According to the novel, this DNA could be manipulated to produce living individuals of long-extinct organisms such as velociraptors and the famous *Tyrannosaurus rex*.

The late Michael Crichton got the idea for his novel from an actual scientific paper in which the authors cracked open amber that was 40 million years old and extracted DNA from a fossilized bee that had been trapped inside. Other scientists had reported on ancient DNA from amber-trapped termites and gnats. Then several reports emerged of DNA from 80-million-year-old dinosaur bones. Unfortunately, upon additional study, these "preserved" DNAs turned out to be contamination—either from microorganisms living in the surrounding soil or even from the scientists studying the samples. In fact, one of the supposed dinosaur DNAs turned out to be from the human Y chromosome.

It is unlikely that any long DNA polymers would survive over millions of years. The oldest fossilized insects in amber are about 40 million years old, and the dinosaurs died out about 65 million years ago. Nevertheless, the huge success of Crichton's book brought ancient DNA to the attention of millions of people, including biologists who study the evolution of life on Earth. DNA samples have been isolated from the remains of entire ecosystems of organisms preserved for many thousands of years in permafrost. With improved methods for DNA analysis, large portions of these organisms' genomes are being sequenced.

Methods to replicate tiny amounts of DNA and keep it from contamination have improved, and attention has turned to ancient human DNA. For example, DNA samples have been studied from people whose bodies were preserved in ice, such as the "Tyrolean Iceman" who died in the Austrian Alps 5,300 years ago. There is even a Neandertal Genome Project to analyze the DNA from preserved specimens of *Homo neanderthalensis*, a species that lived in Europe at the same time as early humans, between 350,000 and 30,000 years ago.

Q What can we learn from ancient DNA?

You will find the answer to this question on page 192.

CONCEPT 9.1 DNA Structure Reflects Its Role as the Genetic Material

In Chapter 8 we described Mendel's experiments in the 1860s demonstrating that genes are physically distinct entities, and other work in the early twentieth century showing that groups of genes are linked together. By the early twentieth century, a "chromosomal theory of inheritance" had been developed, proposing that Mendel's genes are present in the chromosomes of the cell nucleus. This theory came partly from observations of sea urchins: it was shown that an entire set of chromosomes must be present for a sea urchin embryo to grow and develop. Scientists also observed that homologous chromosomes are paired during meiosis, that crossing over occurs during meiosis I (see Figure 8.14), and that the chromosome pairs separate independently at anaphase I. Thus the behavior of chromosomes accounted for Mendel's laws of segregation and independent assortment, as well as the later discoveries of linkage and recombination.

We now turn to the actual chemical nature of genes, beginning with the evidence that DNA is the carrier of heritable information. Scientists used two types of evidence to show that DNA is the genetic material: circumstantial and experimental. We will provide examples of both types.

Circumstantial evidence suggested that DNA is the genetic material

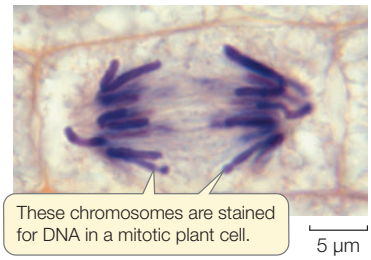
Early observations pointed to the possibility that DNA is the genetic material. Scientists found that DNA:

- is present in the cell nucleus and in condensed chromosomes
- doubles during S phase of the cell cycle
- is twice as abundant in the diploid cells as in the haploid cells of a given organism.

Let's look at some of these lines of evidence.

DNA IN THE NUCLEUS DNA was first isolated in 1868 by the young Swiss researcher–physician Friedrich Miescher, who isolated cell nuclei from white blood cells in pus from the bandages of wounded soldiers. When he treated these nuclei chemically, a fibrous substance came out of solution. He called it "nuclein" and found that it contained the elements C, H, O, N, and P. With no evidence except for finding it in the nucleus, Miescher boldly proposed that nuclein was the genetic material. His supervising professor was so astounded by Miescher's work that he repeated it himself in the laboratory, and finally allowed his student to publish it in a scientific journal.

DNA IN THE CHROMOSOMES In the early twentieth century dyes were developed that react specifically with DNA, only showing color when they bind to it. This allowed individual cells to be examined for the location and amount of DNA they contained. When dividing cells were stained with such a dye, only the chromosomes were stained:



DNA AMOUNTS The amount of dye binding to DNA, and hence the intensity of color observed, was directly related to the amount of DNA present: the greater the intensity, the more DNA. This allowed scientists to analyze DNA amounts in individual cells during the cell cycle (see Concept 7.2).

When a population of actively dividing cells was stained with dye, the amount of DNA in each cell could be quantified by passing the cells one by one through an instrument called a flow cytometer. In general, two populations of cells were seen: most cells were in G1 and contained half the amount of DNA that was in the remaining cells, which were in S, G2, or M (**FIGURE 9.1**). Such staining experiments confirmed two other predictions for DNA as the genetic material:

- Virtually all nondividing somatic cells of a particular organism have the same amount of nuclear DNA. This amount varies from species to species.
- Similar experiments showed that after meiosis, gametes have half the amount of nuclear DNA as somatic cells.

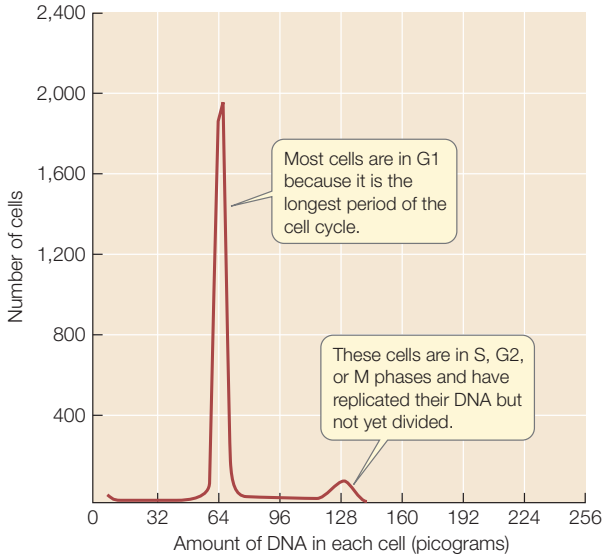


FIGURE 9.1 DNA in the Cell Cycle When dividing cells are stained and analyzed by flow cytometry, there are two populations in terms of DNA content, seen as two peaks in the above graph.

Experimental evidence confirmed that DNA is the genetic material

Circumstantial evidence can show correlations between two phenomena. However, *scientists rely on experiments to provide evidence of a cause-and-effect relationship*. Chromosomes in eukaryotic cells contain DNA, but they also contain proteins that are bound to DNA. Therefore it was difficult to rule out the possibility that genetic information might be carried in proteins. In order to confirm that DNA was the genetic material, biologists used model organisms such as bacteria in **transformation** experiments. They found, for example, that the addition of DNA from one strain of bacterium could genetically transform another strain:

Bacterium strain A + strain B DNA → bacterium strain B

Viruses provided another system to explore this question. Many viruses, including **bacteriophage** (viruses that infect bacteria), are composed of DNA and only one or a few kinds of protein. When a bacteriophage infects a bacterium, it takes about 20 minutes for the virus to hijack the bacterium's metabolic capabilities and turn the bacterium into a virus factory. Minutes later, the bacterium is dead and hundreds of viruses are released.

The transition from bacterium to virus producer is a change in the genetic program of the bacterial cell, resulting in a change of phenotype. Experiments showed that *only the viral DNA* is injected into the cell during infection (**FIGURE 9.2**). Since the viral DNA genetically transformed the bacteria, this was further evidence that DNA and not protein is the genetic material.

The transformation of mammalian cells with a gene for antibiotic resistance provided another model system for showing that DNA is the genetic material (**FIGURE 9.3**). When cultured mammalian cells were treated with DNA containing a gene for resistance to the antibiotic neomycin, the cells were able to grow on media containing the antibiotic.

Many kinds of cells can be transformed in this way—even egg cells. In this case, a whole new genetically transformed organism can result. The fertilized egg can develop into a new multicellular organism through mitosis; such an organism is referred to as **transgenic**. These methods form the basis of much applied research, including biotechnology and genetic engineering. The transformation of multicellular eukaryotes provides powerful experimental evidence for DNA as the genetic material.

THE DISCOVERY OF THE THREE-DIMENSIONAL STRUCTURE OF DNA WAS A MILESTONE IN BIOLOGY Mendel showed that genes are physically distinct entities, and further research identified DNA as the genetic material. The history of how the actual structure of DNA was deciphered is worth considering, as it represents not only talented scientists working together, but also a landmark in our understanding of biology.

By the mid-twentieth century, the chemical makeup of DNA, as a polymer made up of nucleotide monomers, had been known for several decades. In determining the structure of DNA, scientists hoped to answer two additional questions:

- How is DNA replicated between cell divisions?
- How does it direct the synthesis of specific proteins?

They were eventually able to answer both questions. The structure of DNA was deciphered only after many types of experimental evidence were considered together.

X-RAY CRYSTALLOGRAPHY PROVIDED CLUES TO DNA'S STRUCTURE The most crucial evidence was obtained using X-ray crystallography. Some chemical substances, when they are isolated and purified, can be made to form crystals. The positions of atoms in a crystallized substance can be inferred from the diffraction pattern of X rays passing through the substance (**FIGURE 9.4A**). The structure of DNA would not have been characterized without the crystallographs prepared in the early 1950s by the English chemist Rosalind Franklin (**FIGURE 9.4B**). Franklin's work, in turn, depended on the success of the English biophysicist Maurice Wilkins, who prepared samples containing very uniformly oriented DNA fibers. These fibers and the crystallographs Franklin prepared from them suggested a spiral or helical molecule.

THE NUCLEOTIDE COMPOSITION OF DNA WAS KNOWN The chemical composition of DNA also provided important clues

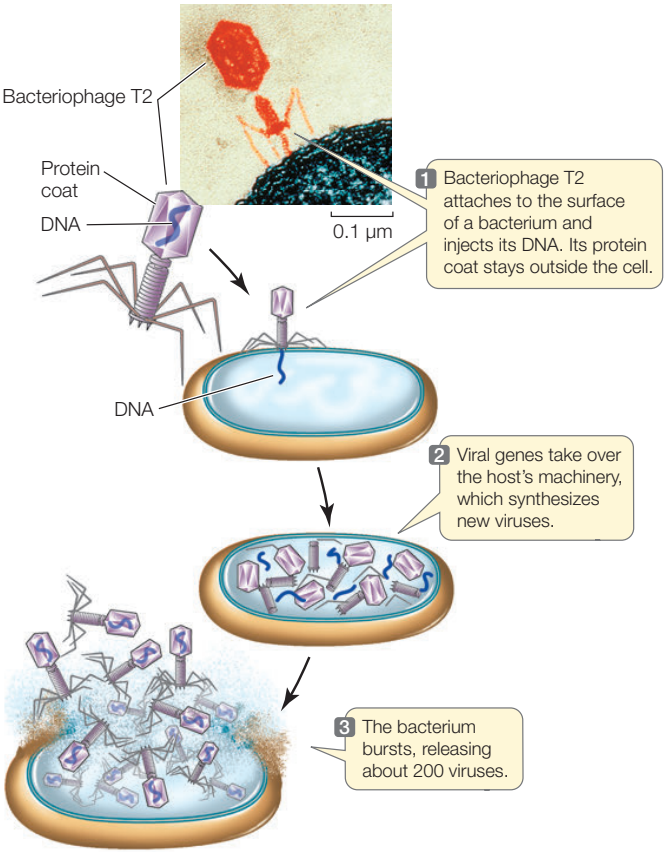


FIGURE 9.2 Viral DNA and Not Protein Enters Host Cells
Bacteriophage T2 infects *E. coli* and depends on the bacterium to produce new viruses. The bacteriophage consists entirely of DNA contained within a protein coat. When the virus infects an *E. coli* cell, its DNA, but not its protein coat, is injected into the host bacterium.

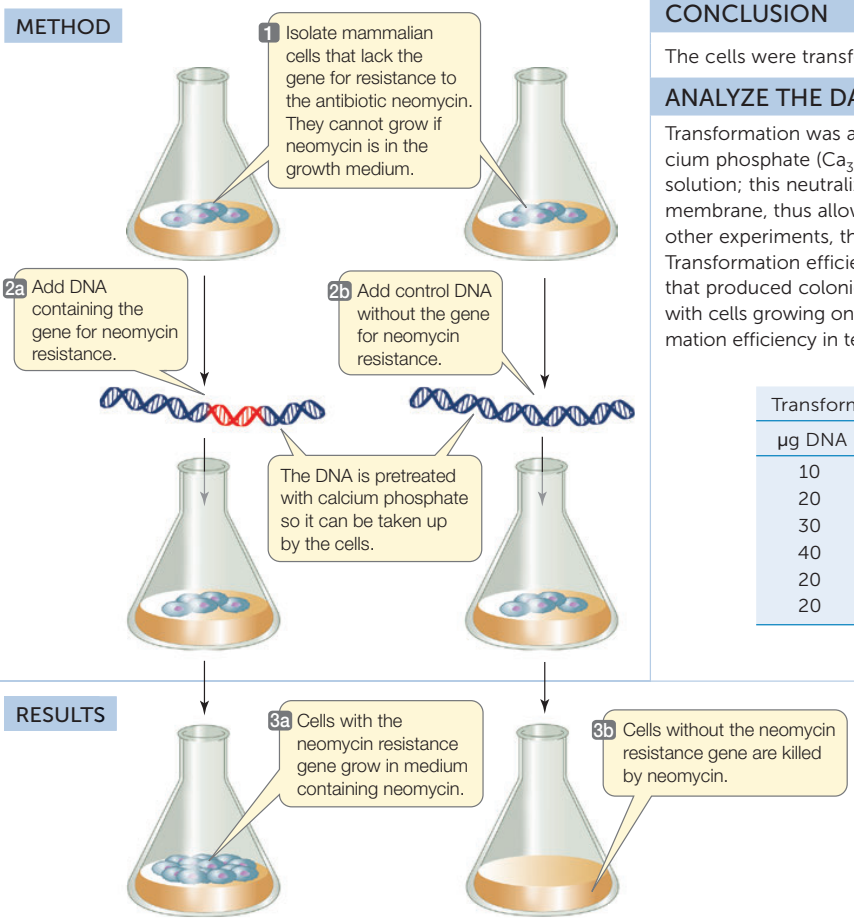
INVESTIGATION

FIGURE 9.3 Transformation of Eukaryotic Cells A DNA molecule can be treated chemically so that it is taken up from a solution by mammalian cells. The inclusion of an antibiotic resistance gene shows that the cells have been genetically transformed by the DNA.[a]

HYPOTHESIS

DNA can transform eukaryotic cells.

METHOD



RESULTS

CONCLUSION

The cells were transformed by DNA.

ANALYZE THE DATA

Transformation was achieved by adding the DNA in a solution of calcium phosphate ($Ca_3[PO_4]_2$) at pH 6.95. $Ca_3(PO_4)_2$ produces Ca^{2+} in solution; this neutralizes negative charges on the DNA and on the cell membrane, thus allowing the DNA to pass through the membrane. In other experiments, the type or amount of DNA and pH were varied. Transformation efficiency was calculated as the percentage of cells that produced colonies on a medium containing neomycin, compared with cells growing on medium without neomycin. Explain the transformation efficiency in terms of the conditions given in the data.

Transformation conditions		Transformation efficiency (%)
μg DNA	pH	
10	6.95	15
20	6.95	50
30	6.95	10
40	6.95	7
20	6.83	0
20	7.12	2

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]C. Chen and H. Okayama. 1987. *Molecular and Cellular Biology* 7: 2745–2752.

(A) (B)

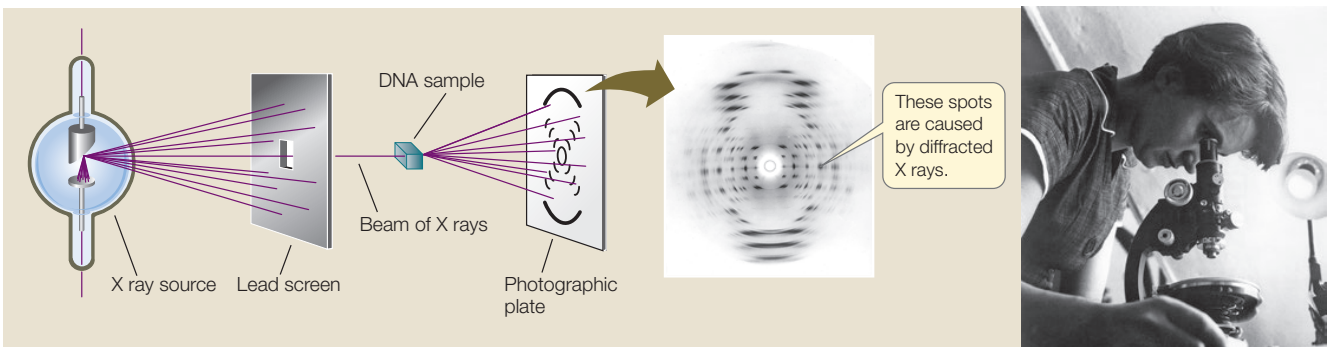


FIGURE 9.4 X-Ray Crystallography Helped Reveal the Structure of DNA (A) The positions of atoms in a crystallized chemical substance can be inferred by the pattern of diffraction of X rays passed through it. The pattern of DNA is both highly regular and repetitive. (B) Rosalind Franklin's crystallographs helped other scientists visualize the helical structure of the DNA molecule.

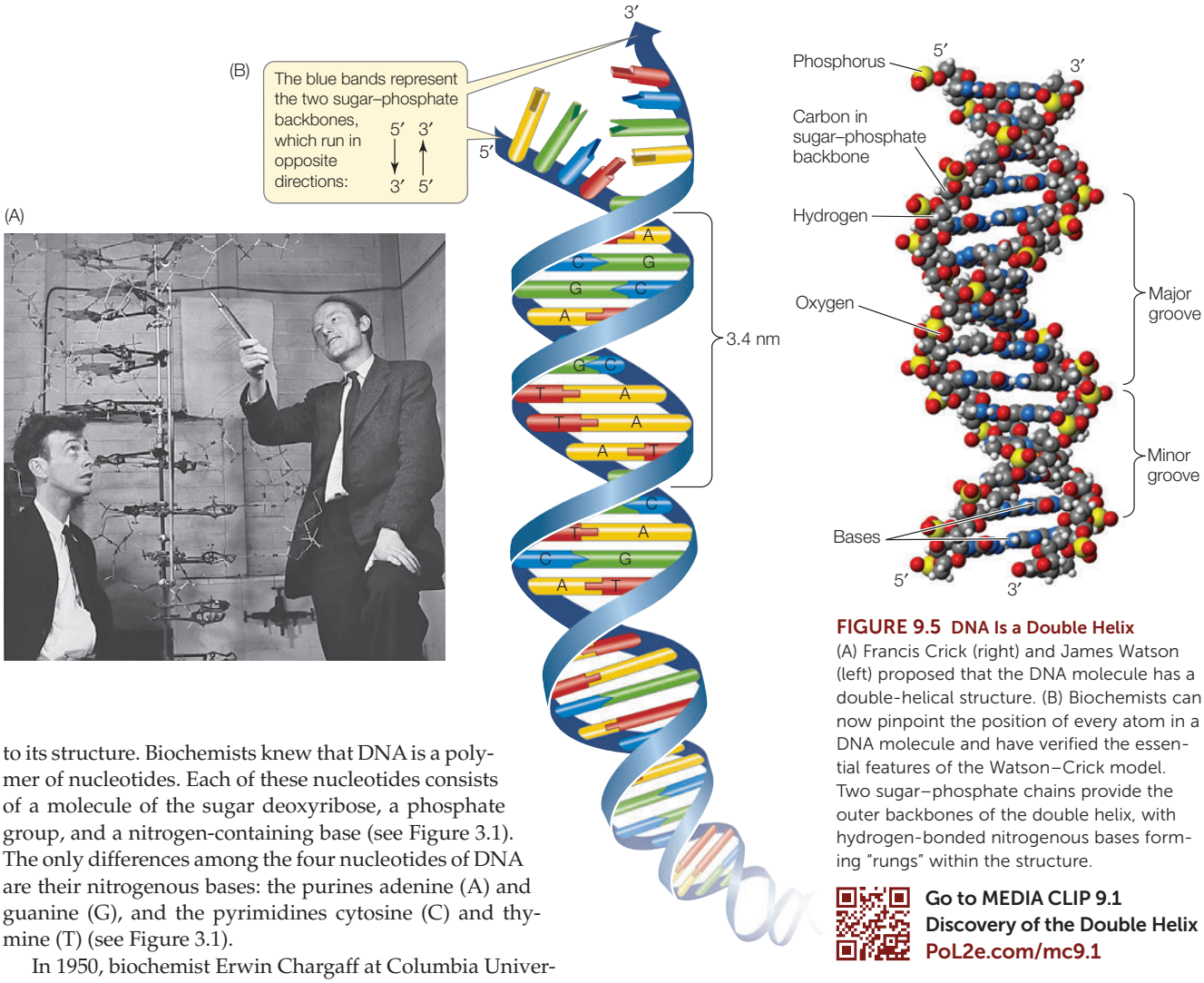
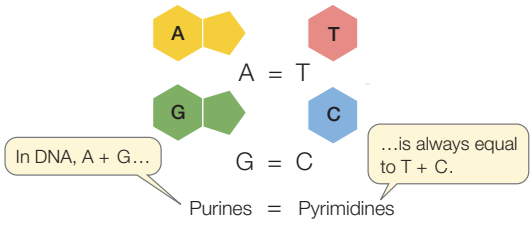


FIGURE 9.5 DNA Is a Double Helix
(A) Francis Crick (right) and James Watson (left) proposed that the DNA molecule has a double-helical structure. (B) Biochemists can now pinpoint the position of every atom in a DNA molecule and have verified the essential features of the Watson–Crick model. Two sugar–phosphate chains provide the outer backbones of the double helix, with hydrogen-bonded nitrogenous bases forming "rungs" within the structure.

Go to MEDIA CLIP 9.1
Discovery of the Double Helix
PoL2e.com/mc9.1

to its structure. Biochemists knew that DNA is a polymer of nucleotides. Each of these nucleotides consists of a molecule of the sugar deoxyribose, a phosphate group, and a nitrogen-containing base (see Figure 3.1). The only differences among the four nucleotides of DNA are their nitrogenous bases: the purines adenine (A) and guanine (G), and the pyrimidines cytosine (C) and thymine (T) (see Figure 3.1).

In 1950, biochemist Erwin Chargaff at Columbia University reported an important observation. He and his colleagues had found that DNA samples from many different species—and from different sources within a single organism—exhibited certain regularities. The following rule held for each sample: the amount of adenine equaled the amount of thymine (A = T), and the amount of guanine equaled the amount of cytosine (G = C). As a result, the total abundance of purines (A + G) equaled the total abundance of pyrimidines (T + C):

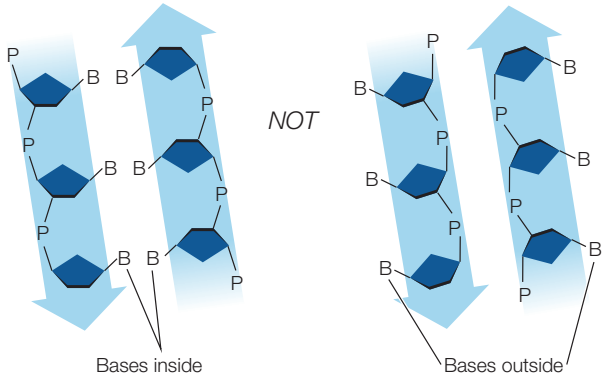


The structure of DNA could not have been worked out without this observation, now known as Chargaff's rule.

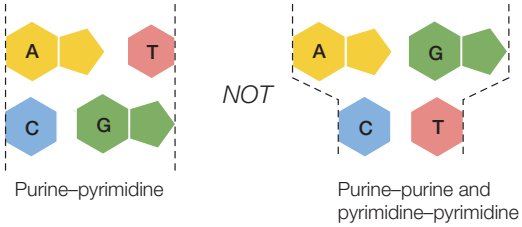
WATSON AND CRICK DESCRIBED THE DOUBLE HELIX Chemical model building is the assembly of three-dimensional structures using known relative molecular dimensions and known bond angles. The English physicist Francis Crick and the American geneticist James D. Watson (FIGURE 9.5A), both then at the Cavendish Laboratory of Cambridge University, used model building to solve the structure of DNA. Rosalind Franklin's crystallography results convinced them that the DNA molecule must be **helical**—it must have a spiral shape like a spring. Density measurements and previous model building experiments suggested that there are two polynucleotide chains in the molecule. Modeling studies also showed that the strands run in opposite directions, that is, they are **antiparallel**. The two strands would not fit together in the model if they were parallel.

How are nucleotides oriented in DNA chains? Watson and Crick suggested that:

- the nucleotide bases are on the interior of the two strands, with a sugar–phosphate backbone on the outside. The strands would not fit together otherwise:



- to satisfy Chargaff's rule (purines = pyrimidines), a purine on one strand is always paired with a pyrimidine on the opposite strand. These **base pairs** (A-T and G-C) have the same width down the double helix, a uniformity shown by X-ray diffraction:



In late February of 1953, Crick and Watson built a model out of tin that established the general structure of DNA. This structure explained all the known chemical properties of DNA, and it opened the door to understanding its biological functions. There have been minor amendments to that first published structure, but its principal features remain unchanged.

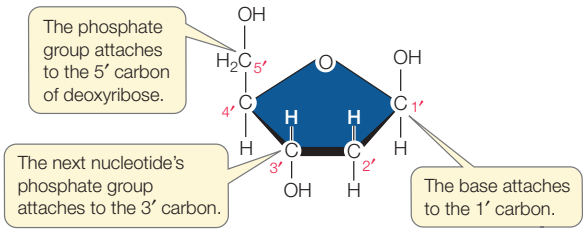
Four key features define DNA structure

Four features summarize the molecular architecture of the DNA molecule (**FIGURE 9.5B**; also review Figure 3.4):

- *DNA is a double-stranded helix of uniform diameter*. The sugar–phosphate backbones of the two chains form a coil on the outside of the helix, and the nitrogenous bases point toward the center. The chains are held together by two chemical forces: hydrogen bonding between the bases

and van der Waals forces between adjacent bases on the same strand. When the base rings come near one another, they tend to stack like poker chips because of these weak attractions.

- *The two DNA strands are antiparallel*. The backbone of each strand contains repeating units of the monosaccharide (sugar) deoxyribose:



The number followed by the prime sign (′) designates the position of a carbon atom in the sugar. In the sugar–phosphate backbone of DNA, the phosphate groups are connected to the 5′ carbon of one deoxyribose molecule and to the 3′ carbon of the next, linking the two sugar molecules together. Thus the two ends of the polynucleotide chain are different. The **5′ end** of the chain is a free (not connected to another nucleotide) phosphate group. The **3′ end** of the chain is a free 3′ hydroxyl (OH) group. In the double helix of DNA, the 5′ end of one strand is paired with the 3′ end of the other strand, and vice versa (see Figure 3.4).

- *In DNA, the outer edges of the nitrogenous bases are exposed in the major and minor grooves*. These grooves exist because the helices formed by the backbones of the two DNA strands are not evenly spaced relative to one another (see Figure 9.5B). **FIGURE 9.6** shows the four possible configurations of the flat, hydrogen-bonded base pairs within the major and minor grooves. The exposed outer edges of the base pairs are accessible for additional hydrogen bonding. Notice that the arrangements of unpaired atoms and groups differ in the A-T base pairs compared with the G-C base pairs. Thus the *surfaces of the A-T and G-C base pairs are chemically distinct*, allowing other molecules, such as proteins, to recognize specific base-pair sequences and bind to them. The atoms and groups in the major groove are more accessible, and tend to bind other molecules more frequently, than those in the minor groove. *This binding of proteins to specific base-pair sequences is the key to protein–DNA interactions*, which are necessary for the replication and expression of the genetic information in DNA.
- *The DNA double helix is right-handed*. Hold your right hand with the thumb pointing up (see Figure 9.10B). Imagine the curve of the helix following the direction of your fingers as it winds upward, and you have the idea.

While DNA usually forms a right-handed helix, it can sometimes be found as a much less stable left-handed helix. So-called Z-DNA ("zig-zag DNA") does not have major and minor grooves and is more elongated and less compact than normal DNA. Z-DNA appears to form in regions of DNA that are being actively transcribed, and it may play a role in stabilizing the DNA during transcription.

The double-helical structure of DNA is essential to its function

The genetic material performs four important functions, and the DNA structure proposed by Watson and Crick was elegantly suited to three of them.

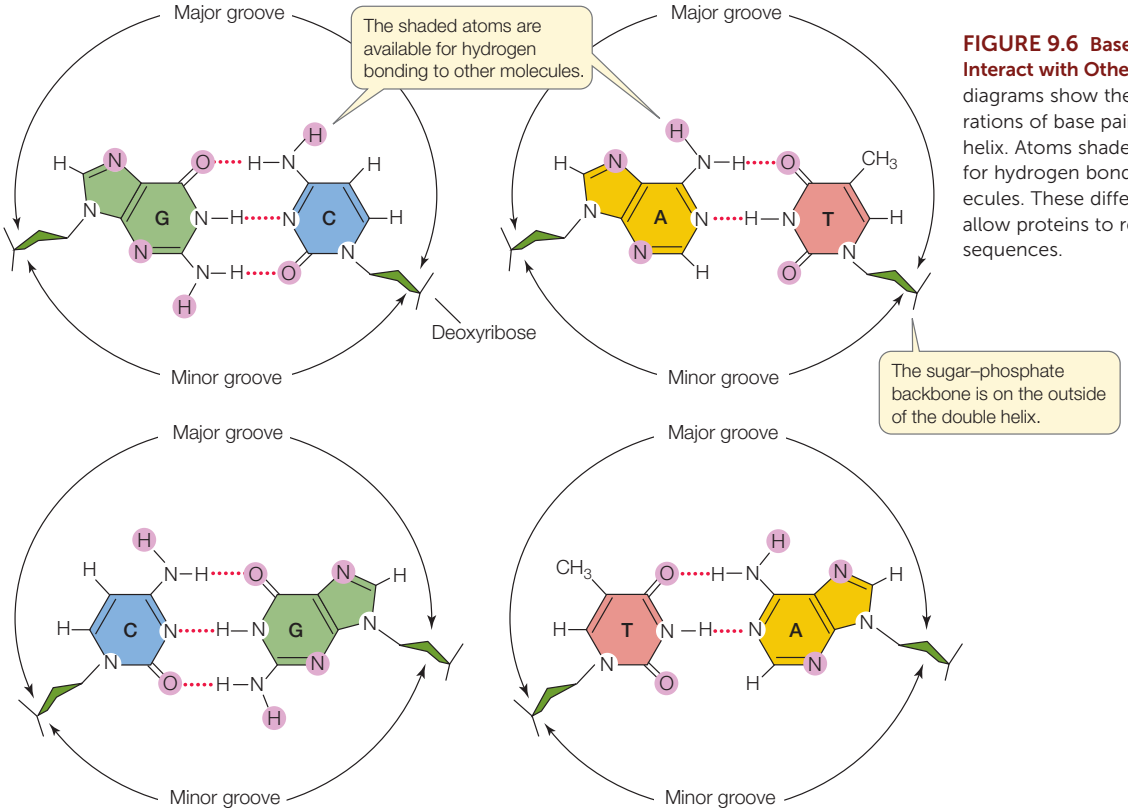


FIGURE 9.6 Base Pairs in DNA Can Interact with Other Molecules These diagrams show the four possible configurations of base pairs within the double helix. Atoms shaded in purple are available for hydrogen bonding with other molecules. These differences in configuration allow proteins to recognize specific DNA sequences.

- *Storage of genetic information*. With its millions of nucleotides, the base sequence of a DNA molecule can encode and store an enormous amount of information. Variations in DNA sequences can account for differences among species and individuals.
- *Precise replication during the cell division cycle*. Replication could be accomplished by complementary base pairing, A with T and G with C. In the original publication of their findings in 1953, Watson and Crick coyly pointed out, "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." We will turn to DNA replication in Concept 9.2.
- *Susceptibility to mutations*. The structure of DNA suggested an obvious mechanism for mutations—stable changes in the genetic material. Mutations might be simple changes in the linear sequence of base pairs. We will discuss mutations in Concept 9.3.
- *Expression of the coded information as phenotypes*. The way this function is accomplished is not obvious in the structure of DNA. However, as we described briefly in Chapter 3 (see Figure 3.5), the nucleotide sequence of DNA can be copied—transcribed—into RNA. The linear sequence of nucleotides that is RNA is then translated into a linear sequence of amino acids that can fold into a protein (see Figure 3.7). Thus the information carried in DNA is ultimately expressed as proteins. The many and varied enzymes and structural proteins encoded by an organism's DNA determine most of the phenotypes of that organism.

LINK

The transcription of DNA into RNA and the translation of RNA into protein are described in detail in **Concepts 10.2** and **10.4**

CHECKpoint CONCEPT 9.1

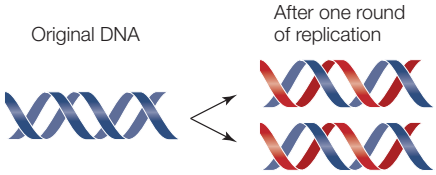
- ✓ Compare circumstantial evidence with experimental evidence in science, using DNA as an example.
- ✓ If a DNA molecule is 20 percent G, what are the percentages of the other three bases?
- ✓ A single-stranded sequence of DNA contains 20 A's, 25 G's, 35 C's, and 24 T's. What is the total base composition of the double-stranded molecule?
- ✓ What circumstantial and experimental evidence would show that protein and not DNA is the genetic material? How would proteins differ from nucleic acids as informational molecules?

We have seen that a DNA molecule consists of long polymers of nucleotides. An individual DNA strand contains thousands or millions (up to about 1 billion) nucleotides in a precise

sequence. How is this huge amount of genetic information replicated before cell division?

CONCEPT 9.2 DNA Replicates Semiconservatively

An important requirement for the genetic material is that it replicates *both completely and accurately* during the cell cycle. The double-helix model of DNA suggested to Watson and Crick how this might be accomplished. **Semiconservative replication** means that each strand of the parental DNA acts as a **template** for a new strand, which is added by base pairing:



There is abundant evidence supporting this mechanism. In a typical experiment, the parental DNA (represented by the blue strands above) is labeled in some fashion (for example, with a radioactive isotope) and then allowed to replicate in cells for a generation. As the new DNA strands (red in the diagram) are made, they are unlabeled. A *conservative* mode of replication would show the parental DNA intact with both strands labeled, and the new DNA with both strands unlabeled. This does not occur. Instead, the resulting DNA molecules are always "hybrids" (one labeled strand and one unlabeled strand), supporting the semiconservative model of replication.

Go to ANIMATED TUTORIAL 9.1
The Hershey-Chase Experiment
PoL2e.com/at9.1

Go to ANIMATED TUTORIAL 9.2
Experimental Evidence for Semiconservative DNA Replication
PoL2e.com/at9.2

DNA replication (**FIGURE 9.7**) involves several different enzymes and other proteins. It takes place in two general steps:

1. The DNA double helix is unwound to separate the two template strands and make them available for new base pairing.
2. As newly added nucleotides form complementary (A with T and G with C) base pairs with template DNA, they are covalently linked together by phosphodiester bonds (see Fig-

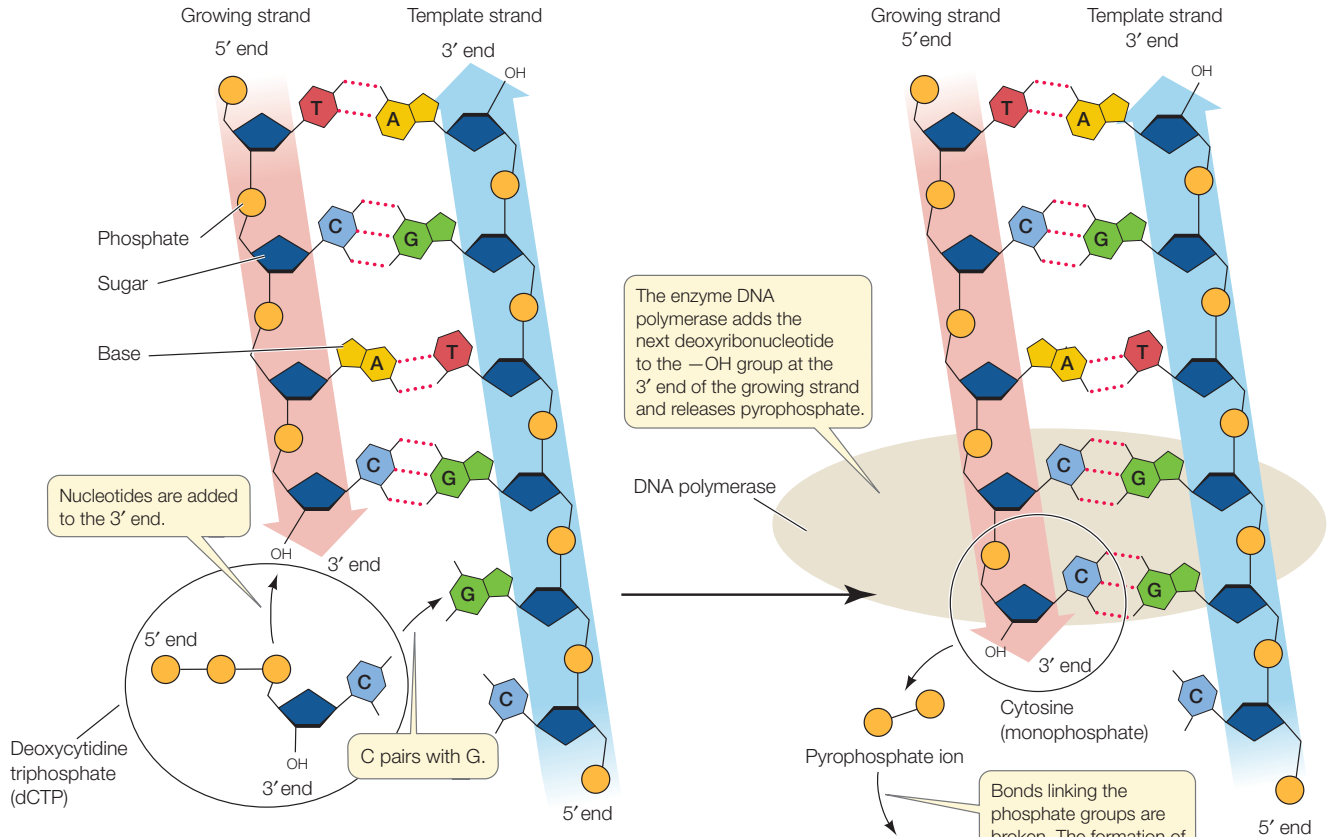


FIGURE 9.7 Each New DNA Strand Grows by the Addition of Nucleotides to Its 3′ End
The DNA strand shown in blue is the template for the synthesis of a complementary strand growing at the left (pink). Here the deoxyribonucleoside triphosphate dCTP (circled) is being added.

APPLY THE CONCEPT

DNA replicates semiconservatively

Each eukaryotic chromosome is composed of a double-stranded DNA molecule. It is possible to use a microscope to distinguish between a chromosome that has been labeled with radioactivity and one that is unlabeled. Plant cells were grown in the lab in the presence of radioactive thymidine for many generations. The cells were then grown for three cell doublings in nonradioactive thymidine so that any new DNA strands would be unlabeled. These cultured cells had synchronized cell cycles. They were examined when they were in anaphase of mitosis.[a]

DOUBLINGS	LABELED CHROMOSOMES (%)
0	100
1	100
2	50
3	25

1. Use diagrams of double-stranded DNA to explain these data.
2. What would the data and diagrams look like if DNA replicated conservatively?

[a] Based on J. H. Taylor et al. 1957. *Proceedings of the National Academy of Sciences* 43: 122–127.

ure 3.2), forming a polymer whose base sequence is complementary to the bases in the template strand. Replication enzymes read the template DNA in the 3′-to-5′ direction.

During DNA synthesis, *nucleotides are added to the 3′ end of the growing new strand*—the end at which the DNA strand has a free hydroxyl (—OH) group on the 3′ carbon of its terminal deoxyribose. In summary, the DNA template is read 3′ to 5′, while the new strand of DNA is generated 5′ to 3′, forming an *antiparallel* double helix.

As we noted in Concept 3.1, a free nucleotide can have one, two, or three phosphate groups attached to its pentose sugar. The raw materials for DNA synthesis are the nucleotides deoxyadenosine triphosphate (dATP), deoxythymidine triphosphate (dTTP), deoxycytidine triphosphate (dCTP), and deoxyguanosine triphosphate (dGTP)—collectively referred to as **deoxyribonucleoside triphosphates** (**dNTPs**) or deoxyribonucleotides. As their names imply, deoxyribonucleoside triphosphates each carry three phosphate groups. During DNA synthesis, the two outer phosphate groups are released in an exergonic reaction, so that the final nucleotide is a monophosphate (adenine, thymine, cytosine, or guanine; see Figure 9.7). The release of the two outer phosphate groups provides energy for the formation of a phosphodiester bond between the single remaining phosphate group of the incoming nucleotide and the 3′ carbon on the sugar at the end of the DNA chain (see Figure 3.2).

Go to ANIMATED TUTORIAL 9.3
DNA Replication Part I: Replication of a Chromosome and DNA Polymerization
PoL2e.com/at9.3

DNA polymerases add nucleotides to the growing chain

DNA replication begins with the binding of a large protein complex (the pre-replication complex) to a specific site on the DNA molecule. This complex contains several different proteins, among them the enzyme **DNA polymerase**, which catalyzes the addition of nucleotides as the new DNA chain grows. All chromosomes have at least one region called the **origin of replication** (*ori*), to which the pre-replication complex binds. Binding occurs when proteins in the complex recognize specific DNA sequences within the *ori*.

ORIGINS OF REPLICATION The single circular chromosome of the bacterium *Escherichia coli* has 4×10^6 base pairs (bp) of DNA. The chromosome has a single 245 bp *ori* sequence. Once the pre-replication complex binds to it, the DNA unwinds and replication proceeds in both directions around the circle, forming two **replication forks** (FIGURE 9.8A). The opening of each fork is catalyzed by an enzyme called **DNA helicase**, which uses free energy from ATP hydrolysis to change shape and wedge into the DNA, locally breaking hydrogen bonds between bases on the two strands and separating them.

The replication rate in *E. coli* is approximately 1,000 bp per second, so it takes about 40 minutes to fully replicate the chromosome (with two replication forks). Rapidly dividing *E. coli* cells divide every 20 minutes. In these cells, new rounds of replication begin at the *ori* of each new chromosome before the first chromosome has fully replicated. In this way the cells can divide in less time than the time needed to finish replicating the original chromosome.

Eukaryotic chromosomes are much longer than those of prokaryotes—up to a billion bp—and are linear, not circular. If replication occurred from a single *ori*, it would take weeks to fully replicate a chromosome. So eukaryotic chromosomes have multiple origins of replication, scattered at intervals of 10,000–40,000 bp (FIGURE 9.8B).

DNA REPLICATION IS INITIATED WITH A PRIMER DNA polymerase elongates a polynucleotide strand by covalently linking new nucleotides to a preexisting strand. However, it cannot begin this process without a short "starter" strand, called a **primer**. In most organisms this primer is a short single strand of RNA (FIGURE 9.9), but in some viruses it is DNA. The primer is complementary to the DNA template and is synthesized one nucleotide at a time by an enzyme called a **primase**. The DNA polymerase then adds nucleotides to the 3′ end of the primer and continues until the replication of that section of DNA has been completed. Then the RNA

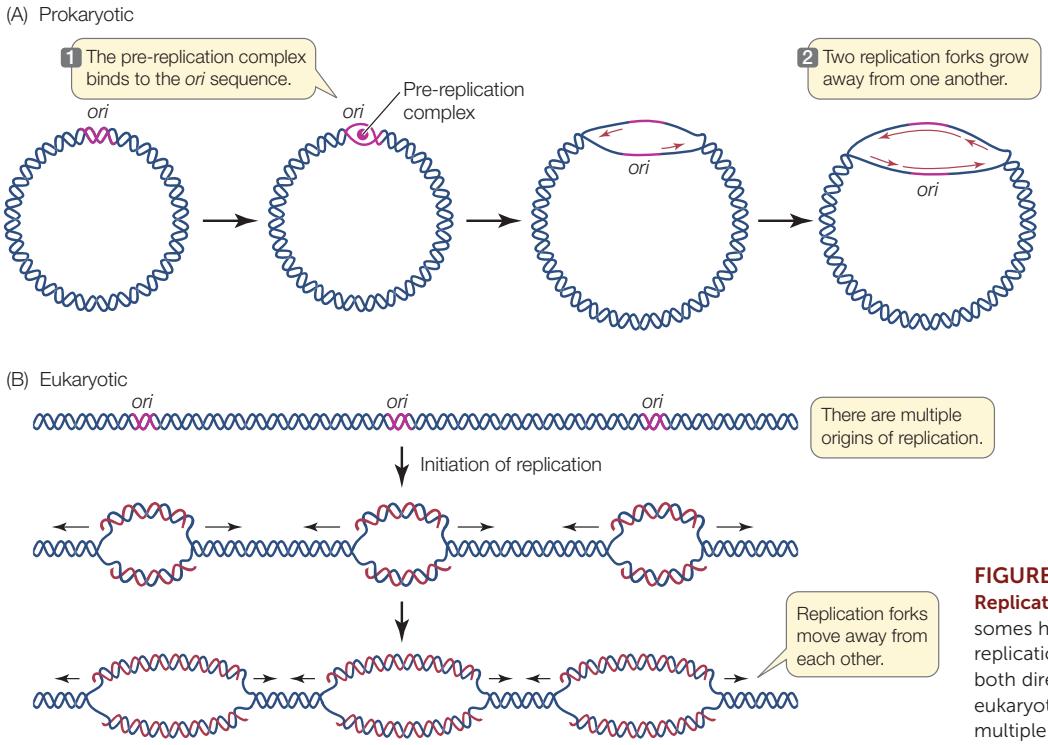


FIGURE 9.8 The Origin of DNA Replication (A) Prokaryotic chromosomes have a single origin where DNA replication starts and then proceeds in both directions. (B) In the much larger eukaryotic chromosomes there are multiple origins of replication.

primer is degraded, DNA is added in its place, and the resulting DNA fragments are connected by the action of another enzyme. When DNA replication is complete, each new strand consists only of DNA.

DNA POLYMERASES ARE LARGE DNA polymerases are much larger than their substrates, the dNTPs, and the template DNA, which is very thin (**FIGURE 9.10A**). Molecular models of the enzyme–substrate–template complex show that the enzyme is shaped like a right hand with a palm, a thumb, and fingers (**FIGURE 9.10B**). Within the "palm" is the active site of the enzyme, which binds the template DNA strand and the new, growing DNA strand. The "fingers" region has precise pockets that can only fit the specific shapes of correct A-T and G-C base pairs. When an incoming nucleotide correctly pairs with a nucleotide on the template strand at the active site of the enzyme, the base pair is recognized by the fingers region. When this occurs, the enzyme undergoes a conformational change (a change in shape), and then catalyzes the condensation reaction that results in the formation of a new phosphodiester bond (see Figures 3.2 and 9.7).

Most cells contain more than one kind of DNA polymerase, but only one of them is responsible for chromosomal DNA replication. The others are involved in primer removal and DNA repair. Fifteen DNA polymerases have been identified in humans; the bacterium *E. coli* has five DNA polymerases.

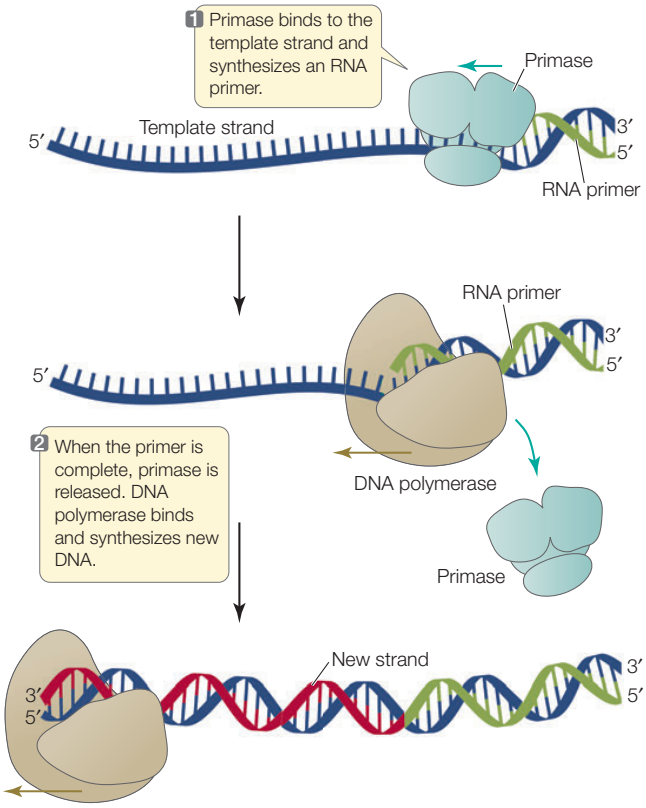


FIGURE 9.9 DNA Forms with a Primer The DNA polymerase requires a primer: a "starter" strand of RNA (or DNA) that provides a 3′—OH end to which new nucleotides can be added.

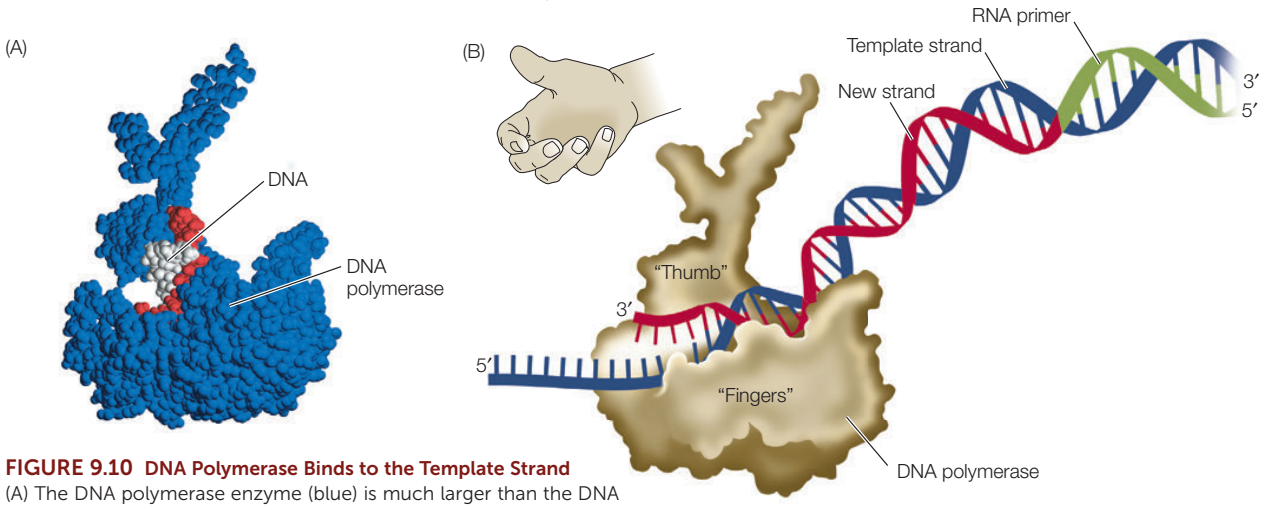


FIGURE 9.10 DNA Polymerase Binds to the Template Strand
(A) The DNA polymerase enzyme (blue) is much larger than the DNA molecule (red and white). (B) DNA polymerase is shaped like a hand, and in this side-on view, its "fingers" can be seen curling around the DNA. These fingers can recognize the distinctive shapes of the four bases.

Go to ACTIVITY 9.1 DNA Polymerase
PoL2e.com/ac9.1

THE TWO DNA STRANDS GROW DIFFERENTLY AT THE REPLICATION FORK A single replication fork opens up in one direction. Study **FIGURE 9.11** and try to imagine what is happening over a short period of time. For the purpose of understanding the process, imagine that the DNA opens from one end like a zipper (even though, as we have seen, it actually opens up from within the molecule and replication extends in both directions). Remember two things:

- The two DNA strands are *antiparallel*—that is, the 3′ end of one strand is paired with the 5′ end of the other.
- DNA polymerase replicates DNA by adding nucleotides only to the 3′end of each *growing strand*.

One newly synthesized growing strand—the **leading strand**—is oriented so that it can grow continuously at its 3′ end as the

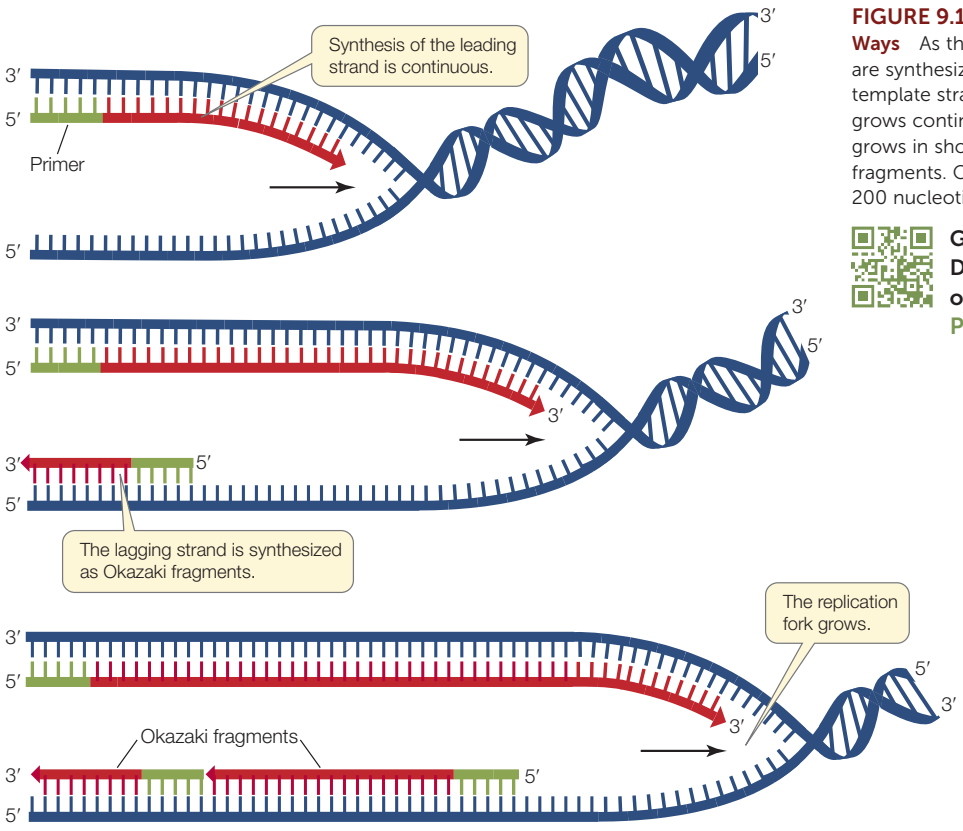


FIGURE 9.11 The Two New Strands Form in Different Ways As the parent DNA unwinds, both new strands are synthesized in the 5′-to-3′ direction, although their template strands are antiparallel. The leading strand grows continuously forward, but the lagging strand grows in short, discontinuous stretches called Okazaki fragments. Okazaki fragments in eukaryotes are 100 to 200 nucleotides long.

Go to ANIMATED TUTORIAL 9.4
DNA Replication, Part 2: Coordination of Leading and Lagging Strand Synthesis
PoL2e.com/at9.4

fork opens up. The other new strand—the **lagging strand**—must be synthesized differently because it grows in the direction away from the replication fork.

Synthesis of the lagging strand requires the synthesis of relatively short, discontinuous stretches of sequence (100–200 nucleotides in eukaryotes; 1,000–2,000 nucleotides in prokaryotes). These discontinuous stretches are synthesized just as the leading strand is, by the addition of new nucleotides one at a time to the 3′ end, but the new strand grows away from the replication fork. These stretches of new DNA are called **Okazaki fragments** after their discoverer, the Japanese biochemist Reiji Okazaki. To summarize, while the leading strand grows continuously "forward," the lagging strand grows in shorter, "backward" stretches with gaps between them.

A single primer is needed to initiate synthesis of the leading strand, but each Okazaki fragment requires its own primer to be synthesized by the primase. DNA polymerase then synthesizes an Okazaki fragment by adding nucleotides to one primer until it reaches the primer of the previous fragment. At this point, a different DNA polymerase removes the old primer and replaces it with DNA. Left behind is a tiny nick—the final phosphodiester linkage between the adjacent Okazaki fragments is missing. The enzyme **DNA ligase** catalyzes the formation of that bond, linking the fragments and making the lagging strand whole (**FIGURE 9.12**).

DNA replication may appear complex (we have simplified it considerably), but it occurs with astonishing speed and accuracy. As we mentioned earlier, the rate of replication in *E. coli* is about 1,000 base pairs per second, yet the polymerase commits very few errors—less than 1 base in a million. How do DNA polymerases work so fast? We saw in Concept 3.3 that an enzyme catalyzes a chemical reaction through a series of events:

Substrate binds to enzyme → one product is formed → enzyme is released → cycle repeats

DNA replication would not proceed as rapidly as it does if DNA polymerase went through such a cycle for each nucleotide. Instead, DNA polymerase is **processive**—that is, it *catalyzes many sequential polymerization reactions each time it binds to a DNA molecule*:

Substrates bind to enzyme → many products are formed → enzyme is released → cycle repeats

Typically, a DNA polymerase can add thousands of nucleotides before it detaches from DNA.

Telomeres are not fully replicated in most eukaryotic cells

As we have just seen, replication of the lagging strand occurs by the addition of Okazaki fragments to RNA primers. When the terminal RNA primer is removed from the replicating end of a linear eukaryotic chromosome, no DNA can be synthesized to replace it because there is no 3′ end to extend. So the new chromosome has a bit of single-stranded DNA at each end. This situation activates a mechanism for cutting off the single-stranded region, along with some of the intact double-stranded

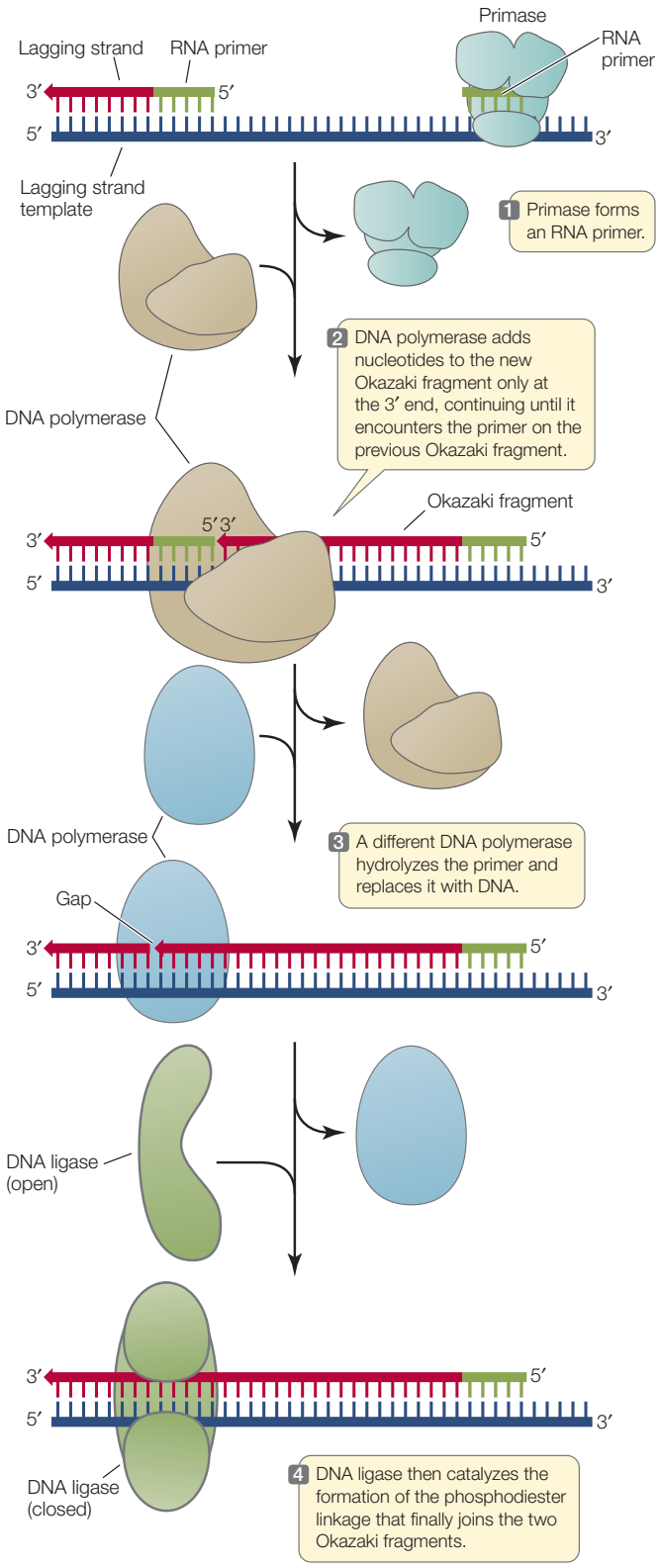


FIGURE 9.12 The Lagging Strand Story In bacteria, DNA polymerases and DNA ligase cooperate to complete the complex task of synthesizing the lagging strand.

DNA. Thus the chromosome becomes slightly shorter with each cell division (FIGURE 9.13A).

Another problem with chromosome ends is that they must be protected from being joined to other chromosomes by the DNA repair system. When DNA is damaged by external or internal agents (e.g., radiation), it is repaired by a combination of DNA polymerase and DNA ligase activities. This system might mistakenly recognize chromosome ends as breaks and join two chromosomes together. This would create havoc with genomic integrity.

To prevent chromosomes from joining, many eukaryotes have strings of repetitive sequences at the ends of their chromosomes called **telomeres** (FIGURE 9.13B). In humans and other vertebrates, the repeated sequence is TTAGGG, and in humans it is repeated about 2,500 times. These repeats bind a protein complex, appropriately named the shelterin complex, which protects the ends from being joined together by the DNA repair system. In addition, the repeats form loops that have a similar protective role. So telomeres act like the plastic tips of shoelaces to prevent fraying (FIGURE 9.13C).

Each human chromosome can lose 50–200 bp of telomeric DNA after each round of DNA replication and cell division. After 20–30 cell divisions, the chromosome ends become short enough to lose their protective role, and the chromosomes lose their integrity. Apoptosis (programmed cell death) ensues, and the cell dies. This phenomenon explains, in part, why many cell lineages do not last the entire lifetime of the organism: their telomeres are lost. Yet continuously dividing cells, such as bone marrow stem cells and gamete-producing cells, maintain their telomeric DNA. An enzyme called **telomerase** catalyzes the addition of any lost telomeric sequences in these cells. Telomerase contains an RNA sequence that acts as a template for the telomeric DNA repeat sequence.

There is a relationship between telomere length and aging: the average telomere length is shorter in older individuals. Furthermore, when a gene expressing high levels of telomerase is added to human cells in culture, their telomeres do not shorten. Instead of living 20–30 cell generations and then dying, the cells become immortal. This phenomenon is also seen in mice that overexpress (make more than the usual amount of) telomerase—they live longer. Moreover, cancer cells also overexpress telomerase and can stay alive for longer periods than normal cells. Not surprisingly, drugs that affect telomerase activity are being developed, with the hope of both treating cancer and preventing aging.

Errors in DNA replication can be repaired

We have stressed that DNA must be accurately replicated; the accurate transmission of genetic information is essential for the proper functioning and even the life of a single cell or multicellular organism. Yet the replication of DNA is not perfectly accurate. DNA polymerases sometimes insert a base that is not complementary to the template (for example, putting an A in the new DNA strand opposite a C in the template strand). In

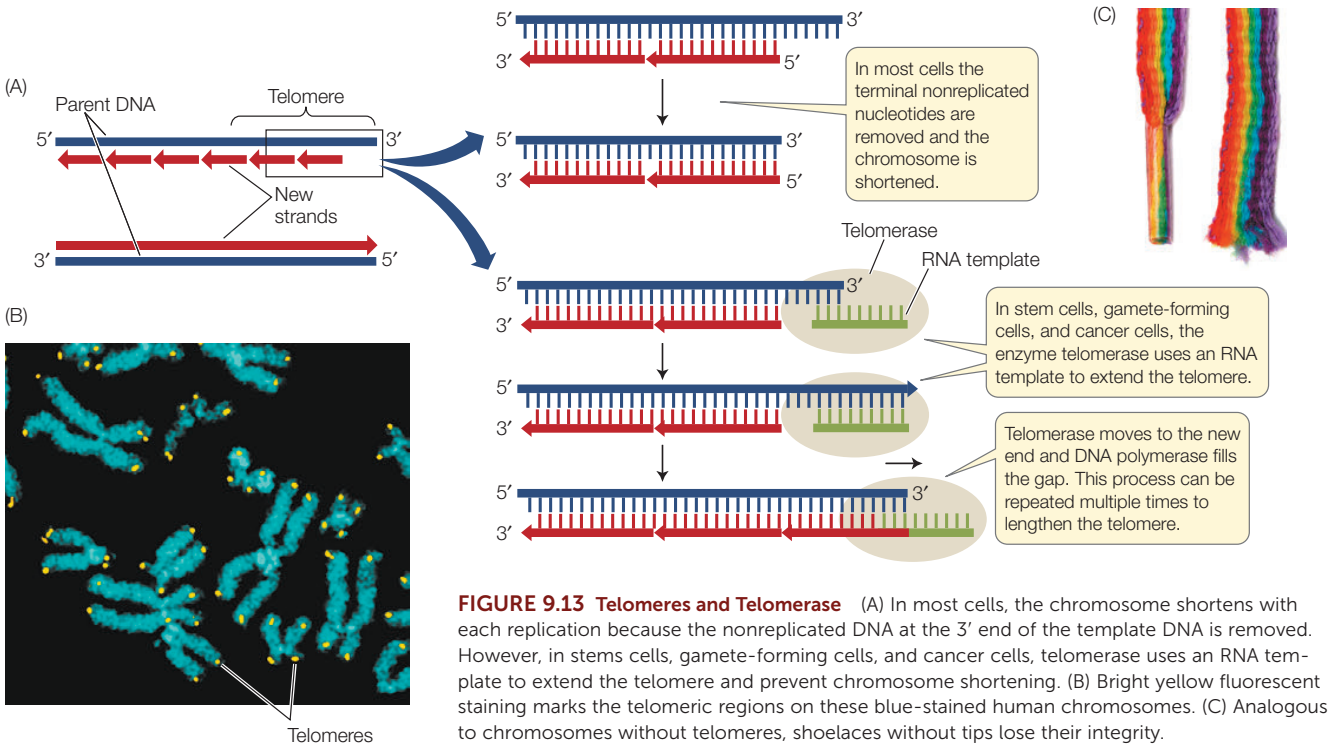


FIGURE 9.13 Telomeres and Telomerase (A) In most cells, the chromosome shortens with each replication because the nonreplicated DNA at the 3′ end of the template DNA is removed. However, in stems cells, gamete-forming cells, and cancer cells, telomerase uses an RNA template to extend the telomere and prevent chromosome shortening. (B) Bright yellow fluorescent staining marks the telomeric regions on these blue-stained human chromosomes. (C) Analogous to chromosomes without telomeres, shoelaces without tips lose their integrity.

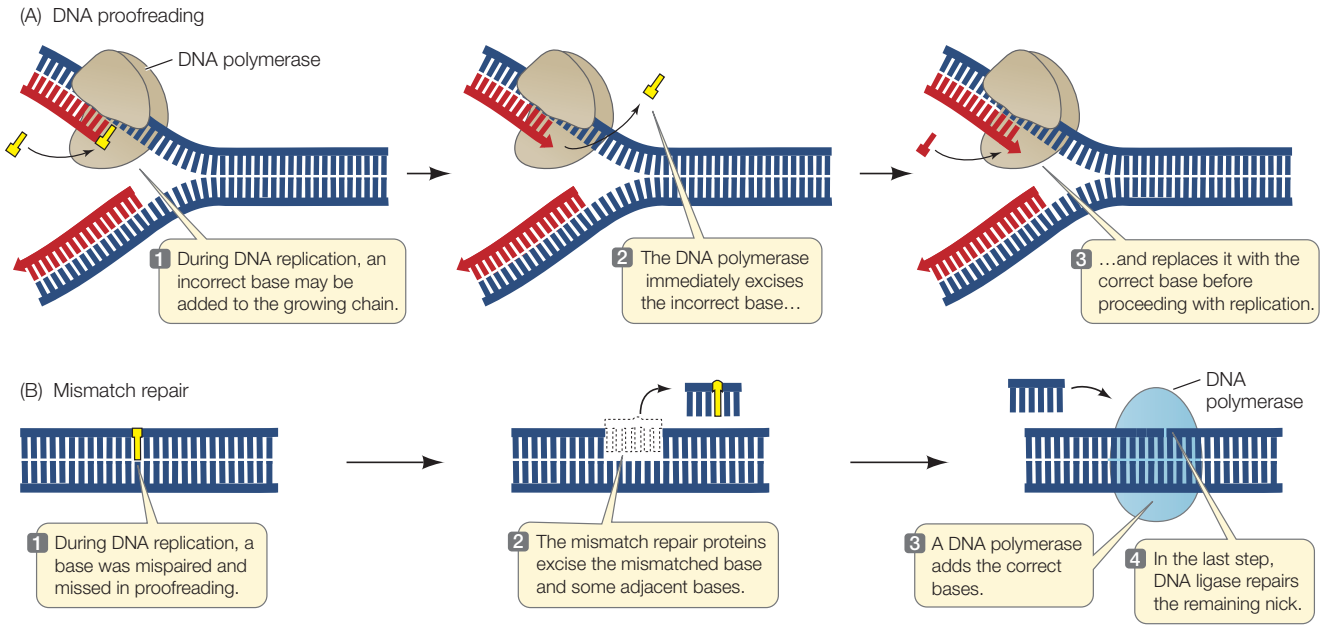


FIGURE 9.14 DNA Repair Mechanisms (A) During replication, the DNA polymerase checks for incorrect bases in the new DNA strand and immediately replaces them with correct ones. This process is called proofreading. (B) After replication, mismatch repair proteins search for incorrect bases that were missed by DNA polymerase and replaces them.

eukaryotes, the error rate is about 1 incorrect base in 100,000 (a 10^{-5} error rate). With a genome size of 3×10^9 bp, this would produce 60,000 errors after every cell division. This is an intolerable mutation rate for survival in the long term. However, if eukaryotic DNA sequences are studied before and after a cell cycle, the actual frequency of DNA errors is 10^{-10} per cell cycle. This means that most of the errors are repaired. There are two major repair mechanisms:

- *Proofreading* occurs right after DNA polymerase inserts a nucleotide (**FIGURE 9.14A**). When a DNA polymerase recognizes a mispairing of bases, it removes the improperly introduced nucleotide and tries again.
- *Mismatch repair* occurs after DNA has been replicated (**FIGURE 9.14B**). A second set of proteins surveys the newly replicated molecule and looks for mismatched base pairs that were missed in proofreading. A portion of the DNA including the incorrect nucleotide is removed, and then a DNA polymerase inserts the correct sequence.

The basic mechanisms of DNA replication can be used to amplify DNA in a test tube

The principles underlying DNA replication in cells have been used to develop a laboratory technique that has been vital in analyzing DNA, genes, and genomes. The **polymerase chain reaction**, or **PCR**, allows researchers to make multiple copies of short DNA sequences in a test tube—a process referred to as DNA amplification. The PCR technique uses:

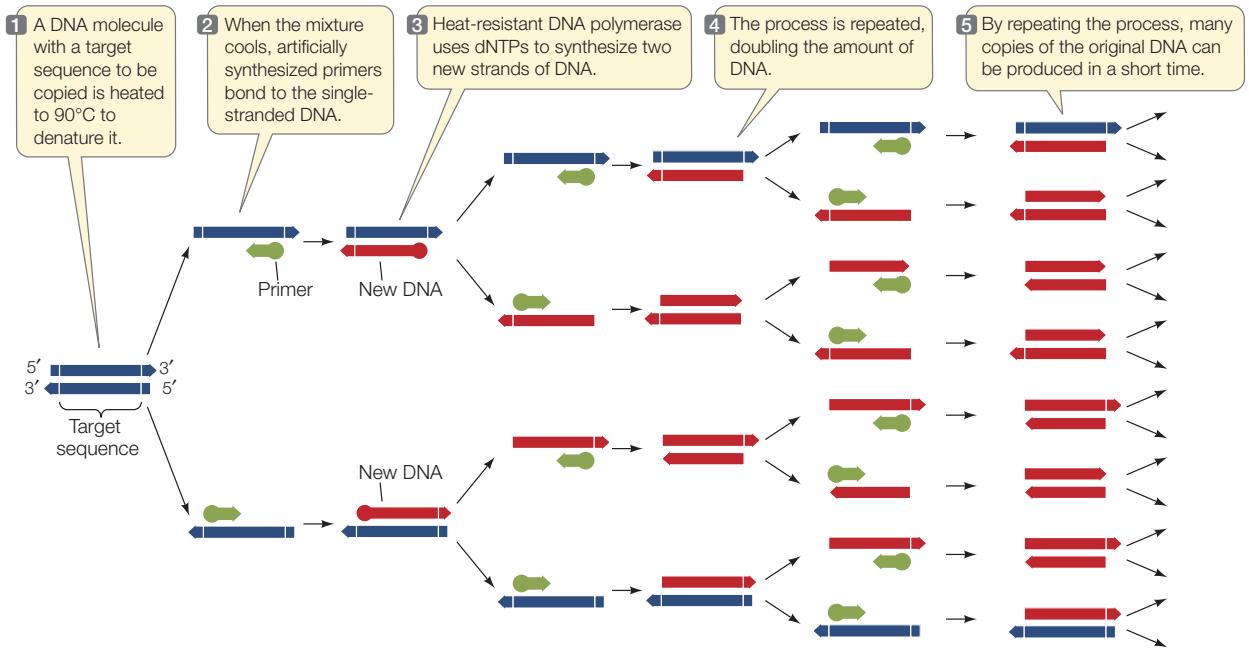
- a sample of double-stranded DNA to act as the template
- two short, artificially synthesized primers that are complementary to the ends of the sequence to be amplified
- the four dNTPs (dATP, dTTP, dCTP, and dGTP)
- a DNA polymerase that can tolerate high temperatures without becoming denatured
- salts and a buffer to maintain a near-neutral pH.

PCR is a cyclic process in which a sequence of steps is repeated over and over again (**FIGURE 9.15**). Since DNA replication is fast even in a test tube, it takes only a short time to go from two to four to millions of short segments of DNA. The PCR technique requires that the base sequences at each end of the amplified fragment be known ahead of time, so that complementary primers, usually 15–30 bases long, can be made in the laboratory. Because of the uniqueness of DNA sequences, a pair of primers this length will usually bind to only a single region of DNA in an organism's genome. This specificity is a key to the power of PCR to amplify just a small part of a larger DNA molecule. Some of the most striking applications of PCR will be described in Chapters 12 and 13. These applications range from the identification of individuals by their DNA to the detection of diseases.

Go to ANIMATED TUTORIAL 9.5
Polymerase Chain Reaction Simulation
PoL2e.com/at9.5

RESEARCH TOOLS

FIGURE 9.15 The Polymerase Chain Reaction The steps in this cyclic process are repeated many times to produce millions of identical copies of a DNA fragment. This makes enough DNA for chemical analysis and genetic manipulations.



CHECKpoint CONCEPT 9.2

- ✓ What is semiconservative DNA replication?
- ✓ Why does the leading strand in DNA replicate continuously and the lagging strand discontinuously?
- ✓ Cells from older people have shorter telomeres than cells from younger people. How might this relate to aging?
- ✓ If you have a small amount of a large chromosome of 20 million bp and want to amplify a short sequence of 1,000 bp, how would you do it? Explain the role of primers in this process.

We have now described (1) the lines of evidence for DNA as the genetic material and (2) the precise replication of DNA during cell division. A less obvious requirement for DNA as the genetic material is its ability to mutate. Mutation creates variability in DNA, which is the raw material for evolution.

CONCEPT 9.3 Mutations Are Heritable Changes in DNA

In Chapter 8 we described mutations as stable and inherited changes in the genetic material. A mutation may result in a new allele of a gene, and different alleles may produce different phenotypes (for example, pea plants with wrinkled seeds versus round seeds). With reference to the chemical nature of genes, we can state that *mutations are changes in the nucleotide sequence of DNA that are passed on from one cell or organism to another.*

Mutations occur by a variety of processes. For example, in Concept 9.2 we described how DNA polymerases can make errors. Repair systems such as proofreading are in place to correct them, but some errors escape being corrected and are passed on to daughter cells.

Mutations in multicellular organisms can be divided into two types:

- **Somatic mutations** occur in the somatic (body) cells of a multicellular organism. These mutations are passed on to the daughter cells during mitosis, and in turn to the offspring of those cells. For example, a mutation in a single skin cell could result in a patch of skin cells that all have the same mutation. However, somatic mutations are not passed on to sexually produced offspring. (Exceptions occur in plants, where germline cells can arise from somatic cells and thus pass on somatic mutations.)
- **Germline mutations** occur in the cells of the germ line—the specialized cells that give rise to gametes (the eggs and sperm of sexual reproduction). A gamete with the mutation passes it on to a new organism at fertilization.

In either case, the mutations may or may not affect the phenotype.

Mutations can have various phenotypic effects

An organism's genome is the total DNA sequence present in all of its chromosomes (or in its single chromosome, in the case of prokaryotes). Depending on the organism, it can consist of millions or billions of base pairs of DNA. Most genomes include both genes and regions of DNA that are not expressed.

As we discussed briefly in Concept 3.1, gene expression involves the transcription of DNA into RNA, followed by the translation of the RNA into a polypeptide. (Some RNAs are not translated but have catalytic or other roles in the cell.) We will discuss gene expression in much more detail in Chapter 10. For now we will look at some of the ways that mutations can affect gene expression and phenotypes. It is also the case that many mutations have no effects on phenotypes.

Mutations are often discussed in terms of their effects on protein-coding genes and their functions (**FIGURE 9.16**):

- **Silent mutations** do not affect gene function (see Figure 9.16B). They can be mutations in DNA that is not expressed, or mutations within an expressed region that do not have any effect on the encoded protein. *Most mutations in large genomes are silent.*

LINK

Silent mutations are a source of neutral alleles that can be acted upon by the mechanisms of evolution, as described in **Concept 15.2**

- **Loss-of-function mutations** can result in either the loss of expression of a gene or in the production of a nonfunctional protein or RNA. Some loss-of-function mutations prevent a gene from being transcribed or cause transcription to terminate too soon. In other cases the gene is transcribed and translated, but the resulting protein no longer works as a structural protein or enzyme (as illustrated in Figure 9.16C). Loss-of-function mutations almost always show recessive inheritance in a diploid organism, because the presence of one wild-type allele usually results in sufficient functional protein for the cell. For example, the wrinkled seed phenotype studied by Mendel (see Figure 8.1) is due to a recessive loss-of-function mutation in the gene for *s*tarch *b*ranching *e*nzyme 1 (*SBE1*). Even in plants with only one copy of the wild-type allele, there is enough SBE1 enzyme to produce the wild-type round phenotype.
- **Gain-of-function mutations** lead to a protein with an altered function (see Figure 9.16D). This kind of mutation usually shows dominant inheritance, because the presence of the wild-type allele does not prevent the mutant allele from functioning. This type of mutation is common in cancer. For example, a receptor for a growth factor normally requires binding of the growth factor (the ligand) to activate the cell division cycle. Some cancers are caused by mutations in genes that encode these receptors so that the receptors are "always on," even in the absence of their ligands. This leads to the unrestrained cell proliferation that is characteristic of cancer cells.

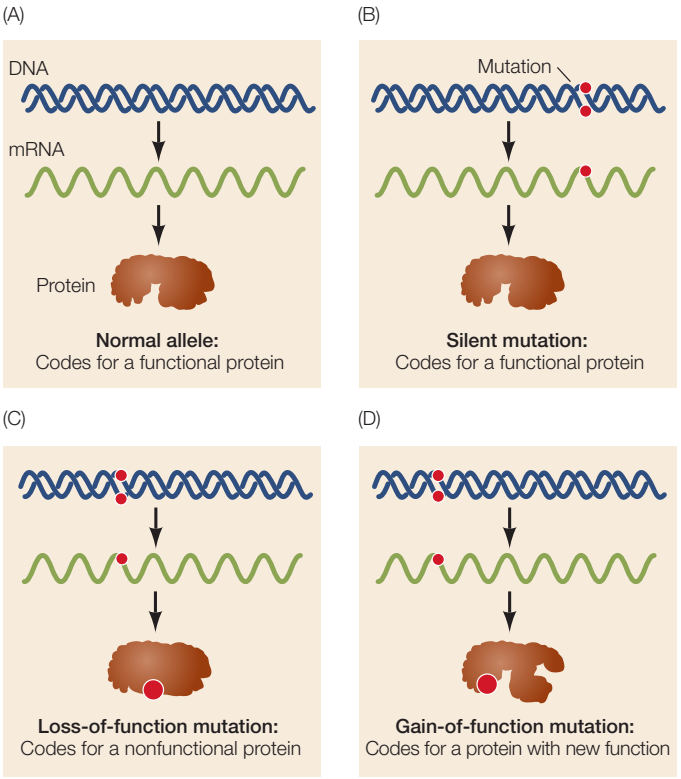


FIGURE 9.16 Mutation and Phenotype Mutations may or may not affect the function and phenotype of a protein.

- **Conditional mutations** cause their phenotypes only under certain *restrictive* conditions. The wild-type phenotype is expressed under other, *permissive* conditions. Many conditional mutants are temperature-sensitive; that is, they show the altered phenotype only at a certain temperature. For example, in Chapter 8 (see p. 161) we described a temperature-sensitive mutation that affects coat color in rabbits and cats. In warmer parts of the body, the mutant protein is inactive, resulting in pale fur. In the cooler extremities of the body, the protein is active, producing dark fur.

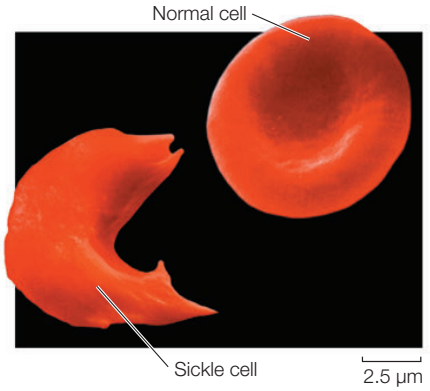
All mutations are alterations in the nucleotide sequence of DNA. They can be small-scale mutations that alter only one or a few nucleotides, or they can be large-scale mutations in which entire segments of DNA are rearranged, duplicated, or irretrievably lost. Next we will consider small-scale mutations, in particular, point mutations.

Point mutations are changes in single nucleotides

A **point mutation** is the addition or subtraction of a single nucleotide base, or the substitution of one base for another. Point mutations can arise because of errors in DNA replication that are not corrected during proofreading, or they may be caused by environmental **mutagens**: substances that cause mutations, such as radiation or certain chemicals.

Some point mutations that occur within genes are loss-of-function mutations because they prevent the gene from being properly transcribed. In other cases the gene is transcribed normally. A point mutation in the coding region of a gene may result in changes in the RNA, but changes in the RNA may or may not result in a change in the amino acid sequence of the protein. If the protein is not changed, the mutation is silent.

Other mutations result in altered amino acid sequences, and in some cases these changes can have drastic phenotypic effects. An example is the mutation that causes sickle-cell disease, a heritable blood disorder. The disease occurs in people who carry two copies of the sickle allele of the gene for human β-globin (a subunit of hemoglobin, the protein in human blood that carries oxygen; see p. 47). The sickle allele differs from the normal allele by one base pair, resulting in a polypeptide that differs by one amino acid from the normal protein. Individuals who are homozygous for this recessive allele have defective, sickle-shaped red blood cells:



Go to MEDIA CLIP 9.2
Sickle Cells: Deformed by a Mutation
PoL2e.com/mc9.2

The deformed cells tend to block narrow capillaries, which results in tissue damage.

Not all changes in the amino acid sequence of a protein affect its function. For example, a hydrophilic amino acid may be substituted for another hydrophilic amino acid, so that the shape of the protein is unchanged. Or a mutation might result in a protein that has reduced efficiency but is not completely inactivated. Individuals homozygous for a point mutation of this type may show no change in phenotype if enough of the protein's function is retained.

LINK

Review **Concept 3.2** to better understand how a mutation affecting amino acid sequence (the primary protein structure) can affect higher levels of protein structure and thus the protein phenotype

In some cases, gain-of-function point mutations occur. An example is a class of mutations in the human gene *TP53*, which encodes the tumor suppressor protein p53. The p53 protein normally functions to inhibit the cell cycle, but certain mutations cause the protein to promote the cell cycle and prevent programmed cell death. So a p53 protein mutated in this way has a gain of oncogenic (cancer-causing) function.

Chromosomal mutations are extensive changes in the genetic material

In addition to point mutations there are other kinds of mutations that affect longer sequences of DNA. The most dramatic changes that can occur in the genetic material are **chromosomal mutations**. Whole chromosomes can break and rejoin, grossly disrupting the sequences of genes. There are four types of chromosomal mutations: deletions, duplications, inversions, and translocations. This kind of severe damage to chromosomes can result from mutagens or from drastic errors in chromosome replication. Like point mutations, chromosome mutations provide new combinations of genes and genetic diversity important to evolution by natural selection.

- **Deletions** result in the removal of part of the genetic material (**FIGURE 9.17A**). Their consequences can be severe or even fatal. It is easy to imagine one mechanism that could

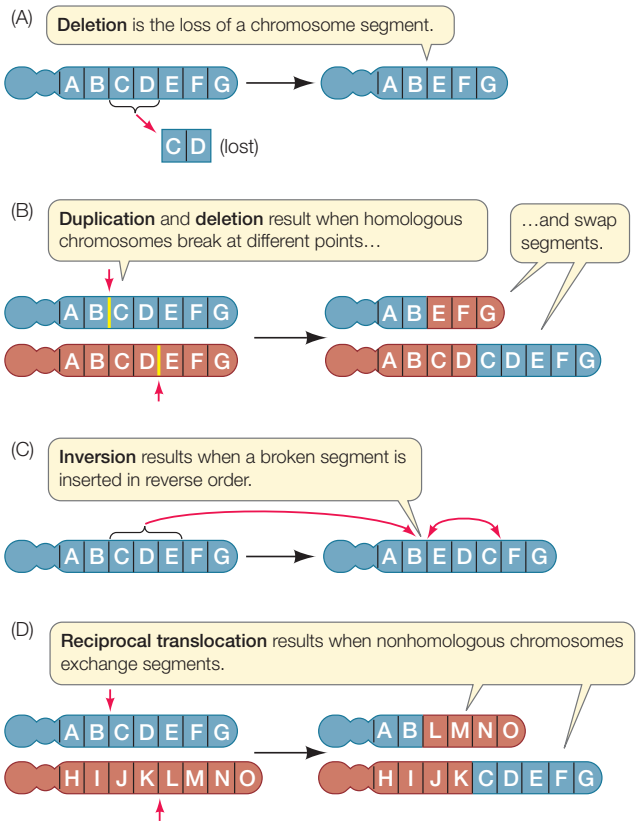


FIGURE 9.17 Chromosomal Mutations Chromosomes may break during replication, and parts of chromosomes may then rejoin incorrectly. The letters on these illustrations represent large chromosomal segments containing anywhere from zero to hundreds or thousands of genes.

APPLY THE CONCEPT

Mutations are heritable changes in DNA

Nitrosamines (R–N–NO_2) are potent mutagens. They can be formed from the reaction of nitrites (R–NO_2) with amino groups in proteins. So there are concerns about using nitrites to preserve meats, which contain amino groups. An experiment was performed to test the effect of vitamin C (ascorbate) on mutagenesis (induction of mutations) caused by meats cured with nitrites. Bacterial cells were incubated with cured meat extracts in the presence or absence of ascorbate. The rates of mutation (number of mutant bacteria per total bacteria) are shown in the table.

MEAT EXTRACT (μg/ml)	AMOUNT OF ASCORBATE (μg/ml)	RATE OF MUTATION ($\times 10^{-5}$)
0	0	2
0	50	2
10	0	5
10	50	2
20	0	14
20	50	2
30	0	35
30	50	2

1. What did the experiment with no extract and no ascorbate show?
2. What did the experiments with increasing amounts of extract and no ascorbate show?
3. What was the effect of ascorbate on the mutation rate?
4. In the bacterium tested, the wild-type DNA had the sequence 5′-ACTTAT-3′, and the mutated strain had the sequence 5′-ATTTAT-3′. What does this tell you about the nature of the mutation? Outline the steps in mutagenesis, noting DNA replication(s). (Hint: see Figure 9.18.)

produce a deletion: a DNA molecule might break at two points and the two end pieces might rejoin, leaving out the DNA between the breaks.

- **Duplications** can be produced at the same time as deletions (**FIGURE 9.17B**). A duplication would arise if homologous chromosomes broke at different positions and then reconnected to the wrong partners. One of the two chromosomes produced by this mechanism would lack a segment of DNA (it would have a deletion), and the other would have two copies (a duplication) of the segment that was deleted from the first chromosome.
- **Inversions** can also result from the breaking and rejoining of chromosomes. A segment of DNA may be removed and reinserted into the same location in the chromosome, but "flipped" end over end so that it runs in the opposite direction (**FIGURE 9.17C**). If either break site occurs within a gene, it is likely to cause a loss-of-function mutation in that gene.
- **Translocations** result when segments of chromosomes break off and become joined to different chromosomes. Translocations may involve reciprocal exchanges of chromosome segments, as in **FIGURE 9.17D**. Translocations often lead to duplications and deletions and may result in sterility if normal chromosome pairing cannot occur during meiosis.

LINK

Mobile DNA elements called transposons are another source of mutation; see **Figure 12.7**

Mutations can be spontaneous or induced

When thinking about the causes of mutations, it is useful to distinguish between mutations that are spontaneous and those that are induced.

Spontaneous mutations are permanent changes in the genetic material that occur without any outside influence. In other words, they occur simply because cellular processes are imperfect. Spontaneous mutations may occur by several mechanisms:

- *DNA polymerase can make errors in replication*. Most of these errors are repaired by the proofreading function of the replication complex, but some errors escape detection and become permanent.
- *The four nucleotide bases of DNA have alternate structures that affect base pairing*. Each nucleotide can exist in two different forms (called tautomers), one of which is common and one rare. When a base temporarily forms its rare tautomer, it can pair with the wrong base (**FIGURE 9.18A,C**).
- *Bases in DNA may change because of spontaneous chemical reactions*. One such reaction is the deamination (conversion of an amino group to a keto group) in cytosine to form the base uracil, which pairs with A rather than G. Usually these errors are repaired, but since the repair mechanism is not perfect, the altered nucleotide will sometimes remain and cause a permanent base change after replication.
- *Meiosis is not perfect*. Sometimes errors occur during the complex process of meiosis. This can result in nondisjunction and aneuploidy (see Concept 7.4) or chromosomal breakage and rejoining (discussed above).
- *Gene sequences can be disrupted*. Random chromosome breakage and rejoining can produce deletions, duplications, inversions, or translocations.

Induced mutations occur when some agent from outside the cell—a mutagen—causes a permanent change in the DNA sequence:

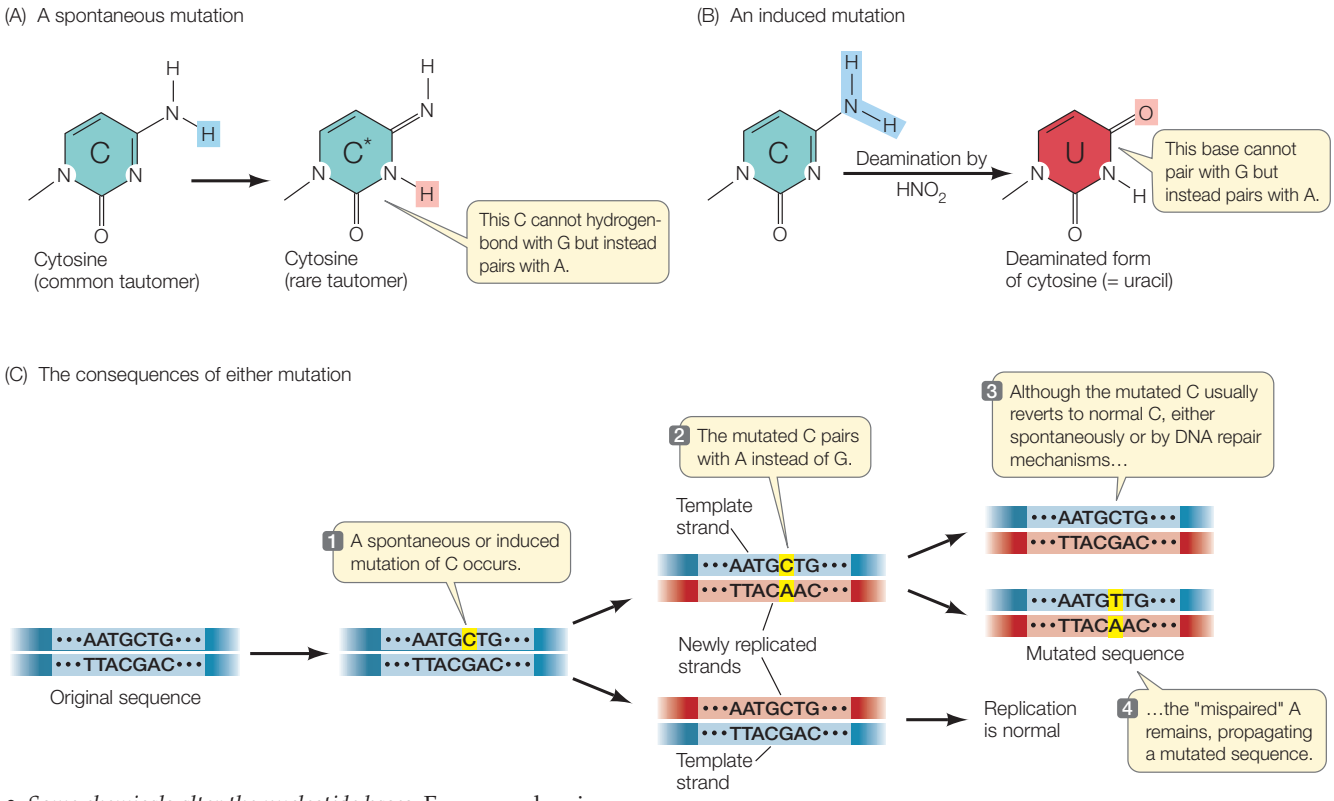
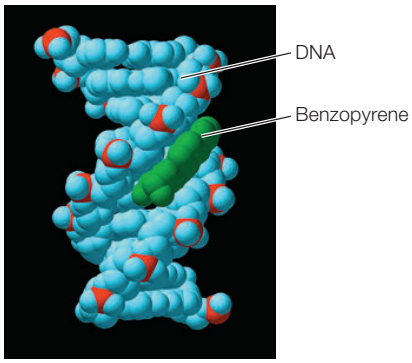


FIGURE 9.18 Spontaneous and Induced Mutations (A) Each of the nitrogenous bases exists in both a common (prevalent) form and a rare form. When a base spontaneously switches to its rare tautomer, it can pair with a different base. (B) Mutagens such as nitrous acid can induce changes in the bases. (C) The results of both spontaneous and induced mutations are permanent changes in the DNA sequence following replication.

- *Some chemicals alter the nucleotide bases*. For example, nitrous acid (HNO_2) reacts with cytosine and converts it to uracil by deamination (**FIGURE 9.18B**). This alteration has the same result as spontaneous deamination: instead of a G, DNA polymerase inserts an A (see Figure 9.18C).
- *Some chemicals add groups to the bases*. An example is benzopyrene, a component of cigarette smoke that adds a large chemical group to guanine, making it unavailable for base pairing. When DNA polymerase reaches such a modified guanine, it inserts any one of the four bases, resulting in a high frequency of mutations.



- *Radiation damages the genetic material*. Radiation can damage DNA in three ways. First, ionizing radiation (including X rays, gamma rays, and particles emitted by unstable isotopes) can detach electrons from atoms or molecules and produce highly reactive chemicals called free radicals. Free radicals can change bases in DNA to forms that are not recognized by DNA polymerase. Second, ionizing radiation can also break the sugar–phosphate backbone of DNA, causing chromosomal abnormalities. And third, ultraviolet radiation (from the sun or a tanning lamp) can cause thymine bases to form covalent bonds with adjacent thymines. This, too, plays havoc with DNA replication by distorting the double helix, and can result in a mutation.

Some base pairs are more vulnerable than others to mutation

DNA sequencing has revealed that mutations occur most often at certain base pairs. These "hotspots" are often located where cytosine has been methylated to 5-methylcytosine. (Methylation of DNA is a normal process in the regulation of chromatin structure and gene expression; see Concept 11.3.)

As we discussed above, unmethylated cytosine can lose its amino group to form uracil (see Figure 9.18B). This error is usually repaired because uracil (which normally occurs in

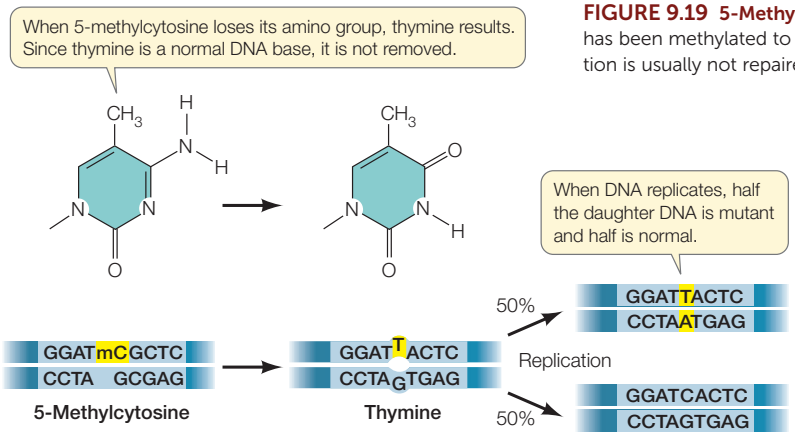


FIGURE 9.19 5-Methylcytosine Is a "Hotspot" for Mutations If a cytosine in DNA has been methylated to 5-methylcytosine and then becomes deaminated, the mutation is usually not repaired, and a C-G base pair is replaced with a T-A base pair.

RNA) is recognized by the repair mechanism as an inappropriate component of DNA. When 5-methylcytosine loses its amino group, however, the product is thymine, a natural base in DNA. The DNA repair mechanism ignores the thymine, and in this case, half of the new DNA molecules contain A-T base pairs instead of the original G-C pair (**FIGURE 9.19**). Thus the frequency of mutation is much greater in regions containing 5-methylcytosine.

Mutagens can be natural or artificial

Many people associate mutagens with materials made by humans, but there are also many naturally occurring mutagens. Plants (and to a lesser extent animals) make thousands of small molecules that serve a range of purposes, including defense against pathogens (see Concept 28.1). Some of these are mutagenic and potentially carcinogenic. An example of a naturally occurring mutagen is aflatoxin, which is made by many species of the mold *Aspergillus*. When mammals ingest the mold, the aflatoxin is converted into a product that binds to guanine and causes mutations, just as benzopyrene from cigarette smoke does.

Radiation can be human-made or natural. Some of the isotopes made in nuclear reactors and nuclear bomb explosions are certainly harmful—as was shown by the increased mutation rates in survivors of the atom bombs dropped on Japan in 1945. A certain amount of radiation comes from space in the form of cosmic radiation, and as mentioned above, natural ultraviolet radiation in sunlight also causes mutations. There is a well-established link between excessive exposure to the sun in light-skinned people and skin cancer.

Biochemists have estimated how much DNA damage occurs in the human genome under normal circumstances: among the haploid genome's 3.2 billion base pairs, there are about 16,000 DNA-damaging events per cell per day, of which 80 percent are repaired.

Mutations have both benefits and costs

What is the overall effect of mutation? For a species as a whole, the evolutionary benefits are clear. But there are costs as well as benefits for individual organisms.

BENEFITS OF MUTATIONS Mutations are the raw material of evolution. As you will see in Part 3 of this book, mutation alone does not drive evolution, but it provides the genetic diversity that makes natural selection possible. A mutation in a germline cell may have no immediate selective advantage to the organism, but it may cause a phenotypic change in the offspring. If the environment changes in a later generation, that mutation may be advantageous, enabling the species as a whole to adapt to changing conditions. A mutation in a somatic cell can sometimes benefit the individual organism, particularly if it occurs in a stem cell that produces a large number of offspring cells.

COSTS OF MUTATIONS Mutations can be harmful if they result in the loss of function of genes (and their protein products) or other DNA sequences that are needed for survival. A harmful mutation in a germline cell may be inherited in heterozygous form by the organism's descendants. If two individuals carrying the mutation mate, some of the offspring may be homozygous for the mutation. In their extreme form, such mutations produce phenotypes that are lethal. Lethal mutations can kill an organism during early development, or the organism may die before it matures and reproduces.

In Chapter 7 we described how mutations in somatic cells can lead to cancer. Typically these are mutations in oncogenes that result in the stimulation of cell division, or mutations in tumor suppressor genes that result in a lack of inhibition of cell division. These mutations can occur spontaneously or they can be induced.

We attempt to minimize our exposure to mutagens

Spontaneous mutagenesis is not in our control, but we can certainly try to avoid mutagenic substances and radiation. Not surprisingly, many things that cause cancer (carcinogens) are also mutagens. A good example is benzopyrene (discussed above), which is found in coal tar, car exhaust fumes, and charbroiled foods, as well as in cigarette smoke.

A major public policy goal is to minimize the effects of both human-made (anthropogenic) and natural mutagens on human health. For example, the Montreal Protocol (the only international environmental agreement signed and adhered to by all nations) bans chlorofluorocarbons (CFCs) and other substances that deplete the ozone layer in Earth's upper atmosphere. The ozone layer screens out ultraviolet radiation from the sun—radiation that can cause somatic mutations that lead to skin cancer. Similarly, bans on cigarette smoking have rapidly spread throughout the world. Cigarette smoking causes cancer because of the increased exposure of lung and other cells to benzopyrene and other carcinogens.

CHECKpoint CONCEPT 9.3

- ✓ What are the differences between a germline mutation and a somatic mutation?
- ✓ Describe how the same base—cytosine—can be mutated spontaneously or by a mutagen. In which case would the mutation be efficiently repaired?
- ✓ The *Bar eyes* mutation in fruit flies was found in genetic crosses to be inherited as an autosomal (non-X chromosome) dominant. Would you expect the mutant allele to be a loss-of-function or a gain-of-function mutation? Explain your answer.

Q What can we learn from ancient DNA?

ANSWER As we described in the opening story of this chapter, studying the DNA of fossils is a challenge. DNA is a large polymer that is rather stable (Concept 9.1). However, during the formation of fossils, most of the soft tissues and cells are broken down and their contents consumed by organisms in the environment. So it is not surprising that research on ancient DNA has been focused on places where DNA is likely to be preserved intact, such as frozen specimens and the interior of bones.

The polymerase chain reaction (PCR; Concept 9.2) has been invaluable in amplifying tiny amounts of ancient DNA to be sequenced. However, even a single molecule of contaminating DNA can ruin the experiment, because it too will be amplified. This has been a major challenge in studies of ancient DNA related to humans.

The Neandertal Genome Project involves an international team of scientists who are extracting DNA from the bones of skeletons of Neanderthals who lived in Europe more than 50,000 years ago, amplifying it by PCR, and then examining the genes. The entire DNA sequence has been completed. It is more than 99 percent identical to our human DNA, justifying the classification of Neanderthals as part of the same genus, *Homo*.

Comparisons of humans and Neanderthals with regard to specific genes and mutations (Concept 9.3) is ongoing and has already shown several interesting facts. For example:

- The gene *MC1R* is involved in skin and hair pigmentation. A point mutation found in Neanderthals but in not humans caused lower activity of the MC1R protein when it was induced in cell cultures. Such lower activity of MC1R is known to result in fair skin and red hair in humans. So it appears that at least some Neanderthals may have had pale skin and red hair (**FIGURE 9.20**).
- The gene *FOXP2* is involved in vocalization in many organisms, including birds and mammals. Mutations in this gene result in severe speech impairment in humans. The Neanderthal *FOXP2* gene is identical to that of humans, whereas that of chimpanzees is slightly different. This has led to speculation that Neanderthals may have been capable of speech.

FIGURE 9.20 A Neanderthal Child This reconstruction of a Neanderthal child who lived about 60,000 years ago was made using bones recovered at Gibraltar, as well as phenotypic projections made from DNA analyses.

While the two genome sequences are very similar, there are differences in many point mutations and larger chromosomal arrangements (Concept 9.3). There are distinctive "human" DNA sequences and also distinctive "Neanderthal" sequences. Some Neanderthal sequences have been found in modern humans, indicating that ancient humans and Neanderthals may have interbred. Such reconstructions of ancient DNAs and their phenotypic expression are being repeated for many other species. These studies underline the universality of DNA as the genetic material.

SUMMARY

CONCEPT 9.1 DNA Structure Reflects Its Role as the Genetic Material

- Circumstantial evidence for DNA as the genetic material includes its presence in the nucleus, its doubling during S phase of the mitotic cell cycle, and its injection into host cells by viruses. **Review Figures 9.1 and 9.2**
- Experimental evidence for DNA as the genetic material is provided by the **transformation** of one genotype into another by adding DNA. **Review Figure 9.3**
- In DNA, the amount of A equals the amount of T and the amount of G equals the amount of C. This observation, along with X-ray crystallography data, helped Watson and Crick unravel the **helical** structure of DNA. **Review Figures 9.4 and 9.5**

CONCEPT 9.2 DNA Replicates Semiconservatively

- DNA exhibits **semiconservative replication**. Each parent strand acts as a **template** for the synthesis of a new strand; thus the two replicated DNA molecules each contain one parent strand (the template) and one newly synthesized strand. **Review ANIMATED TUTORIALS 9.1 and 9.2**
- In DNA replication, the enzyme **DNA polymerase** catalyzes the addition of nucleotides to the 3′ end of each new strand. **Review Figure 9.7**
- Replication proceeds in both directions from the **origin of replication**. The parent DNA molecule unwinds to form a **replication fork**. **Review Figure 9.8 and ANIMATED TUTORIAL 9.3**
- **Primase** catalyzes the synthesis of a short RNA **primer** to which nucleotides are added by DNA polymerase. **Review Figure 9.9 and ACTIVITY 9.1**
- The **leading strand** is synthesized continuously. The **lagging strand** is synthesized in pieces called **Okazaki fragments**. The fragments are joined together by **DNA ligase**. **Review Figures 9.11 and 9.12 and ANIMATED TUTORIAL 9.4**
- Eukaryotic chromosomes have repetitive sequences at each end called **telomeres**. DNA replication leaves a short, unreplicated sequence at the 5′ end of each new DNA strand. Unless the enzyme **telomerase** is present, the sequence is removed. After multiple cell cycles the telomeres shorten, leading to chromosome instability and cell death. **Review Figure 9.13**
- DNA polymerases make errors, which can be repaired by proofreading and mismatch repair. **Review Figure 9.14**
- The **polymerase chain reaction**, or **PCR** technique uses DNA polymerase to make multiple copies of DNA in the laboratory. **Review Figure 9.15 and ANIMATED TUTORIAL 9.5**

CONCEPT 9.3 Mutations Are Heritable Changes in DNA

- **Somatic mutations** occur in the body cells of an individual and are passed on to daughter cells during mitosis. Only **germline mutations** (mutations in the cells that give rise to gametes) can be passed on to sexually produced offspring.
- **Point mutations** are alterations in single base pairs of DNA. **Silent mutations** can occur in genes or nontranscribed regions and do not affect the amino acid sequences of proteins. A mutation in a protein-coding region can lead to an alteration in the amino acid sequence of the protein. **Review Figure 9.16**
- Chromosomal mutations (**deletions**, **duplications**, **inversions**, and **translocations**) involve large regions of chromosomes. **Review Figure 9.17**
- **Spontaneous mutations** occur because of instabilities in DNA or chromosomes. **Induced mutations** occur when a mutagen damages DNA. **Review Figure 9.18**

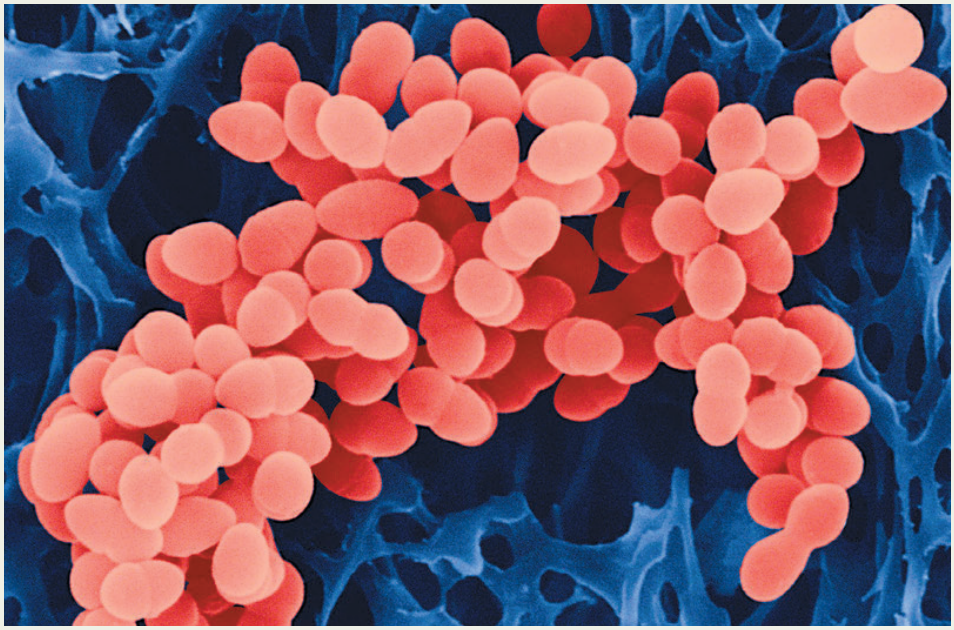
Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is9

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

10

From DNA to Protein: Gene Expression

KEY CONCEPTS



Although these *Staphylococcus aureus* cells look like normal bacteria, they have genes for resistance to multiple antibiotics and are difficult to eradicate. Antibiotic-resistant bacteria pose an ever increasing challenge to public health.

Humans have more prokaryotic cells on and in their bodies than they have eukaryotic cells of their own. Among the billions of bacteria that inhabit the skin and noses of many people is *Staphylococcus aureus*. Healthy people can carry this bacterium without symptoms, but sometimes, especially when the immune system has been weakened by age or disease, *S. aureus* can cause major skin infections and may even enter the body through the nose or a wound site. In these cases, much more serious infections of organs such as the heart and lungs can occur, and may result in death.

Until recently, most *S. aureus* infections were successfully treated with penicillin and related drugs, including methicillin. These antibiotics bind and inactivate several related enzymes (called penicillin-binding proteins) that are involved in the assembly of bacterial cell walls. Bacteria treated with these antibiotics have defective cell walls, and because of this, new cells cannot survive after cell division. Unfortunately, some *S. aureus* strains have acquired mutant versions of a penicillin-binding protein that can catalyze the assembly of cell walls in the presence of the antibiotics, thus conferring antibiotic resistance to these strains. The mutant penicillin-binding protein has an altered shape that doesn't bind the antibiotics. This protein is encoded by the *mecA* gene, which can be passed from one bacterium to another by bacterial conjugation (see Concept 8.4). At a more general level, the mutant phenotype demonstrates that *a gene is expressed as a protein*.

By the late 1990s these bacterial strains were being called "superbugs," with the formal name "methicillin-resistant *S. aureus*," or MRSA. The first decade of the new millennium saw a dramatic rise in MRSA infections. At first, most cases occurred in hospitals and nursing homes, but more recently MRSA has occurred in communities as well. Resistant strains have a selective advantage because of the extensive use of antibiotics in health care. With close to 100,000 serious MRSA infections and 20,000 deaths in the United States each year, more people are dying from MRSA than from AIDS.

MRSA can be treated if detected early. Antibiotics such as tetracycline, which targets bacterial protein synthesis, can be effective in some strains. But there is reasonable concern that MRSA may become resistant to these antibiotics as well.

Q How do antibiotics target bacterial protein synthesis?

You will find the answer to this question on page 213.

CONCEPT 10.1 Genetics Shows That Genes Code for Proteins

Following Mendel's definition of the gene as a physically distinct entity (see Concept 8.1), biologists identified the genetic material as DNA (see Concept 9.1). In this chapter we will show that in most cases, genes code for proteins, and it is proteins that determine phenotypes. The connection between protein and phenotype was made before it was known that DNA is the genetic material.

Observations in humans led to the proposal that genes determine enzymes

The identification of a gene product as a protein began with a mutation. In the early twentieth century, the English physician Archibald Garrod saw several children with a rare disease. One symptom was that the urine turned dark brown or black in air, and for this reason the disease was named alkaptonuria ("black urine").

Garrod noticed that the disease was most common in children whose parents were first cousins. Mendelian genetics had just been "rediscovered," and Garrod realized that because first cousins can inherit some alleles that are the same from their shared grandparents, their children are more likely than others to inherit a rare mutant allele from both parents—and therefore are more likely to be homozygous recessive for rare genetic conditions (see Figure 8.8B). Garrod proposed that alkaptonuria was a phenotype caused by a recessive mutant allele.

LINK

You can review the inheritance patterns of recessive alleles in **Concept 8.1**; see especially **Figure 8.8**

Garrod took the analysis a step further by identifying the biochemical abnormality in the affected children. He isolated from them an unusual substance, homogentisic acid, which accumulated in the blood, joints (where it crystallized and caused severe pain), and urine (where it turned black when exposed to air).

Enzymes as biological catalysts had just been discovered, and Garrod proposed that in healthy individuals, homogentisic acid might be broken down to a harmless product by an enzyme:

Normal allele
↓
Active enzyme
↓
Homogentisic acid ⟶ harmless product

Garrod speculated that the synthesis of the active enzyme is determined by the dominant wild-type allele of the gene that was mutated in alkaptonuria patients. These and other studies led him to correlate one gene to one enzyme, and to coin the term "inborn error of metabolism" to describe this kind of genetically determined biochemical disease.

But Garrod's hypothesis needed direct confirmation by the identification of the specific enzyme and the specific gene mutation involved. In 1958 the enzyme was identified as homogentisic acid oxidase, which breaks down homogentisic acid to a harmless product, just as Garrod predicted (**FIGURE 10.1**). The specific DNA mutation leading to alkaptonuria was described in 1996.

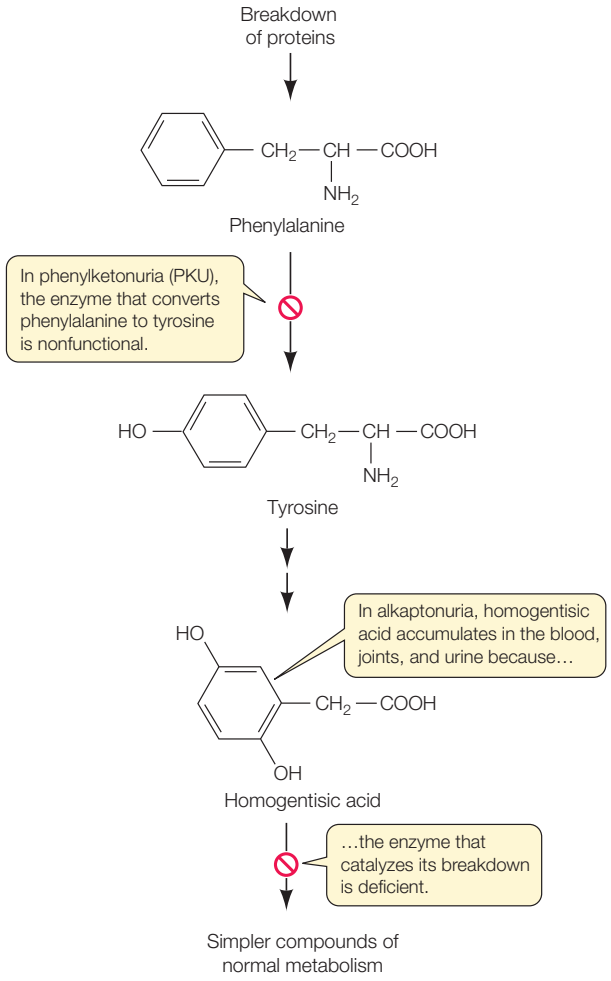


FIGURE 10.1 Metabolic Diseases and Enzymes Both phenylketonuria and alkaptonuria are caused by abnormalities in specific enzymes in a pathway that breaks down proteins.

Homogentisic acid is part of a biochemical pathway that catabolizes proteins, with the amino acids phenylalanine and tyrosine as intermediate products. Phenylketonuria, another genetic disease involving the same pathway, was discovered several decades after Garrod did his work. In phenylketonuria, the enzyme that converts phenylalanine to tyrosine is nonfunctional (see Figure 10.1). If left untreated, this disease leads to significant intellectual disability. Fortunately, the accumulation of phenylalanine can be easily detected in the blood of a newborn infant, and if the child consumes a diet low in proteins containing phenylalanine, intellectual disability is avoided.

The concept of the gene has changed over time

The phenotypic expression of mutations underlying alkaptonuria and phenylketonuria led to the "one gene–one enzyme" hypothesis. Once it was known that proteins (including enzymes) are polymers of amino acids, and that the sequence

APPLY THE CONCEPT

Genetics shows that genes code for proteins

Wild-type bacteria can synthesize the amino acid tryptophan (T) using a biochemical pathway that begins with chorismate (C) and involves four intermediates that we will call D, E, F, and G. Bacterial strains with mutant alleles for the five enzymes (1–5) involved in this pathway cannot synthesize tryptophan, and it must be supplied as a nutrient in the growth medium. The table gives the phenotypes of five mutant strains, each of which has a mutation in a gene for a different enzyme of the five. A "+" means the strain grew when the indicated compound was added to the medium, and a "0" means the strain did not grow. Based on these data, order the compounds (C, D, E, F, G, and T) and the enzymes (1, 2, 3, 4, and 5) in a biochemical pathway, as in the (incorrect) depiction here:

$$C \xrightarrow{1} D \xrightarrow{2} E \xrightarrow{3} F \xrightarrow{4} G \xrightarrow{5} T$$

Hint: Mutant strain 5 will not grow if any compound other than T is supplied, so it must carry a loss-of-function mutation in the enzyme that transforms another compound (either C, D, E, F, or G) into T. Thus enzyme 5 is the final enzyme in the pathway.

MUTANT STRAIN	ADDITION TO THE MEDIUM					
	C	D	E	F	G	T
1	0	0	0	0	+	+
2	0	+	+	0	+	+
3	0	+	0	0	+	+
4	0	+	+	+	+	+
5	0	0	0	0	0	+

of amino acids determines protein function, it became clear that *a mutant phenotype arises from a change in the protein's amino acid sequence*. However, scientists soon realized that the one gene–one enzyme (or one protein) hypothesis was an oversimplification. Once again, studies of human mutations were a key to this realization.

In humans, the oxygen-carrying protein hemoglobin has a quaternary structure of four polypeptide chains—two α-chains and two β-chains (see Concept 3.2, p. 47). Concept 9.3 introduced sickle-cell disease, which is caused by a point mutation in the gene for β-globin and is inherited as an autosomal recessive (carried on an autosome rather than a sex chromosome). In sickle-cell disease, one of the 146 amino acids in the β-globin chain is abnormal: at position 6, the normal glutamic acid has been replaced by valine. This replacement changes the charge of the protein (glutamic acid is negatively charged and valine is neutral), causing it to form long, needlelike aggregates in the red blood cells. The phenotypic result is anemia, an impaired ability of the blood to carry oxygen.

Because hemoglobin is easy to isolate and study, its variations in the human population have been extensively documented (**FIGURE 10.2**). Hundreds of single amino acid alterations in β-globin have been reported. For example, at the same position that is mutated in sickle-cell disease (resulting in hemoglobin S), the normal glutamic acid may be replaced by lysine, causing hemoglobin C disease. In this case, the resulting anemia is usually not severe. Many alterations of hemoglobin do not affect the protein's function. That is fortunate, because about 5 percent of all humans are carriers for one of these variants.

Studies of proteins that are made up of multiple polypeptides (such as hemoglobin) resulted in a modification of the one gene–one enzyme hypothesis. Scientists began to think of the relationship as **one gene–one polypeptide**, which remains a powerful and useful concept today. However, as you will see in this and later chapters, we are learning that this, too, is an oversimplification.

LINK

Exceptions to the one gene–one polypeptide relationship include the alternative splicing of RNA, which can produce multiple functional polypeptides from a single gene; see **Concept 11.4**

Mutations such as those that cause alkaptonuria and phenylketonuria result in alterations in amino acid sequences. But not all genes code for polypeptides. As we will see below and in Chapter 11, some DNA sequences are transcribed into RNA molecules that are *not* translated into polypeptides, but instead have other functions. Like all other DNA sequences, these RNA genes are subject to mutations, which may or may not affect the functions of the RNAs they produce.

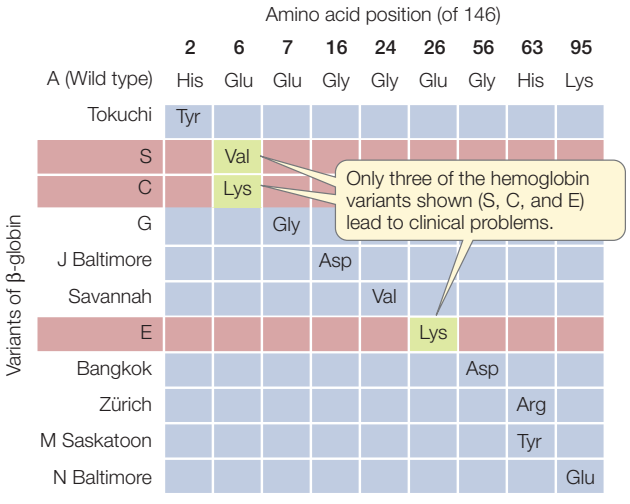


FIGURE 10.2 Gene Mutations and Amino Acid Changes Each of these mutant alleles (e.g., hemoglobin Tokuchi) codes for a β-globin polypeptide with an alteration in one of its 146 amino acids (e.g., a change from His to Tyr at position 2 of the polypeptide).

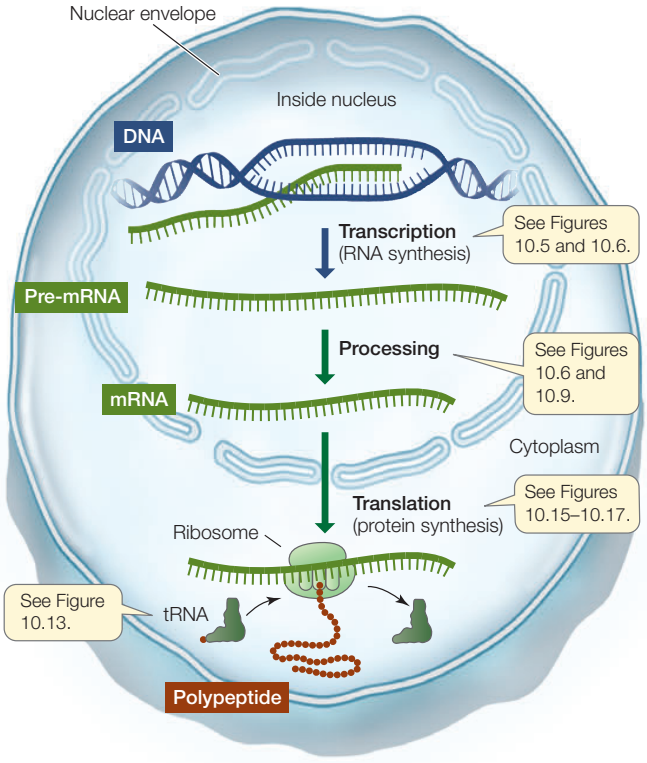


FIGURE 10.3 From Gene to Protein This diagram summarizes the processes of gene expression in eukaryotes.

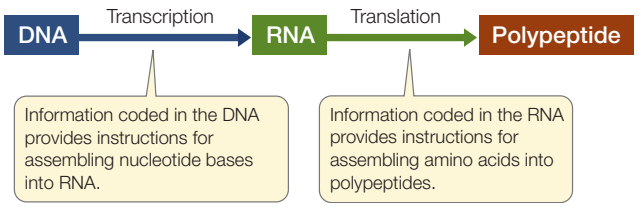
Go to ACTIVITY 10.1 Eukaryotic Gene Expression
PoL2e.com/ac10.1

Our understanding of genes and how they are expressed has increased dramatically over the past 60 years, since Watson and Crick first worked out the structure of DNA (see Concept 9.1). Let's begin our discussion of gene expression with an overview of the kinds of RNA molecules that are involved in transcription and translation.

Genes are expressed via transcription and translation

Molecular biology is the study of nucleic acids and proteins, and it often focuses on gene expression. As we described briefly in Chapter 3, genes are expressed as RNAs, many of which are translated into proteins. This process involves two steps:

- During **transcription**, the information in a DNA sequence (a gene) is copied into a complementary RNA sequence.
- During **translation**, this RNA sequence is used to create the amino acid sequence of a polypeptide.



Here we will consider three types of RNA with regard to their roles in protein synthesis:

- *Messenger RNA and transcription*: When a particular gene is expressed, the two strands of DNA unwind and separate into a **coding strand** and a **template strand**. The template strand is then transcribed to produce an RNA strand by complementary base pairing. The RNA strand is then modified to produce a **messenger RNA** (**mRNA**). In eukaryotic cells, the mRNA is processed in the nucleus and then moves to the cytoplasm, where it is translated into a polypeptide (**FIGURE 10.3**). The nucleotide sequence of the mRNA determines the ordered sequence of amino acids in the polypeptide chain, which is built by a ribosome.
- *Ribosomal RNA and translation:* The **ribosome** is essentially a protein synthesis factory with multiple proteins and several **ribosomal RNAs** (**rRNAs**). One of the rRNAs catalyzes peptide bond formation between amino acids, to form a polypeptide.
- *Transfer RNA mediates between mRNA and protein:* A third kind of RNA called **transfer RNA** (**tRNA**) can both bind a specific amino acid and recognize a specific sequence of nucleotides in mRNA, by complementary base pairing (A with U, and G with C). It is the tRNA that recognizes which amino acid should be added next to a growing polypeptide chain (see Figure 10.3).

In Chapter 11 we will consider other RNAs, which play roles in the regulation of gene expression.

CHECKpoint CONCEPT 10.1

- ✓ What is the difference between the "one gene–one protein" and "one gene–one polypeptide" hypotheses?
- ✓ What is the difference between gene transcription and translation?
- ✓ Could a person inherit homozygous recessive alleles for both alkaptonuria and phenylketonuria? If so, what would the symptoms be?
- ✓ Defining phenotype as the presence of a polypeptide chain of a particular amino acid sequence, would you expect the Zürich variant of β-globin (see Figure 10.2) to be inherited as a dominant, recessive, or codominant? Explain your answer.

In this section we have shown how the connection between genes and phenotypes can be understood in terms of DNA and proteins. We will now turn to some details of the process of gene expression, which is at the heart of what genes do.

CONCEPT 10.2 DNA Expression Begins with Its Transcription to RNA

Transcription—the formation of a specific RNA sequence from a specific DNA sequence—requires several key components:

- A DNA template for complementary base pairing
- The four ribonucleoside triphosphates (ATP, GTP, CTP, and UTP) to act as substrates
- An RNA polymerase enzyme

The same transcription process is responsible for the synthesis of mRNA, tRNA, and rRNA. Like mRNA, tRNA and rRNA are encoded by specific genes; their important roles in protein synthesis will be described in Concepts 10.3 and 10.4. There are also other kinds of RNA in the cell, with functions other than protein synthesis.

LINK

Small nuclear RNAs are involved in processing mRNA after it is transcribed, and microRNAs play important roles in stimulating or inhibiting gene expression; see **Concept 11.4**

RNA polymerases share common features

RNA polymerases from both prokaryotes and eukaryotes catalyze the synthesis of RNA from the DNA template. There is only one kind of RNA polymerase in bacteria and archaea, whereas there are several kinds in eukaryotes. However, they all share a common structure (**FIGURE 10.4**). Like DNA polymerases, RNA polymerases are processive; that is, a single enzyme–template binding event results in the polymerization of hundreds of RNA nucleotides. But unlike DNA polymerases, RNA polymerases do not require a primer.

Transcription occurs in three steps

Transcription can be divided into three distinct processes: initiation, elongation, and termination. You can follow these processes in **FIGURE 10.5**.

INITIATION Transcription begins with initiation, which requires a **promoter**, a special DNA sequence to which the RNA polymerase binds very tightly (see Figure 10.5A). Promoters are control sequences that "tell" the RNA polymerase two crucial things:

- Where to start transcription
- Which of the two DNA strands to transcribe and under what conditions

The promoter has a nucleotide sequence that can be "read" in a particular direction and orients the RNA polymerase, thus "aiming" it at the correct strand to use as a template. Part of each promoter is the **transcription initiation site**, where transcription begins. Groups of nucleotides lying "upstream" from the initiation site (5′ on the coding strand and 3′ on the template strand) are bound by other proteins, which help the RNA polymerase bind. These other proteins, called sigma factors and transcription factors, help determine which genes are expressed at a particular time in a particular cell.

LINK

The roles of sigma factors and transcription factors are described in **Concepts 11.1 and 11.2**

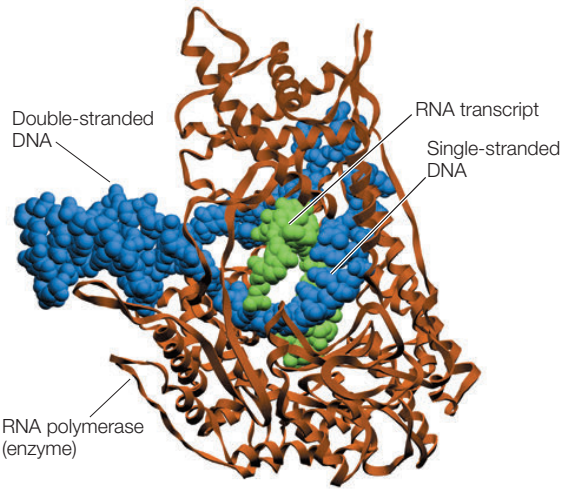


FIGURE 10.4 RNA Polymerase This enzyme from a virus is smaller than most other RNA polymerases, but its active site is similar to those of bacterial and eukaryotic RNA polymerases. The enzyme, shown in brown, is bound to the DNA it is transcribing (blue). The RNA transcript is shown in green.

Every gene has a promoter, but not all promoters are identical; some promoters are more effective at transcription initiation than others. Furthermore, bacteria, archaea, and eukaryotes differ in the details of transcription initiation. Despite these variations, the basic mechanisms of initiation are the same throughout the living world and provide further evidence of the biochemical unity of life on Earth.

ELONGATION Once RNA polymerase has bound to the promoter, it begins the process of **elongation** (see Figure 10.5B). RNA polymerase unwinds the DNA about 13 base pairs at a time and reads the template strand in the 3′-to-5′ direction. Like DNA polymerase, RNA polymerase adds new nucleotides to the 3′ end of the growing strand, beginning with the first nucleotide at the transcription initiation site. Thus the first nucleotide in the new RNA forms its 5′ end, and the RNA transcript is antiparallel to the DNA template strand.

RNA polymerase adds new nucleotides to the RNA molecule by complementary base pairing with nucleotides in the template strand of the DNA. This process is similar to DNA replication except that the base uracil (rather than thymine) in the RNA molecule is paired with adenine in the DNA molecule. In a mechanism very similar to that used by DNA polymerase, the RNA polymerase uses the ribonucleoside triphosphates ATP, UTP, GTP, and CTP as substrates and catalyzes the formation of phosphodiester bonds between them, releasing pyrophosphate in the process (see Figure 9.7). As transcription progresses, the two DNA strands rewind and the RNA grows as a single-stranded molecule (see Figure 10.5B).

Like DNA polymerases, RNA polymerases and associated proteins have mechanisms for proofreading during transcription, but these mechanisms are not as efficient as those for DNA. Transcriptional errors occur at rates of about 1 for every 10^4 to 10^5 bases. Because many copies of RNA are made, however, and because they often have relatively short life spans, these errors are not as potentially harmful as mutations in DNA.

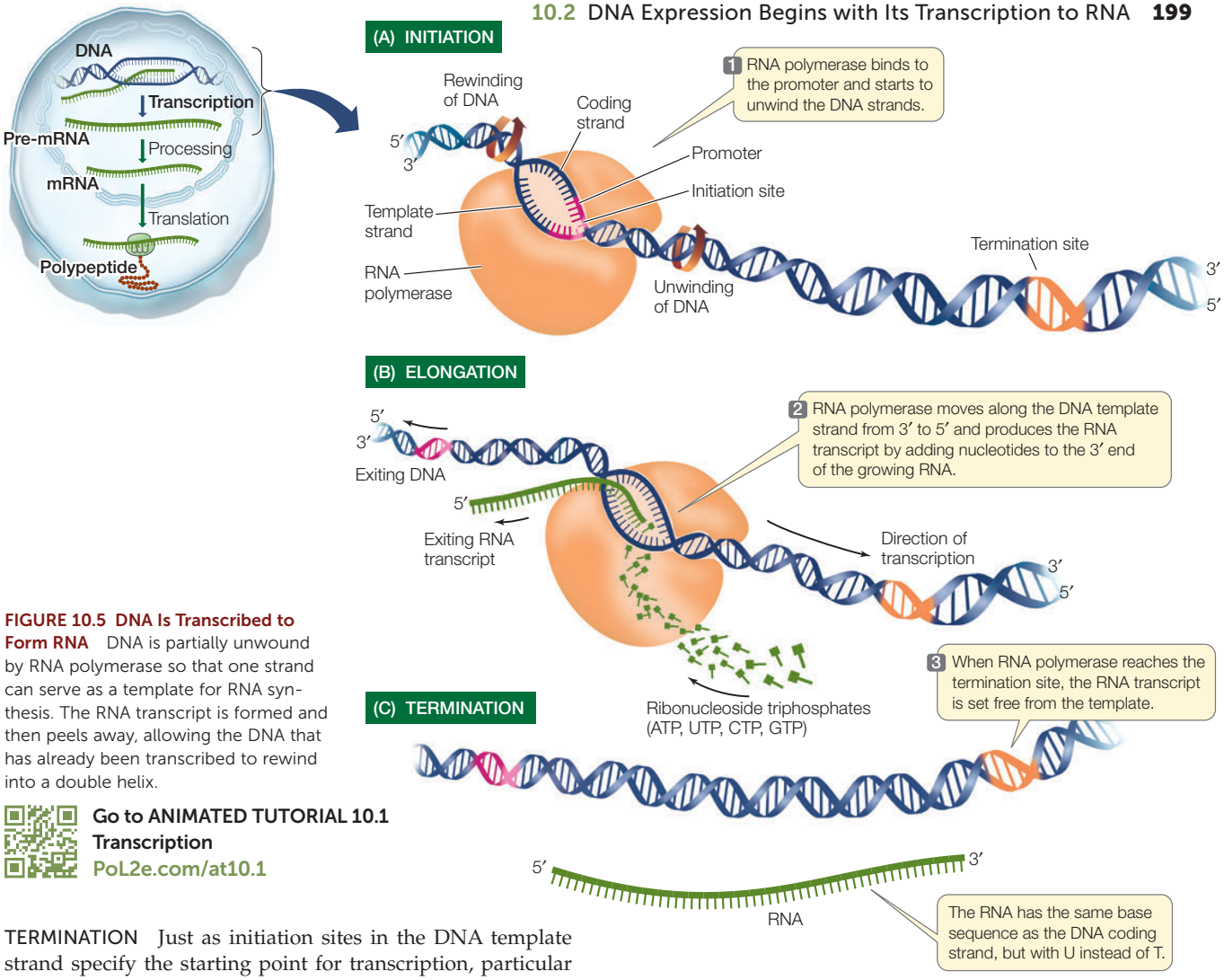


FIGURE 10.5 DNA Is Transcribed to Form RNA DNA is partially unwound by RNA polymerase so that one strand can serve as a template for RNA synthesis. The RNA transcript is formed and then peels away, allowing the DNA that has already been transcribed to rewind into a double helix.

Go to ANIMATED TUTORIAL 10.1
Transcription
PoL2e.com/at10.1

TERMINATION Just as initiation sites in the DNA template strand specify the starting point for transcription, particular base sequences specify its **termination** (see Figure 10.5C). The mechanisms of termination are complex and vary among different genes and organisms. In eukaryotes, multiple proteins are involved in recognizing the transcription termination site and separating the newly formed RNA strand from the DNA template and the RNA polymerase.

Eukaryotic coding regions are often interrupted by introns

Coding regions are sequences within a DNA molecule that are eventually translated as proteins. The coding region on the DNA template strand is transcribed into a complementary mRNA molecule, which has the same base sequence (with U's instead of T's) as the DNA coding strand. In prokaryotes, most of the genomic DNA is made up of coding regions, and the mRNAs are usually co-linear with them. That is, the mRNA sequence (e.g., 5′-AUGAUAGCCCC….) can be found in the DNA coding strand (e.g., 5′-ATGATAGCCCC….) without interruptions. In eukaryotes the situation is often different (**TABLE 10.1**).

A diagram of the structure and transcription of a typical eukaryotic gene is shown in **FIGURE 10.6**. In prokaryotes and viruses several adjacent genes sometimes share one promoter,

TABLE 10.1 Differences between Prokaryotic and Eukaryotic Gene Expression

Characteristic	Prokaryotes	Eukaryotes
Transcription and translation occurrence	At the same time in the cytoplasm	Transcription in the nucleus, then translation in the cytoplasm
Gene structure	DNA sequence usually not interrupted by introns	Transcribed regions often interrupted by noncoding introns
Modification of mRNA after initial transcription but before translation	Usually none	Introns spliced out; 5′ cap and 3′ poly A tail added

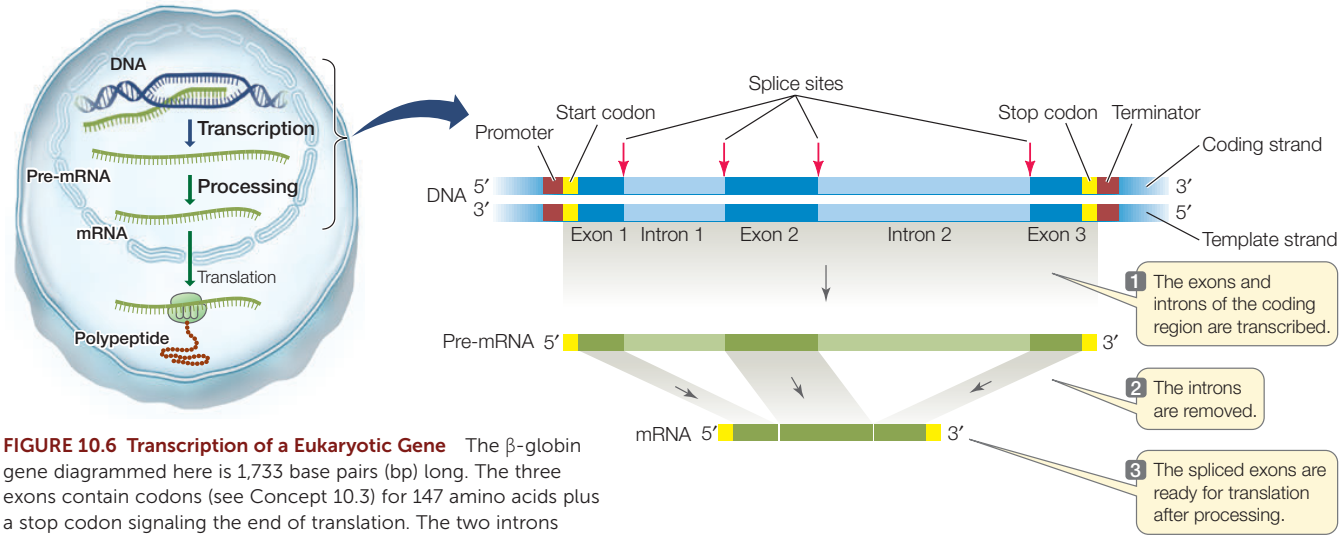


FIGURE 10.6 Transcription of a Eukaryotic Gene The β-globin gene diagrammed here is 1,733 base pairs (bp) long. The three exons contain codons (see Concept 10.3) for 147 amino acids plus a stop codon signaling the end of translation. The two introns (noncoding sequences containing almost 1,000 bp between them) are initially transcribed but then are spliced out of the pre-mRNA transcript.

but in eukaryotes each gene has its own promoter. And while the coding region of a prokaryotic gene is usually continuous (with no interruptions), a eukaryotic gene may contain noncoding sequences called **introns** (*int*ervening *regions*) that interrupt the coding region. The transcribed regions that are interspersed with the introns are called **exons** (*ex*pressed regi*ons*). Both introns and exons appear in the primary mRNA transcript, called the **precursor RNA**, or **pre-mRNA**, but the introns are removed by the time the mature mRNA leaves the nucleus. Pre-mRNA processing involves cutting introns out of the pre-mRNA transcript and splicing together the exon transcripts (see Figure 10.6). If this seems surprising to you, you are in good company. For scientists who were familiar with prokaryotic genes and gene expression, the discovery of introns in eukaryotic genes was entirely unexpected.

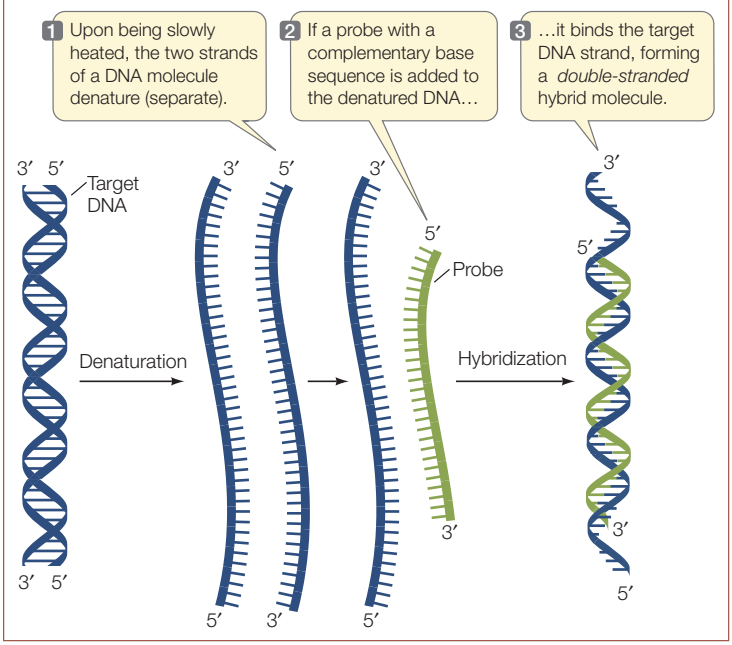
How can we locate introns within a eukaryotic gene? One way is by **nucleic acid hybridization**, the method that originally revealed the existence of introns. This method, outlined in FIGURE 10.7, has been crucial for studying the relationship between eukaryotic genes and their transcripts and is widely used in many applications. It involves two steps:

- The DNA to be analyzed is denatured by heat to break the hydrogen bonds between the base pairs and separate the two strands.
- A single-stranded nucleic acid from another source (called a **probe**) is incubated with the denatured DNA. If the probe has a base sequence complementary to the target DNA, a probe–target double helix forms by hydrogen bonding between the bases. Because the two strands are from different sources, the resulting double-stranded molecule is called a hybrid.

Biologists used this technique to examine the β-globin gene, which encodes a hemoglobin subunit (FIGURE 10.8). The researchers first denatured DNA containing the gene by heating it slowly, then used previously isolated β-globin mRNA as a probe. They were able to view the hybridized molecules using electron microscopy. As expected, the mRNA bound to the template DNA by complementary base pairing. The researchers expected to obtain a linear (1:1) matchup of the mRNA to the

RESEARCH TOOLS

FIGURE 10.7 Nucleic Acid Hybridization Base pairing permits the detection of a sequence that is complementary to the probe.



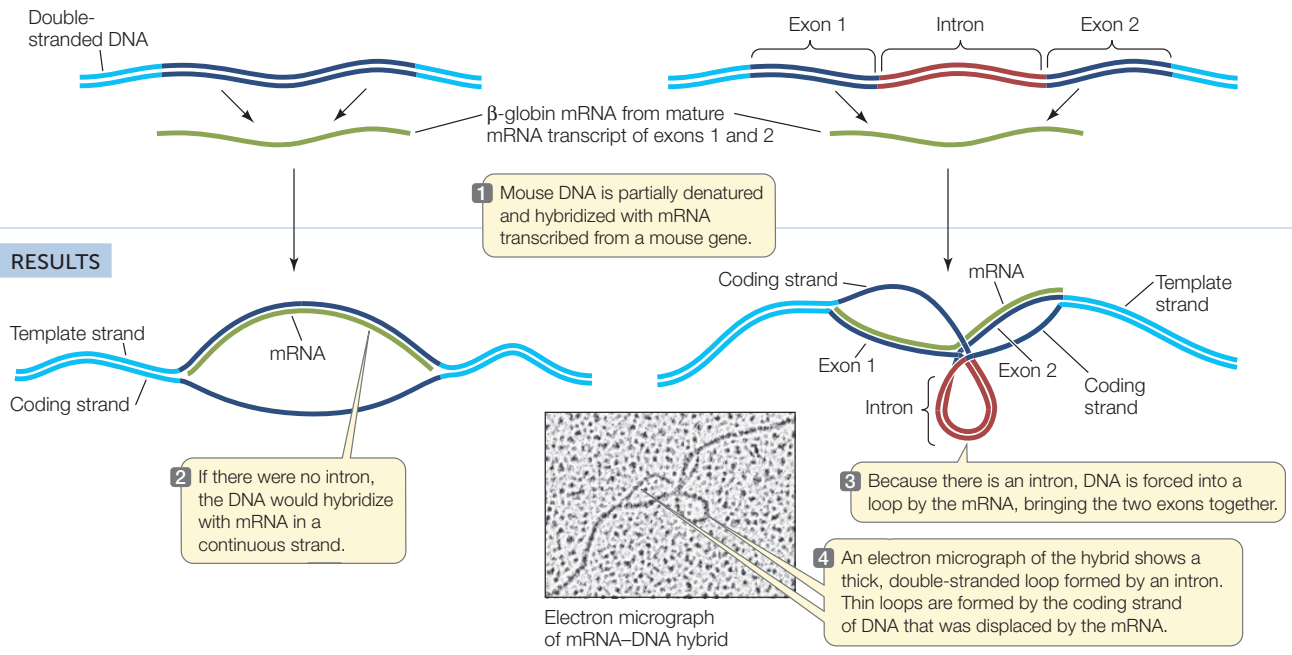
INVESTIGATION

FIGURE 10.8 Demonstrating the Existence of Introns When an mRNA transcript of the β-globin gene was hybridized with the double-stranded DNA of that gene, the introns in the DNA "looped out." This demonstrated that the coding region of a eukaryotic gene can contain noncoding DNA that is not present in the mature mRNA transcript.[a]

HYPOTHESIS

All regions within the coding sequence of a gene end up in its mRNA.

METHOD



RESULTS

CONCLUSION

The DNA contains noncoding regions within the genes that are not present in the mature mRNA.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]S. M. Tilghman et al. 1978. *Proceedings of the National Academy of Sciences USA* 75: 725–729.

coding DNA. That expectation was only partially met. There were indeed stretches of RNA–DNA hybrid, but some *unexpected looped structures* were also visible. These loops turned out to be introns, stretches of DNA that did not have complementary base sequences on the mature mRNA.

When pre-mRNA was used instead of mature mRNA to hybridize to the DNA, there was complete hybridization with *no loops*, revealing that the introns were part of the pre-mRNA transcript. Somewhere on the path from primary transcript (pre-mRNA) to mature mRNA, the introns had been removed, and the exons had been spliced together. We will examine this splicing process in the next section.

Introns *interrupt, but do not scramble*, the DNA sequence of a gene. The base sequences of the exons in the template strand, if joined and taken in order, form a continuous sequence that is complementary to that of the mature mRNA. Most (but not all) eukaryotic genes contain introns, and in rare cases, introns are also found in prokaryotes. The largest human gene encodes a muscle protein called titin; it has 363 exons, which together code for 38,138 amino acids. Can you deduce how many introns the titin gene has?

Eukaryotic gene transcripts are processed before translation

The primary transcript of a eukaryotic gene is modified in several ways before it leaves the nucleus: introns are removed, and both ends of the pre-mRNA are chemically modified.

SPLICING TO REMOVE INTRONS After the pre-mRNA is made, its introns must be removed. If this did not happen, the extra nucleotides in the mRNA would be translated at the ribosome and a nonfunctional protein would result. A process called **RNA splicing** removes the introns and splices the exons together.

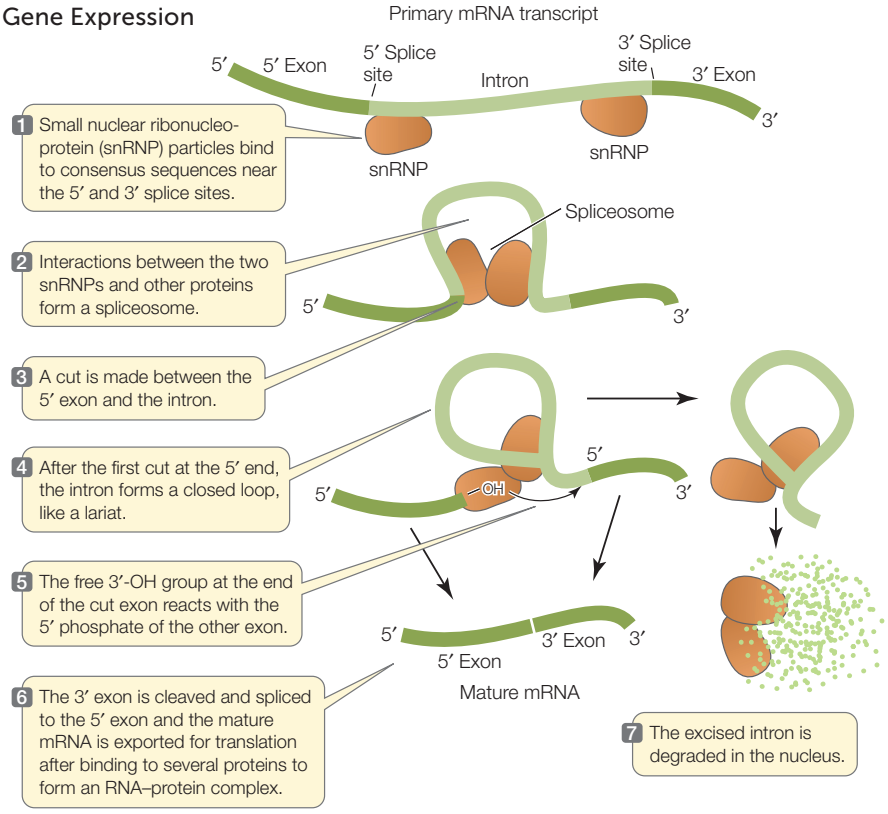


FIGURE 10.9 The Spliceosome: An RNA Splicing Machine Small nuclear ribonucleoprotein particles (snRNPs) bind to consensus sequences bordering the introns on pre-mRNA transcripts. Other proteins then bind, forming a large complex called a spliceosome. This structure determines the exact position of each cut in the pre-mRNA with great precision.

Go to ANIMATED TUTORIAL 10.2
RNA Splicing
PoL2e.com/at10.2

At the boundaries between introns and exons are **consensus sequences**—short stretches of DNA that appear with little variation ("consensus") in many different genes. As soon as the pre-mRNA is transcribed, these consensus sequences are bound by several **small nuclear ribonucleoprotein particles** (**snRNPs**) (FIGURE 10.9). The RNA in one of the snRNPs has a stretch of bases complementary to the consensus sequence at the 5′ exon–intron boundary, and it binds to the pre-mRNA by complementary base pairing. Another snRNP binds to the pre-mRNA near the 3′ intron–exon boundary, and then other proteins accumulate to form a large RNA–protein complex called a **spliceosome**. This complex cuts the pre-mRNA, releases the introns, and joins the ends of the exons together to produce mature mRNA.

Molecular studies of human genetic diseases have provided insights into intron consensus sequences and the splicing machinery. For example, people with the genetic disease beta thalassemia have a defect in the production of one of the hemoglobin subunits. These people suffer from severe anemia because they have an inadequate supply of red blood cells. In some cases, the genetic mutation that causes the disease occurs at an intron consensus sequence in the β-globin gene. Consequently, the β-globin pre-mRNA cannot be spliced correctly, and a defective β-globin mRNA is made. This finding offers another example of how biologists can *use mutations to elucidate cause-and-effect relationships*.

MODIFICATION AT BOTH ENDS While the pre-mRNA is still in the nucleus it undergoes two processing steps, one at each end of the molecule:

- A **5′ cap** (or G cap) is added to the 5′ end of the pre-mRNA as it is transcribed. The 5′ cap is a chemically modified molecule of guanosine triphosphate (GTP). It facilitates the binding of mRNA to the ribosome for translation, and it protects the mRNA from being digested by ribonucleases (enzymes that break down RNAs).
- A **poly A tail** is added to the 3′ end of the pre-mRNA at the end of transcription. This sequence of 100–300 adenine nucleotides assists in the export of the mRNA from the nucleus and is also important for mRNA stability.

CHECKpoint CONCEPT 10.2

✓ Part of a DNA template strand has the sequence 5′-ATGGTGTACG-3′. What will be the sequence of the RNA transcribed from this DNA? (Be careful to specify the 5′ and 3′ ends.)

✓ What would be the consequences of the following?

a. A mutation of a promoter sequence such that the promoter is deleted

b. A mutation of the gene that encodes RNA polymerase, such that the polymerase is not made

c. Deletion of intron consensus sequences from a gene

✓ Refer to the experiment shown in Figure 10.8. What would the result have been if there were five exons and four introns? Sketch what this would look like in an electron micrograph.

The transcription of a gene to produce mRNA is only the first step in gene expression. The next step in the pathway from DNA to RNA to protein is translation, the subject of Concepts 10.3 and 10.4. First we will discuss the genetic code, which enables the base sequence in an mRNA to be translated into a specific amino acid sequence in the resulting polypeptide. Then we will look in more detail at the process of translation.

CONCEPT 10.3 The Genetic Code in RNA Is Translated into the Amino Acid Sequences of Proteins

The translation of the nucleotide sequence of an mRNA into the amino acid sequence of a polypeptide occurs at the ribosome. In prokaryotes, transcription and translation are coupled: there is no nucleus, and ribosomes often bind to an mRNA as it is being transcribed in the cytoplasm. In eukaryotes, the nuclear envelope separates the locations of mRNA production and translation, the latter occurring at ribosomes in the cytoplasm. In both cases, the key event is the decoding of one chemical "language" (the nucleotide sequence) into another (the amino acid sequence).

The information for protein synthesis lies in the genetic code

The genetic information in an mRNA molecule is a series of sequential, nonoverlapping three-letter "words" called **codons**. Each codon specifies a particular amino acid. The "letters" are three adjacent nucleotide bases in the mRNA polynucleotide chain. Each codon in the mRNA is complementary to the corresponding triplet of bases in the template strand of the DNA molecule from which it was transcribed. The genetic code relates codons to their specific amino acids.

CHARACTERISTICS OF THE GENETIC CODE

Molecular biologists "broke" the genetic code in the early 1960s. The problem they addressed was perplexing: how could 20 different amino acids be specified using only four nucleotide bases (A, U, G, and C)? A triplet code with three-letter codons was considered likely because it was the shortest sequence with enough possible variations to encode all 20 amino acids. With four available bases, a triplet codon has $4 \times 4 \times 4 = 64$ variations.

Marshall Nirenberg and Heinrich Matthaei, at the U.S. National Institutes of Health, made the first decoding breakthrough in 1961 when they realized they could use a simple artificial polynucleotide instead of a complex natural mRNA as a template for polypeptide synthesis in a test tube. They could then identify the polypeptide the artificial messenger encoded. This led to the identification of the first two codons, as described in FIGURE 10.10.

Go to ANIMATED TUTORIAL 10.3
Deciphering the Genetic Code
PoL2e.com/at10.3

Other scientists later found that an artificial mRNA only three nucleotides long—amounting to one codon—could bind to a ribosome, and that the resulting complex could bind to a corresponding tRNA carrying a specific amino acid. Thus, for example, a simple UUU mRNA caused the tRNA carrying phenylalanine to bind to the ribosome. After this discovery, the complete deciphering of the genetic code was relatively simple.

The complete genetic code is shown in FIGURE 10.11. Notice that there are many more codons than there are different amino acids in proteins. All possible combinations of the four available bases give 64 (4^3) different three-letter codons, yet these codons determine only 20 amino acids. AUG, which codes for

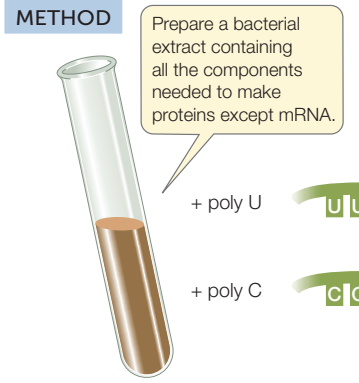
INVESTIGATION

FIGURE 10.10 Deciphering the Genetic Code Nirenberg and Matthaei used a test tube protein synthesis system to determine the amino acids specified by synthetic mRNAs of known compositions.[a]

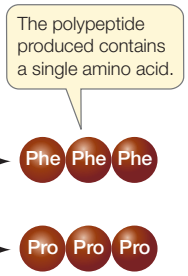
HYPOTHESIS

An artificial mRNA containing only one repeating base will direct the synthesis of a protein containing only one repeating amino acid.

METHOD



RESULTS



CONCLUSION

Poly U contains codons for phenylalanine only.
Poly C contains codons for proline only.

ANALYZE THE DATA

Poly U, an artificial mRNA, was added to a test tube with all other components for protein synthesis ("Complete system"). Other test tubes differed from the complete system as indicated in the table. Samples were tested for radioactive phenylalanine incorporation with the results shown in the table. Explain the results for each of the conditions.

Condition	Radioactivity in polypeptide
Complete system	29,500
Minus poly U mRNA	70
Minus ribosomes	52
Minus ATP	83
Plus RNase (hydrolyzes RNA)	120
Plus DNase	27,600
Mixture of 5 radioactive amino acids minus phenylalanine	33

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]M. W. Nirenberg and J. H. Matthaei. 1961. *Proceedings of the National Academy of Sciences USA* 47: 1588–1602.

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU, UUC: Phenylalanine UUA, UUG: Leucine	UCU, UCC, UCA, UCG: Serine	UAU, UAC: Tyrosine UAA, UAG: Stop codon Stop codon	UGU, UGC: Cysteine UGA: Stop codon UGG: Tryptophan	U C A G
C	CUU, CUC, CUA, CUG: Leucine	CCU, CCC, CCA, CCG: Proline	CAU, CAC: Histidine CAA, CAG: Glutamine	CGU, CGC, CGA, CGG: Arginine	U C A G
A	AUU, AUC, AUA: Isoleucine AUG: Methionine; start codon	ACU, ACC, ACA, ACG: Threonine	AAU, AAC: Asparagine AAA, AAG: Lysine	AGU, AGC: Serine AGA, AGG: Arginine	U C A G
G	GUU, GUC, GUA, GUG: Valine	GCU, GCC, GCA, GCG: Alanine	GAU, GAC: Aspartic acid GAA, GAG: Glutamic acid	GGU, GGC, GGA, GGG: Glycine	U C A G

FIGURE 10.11 The Genetic Code Genetic information is encoded in three-letter units—codons—that are read in the 5′-to-3′ direction on the mRNA. To decode a codon, find its first letter in the left column, then read across the top to its second letter, then read down the right column to its third letter. The amino acid the codon specifies is given in the corresponding row. For example, AUG codes for methionine, and GUA codes for valine.

Go to ACTIVITY 10.2 The Genetic Code
PoL2e.com/ac10.2

methionine, is also the **start codon**, the initiation signal for translation. The AUG codon is somewhat like the capitalized first word of a sentence, indicating how the sequence of words should be read. Three of the codons (UAA, UAG, and UGA) are **stop codons**, or termination signals for translation. When the translation machinery reaches one of these codons, translation stops and the polypeptide is released.

THE GENETIC CODE IS REDUNDANT BUT NOT AMBIGUOUS

The 60 codons that are not start or stop codons are far more than enough to code for the other 19 amino acids—and indeed, there is more than one codon for almost all the amino acids. Thus we say that the genetic code contains redundancies. For example, leucine is represented by six different codons (see Figure 10.11). Only methionine and tryptophan are represented by just one codon each.

A *redundant* code should not be confused with an *ambiguous* code. If the code were ambiguous, a single codon could specify two (or more) different amino acids, and there would be doubt about which amino acid should be incorporated into a growing polypeptide chain. The genetic code is not ambiguous: a given amino acid may be encoded by more than one codon, but each codon encodes only one amino acid.

THE GENETIC CODE IS (NEARLY) UNIVERSAL

The same genetic code is used by all the species on our planet. Thus the code must be an ancient one that has been maintained intact throughout the evolution of living organisms. Exceptions are known: within mitochondria and chloroplasts, the code differs slightly from that in prokaryotes and in the nuclei of eukaryotic cells; and in one group of protists, the codons UAA and UAG encode glutamine rather than functioning as stop codons. The significance of these differences is not yet clear. What is clear is that the exceptions are few.

The common genetic code unifies life, and indicates that all life came ultimately from a common ancestor. The genetic code probably originated early in the evolution of life. The common code also has profound implications for genetic engineering, as we will see in Chapter 13, since it means that the code for a human gene is the same as that for a bacterial gene. It is therefore impressive, but not surprising, that a human gene can be expressed in *Escherichia coli* via laboratory manipulations, since these cells speak the same "molecular language."

The codons shown in Figure 10.11 are for mRNA. The base sequence of the template DNA strand is complementary and antiparallel to these codons. Thus, for example, 3′-ACC-5′ in the template DNA strand corresponds to tryptophan (which is encoded by the mRNA codon 5′-UGG-3′). However, the coding DNA strand has the same sequence as the mRNA (but with T's

APPLY THE CONCEPT

The genetic code in RNA is translated into the amino acid sequences of proteins

The double-stranded DNA sequence for the coding region of a short peptide is:

```
5′- A T G T T T T C G A C G T G C G A T T G A -3′
3′- T A C A A A A G C T G C A C G C T A A C T -5′
    1         5          10         15         20
```

1. Which strand of DNA (top or bottom) is transcribed into mRNA? Explain.
2. What is the amino acid sequence of the peptide coded for by the DNA?
3. A mutant strain has a C-G base pair at position 5 instead of the T-A pair shown here. What is the amino acid sequence of the peptide? Explain.
4. A mutation at base pair 15 results in a peptide that is not full length. What point mutation would cause this? Explain.

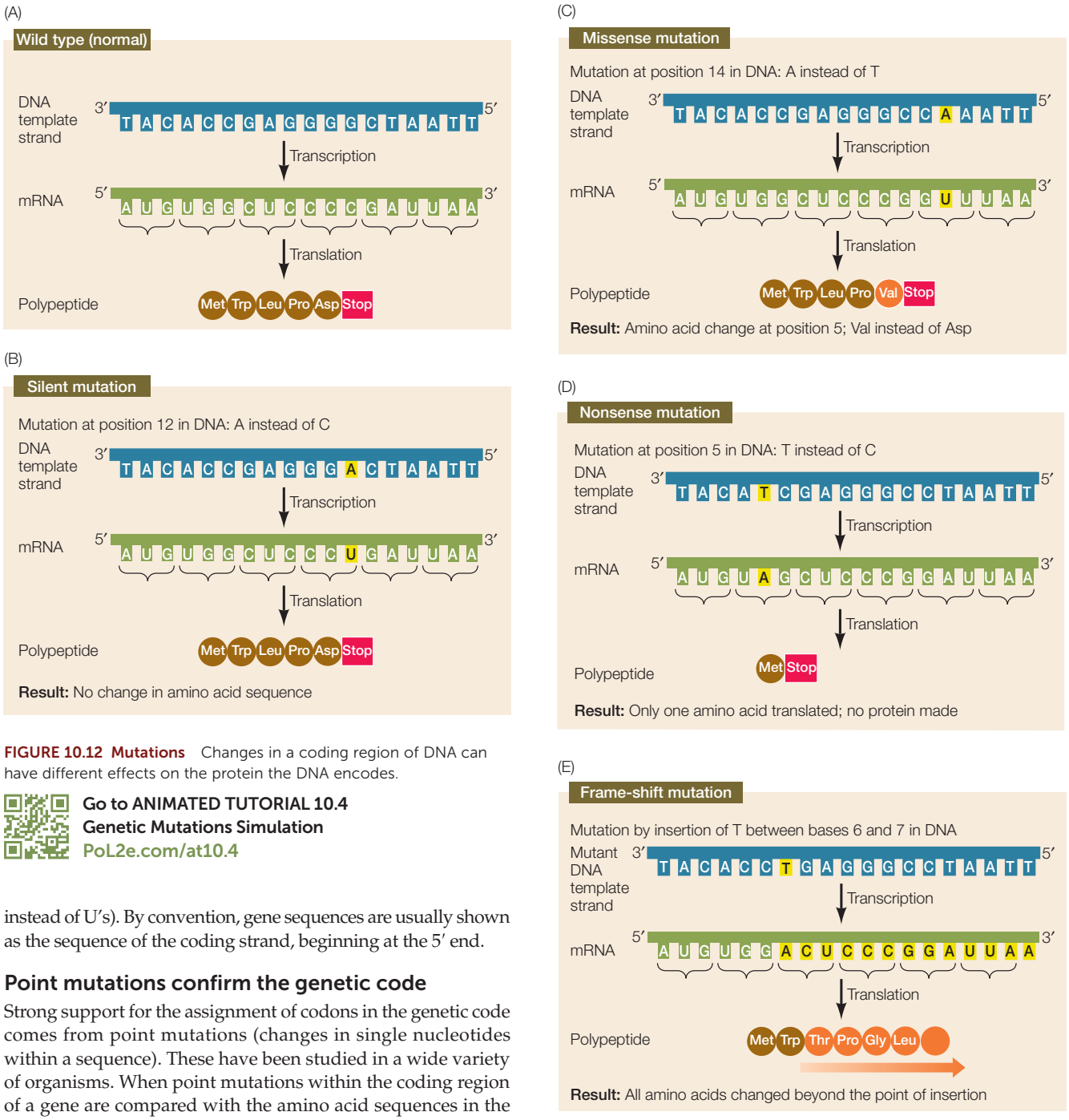


FIGURE 10.12 Mutations Changes in a coding region of DNA can have different effects on the protein the DNA encodes.

Go to ANIMATED TUTORIAL 10.4
Genetic Mutations Simulation
PoL2e.com/at10.4

instead of U's). By convention, gene sequences are usually shown as the sequence of the coding strand, beginning at the 5′ end.

Point mutations confirm the genetic code

Strong support for the assignment of codons in the genetic code comes from point mutations (changes in single nucleotides within a sequence). These have been studied in a wide variety of organisms. When point mutations within the coding region of a gene are compared with the amino acid sequences in the encoded polypeptide, they are consistent with the genetic code.

In Concept 9.3 we discussed mutations in terms of their effects on phenotypes. We can now define types of mutations in terms of their effects on polypeptide sequences (**FIGURE 10.12**):

- *Silent mutations* can occur because of the redundancy of the genetic code. For example, the codons CCG and CCU are both translated from mRNA as proline (Pro). So a change in the template strand of the DNA from 3′-GGC to 3′-GGA (a mutation from C to A) will not cause any change in amino acid sequence.
- *Missense mutations* result in a change in the amino acid sequence. For example, GAU in mRNA is translated as aspartic acid (Asp), whereas a mutation that results in GUU is translated as valine (Val).
- *Nonsense mutations* result in a premature stop codon. For example, the codon UGG is translated as the amino acid tryptophan (Trp). A DNA point mutation could convert this to the stop codon UAG, which acts as a translation termination

signal. If this occurred, the polypeptide chain would end at the amino acid translated just before the stop codon.

- *Frame-shift mutations* result from the insertion or deletion of one or more base pairs within the coding sequence. Since the genetic code is read as sequential, nonoverlapping triplets, this can cause new triplets to be read, and an altered sequence of amino acids in the resulting polypeptide.

CHECKpoint CONCEPT 10.3

- ✓ What are the characteristics of the genetic code?
- ✓ If the artificial mRNA UAUAUAUAUA... is used in a test tube protein synthesis system, what would be the amino acid sequence of the resulting polypeptide chain? Note that in this system translation can begin anywhere on the mRNA.
- ✓ A deletion of two consecutive base pairs in the coding region of DNA causes a frame-shift mutation. But a deletion of three consecutive base pairs causes the deletion of only one amino acid, with the rest of the polypeptide chain intact. Explain.

The mRNA with its coding information is translated into an amino acid sequence at the ribosome. We will now consider this process.

CONCEPT 10.4 Translation of the Genetic Code Is Mediated by tRNAs and Ribosomes

The translation of mRNA into proteins requires molecules that can link the information contained in each mRNA codon with a specific amino acid. That function is performed by a set of transfer RNAs (tRNAs). Two key events must take place to ensure that the protein made is the one specified by the mRNA:

- A tRNA must chemically read each mRNA codon correctly.
- The tRNA must deliver the amino acid that corresponds to the mRNA codon.

Once the tRNAs "decode" the mRNA and deliver the appropriate amino acids, components of the ribosome catalyze the formation of peptide bonds between the amino acids.

Go to ANIMATED TUTORIAL 10.5
Protein Synthesis
PoL2e.com/at10.5

Transfer RNAs carry specific amino acids and bind to specific codons

There is at least one specific tRNA molecule for each of the 20 amino acids. Each tRNA has three functions that are fulfilled by its structure and base sequence (**FIGURE 10.13**):

- *tRNAs bind to particular amino acids*. Each tRNA binds to a specific enzyme that attaches it to only 1 of the 20 amino acids. This covalent attachment is at the 3′ end of the tRNA. We will describe the details of this vital process in the next section. When it is carrying an amino acid, the tRNA is said to be "charged."
- *tRNAs bind to mRNA*. At about the midpoint on the tRNA polynucleotide chain there is a triplet of bases called the **anticodon**, which is complementary to the mRNA codon for the particular amino acid that the tRNA carries. Like the two strands of DNA, the codon and anticodon bind together via noncovalent hydrogen bonds. For example, the

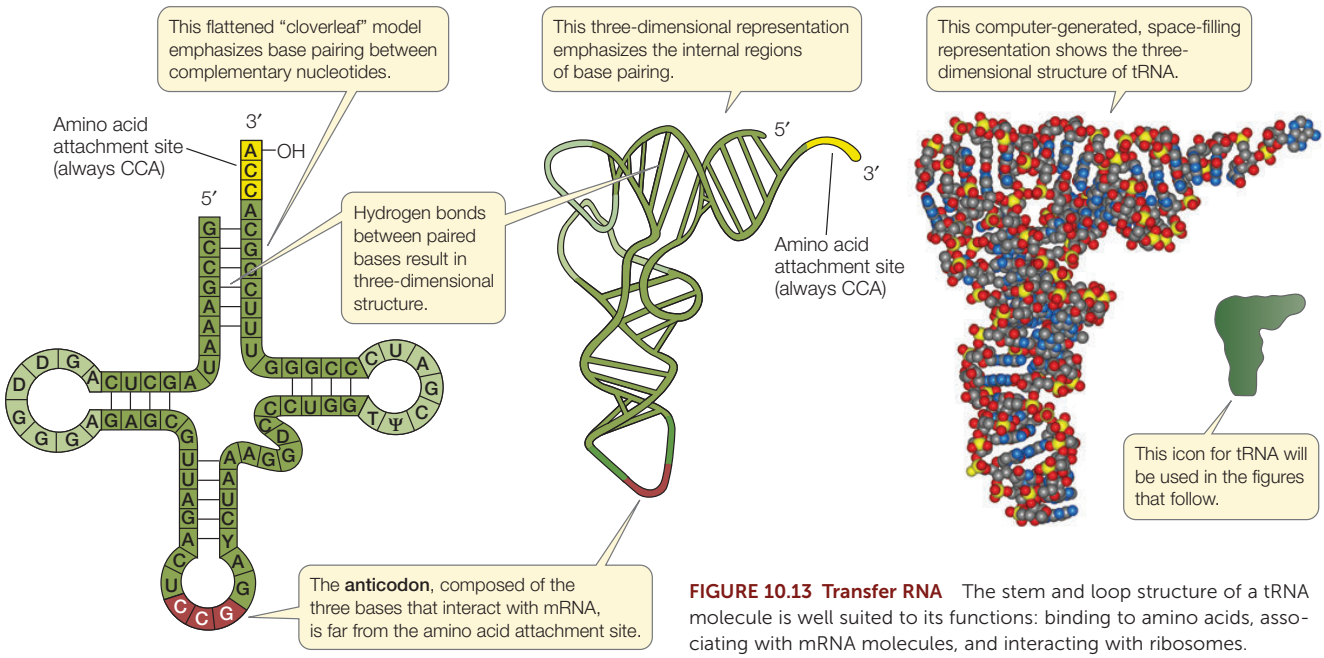


FIGURE 10.13 Transfer RNA The stem and loop structure of a tRNA molecule is well suited to its functions: binding to amino acids, associating with mRNA molecules, and interacting with ribosomes.

mRNA codon for arginine is 5′-CGG-3′, and the tRNA anticodon is 3′-GCC-5′.

- *tRNAs interact with ribosomes.* The ribosome has several sites on its surface that just fit the three-dimensional structure of a tRNA molecule. Interaction between the ribosome and the tRNA is noncovalent.

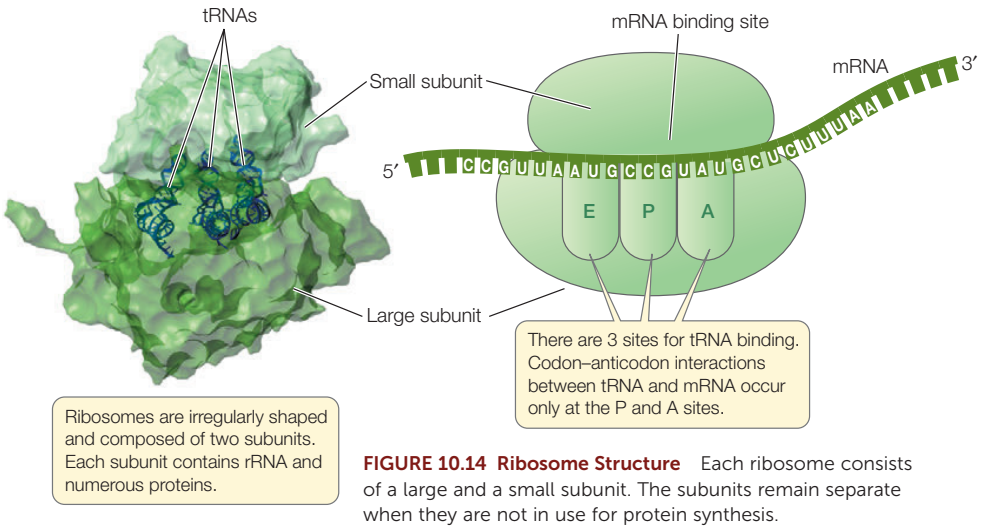
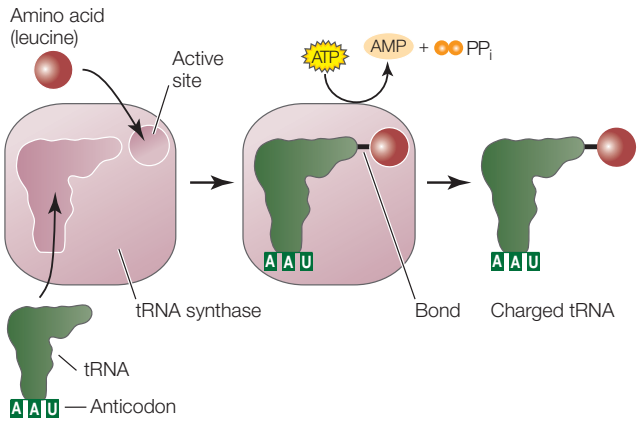


FIGURE 10.14 Ribosome Structure Each ribosome consists of a large and a small subunit. The subunits remain separate when they are not in use for protein synthesis.

Recall that 61 different codons encode the 20 amino acids in proteins (see Figure 10.11). Does this mean that the cell must produce 61 different tRNA species, each with a different anticodon? No. The cell gets by with about two-thirds of that number of tRNA species because the specificity for the base at the 3′ end of the codon (and the 5′ end of the anticodon) is not always strictly observed. This phenomenon is called "wobble," and it occurs because in some cases unusual or modified nucleotide bases occur in the 5′ position of the anticodon. One such unusual base is inosine (I), which can pair with A, C, and U. For example, its presence allows three of the alanine codons—GCA, GCC, and GCU—to be recognized by the same tRNA (with the anticodon 3′-CGI-5′). Wobble occurs in some matches but not in others; of most importance, it does not allow the genetic code to be ambiguous. That is, *each mRNA codon binds to just one tRNA species, carrying a specific amino acid*.

Each tRNA is specifically attached to an amino acid

The charging of each tRNA with its correct amino acid is achieved by a family of enzymes known as aminoacyl-tRNA synthases. Each enzyme is specific for one amino acid and for its corresponding tRNA. The reaction uses the energy in ATP to form a high-energy bond between the amino acid and the tRNA:



The energy in this bond is later used in the formation of peptide bonds between amino acids in a growing polypeptide chain.

Clearly, the specificity between the tRNA and its corresponding amino acid is extremely important. These two reactions, for example, are highly specific:

$$\text{Cysteine} + \text{tRNA}_{\text{cys}}\ (\text{anticodon ACA}) \xrightarrow{\text{Cys-tRNA synthase}} \text{Cys-tRNA}_{\text{cys}}$$

$$\text{Alanine} + \text{tRNA}_{\text{ala}}\ (\text{anticodon CGA}) \xrightarrow{\text{Ala-tRNA synthase}} \text{Ala-tRNA}_{\text{ala}}$$

A clever experiment by Seymour Benzer and his colleagues at Purdue University demonstrated the importance of this specificity. They took the Cys-tRNA$_{cys}$ molecule and chemically converted the cysteine into alanine, resulting in Ala-tRNA$_{cys}$. Which component—the amino acid or the tRNA—would be recognized when this hybrid charged tRNA was put into a protein-synthesizing system? The answer was the tRNA. Everywhere in the synthesized protein where cysteine was supposed to be, alanine appeared instead. The cysteine-specific tRNA had delivered its cargo (alanine) to every mRNA codon for cysteine. This experiment showed that the protein synthesis machinery recognizes the anticodon of the charged tRNA, not the amino acid attached to it.

Translation occurs at the ribosome

The ribosome is the molecular workbench where the translation of mRNA by tRNA is accomplished. All prokaryotic and eukaryotic ribosomes consist of two subunits (**FIGURE 10.14**). In eukaryotes, the large subunit consists of 3 different ribosomal RNA (rRNA) molecules and about 49 protein molecules arranged in a precise pattern. The small subunit consists of one rRNA molecule and about 33 proteins. These two subunits and several dozen other molecules interact noncovalently, fitting together like a jigsaw puzzle. If the hydrophobic interactions between the proteins and RNAs are disrupted, the ribosome falls apart, but it will reassemble perfectly when the disrupting agent is removed. When not active in the translation of mRNA, the ribosome exists as two separate subunits.

On the large subunit of the ribosome there are three sites to which a tRNA can bind, designated the A, P, and E sites (see Figure 10.14). The mRNA and ribosome move in relation to one another, and as they do so, a charged tRNA traverses these three sites in order:

- The *A (amino acid) site* is where the charged tRNA anticodon binds to the mRNA codon, thus lining up the correct amino acid to be added to the growing polypeptide chain.
- The *P (polypeptide) site* is where the tRNA adds its amino acid to the polypeptide chain.
- The *E (exit) site* is where the tRNA, having given up its amino acid, resides before being released from the ribosome and going back to the cytosol to pick up another amino acid and begin the process again.

The ribosome has a fidelity function, which ensures that a charged tRNA with the correct anticodon binds to the appropriate codon in the mRNA. When proper binding occurs, hydrogen bonds form between the three base pairs. The rRNA of the small ribosomal subunit plays a role in validating the three-base-pair match. Any tRNA that does *not* form hydrogen bonds with all three bases of the codon is ejected from the ribosome.

Translation takes place in three steps

Like transcription, translation occurs in three steps: initiation, elongation, and termination.

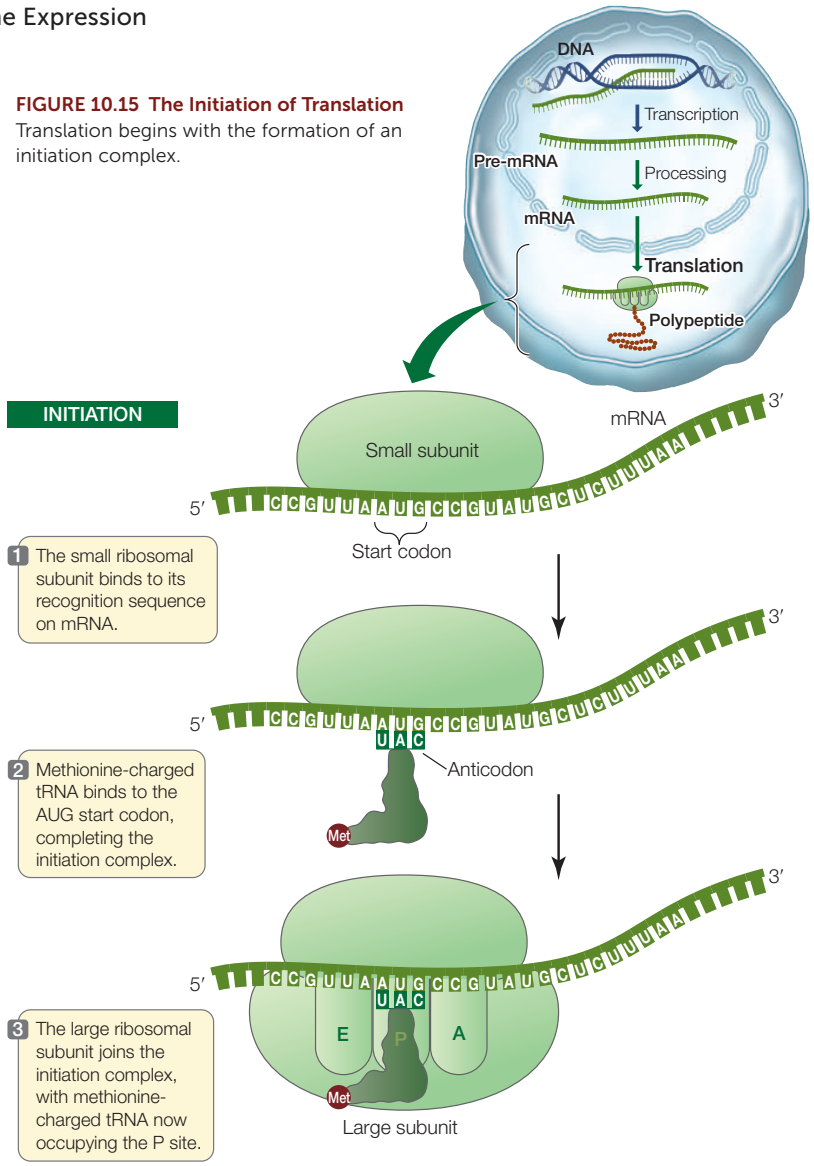
INITIATION The **initiation complex** consists of a charged tRNA and a small ribosomal subunit, both bound to the mRNA (**FIGURE 10.15**). While different organisms have different ways to effect this binding, here is an example from a prokaryote: The rRNA of the small ribosomal subunit binds by base pairing to a complementary sequence on the mRNA, about 8 base pairs upstream of the translation start codon (AUG; see Figure 10.11). After binding, the mRNA and the small subunit are aligned in such a way that the start codon at the beginning of the coding sequence will be aligned with the P site on the large subunit:

mRNA 5′ … A G G A G G … AUG … 3′

rRNA 3′ … U C C U C C … (P site) … 5′

The anticodon of a methionine-charged tRNA binds to the start codon by complementary base pairing to complete the initiation complex. Thus the first amino acid in a new polypeptide chain is always methionine. (In bacteria, but not archaea, the first amino acid is a slightly modified form of methionine called formylmethionine.) However, not all mature proteins have methionine as their first amino acid. In many cases, the initial methionine is removed by an enzyme after translation.

FIGURE 10.15 The Initiation of Translation
Translation begins with the formation of an initiation complex.



After the methionine-charged tRNA has bound to the mRNA, the large subunit of the ribosome joins the complex. The methionine-charged tRNA lies in the P site of the large subunit, and the A site is aligned with the second mRNA codon. These ingredients—mRNA, two ribosomal subunits, and methionine-charged tRNA—are put together properly by a group of proteins called initiation factors.

ELONGATION A charged tRNA whose anticodon is complementary to the second codon of the mRNA now enters the open A site of the large ribosomal subunit (**FIGURE 10.16**). The large subunit then catalyzes two reactions:

- It breaks the bond between the methionine and its tRNA in the P site.
- It catalyzes the formation of a peptide bond between the methionine and the amino acid attached to the tRNA in the A site.

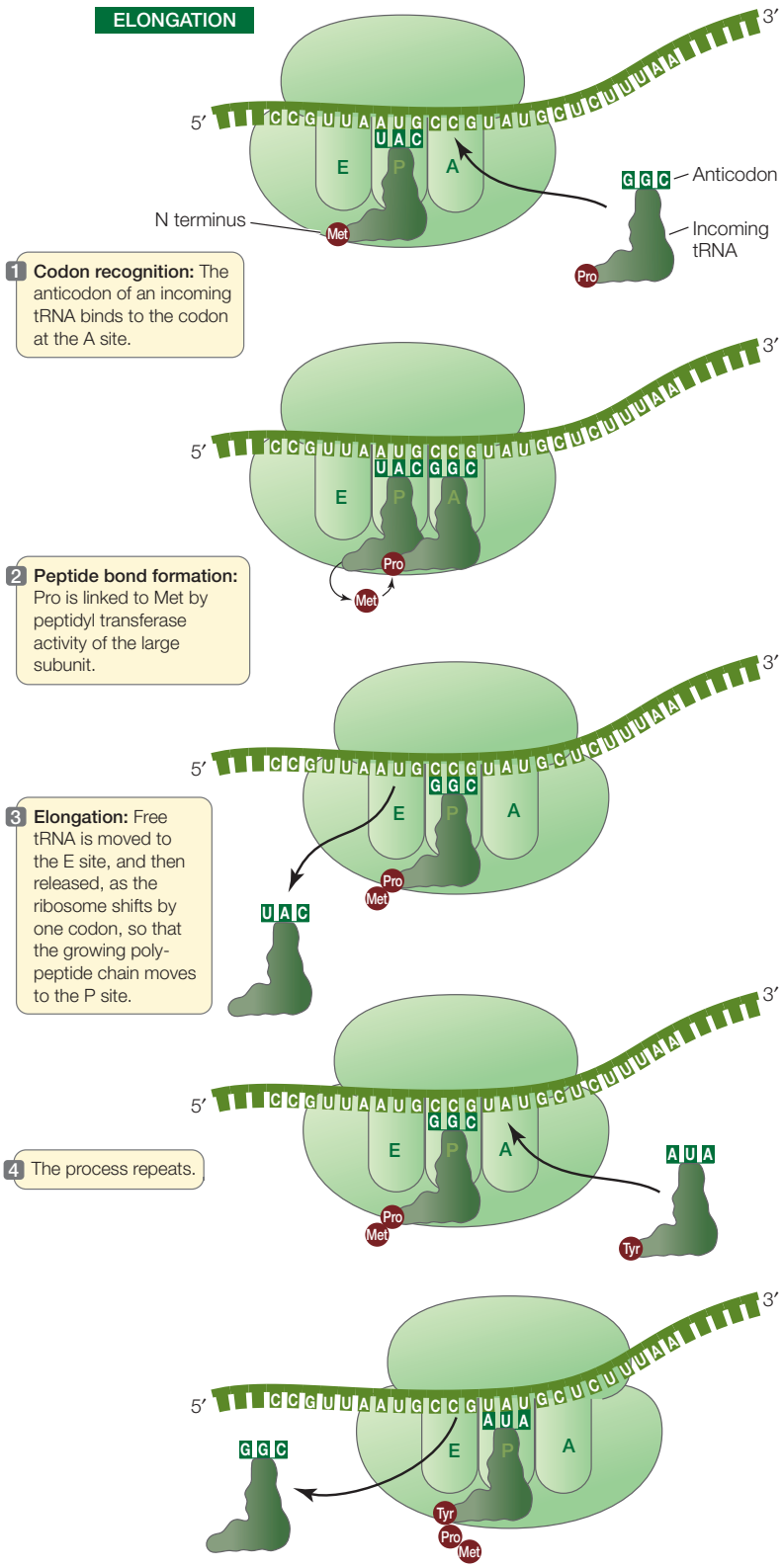


FIGURE 10.16 The Elongation of Translation The polypeptide chain elongates as the mRNA is translated.

Because the large ribosomal subunit performs these two actions, it is said to have **peptidyl transferase** activity. The component with this activity is actually one of the rRNAs in the ribosome, so the catalyst is an example of a **ribozyme** (from *ribo*nucleic acid and en*zyme*).

Methionine thus becomes the amino (N) terminus of the new protein (recall that polypeptides grow in the amino to the carboxyl direction; see Concept 3.2). The second amino acid is now bound to methionine but remains attached to its tRNA at the A site.

After the first tRNA releases its methionine, the ribosome moves so that the first tRNA is at the E site. The tRNA then dissociates from the ribosome and returns to the cytosol to become charged with another methionine. The second tRNA, now bearing a dipeptide (a chain of two amino acids), is shifted to the P site as the ribosome moves one codon along the mRNA in the 5′-to-3′ direction (see Figure 10.16). These steps are repeated, and the polypeptide chain grows as each new amino acid is added.

LINK

You can review the structure and formation of peptide bonds in Concept 3.2, especially Figure 3.6

TERMINATION The elongation cycle terminates at the end of the coding sequence, which is marked by a stop codon: UAA, UAG, or UGA (FIGURE 10.17). When a stop codon enters the A site, it binds a protein release factor, which allows hydrolysis of the bond between the polypeptide chain and the tRNA in the P site. The newly completed polypeptide then separates from the ribosome.

TABLE 10.2 summarizes the nucleic acid signals for initiation and termination of transcription and translation.

Go to MEDIA CLIP 10.1
Protein Synthesis:
An Epic on a Cellular Level
PoL2e.com/mc10.1

Polysome formation increases the rate of protein synthesis

Several ribosomes can simultaneously translate a single mRNA molecule, producing multiple polypeptides at the same time. As soon as the first ribosome has moved far enough from the translation initiation site, a second initiation complex can form, then a third, and so on. An assemblage consisting of a strand of mRNA with its beadlike ribosomes and their growing polypeptide chains is called a **polyribosome**, or **polysome** (FIGURE

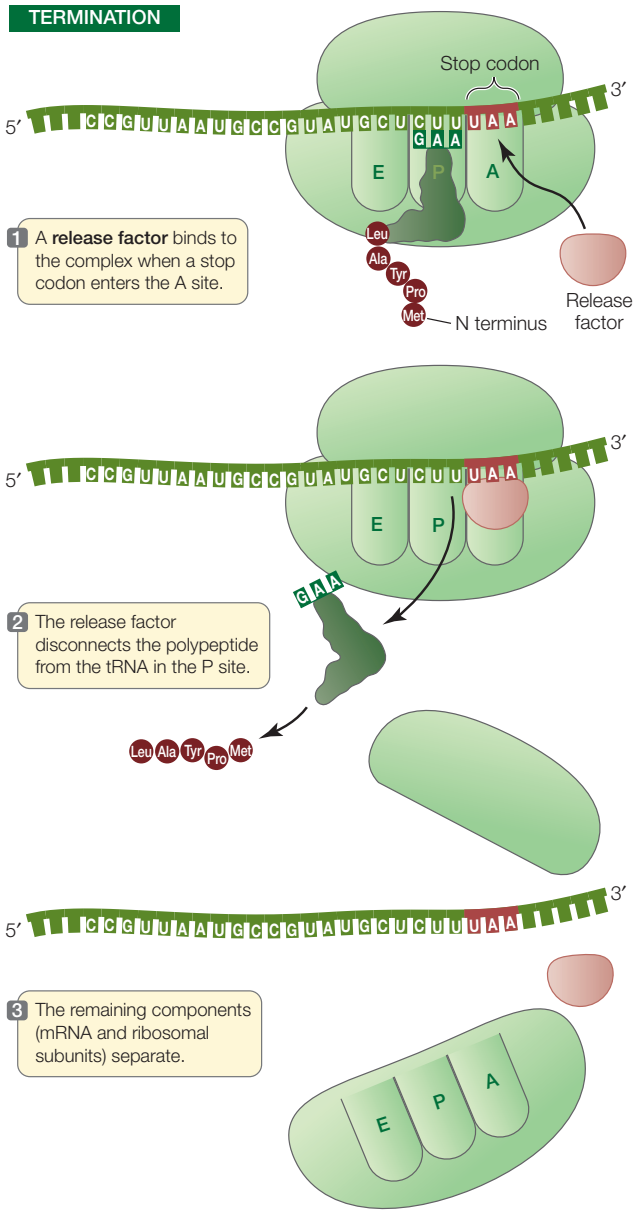


FIGURE 10.17 The Termination of Translation Translation terminates when the A site of the ribosome encounters a stop codon on the mRNA.

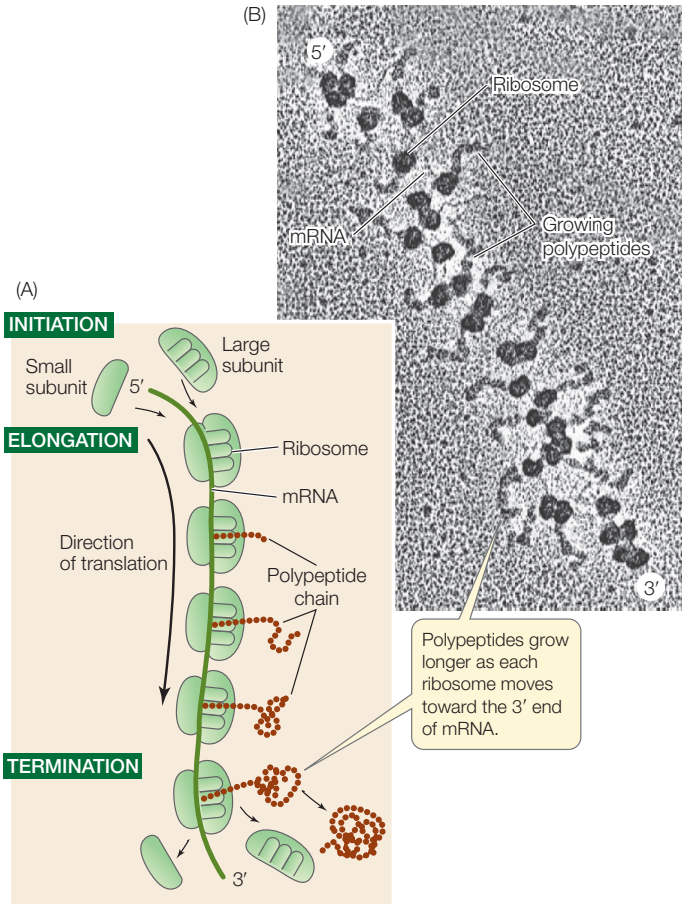


FIGURE 10.18 A Polysome (A) A polysome consists of multiple ribosomes and their growing polypeptide chains moving along an mRNA molecule. (B) An electron micrograph of a polysome.

10.18). Cells that are actively synthesizing proteins contain large numbers of polysomes and few free ribosomes or ribosomal subunits.

TABLE 10.2 Signals that Start and Stop Transcription and Translation

	Transcription	Translation
Initiation	Promoter DNA	AUG start codon in the mRNA
Termination	Terminator DNA	UAA, UAG, or UGA in the mRNA

CHECKpoint CONCEPT 10.4

- ✓ Describe the sequence of events in translation that involve tRNA, from charging with an amino acid to initiation, elongation, and termination.
- ✓ Imagine a polypeptide whose second amino acid is tryptophan. Sketch a ribosome with the mRNA and the first two tRNAs for this polypeptide, noting their positions in the A, P, and E sites.
- ✓ What would happen to polypeptides and cell function if a valine-tRNA synthase lost its specificity and attached any of the 20 amino acids to the 3′ end of the valine tRNA?

The process of protein synthesis usually does not end with translation. Proteins can undergo covalent modifications both during and after translation, with chemical groups being added or parts of the polypeptide chains removed. We now turn to these modifications.

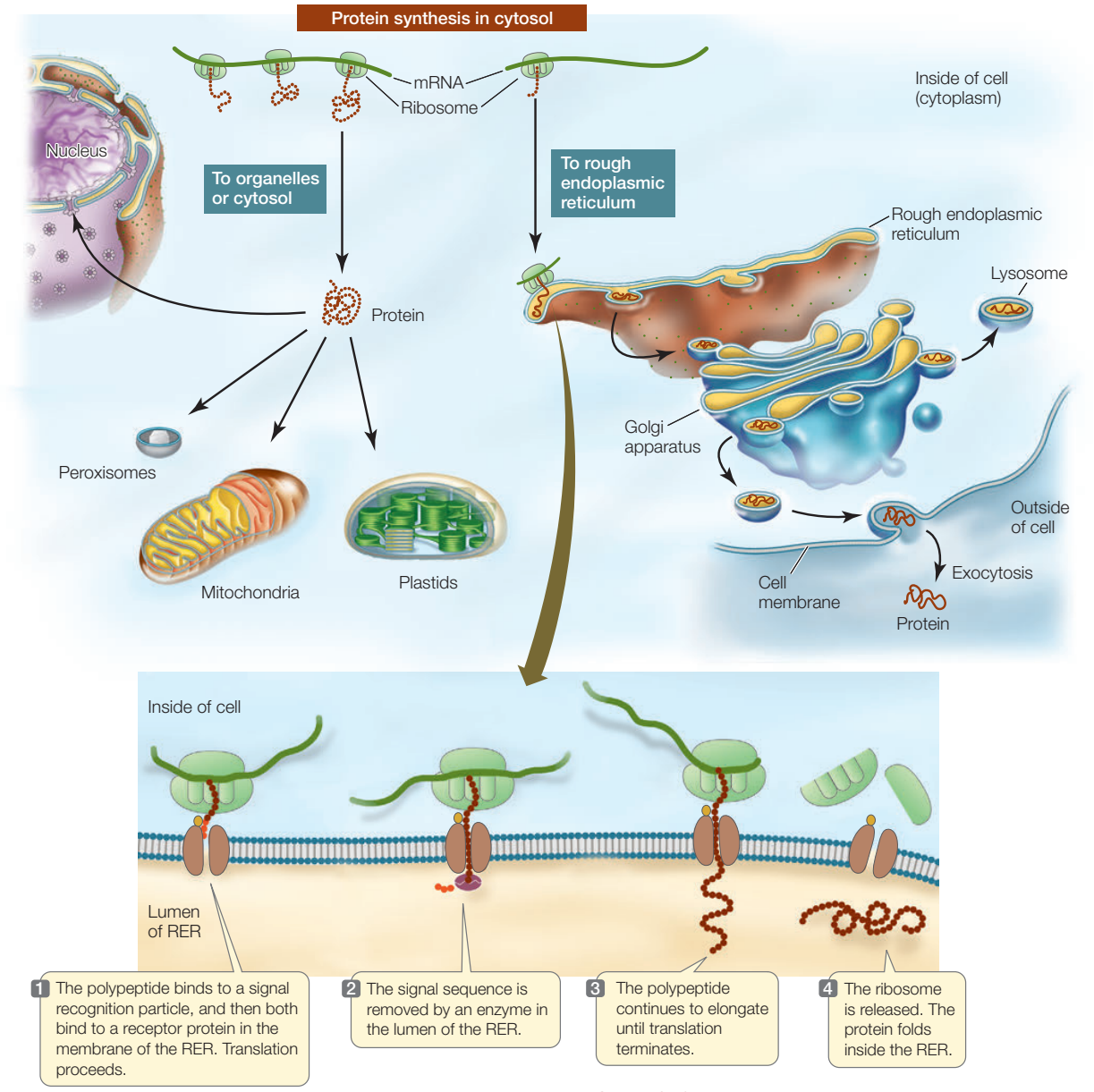


FIGURE 10.19 Destinations for Newly Translated Polypeptides in a Eukaryotic Cell Signal sequences on newly synthesized polypeptides bind to specific receptor proteins on the outer membranes of the organelles to which they are directed. Once the protein has bound to it, the protein enters the organelle through a channel in the membrane.

CONCEPT 10.5 Proteins Are Modified after Translation

The site of a polypeptide's function in the cell may be far away from its point of synthesis at the ribosome. This is especially true for eukaryotes, where a polypeptide may be moved into an organelle. Furthermore, polypeptides are often modified by the addition of new chemical groups that contribute to the function of the mature protein. In this section we examine these posttranslational aspects of protein synthesis.

Signal sequences in proteins direct them to their cellular destinations

Protein synthesis always begins on free ribosomes floating in the cytoplasm, and the "default" location for a protein is the cytosol. As the polypeptide chain emerges from the ribosome it may simply fold into its three-dimensional shape and perform its cellular role in the cytosol. However, a newly formed polypeptide may contain a **signal sequence** (or signal peptide)—a short stretch of amino acids that indicates where in the cell the polypeptide belongs. Proteins destined for different locations have different signals.

In the absence of a signal sequence, the protein will remain in the same cellular compartment where it was synthesized. Some proteins, however, contain signal sequences that "target" them to the nucleus, mitochondria, plastids, or peroxisomes (**FIGURE 10.19, LEFT**). A signal sequence binds to a specific receptor protein at the surface of the organelle. Once it has

bound, a channel forms in the organelle membrane, allowing the targeted protein to move into the organelle. For example, here is a nuclear localization signal (NLS):

-Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val-

The function of the NLS was established using experiments like the one illustrated in **FIGURE 10.20**. Proteins with or without this peptide were introduced into cells and then located by labeling the proteins with fluorescent dyes. Only proteins with the nuclear localization signal were found in the nucleus.

If a polypeptide carries a particular signal sequence of five to ten hydrophobic amino acids at its N terminus, it will be directed to the rough endoplasmic reticulum (RER) for further processing (**FIGURE 10.19, RIGHT AND BOTTOM**). Translation will pause, and the ribosome will bind to a receptor at the RER membrane. Once the polypeptide–ribosome complex is bound, translation will resume, and as elongation continues, the protein will traverse the RER membrane. Such proteins may be retained in the lumen (the inside) or membrane of the RER, or they may move elsewhere within the endomembrane system (Golgi apparatus, lysosomes, and cell membrane). If the proteins lack specific signals for destinations within the endomembrane system, they are usually secreted from the cell via vesicles that fuse with the cell membrane.

LINK

The endomembrane system and its functions are described in **Concept 4.3**

Many proteins are modified after translation

Most mature proteins are not identical to the polypeptide chains that are translated from mRNA on the ribosomes. Instead, most polypeptides are modified in any of a number of ways after translation (**FIGURE 10.21**). These modifications are essential to the final functioning of the protein.

- *Proteolysis* is the cutting of a polypeptide chain. For example, the ER signal sequence is cut off from the growing polypeptide chain as it enters the ER. Some mature proteins are actually made from polyproteins—long polypeptides containing the primary sequences of multiple distinct proteins—that are cut into final products by enzymes called proteases. Proteases are essential to some viruses, including human immunodeficiency virus (HIV), because the large viral polyprotein cannot fold properly unless it is cut. Certain drugs used to treat acquired immune deficiency syndrome (AIDS) work by inhibiting the HIV protease, thereby preventing the formation of proteins needed for viral reproduction.
- *Glycosylation* is the addition of carbohydrates to proteins to form glycoproteins. In both the ER and the Golgi apparatus, resident enzymes catalyze the addition of various oligosaccharides (short chains of monosaccharides; see Concept 2.3) to certain amino acid R groups on proteins. One such type of "sugar coating" is essential for directing proteins to lysosomes. Other types are important for protein conformation and for recognition functions at the cell surface. As we noted in Chapter 8, different chains of sugars added to red blood cell proteins determine an individual's blood type. In many cases the attached oligosaccharides help stabilize proteins, such as those in the extracellular matrix, and those in the storage vacuoles of plants.
- *Phosphorylation* is the addition of phosphate groups to proteins and is catalyzed by protein kinases. The charged phosphate groups change the conformation of the protein, often exposing the active site of an enzyme or the binding

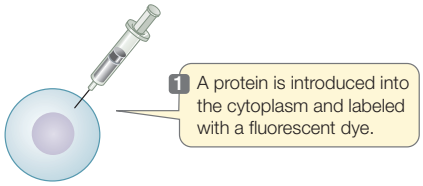
INVESTIGATION

FIGURE 10.20 Testing the Signal A series of experiments were used to test whether a nuclear localization signal (NLS) sequence is all that is needed to direct a protein to the nucleus.[a]

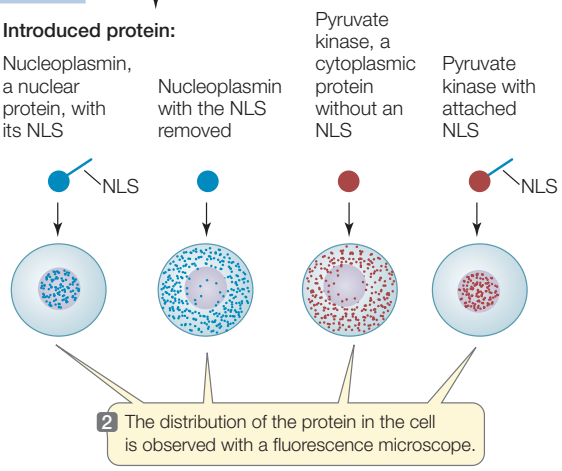
HYPOTHESIS

A nuclear localization signal is necessary for importing a protein into the cell nucleus.

METHOD



RESULTS



CONCLUSION

An NLS is essential for nuclear protein import and will direct a normally cytoplasmic protein to the nucleus.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]C. Dingwall et al. 1988. *Journal of Cell Biology* 107: 841–849.

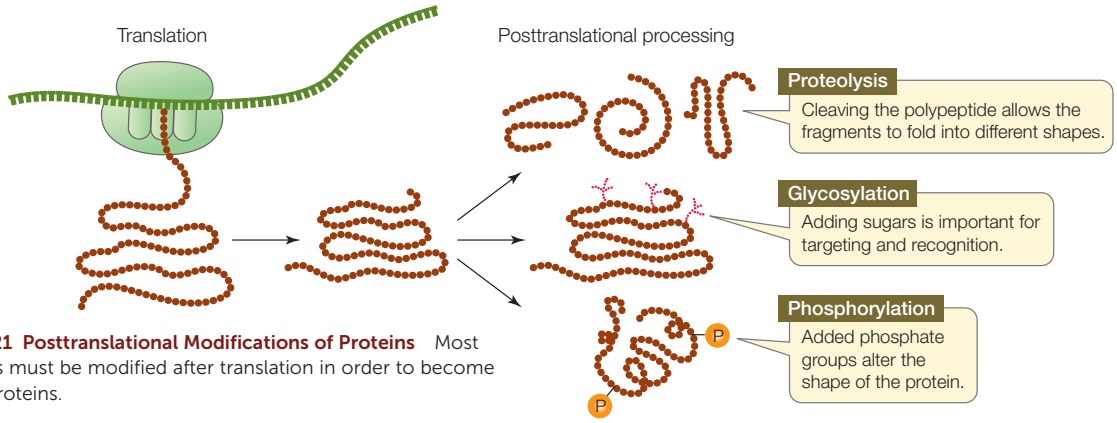


FIGURE 10.21 Posttranslational Modifications of Proteins Most polypeptides must be modified after translation in order to become functional proteins.

site for another protein. Phosphorylation is especially important in cell signaling (see Concepts 5.5 and 5.6).

CHECKpoint CONCEPT 10.5

- ✓ Describe how signal sequences determine where a protein will go after it is made.
- ✓ What are some ways in which posttranslational modifications alter protein structure and function?
- ✓ Describe an experiment you would perform to test a proposed chloroplast-targeting signal sequence. Be specific about the type of cell and the proteins you would use. Describe the results you would expect if the sequence is indeed a chloroplast-targeting signal.

Q How do antibiotics target bacterial protein synthesis?

ANSWER Tetracyclines are antibiotics that are effective against some strains of MRSA and many other bacterial infections. They derive their name from the four hydrocarbon rings that are common to this family of molecules. Tetracyclines kill bacteria by interrupting translation (Concept 10.1). They do this by binding noncovalently to the small subunit of bacterial ribosomes (Concept 10.4), where binding changes ribosome structure such that charged tRNAs can no longer bind to the A site on the ribosome (**FIGURE 10.22**). The specificity of antibiotics for bacterial ribosomes comes from the fact that bacterial and eukaryotic ribosomes have different proteins and RNAs. The target protein for tetracyclines is not present in eukaryotic ribosomes, so these antibiotics disrupt translation in prokaryotes but not in eukaryotes, including humans.

Strains of MRSA with resistance to tetracyclines are emerging. The genes that confer resistance to this group of antibiotics are carried on mobile genetic elements such as plasmids, which can move at high frequencies between bacteria, by bacterial conjugation (see Concept 8.4). Some of the resistance genes encode proteins that transfer the tetracyclines out of the cell, whereas others encode proteins that prevent the antibiotics from binding to the ribosomes. These resistance genes present a major challenge, because MRSA can be lethal. To overcome resistance, new antibiotics are being developed and tried. The evolutionary race between genetically caused drug resistance and new therapies continues. In the meantime, health-care providers and the general public are being advised to take precautions to prevent the spread of MRSA.

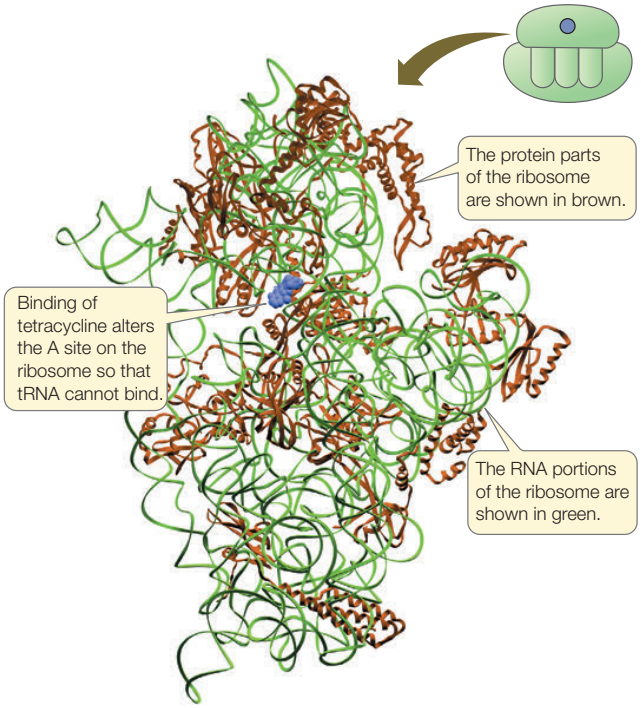


FIGURE 10.22 An Antibiotic at the Ribosome The antibiotic tetracycline binds to the small subunit of bacterial ribosomes. This causes a change in the structure of the A site, preventing tRNAs from binding, and protein synthesis stops.

SUMMARY

CONCEPT 10.1 Genetics Shows That Genes Code for Proteins

- Studies of human genetic diseases such as alkaptonuria linked genes to proteins. **Review Figure 10.1**
- Hemoglobin abnormalities demonstrate that mutations can alter the sequence of amino acids in proteins. **Review Figure 10.2**
- Genes are expressed via transcription and translation. During **transcription**, the information in a gene is copied into a complementary RNA sequence. During **translation**, this RNA sequence is used to create the amino acid sequence of a polypeptide. **Review Figure 10.3 and ACTIVITY 10.1**
- The product of transcription is **messenger RNA** (**mRNA**). **Transfer RNA** (**tRNA**) molecules translate the genetic information in the mRNA into a corresponding sequence of amino acids.
- **Ribosomal RNA** (**rRNA**) helps provide structure to the **ribosome** and acts as a ribozyme that catalyzes peptide bond formation between amino acids during protein synthesis.

CONCEPT 10.2 DNA Expression Begins with Its Transcription to RNA

- In a given gene, only one of the two strands of DNA (the template strand) acts as a template for transcription. **RNA polymerase** is the catalyst for transcription.
- RNA transcription from DNA proceeds in three steps: initiation, **elongation**, and **termination**. Initiation requires a **promoter** to which RNA polymerase binds. Elongation of the RNA molecule proceeds by the addition of nucleotides to the 3′ end of the molecule. **Review Figure 10.5 and ANIMATED TUTORIAL 10.1**
- After transcription, eukaryotic **pre-mRNA** is spliced to remove **introns**. **Review Figures 10.6 and 10.9 and ANIMATED TUTORIAL 10.2**
- Eukaryotic mRNA is also modified by the addition of a **5′ cap** and, at the 3′ end, a **poly A tail**.

CONCEPT 10.3 The Genetic Code in RNA Is Translated into the Amino Acid Sequences of Proteins

- Experiments involving synthetic mRNAs and protein synthesis in the test tube established the genetic code. **Review Figure 10.10 and ANIMATED TUTORIAL 10.3**
- The genetic code consists of triplets of mRNA nucleotide bases (**codons**) that correspond to 20 specific amino acids and to **start codons** and **stop codons**.
- The genetic code is redundant (an amino acid may be represented by more than one codon) but not ambiguous (no single codon represents more than one amino acid). **Review Figure 10.11 and ACTIVITY 10.2**
- Mutations in the coding regions of genes can be silent, missense, nonsense, or frame-shift mutations. **Review Figure 10.12 and ANIMATED TUTORIAL 10.4**

Review ANIMATED TUTORIAL 10.5

CONCEPT 10.4 Translation of the Genetic Code Is Mediated by tRNAs and Ribosomes

- Transfer RNA (tRNA) mediates between mRNA and amino acids during translation at the ribosome.
- Each tRNA species has an amino acid attachment site and an **anticodon** that is complementary to a specific mRNA codon. **Review Figure 10.13**
- A specific synthase enzyme charges each tRNA with its specific amino acid.
- Three sites on the large subunit of the ribosome interact with tRNA anticodons. The A site is where the charged tRNA anticodon binds to the mRNA codon. The P site is where the tRNA adds its amino acid to the growing polypeptide chain. The E site is where the tRNA is released. **Review Figure 10.14**
- Translation occurs in three steps: initiation, elongation, and termination. **Review Figures 10.15–10.17**
- In a **polyribosome**, or **polysome**, more than one ribosome moves along a strand of mRNA at one time. **Review Figure 10.18**

CONCEPT 10.5 Proteins Are Modified after Translation

- **Signal sequences** are short sequences of amino acids that direct polypeptides to their cellular destinations.
- These destinations include the nucleus and other organelles, which proteins enter after being recognized and bound by surface receptors.
- If a ribosome begins translating a polypeptide with an N-terminus RER signal sequence, it pauses and then resumes translation after attachment to a receptor in the RER membrane. **Review Figure 10.19**
- Posttranslational modifications of polypeptides include proteolysis, in which a polypeptide is cut into smaller fragments; glycosylation, in which sugars are added; and phosphorylation, in which phosphate groups are added. **Review Figure 10.21**

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is10

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

11

Regulation of Gene Expression

KEY CONCEPTS

Some rats are genetically programmed to prefer alcohol over plain water.

Many people drink alcoholic beverages, but relatively few of them become addicted to alcohol (alcoholic). Alcoholism is characterized by a compulsion to consume alcohol, tolerance (increasing doses are needed for the same effect), and dependence (abrupt cessation of consumption leads to severe withdrawal symptoms). In most alcoholics, alcohol provides pleasant sensations (positive reinforcement) and alleviates unpleasant ones such as anxiety (negative reinforcement).

Alcoholism is a complex disease. Psychologists sometimes speak of "addictive personalities," and genetic studies indicate there may be inherited factors that predispose people to the disease. One approach to describing the genes involved is to study animal models of alcoholism at the molecular level. James Murphy at Indiana University has bred a genetic strain of alcoholic rats, called P rats, that prefer alcohol when given the choice of alcohol-containing or alcohol-free water. P rats show many of the symptoms of addiction, including compulsive drinking, tolerance, and withdrawal. These rats appear more anxious than wild-type rats, spending more time in a closed rather than an open environment. Drinking alcohol alters this behavior and seems to relieve their anxiety.

There may be a link between a particular protein and alcohol consumption. CREB (*c*yclic AMP *r*esponse element *b*inding protein) is abundant in the brain and regulates the expression of hundreds of genes that are important in metabolism. CREB becomes activated when it is phosphorylated by the enzyme protein kinase A, which in turn is activated by the second messenger cyclic AMP. In an effort to understand the molecular basis of alcoholism and anxiety, neuroscientist Subhash Pandey and his colleagues at the University of Illinois compared CREB levels in the brains of P rats and wild-type rats. They found that P rats have inherently lower levels of CREB in certain parts of the brain. When these rats consumed alcohol, the total levels of CREB did not increase, but the levels of phosphorylated CREB did. It is the phosphorylated version of CREB that regulates gene transcription.

The prospect that CREB, a molecule that regulates gene expression, is a key element in the genetic propensity for alcoholism is important because it begins to explain the molecular nature of a complex behavioral disease. Such understanding may permit more effective treatment of alcohol abuse, or even its prevention. And for our purpose here, it underscores the importance of the regulation of gene expression in biological processes.

Q How does CREB regulate the expression of many genes?

You will find the answer to this question on page 232.

CONCEPT 11.1 Many Prokaryotic Genes Are Regulated in Operons

In Chapter 10 we introduced the concepts of gene expression. DNA is initially expressed as RNA, and in many cases the RNA is then translated into protein by the ribosome. Throughout this book we describe instances where gene expression is altered so that the level of protein produced from a particular gene is altered. Such changes are influenced by environmental conditions and the developmental stage of the cell or organism. Here are a few examples:

- In Chapter 5: When an extracellular signal binds to its receptor on a eukaryotic cell, it sets in motion a signal transduction pathway that may end with some genes being activated (their expression switched on) or others being repressed (their expression switched off).
- In Chapter 7: During the cell cycle, cyclins are synthesized only at specific stages. The genes for cyclins are inactive at other stages in the cycle.
- In Chapter 9: When a virus infects a host cell, it can "hijack" the host gene expression machinery and divert it to viral gene expression.

These and other examples indicate that gene expression is precisely regulated.

Genes are subject to positive and negative regulation

At every step of the way from DNA to protein, gene expression can be regulated (FIGURE 11.1). As we proceed through this chapter, you will see examples of gene regulation at the transcriptional, posttranscriptional, translational, and posttranslational levels. An important form of gene regulation is at the level of transcription.

Gene expression begins at the **promoter**, a region of DNA containing the site where RNA polymerase binds to initiate transcription. As we mentioned above, not all genes are active (being transcribed) at a given time. Two types of regulatory proteins—called **transcription factors**—control whether or not a gene is active: repressors and activators. These proteins bind to specific DNA sequences at or near the promoter (FIGURE 11.2):

- In negative regulation, a **repressor** binds a specific site in or near the promoter to prevent transcription.
- In positive regulation, the binding of an **activator** stimulates transcription.

You will see these mechanisms, or combinations of them, as we examine the regulation of prokaryotic, eukaryotic, and viral genes. Let's begin by looking at the regulation of gene expression in prokaryotes.

LINK

You can review the processes of transcription in Concept 10.2

Regulating gene transcription is a system that conserves energy

Prokaryotes conserve energy and resources by making certain proteins only when they are needed. Because their environments can change abruptly, prokaryotes have evolved mechanisms to rapidly alter the expression levels of certain genes when conditions warrant. An example is the bacterium *Escherichia coli*, which normally inhabits the intestines of humans and other mammals. *E. coli* must be able to adjust to sudden changes in its chemical environment as the foods consumed by its host change (for example, from a meal containing glucose at one time to one containing lactose at another). In many cases *E. coli* responds to such changes by changing the transcription of its genes. To illustrate this we will look at the regulation of the pathway for lactose catabolism in *E. coli*.

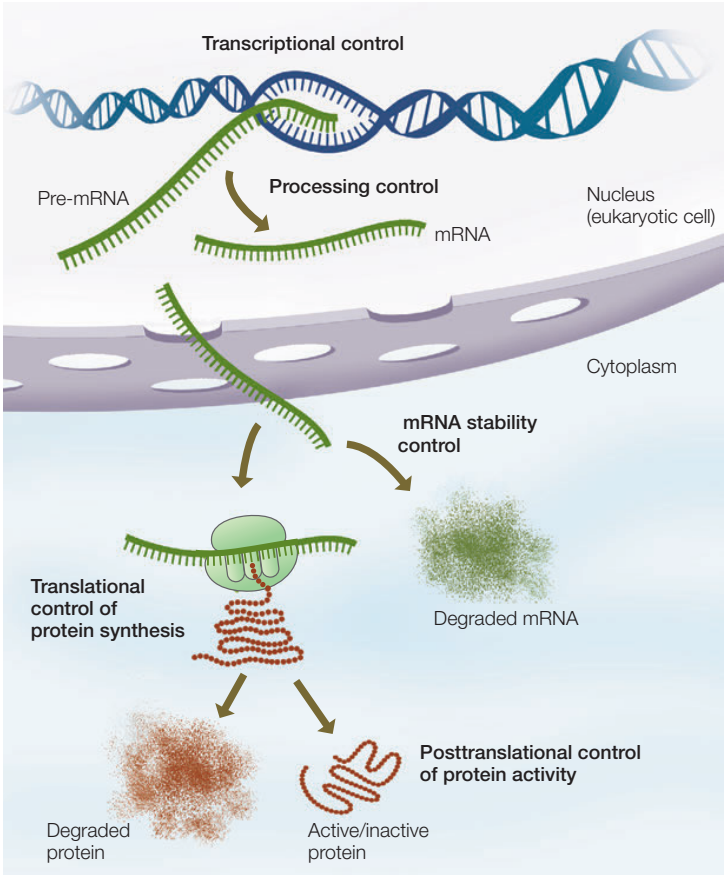


FIGURE 11.1 Potential Points for the Regulation of Gene Expression In a eukaryotic cell, gene expression can be regulated before transcription, during transcription, after transcription but before translation, at translation, or after translation.

Go to ACTIVITY 11.1 Eukaryotic Gene Expression Control Points
PoL2e.com/ac11.1

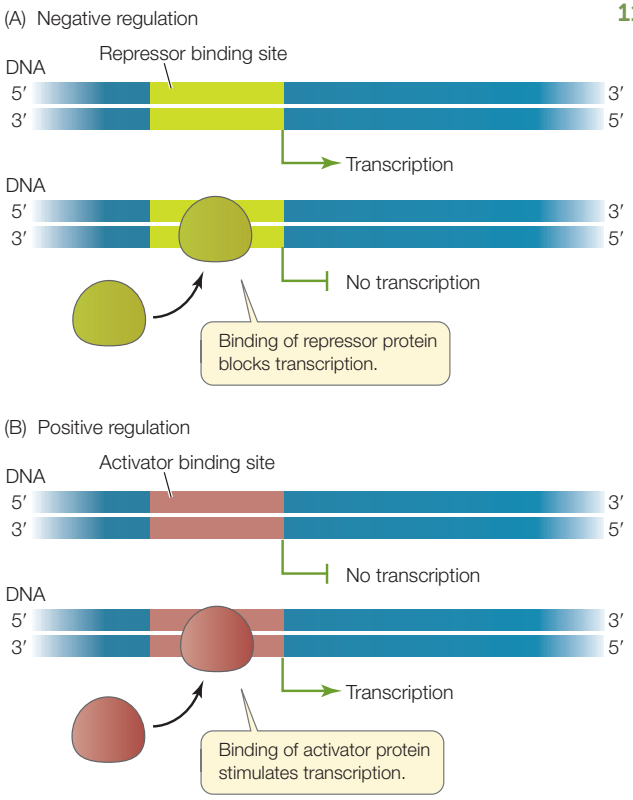
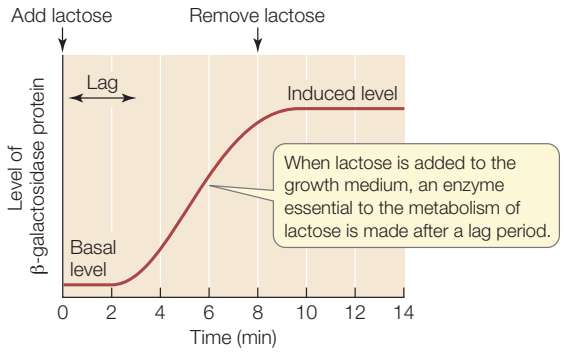


FIGURE 11.2 Positive and Negative Regulation Transcription factors regulate gene expression by binding to DNA and (A) repressing or (B) activating transcription by RNA polymerase.

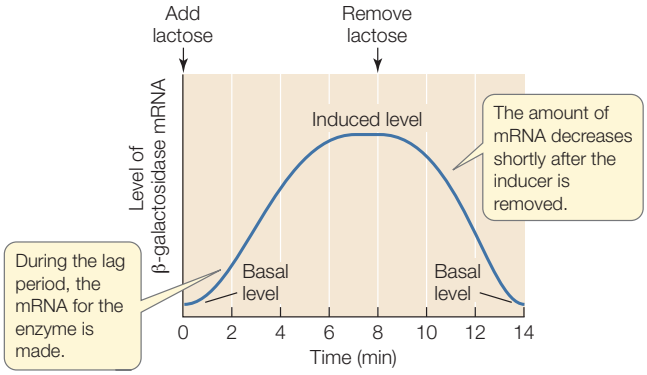
Lactose is a β-galactoside—a disaccharide containing galactose linked to glucose. Three proteins are involved in the initial uptake and metabolism of lactose by *E. coli*:

- β-galactoside permease is a carrier protein in the bacterial cell membrane that moves the sugar into the cell.
- β-galactosidase is an enzyme that hydrolyzes lactose to glucose and galactose.
- β-galactoside transacetylase transfers acetyl groups from acetyl CoA to certain β-galactosides. Its role in the metabolism of lactose is not clear.

When *E. coli* is grown on a medium that contains glucose but no lactose, the basal (uninduced) levels of these three proteins are extremely low—only a few molecules per cell. But if the cells are transferred to a medium with lactose as the predominant sugar, they promptly begin making all three proteins after a short lag period, and within 10 minutes there are about 3,000 of each of these proteins per cell (the induced level):



What causes this dramatic increase? A clue comes from measuring the concentration of mRNA for β-galactosidase. After lactose is added to the medium, the mRNA level increases *before* the level of β-galactosidase protein begins to rise:



The mRNA is produced during the lag phase and then is translated into protein. The high mRNA level depends on the presence of lactose, because if lactose is removed, the mRNA level goes down. The response of the bacteria to lactose is clearly at the level of transcription. (The level of β-galactosidase protein does not go down immediately after the inducer is removed because the protein is more stable than the mRNA.)

Compounds that stimulate the transcription of specific genes are called **inducers**, and genes that can be activated by inducers are called **inducible genes**. In contrast, some other genes are expressed most of the time at a constant rate; these are called **constitutive genes**. The lactose-metabolizing proteins in *E. coli* are encoded by inducible genes. When lactose first enters the cell, some of it is converted to a similar molecule called allolactose. Allolactose is the inducer that switches on the expression of the genes for the lactose-metabolizing proteins.

Operons are units of transcriptional regulation in prokaryotes

The genes that encode the three proteins for processing lactose in *E. coli* lie adjacent to one another on the *E. coli* chromosome. This arrangement—which is common for functionally related genes in prokaryotes—is no coincidence: the genes share a single promoter, and their DNA is transcribed into a single, continuous molecule of mRNA that contains the coding regions for the three proteins. Because this particular mRNA governs the synthesis of all three lactose-metabolizing enzymes, either all or none of these enzymes are made at any particular time.

A cluster of genes with a single promoter is called an **operon**, and the operon that encodes the three lactose-metabolizing enzymes in *E. coli* is called the *lac* operon. The *lac* operon promoter can be very efficient (the maximum rate of mRNA synthesis can be high), but its activity can be reduced when the enzymes are not needed. This example of transcriptional

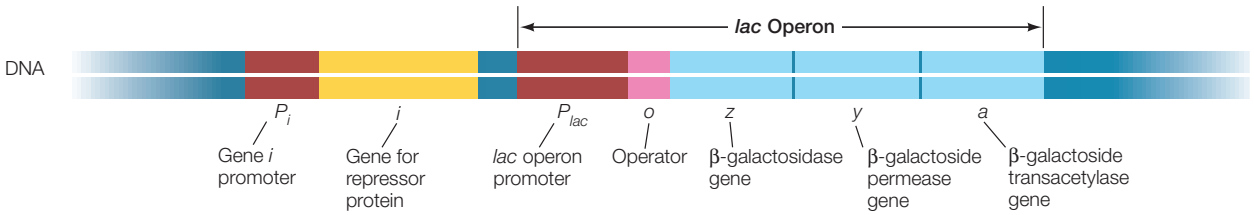


FIGURE 11.3 The *lac* Operon of *E. coli* The *lac* operon is a segment of DNA that includes a promoter, an operator, and the three genes that code for lactose-metabolizing enzymes. In reality, the coding sequences (genes) are much longer than the short regulatory sequences.

regulation, which we explore in more detail below, was worked out in the 1960s by Nobel Prize winners François Jacob and Jacques Monod.

Operator–repressor interactions regulate transcription in the *lac* and *trp* operons

The *lac* operon has a DNA sequence called an operator, which is near the promoter and controls transcription of the *lac* genes (FIGURE 11.3). An **operator** is a repressor-binding site that can bind very tightly with a repressor protein (see Figure 11.2A). Repressors play different roles in different operons:

- An inducible operon is turned off unless needed.
- A repressible operon is turned on unless not needed.

In the case of the inducible *lac* operon, a repressor protein prevents transcription until the *lac*-encoded proteins are needed. In contrast, the *trp* operon (described below) is a repressible operon that is turned off by a repressor only under particular circumstances.

***lac* OPERON** As we described above, the *lac* operon is not transcribed at high levels unless a β-galactoside such as lactose is the predominant sugar available in the cell's environment. A repressor protein normally binds to the operator, preventing RNA polymerase from binding and thereby blocking transcription. When lactose is present, the repressor detaches from the operator, allowing RNA polymerase to bind to the promoter and start transcribing the *lac* genes (FIGURE 11.4).

The key to this regulatory system is the repressor protein. Expressed from a constitutive gene (one that is always active), the repressor is always present in the cell in adequate amounts to occupy the operator and keep the operon turned off. The repressor has a recognition site for the DNA sequence in the operator, and it binds very tightly. However, it also has an allosteric binding site for the inducer. When the inducer (allolactose) binds to the repressor, the repressor changes shape so that it can no longer bind DNA.

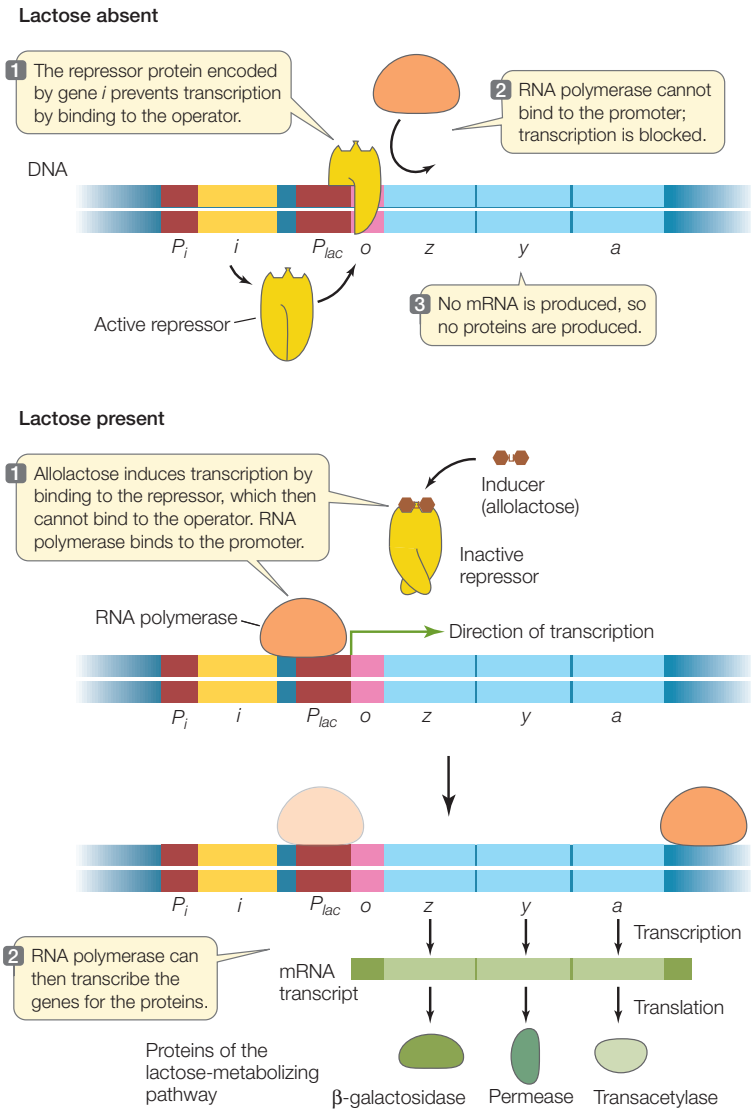


FIGURE 11.4 The *lac* Operon: An Inducible System Allolactose (the inducer) leads to synthesis of the proteins in the lactose-metabolizing pathway by binding to the repressor protein and preventing its binding to the operator.

Go to ANIMATED TUTORIAL 11.1
The *lac* Operon
PoL2e.com/at11.1

The gene for the lac repressor (gene *i* in Figure 11.3) is located upstream of the *lac* operon on the *E. coli* chromosome. The *lac i* gene is referred to as a regulatory gene because it encodes a regulatory protein (a transcription factor). In contrast, a **structural gene** is any gene that encodes a protein that is not directly involved in gene regulation. The three genes that encode the lactose-metabolizing enzymes are structural genes.

***trp* OPERON** Like an inducible operon, a repressible operon is switched off when its repressor is bound to its operator. However, in this case the repressor binds to the DNA only in the presence of a **corepressor**. The corepressor is a molecule that binds to the repressor, causing it to change shape and bind to the operator, thereby inhibiting transcription. An example is the operon whose structural genes catalyze the synthesis of the amino acid tryptophan:

$$\text{Precursor molecules} \xrightarrow{\text{Five enzyme-catalyzed reactions}} \text{Tryptophan}$$

When tryptophan is present in adequate concentrations, it is energy-efficient for the cell to stop making the enzymes for tryptophan synthesis. Therefore tryptophan itself functions as a corepressor: tryptophan binds to the repressor of the *trp* operon, causing the repressor to bind to the *trp* operator to prevent transcription of the enzymes in the pathway (**FIGURE 11.5**).

To summarize the differences between these two regulatory systems:

- In *inducible* systems, the substrate of a metabolic pathway (the inducer) interacts with a transcription factor (the repressor), rendering the repressor incapable of binding to the operator and thus allowing transcription.
- In *repressible* systems, a product of a metabolic pathway (the corepressor) binds to the repressor protein, which is then able to bind to the operator and block transcription.

In general, inducible systems control catabolic pathways (which are turned on only when the substrate is available), whereas repressible systems control anabolic pathways (which are turned on until the concentration of the product becomes sufficient).

LINK

You can review catabolic and anabolic reactions in **Concept 2.5**

In both of the systems described above, the regulatory protein is a repressor that functions by binding to the operator. Transcription in prokaryotes can also be regulated by activator proteins that bind to DNA sequences at or near the promoter and promote transcription (see Figure 11.2). Like repressors, activators can regulate both inducible and repressible systems. Furthermore, many genes and operons are controlled by more than one regulatory mechanism. We will discuss transcription factors in more detail in Concept 11.2.

We have now seen two basic systems for regulating a metabolic pathway. In Concept 3.4 we described the allosteric

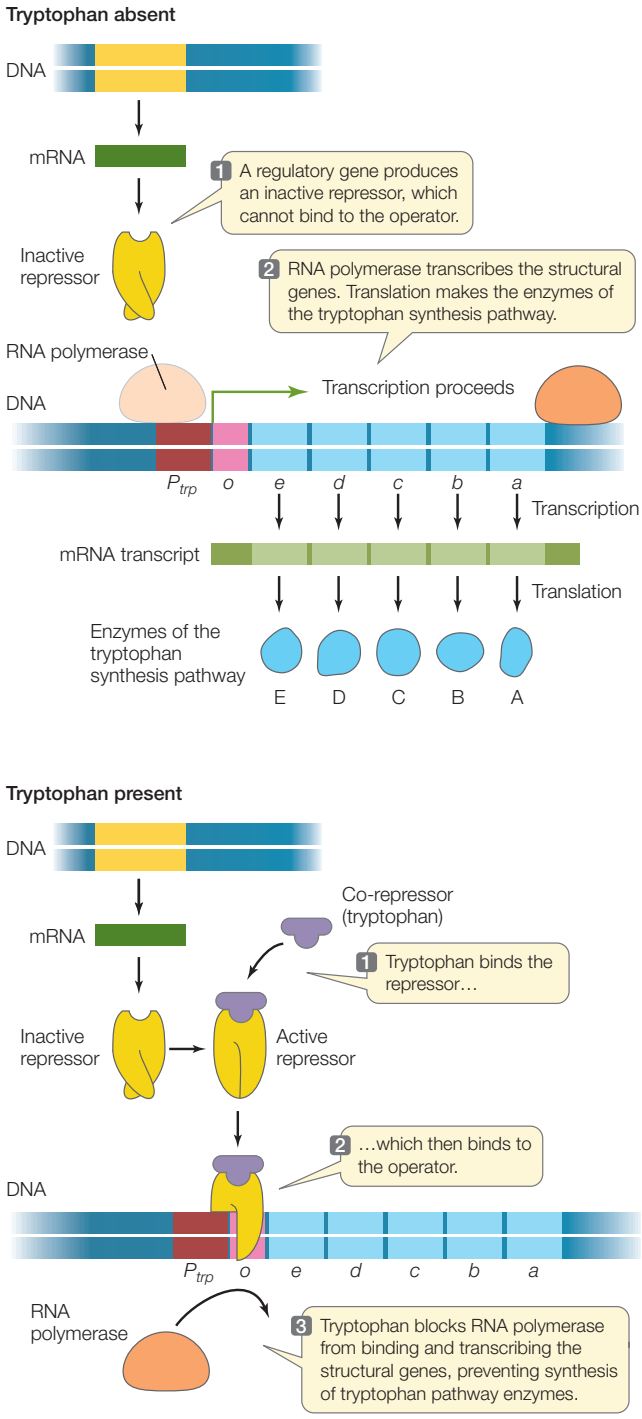


FIGURE 11.5 The *trp* Operon: A Repressible System Because tryptophan activates an otherwise inactive repressor, it is called a corepressor.

Go to ANIMATED TUTORIAL 11.2
The *trp* Operon
PoL2e.com/at11.2

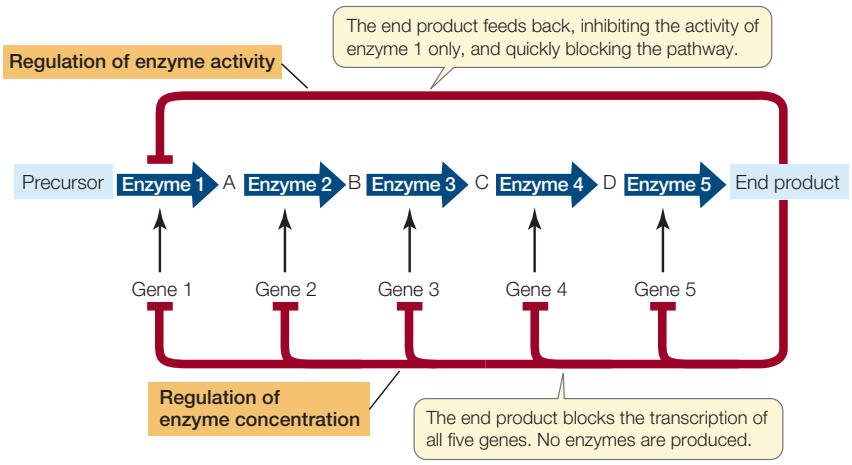


FIGURE 11.6 Systems to Regulate a Metabolic Pathway Feedback from the end product of a metabolic pathway can block enzyme activity (allosteric regulation), or it can stop the transcription of genes that code for the enzymes in the pathway (transcriptional regulation).

regulation of enzyme activity—a mechanism that allows rapid fine-tuning of metabolism. The regulation of transcription is slower but results in greater savings of energy and resources. Protein synthesis is a highly endergonic process; synthesizing mRNA, charging tRNA, and moving the ribosomes along mRNA all require large amounts of energy. FIGURE 11.6 compares allosteric and transcriptional regulation.

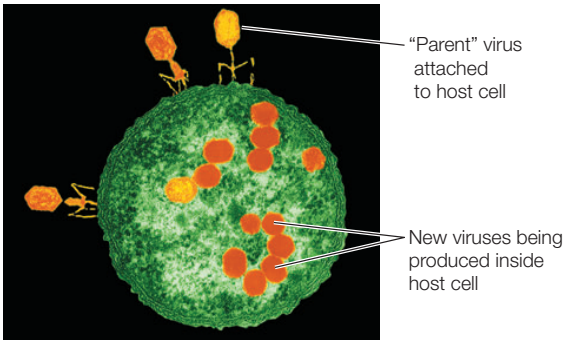
RNA polymerase can be directed to a class of promoters

As noted above and in Chapter 10, RNA polymerase binds to specific DNA sequences at the promoter to initiate transcription. We have just described how repressor proteins can physically block RNA polymerase binding. However, there are other proteins in prokaryotes called **sigma factors** that can bind to RNA polymerase and direct the polymerase to specific promoters.

Genes that encode proteins with related functions may be at different locations in the genome but have the same promoter sequence. This allows them to be expressed at the same time and under the same physiological conditions. For example, some bacteria stop growing when nutrients in their environment are depleted. When this happens, they adopt an alternative lifestyle called sporulation—they reduce their metabolic activity and form a tough spore coat (see Concept 19.2). This process involves the sequential expression of specific classes of genes. Each member of a gene class has a common promoter sequence, and RNA polymerase is directed to the promoter in each case by a specific sigma factor. As we will see in Concept 11.2, this form of global gene regulation by proteins binding to RNA polymerase is also common in eukaryotes.

Viruses use gene regulation strategies to hijack host cells

The immunologist Sir Peter Medawar once described a virus as "a piece of bad news wrapped in protein." As we described in Concept 9.1, a **virus** injects its genetic material into a host cell, and in many cases it turns that cell into a virus factory:



This involves a radical change in gene expression for the host cell, and can result in the death of the cell when new viral particles are released.

Viruses are not cells and do not carry out many of the processes characteristic of life. They are dependent on living

APPLY THE CONCEPT

Many prokaryotic genes are regulated in operons, which include regulatory DNA sequences

Genetic mutations are useful in analyzing the control of gene expression. In the *lac* operon of *E. coli* (see Figures 11.3 and 11.4), gene *i* codes for the repressor protein, P_{lac} is the promoter, *o* is the operator, and *z* is the first structural gene. The superscript "+" designates the wild type; superscript "–" means mutant. Fill in the table, describing the level of transcription in different genetic and environmental conditions. (The first line of the table has been filled in as an example.)

	z TRANSCRIPTION LEVEL	
GENOTYPE	LACTOSE PRESENT	LACTOSE ABSENT
$i^+ P_{lac}{}^+ o^+ z^+$ (wild type)	High	Low
$i^- P_{lac}{}^+ o^+ z^+$		
$i^+ P_{lac}{}^+ o^+ z^-$		
$i^+ P_{lac}{}^- o^+ z^+$		
$i^+ P_{lac}{}^+ o^- z^+$		

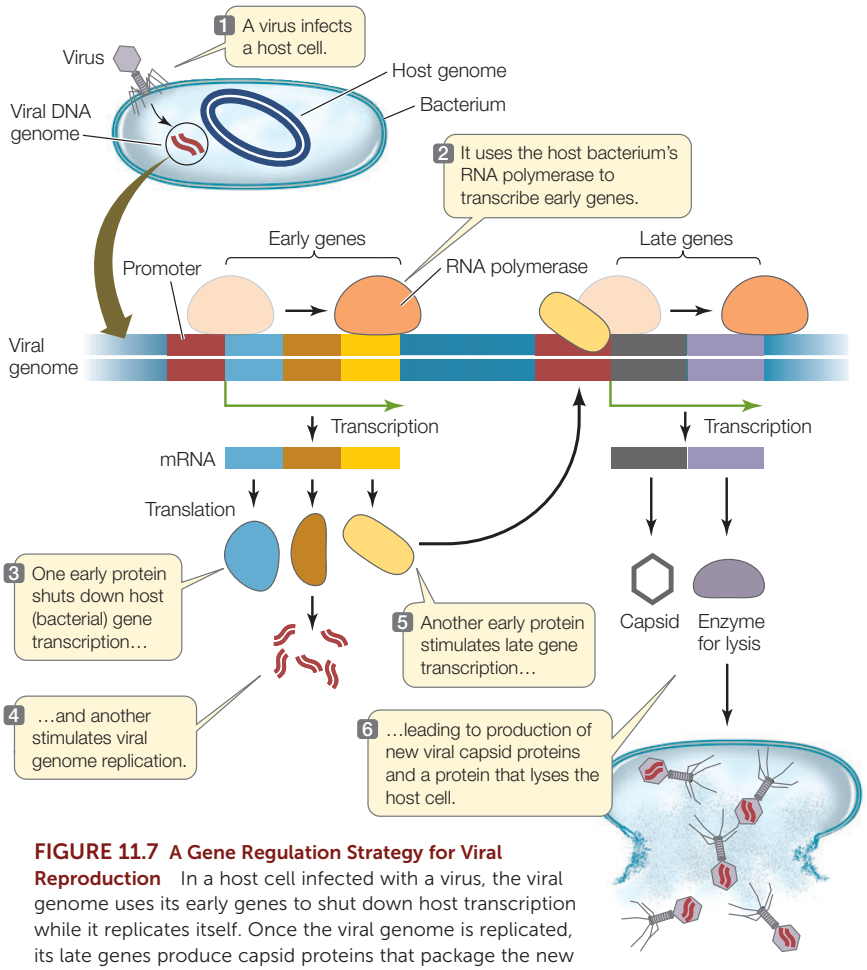


FIGURE 11.7 A Gene Regulation Strategy for Viral Reproduction In a host cell infected with a virus, the viral genome uses its early genes to shut down host transcription while it replicates itself. Once the viral genome is replicated, its late genes produce capsid proteins that package the new genomes, and other proteins that lyse the host cell.

FIGURE 11.7 illustrates molecular events in the lytic life cycle of T4, a typical double-stranded DNA bacteriophage (phage, or bacterial virus). At the molecular level, the lytic cycle has two stages, early and late:

- The viral genome contains a promoter that binds host RNA polymerase. In the early stage, viral genes that lie adjacent to this promoter are transcribed. These early genes encode proteins that shut down expression of host genes, stimulate viral genome replication, and activate the transcription of viral late genes. The host genes are shut down by a posttranscriptional mechanism: a virus-encoded enzyme degrades the host RNA before it can be translated. Another viral nuclease digests the host's chromosome, providing nucleotides for the synthesis of many copies of the viral genome. These processes can occur within a few minutes after the virus first infects the cell.
- In the late stage, viral late genes are transcribed; they encode the viral capsid proteins and enzymes that lyse the host cell to release the new virions.

Under ideal conditions, this entire process—from binding and infection to release of new phage—can be completed in only half an hour.

cells to reproduce. Unlike living cells, not all viruses use double-stranded DNA as the genetic material that is contained within the viral particle and transmitted from one generation to the next. The viral genome may consist of double-stranded DNA, single-stranded DNA, or double- or single-stranded RNA. But whether the genetic material is DNA or RNA, the viral genome takes over the host's protein synthetic machinery within minutes of entering the cell.

Typically, the host cell immediately begins to produce new viral particles (virions), which are released as the cell breaks open, or lyses. This type of prokaryotic viral life cycle is called **lytic**. Some viral life cycles also include a **lysogenic** or dormant phase. In this case the viral genome becomes incorporated into the host cell genome and is replicated along with the host genome. The virus may survive in this way for many host cell generations. Sooner or later, an environmental signal can cause the host cell to begin producing virions—at which point the viral reproductive cycle enters the lytic phase.

LINK

The different types of viruses are described in **Concept 19.4**

CHECKpoint CONCEPT 11.1

- ✓ What is the difference between positive and negative regulation of gene expression?
- ✓ Describe the molecular conditions at the *lac* operon promoter in the presence and absence of lactose.
- ✓ Describe the molecular events at the *trp* operon promoter in the presence and absence of tryptophan.
- ✓ If the *lac* repressor gene were mutated so that the allosteric site on the encoded protein no longer bound allolactose, what would be the effect on transcription of the *lac* operon? What about a similar mutation in the *trp* repressor gene?
- ✓ What would be the effect of a mutation in the gene that encodes RNA polymerase so that it does not bind to the late gene promoter of bacteriophage T4?

Studies of bacteria and bacteriophage provide a basic understanding of the mechanisms that regulate gene expression and of the roles of regulatory proteins in both positive and negative regulation. We will now turn to the control of gene expression in eukaryotes. You will see both negative and positive control of transcription, as well as posttranscriptional mechanisms of regulation.

CONCEPT 11.2 Eukaryotic Genes Are Regulated by Transcription Factors

As we mentioned in Concept 11.1, gene expression can be regulated at several different points in the process of transcribing a gene and translating the mRNA into a protein (see Figure 11.1). In this concept we will describe the mechanisms that result in the selective transcription of specific eukaryotic genes. As in prokaryotes, eukaryotic cells must precisely regulate the expression of their genes. Some genes are constitutive (expressed in most tissues most of the time), whereas others are inducible and expressed only when needed. This is especially important in multicellular organisms with specialized cells and tissues. For example, virtually all of our cells carry the genes encoding keratin (the protein in our hair and nails) and hemoglobin. Yet keratin is made only by epithelial cells such as skin cells, and hemoglobin is made only by developing red blood cells. In contrast, all human cells express the genes that encode enzymes needed for basic metabolic activities (such as glycolysis), and all cells must synthesize certain structural proteins such as actin (a component of the cytoskeleton).

The mechanisms for regulating transcription in eukaryotes are similar conceptually to those of prokaryotes. Both types of cells use DNA–protein interactions to bring about negative and positive control of gene expression. However, there are significant differences, which generally reflect the greater complexity of eukaryotic organisms (TABLE 11.1).

Transcription factors act at eukaryotic promoters

As in bacteria, a eukaryotic promoter is a region of DNA near the 5′ end of a gene where RNA polymerase binds and initiates transcription. Eukaryotic promoters are extremely diverse and difficult to characterize, but they each contain a core promoter sequence to which the RNA polymerase binds. The most common of these is the **TATA box**—so called because it is rich in A-T base pairs.

RNA polymerase II is the polymerase that transcribes the protein-coding genes in eukaryotes. It cannot bind to the promoter and initiate transcription by itself. Rather, it does so only after various **general transcription factors** have bound to the core promoter. General transcription factors bind to most promoters and are distinct from transcription factors that have specific regulatory effects only at certain promoters or classes of promoters. FIGURE 11.8 illustrates the assembly of the resulting transcription complex at a promoter containing a TATA box. First, the protein TFIID ("TF" stands for transcription factor) binds to the TATA box. Binding of TFIID changes both its own shape and that of the DNA, presenting a new surface that attracts the binding of other transcription factors. RNA polymerase II binds only after several other proteins have bound to the complex.

The core promoter sequence is bound by general transcription factors that are needed for the expression of all RNA polymerase II–transcribed genes. Other sequences that are (usually) found in or near promoter regions are specific to only a few genes and are recognized by specific transcription factors. These transcription factors may be positive

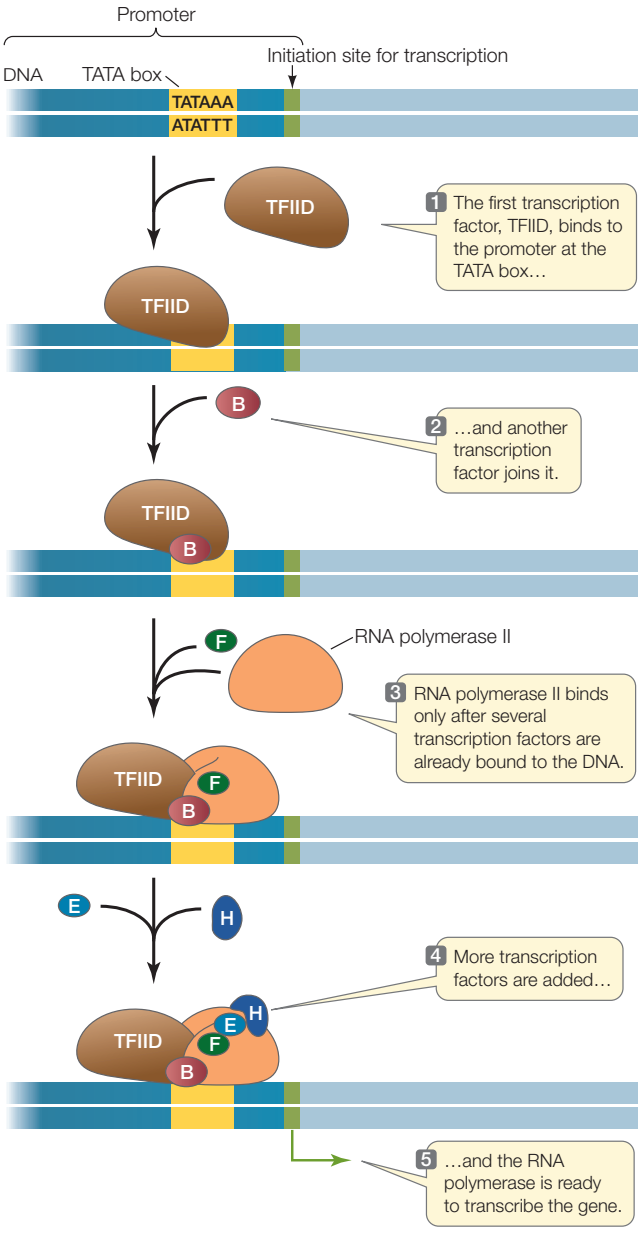


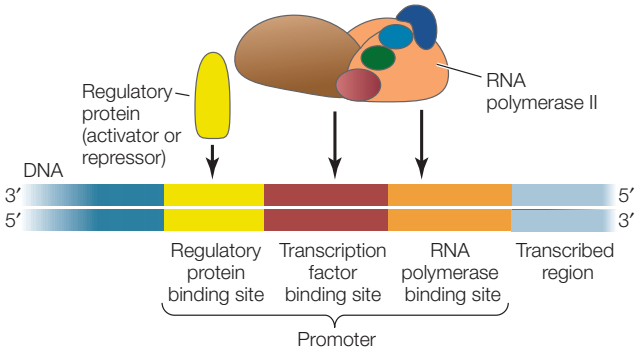
FIGURE 11.8 The Initiation of Transcription in Eukaryotes Apart from TFIID, which binds to the TATA box, each transcription factor in this transcription complex has binding sites only for the other proteins in the complex, and does not bind directly to DNA. B, E, F, and H are general transcription factors.

Go to ANIMATED TUTORIAL 11.3
Initiation of Transcription
PoL2e.com/at11.3

TABLE 11.1 Transcription in Bacteria and Eukaryotes

Characteristic	Bacteria	Eukaryotes
Locations of functionally related genes	Often clustered in operons	Often distant from one another with separate promoters
RNA polymerases	One	Three:
		I: transcribes rRNA
		II: transcribes mRNA
		III: transcribes tRNA and small RNAs
Promoters and other regulatory sequences	Few	Many
Initiation of transcription	Binding of RNA polymerase	Binding of many proteins, including RNA polymerase, to promoter

regulators (activators) or negative regulators (repressors) of transcription:



Such transcription factors may be present only in certain cell types, or they may be present in all cells but activated by specific signals. DNA sequences that bind activators are called enhancers, and those that bind repressors are called silencers. Some enhancers and silencers occur near the core promoter, and others can be as far as 20,000 base pairs away. When the activators or repressors bind to these DNA sequences, they interact with the RNA polymerase complex, causing the DNA to bend. Often many such binding proteins are involved, and the *combination* of factors present determines whether transcription is initiated. With about 2,000 different transcription factors in humans, there are many possibilities for regulation.

How do transcription factors recognize a specific nucleotide sequence in DNA? To answer this question, let's look at a specific example. NFATs (*n*uclear *f*actors of *a*ctivated *T* cells) are a group of transcription factors that control the expression of genes essential for the immune response (see Chapter 39). NFAT proteins bind to a 12-bp recognition sequence near the promoters of these genes, with the sequence CGAGGAAAATTG (**FIGURE 11.9**). Recall that there are atoms in the bases of DNA that are available for hydrogen bonding but are not involved in base pairing (see Figure 9.6). These atoms are important in the interactions between an NFAT and the DNA. In addition, there are hydrophobic interactions between the rings in the DNA bases and some amino acid R groups in the protein. As for an enzyme and its substrate (see Concept 3.3), there is an induced fit between the NFAT and the DNA, such that the protein undergoes a conformational change after binding begins.

The base sequence of a binding site on DNA determines the arrangement of chemical groups available for hydrogen bonding and hydrophobic interactions with DNA-binding proteins; this is the basis of the specificity of DNA–protein interactions.

The expression of transcription factors underlies cell differentiation

During the development of a complex organism from fertilized egg to adult, cells become more and more differentiated (specialized). Differentiation is brought about in many cases by changes in gene expression, resulting from the activation (and inactivation) of transcription factors. We will discuss this topic in more detail in Chapter 14. For now, remember that virtually all differentiated cells contain the entire genome, and that their specific characteristics arise from differential gene expression.

Currently there is great interest in cellular therapy: providing new, functional cells to patients who have diseases that involve the degeneration of certain cell types. An example is Alzheimer's disease, which involves the degeneration of neurons in the brain. Because of the possibility of immune system rejection (see Chapter 39), it would be optimal if patients could receive their own cells, modified in some way to be functional. Since specialized functions are under the control of transcription factors, turning readily available cells into a particular

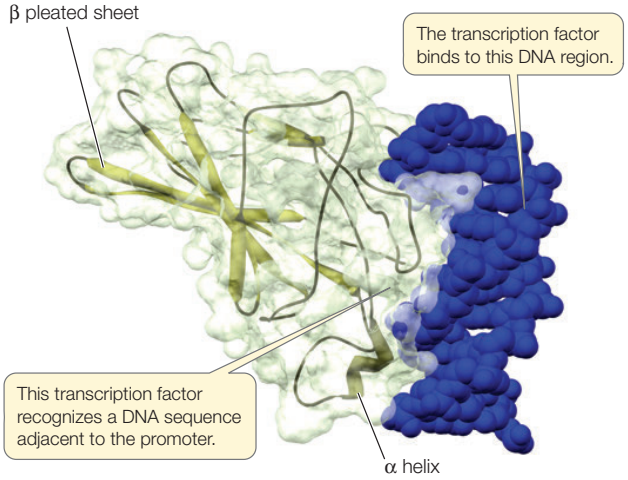


FIGURE 11.9 A Transcription Factor Protein Binds to DNA The transcription factor NFAT activates genes for the immune response by binding to a specific DNA sequence near the promoters of those genes.

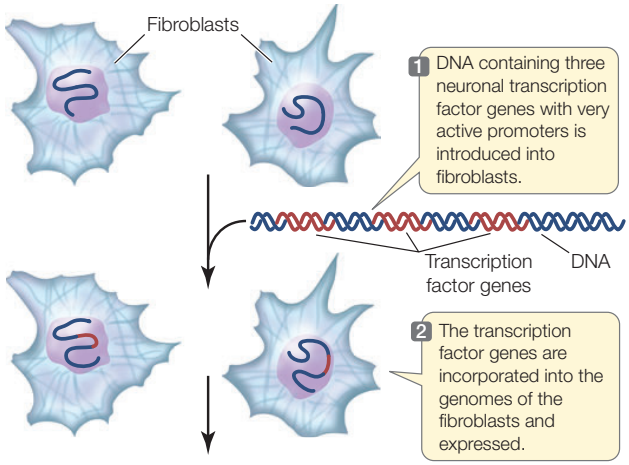
INVESTIGATION

FIGURE 11.10 Expression of Specific Transcription Factors Turns Fibroblasts into Neurons Fibroblasts are cells that secrete abundant extracellular matrix and contribute to the structural integrity of organs. Neurons are highly specialized cells in the nervous system. Marius Wernig and his colleagues performed a series of experiments to find out whether expressing neuronal transcription factors in fibroblasts would be sufficient to cause the fibroblasts to become neurons.[a]

HYPOTHESIS

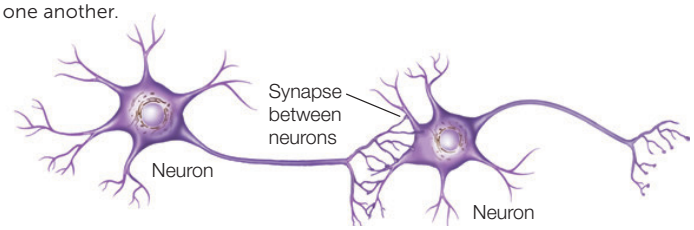
Expression of neuron-specific transcription factors in fibroblasts will turn the latter into neurons.

METHOD



RESULTS

After 6 days, the fibroblasts develop into functional neurons, which form characteristic synapses with one another.



CONCLUSION

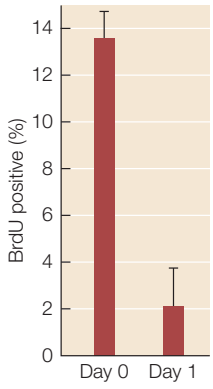
The expression of just three transcription factors is sufficient to transform a fibroblast into a neuron.

ANALYZE THE DATA

Fibroblasts are active in cell division; neurons are not. In addition to morphology, the lack of cell division was used as a criterion to show that the transformed cells were neurons. The rate of cell division in the transformed cells was measured by the incorporation of the labeled nucleotide bromodeoxyuridine (BrdU) into DNA. The percentage of labeled (hence dividing) cells is shown in the graph.

A. Was cell division stopped in the transformed cells? Explain your answer.

B. The error bars are standard deviations. What statistical test would you use to show whether the difference between the two cell populations was significant? See Appendix B for a Statistics Primer.



Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]T. Vierbuchen et al. 2010. *Nature* 463: 1035–1041.

desired cell type might be achieved by altering transcription factor expression. Marius Wernig and his colleagues at Stanford University have made important progress toward this goal (**FIGURE 11.10**). They took skin fibroblasts from mice and manipulated the expression of transcription factors in the cells to change them into neurons. By repeating their experiments on human fibroblasts, they have brought cellular therapy closer to reality.

LINK

The basis for rejection of nonself cells is described in **Concept 39.5**. Additional approaches used in cellular therapy are discussed in **Concept 14.1**

Transcription factors can coordinate the expression of sets of genes

We have seen that prokaryotes can coordinate the regulation of several genes by arranging them in an operon. In addition, bacteria can coordinate the expression of groups of genes using sigma factors, which guide RNA polymerase to particular classes of promoters. This latter mechanism is also used in eukaryotes to coordinately regulate genes that may be far apart, even on different chromosomes. The expression of genes can be coordinated if they share regulatory sequences that bind the same transcription factors.

This type of coordination is used by organisms to respond to stress—for example, by plants in response to drought. Under conditions of drought stress, a plant must simultaneously synthesize numerous proteins whose genes are scattered throughout the genome. The synthesis of these proteins comprises the stress response. To coordinate expression, each of these genes has a specific regulatory sequence near its promoter called the dehydration response element (DRE). In response to drought, a transcription factor changes so that it binds to this element and stimulates mRNA synthesis (**FIGURE 11.11**). The dehydration response proteins not only help the plant conserve water but also protect the plant against freezing or excess salt in the soil. This finding has considerable importance for agriculture because crops are often grown under less than optimal conditions.

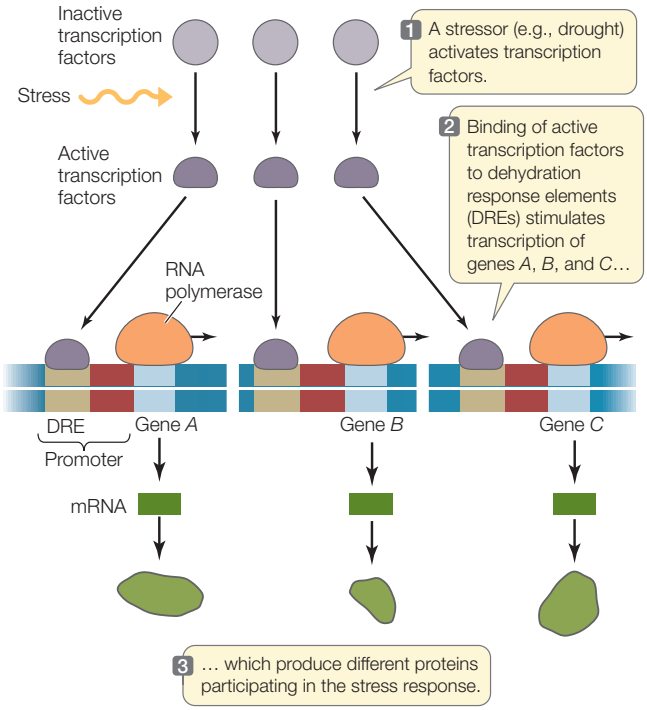


FIGURE 11.11 Coordinating Gene Expression A single environmental signal, such as drought stress, activates a transcription factor that acts on many genes.

LINK

Dehydration response proteins are one of many adaptations to drought stress found among the plants; see **Concept 28.3**

Eukaryotic viruses can have complex life cycles

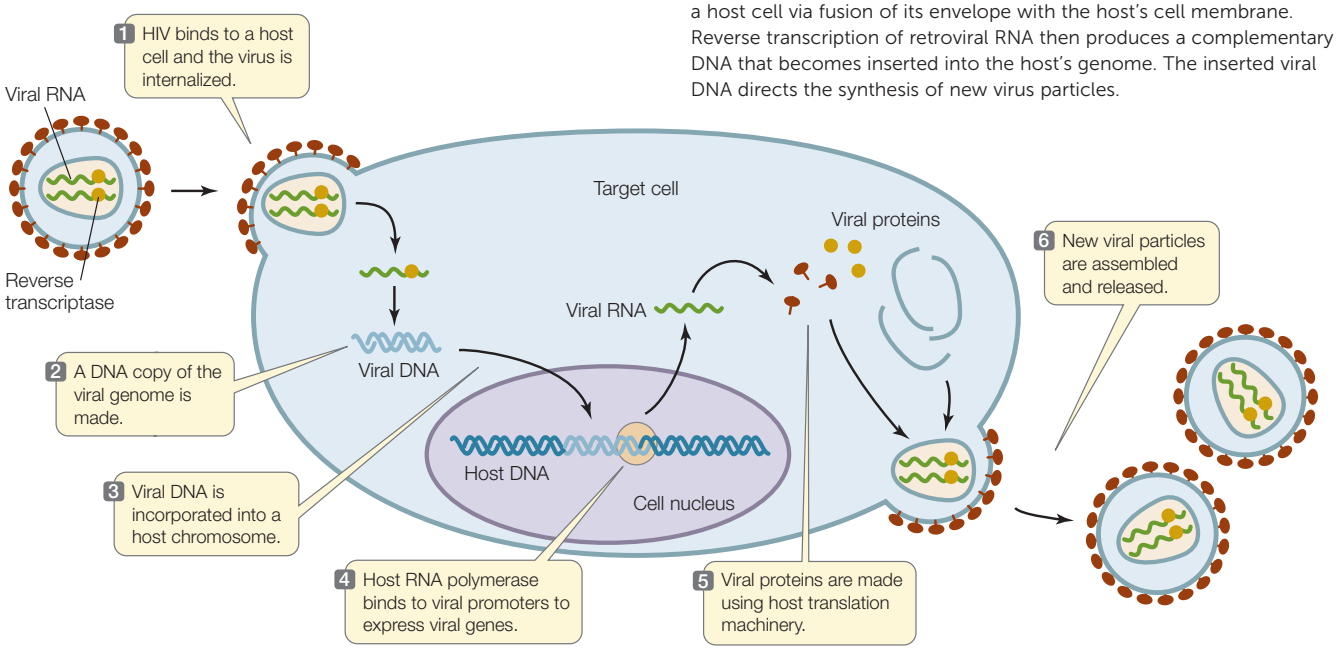
Eukaryotes are susceptible to infections by various kinds of viruses that have a variety of life cycle strategies. These viral life cycles can be very efficient—for example, the poliovirus completes its life cycle (from infection to release of new particles) in 4–6 hours, and each dying host cell can release up to 10,000 new particles. Compare this with the 24-hour cell cycle typical of dividing human cells.

The viruses that infect eukaryotic cells may have genomes of single- or double-stranded DNA or RNA. Some viral life cycles can be quite complex. As an example, we focus here on **human immunodeficiency virus** (**HIV**), the infective agent that causes acquired immunodeficiency syndrome (AIDS) in humans. HIV typically infects only cells of the immune system that express a surface receptor called CD4. The virion is enclosed within a phospholipid membrane derived from its previous host cell. Proteins in the membrane are involved in the infection of new host cells, which HIV enters by direct fusion of the viral envelope with the host's cell membrane (**FIGURE 11.12**).

HIV is a retrovirus: its genome is single-stranded RNA, and it carries within the virion an enzyme called reverse transcriptase. Shortly after infection, the **reverse transcriptase** makes a DNA strand that is complementary to the RNA, while at the same time degrading the RNA and making a second DNA strand that is complementary to the first. The resulting double-stranded DNA becomes integrated into the host's chromosome. The integrated viral DNA is called a provirus.

The provirus resides permanently in the host chromosome and can remain in an inactive state for years. During this time transcription of the viral DNA is initiated, but host cell proteins called termination factors prevent the RNA from elongating, and transcription is terminated prematurely (**FIGURE 11.13A**). Under

FIGURE 11.12 The Reproductive Cycle of HIV This retrovirus enters a host cell via fusion of its envelope with the host's cell membrane. Reverse transcription of retroviral RNA then produces a complementary DNA that becomes inserted into the host's genome. The inserted viral DNA directs the synthesis of new virus particles.



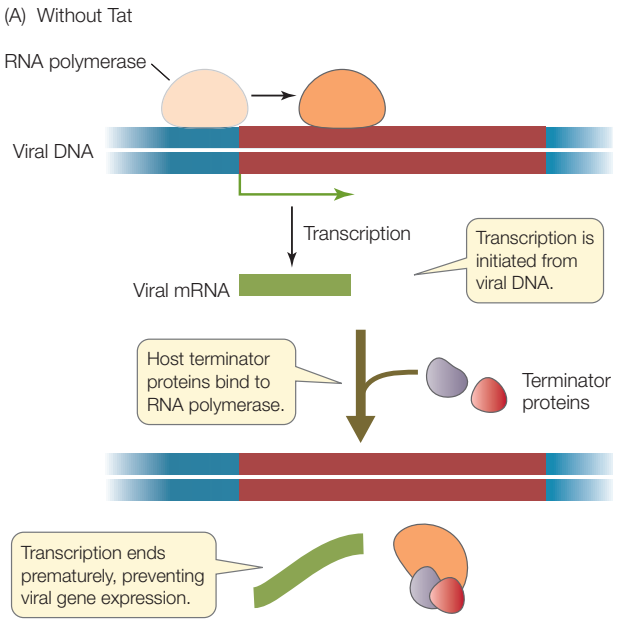


FIGURE 11.13 Regulation of Transcription by HIV The Tat protein acts as an antiterminator, allowing transcription of the HIV genome.

some circumstances, the level of transcription initiation increases and some viral RNA is made. One of the viral genes encodes a protein called Tat (*Trans*activator of *t*ranscription), which binds to the 5′ end of the viral RNA. As a result of Tat binding, the production of full-length viral RNA is dramatically increased (**FIGURE 11.13B**), and the rest of the viral reproductive cycle is able to proceed. It was only after the discovery of this mechanism in HIV and similar viruses that researchers found that many eukaryotic genes are regulated at the level of transcription elongation.

CHECKpoint CONCEPT 11.2

- ✓ How do transcription factors regulate gene expression? How do the roles of transcription factors compare with the roles of proteins that regulate prokaryotic operons?
- ✓ What would be the effect of the inhibition of reverse transcriptase on infection of a cell with HIV?

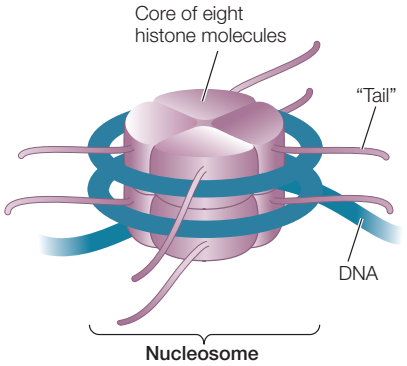
We have discussed some of the mechanisms that cells and viruses use to control gene transcription. These mechanisms involve the interaction of regulatory proteins with specific DNA sequences. However, eukaryotes have other mechanisms for controlling gene expression that do not depend on specific DNA sequences. We will discuss these mechanisms in the next concept.

CONCEPT 11.3 Gene Expression Can Be Regulated via Epigenetic Changes to Chromatin

So far we have focused on regulatory events that involve specific DNA sequences at or near a gene's promoter. Eukaryotic cells are also able to regulate transcription via reversible, non-sequence-specific alterations to either the DNA or the chromosomal proteins that package the DNA in the nucleus. These alterations can be passed on to daughter cells after mitosis or meiosis. They are called **epigenetic** changes to distinguish them from mutations, which involve irreversible changes to the DNA's base sequence (see Concept 9.3).

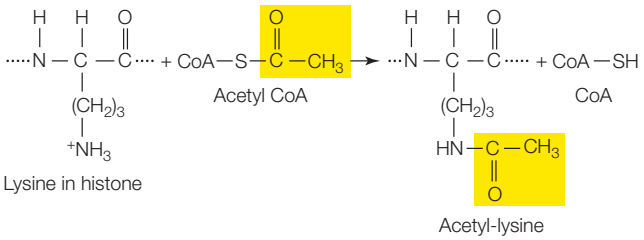
Modification of histone proteins affects chromatin structure and transcription

Epigenetic gene regulation can occur via the alteration of chromatin structure, or **chromatin remodeling**. Large amounts of DNA (nearly 2 meters in humans!) are packed within the nucleus (which has a diameter of about 5 μm). The basic unit of DNA packaging in eukaryotes is the nucleosome, a core of positively charged **histone** proteins around which DNA is wound:



Each histone protein has a "tail" of approximately 20 amino acids at its N terminus that sticks out of the compact structure and contains certain positively charged amino acids, notably lysine. Ordinarily there is strong ionic attraction between the positively charged histone proteins and DNA, which is negatively charged because of its phosphate groups. Because of this attraction, nucleosomes can make DNA physically inaccessible to RNA polymerase and the rest of the transcription apparatus. However, a variety of transcription factors and

RNA polymerase can bind enzymes called **histone acetyltransferases**, which add acetyl groups to these positively charged amino acids and neutralize their charges:



Reducing positive charges on the histone tails reduces the affinity of the histones for the DNA, loosening the compact nucleosome (FIGURE 11.14). The majority of histone acetylation is found near gene promoters, but acetylated histones are also found throughout the transcribed regions of genes. Thus histone acetylation promotes both transcription initiation and elongation. Histone acetylases can be recruited to promoters by transcription factors. An example is CREB, the transcription factor that is associated with addiction (see the opening story of this chapter). CREB binds specific acetyltransferases that participate in the activation of CREB-responsive genes.

Another class of chromatin remodeling proteins, **histone deacetylases**, can remove the acetyl groups from histones and thereby *repress* transcription. Histones can also be modified in other ways, including methylation (the addition of a methyl group) and phosphorylation (the addition of a phosphate group). Histone methylation can contribute to either the activation or repression of gene expression, depending on which lysine residue is methylated. Histone phosphorylation is involved in chromosome condensation during mitosis and meiosis, as well as affecting gene regulation. All of these effects are reversible, and so the transcriptional activity of a eukaryotic gene may be determined by varying patterns of histone modification.

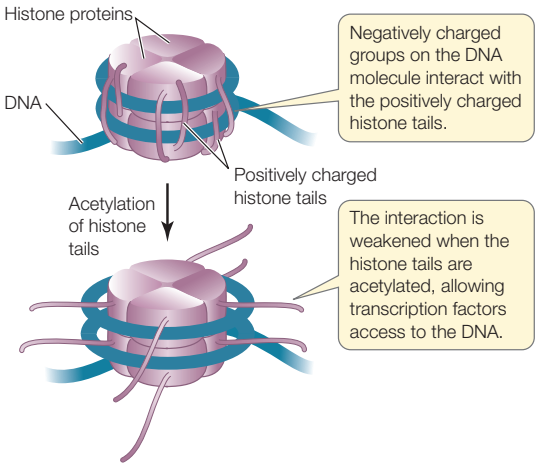
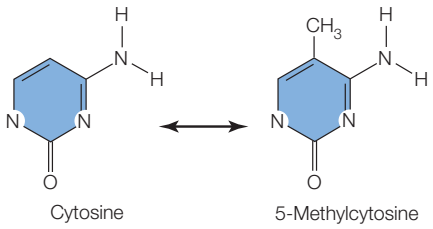


FIGURE 11.14 Epigenetic Remodeling of Chromatin for Transcription Initiation of transcription requires that nucleosomes change their structure, becoming less compact. This chromatin remodeling makes DNA accessible to the transcription complex (see Figure 11.8).

DNA methylation affects transcription

Depending on the organism, from 1 to 5 percent of cytosines in the DNA are chemically modified by the addition of a methyl group (—CH_3), to form 5-methylcytosine (FIGURE 11.15). This



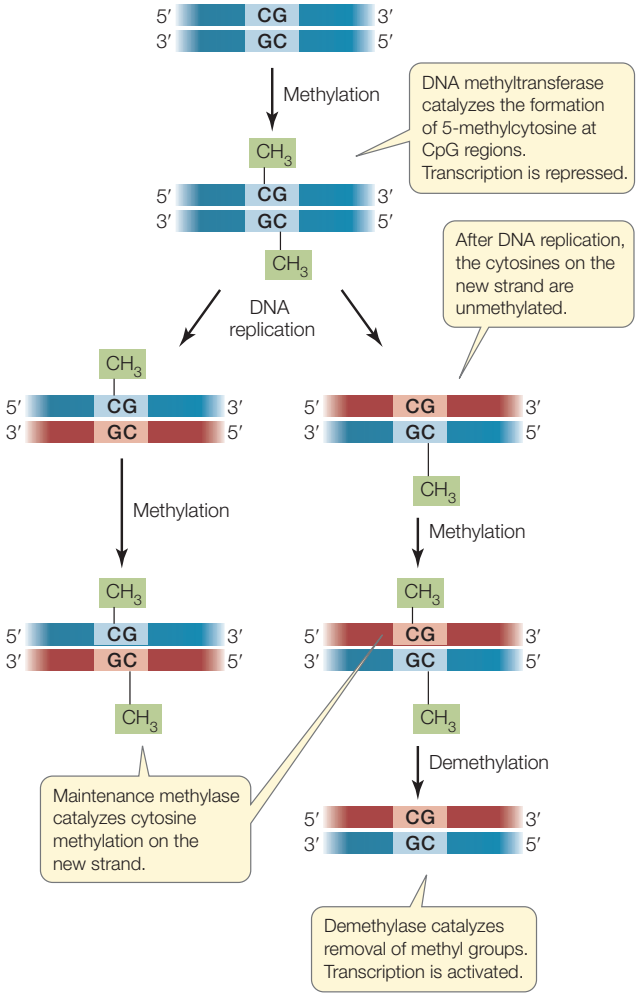


FIGURE 11.15 DNA Methylation: An Epigenetic Change The reversible formation of 5-methylcytosine in DNA can alter the rate of transcription.

covalent addition is catalyzed by the enzyme **DNA methyltransferase** and, in mammals, usually occurs on cytosines (C) that are adjacent to guanines (G). Virtually all of these CpG ("p" is for the phosphate in the DNA backbone) sites are methylated, apart from those found in and near transcriptionally active promoters. Promoters usually contain regions of DNA that are rich in CpG sites, and such regions are called **CpG islands**.

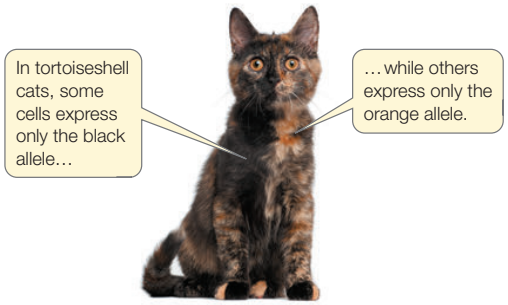
Methylated DNA binds specific proteins that are involved in the repression of transcription; thus heavily methylated genes tend to be inactive (silenced). Whereas histone acetylation/deacetylation (see above) are dynamic processes resulting in short-term changes in gene expression, DNA methylation is usually a stable, long-term silencing mechanism. When DNA is replicated, a **maintenance methyltransferase** catalyzes the formation of 5-methylcytosine in the new DNA strands. However, the pattern of cytosine methylation can also be altered, because methylation is reversible: a third enzyme, appropriately called **demethylase**, catalyzes the removal of the methyl group from cytosine (see Figure 11.15). In ways that are not fully understood, the enzymes involved in histone modification and DNA methylation/demethylation interact to ensure that genes whose products are needed in the cell are kept unmethylated, and their associated histones acetylated.

Sometimes, large stretches of DNA or almost entire chromosomes are methylated. Under a microscope, two kinds of chromatin can be distinguished in the stained interphase nucleus: euchromatin and heterochromatin. The **euchromatin** appears diffuse and stains lightly; it contains the DNA that is transcribed into mRNA. **Heterochromatin** is condensed and stains darkly; any genes it contains are generally not transcribed.

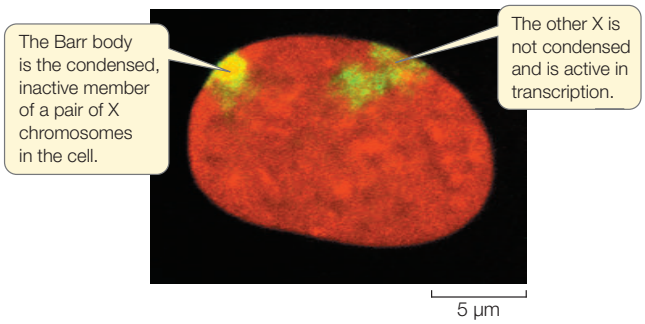
A dramatic example of heterochromatin is the X chromosome in female mammals. A normal female mammal has two X chromosomes, whereas a normal male has an X and a Y (see Concept 8.3). Because each female cell has two copies of each X chromosome gene, the female should have the potential to produce twice as much of each protein product as the male. Nevertheless, for 75 percent of the genes on the X chromosome, the total amount of mRNA produced is generally the same in males and in females. How does this happen?

In the early female embryo, one copy of X becomes heterochromatic and transcriptionally inactive in each cell, and the same X remains inactive in all of that cell's descendants. In a given female embryo cell, the "choice" of which X to inactivate is random. Recall that one X in a female comes from her father and one from her mother. Thus in one embryonic cell the paternal X might be inactivated, but in a neighboring cell the maternal X might be inactivated.

A familiar example of a phenotype caused by X chromosome inactivation is the tortoiseshell cat. In cats, two alleles of an X-linked gene that contributes to coat color are orange (X^B) and black (X^b). In the embryo of a heterozygous female (X^BX^b), there is random X-inactivation such that in some cells only the orange allele is expressed, while in others it is the black allele. These cells then form the progenitors of patches of skin in the animal.



The inactive X is identifiable within the nucleus as a heterochromatic Barr body (named for its discoverer, Murray Barr):



In this micrograph, both X chromosomes are stained yellow-green; the Barr body, which consists of heavily methylated DNA, is more condensed than the active X and therefore appears brighter.

Having the right "dosage" of transcriptionally active genes is important. This is illustrated by the fact that aneuploidy—an unusual number of a particular chromosome—is a harmful condition that often results in embryo death (see Concept 7.4).

Epigenetic changes can be induced by the environment

Methylation patterns are stable and can be passed on from one generation to the next. However, a recent study of human monozygotic (identical) twins shows that these patterns can be altered over time.

Monozygotic twins come from a single fertilized egg that divides to produce two separate cells; each of these develops into a separate individual. Identical twins thus have generally identical genomes. But are they identical in their "epigenomes"? A comparison of DNA in hundreds of such twin pairs shows that in tissues of 3-year-olds, DNA methylation patterns are virtually the same. But by age 50—when twins have usually been living apart and in different environments for decades—the patterns are quite different, and different genes are expressed. This indicates that the *environment plays an important role in epigenetic modifications*, and therefore in the regulation of genes that these modifications affect.

Go to MEDIA CLIP 11.1
The Surprising Epigenetics of Identical Twins
PoL2e.com/mc11.1

What factors in the environment lead to epigenetic changes? Chemicals such as tobacco smoke and dietary components such as folic acid can affect DNA methylation patterns. Another factor might be stress: when mice are put in a stressful situation, genes that are involved in important brain pathways become heavily methylated (and transcriptionally inactive). Treatment of the stressed mice with an antidepressant drug reverses these changes.

LINK

Stressful experiences can have lifelong effects on behavior patterns; see **Concept 40.2**

DNA methylation can result in genomic imprinting

In mammals, specific patterns of methylation develop for each sex during gamete formation. This happens in two stages: first, the existing methyl groups are removed from the 5′-methylcytosines by a demethylase, and then a DNA methylase adds methyl groups to a new set of cytosines. When the gametes form, they carry this new pattern of methylation.

The DNA methylation pattern in male gametes (sperm) differs from that in female gametes (eggs) at about 200 genes in the mammalian genome. That is, a given gene in this group may be methylated in eggs but unmethylated in sperm. In this case the offspring would inherit a maternal gene that is transcriptionally inactive (methylated) and a paternal gene that is transcriptionally active (demethylated). This is called **genomic imprinting**.

Imprinting of specific genes occurs primarily in mammals and flowering plants. Most imprinted genes are involved with embryonic development. An embryo must have both the paternally and maternally imprinted gene patterns to develop properly. In fact, attempts to make an embryo that has chromosomes from only one sex (for example, by chemically treating an egg cell to double its chromosomes) usually fail. So imprinting has an important lesson for genetics: *males and females may be the same genetically (except for the X and Y chromosomes), but they differ epigenetically*.

CHECKpoint CONCEPT 11.3

- ✓ What is the difference between epigenetic regulation and gene regulation by transcription factors?
- ✓ How can a DNA methylation pattern be inherited?
- ✓ In colorectal cancer, some tumor suppressor genes are inactive. This is an important factor resulting in uncontrolled cell division. Two of the possible explanations for the inactive genes are: (1) a mutation in the coding region, resulting in an inactive protein, and (2) epigenetic silencing at the promoter of the gene, resulting in reduced transcription. How would you investigate each of these possibilities?

Thus far we have examined transcriptional gene regulation in viruses, prokaryotes, and eukaryotes. In the final concept we will focus on the posttranscriptional mechanisms for regulating gene expression in eukaryotes.

CONCEPT 11.4 Eukaryotic Gene Expression Can Be Regulated after Transcription

Gene expression involves transcription and then translation. So far we have described how eukaryotic gene expression is regulated at the transcriptional level. But as Figure 11.1 shows, there are many points at which regulation can occur after the initial gene transcript is made.

Different mRNAs can be made from the same gene by alternative splicing

Most primary mRNA transcripts in eukaryotes contain several introns (see Figure 10.6). We have seen how the splicing mechanism recognizes the boundaries between exons and introns. What would happen if the β-globin pre-mRNA, which has two introns, were spliced from the start of the first intron to the end of the second? The middle exon would be spliced out along with the two introns. An entirely new protein (certainly not a β-globin) would be made, and the functions of normal β-globin would be lost. Such **alternative splicing** can be a deliberate mechanism for generating a family of different proteins with different activities and functions from a single gene (**FIGURE 11.16**).

Two examples of this mechanism are found in HIV and in the fruit fly (*Drosophila*):

- The HIV genome (see Figure 11.12) encodes nine proteins but is transcribed as a single pre-mRNA. Most of the nine proteins are then generated by alternative splicing of this pre-mRNA.
- In *Drosophila*, sex is determined by the *Sxl* gene. This gene has four exons, which we will designate 1, 2, 3, and 4. In the female embryo, splicing generates two active forms of the Sxl protein, containing exons 1 and 2, and 1, 2, and 4. However, in the male embryo, the protein contains all four exons (1, 2, 3, and 4) and is inactive.

Before the human genome was sequenced, most scientists estimated that they would find between 80,000 and 150,000 protein-coding genes. You can imagine their surprise when the actual sequence revealed only about 21,000 genes! In fact, there are many more human mRNAs than there are human genes, and most of this variation comes from alternative splicing. Indeed, recent surveys show that more than 80 percent of all human genes are alternatively spliced.

Alternative splicing may be a key to the differences in levels of complexity among organisms. For example, although humans and chimpanzees have similar-sized genomes, there is more alternative splicing in the human brain than in the brain of a chimpanzee.

MicroRNAs are important regulators of gene expression

As we will discuss in Concept 12.3, only a small fraction of the genome in most plants and animals codes for proteins. Some of the genome encodes ribosomal RNA and transfer RNAs, but until recently biologists thought that the rest of the genome

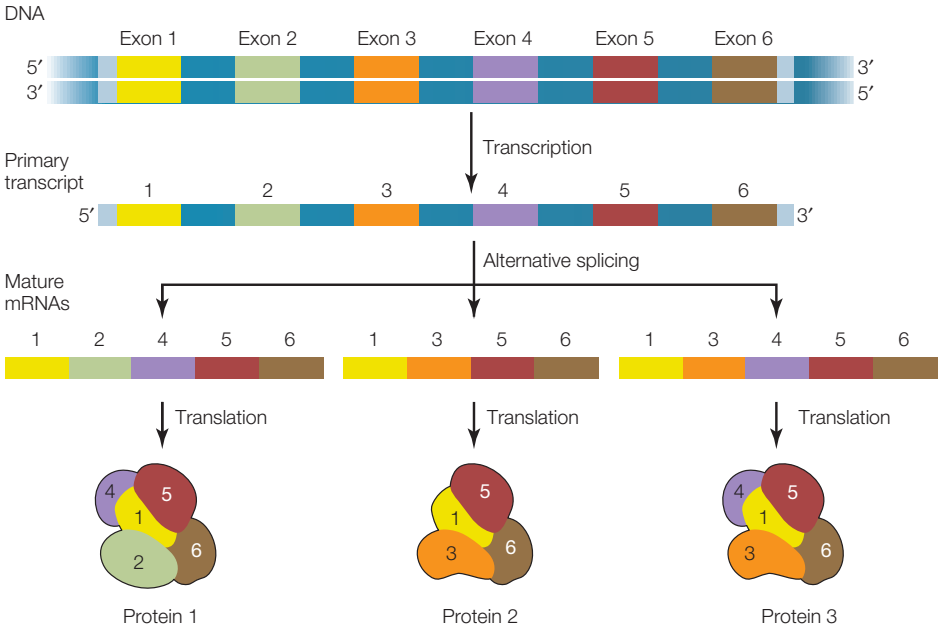


FIGURE 11.16 Alternative Splicing Results in Different Mature mRNAs and Proteins Pre-mRNA can be spliced differently in different tissues, resulting in different proteins.

was not transcribed; some even called it "junk." Recent investigations, however, have shown that some of these noncoding regions are transcribed into tiny RNA molecules called **microRNA** (**miRNA**).

The first miRNA sequences were found in the worm *Caenorhabditis elegans*. This model organism, which has been studied extensively by developmental biologists, goes through several larval stages. Victor Ambros at the University of Massachusetts found *lin* mutations (named for abnormal cell *lin*eage) in two genes that had different effects on progress through these stages:

- *lin-14* mutations cause the larvae to skip the first stage and go straight to the second stage. Thus the gene's normal role is to facilitate events of the first larval stage.
- *lin-4* mutations cause certain cells in later larval stages to repeat a pattern of development normally observed in the first larval stage. It is as if the cells were stuck in that first stage. So the normal role of this gene is to *negatively regulate lin-14*, turning off its expression so the cells can progress to the next stage.

Not surprisingly, further investigation showed that *lin-14* encodes a transcription factor that affects the transcription of genes involved in larval cell progression. It was originally expected that *lin-4*, the negative regulator, would encode a protein that downregulates genes activated by the LIN-14 protein. But this turned out to be incorrect. Instead, *lin-4* encodes a 22-base miRNA that inhibits *lin-14* expression *posttranscriptionally* by binding to its mRNA.

Thousands of miRNAs in a variety of eukaryotes have now been described. Each miRNA is about 22 nucleotides long and usually has dozens of mRNA targets. MicroRNAs are transcribed as longer precursors that fold into double-stranded RNA molecules, which then are processed through a series of steps into single-stranded miRNAs. A protein complex guides the miRNA to its target mRNA, where translation is inhibited and the mRNA is degraded (**FIGURE 11.17**). The remarkable conservation of this gene-silencing mechanism, which is found in most eukaryotes, indicates that it is evolutionarily ancient and biologically important.

Translation of mRNA can be regulated

The amount of a protein in a cell is not determined simply by the amount of its mRNA. For example, in yeast cells only about a third of the genes show clear correlations in the amounts of mRNA and protein; in these cases, more mRNA leads to more protein. For two-thirds of the genes there is no apparent relationship between the two—there may be lots of mRNA and little or no protein, or lots of protein and little mRNA. The concentrations of these proteins must therefore be determined by factors acting after the mRNA is made. Cells do this in two major ways: by regulating the translation of mRNA or by altering how long proteins persist in the cell.

There are three known ways in which the translation of mRNA can be regulated:

- *Inhibition of translation with miRNAs.* This was discussed in the last section (see above).
- *Modification of the 5′ cap.* As noted in Concept 10.2, an mRNA usually has a chemically modified molecule of guanosine triphosphate (GTP) at its 5′ end. An mRNA that is capped with an unmodified GTP molecule is not translated. For example, stored mRNAs in the egg cells of the tobacco hornworm moth are capped with unmodified GTP molecules and are not translated. After the egg is fertilized, however, the caps are modified, allowing the mRNA to be translated to produce the proteins needed for early embryonic development.

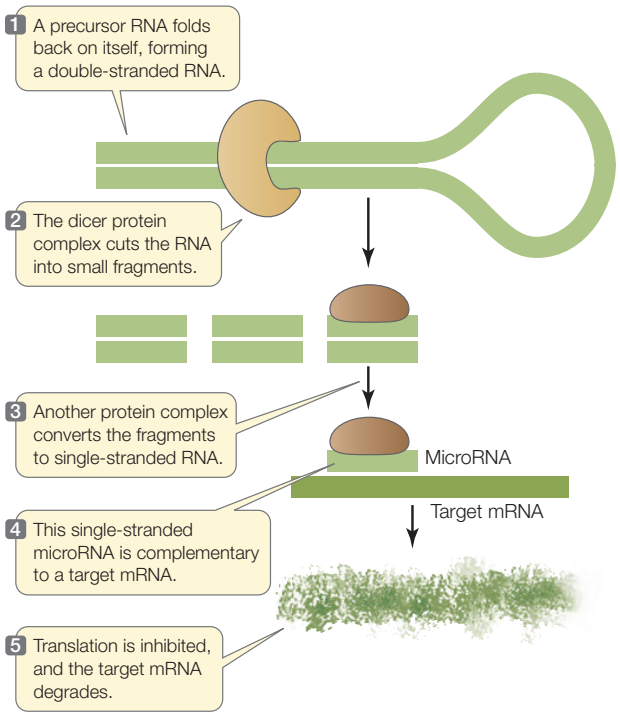


FIGURE 11.17 mRNA Degradation Caused by MicroRNAs
MicroRNAs inhibit the translation of specific mRNAs by causing their premature degradation.

- *Translational repressor proteins.* Such proteins block translation by binding to mRNAs and preventing their attachment to the ribosome. For example, in mammalian cells the rate of translation of the protein ferritin increases rapidly when the level of free iron ions (Fe^{2+}) increases in the cell. Iron is an essential nutrient, but the free ions can be toxic to the cell; ferritin binds the ions and stores them in a safe but accessible form. The amount of ferritin mRNA in the cell remains constant, but when the iron level is low, a **translational repressor** binds to the ferritin mRNA and prevents its translation. When the iron level rises, some of the excess Fe^{2+} ions bind to the repressor and alter its three-dimensional structure, causing the repressor to detach from the mRNA and allowing translation to proceed (**FIGURE 11.18**).

Protein stability can be regulated

The protein content of any cell at a given time is a function of both protein synthesis and protein degradation. Certain proteins can be targeted for destruction in a chain of events that begins when an enzyme attaches a 76–amino acid protein called **ubiquitin** (so named because it is ubiquitous, or widespread) to a lysine residue of the protein to be destroyed. Other ubiquitins then attach to the primary one, forming a polyubiquitin chain. The protein–polyubiquitin complex then binds to a huge protein complex called a **proteasome** (from *protease*

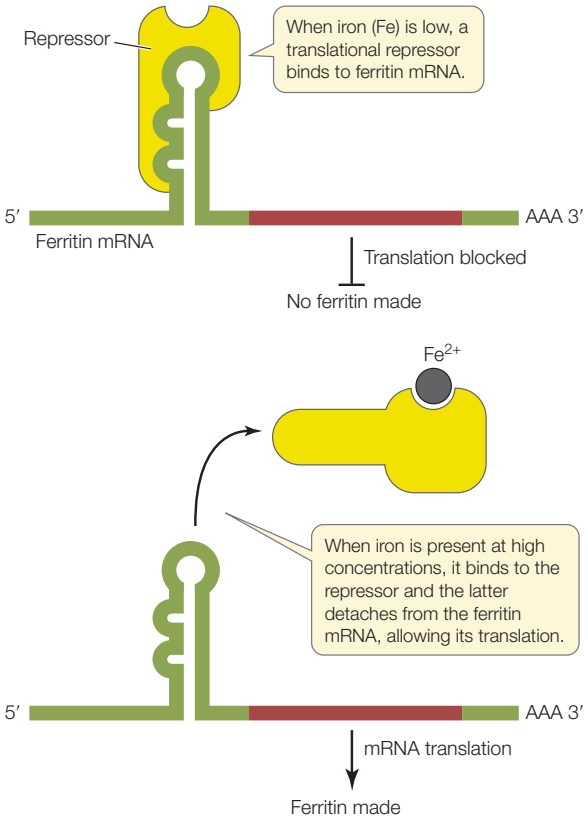


FIGURE 11.18 A Repressor of Translation Binding of a translational repressor to mRNA blocks the mRNA from associating with the ribosome. The repressor can be removed from the mRNA via allosteric regulation.

APPLY THE CONCEPT

Eukaryotic gene expression can be regulated transcriptionally and posttranscriptionally

The enzyme HMG CoA reductase (HR) catalyzes an initial step in the synthesis of cholesterol. The table shows the HR levels in liver cells following various treatments. Explain the results of each of the treatments 1–5.

TREATMENT	AMOUNT OF HR PROTEIN
1. Actinomycin D, a drug that inhibits RNA polymerase II	Reduced
2. Suberoylanilide hydroxamic acid, a histone deacetylase inhibitor	Increased
3. Bortezomib, a proteasome inhibitor	Increased
4. High level of cholesterol	Reduced
5. Azacytidine, inhibitor of DNA methylation	Increased

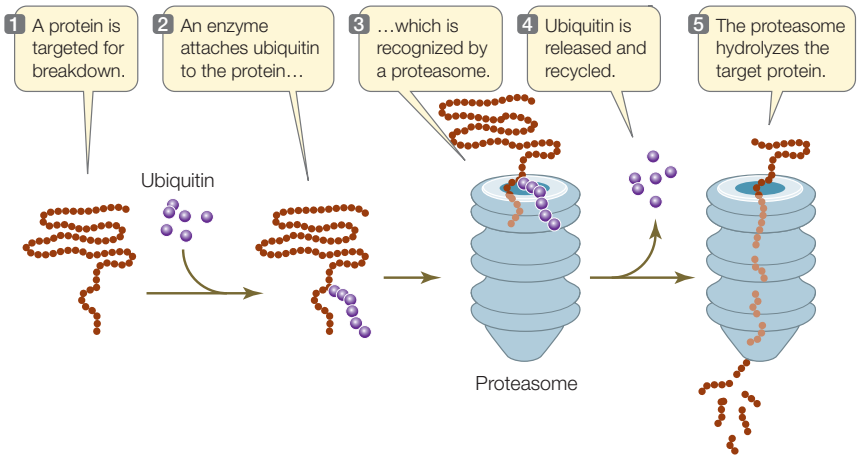


FIGURE 11.19 A Proteasome Breaks Down Proteins Proteins targeted for degradation are bound by ubiquitin, which then directs the targeted protein to a proteasome. The proteasome is a complex structure where proteins are digested by several powerful proteases.

and *soma*, "body"; **FIGURE 11.19**). Upon entering the proteasome, the polyubiquitin is removed and ATP energy is used to unfold the target protein. Three different proteases then digest the protein into small peptides and amino acids. You may recall from Chapter 7 that cyclins are proteins that regulate the activities of key enzymes at specific points in the cell cycle. Cyclins must be broken down at just the right time, and this is done by proteasomes.

CHECKpoint CONCEPT 11.4

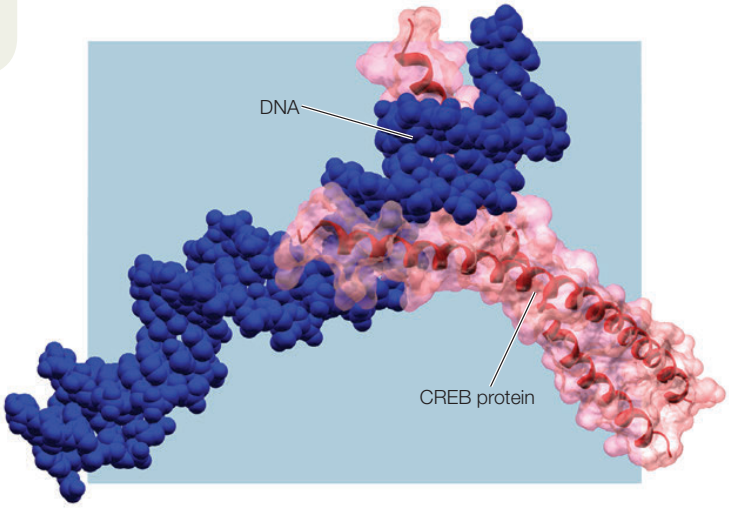
- ✓ How can a single gene transcribed into a single pre-mRNA code for several different proteins?
- ✓ Compare inhibition of translation by miRNA with inhibition by a repressor.
- ✓ You are studying the enzyme protease in germinating seeds. You find that protease activity increases tenfold after treatment of the seeds with a hormone, gibberellic acid. How would you show that this increase is due to:
  a. release of a translational repressor by the hormone?
  b. allosteric inhibition of a transcriptional repressor by the hormone?

Q How does CREB regulate the expression of many genes?

ANSWER CREB is a family of several closely related transcription factors that can either activate or repress gene expression (Concept 11.1). They bind to the *cAMP response element* (*CRE*), a short DNA sequence (GACGTCA) that is found in the promoter regions of many genes (Concept 11.2). CREB proteins have a "leucine zipper" structure that consists of two parallel α helices rich in the amino acid leucine, and fingerlike extensions that fit into the major groove of the DNA double helix (**FIGURE 11.20**).

CREB binding and regulation of gene expression are essential for a number of processes in several organs, including the brain. In addition to its role in drug and alcohol addiction (described in the opening story), CREB has been strongly implicated in long-term memory. Animals with mutations that result in a lack of active CREB can learn a maze test, but they don't remember it later. Similar effects are seen if CREB activity is blocked right after a task is learned. When animals learn a task, imaging studies reveal *CRE*-containing genes becoming active in the hippocampus, a region of the brain that is involved in long-term memory (see Concept 34.5). Thus CREB provides an insight into the molecular biology of memory, linking learning to the regulation of gene expression.

FIGURE 11.20 An Explanation for Alcoholism? The transcription factor CREB binds to DNA and activates the promoters of genes involved in addictive behaviors.



SUMMARY

CONCEPT 11.1 Many Prokaryotic Genes Are Regulated in Operons

- Gene expression can be regulated at the levels of transcription, RNA processing, translation, or posttranslation. **Review Figure 11.1 and ACTIVITY 11.1**
- Some genes are always expressed (**constitutive genes**), whereas others are expressed only at certain times and in certain cells (**inducible genes**).
- **Transcription factors** are regulatory proteins that bind DNA and regulate gene expression. **Activators** positively regulate gene expression. **Repressors** negatively regulate gene expression. **Review Figure 11.2**
- In prokaryotes, several genes can be part of a single transcriptional unit called an **operon**, which consists of a **promoter**, an **operator**, and two or more **structural genes**. **Review Figure 11.3**
- An inducible operon is turned off unless its expression is needed, whereas a repressible operon is turned on unless its expression is not needed. When an operon is turned off, it has a repressor protein bound to its operator, preventing transcription.
- The *lac* operon is an example of an inducible system, whereas the *trp* operon is an example of a repressible system. **Review Figures 11.4 and 11.5 and ANIMATED TUTORIALS 11.1 and 11.2**
- A metabolic pathway can be regulated either by allosteric regulation of an enzyme or by regulation of enzyme synthesis. **Review Figure 11.6**
- **Sigma factors** direct RNA polymerase to specific promoters in prokaryotes.
- **Viruses** provide examples of gene regulation as they convert the host cell into a virus factory. **Review Figure 11.7**

CONCEPT 11.2 Eukaryotic Genes Are Regulated by Transcription Factors

- Eukaryotic gene expression is regulated both during and after transcription.
- **General transcription factors** bind to the core promoter sequences of protein-coding genes and direct RNA polymerase II to the promoter. **Review Figure 11.8 and ANIMATED TUTORIAL 11.3**
- Specific transcription factors (activators and repressors) bind to specific DNA elements near the promoter and affect the rate of transcription initiation. **Review Figure 11.9**

CONCEPT 11.3 Gene Expression Can Be Regulated via Epigenetic Changes to Chromatin

- The term **epigenetic** refers to changes in gene expression that do not involve changes in DNA sequences.
- **Chromatin remodeling** via the modification of **histone** proteins in nucleosomes also affects transcription. **Review Figure 11.14**
- Methylation of cytosines in DNA generally inhibits transcription. **Review Figure 11.15**
- Epigenetic changes can be induced by the environment and can be inherited.

CONCEPT 11.4 Eukaryotic Gene Expression Can Be Regulated after Transcription

- **Alternative splicing** of pre-mRNA can produce different proteins. **Review Figure 11.16**
- A **microRNA** (**miRNA**) is a small noncoding RNA that inhibits the translation of specific mRNAs by causing their premature degradation. **Review Figure 11.17**
- The translation of mRNA to proteins can be regulated by **translational repressors**. **Review Figure 11.18**
- A **proteasome** can break down proteins, thus affecting protein longevity. **Review Figure 11.19**

See **ACTIVITY 11.2** for a concept review of this chapter.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is11

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

12 Genomes

KEY CONCEPTS

The Papillon and the Great Dane are the same species—*Canis lupus familiaris*—yet they show great variation in size. Genome sequencing has provided insights into how size is controlled by genes.

Canis lupus familiaris, the dog, was domesticated by humans from the gray wolf more than 10,000 years ago. There are several kinds of wolves, and they all look more or less the same. Not so with "man's best friend." The American Kennel Club recognizes about 155 different breeds, varying greatly in size, shape, coat color, hair length, and even behavior. For example, an adult Chihuahua weighs just 1.5 kilograms, whereas a Scottish Deerhound weighs 70 kilograms. No other mammalian species shows such large phenotypic variation. Furthermore, we know of hundreds of genetic diseases in dogs, and many of these diseases have counterparts in humans. Biologists are curious about the molecular basis of canine phenotypic variation, and they view dogs as models for studying genetic diseases. For these reasons, the Dog Genome Project began in the late 1990s. Since then the sequences of several dog genomes have been published.

Two dogs—a boxer and a poodle—were the first of their species to have their entire genomes sequenced. The dog genome contains 2.8 billion base pairs of DNA in 39 pairs of chromosomes. There are 22,000 protein-coding genes, most of them with close counterparts in other mammals, including humans. The entire genome sequence made it easy to create a map of genetic markers—specific nucleotides or short sequences of DNA at particular locations on the genome that differ among individual dogs or breeds.

Genetic markers are being used to map the locations of genes that control particular traits. To do this, scientists must extract DNA from many individual dogs that vary in just one or a few characters. Taking samples of cells for DNA isolation is relatively easy: a cotton swab is swept over the inside of the dog's cheek. As one scientist conducting genomic analyses of dogs said, the dogs "didn't care, especially if they were going to get a treat or if there was a tennis ball in our other hand."

The molecular methods used to analyze dogs have been applied to many other animals as well as to plants of economic and social importance to humans. And of course the human genome itself has been sequenced and is being studied intensively.

Q What does genome sequencing reveal about dogs and other animals?

You will find the answer to this question on page 251.

CONCEPT 12.1 There Are Powerful Methods for Sequencing Genomes and Analyzing Gene Products

Genome sequencing involves determining the nucleotide base sequence of the entire genome of an organism. For a prokaryote with a single chromosome, the genome sequence is one continuous string of base pairs (bp). In a eukaryote, there are separate sequences for each chromosome. Scientists can use this genomic information in several ways:

- The genomes of different species can be compared to find out how they differ at the DNA level, and this can be used to trace evolutionary relationships.
- The sequences of individuals within a species and even tissues within an organism can be compared to identify mutations that affect particular phenotypes.
- The sequence information can be used to identify particular traits, such as genes associated with diseases.

The notion of sequencing the entire genome of a complex organism was not contemplated until 1986. The Nobel laureate Renato Dulbecco and others proposed at that time that the world scientific community be mobilized to undertake the sequencing of the entire human genome. One motive was to detect DNA damage in people who had survived the atomic bomb attacks and been exposed to radiation in Japan during the Second World War. But in order to detect changes in the human genome, scientists first needed to know its normal sequence.

The result was the publicly funded **Human Genome Project**, an enormous undertaking that was successfully completed in 2003. This effort was aided and complemented by privately funded groups. The project benefited from the development of many new methods that were first used in the sequencing of smaller genomes—those of prokaryotes and simple eukaryotes, the model organisms you are familiar with from studies in genetics and cell biology. Many of these methods are still applied widely, and powerful new methods for sequencing genomes have emerged. These are complemented by new ways to examine phenotypic diversity in a cell's proteins and in the metabolic products of the cell's enzymes.

Go to ANIMATED TUTORIAL 12.1
Sequencing the Genome
PoL2e.com/at12.1

Methods have been developed to rapidly sequence DNA

Many prokaryotes have a single chromosome, whereas eukaryotes have many. Because of their differing sizes, chromosomes can be separated from one another, identified, and experimentally manipulated. It might seem that the most straightforward way to sequence a chromosome would be to start at one end and simply sequence the DNA molecule one nucleotide at a time. The task is somewhat simplified because only one of the two strands needs to be sequenced, the other being complementary. However, this large-polymer approach is not practical. Using current methods, only several hundred bp can be sequenced at a time, whereas a human chromosome (for example) may be hundreds of millions of bp long.

As you will see, the key to determining genome sequences is to perform many sequencing reactions simultaneously, after first breaking the DNA up into millions of small, overlapping fragments.

In the 1970s, Frederick Sanger and his colleagues invented a way to sequence DNA by using chemically modified nucleotides that were originally developed to stop cell division in cancer. Variations of this method were used to obtain the first human genome sequence as well as those of several model organisms. However, it was relatively slow, expensive, and labor-intensive. The first decade of the new millennium saw the development of faster and less expensive methods, often referred to under the general term **high-throughput sequencing**. These methods use miniaturization techniques first developed for the electronics industry, as well as the principles of DNA replication and the polymerase chain reaction (PCR).

LINK

You can review the processes of DNA replication and PCR in **Concept 9.2**

High-throughput sequencing methods are rapidly evolving. Just one of the many approaches is outlined here and illustrated in **FIGURE 12.1**. First the DNA is prepared for sequencing:

1. A large molecule of DNA is cut into small fragments of about 100 bp each. This can be done physically, using mechanical forces to shear (break up) the DNA, or by using enzymes that hydrolyze the phosphodiester bonds between nucleotides at intervals in the DNA backbone.
2. The DNA is denatured by heat, breaking the hydrogen bonds that hold the two strands together. Each single strand acts as a template for the synthesis of new, complementary DNA.
3. Short, synthetic adapter sequences (oligonucleotides) are attached to each end of each fragment, and the fragments are attached to a solid support. The support can be a microbead or a flat surface.
4. Primers complementary to the adapters are used in PCR reactions to produce many (approximately 1,000) copies of each DNA fragment. The multiple copies at a single location allow for easy detection of added nucleotides during the sequencing steps.

Once the DNA has been attached to a solid substrate and amplified, it is ready to be used as a template for sequencing (see Figure 12.1B):

1. At the beginning of each sequencing cycle, the DNA fragments are heated to denature them. A universal primer, DNA polymerase, and the four deoxyribonucleoside triphosphates (dNTPs: dATP, dGTP, dCTP, and dTTP)

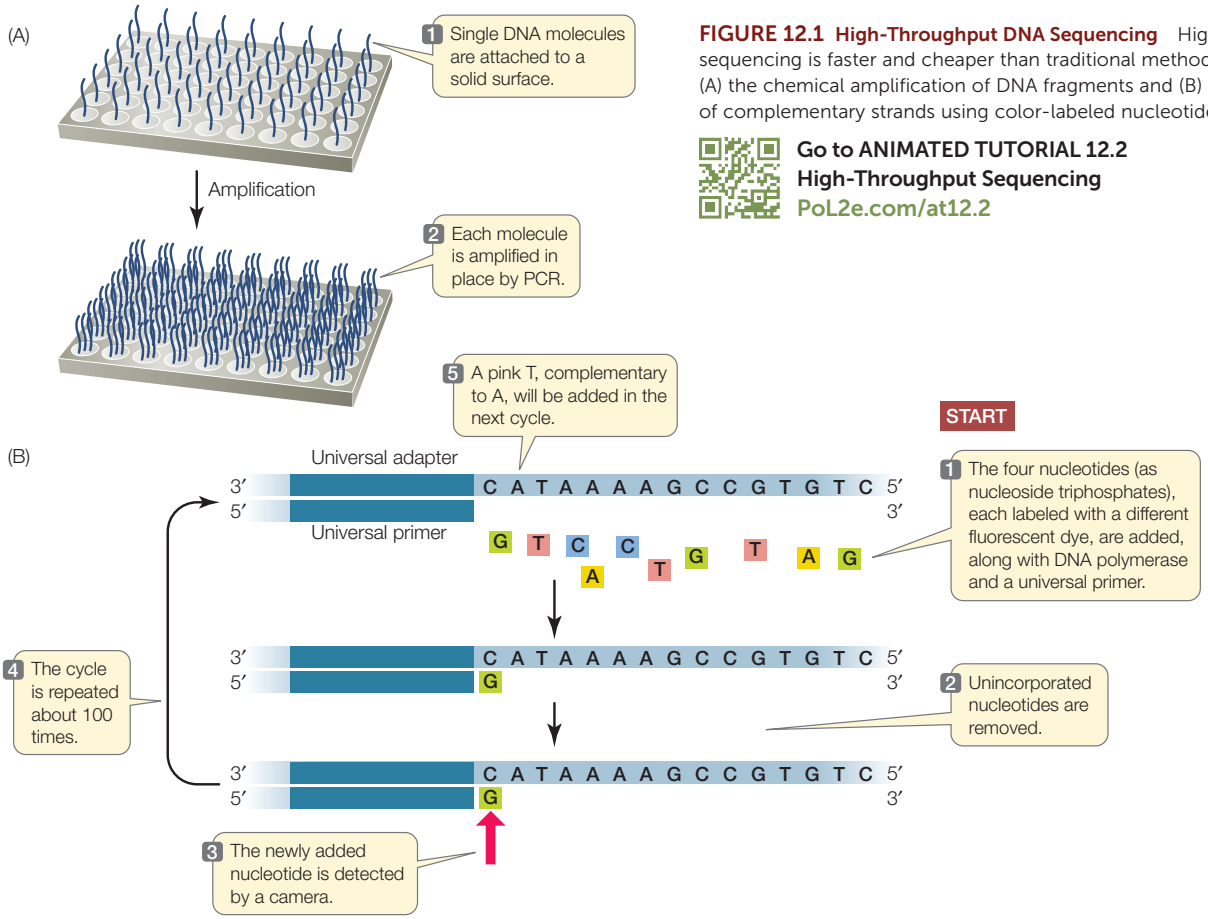


FIGURE 12.1 High-Throughput DNA Sequencing High-throughput sequencing is faster and cheaper than traditional methods. It involves (A) the chemical amplification of DNA fragments and (B) the synthesis of complementary strands using color-labeled nucleotides.

Go to ANIMATED TUTORIAL 12.2
High-Throughput Sequencing
PoL2e.com/at12.2

are added. Each of the four nucleotides (i.e., the dNTPs) is tagged with a different fluorescent dye. The universal primer is complementary to the adapter sequence at one end of each DNA fragment.

2. The DNA sequencing reaction is set up so that only one nucleotide at a time is added to the new DNA strand, which is complementary to the template strand. After each addition, the unincorporated nucleotides are removed.
3. The fluorescence of the new nucleotide at each location is detected with a camera. The color of the fluorescence indicates which of the four nucleotides was added.
4. The fluorescent tag is removed from the nucleotide that is already attached, and then the sequencing cycle is repeated. Images are captured after each nucleotide is added. The series of colors at each location indicates the sequence of nucleotides in the growing DNA strand at that location.

The power of this method derives from several factors:

- It is fully automated and miniaturized.
- Millions of different fragments are sequenced at the same time.
- It is an inexpensive way to sequence large genomes. For example, at the time of this writing, a complete human genome could be sequenced in a few days for several thousand dollars. In contrast, the Human Genome Project took 13 years and $2.7 billion to sequence one genome!

The technology used to sequence millions of short DNA fragments is only half the story, however. Once these sequences have been determined, the problem becomes how to put them together. In other words, how are they arranged in the chromosomes from which they came? Imagine if you cut out every word in this book (there are about 500,000 of them), put them on a table, and tried to arrange them in their original order! The enormous task of determining DNA sequences is possible because the original DNA fragments are *overlapping*.

Let's illustrate the process using a single 10-bp DNA molecule. (This is a double-stranded molecule, but for convenience we show only the sequence of the noncoding strand.) The molecule is cut three ways using three different enzymes. The first cut generates the fragments:

TG, ATG, and CCTAC

The second cut of the same molecule generates the fragments:

AT, GCC, and TACTG

The third cut results in:

CTG, CTA, and ATGC

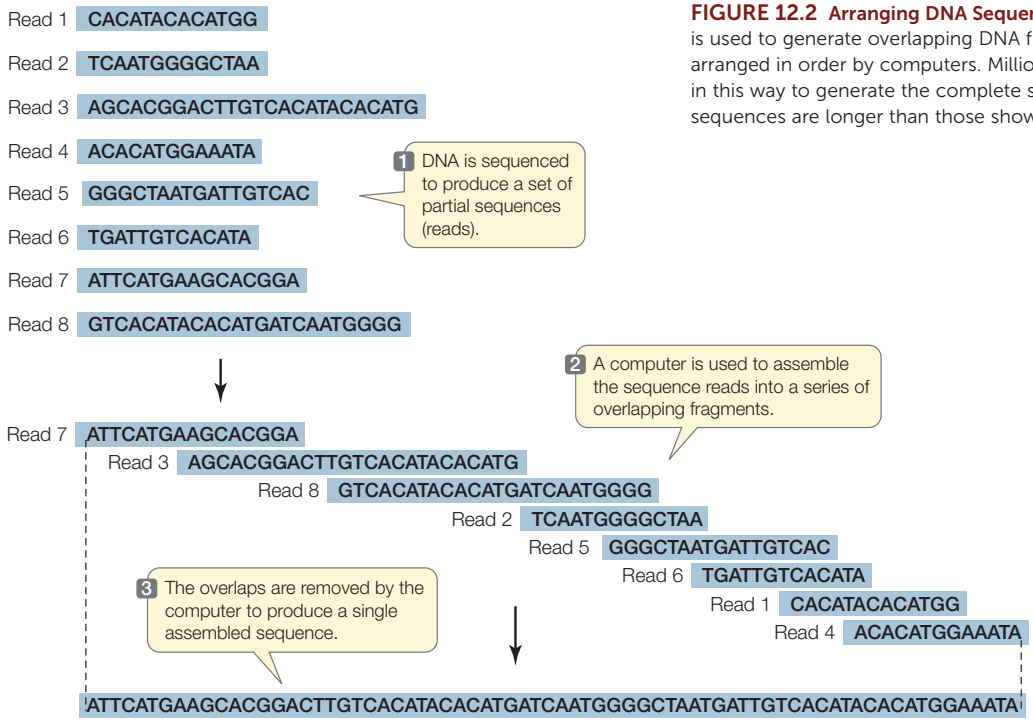
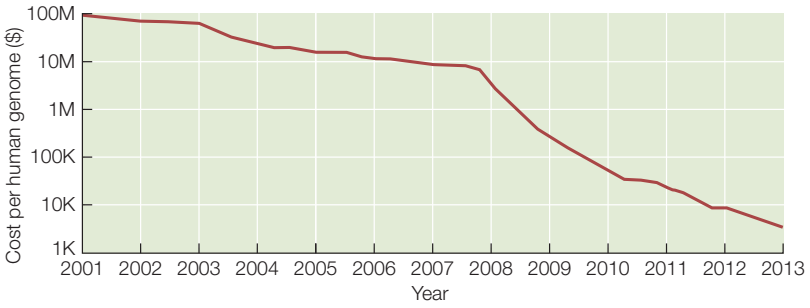


FIGURE 12.2 Arranging DNA Sequences A series of different cuts is used to generate overlapping DNA fragments. Their sequences are arranged in order by computers. Millions of sequences are arranged in this way to generate the complete sequence of a genome. Actual sequences are longer than those shown here.

Can you put the fragments in the correct order? (The answer is ATGCCTACTG.) For genome sequencing, the fragments are called "reads" (FIGURE 12.2). Of course, the problem of ordering 2.5 million fragments of 100 bp from human chromosome 1 (246 million bp) is more challenging than our 10-bp example above. The field of **bioinformatics** was developed to analyze DNA sequences using complex mathematics and computer programs. The cost of genome sequencing and analysis has gone down rapidly:



Genome sequences yield several kinds of information

New genome sequences are being published at an accelerating pace, creating a torrent of biological information. This information is used in two related fields of research, both focused on studying genomes. In **functional genomics**, biologists use sequence information to identify the functions of various parts of genomes (FIGURE 12.3). These parts include:

- **Open reading frames**, which are sequences of DNA that contain no stop codons, and thus may encode parts of proteins. An open reading frame that begins with a start codon or an intron consensus sequence (boundary between exon and intron) and ends with a stop codon or an intron consensus sequence may be an exon, which encodes part or all of a polypeptide. See Concepts 10.2 and 10.3 to review introns, exons, and the genetic code.
  - Regulatory sequences, such as promoters and terminators for transcription. These are identified by their proximity to open reading frames and because they contain consensus sequences for the binding of specific transcription factors and RNA polymerase.
  - Regions of DNA that have regulatory sequences at each end and one or more open reading frames; these may be protein-coding genes. The amino acid sequence of a protein can be deduced by applying the genetic code (see Figure 10.11) to the DNA sequences of the open reading frames within the gene. A major goal of functional genomics is to identify, and understand the function of, every protein-coding gene in each genome.
- RNA genes, including genes for rRNA, tRNA, and miRNA (see Concept 11.4).
- Other noncoding sequences that can be classified into various categories, including centromeric regions (see Concept 7.2), telomeric regions (see Concept 9.2), transposons (see Concept 12.2), and other repetitive sequences.

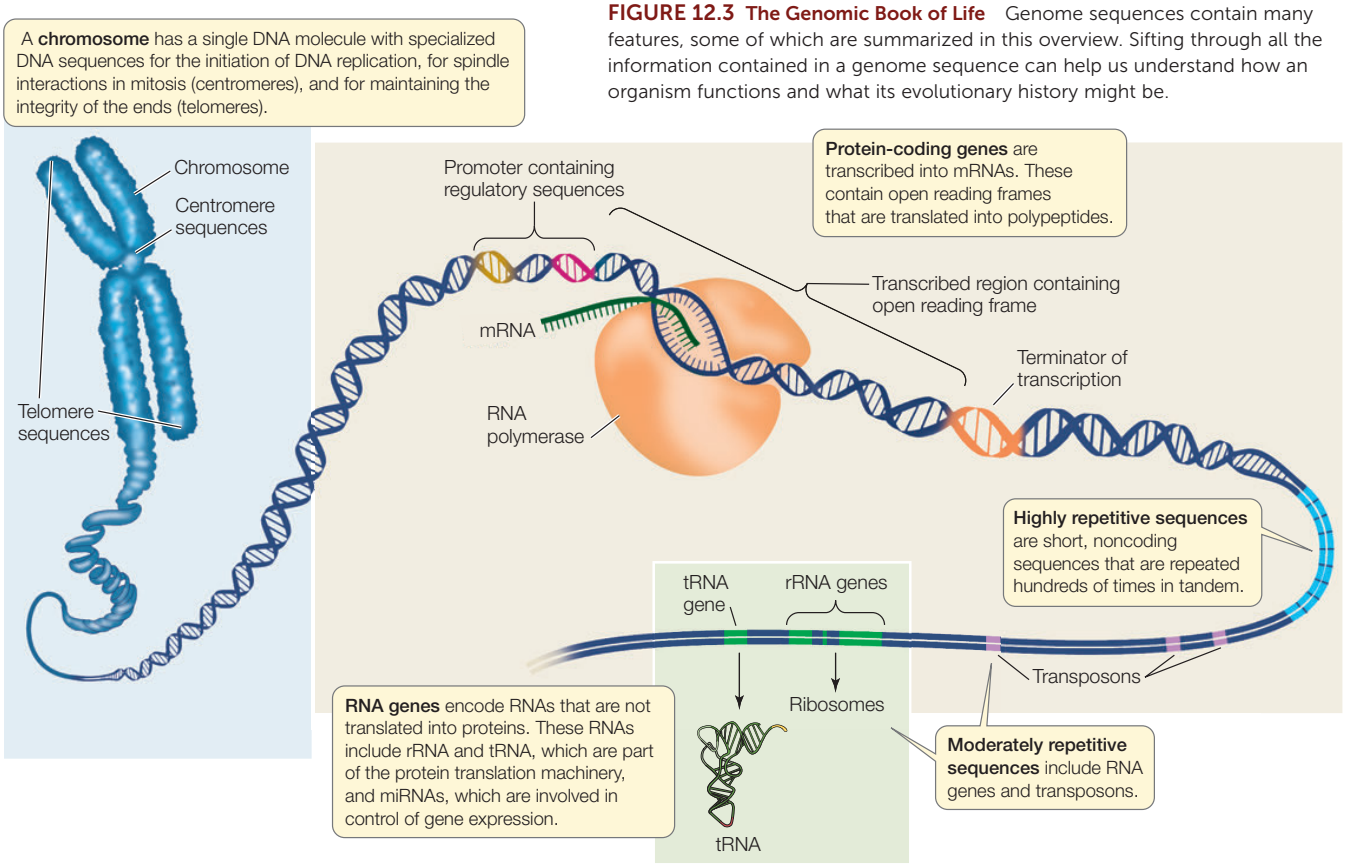


FIGURE 12.3 The Genomic Book of Life Genome sequences contain many features, some of which are summarized in this overview. Sifting through all the information contained in a genome sequence can help us understand how an organism functions and what its evolutionary history might be.

Functional regions in a newly described genomic sequence can be identified by searching DNA databases for similar or identical sequences in other organisms. There are now massive databases (accessible online) containing DNA sequences and their known or possible functions.

Sequence information is also used in **comparative genomics**: the comparison of a newly sequenced genome (or parts thereof) with sequences from other organisms. This can provide further information about the functions of sequences and can be used to trace evolutionary relationships among different organisms.

LINK

The application of genome sequencing to reconstructing phylogenies (evolutionary trees) is described in **Concept 16.2**

Phenotypes can be analyzed using proteomics and metabolomics

"The human genome is the book of life." Statements like this were common when the human genome sequence was first revealed. They reflect the concept of genetic determinism—the idea that a person's phenotype is determined solely by his or her genotype. But is an organism just a product of gene expression? We know that it is not. The proteins and small molecules present in any cell at a given point in time reflect not just gene expression but modifications caused by the intracellular and extracellular environments. Two new fields have emerged to complement genomics and take a more complete snapshot of a cell or organism: proteomics and metabolomics.

PROTEOMICS Many genes encode more than a single protein (**FIGURE 12.4A**). As we described in Concept 11.4, alternative splicing leads to different combinations of exons in the mature mRNAs transcribed from a single gene. Posttranslational modifications also increase the number and the structural and functional diversity of proteins derived from one gene (see Figures 12.4A and 10.21). The **proteome** is the sum total of the proteins produced by an organism, and it is more complex than the organism's genome.

Several approaches are commonly used to analyze proteins and the proteome:

- Because of their unique amino acid compositions (primary structures), most proteins have unique combinations of electric charge and size. On the basis of these two properties, they can be separated by two-dimensional gel electrophoresis (**FIGURE 12.4B**).
- Once they have been isolated, individual proteins can be analyzed by mass spectrometry. This technique uses electromagnets to identify molecules by the masses of their atoms, and it can also be used to determine the structures of molecules.
- Antibodies can also be used to isolate specific proteins, or to detect the proteins in cells or tissues.

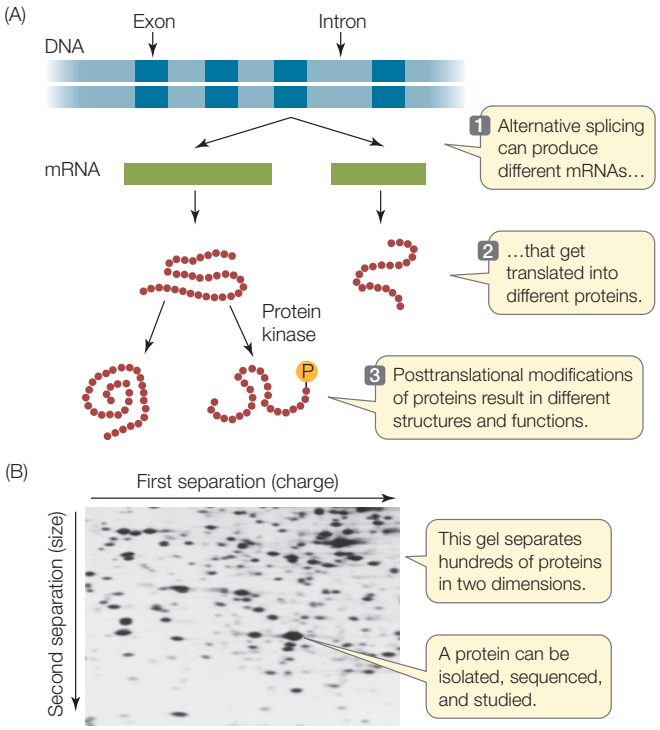


FIGURE 12.4 Proteomics (A) A single gene can code for multiple proteins. (B) Some of a cell's proteins can be separated on the basis of charge and size by two-dimensional gel electrophoresis. The two separations can distinguish some proteins from one another. Further analysis of each spot by mass spectrometry identifies different proteins.

Whereas genomics seeks to describe the genome and its expression, **proteomics** seeks to identify and characterize all of the expressed proteins. Its ultimate aim is just as ambitious as that of genomics. Comparisons of proteomes among organisms have revealed common sets of proteins that can be categorized into groups with similar amino acid sequences and functions. Often these share three-dimensional structural regions called domains (for example, the heme-binding domain of hemoglobin). While a particular organism may have many unique proteins, those proteins are often just unique combinations of domains that exist in proteins of other organisms. This reshuffling of the genetic deck is a key to evolution.

METABOLOMICS Studying genes and proteins gives a limited picture of what is going on in a cell. But as we have seen, both gene function and protein function are affected by a cell's internal and external environments. Many proteins are enzymes, and their activities affect the concentrations of their substrates and products. So as the proteome changes, so do the abundances of these (often small) molecules, called metabolites. The **metabolome** is the complete set of small molecules present in a cell, tissue, or organism. These include:

- *Primary metabolites* that are involved in normal processes, such as intermediates in pathways such as glycolysis. This category also includes hormones and other signaling molecules.
- *Secondary metabolites*, which are often unique to particular organisms or groups of organisms. They are often involved in special responses to the environment. Examples are antibiotics made by microbes, and the many chemicals made by plants that are used in defense against pathogens and herbivores.

Metabolomics aims to describe the metabolic profile of a tissue or organism under particular environmental conditions. Measuring metabolites involves sophisticated analytical instruments. If you have studied organic or analytical chemistry, you may be familiar with gas chromatography and high-performance liquid chromatography, which are used to separate molecules with different chemical properties. Mass spectrometry and nuclear magnetic resonance spectroscopy are used to identify molecules. These measurements result in "chemical snapshots" of cells or organisms, which can be related to physiological states.

Plant biologists are far ahead of medical researchers in the field of metabolomics. Tens of thousands of secondary metabolites, many of them made in response to environmental challenges, have been identified in plants. The metabolomes of agriculturally important plants are being described, and this information may be important in optimizing plant growth for food production.

LINK

Some of the many secondary metabolites made by plants are discussed in **Concept 28.2**

Taken together, the genome, proteome, and metabolome can move biologists toward a more comprehensive picture of an organism's genotype and phenotype (**FIGURE 12.5**).

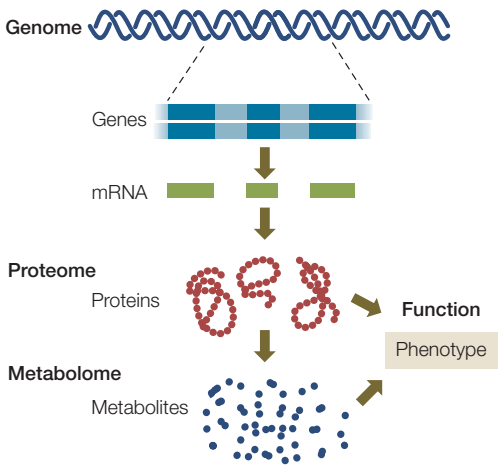


FIGURE 12.5 Genomics, Proteomics, and Metabolomics A combination of these approaches can give more comprehensive information about genotypes and phenotypes.

CHECKpoint CONCEPT 12.1

- ✓ Using a table, compare genomics, proteomics, and metabolomics with regard to the methods used and results obtained.
- ✓ A DNA molecule is cut into the following fragments that are sequenced: AGTTT, TAGG, CGAT, and CCT. The same molecule is cut in a different way to produce TTCGA, TCCT, AGT, GG, and TA. A third cut produces TTCGAT, CCTT, AGG, and AGT. What is the sequence of this DNA?
- ✓ If you were designing a computer program to recognize important sequences in DNA, what types of sequences would you include? What would you learn from finding these sequences in the DNA? Be as specific as you can.

The first cellular genomes to be fully sequenced were those of prokaryotes. In the next concept we will discuss these relatively small, compact genomes.

CONCEPT 12.2 Prokaryotic Genomes Are Small, Compact, and Diverse

When DNA sequencing became possible in the late 1970s, the first life forms to be sequenced were the simplest viruses. The sequences quickly provided new information about how viruses infect their hosts and reproduce. The next genomes to be fully sequenced were those of prokaryotes. We now have genome sequences for many microorganisms, to the great benefit of microbiology and medicine.

Prokaryotic genomes are compact

In 1995 a team led by Craig Venter and Hamilton Smith published the first complete genomic sequence of a free-living cellular organism, the bacterium *Haemophilus influenzae*. Many more prokaryotic sequences have followed. These sequences reveal not only how prokaryotic genes are organized to perform different cellular functions, but also how certain specialized functions of particular organisms are carried out.

There are several notable features of bacterial and archaeal genomes:

- They are relatively small. Prokaryotic genomes range in size from about 160,000 to 12 million bp and are usually organized into a single circular chromosome.
- They are compact. Typically, more than 85 percent of the DNA consists of protein-coding regions or RNA genes, with only short sequences between genes.
- Their genes usually do not contain introns. An exception is the rRNA and tRNA genes of archaea, which are frequently interrupted by introns.
- In addition to the main chromosome, they often carry smaller, circular DNA molecules called plasmids, which may be transferred between cells (see Concept 8.4).

Beyond these broad similarities, there is great diversity among these single-celled organisms, reflecting the huge variety of the environments where they are found.

Let's look in more detail at a few prokaryotic genomes in terms of functional and comparative genomics.

FUNCTIONAL GENOMICS As mentioned above, functional genomics is a biological discipline that assigns functions to DNA sequences. This field is less than 20 years old but is now a major occupation of biologists. You can see the various functions encoded by the genomes of three prokaryotes (in this case, all bacteria) in TABLE 12.1.

H. influenzae lives in the upper respiratory tracts of humans and can cause ear infections and (more seriously) meningitis. Its single circular chromosome has 1,830,138 bp. In addition to its origin of replication and the RNA genes, this bacterial chromosome has 1,727 open reading frames.

When this sequence was first announced, only 1,007 (58 percent) of the open reading frames encoded proteins with known functions. Since then scientists have identified the role of almost every protein encoded by the *H. influenzae* genome. All of the major biochemical pathways and molecular functions are represented.

COMPARATIVE GENOMICS Soon after the sequence of *H. influenzae* was announced, the smaller *Mycoplasma genitalium* (580,073 bp) and the larger *E. coli* (4,639,221 bp) genomic sequences were completed. Thus began the new era of comparative genomics. Scientists can identify genes that are present in one bacterium and missing in another, allowing them to relate these genes to bacterial function.

For example, *E. coli* has more genes than *H. influenzae* in each of the functional groups listed in Table 12.1. This suggests

TABLE 12.1 Gene Functions in Three Bacteria

	Number of genes in:		
Category	***E. coli***	***H. influenzae***	***M. genitalium***
Total protein-coding genes	4,288	1,727	482
Biosynthesis of amino acids	131	68	1
Biosynthesis of cofactors	103	54	5
Biosynthesis of nucleotides	58	53	19
Cell envelope proteins	237	84	17
Energy metabolism	243	112	31
Intermediary metabolism	188	30	6
Lipid metabolism	48	25	6
DNA replication, recombination, and repair	115	87	32
Protein folding	9	6	7
Regulatory proteins	178	64	7
Transcription	55	27	12
Translation	182	141	101
Uptake of molecules from the environment	427	123	34

that there may be more biochemical pathways in *E. coli* than in *H. influenzae*. *M. genitalium* lacks most of the enzymes needed to synthesize amino acids, which *E. coli* and *H. influenzae* both possess (see Table 12.1). This finding reveals that *M. genitalium* must obtain its amino acids from its environment (usually the human urogenital tract). Furthermore, *E. coli* has dozens of genes for regulatory proteins that encode transcriptional activators or repressors; *M. genitalium* has only seven such genes. This suggests that the biochemical flexibility of *M. genitalium* is limited by its relative lack of control over gene expression.

Metagenomics reveals the diversity of viruses and prokaryotic organisms

If you take a microbiology laboratory course, you will learn how to identify various prokaryotes on the basis of their growth in lab cultures. Microorganisms can be identified by their nutritional requirements or the conditions under which they will grow (such as aerobic versus anaerobic). For example, staphylococci are a group of bacteria that inhabit skin and nasal passages. Unlike many bacteria, staphylococci can use the sugar alcohol mannitol as an energy source and thus can grow on a special medium containing mannitol. Often a dye is included in the medium, which changes color if the bacteria are pathogenic (disease-causing). Such culture methods have been the mainstay of microbial identification for more than a century and are still useful and important. However, scientists can now use PCR and modern DNA analysis techniques to analyze microbes *without* culturing them in the laboratory.

In 1985 Norman Pace, then at Indiana University, came up with the idea of isolating DNA directly from environmental samples. He used PCR to amplify specific sequences from the samples to determine whether particular microbes were present. The PCR products were sequenced to explore their diversity. The term **metagenomics** was coined to describe this approach of analyzing genes without isolating the intact organism. It is now possible to do DNA sequencing with samples from almost any environment. The sequences can be used to detect the presence of previously unidentified organisms as well as known microbes (**FIGURE 12.6**). For example:

- Sequencing of DNA from 200 liters of seawater indicated that it contained 5,000 different viruses and 2,000 different bacteria, many of which had not been described previously.
- Samples from the intestines of chickens and turkeys from different flocks have led to the identification of viral causes of serious diseases in these domesticated birds.
- Water runoff from a mine contaminated with toxic chemicals contained many new species of prokaryotes thriving in this apparently inhospitable environment. Some of these organisms exhibited metabolic pathways that were previously unknown to biologists. These organisms and their capabilities may be useful in cleaning up pollutants from the water.
- Gut samples from 124 Europeans revealed that each person harbored at least 160 species of bacteria (constituting their gut microflora or microbiome). Many of these species were found in all of the individuals, but the presence of other bacteria varied from person to person. Such variations in gut microflora may be associated with obesity or bowel diseases.

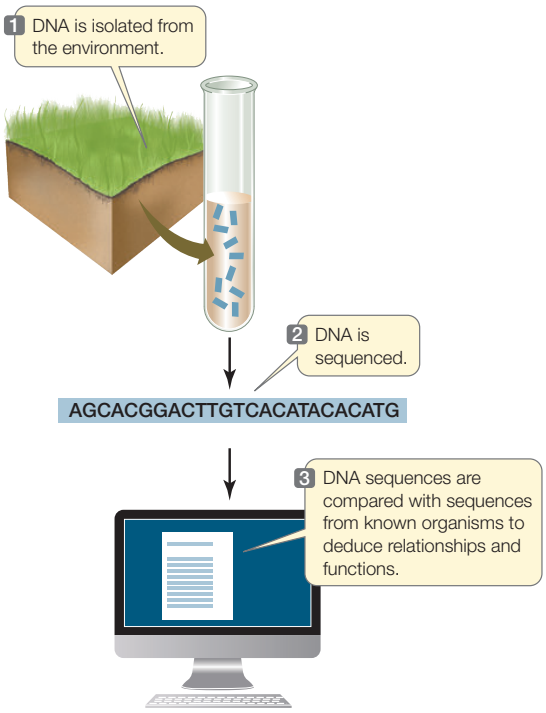


FIGURE 12.6 Metagenomics Microbial DNA extracted from the environment can be amplified and sequenced directly. This has led to the description of many new genes and species.

LINK

For more on the complex microbial ecosystem inside the human gut, see **Figures 19.20 and 41.1**

These and other discoveries are truly extraordinary and potentially very important. It is estimated that 90 percent of the microbial world has been invisible to biologists, in part because the cells could not be grown in the laboratory. These organisms are only now being revealed by metagenomics. Entirely new ecosystems of bacteria and viruses are being discovered in which, for example, one species produces a molecule that another metabolizes. It is hard to overemphasize the importance of such an increase in our knowledge of the hidden world of microbes. This new knowledge underscores the remarkable diversity among prokaryotic organisms, and will further our understanding of natural ecological processes. Furthermore, it has the potential to help us find better ways to manage environmental catastrophes such as oil spills, or to remove toxic heavy metals from soil and water.

Some sequences of DNA can move about the genome

Genome sequencing allowed scientists to study more broadly a class of DNA sequences that had been discovered by geneticists

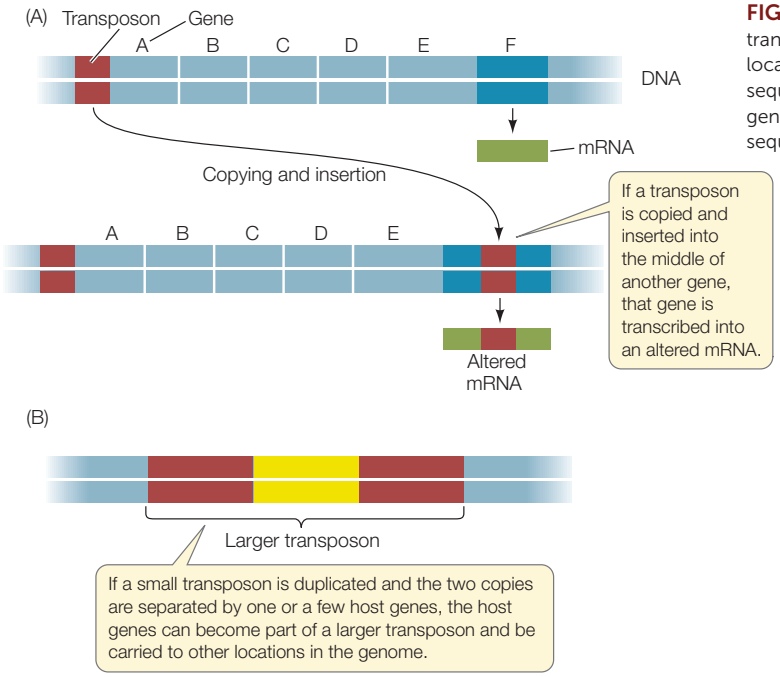


FIGURE 12.7 DNA Sequences That Move Transposons (or transposable elements) are DNA sequences that move from one location to another. (A) In one method of transposition, the DNA sequence is replicated and the copy inserts elsewhere in the genome. (B) Transposons can evolve to carry additional genomic sequences.

decades earlier. Segments of DNA called **transposons** (or transposable elements) can move from place to place in the genome and can even move from one piece of DNA (such as a chromosome) to another (such as a plasmid) in the same cell. A transposon might be at one location in the genome of one *E. coli* cell, and at a different location in another cell. The insertion of this movable DNA sequence from elsewhere in the genome into the middle of a protein-coding gene disrupts that gene (**FIGURE 12.7A**). Any mRNA expressed from the disrupted gene will have the extra sequence, and the protein will be abnormal. Consequently transposons can produce significant phenotypic effects by inactivating genes.

Transposons are often short sequences of 1,000–2,000 bp and are found at many sites in prokaryotic genomes. The mechanisms that allow them to move vary. For example, the transposon may be replicated, and then the copy inserted into another site in the genome. Or the transposon might splice out of one location and move to another location.

If a transposon becomes duplicated, with two copies separated by one or a few genes, the result may be a single larger transposon (up to about 5,000 bp). In this case, the additional genes can be carried to different locations in the genome (**FIGURE 12.7B**). Some of these transposons carry genes for antibiotic resistance. We will discuss transposons again in Concept 12.3.

Will defining the genes required for cellular life lead to artificial life?

When the genomes of prokaryotes and eukaryotes are compared, a striking conclusion arises: certain genes are present in all organisms (universal genes). There are also some (nearly) universal gene segments that are present in many genes in many organisms. One example is a sequence encoding an ATP binding site, which is a domain found in many proteins. These findings suggest that there is some ancient, minimal set of DNA sequences common to all cells. One way to identify these sequences is to look for them in computer analyses of sequenced genomes.

Another way to define the minimal genome is to take an organism with a simple genome and deliberately mutate one gene at a time to see what happens. *M. genitalium* has one of the smallest known genomes, with only 482 protein-coding genes. Even so, some of its genes are dispensable under some circumstances. For example, it has genes for metabolizing both glucose and fructose, but it can survive in the laboratory on a medium containing only one of these sugars. Under such circumstances, the bacterium doesn't need the genes for metabolizing the other sugar.

What about other genes? Researchers addressed this question using transposons as mutagens. When transposons in the bacterium were activated, they inserted themselves into genes at random, mutating and inactivating the genes (**FIGURE 12.8**). The mutated bacteria were tested for growth and survival, and DNA from interesting mutants was sequenced to find out which genes were mutated. The astonishing results of these studies suggested that only 382 of the 482 *M. genitalium* protein-coding genes were needed for survival in the laboratory!

One application of the research might be to design organisms with specific uses. The next step toward that goal is to create an artificial genome and insert it into bacterial cells. As we described in the opening story of Chapter 4, this was recently accomplished, using a synthetic genome based on that of the bacterium *Mycoplasma mycoides*. This research has promise for making organisms with novel functions, such as the synthesis of plastics polymers or the ability to break down environmental pollutants.

CHECKpoint CONCEPT 12.2

- ✓ What are the characteristics of most prokaryotic genomes?
- ✓ Examine Table 12.1 and Figure 12.8. What gene functions would you predict are nonessential for *M. genitalium* as determined by transposon-mediated inactivation?
- ✓ You want to isolate a prokaryote that can live on discarded Styrofoam cups. Such an organism might live in a landfill where ground-up cups are discarded. How would you use metagenomics to identify such a bacterium?
- ✓ How would you show that the prokaryote's ability to live on Styrofoam is essential, and that it cannot live in another environment?

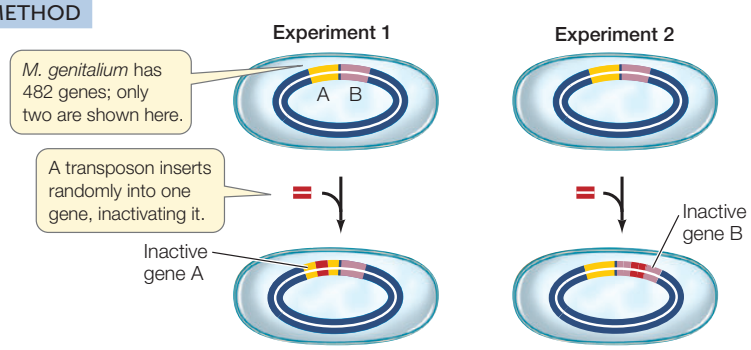
INVESTIGATION

FIGURE 12.8 Using Transposon Mutagenesis to Determine the Minimal Genome
Mycoplasma genitalium has one of the smallest known genomes of any prokaryote. But are all of its genes essential to life? By inactivating the genes one by one, scientists determined which of them are essential for the cell's survival. This research may lead to the construction of artificial cells with customized genomes, designed to perform functions such as degrading oil and making plastics.[a]

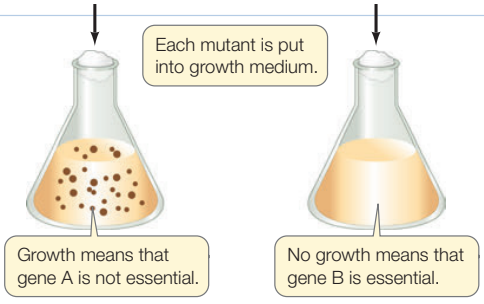
HYPOTHESIS

Only some of the genes in a bacterial genome are essential for cell survival.

METHOD



RESULTS



CONCLUSION

If each gene is inactivated in turn, a "minimal essential genome" can be determined.

ANALYZE THE DATA

The growth of *M. genitalium* strains with insertions in genes (intragenic regions) was compared with the growth of strains with insertions in noncoding (intergenic) regions of the genome:

Type of insertion	Number of strains with insertions	Number of strains that grew
Intragenic	482	100
Intergenic	199	184

A. Explain these data in terms of genes essential for growth and survival. Are all of the genes in *M. genitalium* essential for growth? If not, how many are essential? Why did some of the insertions in intergenic regions prevent growth?

B. If a transposon inserts into the following regions of genes, there might be no effect on phenotype. Explain why in each case:

i. near the 3′ end of the coding region

ii. within a gene coding for rRNA

How does this affect your answer to the first question?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]C. Hutchison et al. 1999. *Science* 286: 2165–2169. J. I. Glass et al. 2006. *Proceedings of the National Academy of Sciences USA* 103: 425–430.

The methods used to sequence and analyze prokaryotic genomes have also been applied to eukaryotic genomes, which we will examine next.

CONCEPT 12.3 Eukaryotic Genomes Are Large and Complex

As genomes have been sequenced and described, a number of major differences have emerged between eukaryotic and prokaryotic genomes:

- *Eukaryotic genomes are larger than those of prokaryotes*, and they have more protein-coding genes. This difference is not surprising given that multicellular organisms have many cell types with specialized functions. As we saw above, one of the simplest prokaryotes, *Mycoplasma*, has several hundred protein-coding genes in a genome of about 0.5 million bp. A rice plant, in contrast, has about 40,383 protein-coding genes.
- *Eukaryotic genomes have more regulatory sequences*—and many more regulatory proteins—than prokaryotic genomes. The greater complexity of eukaryotes requires much more regulation, which is evident in the many points of control associated with the expression of eukaryotic genes (see Concepts 11.2–11.4 and Figure 11.1).
- *Much of eukaryotic DNA does not encode proteins*. Distributed throughout many eukaryotic genomes are various kinds of DNA sequences that are not transcribed into mRNA. Some of these sequences are genes for functional RNAs, such as rRNA, tRNA, and miRNA. Others are introns or regulatory sequences. In addition, eukaryotic genomes contain various kinds of repeated sequences.

Model organisms reveal many characteristics of eukaryotic genomes

Most of our information about eukaryotic genomes has come from model organisms that have been studied extensively. These include the yeast *Saccharomyces cerevisiae*, the nematode (roundworm) *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and the plant *Arabidopsis thaliana* (thale cress). Model organisms have been chosen

APPLY THE CONCEPT

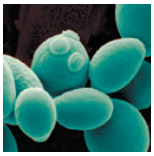
Eukaryotic genomes are large and complex

Repetitive DNA sequences can be classified by nucleic acid hybridization (see Figure 10.7). A genome is initially cut into 300-bp fragments, and these are heated to denature the DNA. If the solution is cooled, the DNA strands will form hydrogen bonds and reassociate into double-stranded structures. If there are many copies of a DNA sequence in the solution (repetitive DNA), it will find its complementary sequence and reassociate faster than if there are only a few copies. The table shows typical results from equal amounts of DNA from three species.

	PERCENTAGE OF DNA REASSOCIATED		
REASSOCIATION TIME (MIN)	*E. COLI*	YEAST	MOUSE
1	0	3	10
10	0	17	35
100	100	100	100

1. Why do yeast and mouse DNAs reassociate faster than *E. coli* DNA?
2. Would you expect human DNA to reassociate faster or slower than yeast DNA?

because they are relatively easy to grow and study in a laboratory, their genetics are well studied, and they exhibit characteristics that represent a larger group of organisms. **TABLE 12.2** shows some characteristics of the genomes of these organisms.



YEAST: THE BASIC EUKARYOTIC CELL MODEL Yeasts are single-celled eukaryotes. Like other eukaryotes, they have membrane-enclosed organelles. They can live as either haploid or diploid organisms, and this is usually determined by environmental conditions: under adverse conditions the diploid cells will undergo meiosis and make spores. Whereas the prokaryote *E. coli* has a single circular chromosome and 4,288 protein-coding genes, *Saccharomyces cerevisiae* has 16 linear chromosomes and 6,275 protein-coding genes. The most striking difference between the yeast genome and that of *E. coli* is in the number of genes involved in secretion or targeting proteins to specific locations within the cell: yeast has 430 such genes; *E. coli* has only 35. Both of these single-celled organisms appear to use about the same number of genes to perform the basic functions of cell survival. It is the compartmentalization of the eukaryotic yeast cell into organelles that requires it to have many more genes. This finding is direct, quantitative confirmation of something we have known for a century: the eukaryotic cell is structurally and functionally more complex than the prokaryotic cell.

THE NEMATODE: UNDERSTANDING CELL DIFFERENTIATION The 1-millimeter-long nematode *Caenorhabditis elegans* normally lives in the soil. It can also live in the laboratory, where it has become a favorite model

TABLE 12.2 Representative Sequenced Genomes

Organism	Haploid genome size (Mb)[a]	Number of protein-coding genes	Percent of genome that codes for proteins	Notable attributes
BACTERIA				
Mycoplasma genitalium	0.58	482	88	Minimal genome
Haemophilus influenzae	1.83	1,727	89	
Escherichia coli	4.6	4,288	88	Well-studied enteric bacterium
YEASTS				Targeting; cell organelles
Saccharomyces cerevisiae	12.2	6,275	70	
Schizosaccharomyces pombe	13.8	4,824	60	
PLANTS				Photosynthesis; cell walls
Arabidopsis thaliana	125	27,416	25	Small plant genome
Oryza sativa (rice)	420	40,838	12	Water tolerance for roots
Glycine max (soybean)	973	46,430	7	Lipid synthesis, storage
ANIMALS				
Caenorhabditis elegans (nematode)	100	20,470	25	Tissue formation
Drosophila melanogaster (fruit fly)	140	13,733	13	Embryonic development
Homo sapiens (human)	3,200	~21,000	1.2	Language

[a]Mb = millions of base pairs

organism of developmental biologists (see Chapter 14). The nematode has a transparent body of about 1,000 somatic cells. It develops over 3 days from a fertilized egg to an adult worm that has a nervous system, digests food, and reproduces sexually. Its genome is 8 times larger than that of yeast and has about 3.3 times as many protein-coding genes. Many of these extra genes encode proteins needed for cell differentiation, for intercellular communication, and for holding cells together to form tissues.

DROSOPHILA MELANOGASTER: UNDERSTANDING GENETICS AND DEVELOPMENT The fruit fly *Drosophila melanogaster* is a famous model organism. Studies of fruit fly genetics resulted in the formulation of many basic principles of genetics (see Concept 8.3). The fruit fly is a much larger organism than *C. elegans* (it has ten times as many cells), and it is much more complex: it undergoes complicated developmental transformations from egg to larva to pupa to adult. These differences are reflected in the fruit fly genome, which has many genes encoding transcription factors needed for complex embryonic development (you will study some of these in Chapter 14). In general, the fruit fly genome has a distribution of coding sequence functions quite similar to those of many other complex eukaryotes (**FIGURE 12.9**).

ARABIDOPSIS: STUDYING THE GENOMES OF PLANTS About 250,000 species of flowering plants dominate the land and fresh water. Although there is generally more interest in the plants we use for food and fiber, scientists first sequenced the genome of a simpler flowering plant with a relatively small genome. *Arabidopsis thaliana*, thale cress, is a member of the mustard family and has long been a favorite model organism of plant biologists. It is small (hundreds could grow and reproduce in the space occupied by this page) and easy to manipulate. Its genome has about 27,400 protein-coding genes, but many of these are duplicates and probably originated by chromosomal rearrangements. When these duplicate genes are subtracted from the total, about 15,000 unique genes are left—similar to the gene numbers found in fruit flies and nematodes. Indeed, many of the genes found in these animals have **orthologs**—genes that are derived from a common ancestral gene—in *Arabidopsis* and other plants, supporting the idea that plants and animals have a common ancestor.

Arabidopsis has some genes, however, that are unique to plants. These include genes involved in photosynthesis, in the transport of water throughout the plant, in the assembly of the cell wall, in the uptake and metabolism of inorganic substances from the environment, and in the synthesis of specific molecules used for defense against microbes and herbivores. These plant defense molecules may be a major reason why the numbers of protein-coding genes in some plants are higher than in many animals. Plants cannot escape their enemies or other adverse conditions as animals can, and they must cope

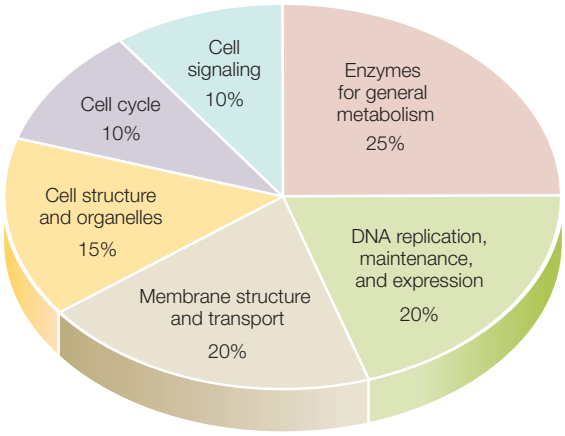


FIGURE 12.9 Functions of the Eukaryotic Genome The distribution of gene functions in *Drosophila melanogaster* shows a pattern that is typical of that of many complex organisms.

with situations where they are. So they make tens of thousands of molecules to help them fight their enemies and adapt to their changing environments (see Chapter 28).

These plant-specific genes are also found in the genomes of other plants, including rice (*Oryza sativa*), the first major crop plant to be fully sequenced. Rice is the world's most important crop—it is a staple in the diets of 3 billion people. The larger genome of rice has a set of genes remarkably similar to that of *Arabidopsis*. The genome of the poplar tree *Populus trichocarpa* was also sequenced, to gain insight into the potential for this rapidly growing tree to be used as a source of fuel. Several more plant genomes have now been sequenced. Comparisons among diverse flowering plant species (including *Arabidopsis*, rice, poplar, and sorghum) suggest that about 12,000 protein-coding genes are shared among all flowering plants (**FIGURE 12.10**). These may comprise the basic plant genome.

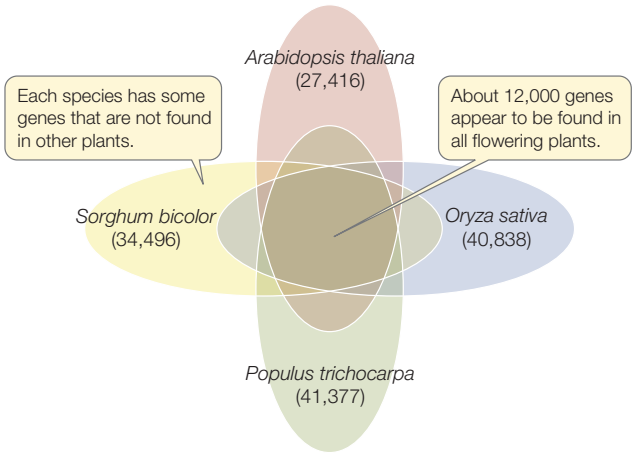


FIGURE 12.10 Plant Genomes Four plant genomes share a common set of approximately 12,000 genes that may comprise the "minimal" plant genome. The total numbers of protein-coding genes are shown in parentheses.

Gene families exist within individual eukaryotic organisms

About half of all eukaryotic protein-coding genes exist as only one copy in the haploid genome (with two alleles in diploid somatic cells). The other half are present in multiple copies that arose from gene duplications. Over evolutionary time, different copies of the genes have undergone separate mutations, giving rise to groups of closely related genes called **gene families**. Some gene families, such as those encoding the globin proteins that make up hemoglobin, contain only a few members in a single organism. Other families, such as the genes encoding the immunoglobulins that make up antibodies, have hundreds of members.

Within a single organism, the genes in a family are usually slightly different from one another. As long as at least one member encodes a functional protein, the other members may mutate in ways that change the functions of the proteins they encode. For evolution, the availability of multiple copies of a gene allows for selection of mutations that provide advantages under certain circumstances. If a mutated gene is useful, it may be selected for in succeeding generations. If the mutated gene is a total loss, the functional copy is still there to carry out its role.

As an example, let's look at the gene family encoding the globins in vertebrates. These proteins are found in hemoglobin and myoglobin (an oxygen-binding protein present in muscle). The globin genes all arose long ago from a single common ancestral gene (see Figure 15.23). In humans there are three functional members of the α-globin cluster and five in the β-globin cluster (**FIGURE 12.11**). Each hemoglobin molecule in an adult human is a tetramer containing two identical α-globin subunits, two identical β-globin subunits, and four heme pigments (see Chapter 3, p. 47).

During human development, different genes of the globin gene cluster are expressed at different times and in different tissues. This differential gene expression has great physiological significance. For example, the hemoglobin in the human fetus contains γ-globin, which binds O_2 more tightly than adult hemoglobin does. This specialized form of hemoglobin ensures that in the placenta, O_2 is transferred from the mother's blood to the developing fetus's blood. Just before birth the liver stops synthesizing fetal hemoglobin and the bone marrow cells take over, making the adult forms (two α and two β). Thus hemoglobins with different binding affinities for O_2 are provided at different stages of human development.

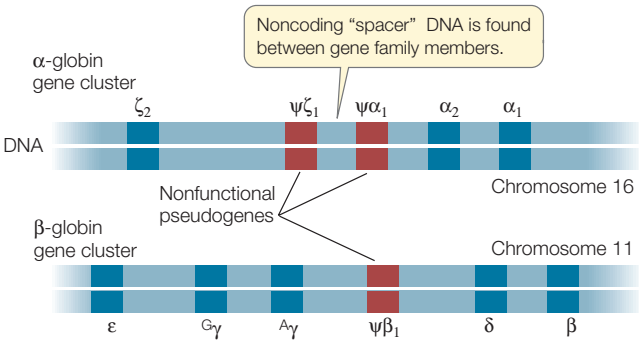


FIGURE 12.11 The Globin Gene Family The α-globin and β-globin clusters of the human globin gene family are located on different chromosomes. The genes of each cluster are separated by noncoding "spacer" DNA. The nonfunctional pseudogenes are indicated by the Greek letter psi (ψ). The γ gene has two variants, $^{A}\gamma$ and $^{G}\gamma$.

LINK

Gene duplication plays a role in the evolution of new protein functions, as described in **Concept 15.6**

In addition to genes that encode functional proteins, many gene families include nonfunctional **pseudogenes**, which are designated with the Greek letter psi (ψ; see Figure 12.11). These pseudogenes result from mutations that cause a loss of function rather than an enhanced or new function. The DNA sequence of a pseudogene may not differ greatly from that of other family members. It may simply lack a promoter, for example, and thus fail to be transcribed. Or it may lack a recognition site needed for the removal of an intron, so that the transcript it makes is not correctly processed into a useful mature mRNA. In some gene families pseudogenes outnumber functional genes. Because some members of the family are functional, there appears to be little selection pressure to preserve the functions of these pseudogenes.

TABLE 12.3 Types of Sequences in Eukaryotic Genomes

Category	Transcribed	Translated
SINGLE-COPY GENES		
Promoters and expression control sequences	No	No
Introns	Yes	No
Exons	Yes	Yes
MODERATELY REPETITIVE SEQUENCES		
rRNA and tRNA genes	Yes	No
Transposons		
I. Retrotransposons		
LTR retrotransposons	Yes	No
SINEs	Yes	No
LINEs	Yes	Yes
II. DNA transposons	Yes	Yes
HIGHLY REPETITIVE SHORT SEQUENCES	No	No

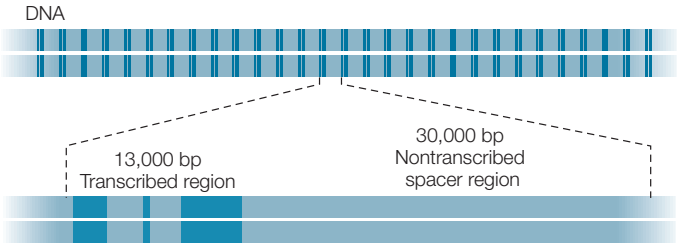
Eukaryotic genomes contain many repetitive sequences

Eukaryotic genomes contain numerous repetitive DNA sequences that do not code for polypeptides. There are highly repetitive sequences and moderately repetitive sequences, which include rRNA genes, tRNA genes, and transposons (**TABLE 12.3**).

Highly repetitive sequences are short (less than 100 bp) sequences that are repeated

thousands of times in tandem (side-by-side) arrangements in the genome. They are not transcribed. Their proportion in eukaryotic genomes varies, from 10 percent in humans to about half the genome in some species of fruit flies. Often they are associated with heterochromatin, the densely packed, largely transcriptionally inactive part of the genome (see Concept 11.3). Other highly repetitive sequences are scattered around the genome. For example, **short tandem repeats** (**STRs**) of 1–5 bp can be repeated up to 100 times at a particular chromosomal location. The copy number of an STR at a particular location varies among individuals and is inherited.

Moderately repetitive sequences are repeated 10–1,000 times in the eukaryotic genome. These sequences include the genes that are transcribed to produce tRNAs and rRNAs, which are used in protein synthesis. The cell makes tRNAs and rRNAs constantly, but even at the maximum rate of transcription, single copies of the tRNA and rRNA genes would be inadequate to supply the large amounts of these molecules needed by most cells. Thus the genome has multiple copies of these genes, in clusters containing transcribed regions (with introns) and nontranscribed "spacers" between the genes. Here, for example, is a region that encodes multiple sets of rRNA genes (in dark blue):



Most moderately repetitive sequences are not stably integrated into the genome but instead are transposons (see Figure 12.7). Transposons make up more than 40 percent of the human genome. There are two main types of transposons in eukaryotes: retrotransposons (Class I transposons) and DNA transposons (Class II; see Table 12.3).

Retrotransposons make RNA copies of themselves, which are then copied back into DNA before insertion at new locations in the genome. They are divided into two categories:

- *LTR retrotransposons* have long terminal repeats (LTRs) of DNA sequence (100–5,000 bp) at each end. LTR retrotransposons constitute about 8 percent of the human genome.
- *Non-LTR retrotransposons* do not have LTR sequences at their ends. They are further divided into two subcategories: SINEs and LINEs. SINEs (*s*hort *in*terspersed *e*lements) are up to 500 bp long and are transcribed but not translated. There are about 1.5 million of them scattered over the human genome, making up about 15 percent of the total DNA content. A single type, the 300-bp *Alu* element, accounts for 11 percent of the human genome; it is present in a million copies. LINEs (*l*ong *in*terspersed *e*lements) are up to 7,000 bp long, and some are transcribed and translated into proteins. They constitute about 17 percent of the human genome.

DNA transposons do not use RNA intermediates. Like some prokaryotic transposons, they are excised from the original location and become inserted at a new location without being replicated.

With so much of the genome made up of transposons, they must have a role in addition to just replicating themselves ("selfish DNA"). One possibly important function occurs when such a sequence moves to a new location within the coding region of a gene and causes mutation by disrupting it. Mutations are the raw material of evolution by natural selection. In most cases, the ability to move about the genome has been suppressed, so transposons are stable and do not move. Some occur within introns, where they may affect alternative splicing of pre-mRNA, causing phenotypic diversity. Others are at or near gene regulatory sequences such as promoters, where they can also affect gene expression.

CHECKpoint CONCEPT 12.3

- ✓ Compare the general properties of the genomes of prokaryotes and eukaryotes.
- ✓ Does the size of a genome determine how much information it contains? Explain in terms of repetitive sequences and protein-coding genes.
- ✓ What is the evolutionary role of eukaryotic gene families?
- ✓ During transposition, an adjacent gene is sometimes transposed along with a retrotransposon. What would be the consequence of making a new copy of this gene at a new location in the genome?

The analysis of eukaryotic genomes has resulted in an enormous amount of useful information, as we have seen. In the next concept we will look more closely at the human genome.

CONCEPT 12.4 The Human Genome Sequence Has Many Applications

During the first decade of this millennium the haploid genomes of more than ten individuals were sequenced and published. With the rapid development of new sequencing technologies, the time is approaching when a human genome can be sequenced for less than $1,000.

The human genome sequence held some surprises

The following are just some of the interesting facts we have learned about the human genome:

- Among the 3.2 billion bp in the haploid human genome, there are about 21,000 protein-coding genes. This was a surprise. Before sequencing began, the diversity of human proteins suggested there would be 80,000–150,000 genes. The actual number—not many more than in a nematode—means that posttranscriptional mechanisms (such as alternative splicing) must account for the observed number of

proteins in humans. It turns out that most human genes encode multiple proteins.

- The average protein-coding gene spans 27,000 bp, and virtually all genes have many introns. Gene sizes vary greatly, from about 1,000 to 2.4 million bp. Variation in gene size is to be expected given that human proteins vary in size, from about 100 to 5,000 amino acids per polypeptide chain.
- More than 50 percent of the genome is made up of transposons and other repetitive sequences. Most transposons are inactive most of the time.
- About 75 percent of the genome is transcribed at some point in some cells. This result came from a recent analysis of genome expression in human cells in culture. A typical specialized cell only transcribes about 25 percent of its genome, so most transcripts are cell-type specific. Many transcripts are noncoding RNAs involved in regulating gene expression.
- Most of the genome (at least 99 percent) is the same in all people. Despite this apparent homogeneity, there are, of course, many individual differences. Current estimates suggest that each haploid genome has variations in about 3.3 million single nucleotide polymorphisms (SNPs; see below), as well as short repeated sequences that are variable in repeat number.
- An individual's genome changes over time, in specific sets of cells, as new mutations occur. These changes can be important if they affect cell function. For example, a mutation in a gene that blocks cell division may result in reduced expression of that gene, and cancer can then occur.

Comparisons among sequenced genomes from prokaryotes and eukaryotes have revealed some of the evolutionary relationships among genes. Some genes are present in both prokaryotes and eukaryotes; others are only in eukaryotes; still others are only in animals, or only in vertebrates (**FIGURE 12.12**).

More comparative genomics is possible now that the genomes of other primates, including all the great apes, have been sequenced. Chimpanzees are our closest living relatives, sharing nearly 99 percent of our DNA sequence. Gorillas and orangutans are next closest, with genomes that are about 98 percent and 97 percent similar to ours. Researchers have identified about 500 protein-coding genes that have undergone accelerated evolution in humans, chimpanzees, and gorillas, including genes involved in hearing and brain development. Further analyses of these sequences may reveal genes that distinguish us from other apes and that "make humans human."

Go to MEDIA CLIP 12.1
A Big Surprise from Genomics
PoL2e.com/mc12.1

Human genomics has potential benefits in medicine

Complex phenotypes are determined not by single genes but by multiple genes interacting with the environment. A single disease-causing allele, such as those associated with phenylketonuria and sickle-cell anemia (see Concept 9.3), does not exist for such common disorders as diabetes, heart disease, and Alzheimer's disease. To understand the genetic bases of these diseases, biologists are now using rapid genotyping technologies to create "haplotype maps" that can be used to identify multiple genes involved in disease.

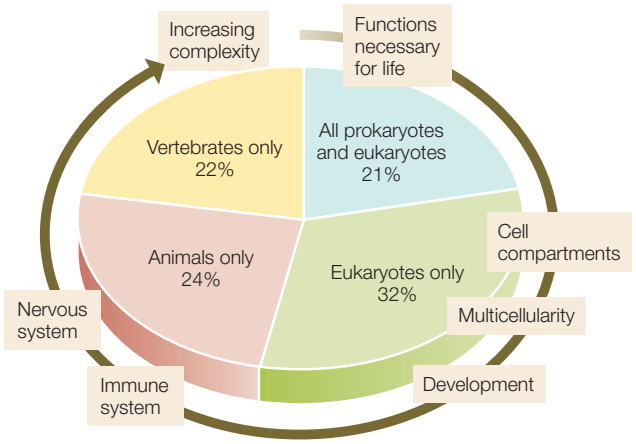


FIGURE 12.12 Evolution of the Genome A comparison of the human and other genomes has revealed how genes with new functions have been added over the course of evolution. Each percentage number refers to genes in the human genome. Thus 21 percent of human genes have orthologs in prokaryotes and other eukaryotes, 32 percent of human genes occur in all other eukaryotes, and so on.

HAPLOTYPE MAPPING Haplotype maps are based on **single nucleotide polymorphisms** (**SNPs**)—DNA sequence variations that involve single nucleotides. SNPs (pronounced "snips") arise as point mutations (see Concept 9.3). Because of these mutations, a single nucleotide in a homologous DNA sequence may vary among individuals or even between alleles in a single individual. Biologists use SNPs to create genetic maps of organisms, to classify organisms and species, and to identify individual organisms carrying specific alleles.

The SNPs that differ among individuals are not necessarily inherited as independent alleles. Rather, a set of SNPs that are close together on a chromosome is inherited as a unit (the SNPs are tightly linked). A piece of chromosome with a set of linked SNPs is called a **haplotype**. You can think of the haplotype as a sentence and the SNP as a word in the sentence. Analyses of haplotypes in humans from all over the world have thus far identified 500,000 common variations.

GENOTYPING TECHNOLOGY AND PERSONAL GENOMICS New technologies are continually being developed to analyze thousands or millions of SNPs in the genomes of individuals. Such technologies include high-throughput sequencing methods and DNA microarrays, which depend on hybridization to identify specific SNPs.

A **DNA microarray** is a grid of microscopic spots of oligonucleotides (short DNA sequences) arrayed on a solid surface. It can be "probed" with a complex mixture of DNA or RNA; if the mixture contains a sequence that is complementary to one of the oligonucleotides, the sequence will hybridize to that spot. Colored fluorescent dyes are used to detect spots that hybridize with components of the probe mixture. The specific

pattern of fluorescent dots reveals the haplotypes of the individual from whom the DNA came. For example, a microarray of 500,000 SNP-containing oligonucleotides has been used to analyze DNA from thousands of people to find out which SNPs are linked to genes associated with specific diseases. The aim is to identify particular alleles that contribute (along with particular alleles of other genes) to each complex disease (**FIGURE 12.13**). The amount of data from 500,000 SNPs and thousands of people with thousands of medical records is prodigious. With so much natural variation, statistical measures of association between a haplotype and a disease need to be very rigorous.

These association tests have revealed haplotypes or alleles that are associated with modestly increased risks for such diseases as breast cancer, diabetes, arthritis, obesity, and coronary heart disease. For example, 12 SNPs are associated with increased incidence of heart attacks, and if considered together, these SNPs can be used to identify individuals who are at increased risk. Indeed, the predictive value of this genetic test is greater than the widely used test for elevated blood cholesterol level. Private companies now offer to scan a human genome for SNP alleles, and the price for this service keeps getting lower. However, at this point it is unclear what a person without symptoms should do with the information, since multiple genes, environmental influences, and epigenetic effects all contribute to the development of these diseases.

PHARMACOGENOMICS Genetic variation can affect how an individual responds to a particular drug. For example, consider an enzyme in the liver that catalyzes the following reaction:

$$\text{Active drug} \xrightarrow{\text{Enzyme A}} \text{less active drug}$$

A mutation in the gene that encodes this enzyme may make the enzyme less active and reduce the rate at which the active drug is modified to a less active form. For a given dose of the drug, a person with the mutation would have more active drug in his or her bloodstream than a person without the mutation. So the dose of the drug needed for the same effect would be lower for this person.

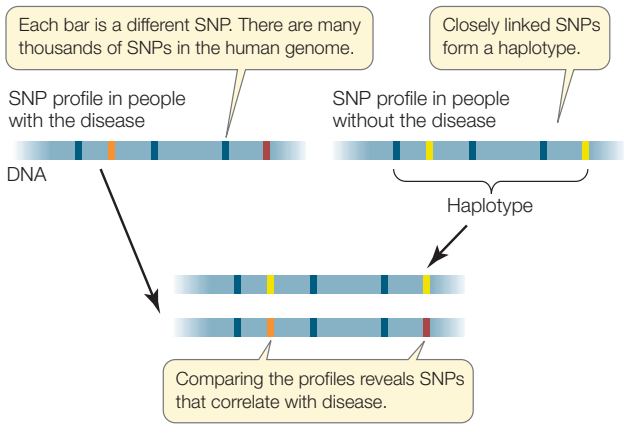


FIGURE 12.13 SNP Genotyping and Disease Scanning the genomes of people with and without particular diseases reveals correlations between SNPs and complex diseases.

Now consider a different case, in which a liver enzyme is needed to make the drug active:

$$\text{Inactive drug} \xrightarrow{\text{Enzyme B}} \text{active drug}$$

A person homozygous for a mutation in the gene encoding this enzyme would not be affected by the drug, since the activating enzyme is not present.

The study of how an individual's genome affects his or her response to drugs or other agents is called **pharmacogenomics**. This type of analysis makes it possible to predict whether or not a drug will be effective. The objective is to personalize drug treatment so that a physician can know in advance whether an individual will benefit from a particular drug (**FIGURE 12.14**). This approach might also be used to reduce the incidence of adverse drug reactions in individuals who metabolize particular drugs slowly.

PROTEOMICS Comparisons of the proteomes of humans and other eukaryotic organisms have revealed a common set of proteins that can be categorized into groups (families) with similar amino acid sequences and similar functions. Forty-six percent of the yeast proteome, 43 percent of the worm proteome, and 61 percent of the fruit fly proteome are shared by the human proteome. Functional analyses indicate that this set of 1,300 proteins provides the basic metabolic functions of a eukaryotic cell—including glycolysis, the citric acid cycle, membrane transport, protein synthesis and targeting, and DNA replication. Of course, these are not the only human

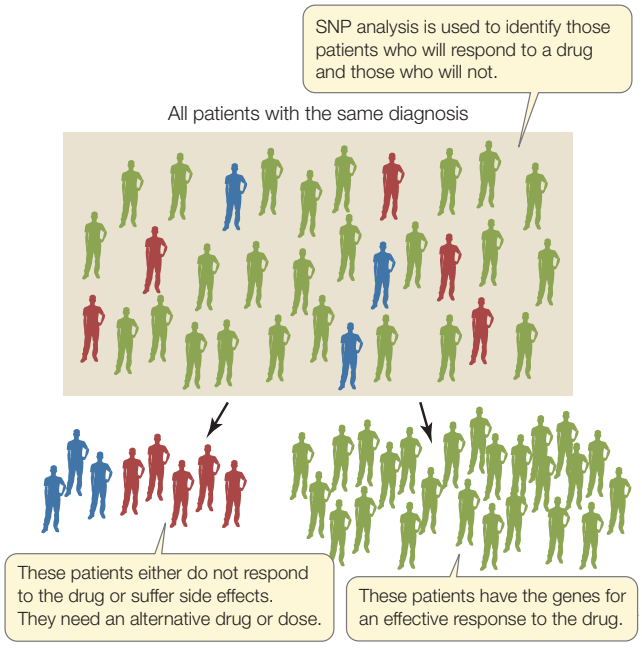


FIGURE 12.14 Pharmacogenomics Correlations between genotypes and responses to drugs will help physicians develop more personalized medical care. SNP analysis is used to identify people who will respond to a drug and those who will not. The different colors indicate individuals with different SNPs.

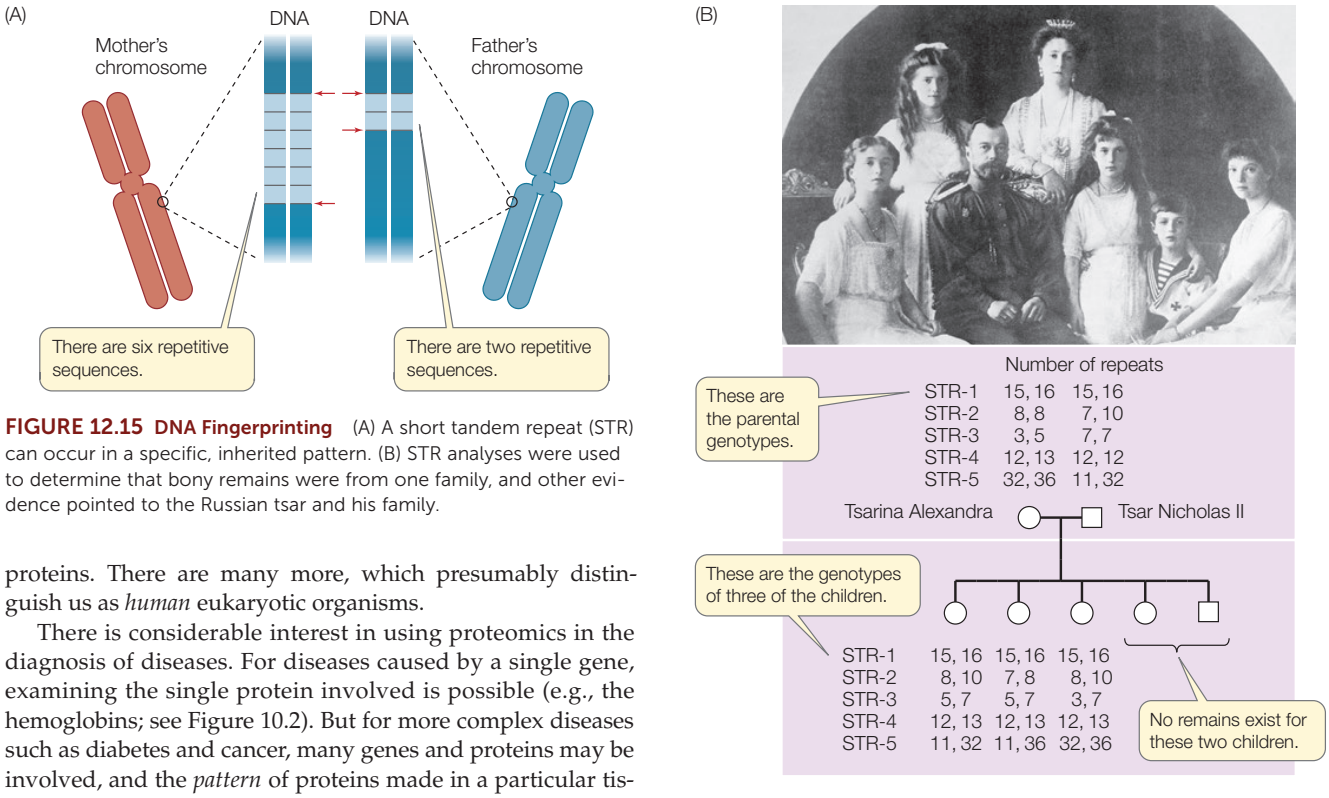


FIGURE 12.15 DNA Fingerprinting (A) A short tandem repeat (STR) can occur in a specific, inherited pattern. (B) STR analyses were used to determine that bony remains were from one family, and other evidence pointed to the Russian tsar and his family.

proteins. There are many more, which presumably distinguish us as *human* eukaryotic organisms.

There is considerable interest in using proteomics in the diagnosis of diseases. For diseases caused by a single gene, examining the single protein involved is possible (e.g., the hemoglobins; see Figure 10.2). But for more complex diseases such as diabetes and cancer, many genes and proteins may be involved, and the *pattern* of proteins made in a particular tissue at a certain time might indicate the presence or likelihood of the disease. Proteomic analyses of tissues (including blood) are being developed to provide early warnings of the presence of particular diseases before symptoms occur.

METABOLOMICS There has been some progress in defining the human metabolome. A database created by David Wishart and colleagues at the University of Alberta contains more than 40,000 entries, including metabolic products of foods and drugs. The challenge now is to relate the levels of these substances to physiology. For example, high levels of glucose in the blood are associated with diabetes, and there may be patterns of metabolites that are diagnostic of other diseases. This could aid in early diagnosis and treatment.

DNA fingerprinting uses short tandem repeats

As noted in Concept 12.3, short tandem repeats (STRs) are blocks of 1–5 bp that can be repeated up to 100 times at particular locations on chromosomes. Since the number of repeats can vary widely, there are usually numerous alleles for a particular STR. For example, at a particular location on human chromosome 15 there might be an STR of AGG. An individual might inherit an allele with six copies of the repeat (AGGAGGAGGAGGAGG) from her mother and an allele with two copies (AGGAGG) from her father (**FIGURE 12.15A**). PCR can be used to amplify DNA fragments containing these repeat sequences, and the number of copies of the repeat can be determined by sizing the DNA fragments.

DNA fingerprinting refers to a group of techniques used to identify particular individuals by their DNA; the most common of these techniques involves STR analysis. When several different STR loci are analyzed, an individual's unique pattern becomes apparent. The U.S. Federal Bureau of Investigation uses 13 STR loci in its Combined DNA Index System (CODIS) database.

DNA fingerprinting is used to resolve questions of paternity, and in forensics (crime investigations) to identify criminals. It also has other uses—for example, it can help in analyses of historical events. In 1918 the Russian tsar Nicholas II and his wife and five children were killed during the Communist revolution. A report that the bodies had been burned was not questioned until 1991, when a shallow grave with several skeletons was discovered a few miles from the presumed execution site. The remains were from a man, a woman, and three female children, and STR analysis indicated that they were all related to one another (**FIGURE 12.15B**). These STR patterns were also related to living descendants of the tsar. The accuracy and specificity of this method gave historical and cultural closure to a major event of the twentieth century.

Genome sequencing is at the leading edge of medicine

With the rapid development of better and better ways to sequence genomes cheaply, we are on the verge of a new era in medicine. The genomes of random cells of people, as well as cells from diseases such as cancers, are being sequenced in two ways:

- Sequencing of the entire genome
- Sequencing of protein-coding exons only

APPLY THE CONCEPT

The human genome sequence has many applications

It is the year 2025. You are taking care of a patient who is worried that he may have an early stage of kidney cancer. His mother died from this disease.

1. Assume that the SNPs linked to genes involved in the development of this type of cancer have been identified. How would you determine if this man has a genetic predisposition for developing kidney cancer? Explain how you would do the analysis.
2. How might you develop a metabolomic profile for kidney cancer and then use it to determine whether your patient has kidney cancer?
3. If the patient is diagnosed with the cancer, how would you use pharmacogenomics to choose the right medications to treat his tumor?

The information these methods provide goes far beyond what SNP and STR genotyping can do. For example:

- The genome sequence of a tumor included mutations in specific genes that drive tumor formation. The proteins encoded by these genes are targets for specific therapy for that tumor.
- The genome sequence of an individual with a genetic disease has led to the identification of the causative gene.

CHECKpoint CONCEPT 12.4

- ✓ The average human gene spans 27,000 bp. The average human polypeptide has 300 amino acids. Explain.
- ✓ What is a haplotype with regard to an STR? How can haplotypes be used to relate DNA to a phenotype?
- ✓ A person has a rare allele for an STR (STR-1) that has a frequency in the population of 1 percent (0.01). The same person has allele frequencies for other STRs as follows: STR-2, 0.005; STR-3, 0.01; STR-4, 0.05; STR-5, 0.01. What is the probability that an individual will have all of these alleles? What does this mean in terms of identifying an individual with this genotype?

Q What does genome sequencing reveal about dogs and other animals?

ANSWER In the opening story of this chapter we described how genome sequencing is being applied to breeds of dogs. For example, high-throughput sequencing methods (Concept 12.1) have allowed biologists to collect data on genes that control body size. This has led to the identification of a SNP (Concept 12.4) in the gene for *i*nsulin-like *g*rowth *f*actor 1 (IGF-1) that is important in determining size. Large breeds have an allele that encodes an active IGF-1, and small breeds have a different allele that encodes a less active version of the protein. In humans, IGF-1 mediates the overall effects of growth hormone, and people with a mutation in the IGF-1 gene have short stature.

Another gene important to phenotypic variation is found in whippets, sleek dogs that run fast and are often raced. A mutation in the gene for myostatin, a protein that inhibits the overdevelopment of muscles, results in a whippet that is more muscular and runs faster (**FIGURE 12.16A**). Comparative genomics shows that this gene is important in other animals as well. In Belgian Blue cattle, individuals homozygous for a particular SNP in the myostatin gene have huge muscles

(A)

(B)

FIGURE 12.16 Muscular Gene (A) These dogs are both whippets, but the muscle-bound dog (left) has a mutation in a gene that normally limits muscle buildup. (B) A similar mutation in Belgian Blue cattle also leads to overgrown muscles.

(FIGURE 12.16B). There is interest in applying this knowledge to humans. For example, in muscular dystrophy the skeletal muscles waste away, and blocking myostatin could be useful in keeping muscles robust. Athletes anxious to have bulkier muscles have also been focusing on this gene and its protein product.

Inevitably, some scientists have set up companies to test dogs for genetic variations using DNA supplied by dog owners and breeders. The black dog rescued from the pound that looks like a Labrador retriever may turn out to be a German pointer. Some traditional breeders frown on this practice, but others say it will bring more joy (and prestige) to owners.

SUMMARY

CONCEPT 12.1 There Are Powerful Methods for Sequencing Genomes and Analyzing Gene Products

- To sequence a genome, the chromosomes are cut into overlapping fragments, which are sequenced. Then the fragment sequences are lined up to assemble the DNA sequence of the chromosome. See ANIMATED TUTORIAL 12.1
- **High-throughput sequencing** involves attaching short, single-stranded DNA fragments to a solid surface. A primer and DNA polymerase are added, and tagged nucleotides are detected by a camera as they are added to the complementary DNA strand. Many sequences can be done in parallel. Review Figure 12.1 and ANIMATED TUTORIAL 12.2
- The analysis of DNA sequences is done by computer. Genomic sequences include protein-coding genes, RNA genes, regulatory sequences, and repeated sequences. Review Figure 12.3
- The **proteome** is the total protein content of an organism. It can be analyzed using chemical methods that separate and identify proteins. These include two-dimensional electrophoresis, mass spectrometry, and techniques involving antibodies. Review Figure 12.4
- The **metabolome** is the total content of small molecules in a tissue under particular conditions. These molecules include intermediates in metabolism, hormones and other signaling molecules, and secondary metabolites. Review Figure 12.5

CONCEPT 12.2 Prokaryotic Genomes Are Small, Compact, and Diverse

- Prokaryotic genomes have been studied using **functional genomics** to determine the roles of various parts of the genome, including the protein-coding genes. **Comparative genomics** is used to compare sequences among organisms. Review Table 12.1
- **Metagenomics** is the identification of DNA sequences in environmental samples without first isolating, growing, and identifying the organisms. Review Figure 12.6
- **Transposons** are sequences of DNA that can move about the genome. Review Figure 12.7
- Transposon mutagenesis can be used to inactivate genes one by one. Then the organism can be tested for survival. In this way, functionally important genes can be identified. Review Figure 12.8

CONCEPT 12.3 Eukaryotic Genomes Are Large and Complex

- Sequences from model organisms have highlighted some common features of eukaryotic genomes. In addition, there are specialized genes such as those for cellular compartmentalization, development, and features unique to plants. Review Figures 12.9 and 12.10 and Table 12.2
- Some genes exist as members of **gene families**. Proteins may be made from these closely related genes at different times and in different tissues. Review Figure 12.11
- Eukaryotic genomes contain various kinds of repeated sequences. Review Table 12.3

CONCEPT 12.4 The Human Genome Sequence Has Many Applications

- The haploid human genome has 3.2 billion bp.
- Only 1.5 percent of the genome codes for proteins; much of the rest consists of repeated sequences.
- Most of the genome is transcribed at some point in some cells.
- Virtually all human genes have introns, and alternative splicing leads to the production of more than one protein per gene.
- Genotyping using **single nucleotide polymorphisms** (**SNPs**) can be used to correlate variations in the genome with diseases or drug sensitivity. It may lead to personalized medicine. Review Figure 12.13
- **Pharmacogenomics** is the analysis of how a person's genetic makeup affects his or her drug metabolism. Review Figure 12.14
- Short tandem repeats (STRs) are DNA sequences that are variable in length. They can be used to identify individuals. Review Figure 12.15

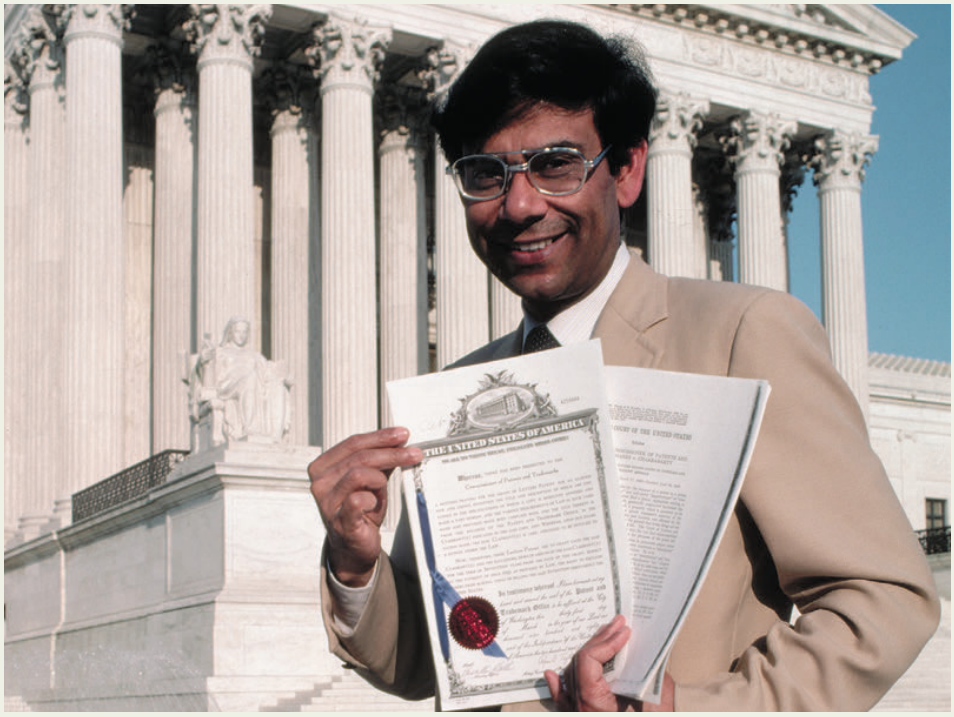
See ACTIVITY 12.1 for a concept review of this chapter.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is12

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

13 Biotechnology

KEY CONCEPTS



Ananda Chakrabarty received the first patent for a genetically modified organism, a bacterium that breaks down crude oil.

The United Nations defines **biotechnology** as "any technological application that uses biological systems, living organisms, or derivatives thereof to make or modify products or processes." This definition encompasses major human activities such as brewing beer (see Chapter 6) and the domestication of animals and plants (see Chapter 12). More recently, biotechnology has become associated with the genetic modification of microorganisms for the production of particular substances, and of a variety of plants and animals used in agriculture.

Industrial biotechnology began in England in 1917, during the First World War. The production of cordite, an explosive used to propel a bullet or shell to its target, required the solvent acetone, $(CH_3)_2CO$. But most acetone was manufactured by England's enemy, Germany. A microbiologist at the University of Manchester, Chaim Weizmann, found that if the bacterium *Clostridium acetylbutylicum* was grown using starch as an energy source, it produced abundant quantities of acetone. The British government set up a factory to grow large vats of these bacteria, and the cordite shortage was solved.

The contemporary era of biotechnology as a major industry dates from June 16, 1980, on the steps of the U.S. Supreme Court. In this case, scientists were studying bacteria not for their ability to make something, but to break it down. Many bacteria have genes that code for unusual enzymes and biochemical pathways, and they can use all sorts of substances as nutrients, including pollutants. Scientists have identified these organisms simply by mixing polluted soil with water and seeing what grows. In 1971, Ananda Chakrabarty at the General Electric Research Center in New York used genetic crosses to develop a single strain of the bacterium *Pseudomonas* that carried genes for the breakdown of various hydrocarbons in oil. He and his company applied for a patent to legally protect their discovery and profit from it. Nine years later, in a landmark case, the U.S. Supreme Court ruled that "a live, human-made microorganism is patentable" under the U.S. Constitution.

The 1980 Supreme Court ruling came at a time when new laboratory methods were being developed to insert specific DNA sequences into organisms by recombinant DNA technology. The resulting flood of patents for DNA sequences and genetically modified organisms, some of them developed to improve the environment, continues to this day and was the subject of another Supreme Court ruling in 2013.

Q How is biotechnology used to alleviate environmental problems?

You will find the answer to this question on page 271.

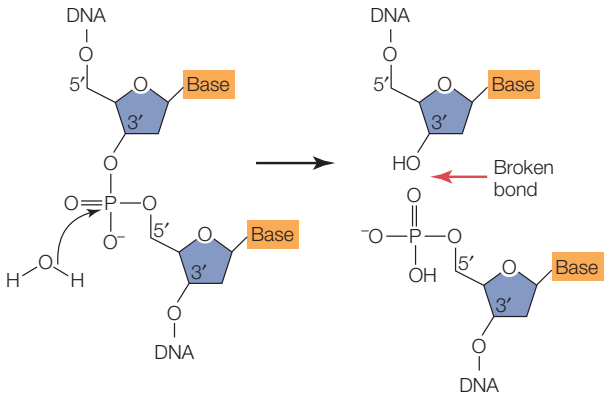
CONCEPT 13.1 Recombinant DNA Can Be Made in the Laboratory

Biotechnology began with the use of organisms with genetic capabilities that occur in nature. For example, existing biochemical pathways in yeast are used to make alcohol, and as noted in the opening story, bacteria can be grown in large quantities to make acetone and other industrial chemicals. More recently, it has become possible to genetically modify organisms with genes from other, distantly related organisms, to create new combinations of genes that would not otherwise occur in the same cell. This technology involves the use of **recombinant DNA**: a single DNA molecule containing DNA sequences from two or more sources. Before the invention of PCR (polymerase chain reaction; see Concept 13.3 and also Concept 9.2), biologists relied on natural molecules and processes to manipulate DNA in the laboratory. Three key tools were, and still are, widely used:

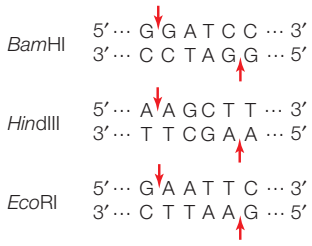
- Restriction enzymes for cutting DNA into pieces (fragments) that can be manipulated
- Gel electrophoresis for the analysis and purification of DNA fragments
- DNA ligase for joining DNA fragments together in novel combinations

Restriction enzymes cleave DNA at specific sequences

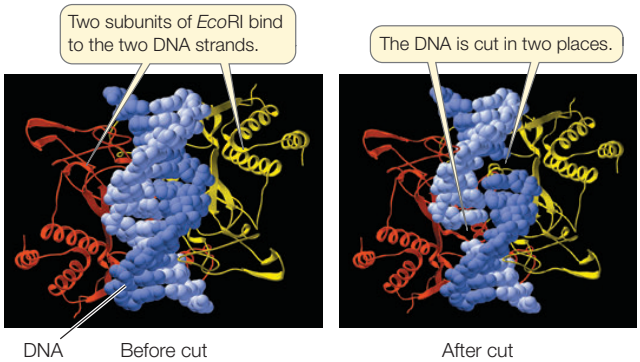
All organisms, including bacteria, must have ways of dealing with their enemies. As we described in Concept 9.1, bacteria are attacked by viruses called bacteriophage (or phage, for short). These viruses inject their genetic material into the host cells and turn them into virus-producing factories, eventually killing the cells. Some bacteria defend themselves against such invasions by producing **restriction enzymes** (also known as restriction endonucleases), which cut double-stranded DNA molecules—such as those injected by bacteriophage—into smaller, noninfectious fragments (FIGURE 13.1). These enzymes break the bonds of the DNA backbone between the 3′ hydroxyl group of one nucleotide and the 5′ phosphate group of the next nucleotide:



There are many different restriction enzymes, each of which cleaves DNA at a specific sequence of bases called a **restriction site** or recognition sequence. Hundreds of these enzymes have been purified from various microorganisms and can be used to cut DNA in the laboratory (by setting up a "restriction digest"). Most restriction sites are four to six base pairs (bp) long, and restriction enzymes catalyze hydrolysis of both strands of DNA. For example, below are three 6-bp sequences, each of which is recognized by a different restriction enzyme. The enzymes cleave the DNA at the sites indicated by the red arrows:



The names of these enzymes reflect their organism of origin. *Eco*RI, for example, was found in *E. coli*, and *Bam*HI comes from the bacterium *Bacillus amyloliquifaciens*. Note that in each of these restriction sites both strands have the same sequence when read from their 5′ ends. This is similar to the word "racecar," which is the same when read in either direction. A word that is the same when read in either direction is called a palindrome. Restriction enzymes have two identical active sites on two subunits, which cleave the two strands simultaneously:



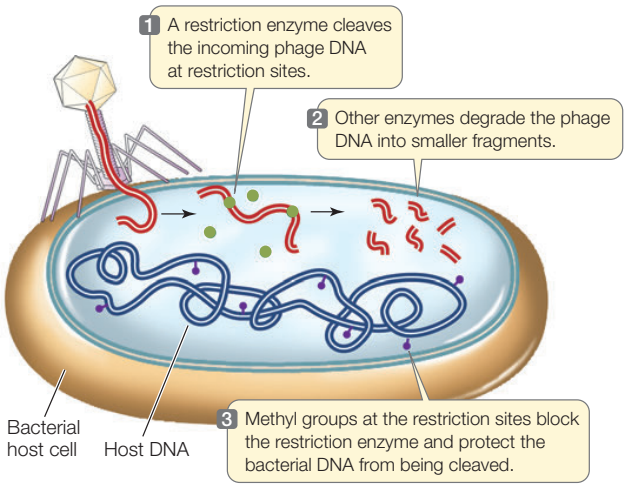
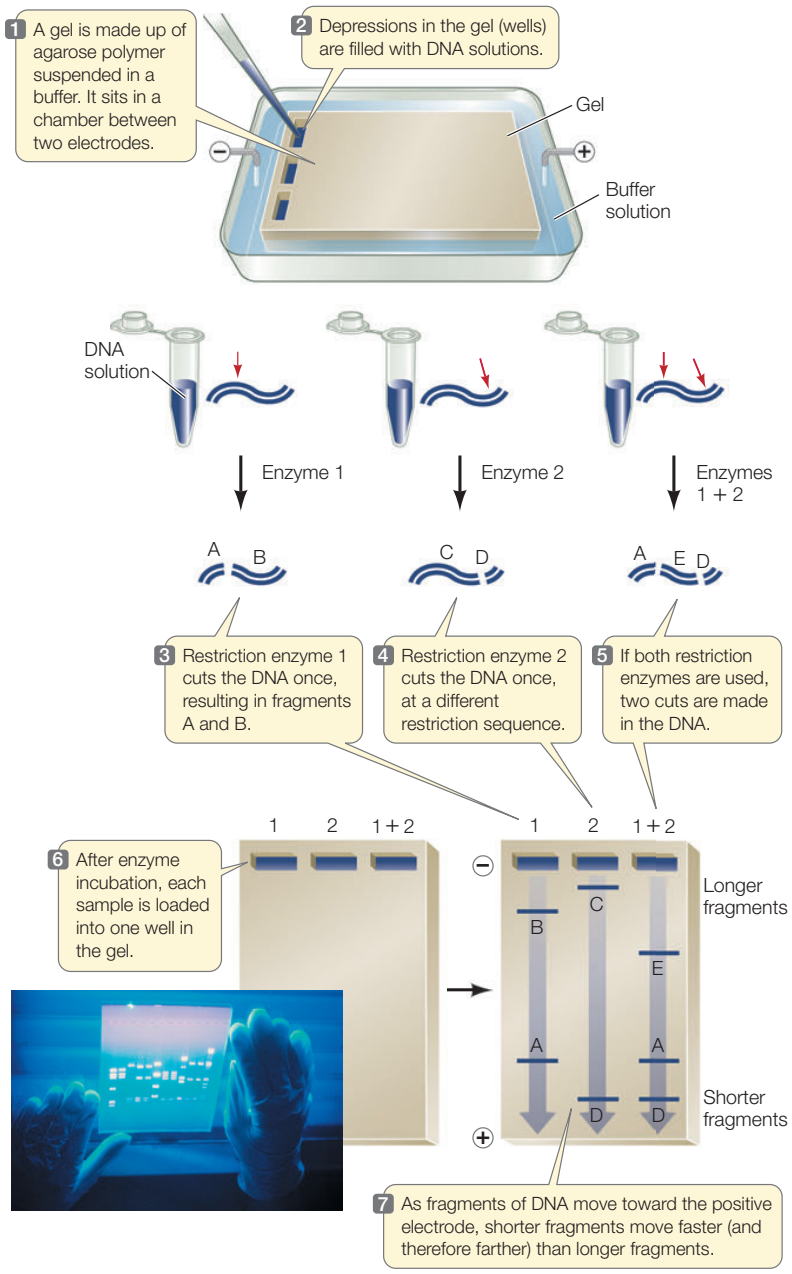


FIGURE 13.1 Bacteria Fight Invading Viruses by Making Restriction Enzymes

RESEARCH TOOLS

FIGURE 13.2 Separating Fragments of DNA by Gel Electrophoresis A mixture of DNA fragments is placed in a gel, and an electric field is applied across the gel. The negatively charged DNA moves toward the positive end of the field, with smaller molecules moving faster (and farther) than larger ones. After minutes to hours for separation, the electric power is shut off and the separated fragments can be analyzed.



Note also that these enzymes cut the two DNA strands in such a way that there will be a short sequence of single-stranded DNA at each cut end. Many restriction enzymes create these single-stranded overhangs, which are referred to as "sticky ends" because they are able to form hydrogen bonds with complementary sequences on other DNA molecules. Other restriction enzymes make cuts directly opposite one another on the two DNA strands, creating "blunt ends."

As shown in Figure 13.1, the restriction enzymes may cleave host bacterial DNA. To prevent cleavage, a "stop sign" in the form of a methyl (—CH_3) group can be placed on restriction sites. This process involves specific DNA methyltransferases. The restriction enzymes do not recognize or cut the methylated restriction sites in the host's DNA. But unmethylated phage DNA is efficiently recognized and cleaved.

The *Eco*RI restriction site recognition sequence occurs, on average, about once in every 4,000 bp in a typical prokaryotic genome, or about once per four prokaryotic genes. So *Eco*RI can chop a large piece of DNA into smaller pieces containing, on average, just a few genes. When *Eco*RI is used in the laboratory to cut a small genome such as that of a virus with, say, 50,000 bp, only a few fragments are obtained. For a huge eukaryotic chromosome with tens of millions of bp, a very large number of fragments are created.

Of course, "on average" does not mean that the enzyme cuts all stretches of DNA at regular intervals. For example, the *Eco*RI restriction site does not occur even once in the 40,000 bp of the T7 phage genome—a fact that is crucial to the survival of this virus, since its host is *E. coli*. Fortunately for *E. coli*, the *Eco*RI restriction site does appear in the DNA of other bacteriophage.

Go to MEDIA CLIP 13.1
Striking Views of Recombinant DNA Being Made
PoL2e.com/mc13.1

Gel electrophoresis separates DNA fragments

After a sample of DNA has been cut with a restriction enzyme, the fragments can be separated from each other to determine the number of fragments and the size (in bp) of each fragment. In this way an individual fragment can be identified, and it can then be purified for further analysis or for use in an experiment.

A convenient way to separate or purify DNA fragments is by **gel electrophoresis**. Samples containing the fragments are placed in wells at one end of a semisolid gel (usually made of agarose or polyacrylamide polymers), and an electric field is applied to the gel (**FIGURE 13.2**). Because of its phosphate groups, DNA is

APPLY THE CONCEPT

Recombinant DNA can be made in the laboratory

The specificity of restriction enzyme recognition can be used to detect mutations. For example, the enzyme *Mst*II cuts DNA at CCTNAGG, where N is any base. Around the sixth codon in the β-globin gene is the sequence CCTGAGG. There are two additional *Mst*II sites on either side of the sixth codon, such that when *Mst*II is used to cut human DNA in this region, two fragments of 1.15 and 0.20 kilobases are obtained (1 kilobase = 1,000 bp). The sickle allele of the β-globin gene causes sickle-cell anemia when it occurs in the homozygous state. In this allele, the sequence around the sixth codon is mutated to CCTGTGG.

1. Identify the point mutation that led to the sickle allele.
2. What fragment(s) would result when *Mst*II is used to cut DNA with the sickle allele?
3. Sketch the patterns of cuts with normal and sickle alleles on a gel.
4. Could an individual have both patterns? Explain.
5. How can this information be used to make a DNA test for these alleles?

negatively charged at neutral pH. Therefore, because opposite charges attract, the DNA fragments move through the gel toward the positive end of the field. Because the spaces between the polymers of the gel are small, small DNA molecules can move through the gel faster than larger ones. Thus DNA fragments of different sizes separate from one another and can be detected with a dye. This gives us three types of information about a DNA sample:

- *The number of fragments*. The number of fragments produced by digestion of a DNA sample with a given restriction enzyme depends on how many times that enzyme's restriction site occurs in the sample. Thus gel electrophoresis can provide some information about the presence of specific DNA sequences in the DNA sample.
- *The sizes of the fragments*. DNA fragments of known size are often placed in one well of the gel to provide a standard for comparison. This tells us how large the DNA fragments in the other wells are. By comparing the fragment sizes obtained with two or more restriction enzymes, the locations of their recognition sequences relative to one another can be worked out (mapped).
- *The relative abundance of a fragment*. In many experiments, the investigator is interested in how much DNA is present. The relative intensity of a band produced by a specific fragment can indicate the amount of that fragment.

After separation on the gel, a slice of gel containing the desired DNA fragment (identified by its size) can be cut out and then be purified by one of a variety of methods. This fragment can

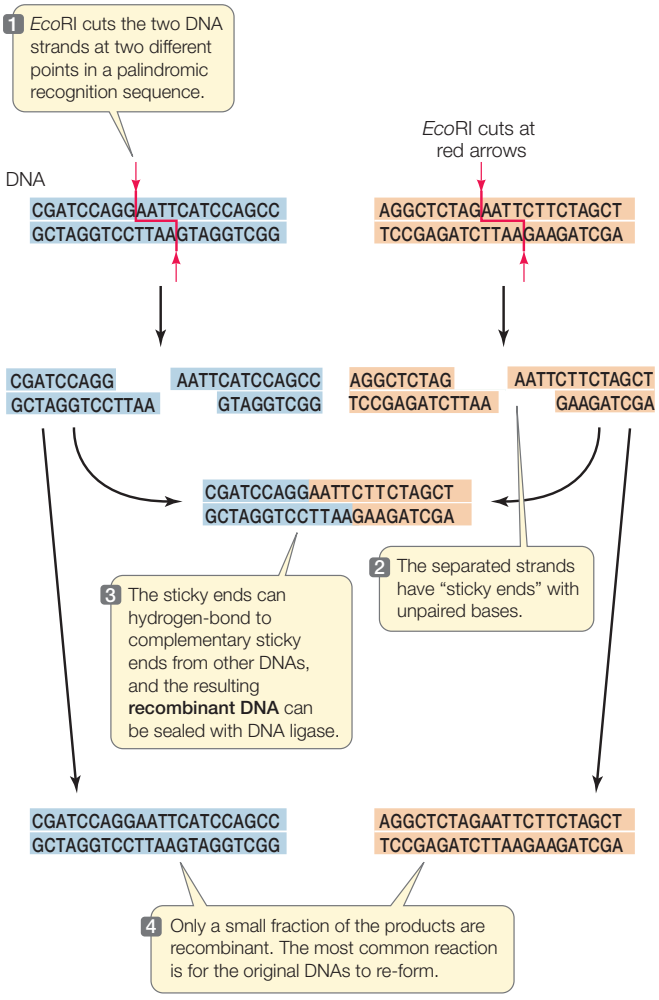


FIGURE 13.3 Cutting and Joining DNA Many restriction enzymes (*Eco*RI is shown here) make staggered cuts in DNA. *Eco*RI can be used to cut two different DNA molecules (blue and orange). The exposed bases can hydrogen-bond with complementary exposed bases on other DNA fragments, forming recombinant DNA. DNA ligase stabilizes the recombinant molecule by forming covalent phosphodiester bonds in the DNA backbone.

then be analyzed to determine its sequence or used to make recombinant DNA.

Go to ANIMATED TUTORIAL 13.1 Separating Fragments of DNA by Gel Electrophoresis
PoL2e.com/at13.1

LINK

Determining and then comparing DNA sequences of the same gene from different species provides information about evolutionary relationships; see **Concepts 16.2 and 16.3**

Recombinant DNA can be made from DNA fragments

Another enzyme that is involved in DNA metabolism in cells is **DNA ligase**, which catalyzes the joining of DNA fragments by making phosphodiester bonds between them. This is the enzyme that joins Okazaki fragments during DNA replication (see Concept 9.2). With restriction enzymes (which *break* bonds) and DNA ligase (which *makes* bonds), scientists can cut DNA into fragments and then join the fragments together in new combinations as recombinant DNA. As shown in **FIGURE 13.3**, two fragments with complementary sticky ends first join by hydrogen bonding, and then the DNA ligase forms a phosphodiester bond in each strand, making a single intact DNA molecule.

In the early 1970s, Stanley Cohen and Herbert Boyer wondered whether recombinant DNA could be a functional carrier of genetic information. They used restriction enzymes to cut sequences from two *E. coli* plasmids (small circular DNAs; see Concept 8.4) containing different antibiotic resistance genes. Then they used DNA ligase to join the fragments together. The resulting plasmid, when inserted into new *E. coli* cells, gave those cells resistance to both antibiotics (**FIGURE 13.4**). A new era of biotechnology was born.

INVESTIGATION

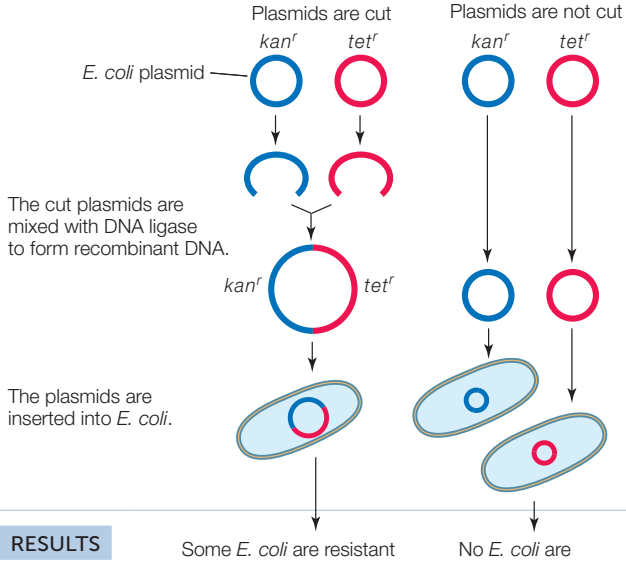
FIGURE 13.4 Recombinant DNA With the discovery of restriction enzymes and DNA ligase, it became possible to combine DNA fragments from different sources in the laboratory. But would such "recombinant DNA" be functional when inserted into a living cell? The results of this experiment completely changed the scope of genetic research, increasing our knowledge of gene structure and function, and ushering in the new field of biotechnology.[a]

HYPOTHESIS

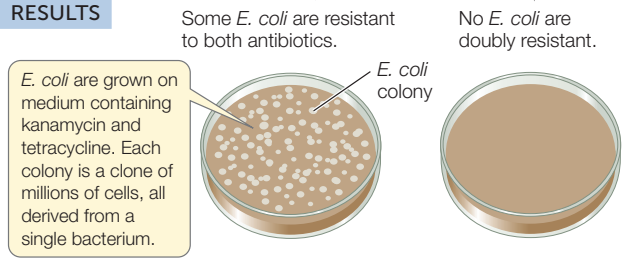
Biologically functional recombinant plasmids can be made in the laboratory.

METHOD

E. coli plasmids carrying a gene for resistance to either the antibiotic kanamycin (*kan*^{r}) or tetracycline (*tet*^{r}) are cut with a restriction enzyme.



RESULTS



CONCLUSION

Two DNA fragments with different genes can be joined to make a recombinant DNA molecule, and the resulting DNA is functional.

ANALYZE THE DATA

Two plasmids were used in this study: pSC101 had a gene for resistance to tetracycline and pSC102 had a gene for resistance to kanamycin. Equal quantities of the plasmids—either intact, cut with *Eco*RI, or cut with *Eco*RI and then sealed with DNA ligase—were mixed and incubated with antibiotic-sensitive *E. coli*. The *E. coli* were then grown on various combinations of the antibiotics. Here are the results:

	Number of resistant colonies		
DNA treatment	Tetracycline only	Kanamycin only	Both antibiotics
None	200,000	100,000	200
*Eco*RI cut	10,000	1,100	70
*Eco*RI, then ligase	12,000	1,300	570

A. Did treatment with *Eco*RI affect the transformation efficiency? Explain.

B. Did treatment with DNA ligase affect the transformation efficiency of each cut plasmid? Which quantitative data support your answer?

C. How did doubly antibiotic-resistant bacteria arise in the "none" treatment? (Hint: see Concept 9.3.)

D. Did the *Eco*RI followed by ligase treatment increase the appearance of doubly antibiotic-resistant bacteria? What data support your answer?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]S. N. Cohen et al. 1973. *Proceedings of the National Academy of Sciences USA* 70: 3240–3244.

CHECKpoint CONCEPT 13.1

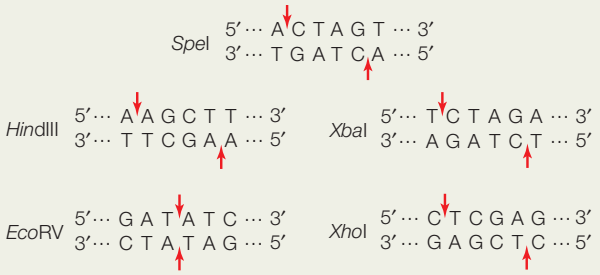
- ✓ Using diagrams of the chemical structure of DNA (see Concepts 3.1 and 9.2), compare the actions of a restriction enzyme and DNA ligase.
- ✓ How is recombinant DNA technology different from genetic recombination that occurs in meiosis?
- ✓ A DNA molecule of 12,000 bp (12 kilobase, or kb) is cut by restriction enzymes and analyzed by gel electrophoresis as shown in the table. The 2-kb band from the double digest (Enzymes A + B) is twice as intense as the 2-kb bands from the single digests.

CONDITION	SIZES OF FRAGMENTS (kb)
Enzyme A	2, 10
Enzyme B	2, 10
Enzymes A + B	2, 8

  a. Sketch the results of the three cuts on an electrophoresis gel.

  b. Indicate on a linear map where each enzyme cuts the DNA.

- ✓ DNA from two different sources is cut with two different restriction enzymes. Fragments from the two sources are then joined together. DNA from source A is cut with *Spe*I. Which of the four restriction enzymes shown below *Spe*I could be used to cut DNA from source B, such that it could join to a fragment from source A?



With the tools described in this concept—restriction enzymes, gel electrophoresis, and DNA ligase—scientists can cut and rejoin different DNA molecules from any and all sources, including artificially synthesized DNA sequences. In the next concept we will examine some of the ways these recombinant DNA molecules are used.

CONCEPT 13.2 DNA Can Genetically Transform Cells and Organisms

One goal of recombinant DNA technology is to **clone**—produce many identical copies of—a particular DNA sequence. We have seen the term "clone" used in the context of whole cells or organisms that are genetically identical to one another (see Concept 7.1). A gene can be cloned by inserting it into a bacterial cell such as *E. coli*. The bacterium is allowed to reproduce and multiply into millions of identical cells, all carrying copies of the gene. Cloning might be done for sequence analysis, to produce a protein product in quantity, or as a step toward creating an organism with a new phenotype.

The process of inserting recombinant DNA into host cells is called **transformation**, or **transfection** if the host cells are derived from an animal. A host cell or organism that contains recombinant DNA is described as **transgenic**. Later in this chapter we will encounter many examples of transgenic cells and organisms, including yeast, mice, rice plants, and even cattle.

Various methods are used to create transgenic cells. Generally, only a few of the cells that are exposed to the recombinant DNA actually become transformed with it. In order to grow only the transgenic cells, **selectable marker** genes, such as genes that confer resistance to antibiotics, are often included as part of the recombinant DNA molecule. Antibiotic resistance genes were the markers used in Cohen and Boyer's experiment (see Figure 13.4).

Genes can be inserted into prokaryotic or eukaryotic cells

In theory, any cell or organism can act as a host for the introduction of recombinant DNA. Most research has been done using model organisms:

- *Bacteria* are easily grown and manipulated in the laboratory. Much of their molecular biology is known, especially for well-studied bacteria such as *E. coli*. Furthermore, bacteria contain plasmids, which are easily manipulated to carry recombinant DNA into the cell. Because the processes of transcription and translation proceed differently in prokaryotes than they do in eukaryotes, however, bacteria might not be suitable as hosts to express eukaryotic genes.
- *Yeasts* such as *Saccharomyces* are commonly used as eukaryotic hosts for recombinant DNA studies. The advantages of using yeasts include rapid cell division (a life cycle completed in 2 hours), ease of growth in the laboratory, and a relatively small genome size. In addition, yeasts have most of the characteristics of other eukaryotes, except for those characteristics involved in multicellularity.
- *Plant cells* are good hosts, because even fully differentiated plant cells can be treated with hormones that make them dedifferentiate into unspecialized stem cells (see Concept 13.3). The unspecialized cells can be transformed with recombinant DNA and then studied in culture, or grown into new plants. There are also methods for making whole transgenic plants without going through the cell culture step. These methods result in plants that carry the recombinant DNA in all their cells, including the germline cells.
- *Cultured animal cells* can be used to study the expression of human or animal genes, for example for medical purposes. Whole transgenic animals can also be created.

A variety of methods are used to insert recombinant DNA into host cells

Methods for inserting DNA into host cells vary. The cells may be chemically treated to make their outer membranes more

permeable, and then mixed with the DNA so it can diffuse into the cells. Another approach is called electroporation: a short electric shock is used to create temporary pores in the membranes through which the DNA can enter. Viruses can be altered so that they carry or insert recombinant DNA into cells. A common method for transforming plants involves a specific bacterium that inserts DNA into cells of some plants. Transgenic animals can be produced by injecting recombinant DNA into the nuclei of fertilized eggs. There are even "gene guns," which "shoot" the host cells with tiny particles carrying the DNA.

The challenge of inserting new DNA into a cell lies not just in getting it into the host cell, but in getting it to replicate as the host cell divides. DNA polymerase does not bind to just any sequence. If the new DNA is to be replicated, it must become part of a segment of DNA that contains an origin of replication (see Concept 9.2). Such a DNA molecule is called a replicon, or replication unit.

There are two general ways in which the newly introduced DNA can become part of a replicon within the host cell:

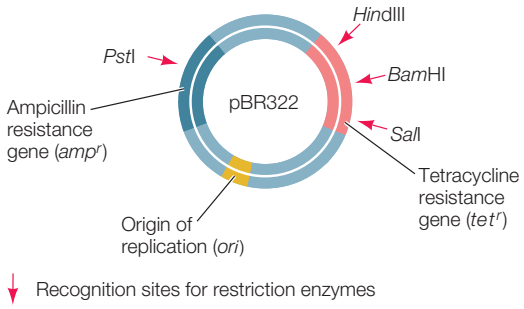
- It may be inserted into a host chromosome. Although the site of insertion is usually random, this is nevertheless a common method of integrating new genes into host cells.
- It can enter the host cell as part of a carrier DNA sequence, called a **vector**, and can either integrate into the host chromosome or have its own origin of DNA replication.

Several types of vectors are used to get DNA into cells.

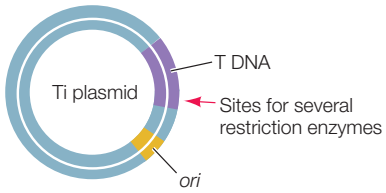
PLASMIDS AS VECTORS As we described in Concept 8.4, plasmids are small, circular DNA molecules that replicate autonomously in many prokaryotic cells. A number of characteristics make plasmids useful as transformation vectors:

- Plasmids are relatively small (an *E. coli* plasmid usually has 2,000–6,000 bp) and are therefore easy to manipulate in the laboratory.
- A typical plasmid has one or more restriction enzyme recognition sequences that each occur only once in the plasmid sequence. These sites make it easy to insert additional DNA into the plasmid before it is used to transform host cells.
- Many plasmids contain genes that confer resistance to antibiotics and thus can serve as selectable markers.
- Plasmids have a bacterial origin of replication (*ori*) and can replicate independently of the host chromosome. It is not uncommon for a bacterial cell to contain hundreds of copies of a recombinant plasmid. For this reason, the power of bacterial transformation to amplify a gene is extraordinary. A 1-liter culture of bacteria harboring the human β-globin gene in a typical plasmid can have as many copies of that gene as there are cells in a typical adult human (10^{14}).

The plasmids used as vectors in the laboratory have been extensively altered to include convenient features: multiple cloning sites with 20 or more unique restriction enzyme sites for cloning purposes; origins of replication for a variety of host cells; and various kinds of reporter genes (see p. 260) and selectable marker genes. An example is pBR322, a plasmid used to transform *E. coli*:



PLASMID VECTORS FOR PLANTS An important vector for carrying new DNA into many types of plants is a plasmid found in the bacterium *Agrobacterium tumefaciens*. This bacterium lives in the soil, infects plants, and causes a disease called crown gall, which is characterized by the presence of growths (or tumors) on the plant. *A. tumefaciens* contains a plasmid called Ti (for *t*umor-*i*nducing). When the bacterium infects a plant cell, a region of the Ti plasmid called the T DNA is inserted into the cell, where it becomes incorporated into one of the plant's chromosomes. The Ti plasmid carries the genes needed for this transfer and incorporation of the T DNA:



The T DNA carries genes that are expressed by the host cell, causing the growth of tumors and the production of specific sugars that the bacterium uses as sources of energy. Scientists have exploited this remarkable natural "genetic engineer" to insert foreign DNA into the genomes of plants. When used as a vector for plant transformation, the tumor-inducing and sugar-producing genes on the T DNA are removed and replaced with foreign DNA. The altered Ti plasmids are first used to transform *Agrobacterium* cells from which the original Ti plasmids have been removed. Then the *Agrobacterium* cells are used to infect plant cells.

VIRUSES AS VECTORS Constraints on plasmid replication limit the size of the new DNA that can be inserted into a plasmid to about 10,000 bp. Although many prokaryotic genes may be smaller than this, most eukaryotic genes—with their introns and extensive flanking sequences—are bigger. A vector that accommodates larger DNA inserts is needed for these genes.

Both prokaryotic and eukaryotic viruses are often used as vectors for eukaryotic DNA. Bacteriophage λ, which infects *E. coli*, has a DNA genome of about 45,000 bp; this is all that fits into the phage head. If the phage genes that cause the host cell to die and lyse—about 20,000 bp—are eliminated, the virus can still attach to a host cell and inject its DNA, but the host cell

APPLY THE CONCEPT

DNA can genetically transform cells and organisms

As shown in Figure 13.5, the β-galactosidase (*lacZ*) gene encodes an enzyme that can convert the colorless substrate X-gal into a bright blue product. A plasmid vector contains a modified version of the *lacZ* gene with multiple restriction sites within its coding sequence, and a gene for resistance (R) to the antibiotic ampicillin. This vector is used with an *E. coli* strain that carries no other functional *lacZ* gene and that is sensitive (S) to ampicillin. A biologist clones a wheat gene by inserting it into the multiple cloning site of the *lacZ* gene in this vector. As a control to ensure that the bacterial cells are viable, she also grows some cells that were not transformed with the ligation products. Remember that after a ligation reaction, only a few of the plasmids are recombinant. Fill in the table with the results of these transformations.

DNA TAKEN UP BY *E. COLI*	PHENOTYPE FOR AMPICILLIN (R OR S)	PHENOTYPE FOR X-GAL (BLUE OR WHITE)
None		
Plasmid only		
Recombinant plasmid		

will not die. The deleted 20,000 bp can be replaced with DNA from another organism. Because viruses infect cells naturally, they offer a great advantage over plasmids, which often require artificial means to coax them to enter host cells.

Reporter genes help select or identify host cells containing recombinant DNA

Even when a population of host cells interacts with an appropriate vector, only a small proportion of the cells actually take up the vector. Furthermore, the process of making recombinant DNA is far from perfect. After a ligation reaction, not all the vector copies contain the foreign DNA. How can we identify or select the host cells that contain the vector with foreign DNA?

As we described above, selectable markers such as antibiotic resistance genes can be used to select cells containing those genes. These cells can be selected because only cells carrying the antibiotic resistance gene can grow in the presence of that antibiotic. Selectable markers are one type of **reporter gene**, which is any gene whose expression is easily assayed. Other reporter genes code for proteins that can be detected visually. Two commonly used reporter genes are *lacZ* and the gene for green fluorescent protein:

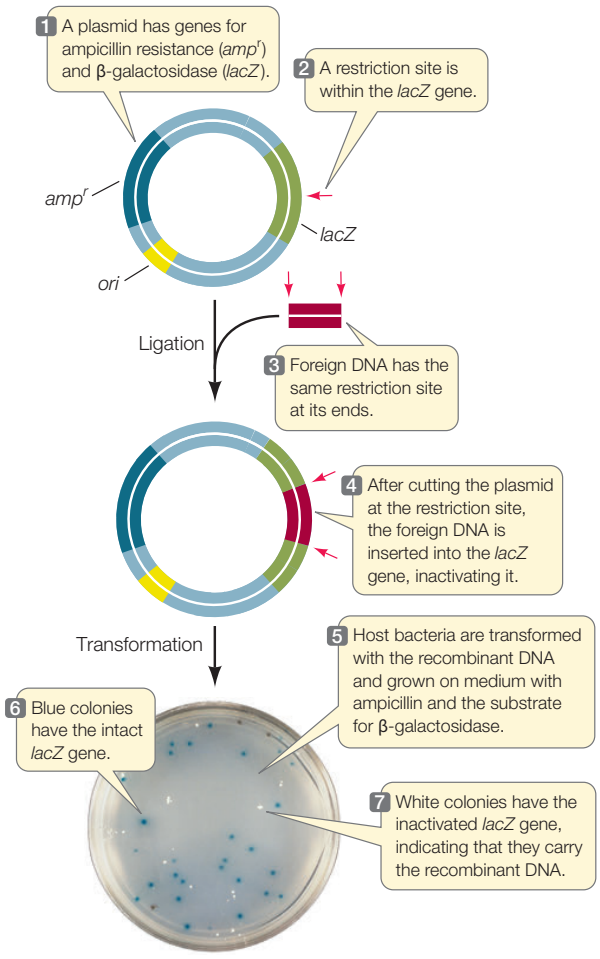
- The β-galactosidase (*lacZ*) gene (see Figure 11.3) codes for an enzyme that can convert the white substrate X-gal into a bright blue product. Many cloning plasmids contain the *lacZ* gene, along with genes for antibiotic resistance. As shown in **FIGURE 13.5**, foreign DNA can be inserted into the *lacZ* gene, inactivating it. Bacteria transformed with the plasmid are selected on a solid medium containing X-gal and the appropriate antibiotic. Clones containing the recombinant plasmid cannot make β-galactosidase, and produce white colonies. Clones that contain the original plasmid with no insert express the *lacZ* gene and make blue colonies.
- Green fluorescent protein (GFP), which normally occurs in the jellyfish *Aequorea victoria*, emits green light when exposed to ultraviolet light. The gene for this protein is now widely used as a reporter gene (**FIGURE 13.6**).

LINK

The *lacZ* gene is part of the inducible *lac* operon in *E. coli*, which encodes proteins required for β-galactoside metabolism; see **Concept 11.1**

RESEARCH TOOLS

FIGURE 13.5 Selection for Recombinant DNA Selectable marker (reporter) genes are used by scientists to select for bacteria that have taken up a plasmid. In a typical experiment, most of the bacteria will not take up any DNA. Of those that do, only a small fraction will take up recombinant DNA.



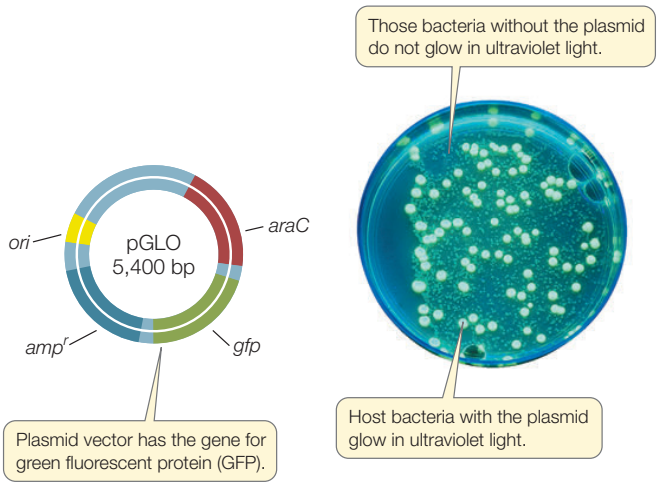


FIGURE 13.6 Green Fluorescent Protein as a Reporter The presence of a plasmid with the gene for green fluorescent protein is readily apparent in transgenic cells because they glow under ultraviolet light.

Such reporters are not just used to select and identify cells carrying recombinant DNA. They can be attached to promoters in order to study how the promoters function under different conditions or in different tissues of a transgenic multicellular organism. Reporters can also be attached to other proteins, to study how and where those proteins become localized within eukaryotic cells.

CHECKpoint CONCEPT 13.2

- ✓ Outline the steps used to create recombinant DNA, transform host cells, and detect cells carrying the recombinant DNA.
- ✓ "Shuttle vectors" have the ability to transform both prokaryotic and eukaryotic cells. What sequences would you expect these vectors to have?
- ✓ What are the advantages of green fluorescent protein over antibiotic resistance as a marker on a plasmid for genetic transformation?
- ✓ Plasmid X has ampicillin and tetracycline resistance genes. The restriction enzyme *Eco*RI cleaves the plasmid once, within the tetracycline resistance gene. Plasmid B has a streptomycin resistance gene and one site for *Eco*RI cleavage that is not within the resistance gene. The two plasmids are cut with *Eco*RI and treated with DNA ligase. The mixture is used to transform an *E. coli* strain that is sensitive to the three antibiotics. Which antibiotic(s) would you add to the bacterial growth medium to select those bacteria carrying a recombinant plasmid?

We have described how DNA can be cut, inserted into a vector, and introduced into host cells. We have also seen how host cells carrying recombinant DNA can be identified. Now let's consider the sources of DNA used for cloning, as well as some molecular methods for manipulating gene expression.

CONCEPT 13.3 Genes Come from Various Sources and Can Be Manipulated

A major goal of molecular cloning experiments is to elucidate the functions of the DNA sequences and the proteins they encode. In this concept we look at ways that specific sequences can be identified and amplified. Millions of copies of a sequence are needed in order to study and manipulate the sequence in the laboratory. Cloned or amplified DNA can be used for various purposes, including the detection of expressed genes in specific cells and the artificial regulation of gene expression.

DNA fragments for cloning can come from several sources

The DNA fragments used in cloning procedures are obtained from a number of sources. In many cases the first step is to create a "library" of DNA fragments: a collection of clones that can be searched for the gene or genes of interest, or analyzed in other ways to learn more about the original source of the DNA fragments.

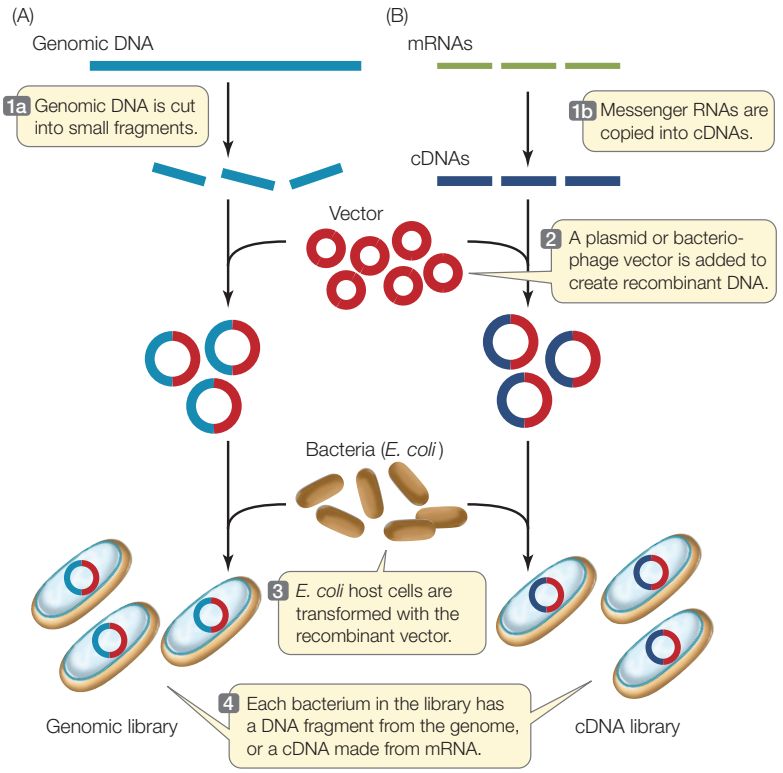
GENOMIC LIBRARIES A **genomic library** is a collection of DNA fragments that together comprise the genome of an organism. This is the starting point for some methods of genome sequencing. Restriction enzymes or other means, such as mechanical shearing, can be used to break chromosomes into smaller pieces (FIGURE 13.7A). Each fragment is inserted into a vector, which is then taken up by a host cell. Proliferation of a single transformed cell on a selective medium (such as for antibiotic resistance) produces a colony of recombinant cells, each of which harbors many copies of the same fragment of DNA. The colonies are grown by spreading the transformed cells over a solid culture medium in petri dishes (small circular plates), which are incubated at a suitable temperature for the host cells to grow.

A single petri dish can hold thousands of bacterial colonies and is easily screened for the presence of a particular DNA sequence. Colonies containing that sequence are identified by DNA hybridization using a probe labeled with complementary fluorescent or radioactive nucleotides. To do this, the petri dish with its bacterial colonies is duplicated, and then the bacteria on one of the plates are treated to expose the DNA for hybridization (see Figure 10.7).

cDNA A much smaller DNA library—one that includes only the genes transcribed in a particular tissue—can be made from **complementary DNA**, or **cDNA** (FIGURE 13.7B). This involves isolating mRNA from cells and making cDNA copies of that mRNA by complementary base pairing. The enzyme reverse transcriptase catalyzes this reaction. This collection of cDNAs from a particular tissue at a particular time is called a **cDNA library**, which is a "snapshot" of the transcription pattern of the cells in the sample. cDNA libraries have been invaluable for comparing gene expression in different tissues at different stages of development. For example, if cDNAs derived from developing red blood cells are examined, the

RESEARCH TOOLS

FIGURE 13.7 Constructing Libraries Intact genomic DNA is too large to be introduced into host cells. (A) A genomic library can be made by breaking the DNA into small fragments, incorporating the fragments into a vector, and then transforming host cells with the recombinant vectors. Each colony of cells contains many copies of a small part of the genome. (B) Similarly, there are many mRNAs in a cell. These can be copied into cDNAs and a library made from them. The DNA in these colonies can then be isolated for analysis.



globin sequences (encoding the subunits of hemoglobin) are prominent. But a cDNA library derived from hair follicles does not contain those sequences.

Reverse transcriptase along with PCR (see below) can be used to create and amplify a specific cDNA sequence without the need to make a library. In this case, RNA is isolated from cells and then reverse transcriptase is used to make cDNA from the RNA. Then PCR is used to amplify a specific sequence directly from the cDNA. This method, called **RT-PCR**, is an invaluable tool for studies of the expression of particular genes in cells and organisms.

Synthetic DNA can be made in the laboratory

In Concept 9.2 (see Figure 9.15) we described the polymerase chain reaction (PCR), a method of amplifying DNA in a test tube. PCR can begin with just a single molecule of DNA, although larger quantities [in the picogram (10^{-12}) to microgram (10^{-6}) range] are more often used. Any fragment of DNA can be amplified as long as appropriate primers are available. This amplified DNA can then be inserted into a plasmid to create recombinant DNA, and cloned in host cells.

The artificial synthesis of DNA by organic chemistry methods is now fully automated. Synthetic oligonucleotides (single-stranded DNA fragments of 20–40 bp) are used as primers in PCR reactions. These primers can be designed to create short new sequences at the ends of the PCR products. This might be done to create a mutation in a recombinant gene, or to add restriction enzyme sites at the ends of the PCR product to aid in ligation reactions. Longer synthetic sequences can be pieced together to construct completely artificial genes that have been designed for specific purposes. For example, a gene might be designed to be highly expressed in a particular cell type, or to encode a highly active enzyme.

LINK

Synthetic DNA was used to create a novel bacterial genome to replace the genome in a host cell, resulting in a new bacterial species; see the opening story of **Chapter 4**

DNA sequences can be manipulated to study cause-and-effect relationships

Mutations that occur in nature have been important in demonstrating cause-and-effect relationships in biology. However, mutations in nature are rare events. Recombinant DNA technology allows us to ask "what if" questions by creating artificial gene constructs. Because synthetic DNA can be made with any desired sequence, it can be manipulated to create specific constructs or mutations, and the resulting phenotypes can be observed when the recombinant DNA is expressed in host cells. Such techniques have revealed thousands of cause-and-effect relationships.

One example involves the auxin response element, a short sequence of DNA that binds a specific transcription factor. This element is found in the promoters of plant genes that are switched on in the presence of the plant hormone auxin (see Concept 26.2). To study the role of the auxin response element in plants, scientists made an artificial promoter containing many copies of the element, and ligated the promoter to a reporter gene. The recombinant DNA was used to transform *Arabidopsis* plants. When the plants were treated with auxin, the reporter gene was switched on at very high levels (higher than those produced by a wild-type auxin-responsive promoter). This experiment helped show that the presence of the auxin response element (the "cause") results in gene expression in response to auxin (the "effect").

Genes can be inactivated by homologous recombination

Another way to understand a gene's function is to inactivate it so it is not transcribed and translated into a protein.

An example of this approach is the use of transposon mutagenesis in experiments designed to describe the minimal genome (see Figure 12.8). In animals, these "knockout" experiments often involve **homologous recombination** rather than transposon mutagenesis. As we saw in Chapter 8, recombination occurs when a pair of homologous chromosomes line up during meiosis. The chromosomes sometimes break and then rejoin in such a way that segments of the two chromosomes are exchanged. A key feature of homologous recombination is that it involves an exchange of DNA between molecules with identical, or nearly identical, sequences.

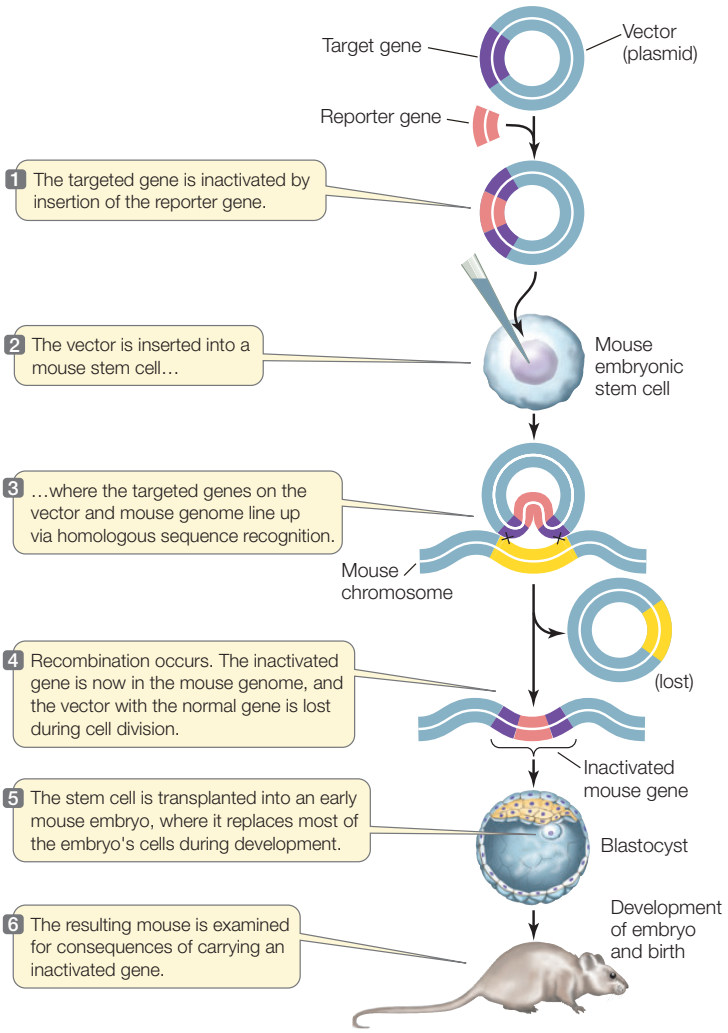
We will focus here on the technique used for mice (FIGURE 13.8). In order to knock out (inactivate) a target gene, the normal allele of the gene is inserted into a plasmid. Restriction enzymes are then used to insert a fragment containing a reporter gene or selectable marker into the middle of the normal gene. This addition of extra DNA disrupts the gene's coding region so that it no longer encodes a functional protein product.

Once the recombinant plasmid has been made, it is used to transfect mouse embryonic stem cells. A **stem cell** is an unspecialized cell that divides and differentiates into specialized cells. The gene sequences in the plasmid tend to line up with their homologous sequences in the mouse chromosome. If recombination occurs, the disrupted, inactive allele is "swapped" with the functional allele in the cell.

The knockout technique has been important in assessing the roles of many genes, and is especially valuable in studying human genetic diseases. Many such diseases (including phenylketonuria; see Concept 10.1) have knockout mouse models: mouse strains with similar diseases that were produced by homologous recombination. These models can be used to study the diseases and to test potential treatments.

RESEARCH TOOLS

FIGURE 13.8 Making a Knockout Mouse Animals carrying mutations are rare. Homologous recombination is used to replace a normal mouse gene with an inactivated copy of that gene, thus "knocking out" the gene. Discovering what happens to a mouse with an inactive gene tells us much about the normal role of that gene.



Complementary RNA can prevent the expression of specific genes

Another way to study the expression of a specific gene is to block the translation of its mRNA. This is yet another example of scientists imitating nature. As described in Concept 11.4, gene expression can be controlled in nature by the production of short, single-stranded RNA molecules (microRNAs or miRNAs) that inhibit the translation of target mRNA sequences. Many complex eukaryotes also produce small interfering RNAs (siRNAs), which are short (20–25 bp) double-stranded RNAs derived from much longer double-stranded RNA molecules. As in the production of miRNAs, these double-stranded siRNA molecules are processed into single-stranded molecules, and then each one is guided by a protein complex to a complementary region on an mRNA. The protein complex then catalyzes the breakdown of the targeted mRNA (FIGURE 13.9). These mechanisms for preventing mRNA translation are called **RNA interference** (**RNAi**).

MicroRNAs and siRNAs are examples of **antisense RNA** because they bind by base pairing to the "sense" bases on the target mRNAs. siRNAs target specific mRNA molecules (from specific genes) because their sequences exactly match the target sequences in the mRNAs. By contrast, miRNAs do not match their targets perfectly, and therefore each one can reduce the expression of multiple, partially matching genes.

RNAi was discovered in the late 1990s, and since then scientists have used synthetic, single-stranded antisense RNAs and double-stranded siRNAs to inhibit the expression of known genes. This technique has been used extensively to block expression of specific genes in the laboratory, as well as in applied situations. For example, macular degeneration is an eye disease that results in near blindness when blood vessels proliferate in the eye. The signaling molecule that stimulates vessel

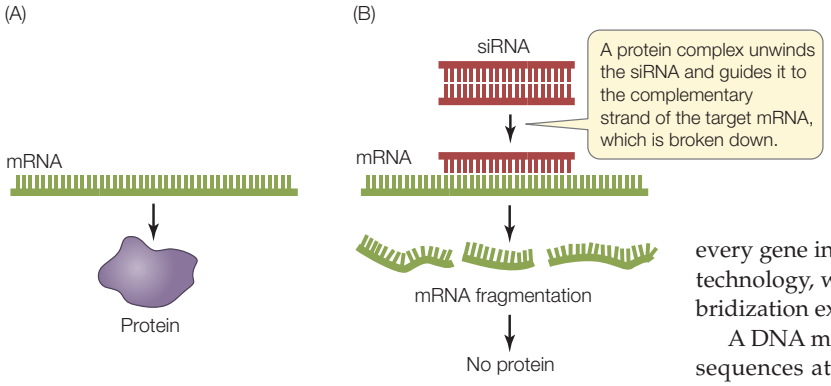


FIGURE 13.9 Using siRNA to Block the Translation of mRNA (A) Normally an mRNA is translated to produce a protein. (B) Translation of a target mRNA can be prevented with a small interfering RNA (siRNA) that is complementary to part of the target mRNA.

proliferation is a growth factor. An RNAi-based therapy is being developed to target this growth factor's mRNA, and the therapy shows promise for stopping and even reversing the progress of the disease.

DNA microarrays reveal RNA expression patterns

The science of genomics faces two major quantitative realities. First, there are very large numbers of genes in eukaryotic genomes. Second, the pattern of gene expression in different tissues at different times is quite distinctive. For example, the cells of a skin cancer at its early stage may have a different set of mRNAs from those of normal skin cells and cells from a more advanced skin cancer.

To find such patterns, scientists could isolate mRNA from a cell and test for the presence of transcripts from each gene by hybridization or RT-PCR. But that would involve many steps and take a long time. It is far simpler to measure expression of every gene in one step. This is possible with DNA microarray technology, which provides large arrays of sequences for hybridization experiments.

A DNA microarray ("gene chip") contains a series of DNA sequences attached to a solid surface. The array is divided into a grid of microscopic spots, each containing thousands of copies of a particular oligonucleotide. A computer controls the addition of these oligonucleotide sequences in a predetermined pattern. Each oligonucleotide can hybridize with only one DNA or RNA sequence, and thus is a unique identifier of a gene. Many thousands of different oligonucleotides can be placed in a single microarray.

Microarrays can be used to examine patterns of gene expression in different tissues and under different conditions, and they can be used to identify individual organisms with particular mutations. You can visualize the concept of microarray analysis by following the example illustrated in **FIGURE 13.10**. Most women with breast cancer are treated with surgery to remove the tumor, and then treated with radiation soon afterward to kill cancer cells that the surgery may have missed. But a few cancer cells may survive in some patients, and these cells eventually form new tumors in the breast or elsewhere in the body. The challenge for physicians is to identify patients with surviving cancer cells so they can be treated aggressively with tumor-killing chemotherapy.

Scientists at the Netherlands Cancer Institute used medical records to identify patients whose cancer recurred or did not recur. They extracted mRNA from the patients' tumors and made cDNA from the samples. The cDNAs were hybridized to microarrays containing sequences derived from 1,000 human genes. The scientists found 70 genes whose expression differed dramatically between tumors from patients whose cancers recurred and tumors from patients whose cancers did not recur. From this information the Dutch group identified "gene expression signatures" that are useful in clinical decision-making: patients with a good prognosis can avoid unnecessary chemotherapy,

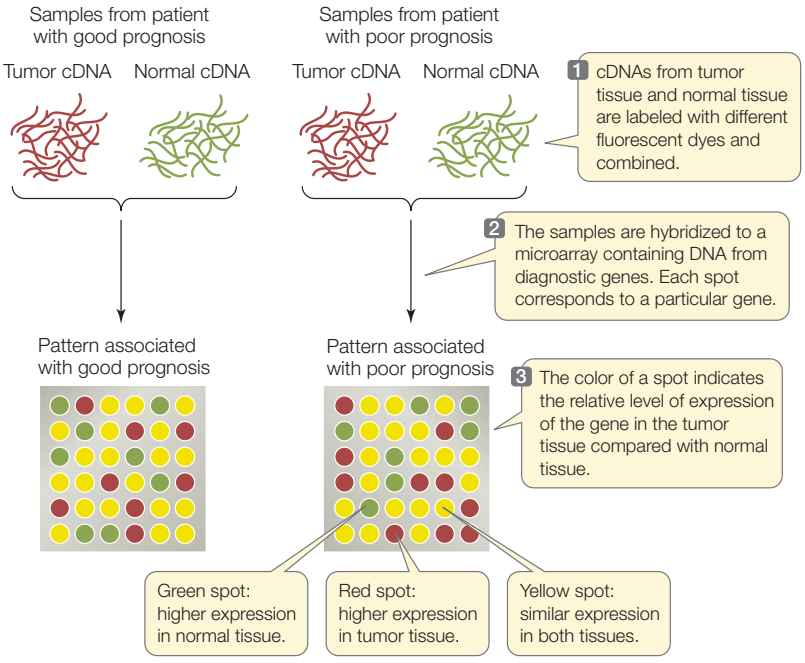


FIGURE 13.10 Using DNA Microarrays for Clinical Decision-Making The pattern of expression of 70 genes in tumor tissues (the pattern of colored spots) indicates whether breast cancer is likely to recur. Actual arrays have more dots than shown here.

Go to ANIMATED TUTORIAL 13.2
DNA Chip Technology
PoL2e.com/at13.2

whereas those with a poor prognosis can receive more aggressive treatment.

CHECKpoint CONCEPT 13.3

- ✓ Outline the steps involved in knocking out a gene in a bacterium and in an animal. What are the uses for these methods?
- ✓ What are the differences between a genomic library and a cDNA library?
- ✓ You hypothesize that when a corn plant is infected with a fungus, a set of genes is turned on that fights the infection. How would you investigate this hypothesis using microarray technology?

We have now seen how recombinant DNA is made, how cells and organisms are transformed, and how gene expression can be manipulated. In the final concept we will look at some of the many applications of biotechnology.

CONCEPT 13.4 Biotechnology Has Wide Applications

In the opening story of this chapter we defined biotechnology as the use of cells or whole living organisms to make or modify materials or processes that are useful to people, such as foods, medicines, and chemicals. Bacteria and yeast cells can be transformed with almost any gene, and they can be induced to express that gene at high levels and to export the protein product out of their cells. This technology has turned these microbes into versatile factories for many important products. Today there is interest in producing nutritional supplements and pharmaceuticals in whole transgenic animals and plants and harvesting the products in large quantities—for example, from cow's milk or rice grains. Another goal is to produce animals and plants with improved characteristics, such as increased nutritional value or better tolerance of harsh environments. Key to this boom in biotechnology has been the development of specialized vectors that not only carry genes into animal and plant cells, but also make those cells express the genes at high levels.

Expression vectors can turn cells into protein factories

Many proteins that are potentially useful to humans come from eukaryotes. But if a eukaryotic gene is inserted into a typical plasmid and used to transform *E. coli*, none of the gene product will be made. Other key prokaryotic DNA sequences must be included with the gene. A bacterial promoter, a signal for transcription termination, and a special sequence that is necessary for ribosome binding on the mRNA must all be included in the transformation vector if the gene is to be expressed in the bacterial cell. In addition, the eukaryotic coding region must be made using cDNA so that it has no introns.

To solve this kind of problem, scientists make **expression vectors** that have all the characteristics of typical vectors, as well as the extra sequences needed for the foreign gene (also called a transgene) to be expressed in the host cell. For bacterial hosts, these additional sequences include the elements named above (FIGURE 13.11). For eukaryotes, they include the poly A addition sequence and a promoter that contains all the elements needed for expression in a eukaryotic cell. An expression vector can include various types of promoters and other features:

- An *inducible promoter*, which responds to a specific signal, can be included. For example, a promoter that responds to hormonal stimulation (e.g., a promoter containing the auxin response element described earlier) can be used so that the transgene will be expressed at high levels only when the hormone is added.
- A *tissue-specific promoter*, which is expressed only in a certain tissue at a certain time, can be used if localized expression is desired. For example, many seed proteins are expressed only in the plant embryo. Coupling a transgene to a seed-specific promoter will allow the gene to be expressed only in seeds and not, for example, in leaves.

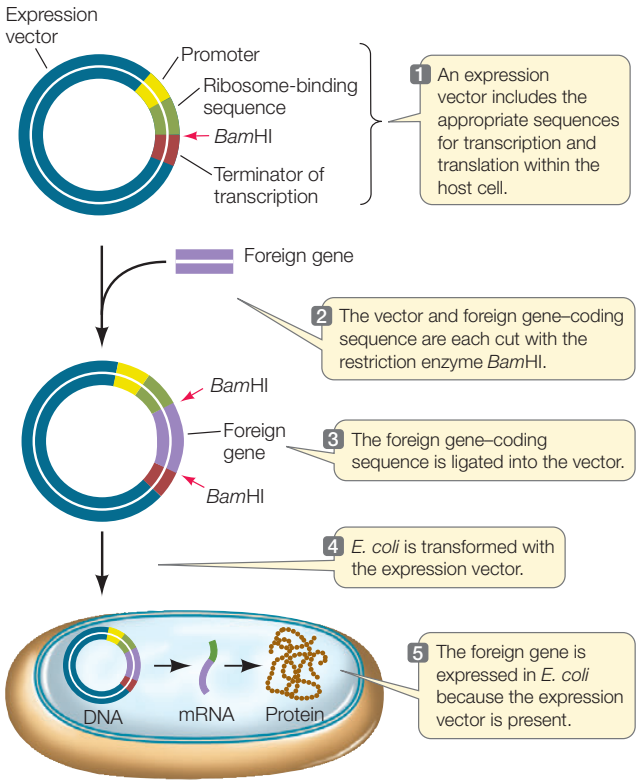


FIGURE 13.11 A Transgenic Cell Can Produce Large Amounts of the Transgene's Protein Product To be expressed in *E. coli*, a gene derived from a eukaryote requires bacterial sequences for transcription initiation (promoter), transcription termination, and ribosome binding. Expression vectors contain these additional sequences, enabling the eukaryotic protein to be synthesized in the prokaryotic cell.

Go to ACTIVITY 13.1 Expression Vectors
PoL2e.com/ac13.1

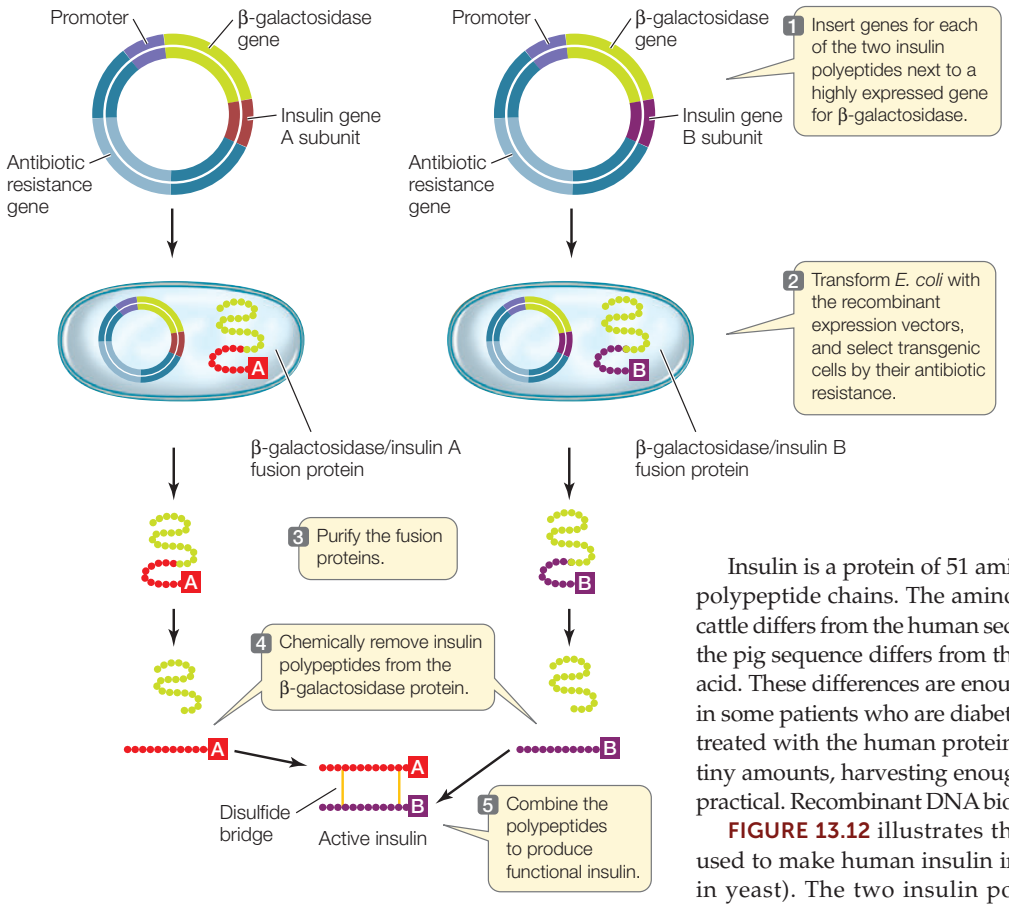


FIGURE 13.12 Human Insulin: From Gene to Drug Human insulin chains are made by recombinant DNA technology and then combined to produce the widely used drug.

- *Signal sequences* can be added so that the gene product is directed to an appropriate destination. For example, when a protein is made by yeast or bacterial cells in a liquid medium, it is economical to include a signal directing the protein to be secreted into the extracellular medium for easier recovery.

LINK

You can review the mechanisms of transcription and mRNA processing in **Concept 10.2**, transcriptional regulation in **Concepts 11.2 and 11.3**, and signal sequences in **Concept 10.5**

Medically useful proteins can be made by biotechnology

Some medically useful products are being made by biotechnology (**TABLE 13.1**), and more are in various stages of development. Human insulin was the first medicine to be made using recombinant DNA, and provides a good illustration of a medical application of biotechnology. Insulin is essential for glucose uptake into cells. People with type I diabetes mellitus cannot make this hormone and must receive insulin injections. Before the advent of biotechnology, the insulin used for this purpose came from cattle and pigs.

Insulin is a protein of 51 amino acids and is made up of two polypeptide chains. The amino acid sequence of insulin from cattle differs from the human sequence by three amino acids, and the pig sequence differs from that of humans by just one amino acid. These differences are enough to cause an immune reaction in some patients who are diabetic, and these patients need to be treated with the human protein. Since the hormone is made in tiny amounts, harvesting enough from deceased persons is not practical. Recombinant DNA biotechnology solved this problem.

FIGURE 13.12 illustrates the strategy that was originally used to make human insulin in *E. coli* (today it is often made in yeast). The two insulin polypeptides were synthesized separately using an expression vector containing the gene for β-galactosidase (*lacZ*). Each insulin gene was inserted into the vector in such a way that it was induced, transcribed, and translated along with the β-galactosidase gene. After extraction and purification of the β-galactosidase–insulin fusion proteins, the insulin polypeptides were cleaved off by chemical treatment. The two insulin peptides were then combined to make a complete, functional human insulin molecule.

Before giving it to human patients, scientists had to be confident that the product made by biotechnology was functional human insulin. Several lines of evidence supported such confidence:

- The synthetic protein is the same size as human insulin.
- It has the same amino acid sequence.
- It has the same shape, as measured by physical techniques.
- It binds to the insulin receptor on cells and stimulates glucose uptake.

LINK

The crucial role of insulin in regulating glucose metabolism is detailed in **Concept 30.5**

Another way of making medically useful products in large amounts is **pharming**: the production of pharmaceuticals in farm animals or plants. For example, a gene encoding a useful

TABLE 13.1 Some Medically Useful Products of Biotechnology

Product	Use
Erythropoietin	Prevents anemia in patients undergoing kidney dialysis and cancer therapy
Colony-stimulating factor	Stimulates production of white blood cells in patients with cancer and AIDS
Bovine/porcine somatotropin	Stimulates growth and milk production in animals
Tissue plasminogen activator	Dissolves blood clots after heart attacks and strokes
Human growth hormone	Replaces missing hormone in people of short stature
Human insulin	Stimulates glucose uptake from blood in patients with type I diabetes mellitus
Factor VIII	Replaces clotting factor missing in patients with hemophilia A
Platelet-derived growth factor	Stimulates wound healing

protein might be placed next to the promoter of the gene that encodes lactoglobulin, an abundant milk protein. Transgenic animals carrying this recombinant DNA will secrete large amounts of the foreign protein into their milk. These natural "bioreactors" can produce abundant supplies of the protein, which can be separated easily from the other components of the milk (**FIGURE 13.13**).

Human growth hormone (hGH) is a protein made in the pituitary gland and has many effects, especially in growing children. People with hGH deficiencies have short stature as well as other abnormalities, a condition known as pituitary dwarfism. In the past they were treated with hGH isolated from the pituitary glands of dead people, but the supply was too limited to meet demand. Recombinant DNA technology was used to coax bacteria into making this protein, but the cost of treatment was high ($30,000 a year). In 2004, a team led by Daniel Salamone at the University of Buenos Aires produced a transgenic cow that secretes hGH in her milk. The yield is prodigious: a mere 15 such cows could meet the worldwide demand of children suffering from this type of dwarfism.

Plants and plant cells can be genetically transformed and induced to make proteins. Recently the first drug produced by plant biotechnology was approved by the U.S. government. The drug is an enzyme used to treat Gaucher's disease, an inherited disorder that affects the breakdown of certain lipids in lysosomes. More than 10,000 patients worldwide may soon be using this enzyme, made by transgenic carrot cells.

DNA manipulation is changing agriculture

The cultivation of plants and the husbanding of animals provide the world's oldest examples of biotechnology, dating back more than 10,000 years. Over the centuries, people have adapted crops and farm animals to their needs. Through selective breeding of these organisms, desirable characteristics such as large seeds, high fat content in milk, or resistance to disease have been selected for and improved.

The traditional way to improve crop plants and farm animals was to identify individuals with desirable phenotypes that existed as a result of natural variation. Through deliberate crosses, the genes responsible for the desirable traits could be introduced into widely used varieties or breeds. Despite some spectacular successes, such as the breeding of high-yielding varieties of wheat, rice, and hybrid corn, such deliberate crossing

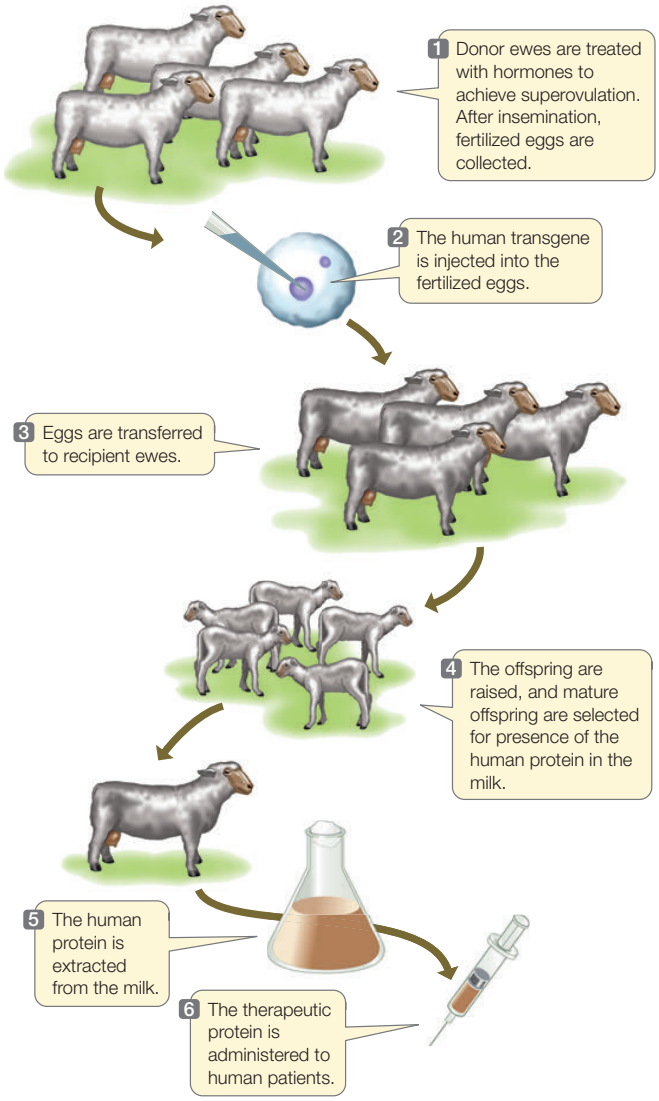


FIGURE 13.13 Pharming An expression vector carrying a desired gene can be put into an animal egg, which is implanted into a surrogate mother. The transgenic offspring produce the new protein in their milk. The milk is easily harvested and the protein isolated, purified, and made clinically available for patients.

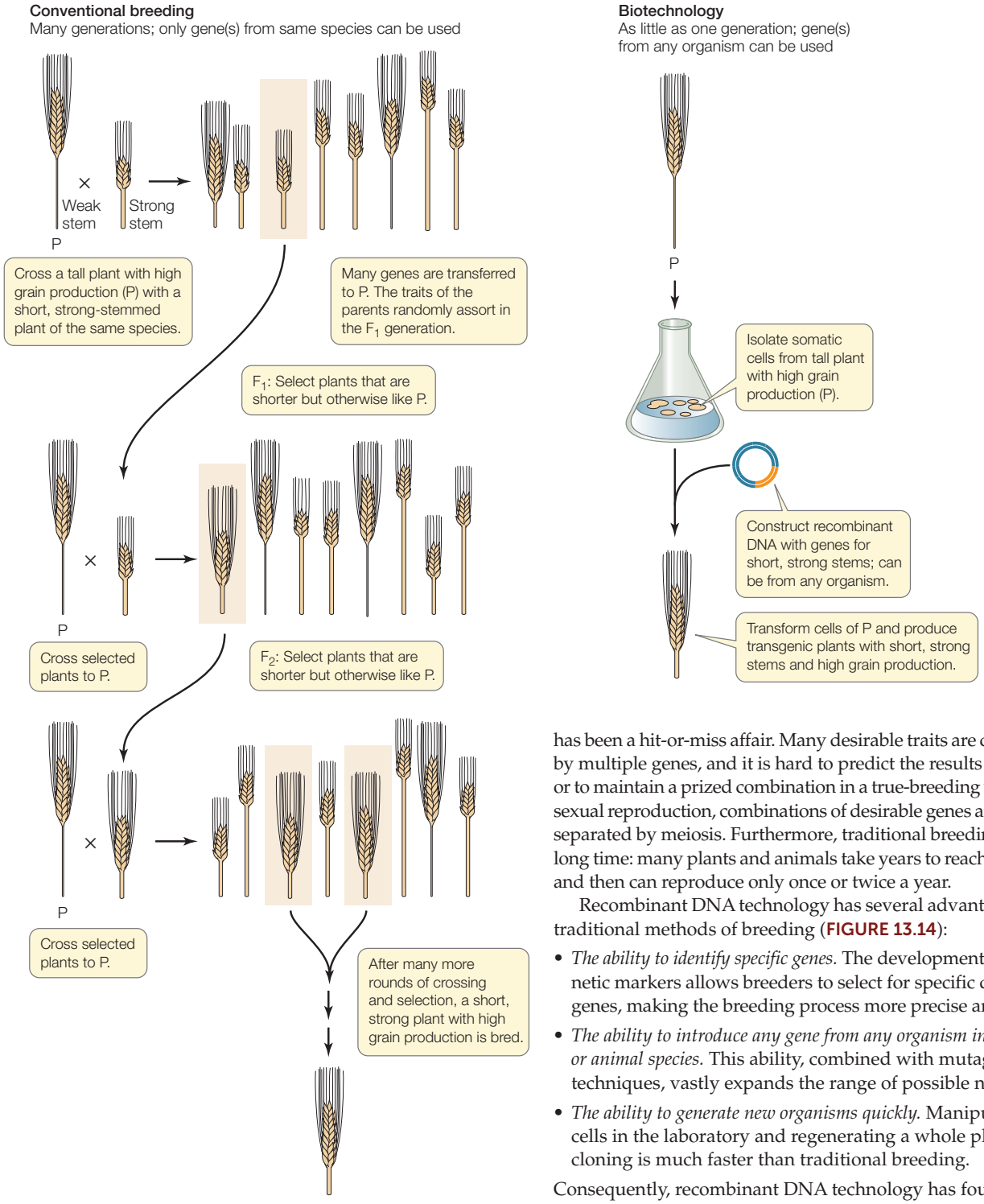


FIGURE 13.14 Genetic Modification of Plants versus Conventional Plant Breeding Plant biotechnology offers many potential advantages over conventional breeding. In the hypothetical example here, the objective is to transfer gene(s) for short, strong stems into a wheat plant that has high grain production but a tall, weak stem.

has been a hit-or-miss affair. Many desirable traits are controlled by multiple genes, and it is hard to predict the results of a cross or to maintain a prized combination in a true-breeding variety. In sexual reproduction, combinations of desirable genes are quickly separated by meiosis. Furthermore, traditional breeding takes a long time: many plants and animals take years to reach maturity and then can reproduce only once or twice a year.

Recombinant DNA technology has several advantages over traditional methods of breeding (**FIGURE 13.14**):

- *The ability to identify specific genes.* The development of genetic markers allows breeders to select for specific desirable genes, making the breeding process more precise and rapid.
- *The ability to introduce any gene from any organism into a plant or animal species.* This ability, combined with mutagenesis techniques, vastly expands the range of possible new traits.
- *The ability to generate new organisms quickly.* Manipulating cells in the laboratory and regenerating a whole plant by cloning is much faster than traditional breeding.

Consequently, recombinant DNA technology has found many applications in agriculture (**TABLE 13.2**). We will describe a few examples to demonstrate the approaches that plant scientists have used to improve crop plants.

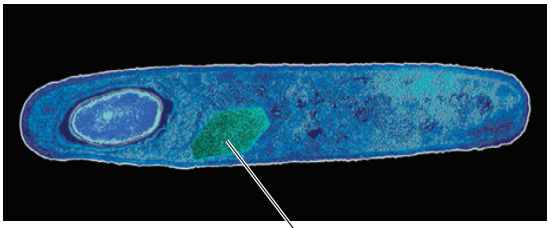
PLANTS THAT MAKE THEIR OWN INSECTICIDES From the locusts of biblical (and modern) times to the cotton boll weevil,

TABLE 13.2 Potential Agricultural Applications of Biotechnology

Problem	Technology/genes
Improving the environmental adaptations of plants	Genes for drought tolerance, salt tolerance
Improving nutritional traits	High-lysine seeds; β-carotene in rice
Improving crops after harvest	Delay of fruit ripening; sweeter vegetables
Using plants as bioreactors	Plastics, oils, and drugs produced in plants

insects have eaten the crops people grow. The development of insecticides improved the situation, but insecticides have their own problems. Many are relatively nonspecific and kill beneficial insects as well as crop pests. Some insecticides have toxic effects on other groups of organisms—including people—and persist in the environment for a long time.

Some bacteria generate proteins that can kill insects. For example, *Bacillus thuringiensis* produces a protein (Bt) that is toxic to the insect larvae that prey on this bacterium:



The Bt protein is 80,000 times more toxic than a typical commercial insecticide. When a hapless insect larva eats the bacteria, Bt is activated and binds specifically to the larva's gut, producing pores and killing the insect. Dried preparations of toxin-containing *B. thuringiensis* have been sold for decades as safe insecticides that break down rapidly in the environment. But the biodegradation of these preparations is also their limitation, because the dried bacteria must be applied repeatedly throughout the growing season.

A more permanent approach is to have the crop plants themselves produce the toxin. The Bt toxin gene has been isolated, cloned, and extensively modified by the addition of a plant promoter and other regulatory sequences. Transgenic (also called genetically modified or GM) corn, cotton, soybeans, tomatoes, and other crops are now being grown successfully with this added gene. Pesticide usage by farmers growing these transgenic crops is greatly reduced.

GRAINS WITH IMPROVED NUTRITIONAL CHARACTERISTICS To remain healthy, humans must consume adequate amounts of β-carotene, which the body converts into vitamin A. About 400 million people worldwide suffer from vitamin A deficiency, which makes them susceptible to infections and blindness. One reason is that rice grains, which do not contain β-carotene, make up a large part of their diets. Other parts

FIGURE 13.15 Transgenic Rice Rich in β-Carotene Right and middle: The grains from these transgenic rice strains are colored because they make the pigment β-carotene, which is converted to vitamin A in the human body. Left: Wild-type rice grains do not contain β-carotene.

of the rice plant—and indeed many plants and other organisms—contain enzymes for the biochemical pathway that leads to β-carotene production.

Plant biologists Ingo Potrykus and Peter Beyer isolated one of the genes for the β-carotene pathway from the bacterium *Erwinia uredovora*, and another from daffodil plants. They added a promoter and other signals for expression in the developing rice grain, and then transformed rice plants with the two genes. The resulting rice plants produce grains that look yellow because of their high β-carotene content. A newer variety with a corn gene replacing the one from daffodils makes even more β-carotene and is golden in color (**FIGURE 13.15**). A daily intake of about 150 grams of this cooked rice can supply all the β-carotene a person needs. This new transgenic strain has been crossed with strains adapted for various local environments, in the hope of improving the diets of millions of people.

CROPS THAT ADAPT TO THE ENVIRONMENT Agriculture depends on ecological management—tailoring the environment to the needs of crop plants and animals. A farm field is an unnatural, human-designed system that must be carefully managed to maintain optimal conditions for crop growth. For example, repeated irrigation meets the water needs of crops but results in increases in soil salinity. The Fertile Crescent, the region between the Tigris and Euphrates rivers in the Middle East where agriculture was practiced 10,000 years ago, is no longer fertile. It is now a desert, largely because the soil has a high salt concentration. Few plants can grow on salty soils, partly because of osmotic effects that result in wilting, and partly because excess salt ions (Na^+) are toxic to plant cells.

Some soils are naturally salty, and some plants can tolerate salty soils because they have a protein that transports Na^+ out of the cytoplasm and into the vacuole, where the ions can accumulate without harming plant growth (see Concept 4.3 for a description of the plant vacuole). Scientists have developed a highly active version of the gene encoding this protein and have used it to transform crop plants—including rapeseed, wheat, and tomatoes—that are normally less tolerant to salt. When the gene was added to tomato plants, they grew in water that was four times more salty than the typical lethal level

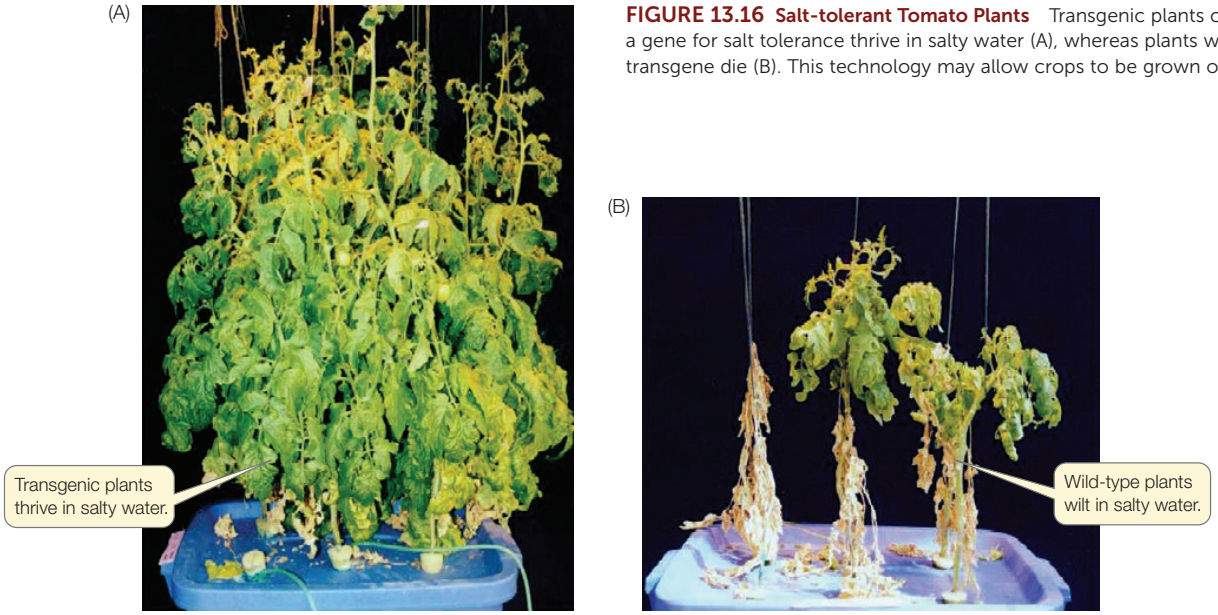


FIGURE 13.16 Salt-tolerant Tomato Plants Transgenic plants containing a gene for salt tolerance thrive in salty water (A), whereas plants without the transgene die (B). This technology may allow crops to be grown on salty soils.

(**FIGURE 13.16**). This finding raises the prospect of growing useful crops on previously unproductive soils.

LINK

Other stress adaptations of plants—some of which could be beneficial if introduced into crop plants—are discussed in **Concept 28.3**

The example described here illustrates what could become a fundamental shift in the relationship between crop plants and the environment. *Instead of manipulating the environment to suit the plant, biotechnology may allow us to adapt the plant to the environment.* As a result, some of the negative effects of agriculture, such as water pollution, could be lessened.

There is public concern about biotechnology

Concerns have been raised about the safety and wisdom of genetically modifying crops and other organisms. These concerns are centered on three claims:

- Genetic manipulation is an unnatural interference with nature.
- Genetically altered foods are unsafe to eat.
- Genetically altered crop plants are dangerous to the environment.

Advocates of biotechnology tend to agree with the first claim. However, they point out that all crops are unnatural in the sense that they come from artificially bred plants growing in a manipulated environment (a farmer's field). Recombinant DNA technology just adds another level of sophistication to these technologies.

To counter the concern about whether genetically engineered crops are safe for human consumption, biotechnology advocates point out that only single genes are added and that these genes are specific for plant function. For example, the *B. thuringiensis* toxin produced by transgenic plants has no effect on people. However, as plant biotechnology moves from adding genes that improve plant growth to adding genes that affect human nutrition, such concerns will become more pressing.

Various negative environmental impacts have been envisaged. There is concern about the possible "escape" of transgenes from crops to other species. If the gene for herbicide resistance, for example, were inadvertently transferred from a crop plant to a closely related weed, that weed could thrive in herbicide-treated areas. Another negative scenario is the possibility that increased use of an herbicide will select for weeds with naturally occurring mutations that make them resistant to that herbicide. This is indeed occurring. Widespread use of glyphosate on fields of glyphosate-resistant crops has resulted in the selection of rare mutations in weeds that make them resistant to glyphosate. To date more than ten resistant weed species have appeared in the U.S.

As we mentioned in the opening story, there are biotechnologically produced microorganisms that are able to break down components of crude oil. These have not been released into the environment because of the unknown effects that such organisms might have on natural ecosystems. However, these organisms potentially provide a way to rapidly clean up catastrophic oil spills.

Because of the potential benefits of biotechnology, most scientists believe that it is wise to proceed, albeit with caution.

CHECKpoint CONCEPT 13.4

- ✓ In addition to the coding sequence for a gene of interest, what other DNA sequences are required for the gene to be expressed in a different host?
- ✓ What is pharming, how is it done, and what are its advantages over more conventional biotechnology approaches?
- ✓ What are the advantages of using biotechnology for plant breeding compared with traditional methods?
- ✓ What are some concerns people might have about agricultural biotechnology?

Q How is biotechnology used to alleviate environmental problems?

ANSWER Among the thousands of species of bacteria, there are many unique enzymes and biochemical pathways. New pathways are continually being discovered as new bacterial species are found (for example, by metagenomics; see Figure 12.6). Bacteria are natural recyclers, thriving on many types of nutrients—including what humans refer to as wastes.

Bioremediation is the use by humans of other organisms to remove contaminants from the environment. Two well-known examples of bioremediation are composting and wastewater treatment. Composting involves the use of bacteria and other microbes to break down large molecules, including carbon-rich polymers and proteins, in waste products such as wood chips, paper, straw, and kitchen scraps. For example, some species of bacteria make cellulase, an enzyme that hydrolyzes cellulose, a major component of plant cell walls and paper. Bacteria are used in wastewater treatment plants to break down human wastes, paper products, and household chemicals.

Bioremediation is also an attractive option for cleaning up oil spills. In 2010 the Deepwater Horizon oil rig in the Gulf of Mexico exploded, and crude oil began gushing out of the well below it. The oil flowed unabated for three months until the wellhead was finally capped; a total of 210 million gallons were released. The oil slick was visible from space—it damaged marine habitats and killed wildlife, washed up on beaches, and shut down fishing and tourism in the area. Efforts to remove the giant slick included collecting the oil, burning it, and dispersing it chemically. In addition, scientists identified bacteria living in the Gulf waters that were able to digest and break down components of the crude oil. Despite the clean-up efforts and the actions of these bacteria, much of the oil remains in the Gulf today.

FIGURE 13.17 The Spoils of War Massive oil spills occurred in Kuwait during the 1990–1991 Gulf War.

Oil spills can be remediated by encouraging the growth of natural microorganisms that digest components of crude oil. After the oil tanker *Exxon Valdez* ran aground near the Alaskan shore in 1989, nitrogen fertilizers were applied to nearby beaches to encourage the growth of oil-eating bacteria. Similar approaches have been tried in Kuwait, where the destruction of oil wells during the 1990–1991 Gulf War led to a massive release of oil (**FIGURE 13.17**). The success of such methods has been limited, however, because the naturally occurring bacteria digest only certain components of the crude oil, and because of technical difficulties in bringing the organisms into contact with the oil.

In the opening story we described how conventional genetic crosses were used to produce (and patent) bacteria that have the capacity to break down oil. Since that time, scientists have isolated the genes that encode the oil-degrading enzymes from such bacteria and cloned them into vectors (Concepts 13.2 and 13.3). This DNA has been used to transform several species of bacteria that live in environments where oil spills have occurred. However, as we mentioned in Concept 13.4, these genetically modified bacteria have not been released into the environment because of environmental concerns.

SUMMARY

CONCEPT 13.1 Recombinant DNA Can Be Made in the Laboratory

- **Biotechnology** is the use of living cells or their derivatives to make or modify materials and processes useful to people.
- **Restriction enzymes** make cuts in double-stranded DNA, creating fragments of various lengths.
- DNA fragments can be separated by size using **gel electrophoresis**. **Review Figure 13.2 and ANIMATED TUTORIAL 13.1**
- DNA fragments from different sources can be used to create **recombinant DNA** by joining them together using **DNA ligase**. **Review Figures 13.3 and 13.4**

CONCEPT 13.2 DNA Can Genetically Transform Cells and Organisms

- One goal of recombinant DNA technology is to **clone** a particular gene, either for analysis or to produce its protein product in quantity.
- Bacteria, yeasts, and cultured plant and animal cells are commonly used as hosts for recombinant DNA. The insertion of foreign DNA into host cells is called **transformation**, or if the host cells are derived from an animal, **transfection**.
- Various methods are used to get recombinant DNA into cells. These include chemical and physical treatments for plasmids and the use of viral **vectors**.
- **Selectable markers** such as genes for antibiotic resistance are used to select for host cells that have taken up a foreign gene. **Review Figure 13.5**
- **Reporter genes** (of which selectable markers are one type) are genetic markers with easily identifiable phenotypes. **Review Figure 13.6**

CONCEPT 13.3 Genes Come from Various Sources and Can Be Manipulated

- DNA fragments from a genome can be inserted into host cells to create a **genomic library**. A **cDNA library** is made by reverse transcribing mRNA to make cDNA. **Review Figure 13.7**
- Synthetic DNA containing any desired sequence can be made in the laboratory.
- Manipulating gene expression is one way to study the functions of particular genes.
- **Homologous recombination** is used to knock out a gene in a living organism. **Review Figure 13.8**
- Gene silencing techniques using miRNA or siRNA are used to prevent the translation of genes. **Review Figure 13.9**
- DNA microarray technology permits the screening of thousands of cDNA sequences at the same time. **Review Figure 13.10 and ANIMATED TUTORIAL 13.2**

CONCEPT 13.4 Biotechnology Has Wide Applications

- **Expression vectors** allow transgenes to be expressed in host cells. **Review Figure 13.11 and ACTIVITY 13.1**
- Recombinant DNA techniques have been used to make medically useful proteins. **Review Figure 13.12**
- **Pharming** is the use of transgenic plants or animals to produce pharmaceuticals. **Review Figure 13.13**
- Transgenic crop plants can be adapted to their environments, rather than vice versa. **Review Figure 13.16**
- There is public concern about the application of recombinant DNA technology to food production.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is13

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

14

Genes, Development, and Evolution

KEY CONCEPTS

Scanning electron micrographs of a normal fruit fly (left) and an *eyeless* mutant fly of the same species (right).

Eyes are not essential for survival; many animals and all plants get by just fine without them. However, more than 90 percent of all animals *do* have eyes or some type of light-sensing organs, and having eyes can confer a selective advantage. About a dozen different kinds of eyes are found among the animals, including the camera-like eyes of humans and the compound eyes of insects, with their hundreds or thousands of individual units. In trying to understand how this variety came about, scientists—starting with Charles Darwin—proposed that eyes evolved independently many times in different animal groups, and that each improvement in the ability of eyes to gather light and form images conferred a selective advantage on their possessor.

Our understanding of the evolution of eyes remained at this level until 1915, when a mutant fruit fly without eyes was found and the gene involved, appropriately called *eyeless*, was mapped on one of the fly's chromosomes. This mutant fly remained a laboratory curiosity until the 1990s, when the Swiss developmental biologist Walter Gehring and his colleagues began looking for genes expressed in the fly embryo and found one that mapped to the *eyeless* locus. By transforming flies with recombinant DNA containing the *eyeless* gene, they showed that this gene's product controls the formation of the eye. Expression of the *eyeless* gene was then manipulated so that it was expressed in different parts of the body. In this way, eye structures were induced on legs, wings, and antennae of different flies. It did not matter where on the body the gene was expressed; if the gene was active, an eye developed there.

An even bigger surprise was in store when database searches revealed that the *eyeless* gene sequence was similar to that of *Pax6*, a mouse gene that, when mutated, leads to the development of very small eyes. Could the very different eyes of flies and mice result from simple variations on a common developmental theme? To test for functional similarity between the insect and mammalian genes, Gehring's team repeated their experiments on flies, but using the mouse *Pax6* gene instead of the fly *eyeless* gene. Once again, eyes developed. A gene whose expression normally leads to the development of a mammalian "camera" eye now led to the development of an insect's compound eye—a very different eye type. Thus a single gene product appears to function as a molecular switch that turns on eye development in diverse animals. Because of this and other recent findings, it now seems likely that all eye types evolved from a common origin very early in animal evolution.

Q How do gene products control the development of the eye?

You will find the answer to this question on page 294.

CONCEPT 14.1 Development Involves Distinct but Overlapping Processes

Development is the process by which a multicellular organism, beginning with a single cell, goes through a series of changes, taking on the successive forms that characterize its life cycle (FIGURE 14.1). After the egg is fertilized it is called a **zygote**, and in the earliest stages of development a plant or animal is called an **embryo**. Progress through a series of embryonic stages precedes emergence of the new, independent organism. Many organisms continue to develop throughout their lives, with development ceasing only at death.

Four key processes underlie development

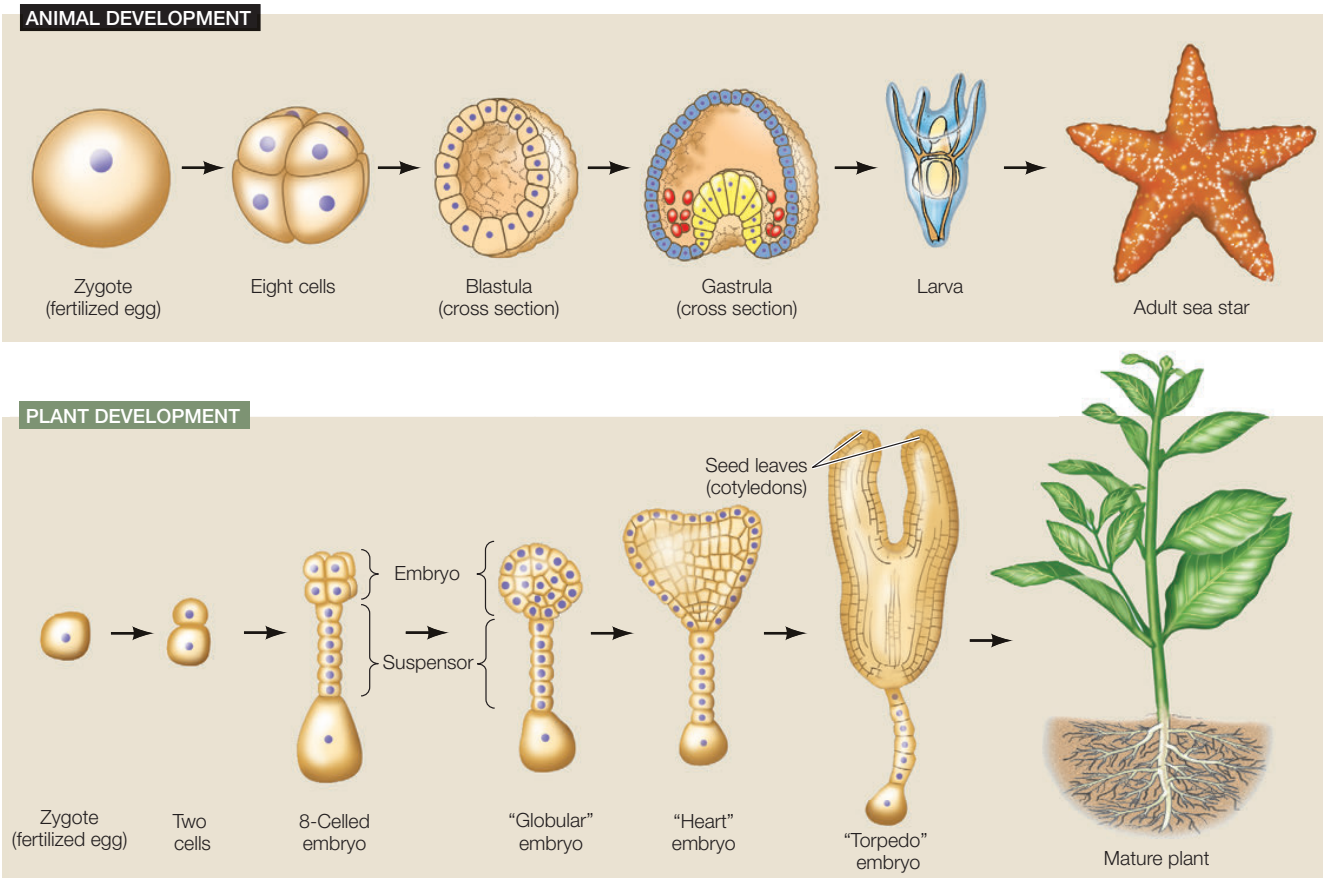
The developmental changes an organism undergoes as it progresses from an embryo to mature adulthood involve four processes:

- **Determination** sets the developmental fate of a cell—what type of cell it will become—even before any characteristics of that cell type are observable. For example, in a developing mammalian embryo, as well as in some adult organs, there are mesenchymal stem cells that look unspecialized. But their fate to become muscle, fat, tendon, or other connective tissue cells has already been determined.
- **Differentiation** is the process by which different types of cells arise from less specialized cells, leading to cells with specific structures and functions. For example, mesenchymal stem cells differentiate to become the cells listed above.
- **Morphogenesis** (Greek for "origin of form") is the organization and spatial distribution of differentiated cells into the multicellular body and its organs. Morphogenesis can occur by cell division, cell expansion (especially in plants), cell movements, and apoptosis (programmed cell death).
- **Growth** is the increase in size of the body and its organs by cell division and cell expansion. Growth can occur by an increase in the number of cells or by the enlargement of existing cells. Growth continues throughout the individual's life in some organisms, but reaches a more or less stable end point in others.

All of these processes involve differential gene expression. The cells that arise from repeated mitoses in the early embryo may look the same superficially, but they soon begin to differ

FIGURE 14.1 Development Selected stages and processes of development from zygote to maturity are shown for an animal and for a plant. The blastula is a hollow sphere of cells; the gastrula has three cell layers (indicated by blue, red, and yellow).

Go to ACTIVITY 14.1 Stages of Development
PoL2e.com/ac14.1



in terms of which genes they express. These processes also involve signals between cells. For example, within a developing embryo there are gradients of signaling molecules called morphogens that help determine cell fate and trigger cell differentiation.

Cell fates become progressively more restricted during development

A zygote is a single cell that gives rise to all the cells in the organism that will develop from it. As the zygote divides to form a multicellular embryo, each of the embryo's undifferentiated cells are destined to become part of a particular type of tissue—this is referred to as the **cell fate** of that undifferentiated cell.

When is the fate determined? At what point does a cell become committed to a particular fate and no other? One way to find out is to transplant cells from one embryo to a different region of a recipient embryo (FIGURE 14.2). A dye is used to mark the transplanted cells so that their subsequent development can be followed. The question is, will the transplanted cells adopt the differentiation pattern of their new surroundings, or will they continue on their own path, with their fate already sealed?

Experiments with frog embryos give an answer for that organism: If the donor tissue is from an early-stage embryo (blastula), it adopts the fate of its new surroundings; its fate *has not* been sealed. But if the donor tissue is from an older embryo (gastrula), it continues on its own path; its fate *has* been sealed. Determination is influenced by changes in gene expression as well as by the extracellular environment and is not something that is visible under the microscope—cells do not change their appearance when they become determined. Determination is followed by differentiation—the actual changes in biochemistry, structure, and function that result in cells of different types. *Determination is a commitment; the final realization of that commitment is differentiation.*

During animal development, cell fate becomes progressively restricted. This can be thought of in terms of **cell potency**, which is a cell's potential to differentiate into other cell types:

- The cells of an early embryo are **totipotent** (*toti*, "all"; *potent*, "capable"); they have the potential to differentiate into any cell type, including more embryonic cells.
- In later stages of the embryo, many cells are **pluripotent** (*pluri*, "many"); they have the potential to develop into most other cell types, but they cannot form new embryos.
- Through later developmental stages, including adulthood, certain stem cells are **multipotent**; they can differentiate into several different, related cell types. Mesenchymal stem cells (see above) are one kind of multipotent stem cell.
- Many cells in the mature organism are **unipotent**; they can produce only one cell type—their own.

Go to ANIMATED TUTORIAL 14.1
Cell Fates
PoL2e.com/at14.1

Cell differentiation is sometimes reversible

Once a cell's fate is determined, the cell differentiates. However, under the right experimental conditions, a determined or differentiated cell can become undetermined again. In some cases the cell can even become totipotent, meaning it is able to form the entire organism, with all of its differentiated cells. Normally this is a property of only the zygote or, in some cases, the first few cells of the early embryo.

PLANT CELL TOTIPOTENCY A carrot root cell normally faces a dark future. It cannot photosynthesize and generally does not give rise to new carrot plants. However, in 1958 Frederick Steward at Cornell University showed that if he isolated cells from a carrot root and maintained them in a suitable nutrient medium, he could induce them to dedifferentiate—to lose their differentiated characteristics. The cells could divide and give rise to masses of undifferentiated cells called calli (singular callus), which could be maintained in culture indefinitely. Furthermore, if they were provided with the right chemical cues, the cells could develop into embryos and eventually into complete new plants (FIGURE 14.3).

Since the new plants in Steward's experiments were genetically identical to the cells from which they came, they were clones of the original carrot plant. The ability to produce clones is evidence for the **genomic equivalence** of somatic (body) cells; that is, all somatic cells in a plant have a complete genome and thus have all the genetic information needed to become any cell in the plant.

Many types of cells from other plant species show similar behavior in the laboratory.

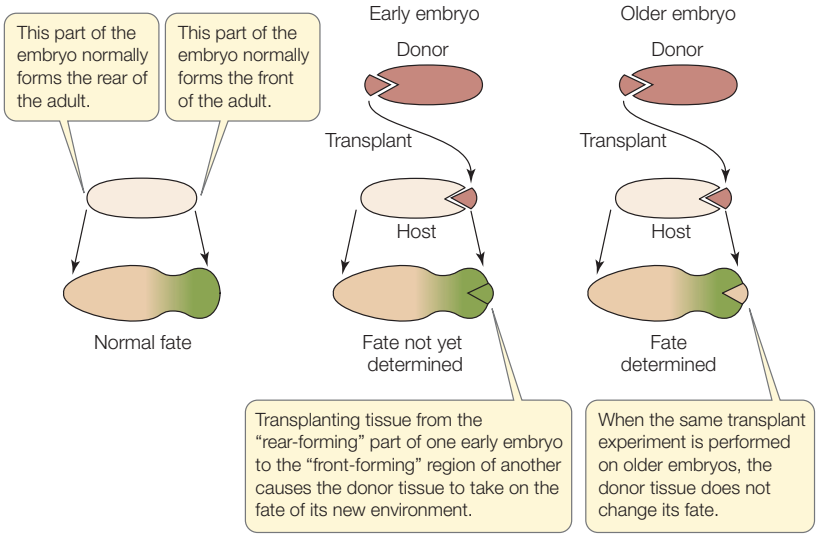


FIGURE 14.2 A Cell's Fate Is Determined in the Embryo Transplantation experiments using frog embryos show that the fate of cells is determined as the early embryo develops.

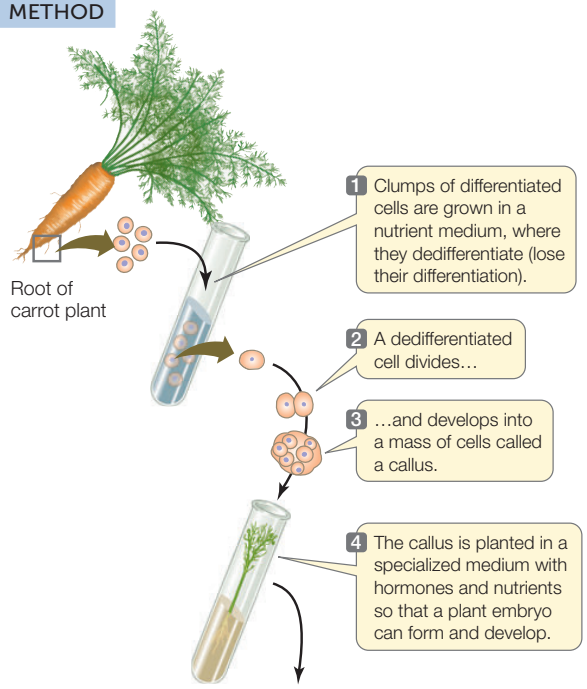
INVESTIGATION

FIGURE 14.3 Cloning a Plant When cells were removed from a plant and put into a medium with nutrients and hormones, they lost many of their specialized features—they dedifferentiated and became totipotent.[a]

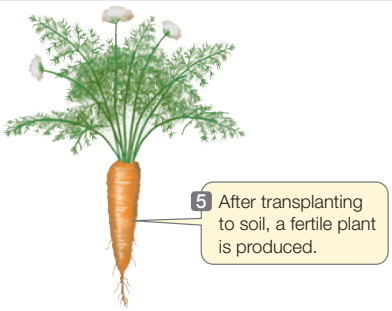
HYPOTHESIS

Differentiated plant cells can be totipotent and can be induced to generate an entire new plant.

METHOD



RESULTS



CONCLUSION

Differentiated plant cells can be totipotent.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]F. C. Steward et al. 1958. *American Journal of Botany* 45: 705–709.

This ability to generate a whole plant from groups of cells or even a single cell has been invaluable in agriculture and forestry. For example, trees from planted forests are used in making paper, lumber, and other products. To replace the trees reliably, forestry companies regenerate new trees from the leaves of selected trees with desirable traits. The characteristics of these clones are more uniform and predictable than those of trees grown from seeds.

NUCLEAR TOTIPOTENCY IN ANIMALS Animal somatic cells cannot be manipulated as easily as plant cells can. Until recently, it was not possible to induce a cell from a fully developed animal to dedifferentiate and then redifferentiate into another cell type. However, nuclear transfer experiments have shown that the genetic information from a differentiated animal cell can be used to create cloned animals. The nucleus from an unfertilized egg is removed, forming an enucleated egg. A donor nucleus from a somatic cell is then introduced into the "empty" egg. If it is then stimulated to divide, the egg forms an embryo that can develop into an adult with the genetic composition of its nuclear donor. This is the basis of cloning animals. Dolly the sheep was the first experimentally produced mammalian clone, born in 1996 (**FIGURE 14.4**).

Many other animal species, including cats, dogs, horses, pigs, rabbits, and mice, have since been cloned by nuclear transfer. As in plants, the cloning of animals has shown that their differentiated cells have genomic equivalence. Cloning of animals has practical uses as well:

- Expansion of the numbers of valuable animals: One goal of the researchers who produced Dolly the sheep was to develop a method of cloning transgenic animals with useful phenotypes (see Concept 13.4). For example, a cow that was genetically engineered to make human growth hormone in milk has been cloned to produce two more cows that do the same thing. Only 15 such cows could supply the world's need for this protein, which is used to treat short stature that is due to growth hormone deficiency.
- Preservation of endangered species: The banteng, a relative of the cow, was the first endangered animal to be cloned and survive. The banteng was made using the enucleated egg from a cow, the nucleus from a banteng cell, and a cow surrogate mother. Cloning may be the only way to save endangered species with low rates of natural reproduction.
- Resurrection of extinct species. With the discoveries of intact DNA in fossils, the once-fictional idea of cloning an extinct species is becoming a possibility. In 2000, the last Pyrenian ibex—a type of mountain goat—died and the species became extinct. But in 2009, scientists used DNA from the dead animal to replace the DNA in a domestic goat egg and cloned a new ibex. Although this animal died shortly after birth, the resurrection of an extinct species was proven in principle. This has led to proposals to resurrect other extinct species, including the wooly mammoth and the Neanderthal. The genomes of both these species are available and have been sequenced.

Stem cells differentiate in response to environmental signals

The processes of development do not occur only in embryos. In adult plants, the growing regions at the tips of roots and stems

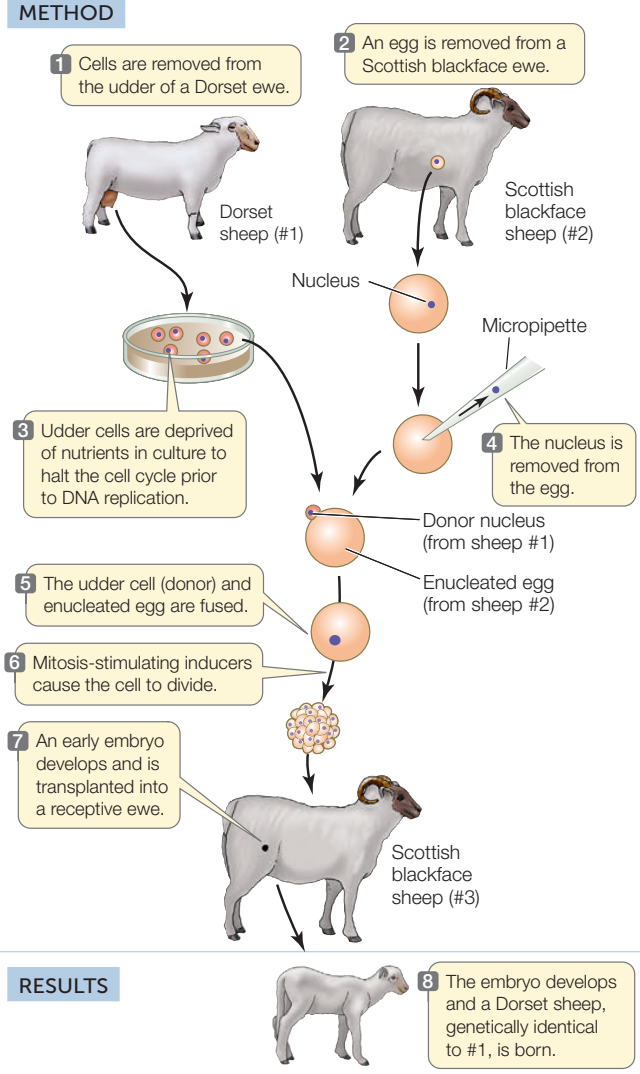
INVESTIGATION

FIGURE 14.4 Cloning a Mammal The experimental procedure described here was used to produce the first cloned mammal, a Dorset sheep named Dolly (shown on the left in the photo). As an adult, Dolly mated and subsequently gave birth to a normal offspring (the lamb on the right), thus proving the genetic viability of cloned mammals.[a]

HYPOTHESIS

Differentiated animal cells are totipotent.

METHOD



RESULTS

CONCLUSION

Differentiated animal cells are totipotent in nuclear transplant experiments.

ANALYZE THE DATA

The team that cloned Dolly the sheep used a nucleus from a mammary epithelium (ME) cell. They also tried cloning by transplanting nuclei from fetal fibroblasts (FB) and embryos (EC), with the results shown in the table.

	Number of attempts that progressed to each stage		
Stage	ME	FB	EC
Egg fusions	277	172	385
Embryos transferred to recipients	29	34	72
Pregnancies	1	4	14
Live lambs	1	2	4

A. Calculate the percentage survival of eggs from fusion to birth. What can you conclude about the efficiency of cloning?

B. Compare the efficiencies of cloning using different nuclear donors. What can you conclude about the ability of nuclei at different stages to be totipotent?

C. What statistical test would you use to show whether the differences in A and B were significant (see Appendix B)?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a] I. Wilmut et al. 1997. *Nature* 385: 810–813.

contain meristems, which are clusters of undifferentiated, rapidly dividing stem cells. These cells can differentiate into the 15–20 specialized cell types that make up roots, stems, leaves, and flowers. As you will see in Chapter 26, the plant body undergoes constant growth and renewal, with new organs forming often. (Think of flowers and leaves in the spring.)

In adult mammals, stem cells persist in many tissues, where they are used as a pool of cells that can differentiate and replace

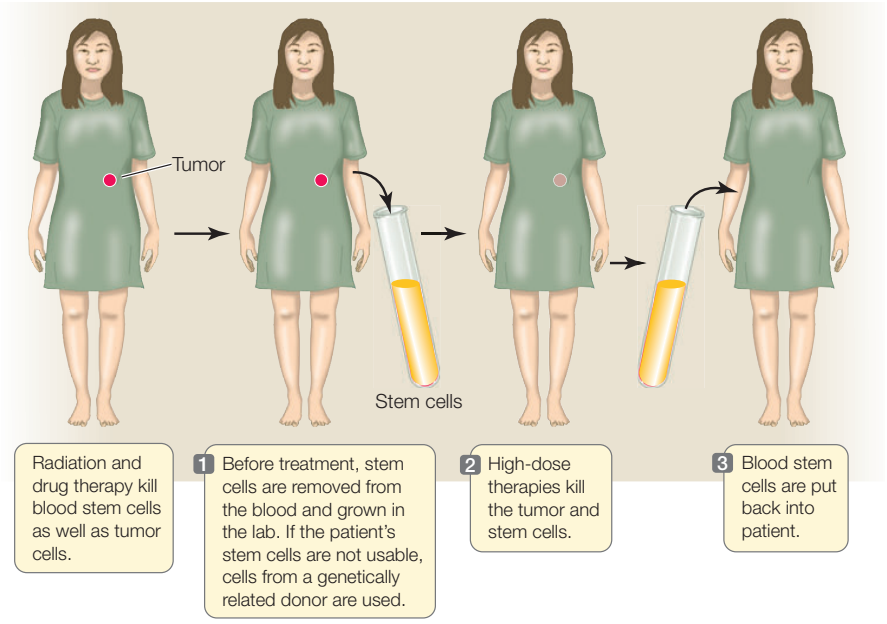


FIGURE 14.5 Multipotent Stem Cells In hematopoietic stem cell transplantation, blood stem cells are used to replace stem cells destroyed by cancer therapy.

cells that are lost by "wear and tear" (necrosis) and programmed cell death (apoptosis). This is especially evident in tissues such as the skin, inner lining of the intestine, and blood. There are about 300 different cell types in a mammal.

MULTIPOTENT STEM CELLS Stem cells in particular mammalian tissues are multipotent, meaning they can form a limited repertoire of differentiated cells. For example, there are two types of multipotent stem cells in bone marrow. One type (called hematopoietic stem cells) produces the various kinds of red and white blood cells. The other type (mesenchymal stem cells) produces the cells that make bone and surrounding tissues, including muscle.

The proliferation and differentiation of multipotent stem cells is "on demand." Hematopoeitic stem cells in the bone marrow, for example, differentiate in response to specific signals. These signals can come from either adjacent bone marrow cells or from the circulating blood. This ability to differentiate is the basis of an important cancer therapy called hematopoietic stem cell transplantation (HSCT; **FIGURE 14.5**). Some treatments that kill cancer cells also kill other dividing cells, including the stem cells in the bone marrow of patients exposed to these treatments. To circumvent this problem, stem cells are harvested from the blood or bone marrow of the patient prior to treatment or from a donor; the cells are injected back into the patient after cancer treatment. Before the cells are harvested, the patient (or donor) receives injections of a growth factor that stimulates proliferation of the

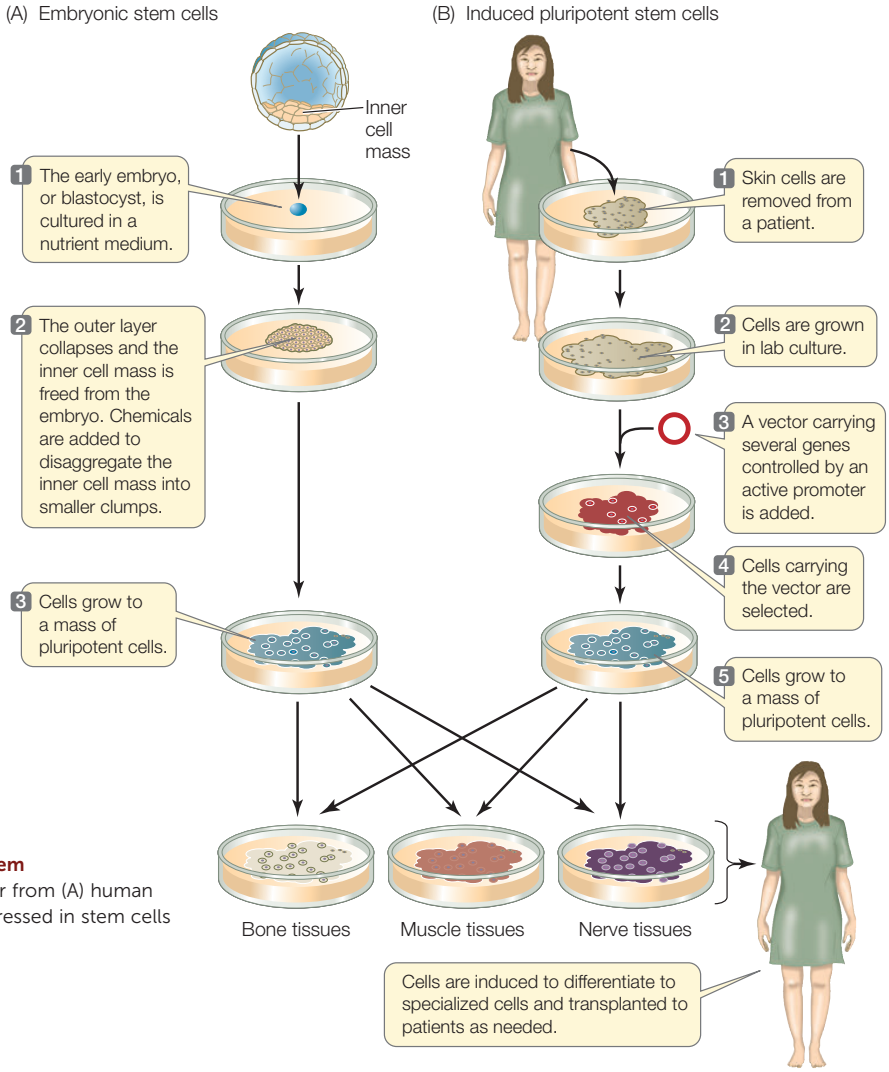


FIGURE 14.6 Two Ways to Obtain Pluripotent Stem Cells Pluripotent stem cells can be obtained either from (A) human embryos or (B) by adding genes that are highly expressed in stem cells to skin cells to transform them into stem cells.

Go to ANIMATED TUTORIAL 14.2
Embryonic Stem Cells
PoL2e.com/at14.2

hematopoietic stem cells. The stored stem cells retain their ability to differentiate in the bone marrow environment. By allowing the use of high doses of treatment to kill tumors, HSCT saves thousands of lives each year.

PLURIPOTENT STEM CELLS In mammals, totipotent stem cells that can individually give rise to an organism are found only in very early embryos. In both mice and humans, the last embryonic stage before differentiation occurs is called a blastocyst (the term for a mammalian blastula; see Figures 14.1 and 38.8). Although they cannot form an entire embryo, a group of cells in the blastocyst still retains the ability to form any cell type in the body; these cells are pluripotent. These **embryonic stem cells** (**ESCs**) can be removed from the blastocyst and grown in laboratory culture almost indefinitely if provided with the right conditions. They can also be induced to express appropriate genes and differentiate in a particular way if the right signal is provided (FIGURE 14.6A). For example, treatment of mouse ESCs with a derivative of vitamin A causes them to form neurons (nerve cells), whereas other growth factors induce them to form blood cells. Such experiments demonstrate both the cells' developmental potential and the roles of environmental signals. This finding raises the possibility of using ESC cultures as sources of differentiated cells to repair specific tissues, such as a damaged pancreas in diabetes, or a brain that malfunctions in Parkinson's disease.

ESCs can be harvested from human embryos conceived by in vitro ("under glass"—in the laboratory) fertilization, with the consent of the donors. Since more than one embryo is usually conceived in this procedure, embryos not used for reproduction might be available for embryonic stem cell isolation. These cells could then be grown in the laboratory and used as sources of tissues for transplantation into patients with tissue damage. There are two problems with this approach:

- Some people object to the destruction of human embryos for this purpose.
- The stem cells, and tissues derived from them, would provoke an immune response in a recipient (see Chapter 39).

Shinya Yamanaka and coworkers at Kyoto University in Japan developed another way to produce pluripotent stem cells that applies the concepts of gene expression and development (FIGURE 14.6B). Instead of extracting ESCs from blastocysts, they make **induced pluripotent stem cells** (**iPS cells**) from skin cells. This approach destroys no embryos and allows tissues to be made from skin cells of any individual, thus preventing an immune response. The scientists developed this method systematically (see Chapter 13 for more information on the techniques discussed here):

1. First, they used microarrays to compare the genes expressed in ESCs with those expressed in nonstem cells. They found several genes that were uniquely expressed at high levels in ESCs. These genes encode transcription factors believed to be essential to the undifferentiated state and function of stem cells. Recall that transcription factors are DNA binding proteins that regulate the expression of specific genes.
2. Next, they isolated the genes and inserted them into a vector for genetic transformation of skin cells. They found that the skin cells now expressed the newly added genes at high levels.
3. Finally, they showed that the transformed cells were pluripotent and could be induced to differentiate into many tissues—they had become iPS cells.

Yamanaka was awarded the Nobel Prize for his work. The ultimate aim is to use the cells for research and therapy in diseases.

CHECKpoint CONCEPT 14.1

- ✓ Describe the four major processes of development.
- ✓ Not all the DNA in a cell is in the nucleus. What are the genetic differences between cloning in carrot plants and cloning in sheep? How would you show this?
- ✓ Identical twins are formed when a zygote divides once by mitosis and then each mitotic product forms an embryo. Are identical twins clones? Explain your answer.

Having considered the general principles of development, we will now turn to the mechanisms that govern developmental events. Not surprisingly, these mechanisms have been studied at the molecular level and involve changes in gene expression and the activities of specific proteins.

CONCEPT 14.2 Changes in Gene Expression Underlie Cell Fate Determination and Differentiation

Virtually every cell of an individual organism contains a complete copy of the organism's genome. Each cell, however, expresses only a subset of these DNA sequences. For example, certain cells in hair follicles produce keratin, the protein that makes up hair, whereas other cell types in the body do not. In Chapter 5 we discussed cell-signaling pathways, many of which result in changes in gene expression. In Chapter 11 we described several mechanisms by which cells control gene expression—by controlling transcription and translation, and by making posttranslational protein modifications. As we mentioned in Concept 14.1, all four processes of development—determination, differentiation, morphogenesis, and growth—involve changes in gene expression, and these changes often result from signaling between cells. In this concept we focus on the processes of cell fate determination and differentiation.

The most fundamental decisions in development are generally controlled at the level of transcription. Genes that determine cell fate and trigger differentiation (often by regulating the expression of other genes on other chromosomes) usually encode transcription factors. In some cases, a single transcription factor can cause a cell to differentiate in a certain way. In other cases, multiple interactions between genes and proteins set off a sequence of transcriptional events that leads to

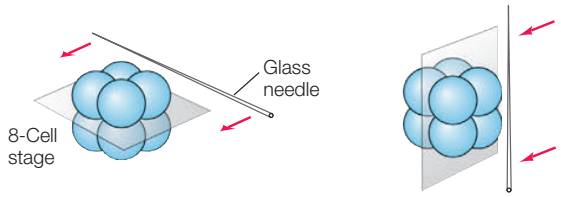
differential gene expression. There are two ways in which cell fate can be determined:

- by the asymmetrical distribution of cytoplasmic factors inside a cell, so that its two progeny cells receive unequal amounts of the factors, or
- by the differential exposure of two cells to an external signal (an inducer; see p. 282).

Cell fates can be determined by cytoplasmic polarity

An early event in development is often the establishment of axes that relate to the body plan of the organism. For example, an embryo may develop a distinct "top" and "bottom" corresponding to what will become opposite ends of the mature organism; such a difference is called **polarity**. Many examples of polarity are observed as development proceeds. Our heads are distinct from our rear ends, and the distal (far from the center) ends of our arms and legs (wrists, ankles, fingers, toes) differ from the proximal (near) ends (shoulders and hips).

Polarity may develop early; even within the egg, the yolk and other factors are often distributed asymmetrically. During early development in animals, polarity is specified by an "animal pole" at the top of the zygote and a "vegetal pole" at the bottom. This polarity can lead to the determination of cell fates at a very early stage of development. For example, sea urchin embryos can be bisected at the eight-cell stage in two different ways:



If the two halves (with four cells each) of these embryos are allowed to develop, the results are dramatically different for the two different cuts:

- For an embryo cut into a top half and a bottom half (left, above), the bottom half develops into a small sea urchin and the top half does not develop at all.
- For an embryo cut into two side halves (right, above), both halves develop into normal, though smaller, sea urchins.

These results indicate that the top and bottom halves of an eight-cell sea urchin embryo have already developed distinct fates. These observations led to the model of **cytoplasmic segregation**, which states that certain materials, called **cytoplasmic determinants**, are distributed unequally in the egg cytoplasm (FIGURE 14.7). During the early cell divisions of the embryo's development, the progeny cells receive unequal amounts of these determinants. Cytoplasmic determinants include specific transcription factors that promote differential gene expression in the two daughter cells. They also include small regulatory RNAs and mRNAs, which also contribute to differential gene expression. What accounts for the unequal distribution of these determinants?

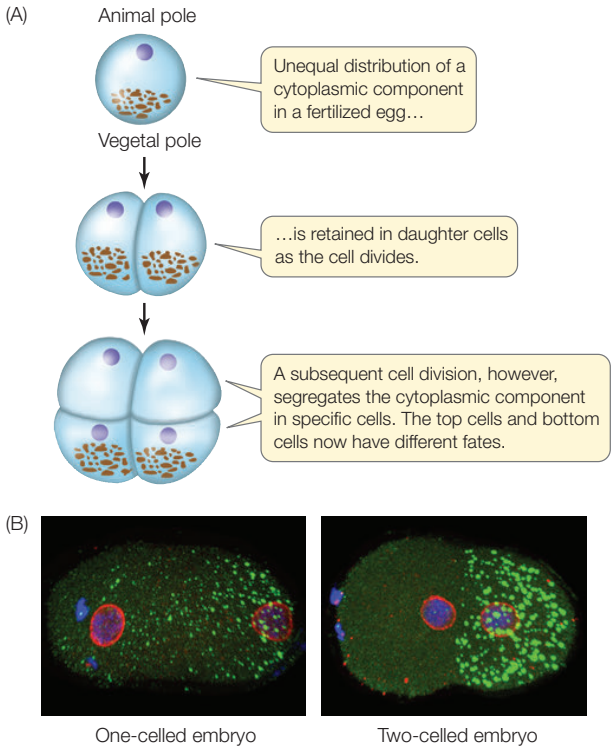


FIGURE 14.7 The Concept of Cytoplasmic Segregation (A) The unequal distribution of cytoplasmic determinants in a fertilized egg determines the fates of its descendants. (B) The zygote of the nematode worm *Caenorhabditis elegans* (left) shows an asymmetrical distribution of cytoplasmic particles (stained green). The progeny of the first cell division (right) receive unequal amounts of the particles.

Go to ANIMATED TUTORIAL 14.3
Early Asymmetry in the Embryo
PoL2e.com/at14.3

It turns out that the cytoskeleton contributes to the asymmetrical distribution of these determinants in the egg. Recall from Concept 4.4 that an important function of the microtubules and microfilaments in the cytoskeleton is to help move materials in the cell. Two properties allow these structures to accomplish this:

- Microtubules and microfilaments have polarity—they grow by adding subunits to the plus end.
- Cytoskeletal elements can bind specific proteins, which can be used in the transport of mRNA.

For example, in the sea urchin egg, a protein binds to both the growing (+) end of a microfilament and to an mRNA encoding a cytoplasmic determinant. As the microfilament grows toward one end of the cell, it carries the mRNA along with it. The asymmetrical distribution of the mRNA leads to asymmetrical distribution of the protein it encodes—a transcription factor.

Inducers passing from one cell to another can determine cell fates

The term "induction" has different meanings in different contexts. In biology it can be used broadly to refer to the initiation of, or cause of, a change or process. But in the context of cellular differentiation, it refers to the signaling events by which

cells in a developing embryo communicate and influence one another's developmental fate. Induction involves chemical signals called **inducers** and the signal transduction pathways that are triggered by these signals. Exposure to different amounts of inductive signals can lead to differences in gene expression among cells in a developing organism.

LINK

Signal transduction pathways are described in **Concepts 5.5 and 5.6**

The nematode worm *Caenorhabditis elegans* was one of the first model eukaryotic organisms to have its entire genome sequenced (see Concept 12.3). This worm develops from fertilized egg to larva in only about 8 hours and reaches the adult stage in just 3.5 days. The process is easily observed using a low-magnification dissecting microscope because the body covering is transparent (**FIGURE 14.8A**). To illustrate the principles of induction, we focus here on the development of one part of the *C. elegans* body: the vulva (**FIGURE 14.8B**).

The adult nematode is hermaphroditic, containing both male and female reproductive organs. It lays eggs through a pore called the vulva on its ventral (lower) surface. During development, a single cell, called the anchor cell, induces the vulva to form from six cells on the worm's ventral surface. In this case there are two molecular signals, the primary inducer and the secondary inducer. Each of the six ventral cells has three possible fates: it may become a primary vulval precursor cell, a secondary vulval precursor cell, or simply become part of the worm's skin—an epidermal cell. You can follow the sequence of events in Figure 14.8B. The concentration of the primary inducer, LIN-3, is key: the anchor cell produces LIN-3, which diffuses out of the cell and forms a concentration gradient with respect to adjacent cells. The three cells that are closest to the anchor cell receive the most LIN-3 and become vulval precursor cells; cells slightly farther from the anchor cell receive less LIN-3 and become epidermal cells. A second induction event results in the two classes of vulval precursor cells: primary and secondary. Induction involves the activation or inactivation of specific sets of genes through

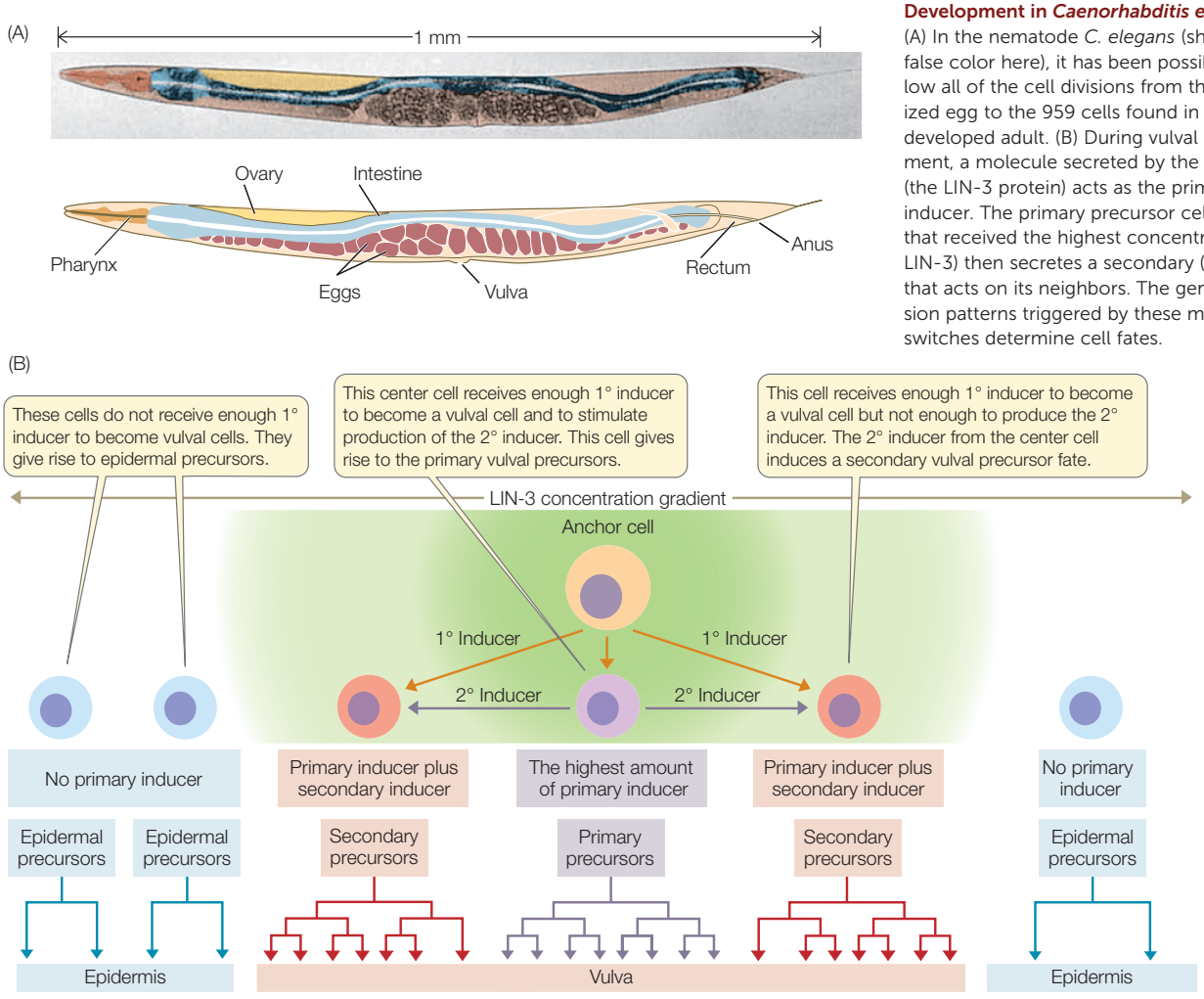


FIGURE 14.8 Induction during Vulval Development in *Caenorhabditis elegans*
(A) In the nematode *C. elegans* (shown in false color here), it has been possible to follow all of the cell divisions from the fertilized egg to the 959 cells found in the fully developed adult. (B) During vulval development, a molecule secreted by the anchor cell (the LIN-3 protein) acts as the primary (1°) inducer. The primary precursor cell (the one that received the highest concentration of LIN-3) then secretes a secondary (2°) inducer that acts on its neighbors. The gene expression patterns triggered by these molecular switches determine cell fates.

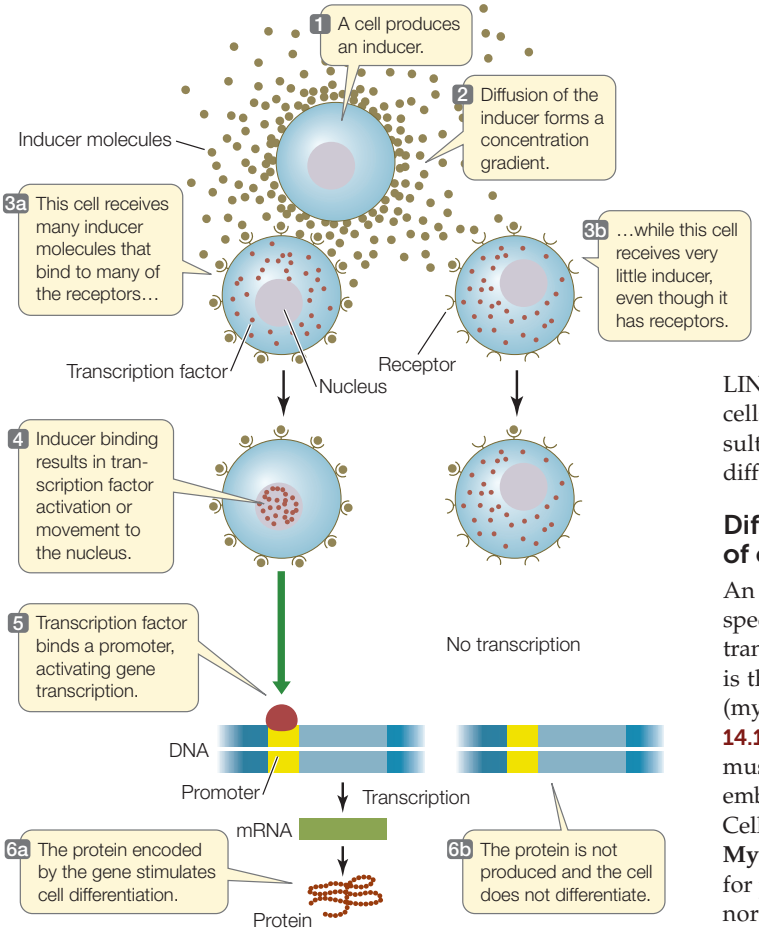


FIGURE 14.9 The Concept of Embryonic Induction The concentration of an inducer directly affects the degree to which a transcription factor is activated. The inducer acts by binding to a receptor on the target cell. This binding is followed by signal transduction involving transcription factor activation or movement from the cytoplasm to the nucleus. In the nucleus, the transcription factor acts to stimulate the expression of genes involved in cell differentiation.

signal transduction cascades in the responding cells (**FIGURE 14.9**).

This example from nematode development illustrates an important observation: *much of development is controlled by molecular switches that allow a cell to proceed down one of two alternative paths*. One challenge for developmental biologists is to find these switches and determine how they work. The primary inducer, LIN-3, released by the *C. elegans* anchor cell, is a growth factor similar in gene and protein sequence to a vertebrate growth factor called EGF (*e*pidermal *g*rowth *f*actor). LIN-3 binds to a receptor on the surfaces of vulval precursor cells, setting in motion a signal transduction cascade that results in increased transcription of the genes involved in the differentiation of vulval cells.

Differential gene transcription is a hallmark of cell differentiation

An important mechanism by which cells differentiate into specific cell types, with specific functions, is differential gene transcription. One well-studied example of cell differentiation is the conversion of undifferentiated muscle precursor cells (myoblasts) into the cells that make up muscle fibers (**FIGURE 14.10**). A key event in the commitment of these cells to become muscle is that they stop dividing. Indeed, in many parts of the embryo, *cell division and cell differentiation are mutually exclusive*. Cell signaling activates the gene for a transcription factor called **MyoD** (*myo*blast-*d*etermining gene). MyoD activates the gene for p21, an inhibitor of cyclin-dependent kinases (CDKs) that normally stimulate the cell cycle at G1 (see Figure 7.10). Expression of the *p21* gene causes the cell cycle to stop, and other transcription factors then enter the picture so that myoblasts can differentiate into muscle cells.

CHECKpoint CONCEPT 14.2

- ✓ What would be the effect on the embryo of injecting an inhibitor of microtubule polymerization into a fertilized sea urchin egg?
- ✓ Compare the internal and external stimuli that lead to differential gene expression in embryonic cells.
- ✓ What would be the consequences of a homozygous deletion mutation of *lin-3*, the gene that encodes LIN-3?

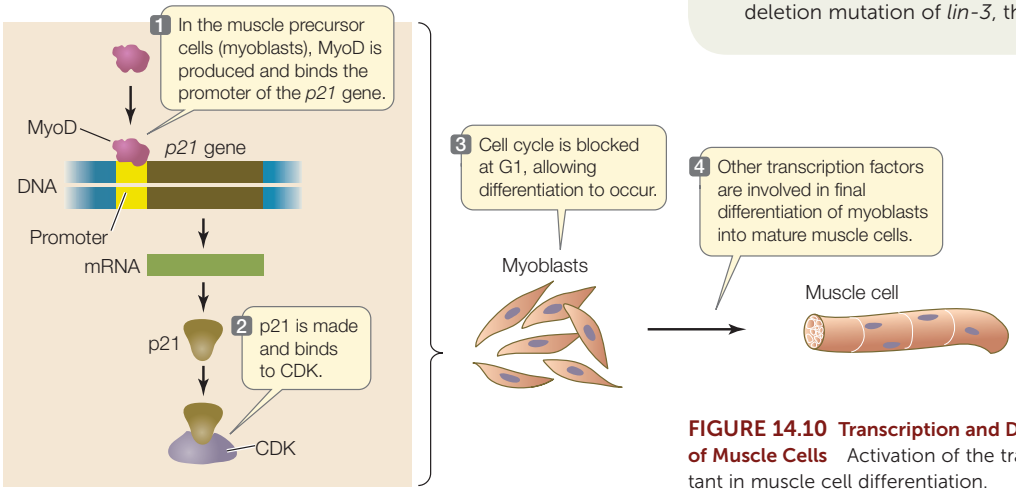


FIGURE 14.10 Transcription and Differentiation in the Formation of Muscle Cells Activation of the transcription factor MyoD is important in muscle cell differentiation.

Cytoplasmic polarity and inducers affect the expression of genes that determine cell fates. We will now look in more detail at how spatial differences in gene expression affect cell fate determination and the formation of tissues and organs.

CONCEPT 14.3 Spatial Differences in Gene Expression Lead to Morphogenesis

Pattern formation is the developmental process that results in the spatial organization of a tissue or organism. An example is the development of the vulva in *C. elegans* (see Figure 14.8). Pattern formation is inextricably linked to morphogenesis, the development of body form. Underlying both of these processes are spatial differences in gene expression, which determine whether, for example, a particular piece of tissue will become a leg or a wing or a flower petal. These instances where different genes are expressed in different places in the developing organism, in turn, depend on two cellular processes:

- The cells in the tissue must "know" where they are in relation to rest of the body.
- The cells must activate the pattern of gene expression that is appropriate for their location.

In the sections that follow, we will explore the mechanisms used by various organisms to direct pattern formation and morphogenesis.

Morphogen gradients provide positional information during development

During development, the key cellular question "What am I (or what will I be)?" is often answered in part by "Where am I?" Think of the cells in the developing nematode, which develop into different parts of the vulva depending on their positions relative to the anchor cell. The same is true for the cells between the digits of a developing hand and in different whorls of a developing flower. This spatial "sense" is called **positional information**.

Positional information often comes in the form of an inducer called a **morphogen**, which diffuses from one group of cells to surrounding cells, setting up a concentration gradient. There are two requirements for a signal to be considered a morphogen:

- It must specifically affect target cells.
- Different concentrations of the signal must cause different effects.

Developmental biologist Lewis Wolpert uses the "French flag model" to explain the action of morphogens (**FIGURE 14.11A**). This model can be applied to the differentiation of the vulva in *C. elegans* (see Figure 14.8), which relies on a gradient of LIN-3. Another example can be seen in the development of vertebrate limbs.

The vertebrate limb develops from a paddle-shaped limb bud (**FIGURE 14.11B**). The cells that develop into different digits must receive positional information; if they do not, the limb will not be organized properly (imagine a hand with only thumbs or only little fingers). How do the cells know where they are? A group of cells at the posterior base of the limb bud, just where it joins the body wall, is called the zone of polarizing activity (ZPA). The cells of the ZPA secrete a morphogen called *S*onic *h*edge*h*og (Shh), which forms a gradient that determines the posterior–anterior (little finger to thumb) axis of the developing limb. The cells getting the highest dose of Shh form the little finger; those getting the lowest dose develop into the thumb. Recall the French flag model when considering the gradient of Shh.

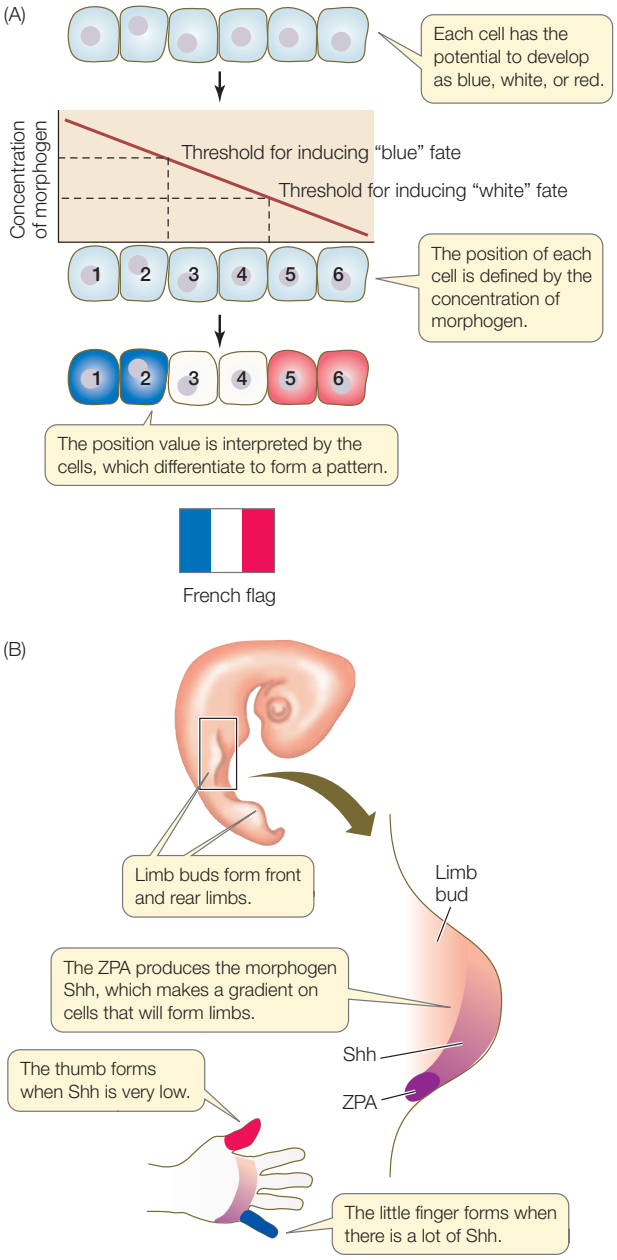
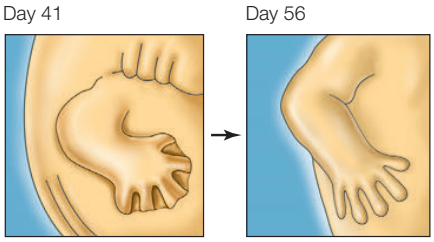


FIGURE 14.11 The French Flag Model (A) In the "French flag model," a concentration gradient of a diffusible morphogen signals each cell to specify its position. (B) The zone of polarizing activity (ZPA) in the limb bud of the embryo secretes the morphogen Sonic hedgehog (Shh). Cells in the bud form different digits depending on the concentration of Shh.

Multiple proteins interact to determine developmental programmed cell death

You might expect morphogenesis to involve a lot of cell division, followed by differentiation—and it does. But what you might not expect is the amount of programmed cell death—apoptosis—that occurs during morphogenesis. For example, in an early human embryo, the hands and feet look like tiny paddles: the tissues that will become fingers and toes are linked by connective tissue. Between days 41 and 56 of development, the cells between the digits die, freeing the individual fingers and toes:

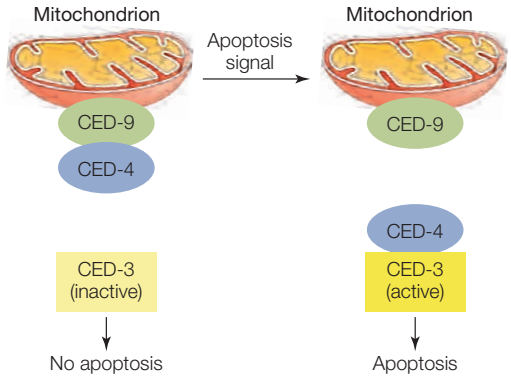


Many cells and structures form and then disappear during development, in processes involving apoptosis.

LINK

Concept 7.5 describes some of the cellular events of apoptosis

Model organisms have been very useful in studying the genes and proteins involved in apoptosis. For example, the nematode worm *C. elegans* produces precisely 1,090 somatic cells as it develops from a fertilized egg into an adult, but 131 of those cells die (leaving 959 cells in the adult worm). The sequential activation of two proteins called CED-4 and CED-3 (for *ce*ll *d*eath) is essential to this programmed cell death. A third protein called CED-9, which is bound to the outside of the mitochondria, inhibits apoptosis in cells that are not programmed to die. In these cells, CED-9 binds CED-4 and prevents it from activating CED-3. If the cell receives a signal for apoptosis, CED-9 releases CED-4, which then activates CED-3:



CED-3 is a caspase (a protease involved in apoptosis) that turns out to be similar to the caspase protein in humans. Several other proteins involved in the nematode apoptosis pathway (including CED-4 and CED-9) also have relatives in humans. So humans and *C. elegans*, two species separated by more than 600 million years of evolutionary history, have similar genes (encoding similar proteins) that control programmed cell death. The commonality of this pathway indicates its importance: most mutations in the genes that control this pathway are harmful and evolution selects against them. We will return to other examples of links between evolution and development in Concept 14.4.

Our example of apoptosis in the development of fingers and toes shows one of the many ways that the behavior of cells can give rise to body form during development. It also illustrates the two cellular processes underlying pattern formation: only cells in a particular place (between the digits) activate a specific pattern of gene expression (to trigger apoptosis).

Expression of transcription factor genes determines organ placement in plants

Like animals, plants have organs—for example, leaves and roots. Many plants form flowers, and many flowers are composed of four types of organs: sepals, petals, stamens (male reproductive organs), and carpels (female reproductive organs). These floral organs occur in concentric whorls (rings), with groups of each organ type encircling a central axis. The sepals are on the outside and the carpels are on the inside (**FIGURE 14.12A**).

In the model plant *Arabidopsis thaliana* (thale cress), flowers develop in a radial pattern around the top of the stem as it develops and elongates. At the shoot apex and in other parts of the plant where growth and differentiation occur (such as the root tip), there are groups of undifferentiated, rapidly dividing cells called **meristems** (see Concept 24.1). Each flower begins as a floral meristem of about 700 undifferentiated cells arranged in a dome, and the four whorls develop from this meristem. How is the identity of a particular whorl determined? Three classes of genes called **organ identity genes** encode proteins that act in combination to produce specific whorl features (**FIGURE 14.12B,C**):

- Genes in class A are expressed in whorls 1 and 2 (which form sepals and petals, respectively).
- Genes in class B are expressed in whorls 2 and 3 (which form petals and stamens).
- Genes in class C are expressed in whorls 3 and 4 (which form stamens and carpels).

APPLY THE CONCEPT

Gene expression and morphogenesis

Molecular biologists can attach genes to active promoters and insert them into cells. This results in higher than normal expression (overexpression) of the genes. What do you think would happen in each case if the four genes listed were to be overexpressed in the specified tissues? Explain your answers.

1. *ced-3* in embryonic neuron precursors of *C. elegans*
2. *myoD* in undifferentiated myoblasts
3. *Sonic hedgehog* in a chick limb bud
4. *LEAFY* in a leaf bud meristem of *Arabidopsis*

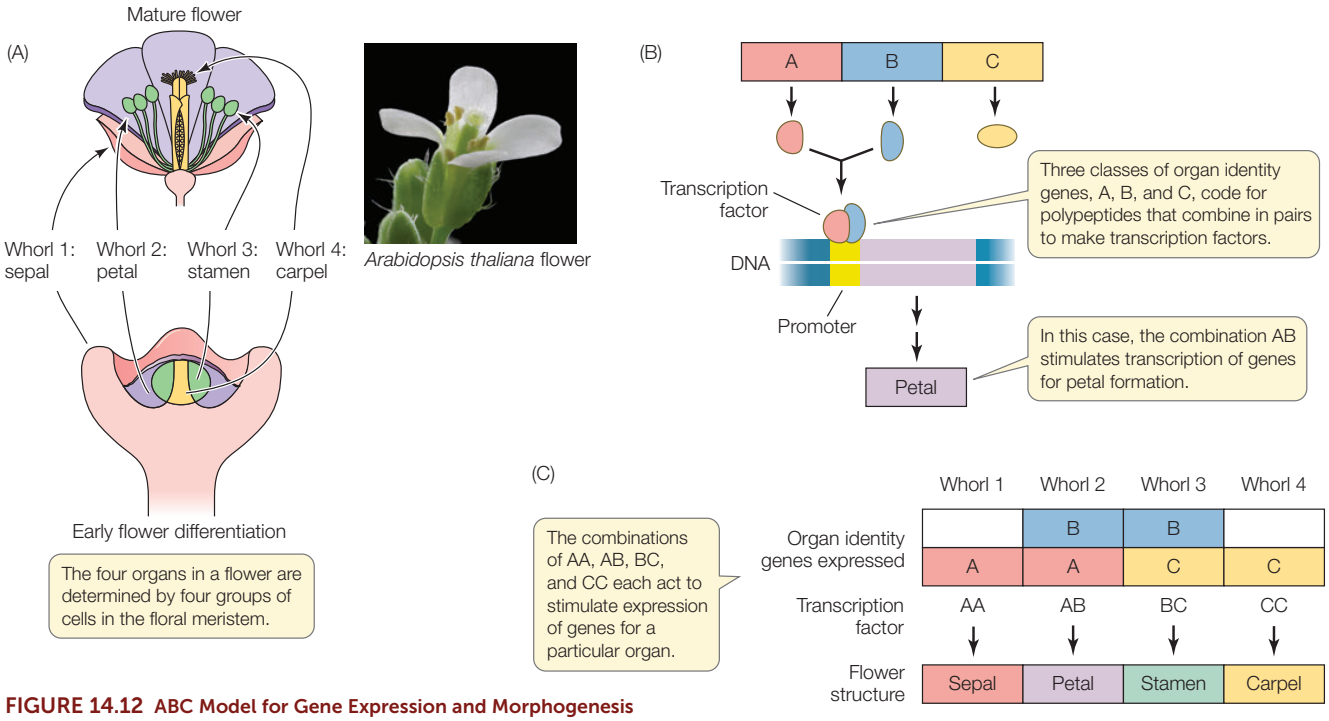


FIGURE 14.12 ABC Model for Gene Expression and Morphogenesis in *Arabidopsis thaliana* Flowers (A) The four organs of a flower—sepals (pink), petals (purple), stamens (green), and carpels (yellow)—grow in whorls that develop from the floral meristem. (B) Floral organs are determined by three classes of genes whose polypeptide products combine in pairs to form transcription factors. (C) Combinations of polypeptide subunits in transcription factors activate gene expression for specific organs.

These genes encode transcription factors that are active as dimers, that is, proteins with two polypeptide subunits. The composition of the dimer determines which genes the transcription factor activates. For example, a dimer made up of two class A monomers activates transcription of the genes that make sepals; a dimer made up of a class A monomer and a class B monomer results in petals, and so forth.

Two lines of experimental evidence support this model for floral organ determination:

- *Loss-of-function mutations*: for example, a mutation in a class A gene results in no sepals or petals.
- *Gain-of-function mutations*: for example, a promoter for a class C gene can be artificially coupled to a class A gene. In this case, the class A gene is expressed in all four whorls, resulting in only sepals and petals. In any organism, the replacement of one organ by another is called homeosis, and this type of mutation is a **homeotic mutation**.

Transcription of the floral organ identity genes is controlled by other gene products, including the LEAFY protein. Plants with loss-of-function mutations in the *LEAFY* gene make stems instead of flowers, with increased numbers of modified leaves called bracts. The wild-type LEAFY protein is a transcription factor that stimulates expression of the class A, B, and C genes so that they produce flowers. This finding has practical applications. It usually takes 6–20 years for a citrus tree to produce flowers and fruits. Scientists have made transgenic orange trees expressing the *LEAFY* gene coupled to a strongly expressed promoter. These trees flower and fruit years earlier than normal trees.

A cascade of transcription factors establishes body segmentation in the fruit fly

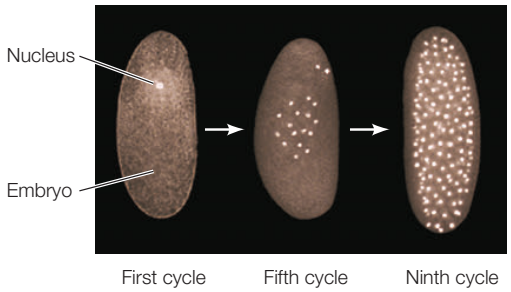
A major achievement in studies of developmental biology has been the ever-advancing description of how morphogens act in another model organism, the fruit fly *Drosophila melanogaster*. As you will see in Concept 14.4, the molecular events that underlie fruit fly development turn out to be similar to events that occur in many other organisms, including ourselves. So they merit examination in some detail.

The insect body is made up of segments that differ from one another. The adult fly has an anterior head (composed of several fused segments), three different thoracic segments, and eight abdominal segments at the posterior end. Each segment gives rise to different body parts: for example, antennae and eyes develop from head segments, wings from the thorax, and so on.

The life cycle of *Drosophila* from fertilized egg to adult takes about 2 weeks. The egg hatches into a larva, which then forms a pupa, which finally is transformed into the adult fly. By the time a larva appears—about 24 hours after fertilization—there are recognizable segments. The thoracic and abdominal segments all look similar, but *the fates of the cells to become different adult segments are already determined*.

As with other organisms, fertilization in *Drosophila* leads to a rapid series of mitoses. However, the first 12 nuclear divisions are not accompanied by cytokinesis. So a *multinucleate* embryo

forms instead of a *multicellular* embryo. The nuclei are brightly stained in the micrographs below:



With no cell membranes to cross, morphogens can diffuse easily within the embryo. As you will see, many of these morphogens affect transcription in the cell nuclei. We focus here on cell fate determination events that occur in the first 24 hours, which were elucidated in *Drosophila* using genetics:

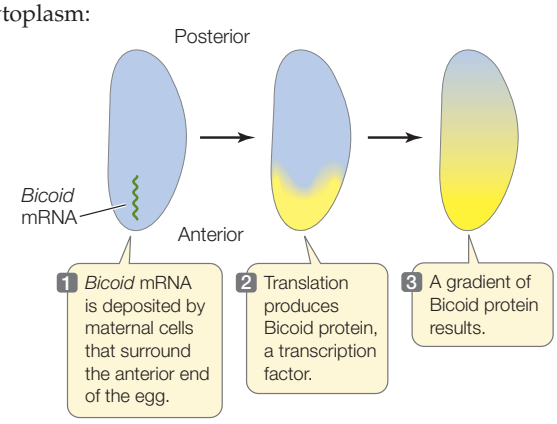
- First, developmental mutations were identified. For example, a mutant strain might produce larvae with two heads or missing certain segments.
- Second, each mutant was compared with wild-type flies, and the gene responsible for the developmental mistake, and its protein product (if appropriate), was isolated.
- Finally, experiments with the gene (making transgenic flies) and protein (injecting the protein into an egg or embryo) were done to confirm the proposed developmental pathway.

Together, these approaches revealed a sequential pattern (cascade) of gene expression that results in the determination of each segment within 24 hours after fertilization. Several classes of genes are involved:

- Maternal effect genes, which set up the major axes (anterior–posterior and dorsal–ventral) of the egg
- Segmentation genes, which determine the boundaries and polarity of each of the segments
- Hox genes, which determine what organ will be made at a given location

MATERNAL EFFECT GENES Like the eggs and early embryos of many other organisms (see Figure 14.7), *Drosophila* eggs and larvae are characterized by unevenly distributed cytoplasmic determinants. These molecular determinants, which include both mRNAs and proteins, are the products of specific **maternal effect genes**. These genes are transcribed in the cells of the mother's ovary, and the mRNAs are passed to the egg via cytoplasmic bridges. Two maternal effect genes, called *bicoid* and *nanos*, help determine the anterior–posterior axis of the egg. (The dorsal–ventral, or back–belly, axis is determined by other maternal effect genes that we will not describe here.)

The mRNAs for *bicoid* and *nanos* diffuse from the mother's cells into what will be the anterior (head) end of the egg. After fertilization, the *bicoid* mRNA is translated to produce Bicoid protein, a transcription factor that diffuses away from the anterior end, establishing a concentration gradient in the egg cytoplasm:



At this point, the egg is in its multinucleate stage.

Where it is present in sufficient concentration, Bicoid stimulates the transcription of the *hunchback* gene in the early embryo. Consequently, the nuclei nearest the anterior end are most active in the transcription of the *hunchback* gene, and the resulting gradient of Hunchback protein (itself a transcription factor) establishes the head, or anterior, region.

Meanwhile, the egg's cytoskeleton transports the *nanos* mRNA from the anterior end of the egg to the posterior (tail) end, where it is translated after fertilization. This results in a gradient of the Nanos protein, with the highest concentration at the posterior end. At that end, the Nanos protein inhibits the translation of *hunchback* mRNA, preventing accumulation of Hunchback protein. Thus the actions of both Bicoid and Nanos establish a Hunchback protein gradient, which determines the anterior and posterior ends of the embryo by influencing the gene expression patterns of the nuclei along the gradient.

The events involving *bicoid*, *nanos*, and *hunchback* begin before fertilization and continue after it, during the multinucleate stage, which lasts a few hours. At this stage the embryo looks like a bunch of indistinguishable nuclei under the light microscope. But the fates of the individual nuclei and the cells they will occupy have already begun to be determined. After the anterior and posterior ends have been established, the next step in pattern formation in fruit flies is the determination of segment number and locations.

SEGMENTATION GENES The number, boundaries, and polarity of the *Drosophila* larval segments are determined by proteins encoded by the **segmentation genes**. These genes are expressed when there are about 6,000 nuclei in the embryo (about 3 hours after fertilization). Three classes of segmentation genes act one after the other to regulate finer and finer details of the segmentation pattern (**FIGURE 14.13**):

- **Gap genes** organize broad areas along the anterior–posterior axis. Mutations in gap genes result in gaps in the body plan—the omission of several consecutive larval segments.
- **Pair rule genes** divide the embryo into units of two segments each. Mutations in pair rule genes result in embryos missing every other segment.

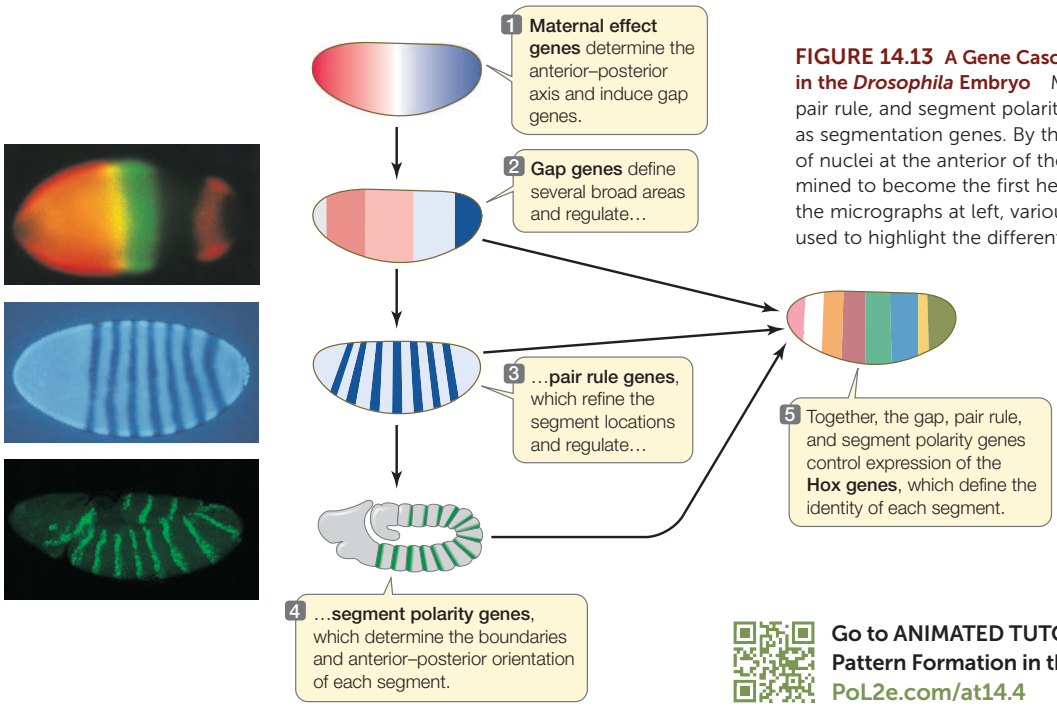


FIGURE 14.13 A Gene Cascade Controls Pattern Formation in the *Drosophila* Embryo Maternal effect genes induce gap, pair rule, and segment polarity genes—collectively referred to as segmentation genes. By the end of this cascade, a group of nuclei at the anterior of the embryo, for example, is determined to become the first head segment in the adult fly. In the micrographs at left, various staining methods have been used to highlight the different gene products.

Go to ANIMATED TUTORIAL 14.4
Pattern Formation in the *Drosophila* Embryo
PoL2e.com/at14.4

- **Segment polarity genes** determine the boundaries and anterior–posterior organization of the individual segments. Mutations in segment polarity genes can result in segments in which posterior structures are replaced by reversed (mirror-image) anterior structures.

By the end of this part of the cascade, nuclei throughout the embryo "know" which segment they will be part of in the adult fly. The next set of genes in the cascade determines the form and function of each segment.

Go to MEDIA CLIP 14.1
Spectacular Fly Development in 3D
PoL2e.com/mc14.1

HOX GENES **Hox genes** encode a family of transcription factors that are expressed in different combinations along the length of the embryo, and help determine cell fates within each segment. Expression of certain Hox genes leads to the development of antennae in the head segment, whereas other Hox genes are expressed in the thorax to make wings, and so on. Hox genes are homeotic genes that are shared by most animals, and they are functionally similar to the organ identity genes of plants (see Figure 14.12).

How do we know that the Hox genes determine segment identity? A clue comes from homeotic mutations in *Drosophila*. A gain-of-function mutation in the Hox gene *Antennapedia* causes legs to grow on the head in place of antennae (**FIGURE 14.14**). When another Hox gene, *bithorax*, is mutated, an extra pair of wings grows in a thoracic segment where wings do not normally occur. So the normal (wild-type) functions of the Hox genes must be to "tell" a segment what organ to form. Hox genes encode transcription factors and have a conserved 180-base-pair sequence called the **homeobox** (from which the

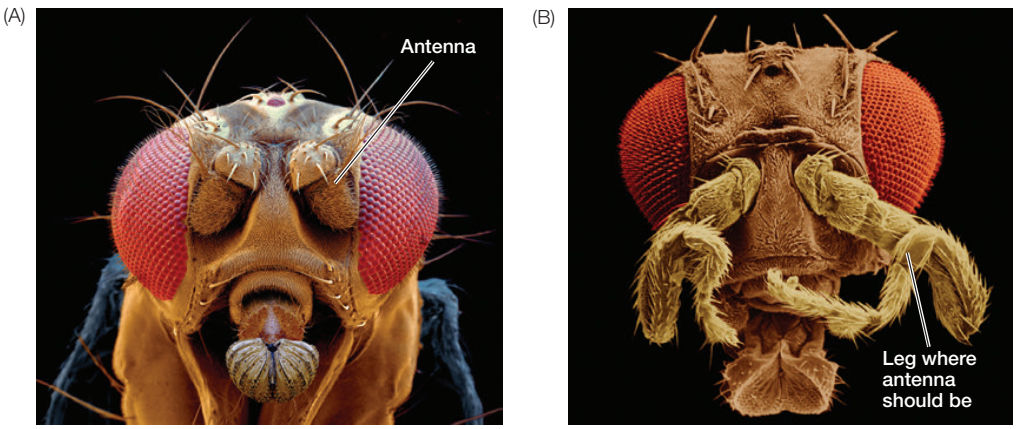


FIGURE 14.14 A Homeotic Mutation in *Drosophila* Mutations of the Hox genes cause body parts to form on inappropriate segments. (A) A wild-type fruit fly. (B) An *Antennapedia* mutant fruit fly. Mutations such as this reveal the normal role of the *Antennapedia* gene in determining segment function.

genes get their name). The homeobox encodes a 60-amino acid sequence called the homeodomain. The homeodomain recognizes and binds to a specific DNA sequence in the promoters of its target genes. As you will see in Concept 14.4, this domain is found in transcription factors that regulate development in many other animals with an anterior–posterior axis.

LINK

To review transcriptional regulation, see **Concept 11.2**

CHECKpoint CONCEPT 14.3

- ✓ Outline the steps that determine that a nucleus and cell in the developing *Drosophila* embryo will be part of an antenna.
- ✓ Compare the determination of organ identity in *Arabidopsis* and *Drosophila*.
- ✓ How does the "French flag model" apply to development in *Drosophila*?
- ✓ In the nematode nervous system, 302 neurons come from 405 precursors. How would you investigate the fate of the 103 "missing" cells? What gene(s) might be involved?

We have seen how positional information leads to changes in the expression of key developmental genes, which in turn control morphogenesis. It turns out that there are remarkable similarities in the genes used to guide development in diverse organisms, and this has led to a new way to look at the evolution of development.

CONCEPT 14.4 Changes in Gene Expression Pathways Underlie the Evolution of Development

The discovery of the genes that control the development of *Drosophila* provided biologists with tools to investigate the development of other organisms. For example, when scientists used homeobox DNA as a hybridization probe (see Figure 10.7) to search for similar genes elsewhere, they found that the homeobox is present in many genes in many other organisms. This, and other astounding discoveries that followed, showed a similarity in the molecular events underlying morphogenesis in organisms ranging from flies to fish to mammals. These results suggested that just as the forms of organisms evolved through descent with modification from a common ancestor, so did the molecular mechanisms that produce those forms. Biologists started to ask new questions about the interplay between evolutionary and developmental processes—a field of study called **evolutionary developmental biology** (**evo-devo**). The major findings of evo-devo are:

- Organisms share similar molecular mechanisms for development, including a "toolkit" of regulatory molecules that control the expression of genes.
- These regulatory molecules are able to act independently in different tissues and regions of the body, so that evolutionary change can occur in independent "modules."
- Developmental differences can arise from changes in the timing of action of a regulatory molecule, the location of its action, or the quantity of its action.
- Developmental changes can arise from environmental influences on developmental processes.

The development of a multicellular organism from a fertilized egg—a single cell—involves an intricate pattern of sequential gene expression. When developmental biologists began to describe the events responsible for the differentiation and controlled proliferation of cells and tissues at the molecular level, they found common regulatory genes and pathways in organisms that don't appear similar at all, such as fruit flies and mice.

Go to ACTIVITY 14.2 Plant and Animal Development
PoL2e.com/ac14.2

Developmental genes in distantly related organisms are similar

Initially through hybridization with a homeobox probe, and then by genome sequencing and comparative genomics (see Concept 12.2), biologists have found that diverse animals share numerous molecular pathways that govern gene expression during development. For example, fruit fly homeotic genes such as *Antennapedia* and *bithorax* are similar to mouse (and human) genes that play similar developmental roles. This means that the positional information controlled by these genes has been conserved, even as the structures formed at each position

APPLY THE CONCEPT

Changes in gene expression pathways underlie the evolution of development

The control of eye formation during the development of many animals is under the control of a genetic switch involving a transcription factor. Partial DNA sequences for the control gene from two organisms are given below.

Mouse *Pax6* gene:
5′-GTA TCC AAC GGT TGT GTG AGT AAA ATT-3′

Fruit fly *eyeless* gene:
5′-GTA TCA AAT GGA TGT GTG AGC AAA ATT-3′

1. Calculate the percentage of identity (the percentage of bases that are identical) between the two DNA sequences.
2. Use the genetic code (see Figure 10.11) to determine the amino acid sequences (see Table 3.2) encoded by the two regions and calculate the percentage of identity between the two amino acid sequences.
3. The fruit fly and the mouse diverged from a common ancestor about 500 million years ago. Comment on your answers to 1 and 2 in terms of the evolution of developmental pathways.

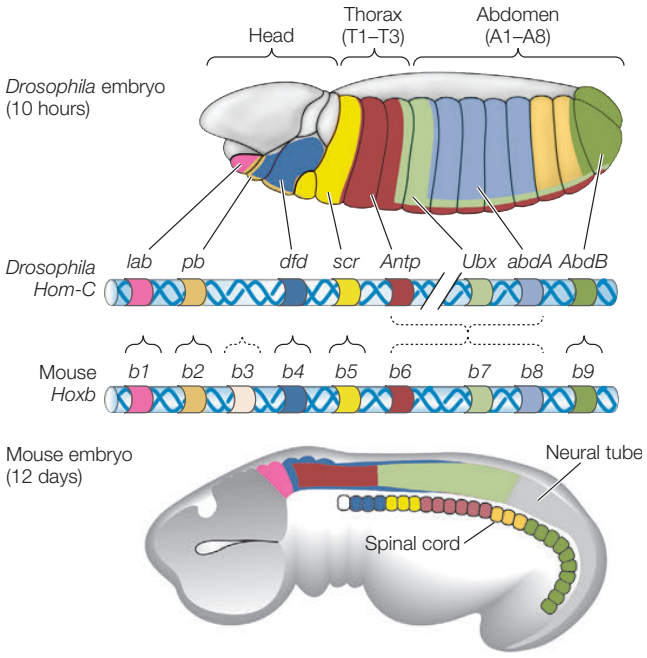


FIGURE 14.15 Regulatory Genes Show Similar Expression Patterns Similar genes encoding similar transcription factors are expressed in similar patterns along the anterior–posterior axis of both insects and vertebrates. Related genes and the locations of their expression are indicated by shared colors. The mouse (and human) Hox genes are actually present in multiple copies; this prevents a single mutation from having drastic effects.

have changed. Over the millions of years that have elapsed since these animals diverged from a common ancestor, the genes in question have mostly been maintained, suggesting that their functions are essential for animal development.

Remarkably, these genes are arranged along a chromosome in both the fruit fly and mouse *in the same order as they are expressed along the anterior–posterior axis of their embryos* (**FIGURE 14.15**). In the mouse and other vertebrates, these genes are switched on sequentially during embryogenesis. The Hox genes controlling the development of anterior structures (such as the head) are expressed earliest, and as a result, anterior structures develop earlier than posterior structures. The spatial organization of the Hox genes on the chromosome is important for the timing of expression of the genes.

These and other examples have led biologists to the idea that certain developmental mechanisms, controlled by specific DNA sequences, have been conserved over long periods during the evolution of multicellular organisms. These sequences comprise a **genetic toolkit**, the contents of which have been modified and reshuffled over the course of evolution to produce the remarkable diversity of plants, animals, and other organisms in the world today.

Genetic switches govern how the genetic toolkit is used

The genetic toolkit is also used to generate diverse structures within a single organism. Different structures can evolve within a single organism using a common set of genetic instructions because there are mechanisms called **genetic switches** (also called molecular switches) that control how the genetic toolkit is used. As we have seen, these mechanisms involve promoters and transcription factors. The signal cascades that converge on and operate these switches determine when and where genes will be turned on and off. Multiple switches control each gene by influencing its expression at different times and in different places. In this way, elements of the genetic toolkit can be involved in multiple developmental processes while still allowing individual modules to develop independently. For example, the morphogenesis of different flower organs is determined by different combinations of three classes of transcription factors (the ABC model; see Figure 14.12) acting at specific times and locations.

During evolution, changes in the functions of genetic switches have led to changes in the forms or functions of organisms. To illustrate this, let's look at the development of wings in *Drosophila* and other insects. *Drosophila* species are members of the insect group Diptera, which means "two wings"—that is, they have a single pair of wings, whereas most insects have two pairs of wings (i.e., four wings). In dipterans, the single pair of wings develops on the second thoracic segment, and a pair of balancing organs called halteres develops on the third thoracic segment. A critical difference between thoracic segments 2 and 3 is that the Hox gene *Ultrabithorax* (*Ubx*) is expressed in segment 3 but not in segment 2. The Ubx transcription factor represses the development of wings in dipterans (**FIGURE 14.16**). If the *Drosophila Ubx* gene is inactivated by mutation,

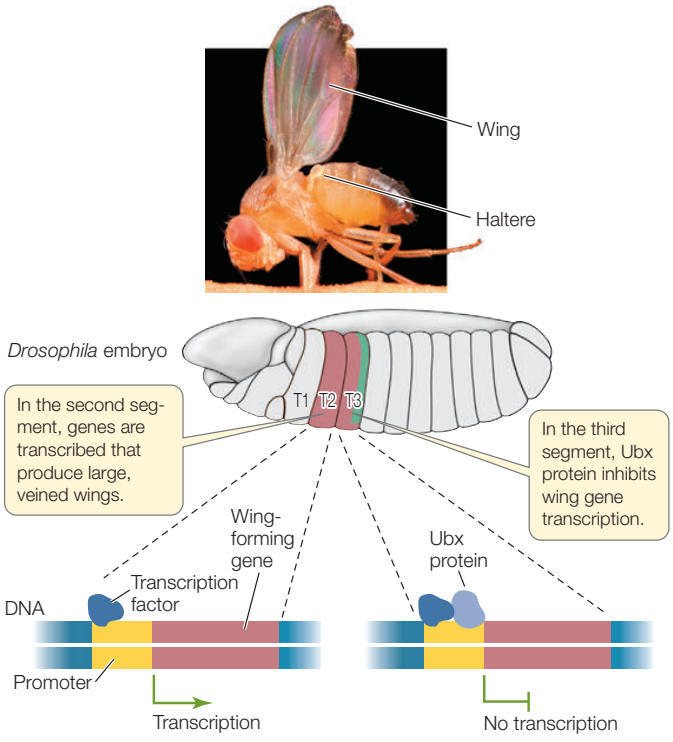


FIGURE 14.16 Segments Differentiate under Control of Genetic Switches The binding of a single protein, Ultrabithorax (Ubx), determines whether a thoracic segment in *Drosophila* produces full wings or halteres (balancers).

a second pair of wings forms in thoracic segment 3. In other insects such as butterflies, Ubx turns *on* the expression of wing-forming genes so that full hindwings develop. Therefore a simple genetic change in the effect of Ubx on genes that promote wing development results in a major morphological difference in the wings of flies and butterflies. This phenomenon—the same switch having different effects on target genes in different species—is important in evolution.

Modularity allows for differences in the pattern of gene expression among organisms

The modularity of development means that the molecular pathways for developmental processes such as organ formation operate independently from one another. For example, an *Antennapedia* mutant grows a leg where an antenna should be (see Figure 14.14), but all of the mutant's other organs develop normally and in their proper places. On an evolutionary time scale, modularity means that the timing and position of a particular developmental process can change without disrupting the whole organism.

Go to ANIMATED TUTORIAL 14.5
Modularity
PoL2e.com/at14.5

TIMING DIFFERENCES The genes regulating the development of a module may be expressed at different developmental stages or for different durations in different species, a phenomenon called **heterochrony**. An example is the evolution of the giraffe's neck. As in virtually all mammals (with the exception of manatees and sloths), there are seven vertebrae in the neck of the giraffe. So the giraffe did not get a longer neck than other mammals by adding vertebrae. Instead, each of the cervical (neck) vertebrae of the giraffe is much longer than those of other mammals (FIGURE 14.17). How does this happen?

Bones grow because of the proliferation of cartilage-producing cells called chondrocytes. Bone growth is stopped by a signal that results in death of the chondrocytes and the accumulation of calcium salts in the bone. In giraffes this signaling process is delayed in the cervical vertebrae, with the result that these vertebrae grow longer. Thus the evolution of longer necks occurred through *changes in the timing of expression* of the genes that control bone formation.

SPATIAL DIFFERENCES Changes in the spatial expression pattern of a developmental gene are known as **heterotopy** and can also result in evolutionary change. For example, the difference in foot webbing in ducks versus chickens is determined by an alteration in the spatial expression of a single gene. The feet of all bird embryos have webs of skin that connect their toes. This webbing is retained in adult ducks (and other aquatic birds) but not in adult chickens (and other non-aquatic birds). The loss of webbing is caused by a signaling protein called bone morphogenetic protein 4 (BMP4) that instructs the cells in the webbing to undergo apoptosis. The death of these cells destroys the webbing between the toes.

Embryonic duck and chicken hindlimbs both express the *BMP4* gene in the webbing between the toes, but they differ in expression of a gene called *Gremlin*, which encodes a BMP *inhibitor* protein (FIGURE 14.18). In ducks, but not chickens, the *Gremlin* gene is expressed in the webbing cells. The Gremlin protein inhibits the BMP4 protein from signaling for apoptosis, and the result is a webbed foot.

FIGURE 14.17 Heterochrony in the Development of a Longer Neck There are seven vertebrae in the neck of the giraffe (left) and human (right; not to scale). But the vertebrae of the giraffe are much longer (25 cm compared with 1.5 cm) because during development, growth continues for a longer period of time. This timing difference is called heterochrony.

CHECKpoint CONCEPT 14.4

- ✓ Describe the major ideas of evolutionary developmental biology.
- ✓ What is the evidence that there was a common origin for the developmental pathways leading to segment identity in insects and the organization of the spinal cord in mice?
- ✓ What is the evidence that changes in the transcription of a single gene can lead to differences in morphogenesis between different regions of developing organ?
- ✓ Examine Figure 14.18 and the related text. If Gremlin protein were added to the webbed region between the developing toes of a chicken, what would be the result?

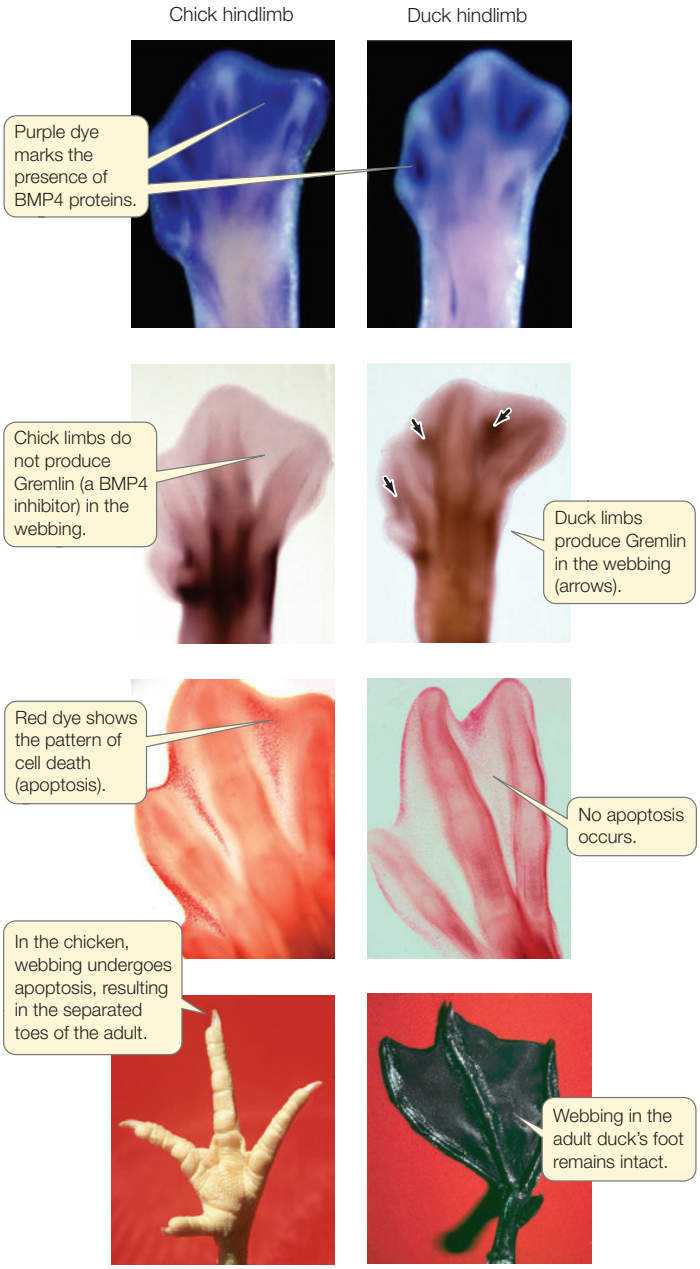


FIGURE 14.18 Changes in Gremlin Expression Correlate with Changes in Hindlimb Structure The left column of photos shows the development of a chicken's foot; the right column shows foot development in a duck. Gremlin protein in the webbing of the duck foot inhibits BMP4 signaling, thus preventing the embryonic webbing from undergoing apoptosis.

We have seen how the genetic toolkit guides morphogenesis in individual organisms, and how differences in genetic switches contribute to differences among species. In the next concept we will discuss further the roles that some of these same tools play in the evolution of new forms and new species.

CONCEPT 14.5 Developmental Genes Contribute to Species Evolution but Also Pose Constraints

The genetic switches that allow different structures to develop in different regions of an embryo can also give rise to major morphological differences among species. We have already seen examples of this: the difference in wing number between dipterans and other insects; the development of cervical vertebrae in giraffes versus other mammals; and the differences in Gremlin expression that determine whether a bird's foot will be webbed or not. Thus changes in the timing and position of a genetic switch can generate morphological variation, which then can be acted on by natural selection.

At the same time, the reliance of development on a genetic toolkit with a limited set of tools places constraints on how radically organisms can differ from one another. Four decades ago, the French geneticist François Jacob made the analogy that evolution works like a tinker, assembling new structures by *combining and modifying the available materials*, and not like an engineer, who is free to develop dramatically different designs (say, a jet engine to replace a propeller-driven engine). The evolution of morphology has not been governed by the appearance of radically new genes, but by modifications of existing genes and their regulatory pathways. Thus developmental genes and their expression constrain evolution in two major ways:

- Nearly all evolutionary innovations are modifications of previously existing structures.
- The genes that control development are highly conserved; that is, the regulatory genes usually change slowly over the course of evolution.

Mutations in developmental genes can cause major morphological changes

Sometimes a major developmental change is due to an alteration in the regulatory molecule itself rather than a change in where, when, or how much it is expressed. This is called **heterotypy** ("different type"). An excellent example of heterotypy is a gene that controls the number of legs in arthropods. Arthropods all have head, thoracic, and abdominal regions with variable numbers of segments. Insects such as *Drosophila* have three pairs of legs, one pair on each of their three thoracic segments, whereas centipedes have many legs on both thoracic and abdominal segments. All arthropods express a gene called *Distalless* (*Dll*) that controls segmental leg development. In insects, *Dll* expression is repressed in abdominal segments by the Hox gene *Ubx*. *Ubx* is expressed in the abdominal segments of all arthropods, but it has different effects in different species. In centipedes, Ubx protein *activates* expression of the *Dll* gene to promote the formation of legs. During the evolution of insects, a change in the *Ubx* gene

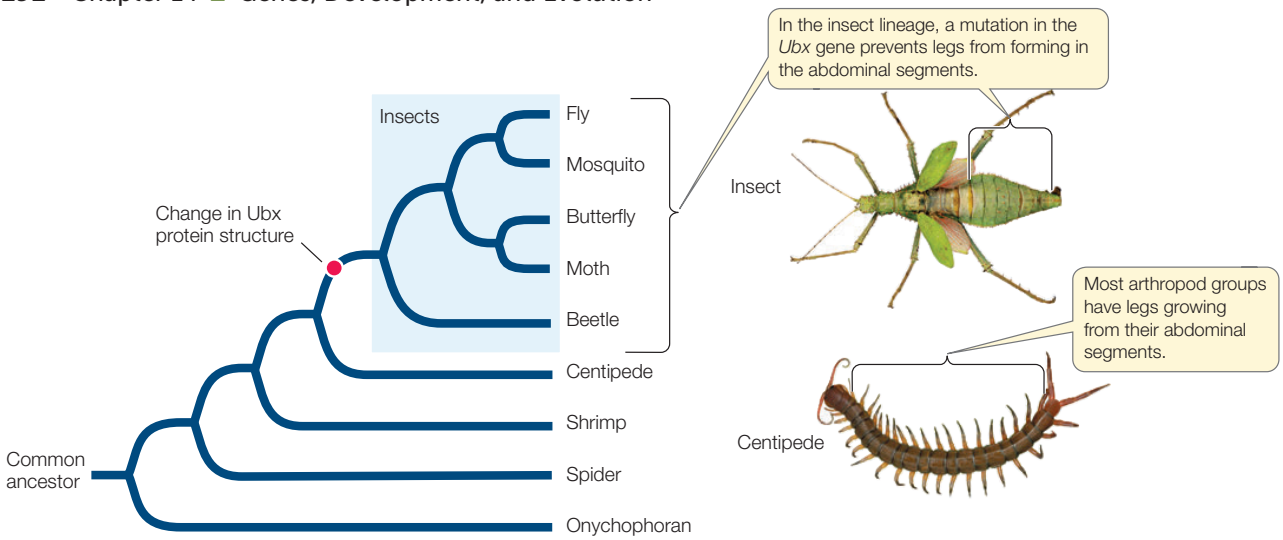


FIGURE 14.19 A Mutation in a Hox Gene Changed the Number of Legs in Insects In the insect lineage (blue box) of the arthropods, a change to the *Ubx* gene resulted in a protein that inhibits the *Dll* gene, which is required for legs to form. Because insects express this modified *Ubx* gene in their abdominal segments, no legs grow from these segments. Other arthropods, such as centipedes, do grow legs from their abdominal segments.

sequence resulted in a modified Ubx protein that *represses Dll* expression in abdominal segments. A phylogenetic tree of arthropods shows that this change in *Ubx* occurred in the ancestor of insects, at the same time that abdominal legs were lost (**FIGURE 14.19**).

LINK

Arthropod evolution and diversity are discussed in **Concept 23.4**

Evolution proceeds by changing what's already there

The features of organisms almost always evolve from preexisting features in their ancestors. New "wing genes" did not suddenly appear in birds and bats; instead, wings arose as modifications of existing structures (**FIGURE 14.20**). In vertebrates, the wings are modified limbs.

Although the wings of birds and bats look different, they are made from the same basic parts. Like limbs, wings have a common structure: a humerus that connects to the body; two longer bones, the radius and ulna, that project away from the humerus; and then metacarpals and phalanges (digits). During development these bones take on different lengths and

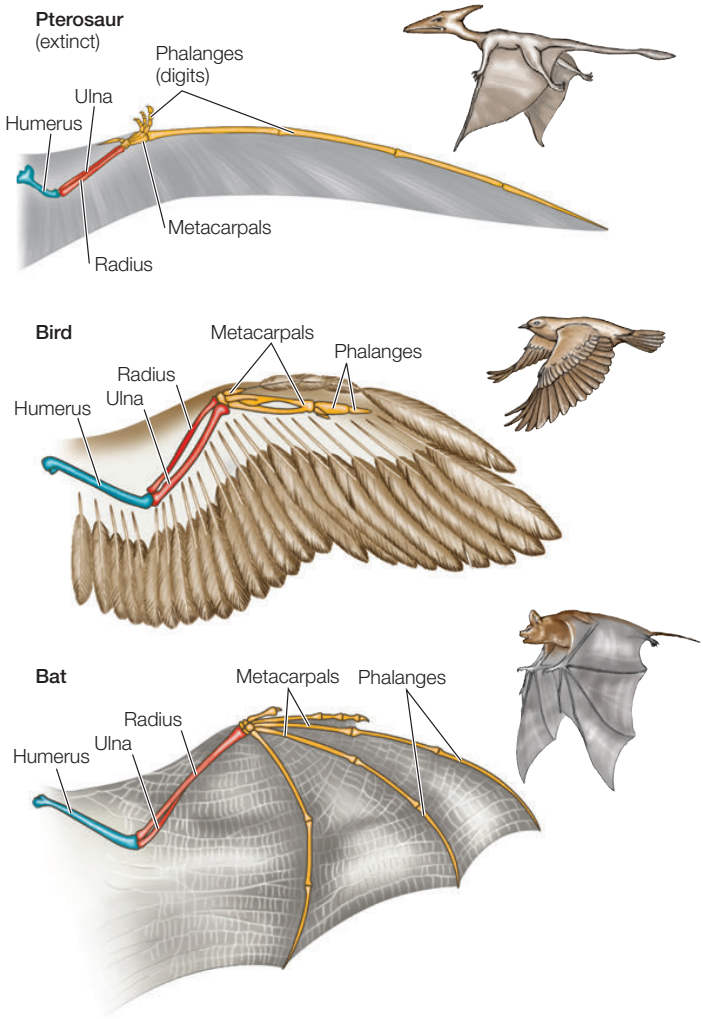


FIGURE 14.20 Wings Evolved Three Times in Vertebrates The wings of pterosaurs (the earliest flying vertebrates, which lived from 265 to 220 million years ago), birds, and bats are all modified forelimbs constructed from the same skeletal components. However, the components have different forms in the different groups of vertebrates.

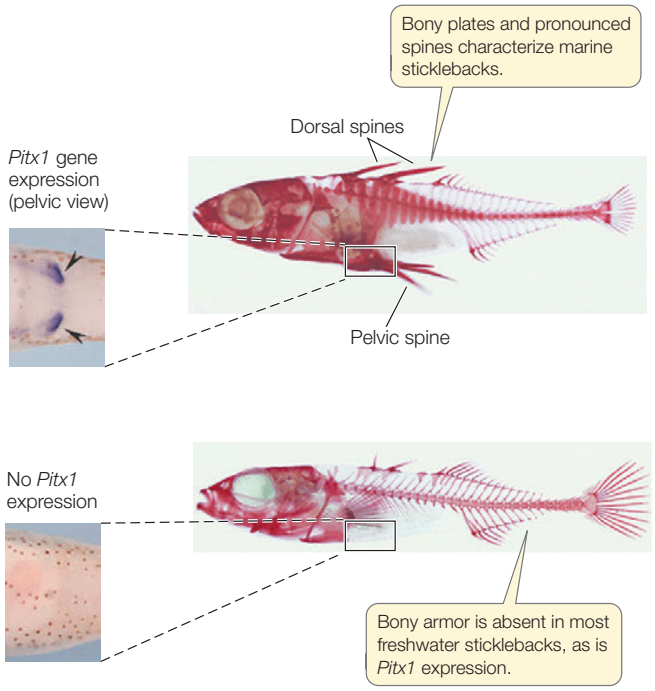


FIGURE 14.21 Parallel Phenotypic Evolution in Sticklebacks
A developmental gene, *Pitx1*, encodes a transcription factor that stimulates the production of plates and spines. This gene is active in marine sticklebacks (indicated by arrowheads in inset at left) but is mutated and inactive in various freshwater populations of the fish. The fact that this mutation is found in geographically distant and isolated freshwater populations is evidence for parallel evolution.

weights in different organisms. For example, the phalanges are relatively short in birds and relatively long in bats. These differences arise from changes in the molecular mechanisms that control development, as we saw for cervical vertebrae in giraffes.

Developmental controls also influence how organisms lose structures. The ancestors of present-day snakes lost their forelimbs as a result of changes in the segmental expression of Hox genes. The snake lineage subsequently lost its hindlimbs by the loss of expression of the *Sonic hedgehog* gene in the limb bud tissue. But some snake species such as boas and pythons still have rudimentary pelvic bones and upper leg bones. Recall that Sonic hedgehog also functions as a morphogen in hand development (see p. 283). This is yet another example of how the same basic genetic tools are used in different ways in different species.

Conserved developmental genes can lead to parallel evolution

As we saw for the Hox genes, the nucleotide sequences of many developmental genes have been highly conserved throughout the evolution of multicellular organisms. In other words, these genes exist in similar form across a broad spectrum of species. The existence of these highly conserved genes makes it likely that similar traits will evolve repeatedly, especially among closely related species, in a phenomenon called **parallel phenotypic evolution**. A good example is provided by a small fish, the three-spined stickleback (*Gasterosteus aculeatus*: "bony stomach with spines").

Sticklebacks are widely distributed across the Atlantic and Pacific oceans and are also found in many freshwater lakes. Marine populations of this species spend most of their lives at sea but return to fresh water to breed. Members of freshwater populations live in lakes and never journey to salt water. Genetic evidence shows that freshwater populations have arisen independently from marine populations many times in different parts of the world, most recently at the end of the last ice age. Marine sticklebacks have structures that protect them from predators: well-developed pelvic bones with pelvic spines, and bony plates. In each of the separate freshwater populations, this body armor is greatly reduced, and dorsal and pelvic spines are much shorter or even lacking (FIGURE 14.21).

The difference between marine and freshwater sticklebacks is not induced by environmental conditions. Marine species that are reared in fresh water still grow spines. Not surprisingly, the difference is due to a gene that affects development. The *Pituitary homeobox transcription factor 1* (*Pitx1*) gene codes for a transcription factor that is normally expressed in regions of the developing embryo that form the head, trunk, tail, and pelvis of the marine stickleback. However, in independent freshwater populations from Canada, the United Kingdom, the United States, and Iceland, the gene is no longer expressed in the pelvis, and the spines do not develop. *This same gene sequence has evolved to produce similar phenotypic changes in several independent populations*, and is thus a good example of parallel evolution. What could be the common selective mechanism in these cases? Possibly, the decreased predation pressure in the freshwater environment allows for increased reproductive success in animals that invest less energy in the development of unnecessary protective structures.

CHECKpoint CONCEPT 14.5

- ✓ How have diverse body forms such as wings evolved by means of modifications in the functioning of existing genes?
- ✓ What would happen at the molecular and phenotypic levels if a *Ubx* gene from an adult insect replaced the *Ubx* gene in a fertilized insect egg?
- ✓ When several freshwater populations of stickleback fish were compared, the coding region of the *Pitx1* gene was identical to that found in marine populations. But in every case, the freshwater fish had mutations in *noncoding* regions of *Pitx1* that led to reduced expression. What might these noncoding-region mutations be?

Q How do gene products control the development of the eye?

ANSWER After reading this chapter, it should not surprise you that the product of both the fruit fly *eyeless* gene and the vertebrate *Pax6* gene is a transcription factor. This protein is produced in the front of the developing brain in a region called the neural plate. The result is a region called the "eye field," where an eye can develop. The separation of this single region into two eyes occurs when cells in the middle of the region produce Shh (Sonic hedgehog—a protein that is also involved in limb specification). Shh is a transcription factor that blocks the synthesis of Pax6, so where there is Shh, there is no eye development. Typically Shh is produced only in a central region of the neural plate. But in cave-dwelling fish, it occurs over a wider area of the eye field and adults have no eyes (**FIGURE 14.22**). If too little or no Shh is made, a single eye is formed; this occurs in the human disorder known as cyclopia.

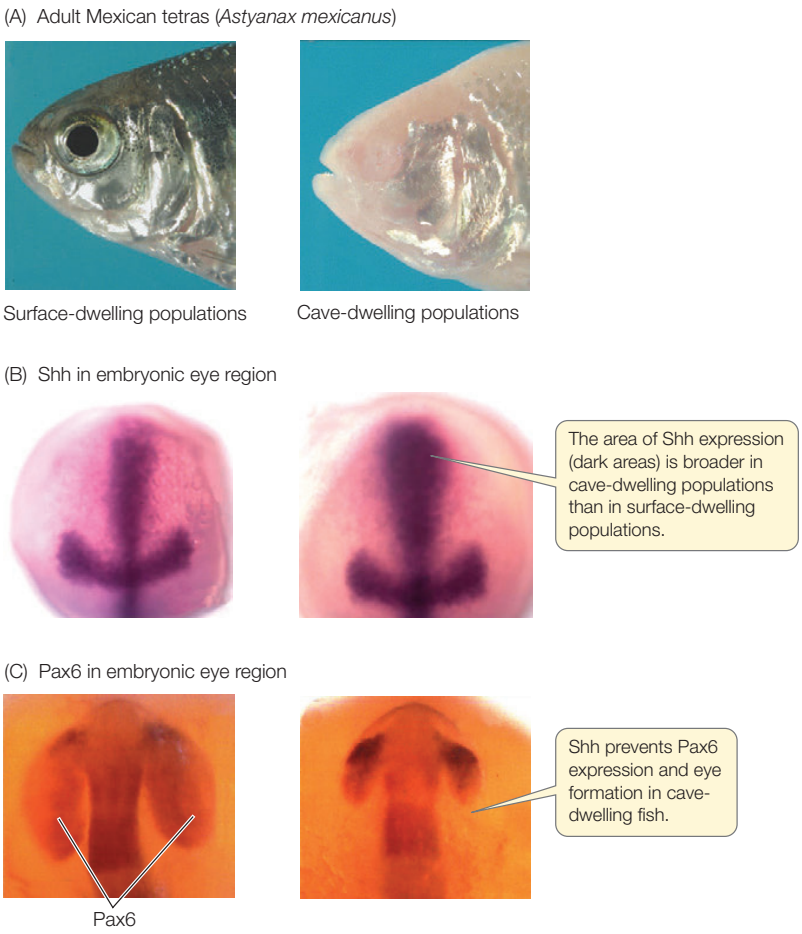


FIGURE 14.22 Inhibition of a Molecular Switch Results in No Eyes (A) In the Mexican tetra fish (*Astyanax mexicanus*), fish dwelling on the surface have two eyes (left), whereas those living in dark caves have no eyes (right). The difference results from overexpression of the *Shh* gene in cave-dwelling fish (B), which inhibits production of the molecular switch made by the *Pax*6 gene (C).

SUMMARY

CONCEPT 14.1 Development Involves Distinct but Overlapping Processes

- A multicellular organism begins its development as an embryo, and several embryonic stages precede the production of an independent organism. **Review Figure 14.1 and ACTIVITY 14.1**
- The processes of development are **determination**, **differentiation**, **morphogenesis**, and **growth**.
- The zygote is **totipotent**; it is capable of producing an entire new organism, with every type of cell in the adult body. **Review ANIMATED TUTORIAL 14.1**
- The ability to create clones from differentiated cells demonstrates the principle of **genomic equivalence**. **Review Figures 14.3 and 14.4**
- **Multipotent** stem cells occur in the growing regions of many tissues in plants and animals. They constantly divide and form a pool of cells that can be used for differentiation to specialized cells. **Review Figure 14.5**
- **Pluripotent** stem cells can form every cell type of a mammal, but not an entire organism. They occur in the embryo and can be induced to form in the laboratory. They may have medical uses. **Review Figure 14.6 and ANIMATED TUTORIAL 14.2**

CONCEPT 14.2 Changes in Gene Expression Underlie Cell Fate Determination and Differentiation

- Differential gene expression results in cell differentiation. Transcription factors are especially important in regulating gene expression during differentiation.
- **Cytoplasmic segregation**—the unequal distribution of **cytoplasmic determinants** in the egg, zygote, or early embryo—can establish **polarity** and lead to cell fate determination. **Review Figure 14.7 and ANIMATED TUTORIAL 14.3**
- Induction is a process by which embryonic animal tissues direct the development of neighboring cells and tissues by secreting chemical signals called **inducers**. **Review Figure 14.9**

CONCEPT 14.3 Spatial Differences in Gene Expression Lead to Morphogenesis

- During development, selective elimination of cells by apoptosis results from the expression of specific genes.
- Both plants and animals use **positional information** in the form of a signal called a **morphogen** to stimulate cell determination. **Review Figure 14.11**
- In plants, **organ identity genes** encode polypeptides that associate to form transcription factors. These proteins determine the formation of flower organs. **Review Figure 14.12**
- In the fruit fly *Drosophila melanogaster*, a cascade of transcriptional activation sets up the axes of the embryo, the development of the segments, and the determination of cell fate in each segment. **Review Figure 14.13 and ANIMATED TUTORIAL 14.4**
- **Hox genes** determine cell fate in the embryos of many animals. The **homeobox** is a DNA sequence found in Hox genes and other genes that code for transcription factors. The sequence of amino acids encoded by the homeobox is called the homeodomain.

CONCEPT 14.4 Changes in Gene Expression Pathways Underlie the Evolution of Development

- **Evolutionary developmental biology** (**evo-devo**) is the modern study of the evolutionary aspects of development, and it focuses on molecular mechanisms.
- Hox genes have evolved from a common ancestor. **Review Figure 14.15 and ACTIVITY 14.2**
- Genes such as Hox genes underlie evolutionary changes in morphology that produce major differences in body forms.
- Evolutionary diversity is produced using a modest number of regulatory genes. **Review Figure 14.16**
- The transcription factors and chemical signals that govern pattern formation in the bodies of multicellular organisms, and the genes that encode them, can be thought of as a **genetic toolkit**.
- The bodies of developing and mature organisms are organized into self-contained units that can be modified independently in space and time. **Review ANIMATED TUTORIAL 14.5**
- Changes in **genetic switches** that determine where and when a set of genes will be expressed underlie the transformation of an individual from egg to adult.

CONCEPT 14.5 Developmental Genes Contribute to Species Evolution but Also Pose Constraints

- Evolutionary innovations are modifications of preexisting structures. **Review Figure 14.20**
- Because many genes that govern development have been highly conserved, similar traits are likely to evolve repeatedly, especially among closely related species. This process is called **parallel phenotypic evolution**. **Review Figure 14.21**

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is14

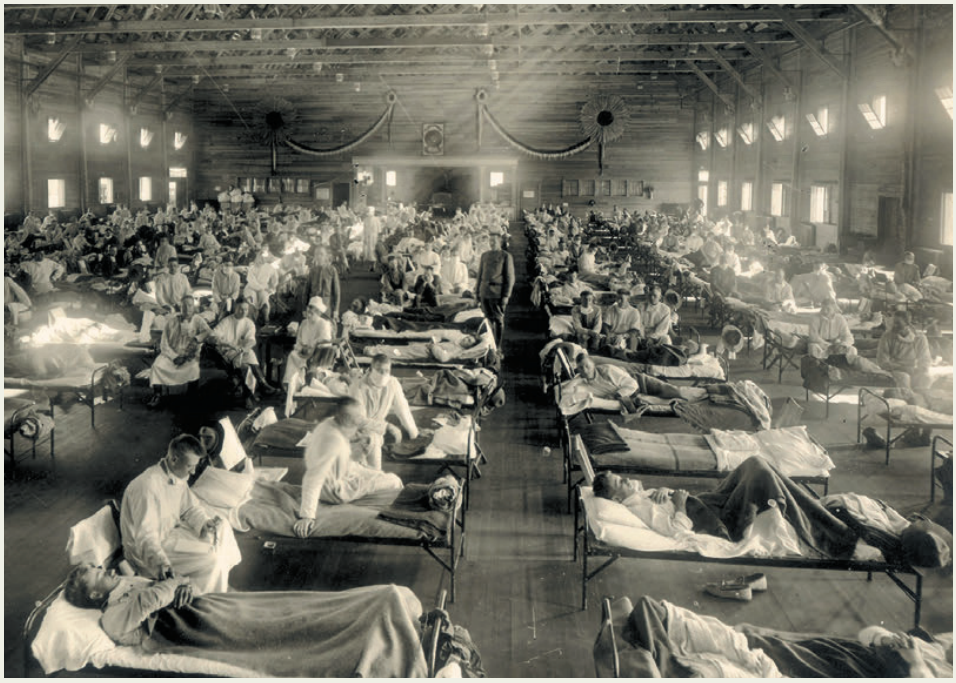
Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

PART 3

Evolution

15 Processes of Evolution

KEY CONCEPTS



Flu victims are treated at a U.S. Army hospital in 1918.

On November 11, 1918, an armistice agreement signed in France signaled the end of World War I. But the death toll from four years of war was soon surpassed by the casualties of a massive influenza epidemic that began in the spring of 1918 among soldiers in a U.S. Army barracks. Over the next 18 months, this particular strain of flu virus spread across the globe, killing more than 50 million people worldwide—more than twice the number of World War I–related combat deaths.

The 1918–1919 pandemic was noteworthy because the death rate among young adults—who are usually less likely to die from influenza than are the elderly or the very young—was 20 times higher than in flu epidemics before or since. Why was that particular virus so deadly, especially to typically hardy individuals? The 1918 flu strain triggered an especially intense reaction in the human immune system. This overreaction meant that people with strong immune systems were likely to be more severely affected.

In most cases, however, our immune system helps us fight viruses; this response is the basis of vaccination. Since 1945, programs to administer flu vaccines have helped keep the number and severity of influenza outbreaks in check. Last year's vaccine, however, will probably not be effective against this year's virus. New strains of flu virus are evolving continuously, ensuring genetic variation in the population. If these viruses did not evolve, we would become resistant to them and annual vaccination would become unnecessary. But because the viruses do evolve, biologists must develop a new and different flu vaccine each year.

Vertebrate immune systems recognize proteins on the viral surface, and changes in these proteins mean that the virus can escape immune detection. Virus strains with the greatest number of changes to their surface proteins are most likely to avoid detection and infect their hosts, and thus have an advantage over other strains. Biologists can observe evolution in action by following changes in influenza virus proteins from year to year.

We learn a great deal about the processes of evolution by examining rapidly evolving organisms such as viruses, and these studies contribute to the development of evolutionary theory. Evolutionary theory, in turn, is put to practical uses, such as the development of better strategies for combating deadly diseases.

Q How do biologists use evolutionary theory to develop better flu vaccines?

You will find the answer to this question on page 322.

CONCEPT 15.1 Evolution Is Both Factual and the Basis of Broader Theory

All biological populations change in their genetic makeup over time. This change in the genetic composition of populations over time is called **evolution**. We can, and do, observe evolutionary change on a regular basis, both in laboratory experiments and in natural populations. We measure the rate at which new mutations arise, observe the spread of new genetic variants through a population, and see the effects of genetic change on the form and function of organisms. In the fossil record, we observe the long-term morphological changes (which are the result of underlying genetic changes) that have occurred among living organisms. These underlying changes in the genetic makeup of populations drive the origin and extinction of species and fuel the diversification of life.

In addition to observing and recording physical changes over evolutionary time, biologists have accumulated a large body of evidence about *how* these changes occur, and about *what* evolutionary changes have occurred in the past. The resulting understanding and application of the processes of evolutionary change to biological problems is known as **evolutionary theory**.

Evolutionary theory has many useful applications. We constantly apply it to the study and treatment of diseases. Evolutionary theory is critical to the development of better agricultural crops and practices, and to the development of industrial processes that produce new molecules with useful properties. At a more basic level, knowledge of evolutionary theory allows biologists to understand how life diversified. It also helps us make predictions about the biological world.

In everyday speech, people tend to use the word "theory" to mean an untested hypothesis, or even a guess. But evolutionary theory does not refer to any single hypothesis, and it certainly is not guesswork. The concept of evolutionary change among living organisms was present among a few scientists even before Charles Darwin so clearly described his observations, presented his conclusions, and articulated the premise of natural selection in his book *On the Origin of Species*. The rediscovery of Mendel's experiments and the subsequent establishment of the principles of genetic inheritance early in the 1900s set the stage for vast amounts of research. By the end of the twentieth century, findings from many fields of biology firmly upheld Darwin's basic premises about the common ancestry of life and the role of natural selection as an important process of evolution. Today a vast and rich array of geological, morphological, behavioral, and molecular data all support and expand the factual basis of evolution. Observations of fossils and natural populations are supported by experiments that demonstrate the basic operation of evolutionary processes.

When we refer to evolutionary theory, we are referring to our understanding of the processes that result in genetic changes in populations over time. We then apply that understanding to interpret the changes we observe in natural populations. We can directly observe the evolution of influenza viruses, but it is evolutionary theory that allows us to apply our observations to the task of developing more effective vaccines. Several processes of evolutionary change are recognized, and the scientific community is continually using evolutionary theory to expand its understanding of how and when these processes apply to particular biological problems.

Go to MEDIA CLIP 15.1
Watching Evolution in Real Time
PoL2e.com/mc15.1

Darwin and Wallace introduced the idea of evolution by natural selection

In the early 1800s, it was not yet evident to many people that populations of living organisms evolve. But several biologists had suggested that the species living on Earth had changed over time—that is, that evolution had taken place. Jean-Baptiste Lamarck, for one, presented strong evidence for the fact of evolution in 1809, but his ideas about *how* it occurred were not convincing. At that time, no one had yet envisioned a viable process for evolution.

Charles Robert Darwin

In the 1820s, a young Charles Darwin became passionately interested in the subjects of geology (with its new sense of Earth's great age) and natural history (the scientific study of how different organisms function and carry out their lives in nature). Despite these interests, he planned, at his father's behest, to become a doctor. But surgery conducted without anesthesia nauseated Darwin, and he gave up medicine to study at Cambridge University for a career as a clergyman in the Church of England. Always more interested in science than in theology, he gravitated toward scientists on the faculty, especially the botanist John Henslow. In 1831, Henslow recommended Darwin for a position on HMS *Beagle*, a Royal Navy vessel that was preparing for a survey voyage around the world.

HMS *Beagle*

Whenever possible during the 5-year voyage (**FIGURE 15.1**), Darwin went ashore to study rocks and to observe and collect plants and animals. He noticed striking differences between the species he saw in South America and those of Europe. He

FIGURE 15.1 The Voyage of the *Beagle* The mission of HMS *Beagle* was to chart the oceans and collect oceanographic and biological information from around the world. The world map indicates the ship's path; the inset map shows the Galápagos Islands, whose organisms were an important source of Darwin's ideas on natural selection.

Go to ACTIVITY 15.1
Darwin's Voyage
PoL2e.com/ac15.1

observed that the species of the temperate regions of South America (Argentina and Chile) were more similar to those of tropical South America (Brazil) than they were to temperate European species. When he explored the islands of the Galápagos archipelago west of Ecuador, he noted that most of the animals were endemic to the islands (that is, unique and found nowhere else), although they were similar to animals found on the mainland of South America. Darwin also observed that the fauna of the Galápagos differed from island to island. He postulated that some animals had come to the archipelago from mainland South America and had subsequently undergone different changes on each of the islands. He wondered what might account for these changes.

When he returned to England in 1836, Darwin continued to ponder his observations. His thoughts were strongly influenced by the geologist Charles Lyell, who had recently popularized the idea that Earth had been shaped by slow-acting forces that are still at work today. Darwin reasoned that similar thinking could be applied to the living world. Within a decade, he had developed the framework of an explanatory theory for evolutionary change based on three major propositions:

- Species are not immutable; they change over time.
- Divergent species share a common ancestor and have diverged from one another gradually over time (a concept Darwin termed **descent with modification**).
- Changes in species over time can be explained by **natural selection**: the increased survival and reproduction of some individuals compared with others, based on differences in their traits.

The first of these propositions was not unique to Darwin; several earlier authors had argued for the fact of evolution. A more revolutionary idea was his second proposition, that *divergent species are related to one another through common descent*. But Darwin is probably best known for his third proposition, that of natural selection.

Darwin realized that many more individuals of most species are born than survive to reproduce. He also knew that, although offspring usually resemble their parents, offspring are not identical to one another or to either parent. Finally, he was well aware of the fact that human breeders of plants and animals often selected their breeding stock based on the occurrence of particular traits. Over time, this selection resulted in dramatic changes in the appearance of the descendants of those plants or animals. In natural populations, wouldn't the individuals with the best chances of survival and reproduction be similarly "selected," and thus pass their traits on to the next generation? Darwin's simple but powerful idea was that nature did the selecting in natural populations on the basis of traits that resulted in greater survival and, eventually, greater likelihood of reproduction.

In 1844, Darwin wrote a long essay describing the role of natural selection as a process of evolution. But he was reluctant to publish it, preferring to assemble more evidence first. Darwin's hand was forced in 1858, when he received a letter and manuscript from another traveling English naturalist, Alfred Russel Wallace, who was studying the plants and animals of the Malay Archipelago. Wallace asked Darwin to evaluate his manuscript, which included an explanation of natural selection almost identical to Darwin's. Darwin was at first dismayed, believing Wallace to have preempted his idea. Parts of Darwin's 1844 essay, together with Wallace's manuscript, were presented to the Linnaean Society of London on July 1, 1858, thereby crediting both men for the idea of natural selection. Darwin then worked quickly to finish his own book, *On the Origin of Species*, which was published the following year.

Go to ANIMATED TUTORIAL 15.1
Natural Selection
PoL2e.com/at15.1

Although Darwin and Wallace independently articulated the concept of natural selection, Darwin developed his ideas first. Furthermore, *On the Origin of Species* proved to be a stunning work of scholarship that provided exhaustive evidence from many fields supporting both the premise of evolution itself and the notion of natural selection as a process of evolution. Thus both concepts are more closely associated with Darwin than with Wallace.

The publication of *On the Origin of Species* in 1859 stirred considerable interest (and controversy) among scientists and the public alike. Scientists spent much of the rest of the nineteenth century amassing biological and paleontological data to test evolutionary ideas and document the history of life on Earth. By 1900, the fact of biological evolution (defined at that time as change in the physical characteristics of populations over time) was established beyond any reasonable doubt. But the *genetic* basis of evolutionary change was not yet understood.

Evolutionary theory has continued to develop over the past century

Shortly after 1900, several individuals rediscovered the work of Gregor Mendel (which had been published in 1866 but rarely read or cited), and the basic processes of genetic inheritance began to be unraveled. In the first decades of the twentieth century, Thomas Hunt Morgan's studies on fruit flies led to his discovery of the role of chromosomes in inheritance. In the 1920s and early 1930s, the major principles of population genetics were established, the genetic basis of new variation (i.e., mutations) began to be understood, and processes of evolution such as genetic drift were described (see Concept 15.2). This work set the stage for a "modern synthesis" of genetics and evolution that took place over the period 1936–1947. Some of the major contributors to this synthesis and a few of their books are listed in **FIGURE 15.2**.

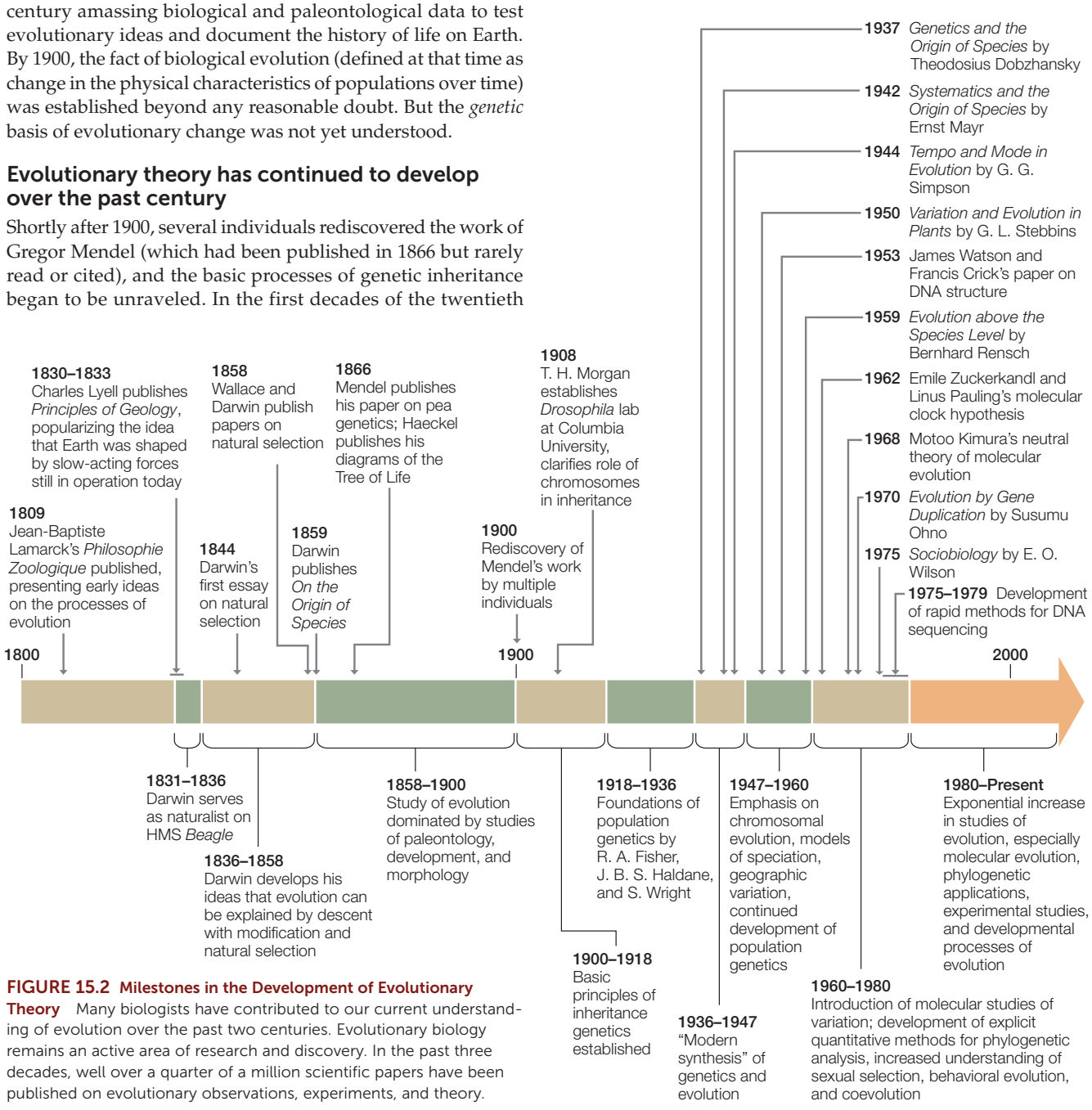


FIGURE 15.2 Milestones in the Development of Evolutionary Theory Many biologists have contributed to our current understanding of evolution over the past two centuries. Evolutionary biology remains an active area of research and discovery. In the past three decades, well over a quarter of a million scientific papers have been published on evolutionary observations, experiments, and theory.

Although chromosomes were understood to be the basis of genetic transmission in eukaryotes by the early 1900s, their molecular structure remained a mystery until soon after the modern synthesis. Then, in 1953, Watson and Crick published their paper on the structure of DNA, opening the door to our current detailed understanding of molecular evolutionary processes. By the 1960s, biologists could study and document changes in allele frequencies in populations over time (see Concept 15.3). Most of this early work necessarily focused on variants of proteins that differed within and between populations and species. Even though the molecular structure of DNA was known, it was not yet practical to sequence long stretches of DNA. Nonetheless, many important advances occurred in evolutionary theory during this time (see Figure 15.2), and these advances were not focused solely on a genetic understanding of evolution. E. O. Wilson's 1975 book *Sociobiology*, for example, invigorated studies of the evolution of behavior (a subject that had fascinated Darwin).

In the late 1970s, several techniques were developed that allowed the rapid sequencing of long stretches of DNA, which in turn allowed researchers to determine the amino acid sequences of proteins. This ability opened a new door for evolutionary biologists, who could now explore the structure of genes and proteins and document evolutionary changes within and between species in ways never before possible.

CHECKpoint CONCEPT 15.1

- ✓ How would you respond to someone who said that evolution was "just a theory"?
- ✓ Why do you think Darwin and Wallace formulated their ideas on natural selection at about the same time?
- ✓ Discuss the significance of each of the following scientific advances for evolutionary theory:
  a. Elucidation of the principles of chromosomal inheritance
  b. The discovery of DNA, its structure, and the universal genetic code
  c. Technology that allows us to sequence long segments of DNA

  Keep your discussion in mind as you continue reading this chapter.

Natural selection is not the only process that drives evolution, although the importance of natural selection to evolution has been confirmed in many thousands of scientific studies. In the next section we'll consider a more complete view of evolutionary processes and how they operate.

CONCEPT 15.2 Mutation, Selection, Gene Flow, Genetic Drift, and Nonrandom Mating Result in Evolution

The word "evolution" is often used in a general sense to mean simply "change," but in a biological context "evolution" refers specifically to change in the genetic makeup of populations over time. Developmental changes that occur in a single individual over the course of the life cycle are not the result of evolutionary change. Evolution is genetic change occurring in a **population**—a group of individuals of a single species that live and interbreed in a particular geographic area at the same time. It is important to remember that *individuals do not evolve; populations do*.

The premise of natural selection was one of Darwin's principal insights and has been demonstrated to be an important process of evolution, but natural selection does not act alone. Additional processes—gene flow, genetic drift, and nonrandom mating—affect the genetic makeup of populations over time. Before we consider how these processes change the frequencies of gene variants in a population, however, we need to understand how mutation brings such variants into existence.

Mutation generates genetic variation

The origin of novel genetic variation is mutation. As described in Concept 9.3, a mutation is any change in the nucleotide sequences of an organism's DNA. The process of DNA replication is not perfect, and some changes appear almost every time a genome is replicated. Mutations occur randomly with respect to an organism's needs; it is natural selection acting on this random variation that results in adaptation. Most mutations are either harmful to their bearers (deleterious mutations) or have no effect (neutral mutations). But a few mutations are beneficial, and even previously deleterious or neutral alleles may become advantageous if environmental conditions change. In addition, mutation can restore genetic variation that other evolutionary processes have removed. Thus mutation both creates and helps maintain genetic variation in populations.

Mutation rates can be high, as we saw in the case of the influenza viruses described at the opening of this chapter, but in many organisms the mutation rate is very low (on the order of 10^{-8} to 10^{-9} changes per base pair of DNA per generation). Even low overall mutation rates, however, create considerable genetic variation, because each of a large number of genes may change, and populations often contain large numbers of individuals. For example, if the probability of a point mutation (an addition, deletion, or substitution of a single base) were 10^{-9} per base pair per generation, then each human gamete—the DNA of which contains 3×10^9 base pairs—would average three new point mutations ($3 \times 10^9 \times 10^{-9} = 3$), and each zygote would carry an average of six new mutations. The current human population of about 7 billion people would thus be expected to carry about 42 billion new mutations (i.e., changes in the nucleotide sequences of their DNA that were not present one generation earlier). So even though the mutation rate in humans is low, human populations still contain enormous genetic variation on which other evolutionary processes can act.

As a result of mutation, different forms of a gene, known as **alleles**, may exist at a particular chromosomal locus. At any particular locus, a single diploid individual has no more than two of the alleles found in the population to which it belongs. The sum of all copies of all alleles at all loci found in a population constitutes its gene pool (**FIGURE 15.3**). (We can also refer

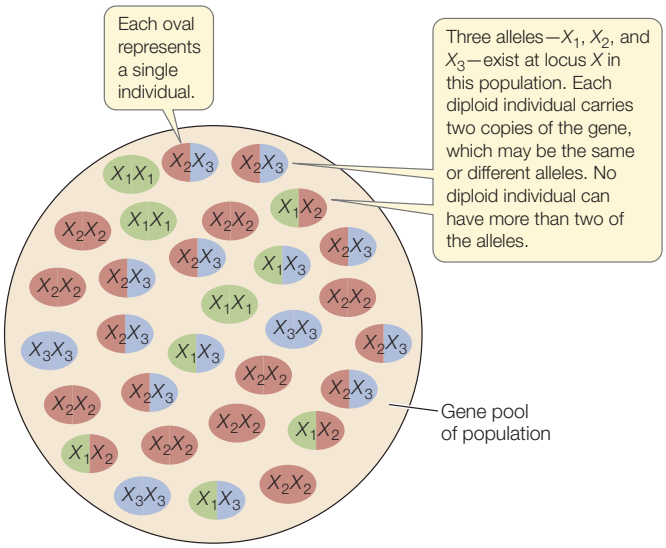


FIGURE 15.3 A Gene Pool A gene pool is the sum of all the alleles found in a population or at a particular locus. This figure shows the gene pool for one locus, X. The allele frequencies in this case are 0.20 for X_1, 0.50 for X_2, and 0.30 for X_3 (see Figure 15.11).

to the gene pool for a particular chromosomal locus or loci.) The gene pool is the sum of the genetic variation in the population. The proportion of each allele in the gene pool is the allele frequency. Likewise, the proportion of each genotype among individuals in the population is the genotype frequency.

A simple experiment demonstrates how mutations accumulate in populations in a continuous, almost constant fashion over time (**FIGURE 15.4**). Lines of the bacterium *E. coli* were grown in the laboratory for 20,000 generations, and the genomes were sequenced from individuals in the experimental lines at least once every 5,000 generations. Over the experiment, the lines accumulated about 45 changes to their genomes, and these changes appeared at a fairly constant rate over time. All populations experience a similar accumulation of mutations over time (although the rate of change differs among species), and these changes provide the raw material for evolution.

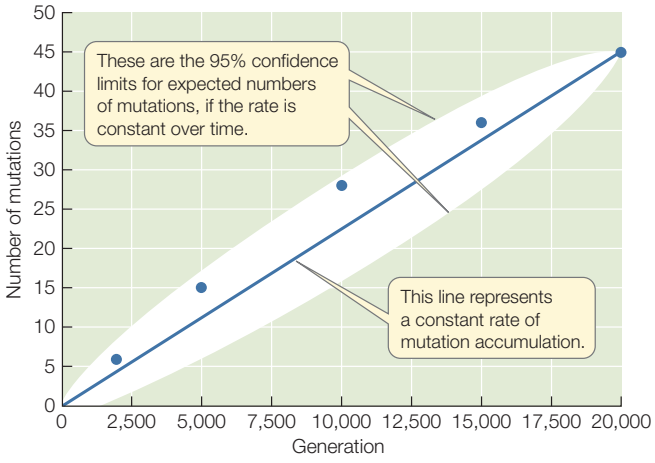


FIGURE 15.4 Mutations Accumulate Continuously An experimental lineage of the bacterium *Escherichia coli* was propagated in the laboratory for 20,000 generations. Genomes were sequenced from individuals sampled at various points during the experiment and were compared with the genome of the ancestral clone. Note that mutations accumulated at a relatively constant rate throughout the experiment.

LINK

Review the nature of alleles and genetic inheritance in **Concepts 8.1 and 8.2**

Selection on genetic variation leads to new phenotypes

As a result of mutation, the gene pools of nearly all populations contain variation for many traits. Selection that favors different traits can lead to many different lineages that descend from the same ancestor. For example, artificial selection on different traits in a single European species of wild mustard produced many important crop plants (**FIGURE 15.5**). Agriculturalists

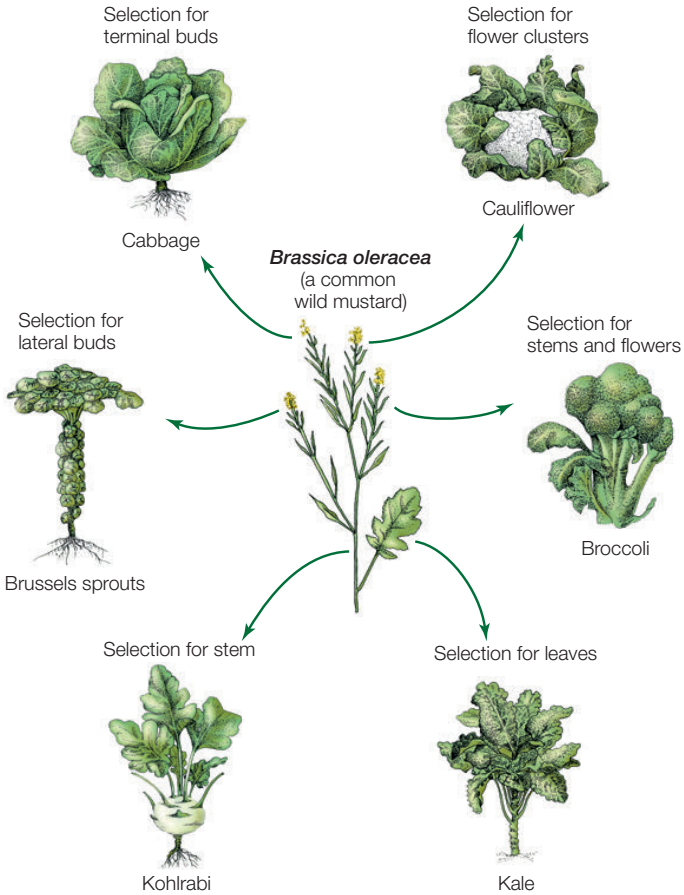


FIGURE 15.5 Many Vegetables from One Species All of the crop plants shown here derive from a single wild mustard species. European agriculturalists produced these crop species by selecting and breeding plants with unusually large buds, stems, leaves, or flowers. The results substantiate the vast amount of variation present in a gene pool.

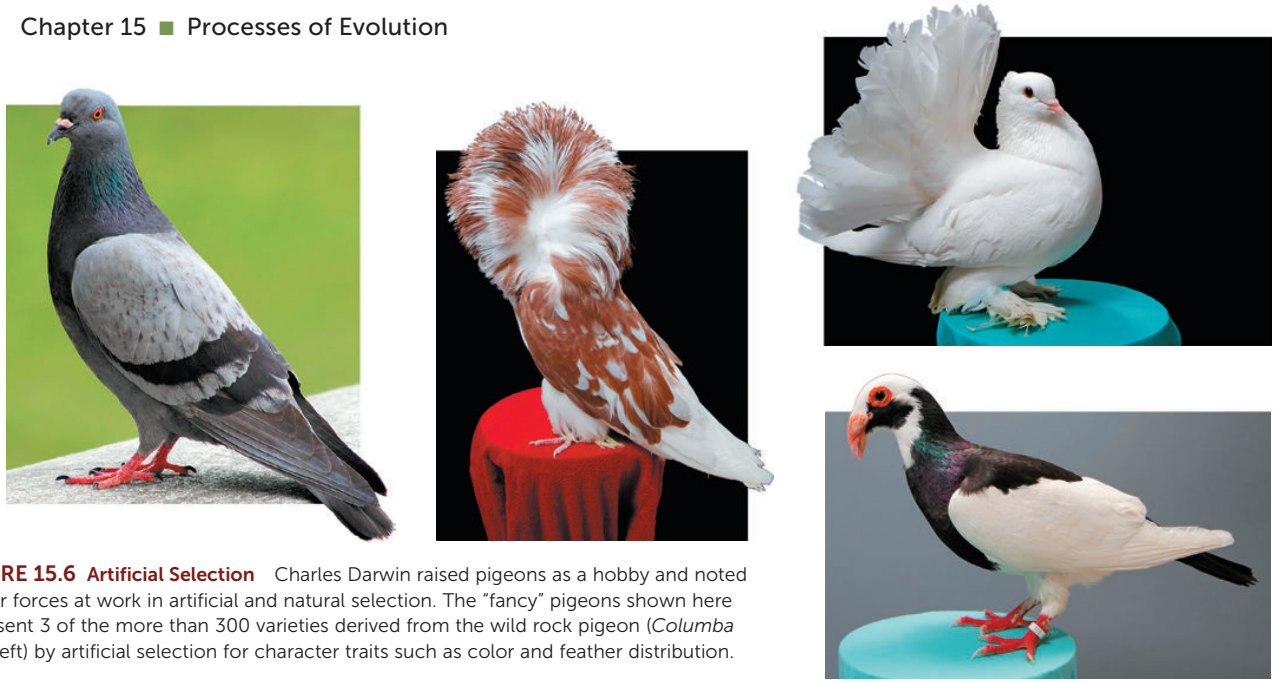


FIGURE 15.6 Artificial Selection Charles Darwin raised pigeons as a hobby and noted similar forces at work in artificial and natural selection. The "fancy" pigeons shown here represent 3 of the more than 300 varieties derived from the wild rock pigeon (*Columba livia*; left) by artificial selection for character traits such as color and feather distribution.

were able to achieve these results because the original mustard population had genetic variation for the characteristics of interest (such as stem thickness or number of leaves).

Darwin compared this artificial selection, which was commonly practiced by animal and plant breeders, with natural selection that occurred in natural populations. Many of Darwin's observations on the nature of variation and selection came from domesticated plants and animals. Darwin bred pigeons and thus knew firsthand the astonishing diversity in color, size, form, and behavior that breeders could achieve (**FIGURE 15.6**). He recognized close parallels between selection by breeders and selection in nature. Whereas artificial selection resulted in traits that were preferred by the human breeders, natural selection resulted in traits that helped organisms survive and reproduce more effectively. In both cases, selection simply increased the frequency of the favored trait from one generation to the next.

Laboratory experiments also demonstrate the existence of considerable genetic variation in populations, and show how this variation can lead to evolution through selection. In one such experiment, investigators bred populations of the fruit fly *Drosophila melanogaster* with high or low numbers of bristles on their abdomens from an initial population with intermediate numbers of bristles. After 35 generations, all flies in both the high- and low-bristle lineages had bristle numbers that fell well outside the range found in the original population (**FIGURE 15.7**). Selection for high and low bristle numbers resulted in new combinations of the many different genes that were present in the original population, so that the phenotypic variation seen in subsequent generations fell outside the phenotypic variation seen in the original population.

Natural selection increases the frequency of beneficial mutations in populations

Darwin knew that far more individuals of most species are born than survive to reproduce. He also knew that, although offspring tend to resemble their parents, the offspring of most organisms are not identical either to their parents or to one another. He suggested that slight differences among individuals affect the chance that a given individual will survive and reproduce, which increases the frequency of the favored trait in the next generation. A favored trait that evolves through natural selection is known as an **adaptation**; this word is used to describe both the trait itself and the process that produces the trait.

Biologists regard an organism as being adapted to a particular environment when they can demonstrate that a slightly

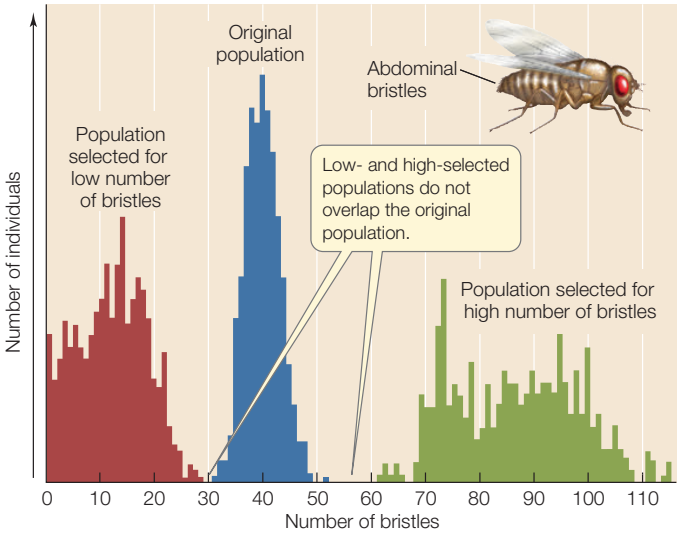


FIGURE 15.7 Artificial Selection Reveals Genetic Variation When investigators subjected *Drosophila melanogaster* to artificial selection for abdominal bristle number, that character evolved rapidly. The graph shows the number of flies with different numbers of bristles in the original population and after 35 generations of artificial selection. The bristle numbers of the selected lineages clearly diverged from those of the original population.

different organism is less likely to survive and reproduce in that environment. To understand adaptation, biologists compare the performances of individuals that differ in their traits.

Natural selection also acts to remove deleterious mutations from populations. Individuals with deleterious mutations are less likely to survive and reproduce, so they are less likely to pass their alleles on to the next generation.

Gene flow may change allele frequencies

Few populations are completely isolated from other populations of the same species. Migration of individuals and movements of gametes (in pollen, for example) between populations—a phenomenon called **gene flow**—can change allele frequencies in a population. If the arriving individuals survive and reproduce in their new location, they may add new alleles to the population's gene pool, or they may change the frequencies of alleles present in the original population.

LINK

If gene flow between two populations stops, those populations may diverge and become different species; see **Concept 17.2**

Genetic drift may cause large changes in small populations

In small populations, **genetic drift**—random changes in allele frequencies from one generation to the next—may produce large changes in allele frequencies over time. Harmful alleles may increase in frequency, and rare advantageous alleles may be lost. Even in large populations, genetic drift can influence the frequencies of neutral alleles (which do not affect the survival and reproductive rates of their bearers).

To illustrate the effects of genetic drift, suppose there are only two females in a small population of normally brown mice, and one of these females carries a newly arisen dominant allele that produces black fur. Even in the absence of any selection, it is unlikely that the two females will produce exactly the same number of offspring. Even if they do produce identical litter sizes and identical numbers of litters, chance events that have nothing to do with genetic characteristics are likely to result in differential mortality among their offspring. If each female produces one litter, but a flood envelops the black female's nest and kills all of her offspring, the novel allele could be lost from the population in just one generation. In contrast, if the brown female's litter is lost, then the frequency of the newly arisen allele (and phenotype) for black fur will rise dramatically in just one generation.

Genetic drift is especially potent when a population is reduced dramatically in size. Even populations that are normally large may occasionally pass through environmental events that only a small number of individuals survive, a situation known as a **population bottleneck**. The effect of genetic drift in such a situation is illustrated in **FIGURE 15.8**, in which red and yellow beans represent two alleles of a gene. Most of the beans in the small sample of the "population" that "survives" the bottleneck event are, just by chance, red, so the new population has a much higher frequency of red beans than the previous generation had. In a real population, the red and yellow allele frequencies would be described as having "drifted."

A population forced through a bottleneck is likely to lose much of its genetic variation. For example, when Europeans first arrived in North America, millions of greater prairie-chickens (*Tympanuchus cupido*) inhabited the midwestern prairies. As a result of hunting and habitat destruction by the new settlers, the Illinois population of this species plummeted from about 100 million birds in 1900 to fewer than 50 individuals in the 1990s. A comparison of DNA from birds collected in Illinois during the middle of the twentieth century with DNA from the surviving population in the 1990s showed that Illinois prairie-chickens have lost most of their genetic diversity. Loss of genetic variation in small populations is one of the problems facing biologists who attempt to protect endangered species.

Genetic drift can have similar effects when a few pioneering individuals colonize a new region. Because of its small size, the colonizing population is unlikely to possess all of the alleles found in the gene pool of its source population. The resulting change in genetic variation, called a **founder effect**, is equivalent to that in a large population reduced by a bottleneck.

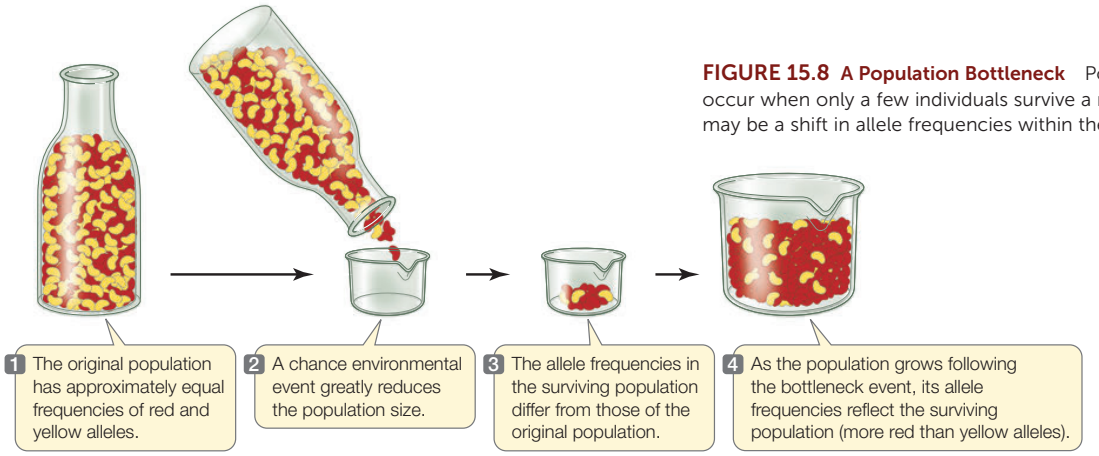


FIGURE 15.8 A Population Bottleneck Population bottlenecks occur when only a few individuals survive a random event. The result may be a shift in allele frequencies within the population.

Nonrandom mating can change genotype or allele frequencies

Mating patterns often alter genotype frequencies because the individuals in a population do not choose mates at random. For example, self-fertilization is common in many groups of organisms, especially plants. Any time individuals mate preferentially with other individuals of the same genotype (including themselves), homozygous genotypes will increase in frequency and heterozygous genotypes will decrease in frequency over time. The opposite effect (more heterozygotes, fewer homozygotes) is expected when individuals mate primarily or exclusively with individuals of different genotypes.

Nonrandom mating systems that do not affect the relative reproductive success of individuals produce changes in genotype frequencies but not in allele frequencies, and thus do not, by themselves, result in evolutionary change in a population. However, nonrandom mating systems that result in different reproductive success among individuals do produce allele frequency changes from one generation to the next. **Sexual selection** occurs when individuals of one sex mate preferentially with particular individuals of the opposite sex rather than at random.

Sexual selection was first suggested by Charles Darwin, who developed the idea to explain the evolution of conspicuous traits that would appear to inhibit survival, such as bright colors and elaborate courtship displays in males of many species. He hypothesized that these features either improved the ability of their bearers to compete for access to mates (intrasexual selection) or made their bearers more attractive to members of the opposite sex (intersexual selection). The concept of sexual selection was either ignored or questioned for many decades, but recent investigations have demonstrated its importance.

Darwin argued that while natural selection typically favors traits that enhance the survival of their bearers or their bearers' descendants, sexual selection is primarily about successful reproduction. An animal that survives but fails to reproduce makes no contribution to the next generation. Thus sexual selection may favor traits that enhance an individual's chances of reproduction even when these traits reduce its chances of survival. For example, females may be more likely to see or hear males with a given trait (and thus be more likely to mate with those males), even though the favored trait also increases the chances that the male will be seen or heard by a predator.

LINK

Some of the animal behaviors that have evolved in response to sexual selection are described in **Concepts 40.5 and 40.6**

One example of a trait that Darwin attributed to sexual selection is the remarkable tail of the male African long-tailed widowbird (*Euplectes progne*), which is longer than the bird's head and body combined (**FIGURE 15.9**). Male widowbirds normally select, and defend from other males, a territory where they perform courtship displays to attract females. To investigate whether sexual selection drove the evolution of widowbird tails, a biologist clipped the tails of some captured male widowbirds and lengthened the tails of others by gluing on additional feathers. He then cut and reglued the tail feathers of still other males, which served as controls. Both short- and long-tailed males successfully defended their display territories, indicating that a long tail does not confer an advantage in male–male competition. However, males with artificially elongated tails attracted about four times more females than did males with shortened tails (**FIGURE 15.10**). Thus males with long tails pass on their genes to more offspring than do males with short tails, which leads to the evolution of this unusual trait.

Euplectes progne

FIGURE 15.9 What Is the Advantage? The extensive tail of the male African long-tailed widowbird actually inhibits its ability to fly. Darwin attributed the evolution of this seemingly nonadaptive trait to sexual selection.

CHECKpoint CONCEPT 15.2

- ✓ How do deleterious, neutral, and beneficial mutations differ?
- ✓ Can you explain how natural selection results in an increase in the frequency of beneficial alleles in a population over time, and a decrease in the frequency of deleterious alleles?
- ✓ How can genetic drift cause large changes in small populations?
- ✓ How do self-fertilization and sexual selection differ in their expected effects on genotype and allele frequencies over time?

The processes of mutation, selection, gene flow, genetic drift, and nonrandom mating can all result in evolutionary change. We will consider next how evolutionary change that results from these processes is measured.

INVESTIGATION

FIGURE 15.10 Sexual Selection in Action Behavioral ecologist Malte Andersson tested Darwin's hypothesis that excessively long tails evolved in male widowbirds because female preference for longer-tailed males increased their mating and reproductive success.[a]

HYPOTHESIS

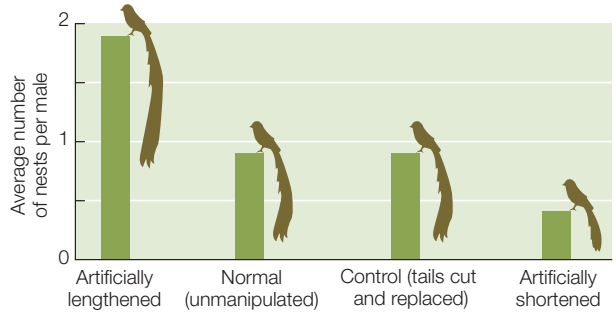
Female widowbirds prefer to mate with the male that displays the longest tail; longer-tailed males thus are favored by sexual selection because they will father more offspring.

METHOD

1. Capture males and artificially lengthen or shorten tails by cutting or gluing on feathers. In a control group, cut and replace tails to their normal length (to control for the effects of tail-cutting).
2. Release the males to establish their territories and mate.
3. Count the nests with eggs or young on each male's territory.

RESULTS

Male widowbirds with artificially shortened tails established and defended display sites sucessfully but fathered fewer offspring than did control or unmanipulated males. Males with artificially lengthened tails fathered the most offspring.



CONCLUSION

Sexual selection in *Euplectes progne* has favored the evolution of long tails in the male.

ANALYZE THE DATA

Are the differences plotted above significantly different? See Analyze the Data 15.1 in **LaunchPad** for a simple method to test the statistical significance of the differences using the following data.

	Number of nests per male		
Group	Shortened tail	Control	Elongated tail
1	0	0	2
2	0	0	2
3	2	3	5
4	1	2	4
5	0	1	2
6	0	1	2
7	0	1	0
8	0	0	0
9	1	0	0

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]M. Andersson. 1982. *Nature* 299: 818–820.

CONCEPT 15.3 Evolution Can Be Measured by Changes in Allele Frequencies

Much of evolution occurs through gradual changes in the relative frequencies of different alleles in a population from one generation to the next. Major genetic changes can also be sudden, as happens when two formerly separated populations merge and hybridize, or when genes within a population are duplicated within the genome (see Concept 15.6). But in most cases, we measure evolution by looking at changes in allele and genotype frequencies in populations over time.

To measure allele frequencies in a population precisely, we would need to count every allele at every locus in every individual in the population. Fortunately, we do not need to make such complete measurements because we can reliably estimate allele frequencies for a given locus by counting alleles in a sample of individuals from the population. The sum of all allele frequencies at a locus is equal to 1, so measures of allele frequency range from 0 to 1.

Go to ANIMATED TUTORIAL 15.2
Genetic Drift Simulation
PoL2e.com/at15.2

An allele's frequency is calculated using the following formula:

$$p = \frac{\text{number of copies of the allele in the population}}{\text{total number of copies of all alleles in the population}}$$

If only two alleles (we'll call them A and a) for a given locus are found among the members of a diploid population, those alleles can combine to form three different genotypes: AA, Aa, and aa (see Figure 15.3). A population with more than one allele at a locus is said to be polymorphic ("many forms") at that locus. Applying the formula above, as shown in **FIGURE 15.11**, we can calculate the relative frequencies of alleles A and a in a population of N individuals as follows:

- Let N_{AA} be the number that are homozygous for the A allele (AA).
- Let N_{Aa} be the number that are heterozygous for the two alleles (Aa).
- Let N_{aa} be the number that are homozygous for the a allele (aa).

Note that $N_{AA} + N_{Aa} + N_{aa} = N$, the total number of individuals in the population, and that the total number of copies of both alleles present in the population is $2N$, because each individual is diploid. Each AA individual has two copies of the A allele, and each Aa individual has one copy of the A allele. Therefore the total number of A alleles in the population is $2N_{AA} + N_{Aa}$. Similarly, the total number of a alleles in the population is

RESEARCH TOOLS

FIGURE 15.11 Calculating Allele and Genotype Frequencies
Allele and genotype frequencies for a gene locus with two alleles in the population can be calculated using the equations in panel 1. When the equations are applied to two populations (panel 2), we find that the frequencies of alleles *A* and *a* in the two populations are the same, but the alleles are distributed differently between heterozygous and homozygous genotypes.

1 In any population, where *N* is the total number of individuals in the population:

$$\text{Frequency of allele } A = p = \frac{2N_{AA} + N_{Aa}}{2N} \qquad \text{Frequency of allele } a = q = \frac{2N_{aa} + N_{Aa}}{2N}$$

Frequency of genotype $AA = N_{AA}/N$
Frequency of genotype $Aa = N_{Aa}/N$
Frequency of genotype $aa = N_{aa}/N$

2 Compute the allele and genotype frequencies for two separate populations of $N = 200$:

Population 1 (mostly homozygotes)	Population 2 (mostly heterozygotes)
$N_{AA} = 90$, $N_{Aa} = 40$, and $N_{aa} = 70$	$N_{AA} = 45$, $N_{Aa} = 130$, and $N_{aa} = 25$
$p = \frac{180 + 40}{400} = 0.55$	$p = \frac{90 + 130}{400} = 0.55$
$q = \frac{140 + 40}{400} = 0.45$	$q = \frac{50 + 130}{400} = 0.45$
Freq. AA = 90/200 = 0.45	*Freq. AA* = 45/200 = 0.225
Freq. Aa = 40/200 = 0.20	*Freq. Aa* = 130/200 = 0.65
Freq. aa = 70/200 = 0.35	*Freq. aa* = 25/200 = 0.125

$2N_{aa} + N_{Aa}$. If p represents the frequency of A, and q represents the frequency of a, then

$$p = \frac{2N_{AA} + N_{Aa}}{2N}$$

and

$$q = \frac{2N_{aa} + N_{Aa}}{2N}$$

The calculations in Figure 15.11 demonstrate two important points. First, notice that for each population, $p + q = 1$, which means that $q = 1 - p$. So when there are only two alleles at a given locus in a population, we can calculate the frequency of one allele and obtain the second allele's frequency by subtraction. If there is only one allele at a given locus in a population, its frequency is 1: the population is then monomorphic at that locus, and the allele is said to be **fixed**.

The second thing to notice is that population 1 (consisting mostly of homozygotes) and population 2 (consisting mostly of heterozygotes) have the same allele frequencies for *A* and *a*. Thus they have the same gene pool for this locus. Because the alleles in the gene pool are distributed differently among individuals, however, the genotype frequencies of the two populations differ.

The frequencies of the different alleles at each locus and the frequencies of the different genotypes in a population describe that population's **genetic structure**. Allele frequencies measure the amount of genetic variation in a population, whereas genotype frequencies show how a population's genetic variation is distributed among its members. Other measures, such as the proportion of loci that are polymorphic, are also used to measure variation in populations. With these measurements, it becomes possible to consider how the genetic structure of a population changes or remains the same over generations—that is, to measure evolutionary change.

Evolution will occur unless certain restrictive conditions exist

In 1908, the British mathematician Godfrey Hardy and the German physician Wilhelm Weinberg independently deduced the conditions that must prevail if the genetic structure of a population is to remain the same over time. If the conditions they identified do not exist, then evolution will occur. The resulting principle is known as **Hardy–Weinberg equilibrium**. Hardy–Weinberg equilibrium describes a model in which allele frequencies do not change across generations and genotype frequencies can be predicted from allele frequencies (**FIGURE 15.12**). The principles of Hardy–Weinberg equilibrium apply only to sexually reproducing organisms. Several conditions must be met for a population to be at Hardy–Weinberg equilibrium. Note that the following conditions correspond inversely to the five principal processes of evolution (discussed in Concept 15.2):

- *There is no mutation*. The alleles present in the population do not change, and no new alleles are added to the gene pool.
- *There is no selection among genotypes*. Individuals with different genotypes have equal probabilities of survival and equal rates of reproduction.
- *There is no gene flow*. There is no movement of individuals into or out of the population or reproductive contact with other populations.
- *Population size is infinite*. The larger a population, the smaller will be the effect of genetic drift.
- *Mating is random*. Individuals do not preferentially choose mates with certain genotypes.

If these idealized conditions hold, two major consequences follow. First, the frequencies of alleles at a locus remain constant from generation to generation. Second, following one generation of random mating, the genotype frequencies occur in the following proportions:

Genotype	*AA*	*Aa*	*aa*
Frequency	p^2	$2pq$	q^2

To understand why these consequences are important, start by considering a population that is *not* in Hardy–Weinberg equilibrium, such as generation I in Figure 15.12. This could occur, for example, if the initial population is founded by migrants from several other populations, thus violating the Hardy–Weinberg assumption of no gene flow. In this example,

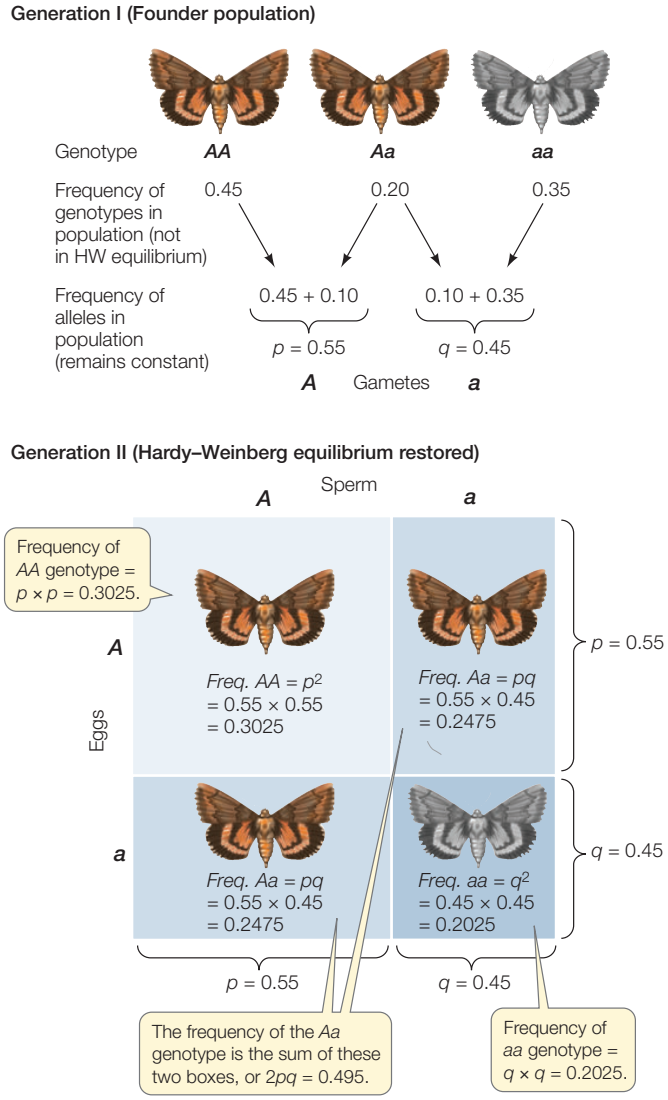


FIGURE 15.12 One Generation of Random Mating Restores Hardy–Weinberg Equilibrium Generation I of this population is made up of migrants from several source populations, and so is not in Hardy–Weinberg equilibrium. After one generation of random mating, the allele frequencies are unchanged, and the genotype frequencies return to Hardy–Weinberg expectations. The lengths of the sides of each rectangle are proportional to the allele frequencies in the population; the areas of the rectangles are proportional to the genotype frequencies.

generation I has more homozygous individuals and fewer heterozygous individuals than would be expected under Hardy–Weinberg equilibrium (a condition known as heterozygote deficiency).

Even with a starting population that is not in Hardy–Weinberg equilibrium, we can predict that after a single generation of random mating, and if the other Hardy–Weinberg assumptions are not violated, the allele frequencies will remain unchanged, but the genotype frequencies will return to Hardy–Weinberg expectations. Let's explore why this is true.

In generation I of Figure 15.12, the frequency of the *A* allele (p) is 0.55. Because we assume that individuals select mates at random, without regard to their genotype, gametes carrying *A* or *a* combine at random—that is, as predicted by the allele frequencies p and q. Thus in this example, the probability that a particular sperm or egg will bear an *A* allele is 0.55. In other words, 55 out of 100 randomly sampled sperm or eggs will bear an *A* allele. Because $q = 1 - p$, the probability that a sperm or egg will bear an *a* allele is $1 - 0.55 = 0.45$.

LINK

You may wish to review the discussion of probability and inheritance in **Concept 8.1**

To obtain the probability of two *A*-bearing gametes coming together at fertilization, we multiply the two independent probabilities of their occurrence:

$$p \times p = p^2 = (0.55)^2 = 0.3025$$

Therefore 0.3025, or 30.25 percent, of the offspring in generation II will have homozygous genotype *AA*. Similarly, the probability of two *a*-bearing gametes coming together is

$$q \times q = q^2 = (0.45)^2 = 0.2025$$

which means that 20.25 percent of generation II will have the *aa* genotype.

There are two ways of producing a heterozygote: an *A* sperm may combine with an *a* egg, the probability of which is $p \times q$; or an *a* sperm may combine with an *A* egg, the probability of which is $q \times p$. Consequently, the overall probability of obtaining a heterozygote is $2pq$, or 0.495. The frequencies of the *AA*, *Aa*, and *aa* genotypes in generation II of Figure 15.12 now meet Hardy–Weinberg expectations, and the frequencies of the two alleles (p and q) have not changed from generation I.

Under the assumptions of Hardy–Weinberg equilibrium, allele frequencies p and q remain constant from generation to generation. If Hardy–Weinberg assumptions are violated and the genotype frequencies in the parental generation are altered (say, by the loss of a large number of *AA* individuals from the population), then the allele frequencies in the next generation will be altered. However, based on the new allele frequencies, another generation of random mating will be sufficient to restore the genotype frequencies to Hardy–Weinberg equilibrium.

Go to ANIMATED TUTORIAL 15.3
Hardy–Weinberg Equilibrium
PoL2e.com/at15.3

Deviations from Hardy–Weinberg equilibrium show that evolution is occurring

You probably have realized that populations in nature never meet the stringent conditions necessary to be at Hardy–Weinberg equilibrium—which explains why all biological populations evolve. Why, then, is this model considered so

APPLY THE CONCEPT

Evolution can be measured by changes in allele frequencies

Imagine you have discovered a new population of curly-tailed lizards established on an island after immigrant lizards have arrived from several different source populations during a hurricane. You collect and tabulate genotype data (right) for the lactate dehydrogenase gene (*Ldh*) for each of the individual lizards. Use the table to answer the following questions.

1. Calculate the allele and genotype frequencies of *Ldh* in this newly founded population.
2. Is the population in Hardy–Weinberg equilibrium? If not, which genotypes are over- or underrepresented? Given the population's history, what is a likely explanation of your answer?
3. Under Hardy–Weinberg assumptions, what allele and genotype frequencies do you predict for the next generation?
4. Imagine that you are able to continue studying this population and determine the next generation's actual allele and genotype frequencies. What are some of the principal reasons you might expect the observed allele and genotype frequencies to differ from the Hardy–Weinberg expectations you calculated in question 3?

INDIVIDUAL NUMBER	SEX	INDIVIDUAL GENOTYPE FOR *Ldh*
1	Male	*Aa*
2	Male	*AA*
3	Female	*AA*
4	Male	*aa*
5	Female	*aa*
6	Female	*AA*
7	Male	*aa*
8	Male	*aa*
9	Female	*Aa*
10	Male	*AA*

important for the study of evolution? There are two reasons. First, the model is useful for predicting the approximate genotype frequencies of a population from its allele frequencies. Second—and crucially—the model allows biologists to evaluate which processes are acting on the evolution of a particular population. The specific patterns of deviation from Hardy–Weinberg equilibrium can help us identify the various processes of evolutionary change.

CHECKpoint CONCEPT 15.3

- ✓ Why is the concept of Hardy–Weinberg equilibrium important even though the assumptions on which it is based are never completely met in nature?
- ✓ Although the stringent assumptions of Hardy–Weinberg equilibrium are never met completely in real populations, the genotype frequencies of many populations do not deviate significantly from Hardy–Weinberg expectations. Can you explain why?
- ✓ Suppose you examine a population of toads breeding in a single pond and find that heterozygous genotypes at several different loci are present at significantly lower frequencies than predicted by Hardy–Weinberg equilibrium. What are some possible explanations?

Our discussion so far has focused on changes in allele frequencies at a single gene locus. Genes do not exist in isolation, however, but interact with one another (and with the environment) to produce an organism's phenotype. What effects can these interactions have on selection?

CONCEPT 15.4 Selection Can Be Stabilizing, Directional, or Disruptive

Until now, we have only discussed traits influenced by alleles at a single locus. Such traits are often distinguished by discrete qualities (black versus white, or smooth versus wrinkled) and so are called **qualitative traits**. Many traits, however, are influenced by alleles at more than one locus. Such traits are likely to show continuous quantitative variation rather than discrete qualitative variation, and so are known as **quantitative traits**. For example, the distribution of body sizes of individuals in a population, a trait that is influenced by genes at many loci as well as by the environment, is likely to resemble a continuous bell-shaped curve.

Natural selection can act on characters with quantitative variation in any one of several different ways, producing quite different results (FIGURE 15.13):

- **Stabilizing selection** preserves the average characteristics of a population by favoring average individuals.
- **Directional selection** changes the characteristics of a population by favoring individuals that vary in one direction from the mean of the population.
- **Disruptive selection** changes the characteristics of a population by favoring individuals that vary in both directions from the mean of the population.

Stabilizing selection reduces variation in populations

If the smallest and largest individuals in a population contribute fewer offspring to the next generation than do individuals

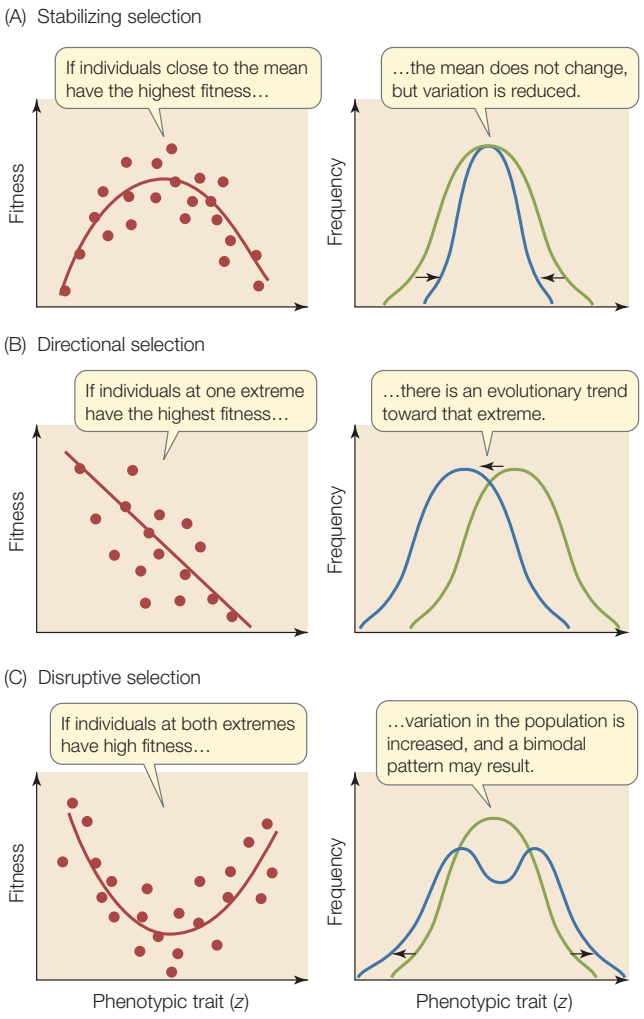


FIGURE 15.13 Natural Selection Can Operate in Several Ways The graphs in the left-hand column show the fitness of individuals with different phenotypes of the same trait. The graphs on the right show the distribution of the phenotypes in the population before (green) and after (blue) the influence of selection.

closer to the average size, then stabilizing selection is operating on size (see Figure 15.13A). Stabilizing selection reduces variation in populations, but it does not change the mean. Natural selection frequently acts in this way, countering increases in variation brought about by sexual recombination, mutation, or gene flow. Rates of phenotypic change in many species are slow because natural selection is often stabilizing. Stabilizing selection operates, for example, on human birth weight. Babies who are lighter or heavier at birth than the population mean die at higher rates than babies whose weights are close to the mean (**FIGURE 15.14**). In discussions of specific genes, stabilizing selection is often called **purifying selection** because there is selection against any deleterious mutations to the usual gene sequence.

Directional selection favors one extreme

Directional selection is operating when individuals at one extreme of a character distribution contribute more offspring to the next generation than other individuals do, shifting the average value of that character in the population toward that extreme. In the case of a single gene locus, directional selection may result in favoring a particular genetic variant—referred to as **positive selection** for that variant. By favoring one phenotype over another, directional selection results in an increase of the frequencies of alleles that produce the favored phenotype (as with the surface proteins of influenza discussed in the opening of this chapter).

If directional selection operates over many generations, an evolutionary trend is seen in the population (see Figure 15.13B). Evolutionary trends often continue for many generations, but they can be reversed if the environment changes and different phenotypes are favored, or halted when an optimal phenotype is reached or trade-offs between different adaptational advantages oppose further change. The character then undergoes stabilizing selection.

The long horns of Texas Longhorn cattle (**FIGURE 15.15**) are an example of a trait that has evolved through directional selection. Texas Longhorns are descendants of cattle brought to the New World by Christopher Columbus, who picked up a few cattle in the Canary Islands and brought them to the island of Hispaniola in 1493. The cattle multiplied, and their descendants were taken to the mainland of Mexico. Spaniards exploring what would become Texas and the southwestern United States brought these cattle with them, some of which escaped and formed feral herds. Populations of feral cattle increased greatly over the next few hundred years, but there was heavy predation from bears, mountain lions, and wolves, especially on the young calves. Cows with longer horns were more successful in protecting their calves against attacks, and

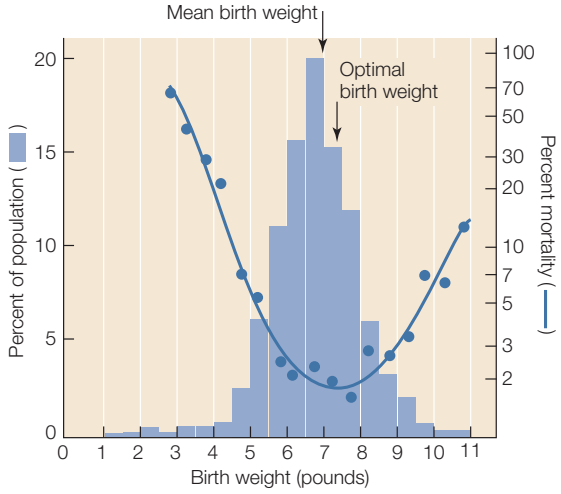


FIGURE 15.14 Human Birth Weight Is Influenced by Stabilizing Selection Babies that weigh more or less than average are more likely to die soon after birth than babies with weights close to the population mean.

FIGURE 15.15 Long Horns Are the Result of Directional Selection Long horns were advantageous for defending young calves from attacks by predators, so horn length increased in feral herds of Spanish cattle in the American Southwest between the early 1500s and the 1860s. The result was the familiar Texas Longhorn breed. This evolutionary trend has been maintained in modern times by ranchers practicing artificial selection.

over a few hundred years the average horn length in the feral herds increased considerably. In addition, the cattle evolved resistance to endemic diseases of the Southwest, as well as higher fecundity and longevity. Texas Longhorns often live and produce calves well into their twenties—about twice as long as many breeds of cattle that have been artificially selected by humans for traits such as high fat content or high milk production (which are examples of artificial directional selection).

Disruptive selection favors extremes over the mean

When disruptive selection operates, individuals at opposite extremes of a character distribution contribute more offspring to the next generation than do individuals close to the mean, which increases variation in the population (see Figure 15.13C).

The strikingly bimodal (two-peaked) distribution of bill sizes in the black-bellied seedcracker (*Pyrenestes ostrinus*), a West African finch (**FIGURE 15.16**), illustrates how disruptive selection can influence populations in nature. The seeds of two types of sedges (marsh plants) are the most abundant food source for these finches during part of the year. Birds with large bills can readily crack the hard seeds of the sedge *Scleria verrucosa*. Birds with small bills can crack *S. verrucosa* seeds only with difficulty; however, they feed more efficiently on the soft seeds of *S. goossensii* than do birds with larger bills. Young finches whose bills deviate markedly from the two predominant bill sizes do not survive as well as finches whose bills are close to one of the two sizes represented by the distribution peaks. Because there are few abundant food sources in the finches' environment, and because the seeds of the two sedges do not overlap in hardness, birds with intermediate-sized bills are less efficient in using either one of the species' principal food sources. Disruptive selection therefore maintains a bimodal bill size distribution.

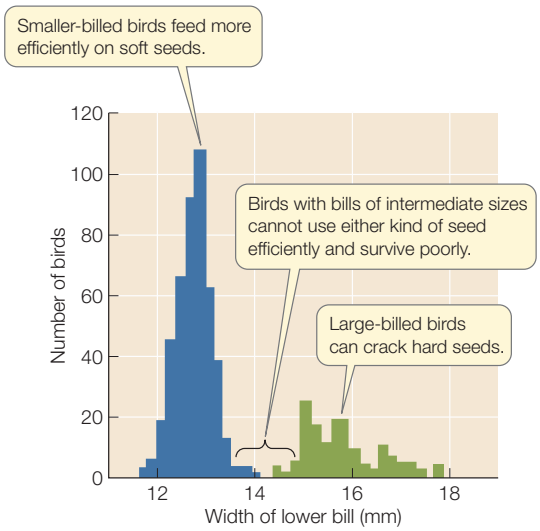


FIGURE 15.16 Disruptive Selection Results in a Bimodal Character Distribution The bimodal distribution of bill sizes in the black-bellied seedcracker of West Africa is a result of disruptive selection, which favors individuals with larger and smaller bill sizes over individuals with intermediate-sized bills.

CHECKpoint CONCEPT 15.4

- ✓ What are the different expected outcomes of stabilizing, directional, and disruptive selection?
- ✓ Why would you expect selection on human birth weight to be stabilizing rather than directional?
- ✓ Can you think of examples of extreme phenotypes in animal or plant populations that could be explained by directional selection?

Our discussion so far has largely focused on the evolution of phenotypes (what organisms look like and how they behave). We will now consider the specific mechanistic processes that operate at the level of genes and genomes.

CONCEPT 15.5 Genomes Reveal Both Neutral and Selective Processes of Evolution

Most natural populations harbor far more genetic variation than we would expect to find if genetic variation were influenced by natural selection alone. This discovery, combined with the knowledge that many mutations do not change molecular function, provided a major stimulus to the development of the field of molecular evolution.

To discuss the evolution of genes, we need to consider the specific types of mutations that are possible. A nucleotide substitution is a change in a single nucleotide in a DNA sequence (a type of point mutation). Many nucleotide substitutions have no effect on phenotype, even if the change occurs in a gene that encodes a protein, because most amino acids are specified by

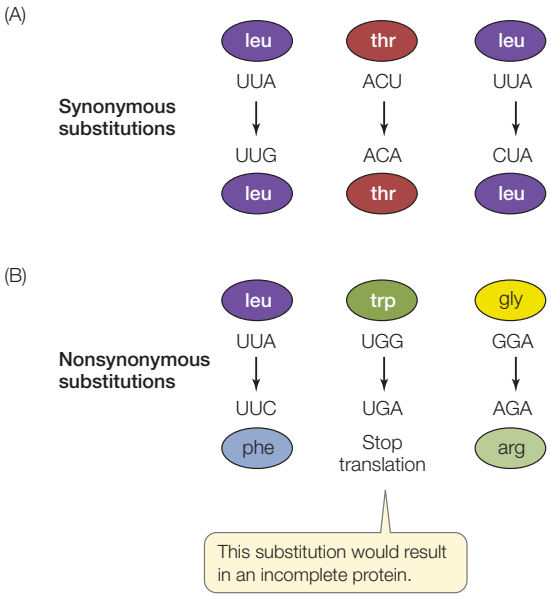


FIGURE 15.17 When One Nucleotide Changes (A) Synonymous substitutions do not change the amino acid specified and do not affect protein function. Such substitutions are less likely to be subject to natural selection, although they contribute greatly to the buildup of neutral genetic variation in a population. (B) Nonsynonymous substitutions do change the amino acid sequence and are likely to have an effect (often deleterious, but sometimes beneficial) on protein function. Such nucleotide substitutions are targets for natural selection.

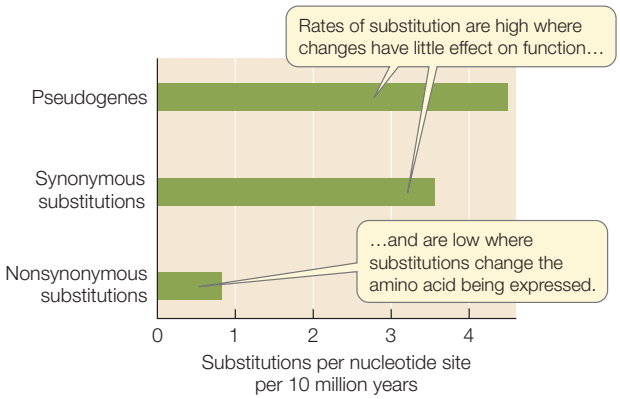


FIGURE 15.18 Rates of Substitution Differ Rates of nonsynonymous substitution are typically much lower than rates of synonymous substitution, and much lower than substitution rates in pseudogenes. This pattern reflects stronger stabilizing selection in functional genes than in pseudogenes.

more than one codon. A substitution that does not change the encoded amino acid is known as a **synonymous substitution** (also called a **silent substitution**; FIGURE 15.17A). Synonymous substitutions do not affect the functioning of a protein (although they may have other effects, such as changes in mRNA stability or translation rates) and are therefore less likely to be influenced by natural selection.

A nucleotide substitution that *does* change the amino acid sequence encoded by a gene is known as a **nonsynonymous substitution** (also called a **missense substitution**; FIGURE 15.17B). In general, nonsynonymous substitutions are likely to be deleterious to the organism. But not every amino acid replacement alters a protein's shape and charge (and hence its functional properties). Therefore some nonsynonymous substitutions are selectively neutral, or nearly so. A third possibility is that a nonsynonymous substitution alters a protein in a way that confers an advantage to the organism, and is therefore favored by natural selection.

LINK

The genetic code determines the amino acid that is encoded by each codon; see **Figure 10.11**

The rate of synonymous substitutions in most protein-coding genes is much higher than the rate of nonsynonymous substitutions. In other words, *substitution rates are highest at nucleotide positions that do not change the amino acid being expressed* (FIGURE 15.18). The rate of substitution is even higher in **pseudogenes**, which are copies of genes that are no longer functional.

Insertions, deletions, and rearrangements of DNA sequences are all mutations that may affect a larger portion of the gene or genome than do point mutations (see Concept 9.3). Insertions and deletions of nucleotides in a protein-coding sequence interrupt its reading frame, unless they occur in multiples of three nucleotides (the length of one codon). Rearrangements may merely change the order of whole genes along chromosomes, or they may rearrange functional domains among individual genes.

When biologists began to examine the details of genetic variation of populations, they soon discovered many gene variants that had little or no effect on function. This gave rise to new ideas about how these neutral variants arise and spread in populations.

Much of molecular evolution is neutral

Motoo Kimura proposed the neutral theory in 1968. He suggested that, at the molecular level, the majority of variants found in most populations are selectively neutral. That is, most gene variants confer neither an advantage nor a disadvantage on their bearers. Therefore these neutral variants must accumulate through genetic drift rather than through positive selection.

We saw in Concept 15.2 that genetic drift of existing gene variants tends to be greatest in small populations. However, the rate of fixation of new neutral mutations by genetic drift is independent of population size. To see why this is so, consider a population of size N and a neutral mutation rate of μ (mu) per gamete per generation at a particular locus. The number of new mutations would be, on average, $\mu \times 2N$, because $2N$ gene copies are available to mutate in a population of diploid organisms. The probability that a given mutation will be fixed

by drift alone is its frequency, which equals $1/(2N)$ for a newly arisen mutation. We can multiply these two terms to get the rate of fixation of neutral mutations (m) in a given population of N individuals:

$$m = 2N\mu\frac{1}{2N}$$

Therefore the rate of fixation of neutral mutations depends only on the neutral mutation rate μ and is independent of population size. Any given mutation is more likely to appear in a large population than in a small one, but any mutation that does appear is more likely to become fixed in a small population. These two influences of population size cancel each other out. Therefore the rate of fixation of neutral mutations is equal to the mutation rate (i.e., $m = \mu$).

As long as the underlying mutation rate is constant, genes and proteins evolving in different populations should diverge from one another in neutral changes at a constant rate. The rate of evolution of particular genes and proteins is indeed often relatively constant over time, and therefore can be used as a "molecular clock" to calculate evolutionary divergence times between species (see Concept 16.3).

Neutral theory does not imply that most mutations have no effect on the individual organism, even though much of the genetic variation present in a population is the result of neutral evolution. Many mutations are never observed in populations because they are lethal or strongly detrimental, and the individuals that carry them are quickly removed from the population through natural selection. Similarly, because mutations that confer a selective advantage tend to be quickly fixed in populations, they also do not result in significant variation at the population level. Nonetheless, if we compare homologous proteins from different populations or species, some amino acid positions will remain constant under purifying selection, others will vary through neutral genetic drift, and still others will differ among species as a result of positive selection for change. How can these evolutionary processes be distinguished?

Go to MEDIA CLIP 15.2
The Ubiquitous Protein
PoL2e.com/mc15.2

Positive and purifying selection can be detected in the genome

As we have just seen, substitutions in a protein-coding gene can be either synonymous or nonsynonymous, depending on whether they change the resulting amino acid sequence of the protein. The relative rates of synonymous and nonsynonymous substitutions are expected to differ in regions of genes that are evolving neutrally, or evolving under positive selection for change, or staying unchanged under purifying selection.

- If a given amino acid in a protein can be one of many alternatives (without changing the protein's function), then an amino acid replacement is *neutral* with respect to the fitness of an organism. In this case, the rates of synonymous and nonsynonymous substitutions in the corresponding DNA sequences are expected to be very similar, so the ratio of the two rates should be close to 1.
- If a given amino acid position is under *positive selection* for change, the observed rate of nonsynonymous substitutions is expected to exceed the rate of synonymous substitutions in the corresponding DNA sequences.
- If a given amino acid position is under *purifying selection*, then the observed rate of synonymous substitutions is expected to be much higher than the rate of nonsynonymous substitutions in the corresponding DNA sequences.

The evolution of lysozyme illustrates how and why particular codons in a gene sequence might be under different modes of selection. The enzyme lysozyme is found in almost all animals. It is produced in the tears, saliva, and milk of mammals and in the albumen (whites) of bird eggs. Lysozyme digests the cell walls of bacteria, rupturing and killing them. As a result, it plays an important role as a first line of defense against invading bacteria. Most animals defend themselves against bacteria by digesting them, which is probably why most animals have lysozyme. Some animals also use lysozyme in the digestion of food.

Among mammals, a mode of digestion called foregut fermentation has evolved twice. In mammals with this mode of digestion, the foregut—consisting of part of the esophagus and/or stomach—has been converted into a chamber in which bacteria break down ingested plant matter by fermentation. Foregut fermenters can obtain nutrients from the otherwise indigestible cellulose that makes up a large proportion of plant tissue. Foregut fermentation evolved independently in ruminants (a group of hoofed mammals that includes cattle) and in certain leaf-eating monkeys, such as langurs. We know that these evolutionary events were independent because both langurs and ruminants have close relatives that are not foregut fermenters.

In both mammalian foregut-fermenting lineages, lysozyme has been modified to play a new, nondefensive role. The modified lysozyme enzyme ruptures some of the bacteria that live in the foregut, releasing nutrients metabolized by the bacteria, which the mammal then absorbs. How many changes in the lysozyme molecule were needed to allow it to perform this function amid the digestive enzymes and acidic conditions of the mammalian foregut? To answer this question, biologists compared the lysozyme-coding sequences in foregut fermenters with those in several of their nonfermenting relatives. They determined which amino acids differed and which were shared among the species (**FIGURE 15.19A**), as well as the rates of synonymous and nonsynonymous substitution in lysozyme genes across the evolutionary history of the sampled species.

For many of the amino acid positions of lysozyme, the rate of synonymous substitution in the corresponding gene sequence was much higher than the rate of nonsynonymous substitution. This observation indicates that many of the amino acids that make up lysozyme are evolving under purifying selection. In other words, there is selection against change in the lysozyme protein at these positions, and the encoded amino acids must therefore be critical for lysozyme function. At other

(A) *Semnopithecus* sp.

Bos taurus

(B) *Opisthocomus hoazin*

The lysozymes of langurs and cattle are convergent for 5 amino acid residues, indicative of the independent evolution of foregut fermentation in these two species.

	Langur	Baboon	Human	Rat	Cattle	Horse
Langur		14	18	38	32	65
Baboon	0		14	33	39	65
Human	0	1		37	41	64
Rat	0	0	0		55	64
Cattle	[5]	0	0	0		71
Horse	0	0	0	0	1	

FIGURE 15.19 Convergent Molecular Evolution of Lysozyme
(A) The numbers of amino acid differences in the lysozymes of several pairs of mammals are shown above the diagonal line; the numbers of similarities that arose from convergence between species are shown below the diagonal. The two foregut-fermenting species (cattle and langur) share five convergent amino acid replacements related to this digestive adaptation. (B) The hoatzin—the only known foregut-fermenting bird species—has been evolving independently from mammals for hundreds of millions of years but has independently evolved modifications to lysozyme similar to those found in cattle and langurs.

positions, several different amino acids function equally well, and the corresponding codons have similar rates of synonymous and nonsynonymous substitution.

The most striking finding was that amino acid replacements in lysozyme happened at a much higher rate in the lineage leading to langurs than in any other primate lineage. The high rate of nonsynonymous substitution in the langur lysozyme gene shows that lysozyme went through a period of rapid change in adapting to the stomachs of langurs. Moreover, the lysozymes of langurs and cattle share five convergent amino

APPLY THE CONCEPT

Genomes reveal both neutral and selective processes of evolution

Analysis of synonymous and nonsynonymous substitutions in protein-coding genes can be used to detect neutral evolution, positive selection, and purifying selection. An investigator compared many gene sequences that encode the protein hemagglutinin (a surface protein of influenza virus) sampled over time, and collected the data at right.[a] Use the table to answer the following questions.

1. Which codon positions are likely evolving under positive selection? Why?
2. Which codon positions are likely evolving under purifying selection? Why?

(Hint: To calculate rates of each substitution type, you will need to consider the number of synonymous and nonsynonymous substitutions *relative to the number of possible substitutions of each type*. There are approximately three times as many possible nonsynonomous substitutions as there are synonymous substitutions.)

CODON POSITION	NUMBER OF SYNONYMOUS SUBSTITUTIONS IN CODON	NUMBER OF NONSYNONYMOUS SUBSTITUTIONS IN CODON
12	0	7
15	1	9
61	0	12
80	7	0
137	12	1
156	24	2
165	3	4
226	38	3

[a]R. M. Bush et al. 1999. *Molecular Biology and Evolution* 16: 1457–1465.

acid replacements, all of which lie on the surface of the lysozyme molecule, well away from the enzyme's active site. Several of these shared replacements are changes from arginine to lysine, which make the protein more resistant to degradation by the stomach enzyme pepsin. By understanding the functional significance of amino acid replacements, biologists can explain the observed changes in amino acid sequences in terms of changes in the functioning of the protein.

A large body of fossil, morphological, and molecular evidence shows that langurs and cattle do not share a recent common ancestor. However, langur and ruminant lysozymes share several amino acids that neither mammal shares with the lysozymes of its own closer relatives. The lysozymes of these two mammals have converged on some of the same amino acids despite their very different ancestry. The amino acids they share give these lysozymes the ability to lyse the bacteria that ferment plant material in the foregut.

The hoatzin, an unusual leaf-eating South American bird (**FIGURE 15.19B**) and the only known avian foregut fermenter, offers another remarkable example of the convergent evolution of lysozyme. Many birds have an enlarged esophageal chamber called a crop. The crop of the hoatzin contains lysozyme and bacteria and acts as a fermentation chamber. Many of the amino acid replacements that occurred in the adaptation of hoatzin lysozyme are identical to those that evolved in ruminants and langurs. Thus even though the hoatzin and foregut-fermenting mammals have not shared a common ancestor in hundreds of millions of years, similar adaptations have evolved in their lysozyme enzymes, enabling both groups to recover nutrients from fermenting bacteria.

Heterozygote advantage maintains polymorphic loci

In many cases, different alleles of a particular gene are advantageous under different environmental conditions. Most organisms, however, experience a wide diversity of environments. A night is dramatically different from the preceding day. A cold, cloudy day differs from a clear, hot one. Day length and temperature change seasonally. For many genes, a single allele is unlikely to perform well under all these conditions. In such situations, a heterozygous individual (with two different alleles) is likely to outperform individuals that are homozygous for either one of the alleles.

Colias butterflies of the Rocky Mountains live in environments where dawn temperatures often are too cold, and afternoon temperatures too hot, for the butterflies to fly. Populations of these butterflies are polymorphic for the gene that encodes

INVESTIGATION

FIGURE 15.20 A Heterozygote Mating Advantage Among butterflies of the genus *Colias*, males that are heterozygous for two alleles of the PGI enzyme can fly farther under a broader range of temperatures than males that are homozygous for either allele. Does this ability give heterozygous males a mating advantage?[a]

HYPOTHESIS

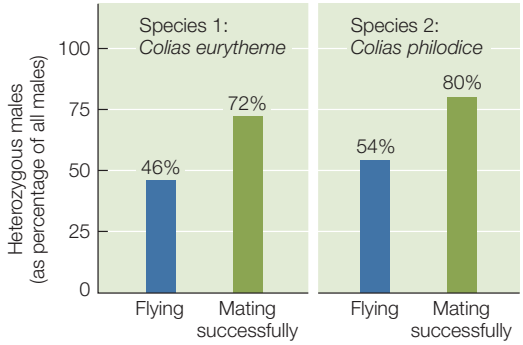
Heterozygous male *Colias* will have proportionally greater mating success than homozygous males.

METHOD

1. For each of two *Colias* species, capture butterflies in the field. In the laboratory, determine their genotypes and allow them to mate.
2. Determine the genotypes of the offspring, thus revealing paternity and mating success of the males.

RESULTS

For both species, the proportion of heterozygous males that mated successfully was higher than the proportion of all males seeking females ("flying").



CONCLUSION

Heterozygous *Colias* males have a mating advantage over homozygous males.

ANALYZE THE DATA

Analyze these sampling data collected during the experiment (only one of several samples is shown for each species).

	All viable males*		Mating Males	
Species	Heterozygous/ total	Percent heterozygous	Heterozygous/ total	Percent heterozygous
C. philodice	32/74	43.2	31/50	62.0
C. eurytheme	44/92	47.8	45/59	76.3

*"Viable males" are all males captured flying with females (hence with the potential to mate).

A. Under the assumption that the proportions of each genotype (heterozygotes and homozygotes) of mating males are the same as the proportions seen among all viable males, calculate the number of mating males expected to be heterozygous and the number expected to be homozygous.

B. Use a chi-square test (see Appendix B) to evaluate the significance of the difference in your expected numbers in (A) and the observed percentages of heterozygous mating males. The critical value ($P = 0.05$) of the chi-square distribution with one degree of freedom is 3.841. Are the observed and expected numbers of heterozygotes and homozygotes among mating males significantly different in these samples?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]W. B. Watt et al. 1985. *Genetics* 109: 157–175.

phosphoglucose isomerase (PGI), an enzyme that influences how well an individual flies at different temperatures. Butterflies with certain PGI genotypes can fly better during the cold hours of early morning; those with other genotypes perform better during midday heat. The optimal body temperature for flight is 35°C–39°C, but some butterflies can fly with body temperatures as low as 29°C or as high as 40°C. Heat-tolerant genotypes are favored during spells of unusually hot weather; during spells of unusually cool weather, cold-tolerant genotypes are favored.

Heterozygous *Colias* butterflies can fly over a greater temperature range than homozygous individuals because they produce two different forms of PGI. This greater range of activity should give them an advantage in foraging and finding mates. A test of this prediction did find a mating advantage in heterozygous males, and further found that this mating advantage maintains the polymorphism in the population (**FIGURE 15.20**). The heterozygous condition can never become fixed in the population, however, because the offspring of two heterozygotes will always include both classes of homozygotes in addition to heterozygotes.

Genome size and organization also evolve

We know that genome size varies tremendously among organisms. Across broad taxonomic categories, there is some correlation between genome size and organismal complexity. The genome of the tiny bacterium *Mycoplasma genitalium* has only 470 genes. *Rickettsia prowazekii*, the bacterium that causes typhus, has 634 genes. *Homo sapiens*, by contrast, has about 21,000 protein-coding genes. **FIGURE 15.21** shows the number of genes from a sample of organisms whose genomes have been fully sequenced, arranged by their evolutionary relationships. As this figure reveals, however, a larger genome does not always indicate greater complexity (compare rice with the other plants, for example). It is not surprising that more complex genetic instructions are needed for building and maintaining a large, multicellular organism than a small, single-celled bacterium. What is surprising is that some organisms, such as lungfishes, some salamanders, and lilies, have about 40 times as much DNA as humans do (**FIGURE 15.22**). Structurally, a lungfish or a lily is not 40 times more complex than a human. So why does genome size vary so much?

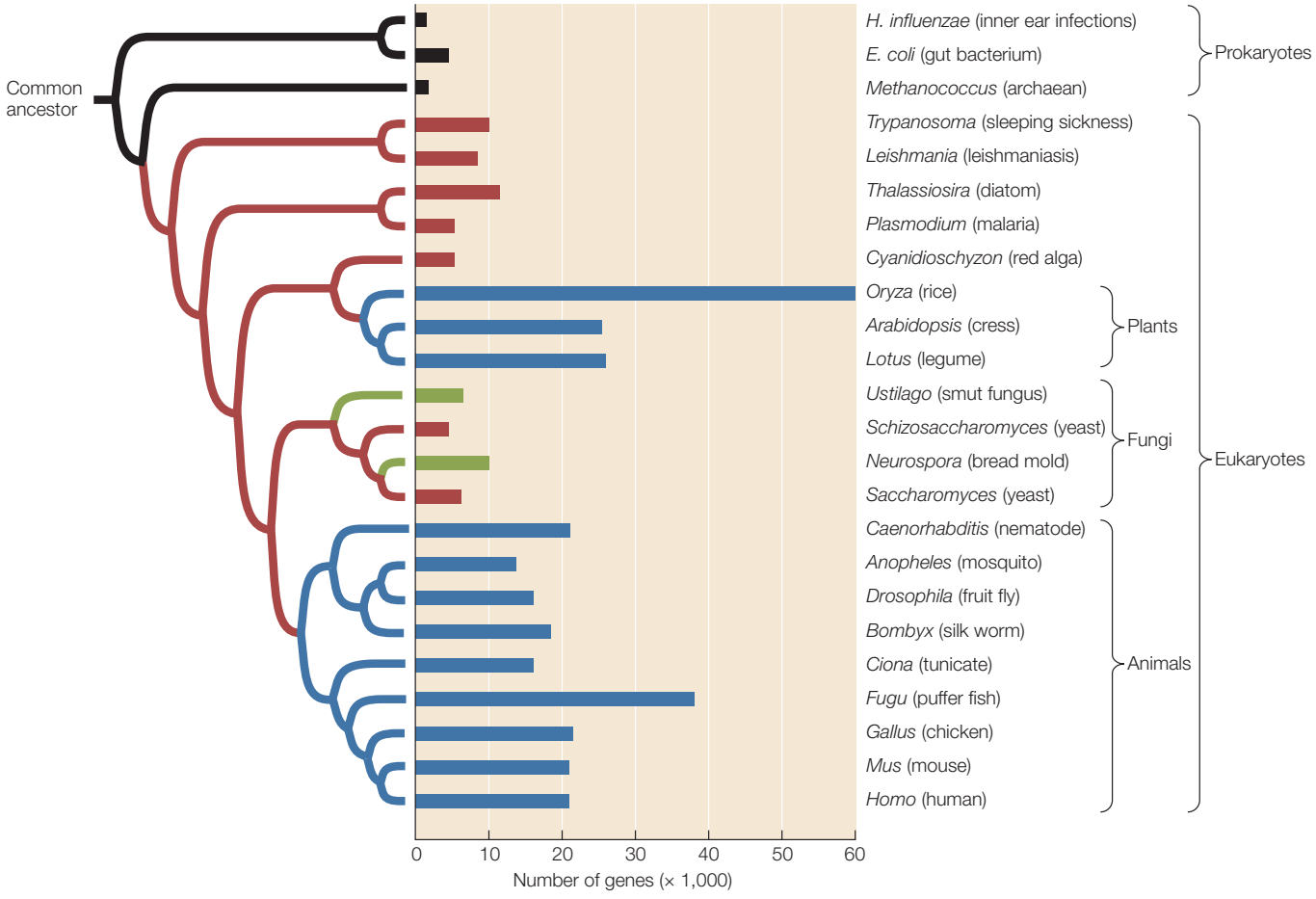


FIGURE 15.21 Evolution of Gene Number This figure shows the number of genes from a sample of organisms whose genomes have been fully sequenced, arranged by their evolutionary relationships. Bacteria and archaea (black branches) typically have fewer genes than most eukaryotes. Among eukaryotes, multicellular organisms with tissue organization (plants and animals; blue branches) have more genes than single-celled organisms (red branches) or multicellular organisms that lack pronounced tissue organization (green branches).

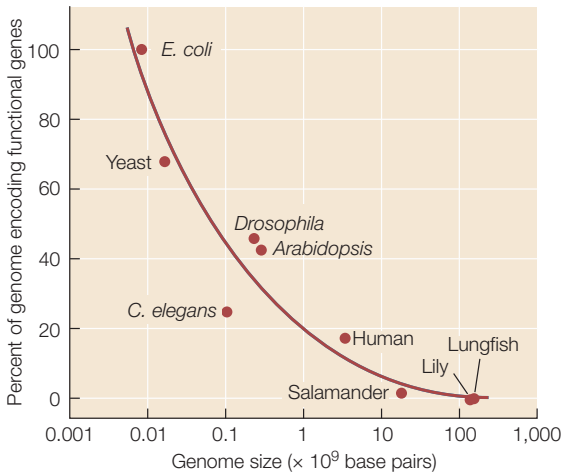


FIGURE 15.22 A Large Proportion of DNA Is Noncoding Most of the DNA of bacteria and yeasts encodes RNAs or proteins, but a large percentage of the DNA of multicellular species is noncoding.

Differences in genome size are not so great if we take into account only the portion of DNA that actually encodes proteins. The organisms with the largest total amounts of nuclear DNA (some ferns and flowering plants) have 80,000 times as much DNA as do the bacteria with the smallest genomes, but no species has more than about 100 times as many protein-coding genes as a bacterium. Therefore much of the variation in genome size lies not in the number of functional genes, but in the amount of noncoding DNA (see Figure 15.22).

Why do the cells of most eukaryotic organisms have so much noncoding DNA? Does this noncoding DNA have a function? Although some of this DNA does not encode proteins, it can alter the expression of the genes surrounding it. The degree or timing of gene expression can vary dramatically depending on the gene's position relative to noncoding sequences that regulate gene expression. Other regions of noncoding DNA consist of pseudogenes (regions that have evolved from functional genes, even though they have no function at present). Pseudogenes are often carried in the genome because the cost of doing so is very small. Occasionally, these pseudogenes become the raw material for the evolution of new genes with novel functions. Other noncoding sequences function in maintaining chromosomal structure. Still others consist of parasitic transposable elements that spread through populations because they reproduce faster than the host genome.

Another hypothesis is that the proportion of noncoding DNA is related primarily to population size. Noncoding sequences that are only slightly deleterious to the organism are likely to be purged by selection most efficiently in species with large population sizes. In species with small populations, the effects of genetic drift can overwhelm selection against noncoding sequences that have small deleterious consequences. Therefore selection against the accumulation of noncoding sequences is most effective in species with large populations, so such species (such as bacteria or yeasts) have relatively little noncoding DNA compared with species with small populations (see Figure 15.22).

CHECKpoint CONCEPT 15.5

- ✓ How can the ratio of synonymous to nonsynonymous substitutions be used to determine whether a particular gene is evolving neutrally, under positive selection, or under stabilizing selection?
- ✓ Why is the rate of fixation of neutral mutations independent of population size?
- ✓ Why do heterozygous individuals sometimes have an advantage over homozygous individuals?
- ✓ Why can a mutation that results in the replacement of one amino acid by another be a neutral event in some cases and in other cases be detrimental or beneficial? (Hint: Review the information about amino acids in Table 3.2 and the details of protein structure in Concept 3.2.)
- ✓ Postulate and contrast two hypotheses for the wide diversity of genome sizes among different organisms.

Most of our discussion so far has centered on changes in existing genes and phenotypes. Next we'll consider how new genes with novel functions arise in populations in the first place.

CONCEPT 15.6 Recombination, Lateral Gene Transfer, and Gene Duplication Can Result in New Features

Several evolutionary processes can result in the acquisition of major new characteristics in populations. Each of these processes results in larger and more rapid evolutionary changes than do single point mutations.

Sexual recombination amplifies the number of possible genotypes

In asexually reproducing organisms, each new individual is genetically identical to its parent unless there has been a mutation. When organisms reproduce sexually, however, offspring differ from their parents because of crossing over and independent assortment of chromosomes during meiosis, as well as the combination of genetic material from two different gametes, as described in Concept 7.4. Sexual recombination generates an endless variety of genotype combinations that increase the evolutionary potential of populations—a long-term advantage of sex. Although some species may reproduce asexually most of the time, most asexual species have some means of achieving genetic recombination.

The evolution of meiosis and sexual recombination was a crucial event in the history of life. Exactly how these processes arose is puzzling, however, because in the short term, sex has at least three striking disadvantages:

- Recombination breaks up adaptive combinations of genes.
- Sex reduces the rate at which females pass genes on to their offspring.
- Dividing offspring into separate genders greatly reduces the overall reproductive rate.

To see why this last disadvantage exists, consider an asexual female that produces the same number of offspring as a sexual female. Assume that both females produce two offspring, but that half of the sexual female's offspring are males. In the next (F_1) generation, then, each of the two asexual F_1 females will produce two more offspring—but there is only one sexual F_1 female to produce offspring. Thus the effective reproductive rate of the asexual lineage is twice that of the sexual lineage. The evolutionary problem is to identify the advantages of sex that can overcome such short-term disadvantages.

A number of hypotheses have been proposed to explain the existence of sex, none of which are mutually exclusive. One is that sexual recombination facilitates repair of damaged DNA, because breaks and other errors in DNA on one chromosome can be repaired by copying the intact sequence from the homologous chromosome.

Another advantage of sexual reproduction is that it permits the elimination of deleterious mutations through recombination followed by selection. As Concept 9.2 described, DNA replication is not perfect, and many replication errors result in lower fitness. Meiotic recombination distributes these deleterious mutations unequally among gametes. Sexual reproduction then produces some individuals with more deleterious mutations and some with fewer. The individuals with fewer deleterious mutations are more likely to survive. Therefore sexual reproduction allows natural selection to eliminate particular deleterious mutations from the population over time.

In asexual reproduction, deleterious mutations can be eliminated only by the death of the lineage or by a rare back mutation (that is, when a subsequent mutation returns a mutated sequence to its original DNA sequence). Hermann J. Muller noted that deleterious mutations in a non-recombining genome accumulate—"ratchet up"—at each replication. Mutations occur and are passed on each time a genome replicates, and these mutations accumulate with each subsequent generation. This accumulation of deleterious mutations in lineages that lack genetic recombination is known as **Muller's ratchet**.

Another explanation for the existence of sex is that the great variety of genetic combinations created in each generation can itself be advantageous. For example, genetic variation can be a defense against pathogens and parasites. Most pathogens and parasites have much shorter life cycles than their hosts and can rapidly evolve counter-adaptations to host defenses. Sexual recombination might give the host's defenses a chance to keep up.

Sexual recombination does not directly influence the frequencies of alleles. Rather, *it generates new combinations of alleles on which natural selection can act*. It expands variation in quantitative characters by creating new genotypes. That is why artificial selection for bristle number in *Drosophila* (see Figure 15.7) resulted in flies with either more or fewer bristles than were present in the flies in the initial population.

Lateral gene transfer can result in the gain of new functions

The tree of life is usually visualized as a branching diagram, with each lineage diverging into two (or more) lineages over time, from one common ancestor to the millions of species that are alive today. Ancestral lineages divide into descendant lineages, and it is those speciation events that the tree of life captures. However, there are also processes that result in **lateral gene transfer**—the horizontal movement of individual genes, organelles, or fragments of genomes from one lineage to another. Some species may pick up fragments of DNA directly from the environment. A virus may pick up some genes from one host and transfer them to a new host when the virus becomes integrated into the new host's genome. Hybridization between species also results in the lateral transfer of large numbers of genes.

Lateral gene transfer can be highly advantageous to the species that incorporates novel genes from a distant relative. Genes that confer antibiotic resistance, for example, are commonly transferred among different species of bacteria. Lateral gene transfer is another way, in addition to mutation and recombination, that species can increase their genetic variation.

The degree to which lateral gene transfer events occur in various parts of the tree of life is a matter of considerable current investigation and debate. Lateral gene transfer appears to be relatively uncommon among most eukaryote lineages, although the two major endosymbioses that gave rise to mitochondria and chloroplasts involved lateral transfers of entire bacterial genomes to the eukaryote lineage. Some groups of eukaryotes, most notably some plants, are subject to relatively high levels of hybridization among closely related species. Hybridization leads to the exchange of many genes among recently separated lineages of plants. The greatest degree of lateral transfer, however, occurs among bacteria. Many genes have been transferred repeatedly among bacteria, to the point that relationships and boundaries among species of bacteria are sometimes hard to decipher.

Many new functions arise following gene duplication

Gene duplication is yet another way that genomes can acquire new functions. When a gene is duplicated, one copy of that gene is potentially freed from having to perform its original function. The identical copies of a duplicated gene can have any one of four different fates:

- Both copies of the gene may retain their original function (which can result in a change in the amount of gene product that is produced by the organism).
- Both copies of the gene may retain the ability to produce the original gene product, but the expression of the genes may diverge in different tissues or at different times in development.
- One copy of the gene may be incapacitated by the accumulation of deleterious mutations and become a functionless pseudogene.
- One copy of the gene may retain its original function while the second copy changes and evolves a new function.

How often do gene duplications arise, and which of these four outcomes is most likely? Investigators have found that rates of gene duplication are fast enough for a yeast or *Drosophila*

population to acquire several hundred duplicate genes over the course of a million years. They have also found that most of the duplicated genes that are still present in these organisms are very young. Many duplicated genes are lost from a genome within 10 million years—an eyeblink on an evolutionary time scale.

Many gene duplications affect only one or a few genes at a time, but in some cases entire genomes may be duplicated. When all the genes are duplicated, there are massive opportunities for new functions to evolve. That is exactly what seems to have happened during the course of vertebrate evolution. The genomes of the jawed vertebrates have four diploid sets of many major genes, which leads biologists to conclude that two genome-wide duplication events occurred in the ancestor of these species. These duplications allowed considerable specialization of individual vertebrate genes, many of which are now highly tissue-specific in their expression.

LINK

See **Concept 14.4** for a discussion of the role of duplicated Hox genes in vertebrate evolution

Several successive rounds of duplication and sequence evolution may result in a **gene family**, a group of homologous genes with related functions, often arrayed in tandem along a chromosome. An example of a group of genes related by gene duplication is the globin gene family (**FIGURE 15.23**). Comparisons of the amino acid sequences among globins strongly suggest that this family of proteins arose via gene duplications.

Hemoglobin is a tetramer (four-subunit molecule) consisting of two α-globin and two β-globin polypeptide chains. It carries oxygen in the blood. Myoglobin, a monomer, is the primary O_2 storage protein in muscle. Myoglobin's affinity for O_2 is much higher than that of hemoglobin, but hemoglobin has evolved to be more diversified in its role. Hemoglobin binds O_2 in the lungs or gills, where the O_2 concentration is relatively high, transports it to deep body tissues, where the O_2 concentration is low, and releases it in those tissues. With its more complex tetrameric structure, hemoglobin is able to carry four molecules of O_2, as well as hydrogen ions and carbon dioxide, in the blood. Hemoglobin and myoglobin are estimated to have arisen through gene duplication about 500 million years ago.

CHECKpoint CONCEPT 15.6

- ✓ What are some of the potential advantages of lateral gene transfer to the organisms that gain new genes by this process?
- ✓ Why is gene duplication considered important for long-term evolutionary change?
- ✓ Why is sexual reproduction so prevalent in nature, despite its having at least three short-term evolutionary disadvantages?

The development of evolutionary theory has helped reveal how biological molecules function, how genetic diversity is created

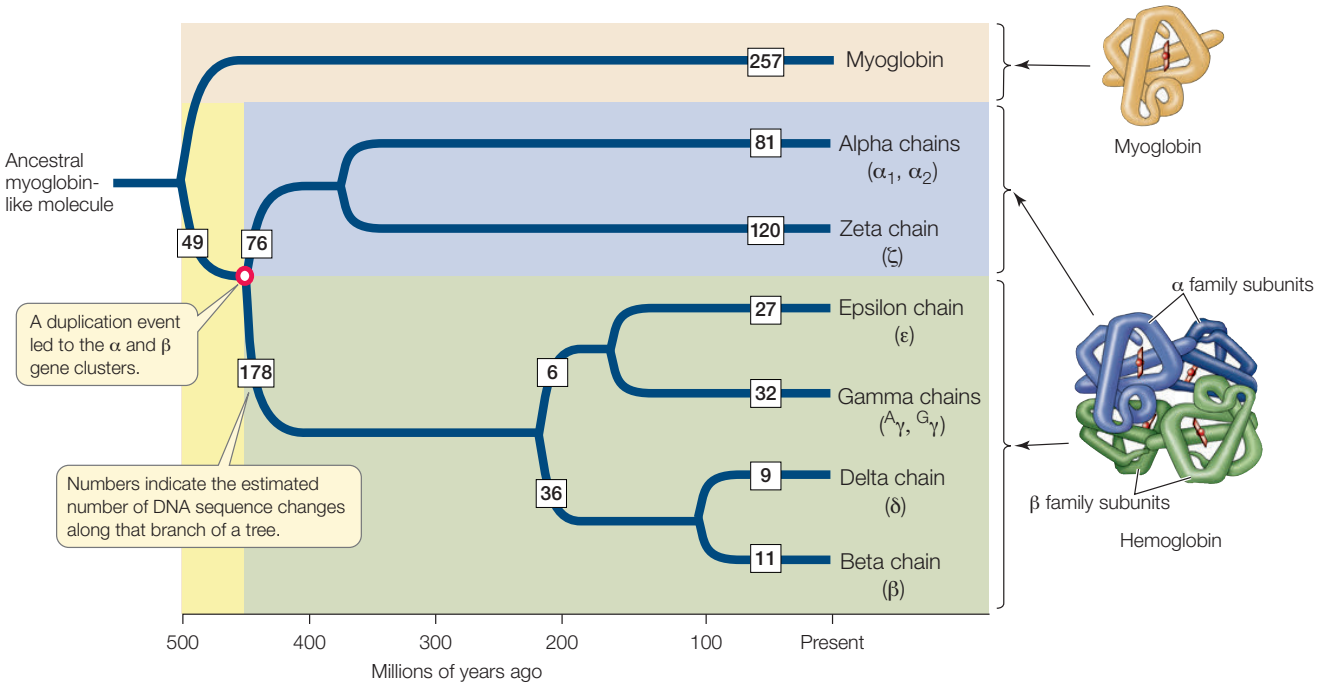


FIGURE 15.23 A Globin Family Gene Tree This gene tree suggests that the α-globin and β-globin gene clusters diverged about 450 million years ago (open circle), soon after the origin of the vertebrates.

Go to ACTIVITY 15.2 Gene Tree Construction
PoL2e.com/ac15.2

and maintained, and how organisms develop new features. Next we will see how biologists put this theory into practice.

CONCEPT 15.7 Evolutionary Theory Has Practical Applications

Evolutionary theory has many practical applications across biology, and new ones are being developed every day. Here we'll discuss a few of these applications to fields such as agriculture, industry, and medicine.

Knowledge of gene evolution is used to study protein function

Earlier in this chapter we discussed some of the ways biologists can detect codons or genes that are under positive selection for change. These methods have greatly increased our understanding of the functions of many genes. Consider, for example, the gated sodium channel genes. Sodium channels have many functions, including the control of nerve impulses in the nervous system (see Concept 34.2). Sodium channels can become blocked when they bind certain toxins, one of which is the tetrodotoxin (TTX) present in puffer fishes and many other animals. A human who eats puffer fish tissues that contain TTX can become paralyzed and die because the toxin-blocked sodium channels prevent nerves and muscles from functioning properly.

But puffer fish themselves have sodium channels, so why doesn't the TTX in their system paralyze them? Nucleotide substitutions in the puffer fish genome have resulted in structural changes in the proteins that form the sodium channels, and those changes prevent TTX from binding to the channel pore. Several different substitutions that result in such resistance have evolved in the various duplicated sodium channel genes of the many species of puffer fish. Many other changes that have nothing to do with the evolution of tetrodotoxin resistance have occurred in these genes as well.

So how does what we have learned about the evolution of TTX-resistant sodium channels affect our lives? Mutations in human sodium channel genes are responsible for a number of neurological pathologies. By studying the function of sodium channels and understanding which changes have produced tetrodotoxin resistance, we are learning a great deal about how these crucial channels work and how various mutations affect them. Biologists do this by comparing rates of synonymous and nonsynonymous substitutions across sodium channel genes in various animals that have evolved TTX resistance. In a similar manner, molecular evolutionary principles are used to understand function and diversification of function in many other proteins.

In vitro evolution produces new molecules

Living organisms produce thousands of compounds that humans have found useful. The search for naturally occurring compounds that can be used for pharmaceutical, agricultural, or industrial purposes has been termed "bioprospecting." These compounds are the result of millions of years of molecular evolution across millions of species of living organisms. Yet biologists can imagine molecules that could have evolved but have not, in the absence of the right combination of selection pressures and opportunities.

For instance, we might want to find a molecule that binds a particular environmental contaminant so that the contaminant can be isolated and extracted from the environment. But if the contaminant is synthetic (not produced naturally), then it is unlikely that any living organism would have evolved a molecule with the function we desire. This problem was the inspiration for the field of **in vitro evolution**, in which new molecules are produced in the laboratory to perform novel and useful functions.

The principles of in vitro evolution are based on principles of molecular evolution that we have learned from the natural world. Consider a new RNA molecule that was produced in the laboratory using the principles of mutation and selection. The new molecule's intended function was to join two other RNA molecules (acting as a ribozyme with a function similar to that of the naturally occurring DNA ligase described in Concept 9.2, but for RNA molecules). The process started with a large pool of random RNA sequences (10^{15} different sequences, each about 300 nucleotides long), which were then selected for displaying any ligase activity (**FIGURE 15.24**). None were very effective ligases, but some were slightly better than others. The most functional of the ribozymes were selected and reverse-transcribed into cDNA (using the enzyme reverse transcriptase). The cDNA molecules were then amplified using the polymerase chain reaction (PCR; see Figure 9.15). PCR amplification is not perfect, and it introduced many new mutations into the pool of RNA sequences. These sequences were then transcribed back into RNA molecules using RNA polymerase, and the process was repeated.

The ligase activity of the RNAs evolved quickly; after ten rounds of in vitro evolution, it had increased by about 7 million times. Similar techniques have been used to create a wide variety of molecules with novel enzymatic and binding functions.

Evolutionary theory provides multiple benefits to agriculture

Well before humans had a clear understanding of evolution, they were selecting beneficial traits in the plants and animals they used for food. Modern agricultural practices have benefitted from a clearer understanding of evolutionary principles. Agriculturists have also used knowledge of evolutionary relationships and principles to incorporate beneficial genes into our food crops from many wild species.

Evolutionary theory has also proved important for understanding how to reduce the threats of pesticide and herbicide resistance. When farmers use the same pesticide over many seasons, the pests they are trying to kill gradually evolve resistance to the pesticide. Each year, a few pest individuals are slightly better at surviving in the presence of the pesticide, and those individuals produce most of the next generation of crop pests. Because their genes allow them to survive at a higher rate, and because they pass these resistant genes on to their offspring, pesticide resistance quickly evolves in the entire

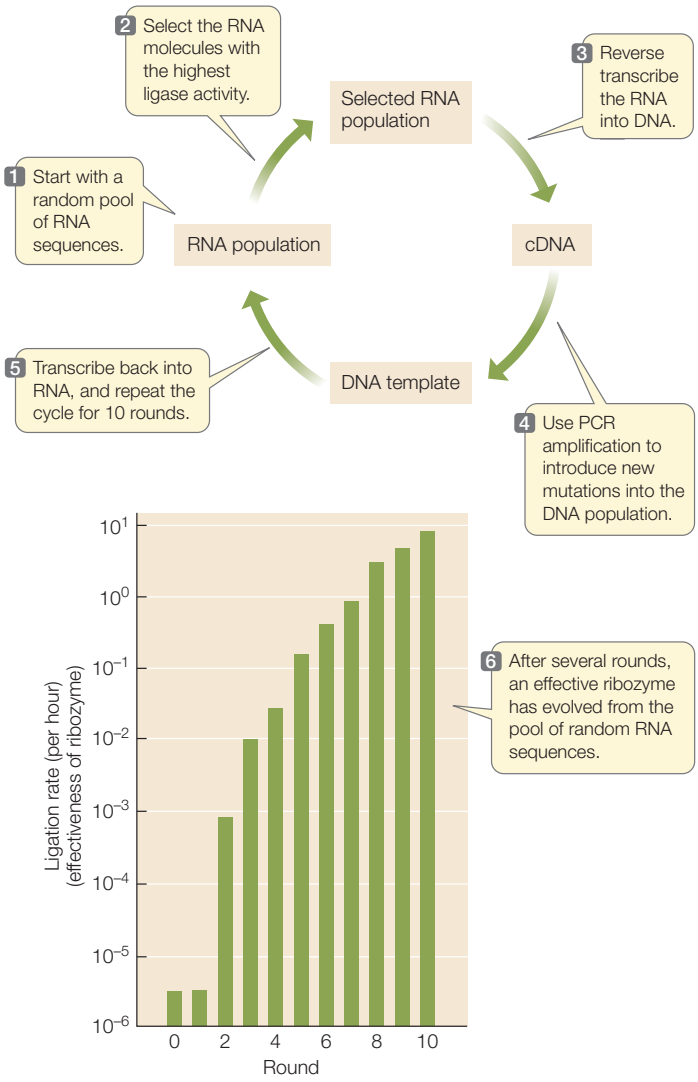


FIGURE 15.24 In Vitro Evolution Starting with a large pool of random RNA sequences, David Bartel and Jack Szostak of Massachusetts General Hospital produced a new ribozyme through rounds of mutation and selection for the ability to ligate (join) RNA sequences.

population. To combat this problem, evolutionary biologists have devised pesticide application and rotation schemes to reduce the rate of evolution of pesticide resistance, thus allowing farmers to use pesticides more effectively for longer periods of time.

Knowledge of molecular evolution is used to combat diseases

Many of the most problematic human diseases are caused by living, evolving organisms that present a moving target for modern medicine, as we described for influenza at the start of this chapter. The control of these and many other human diseases depends on techniques that can track the evolution of pathogenic organisms over time.

During the past century, transportation advances have allowed humans to move around the world with unprecedented speed and increasing frequency. Unfortunately, this mobility has increased the rate at which pathogens are transmitted among human populations, leading to the global emergence of many "new" diseases. Most of these emerging diseases are caused by viruses, and virtually all new viral diseases have been identified by evolutionary comparison of their genomes with those of known viruses. In recent years, rodent-borne hantaviruses have been identified as the source of widespread respiratory illnesses, and the virus that causes sudden acute respiratory syndrome (SARS) has been identified, as has its host, using evolutionary comparisons of genes. Studies of the origins, timing of emergence, and global diversity of many human pathogens (including HIV, the human immunodeficiency virus) depend on evolutionary principles and methods, as do efforts to develop effective vaccines against these pathogens.

At present, it is difficult to identify many common infections (the viral strains that cause "colds," for instance). As genomic databases increase, however, automated methods of sequencing and making evolutionary comparisons of sequences will allow us to identify and treat a much wider array of human (and other) diseases. Once biologists have collected genome data for enough infectious organisms, it will be possible to identify an infection by sequencing a portion of the pathogen's genome and comparing this sequence with other sequences on an evolutionary tree.

CHECKpoint CONCEPT 15.7

- ✓ How can gene evolution be used to study protein function?
- ✓ How are principles of evolutionary biology used to identify emerging diseases?
- ✓ What are the key elements of in vitro evolution, and how do these elements correspond to natural evolutionary processes?

The processes of evolution have produced a remarkable variety of organisms, some of which are adapted to most environments on Earth. In the next chapter we will describe how biologists study the evolutionary relationships across the great diversity of life.

Q How do biologists use evolutionary theory to develop better flu vaccines?

ANSWER Many different strains of influenza virus circulate among human populations and other vertebrate hosts each year, but only a few of those strains survive to leave descendants. Selection among these circulating influenza strains results in rapid evolution of the viral genome. One of the ways that influenza strains differ is in the configuration of

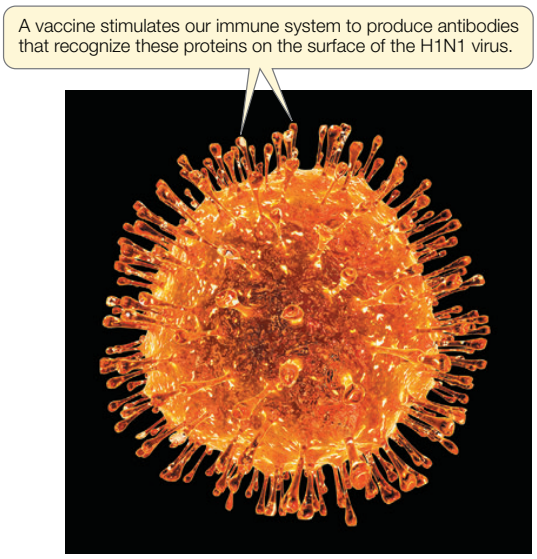


FIGURE 15.25 Evolutionary Analysis of Surface Proteins Leads to Improved Flu Vaccines This computer-generated image is of the H1N1 virus that was the target of a 2009–2010 flu vaccine. Rapidly evolving surface proteins ("spikes" in this illustration) allow flu viruses to escape detection by the host's immune system. Analyzing the surface proteins among current strains of the virus can help biologists anticipate which strains are most likely to be the cause of future epidemics.

proteins on their surface. These surface proteins are the targets of recognition by the host immune system (**FIGURE 15.25**).

When changes occur in the surface proteins of an influenza virus, the host immune system may no longer detect the invading virus, so the virus is more likely to replicate successfully. The viral strains with the greatest number of changes to their surface proteins are most likely to escape detection by the host immune system, and are therefore most likely to spread among the host population and result in future flu epidemics. In other words, there is positive selection for change in the surface proteins of influenza.

By comparing the survival and proliferation rates of virus strains that have different gene sequences coding for their surface proteins, biologists can study adaptation of the viruses over time (Concept 15.2). If biologists can predict which of the currently circulating flu virus strains are most likely to escape host immune detection, then they can identify the strains that are most likely to be involved in upcoming influenza epidemics and can target those strains for vaccine production.

How can biologists make such predictions? By examining the ratio of synonymous to nonsynonymous substitutions in genes that encode viral surface proteins, biologists can detect which codon changes (i.e., mutations) are under positive selection (Concept 15.5). They can then assess which of the currently circulating flu strains show the greatest number of changes in these positively selected codons. It is these flu strains that are most likely to survive and lead to the flu epidemics of the future, so they are the best targets for new vaccines. This practical application of evolutionary theory leads to more effective flu vaccines—and thus fewer illnesses and influenza-related deaths each year.

SUMMARY

CONCEPT 15.1 Evolution Is Both Factual and the Basis of Broader Theory

- **Evolution** is genetic change in populations over time. Evolution can be observed directly in living populations as well as in the fossil record of life.
- **Evolutionary theory** refers to our understanding and application of the processes of evolutionary change.
- Charles Darwin in best known for his ideas on the common ancestry of divergent species and on **natural selection** as a process of evolution. **Review ANIMATED TUTORIAL 15.1 and ACTIVITY 15.1**
- Since Darwin's time, many biologists have contributed to the development of evolutionary theory, and rapid progress in our understanding continues today. **Review Figure 15.2**

CONCEPT 15.2 Mutation, Selection, Gene Flow, Genetic Drift, and Nonrandom Mating Result in Evolution

- Mutation produces new genetic variants (**alleles**).
- Within **populations**, natural selection acts to increase the frequency of beneficial alleles and decrease the frequency of deleterious alleles.
- **Adaptation** refers both to a trait that evolves through natural selection and to the process that produces such traits.
- Migration or mating of individuals between populations results in **gene flow**.
- **Genetic drift**—the random loss of individuals and the alleles they possess—may produce large changes in allele frequencies from one generation to the next and greatly reduce genetic variation.
- **Population bottlenecks** occur when only a few individuals survive a random event, resulting in a drastic shift in allele frequencies within the population and the loss of variation. Similarly, a population established by a small number of individuals colonizing a new region may lose variation via a **founder effect**. **Review Figure 15.8**
- Nonrandom mating may result in changes in genotype frequencies in a population.
- **Sexual selection** results from differential mating success of individuals based on their phenotype. **Review Figure 15.10**

(*continued*)

SUMMARY *(continued)*

CONCEPT 15.3 Evolution Can Be Measured by Changes in Allele Frequencies

- Allele frequencies measure the amount of genetic variation in a population. Genotype frequencies show how a population's genetic variation is distributed among its members. Together, allele and genotype frequencies describe a population's **genetic structure**. Review Figure 15.11 and ANIMATED TUTORIAL 15.2
- **Hardy–Weinberg equilibrium** predicts genotype frequencies from allele frequencies in the absence of evolution. Deviation from these frequencies indicates that evolutionary processes are at work. Review Figure 15.12 and ANIMATED TUTORIAL 15.3

CONCEPT 15.4 Selection Can Be Stabilizing, Directional, or Disruptive

- **Qualitative traits** differ by discrete qualities (e.g., black versus white) and often are determined by alleles of a single gene.
- **Quantitative traits** differ along a continuum (e.g., small to large size), and usually are influenced by variation at multiple genes.
- Natural selection can act on characters with quantitative variation in three different ways. Review Figure 15.13
- **Stabilizing selection** acts to reduce variation without changing the mean value of a trait. When applied to selection that maintains a particular genetic variant in a population, stabilizing selection is called **purifying selection**. Review Figure 15.14
- **Directional selection** acts to shift the mean value of a trait toward one extreme. When applied to selection for change at a single genetic locus, directional selection is called **positive selection**. Review Figure 15.15
- **Disruptive selection** favors both extremes of trait values, resulting in a bimodal character distribution. Review Figure 15.16

CONCEPT 15.5 Genomes Reveal Both Neutral and Selective Processes of Evolution

- **Nonsynonymous substitutions** of nucleotides result in amino acid replacements in proteins, but **synonymous substitutions** do not. Review Figure 15.17
- Rates of synonymous substitution are typically higher than rates of nonsynonymous substitution in protein-coding genes (a result of stabilizing selection). Review Figure 15.18
- Much of the change in nucleotide sequences over time is a result of neutral evolution. The rate of fixation of neutral mutations is independent of population size and is equal to the mutation rate.
- Positive selection for change in a protein-coding gene may be detected by a higher rate of nonsynonymous than synonymous substitution.
- Specific codons within a given gene sequence can be under different modes of selection. Review Figure 15.20
- The total size of genomes varies much more widely across multicellular organisms than does the number of functional genes. Review Figures 15.21 and 15.22
- Even though many noncoding regions of the genome may not have direct functions, these regions can affect the phenotype of an organism by influencing gene expression.
- Functionless **pseudogenes** can serve as the raw material for the evolution of new genes.

CONCEPT 15.6 Recombination, Lateral Gene Transfer, and Gene Duplication Can Result in New Features

- Despite its short-term disadvantages, sexual reproduction generates countless genotype combinations that increase genetic variation in populations.
- In the absence of genetic recombination (as in some asexual organisms), deleterious mutations accumulate with each replication—a phenomenon known as **Muller's ratchet**.
- **Lateral gene transfer** can result in the rapid acquisition of new functions from distantly related species.
- Gene duplications can result in increased production of the gene's product, in divergence of the duplicated genes' expression, in pseudogenes, or in new gene functions. Several rounds of gene duplication can give rise to multiple genes with related functions, known as a **gene family**. Review Figure 15.23 and ACTIVITY 15.2

CONCEPT 15.7 Evolutionary Theory Has Practical Applications

- Protein function can be studied by examining gene evolution. Detection of positive selection can be used to identify molecular changes that have resulted in functional changes.
- Agricultural applications of evolution include the development of new crop plants and domesticated animals, as well as a reduction in the rate of evolution of pesticide resistance.
- **In vitro evolution** is used to produce synthetic molecules with particular desired functions. Review Figure 15.24
- Many diseases are identified, studied, and combated through molecular evolutionary investigations.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is15

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

16

Reconstructing and Using Phylogenies

KEY CONCEPTS

The reef-building coral *Acropora millepora* shows cyan and red fluorescence. This photograph was taken under a fluorescent microscope that affects the colors we see. The colors are perceived differently by marine animals in their natural environment.

Green fluorescent protein (GFP) was discovered in 1962 when Osamu Shimomura, an organic chemist and marine biologist, led a team that was able to purify the protein from the tissues of the bioluminescent jellyfish *Aequorea victoria*. Some 30 years after GFP's initial discovery, Martin Chalfie had the idea (and the technology) to link the gene for GFP to other protein-coding genes, so that the expression of specific genes of interest could be visualized in glowing green within cells and tissues of living organisms (see Figure 13.6). This work was extended by Roger Tsien, who changed some of the amino acids within GFP to create proteins of several distinct colors. Different colored proteins meant that the expression of a number of different proteins could be visualized and studied in the same organism at the same time. These three scientists were awarded the 2008 Nobel Prize in Chemistry for the isolation and development of GFP for visualizing gene expression.

Tsien was able to produce different colored proteins, but he could not produce a *red* protein. This was frustrating; a red fluorescent protein would be particularly useful to biologists because red light penetrates tissues more easily than do other colors. Tsien's work stimulated Mikhail Matz to look for new fluorescent proteins in corals (which are relatives of the jellyfishes). Among the different species he studied, Matz found coral proteins that fluoresced in various shades of green, cyan (blue-green)—and red.

How had fluorescent red pigments evolved among the corals, given that the necessary molecular changes had eluded Tsien? To answer this question, Matz sequenced the genes of the fluorescent proteins and used these sequences to reconstruct the evolutionary history of the amino acid changes that produced different colors in different species of corals.

Matz's work showed that the ancestral fluorescent protein in corals was green, and that red fluorescent proteins evolved in a series of gradual steps. His analysis of evolutionary relationships allowed him to retrace these steps. Such an evolutionary history, as depicted in a tree of relationships among lineages, is called a phylogeny.

The evolution of many aspects of an organism's biology can be studied using phylogenetic methods. This information is used in all fields of biology to understand the structure, function, and behavior of organisms.

Q How are phylogenetic methods used to resurrect protein sequences from extinct organisms?

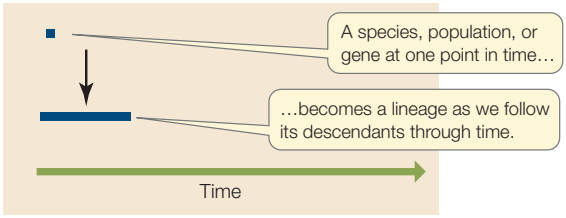
You will find the answer to this question on page 341.

CONCEPT 16.1 All of Life Is Connected through Its Evolutionary History

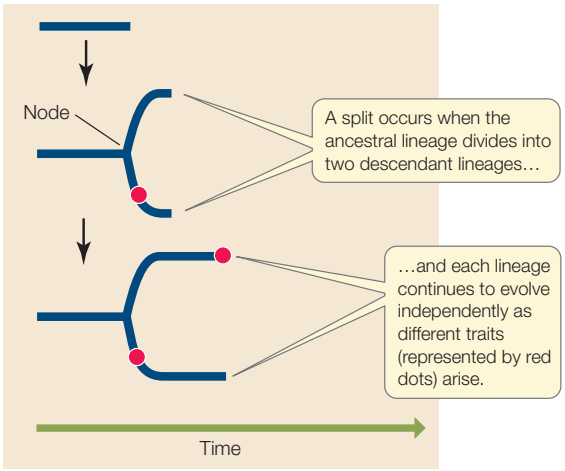
The sequencing of complete genomes from many diverse species has confirmed what biologists have long suspected: all of life is related through a common ancestor. The common ancestry of life explains why the general principles of biology apply to all organisms. Thus we can learn much about how the human genome works by studying the biology of model organisms because we share a common evolutionary history with those organisms. The evolutionary history of these relationships is known as **phylogeny**, and a **phylogenetic tree** is a diagrammatic reconstruction of that history.

Phylogenetic trees are commonly used to depict the evolutionary history of species, populations, and genes. For many years such trees have been constructed based on physical structures, behaviors, and biochemical attributes. Now, as genomes are sequenced for more and more organisms, biologists are able to reconstruct the history of life in ever greater detail.

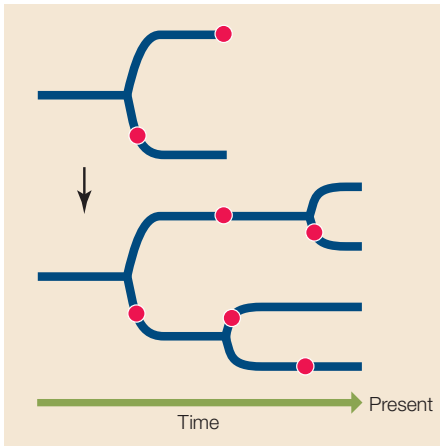
In Chapter 15 we discussed why we expect populations of organisms to evolve over time. Such a series of ancestor and descendant populations forms a **lineage**, which we can depict as a line drawn on a time axis:



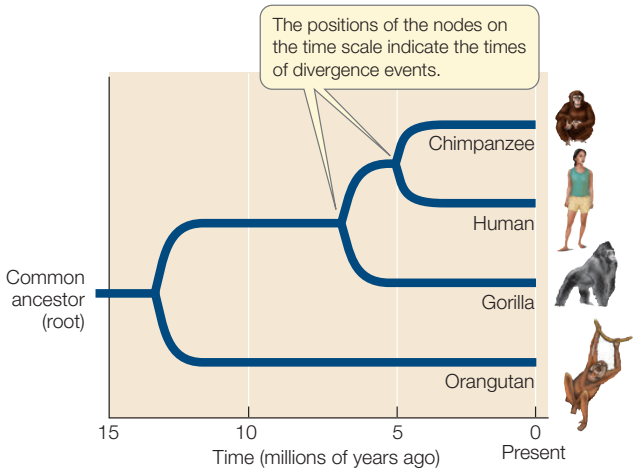
What happens when a single lineage divides into two? For example, a geographic barrier (such as a new mountain range) may divide an ancestral population into two descendant populations that no longer interact with one another. We depict such an event as a split, or **node**, in a phylogenetic tree. Each of the descendant populations gives rise to a new lineage, and as these independent lineages evolve, new traits arise in each:



As the lineages continue to split over time, this history can be represented in the form of a branching tree that can be used to trace the evolutionary relationships from the ancient common ancestor of a group of species, through the various lineage splits, up to the present populations of the organisms:

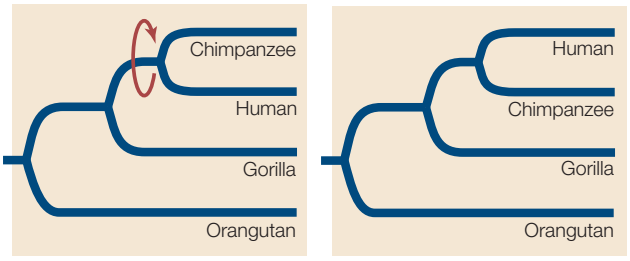


A phylogenetic tree may portray the evolutionary history of all life forms. Phylogenetic trees can also depict the history of a major evolutionary group (such as the insects) or of a much smaller group of closely related species. In some cases, phylogenetic trees are used to show the history of individuals, populations, or genes within a species. The common ancestor of all the organisms in the tree forms the **root** of the tree. The depictions of phylogenetic trees in this book are rooted at the left, with time flowing from left (earliest) to right (most recent):



The timing of splitting events in lineages is shown by the position of nodes on a time axis. These splits represent events where one lineage diverged into two, such as a speciation event (for a tree of species), a gene duplication event (for a tree of genes), or a transmission event (for a tree of viral lineages transmitted through a host population). The time axis may have an explicit scale, or it may simply show the relative timing of divergence events.

In this book's illustrations, the order in which nodes are placed along the horizontal (time) axis has meaning, but the vertical distance between the branches does not. Vertical distances have been adjusted for legibility and clarity of presentation; they do not correlate with the degree of similarity or difference among groups. Note too that lineages can be rotated around nodes in the tree, so the vertical order of lineages is also largely arbitrary:



Any group of species that we designate with a name is a **taxon** (plural *taxa*). Examples of familiar taxa include humans, primates, mammals, and vertebrates; in this series, each taxon is also a member of the next, more inclusive taxon. Any taxon that consists of all the evolutionary descendants of a common ancestor is called a **clade**. Clades can be identified by picking any point on a phylogenetic tree and from that point tracing all the descendant lineages to the tips of the terminal branches (FIGURE 16.1). Two species that are each other's closest relatives are called **sister species**. Similarly, any two clades that are each other's closest relatives are **sister clades**.

Before the 1980s, phylogenetic trees tended to be seen only in the literature on evolutionary biology, especially in the area of **systematics**—the study and classification of biodiversity. But almost every journal in the life sciences published during the last few years contains phylogenetic trees. Trees are widely used in molecular biology, biomedicine, physiology, behavior, ecology, and virtually all other fields of biology. Why have phylogenetic studies become so widespread?

Phylogenetic trees are the basis of comparative biology

In biology we study life at all levels of organization—from genes, cells, organisms, populations, and species to the major divisions of life. In most cases, however, no individual gene or organism (or other unit of study) is exactly like any other gene or organism that we investigate.

Consider the individuals in your biology class. We recognize each person as an individual human, but we know that no two are exactly alike. If we knew everyone's family tree in detail, the genetic similarity of any pair of students would be more predictable. We would find that more closely related students have many more traits in common (from the color of their hair to their susceptibility or resistance to diseases). Likewise, biologists use phylogenies to make comparisons and predictions about shared traits across genes, populations, and species.

The evolutionary relationships among species, as represented in the tree of life, form the basis for biological

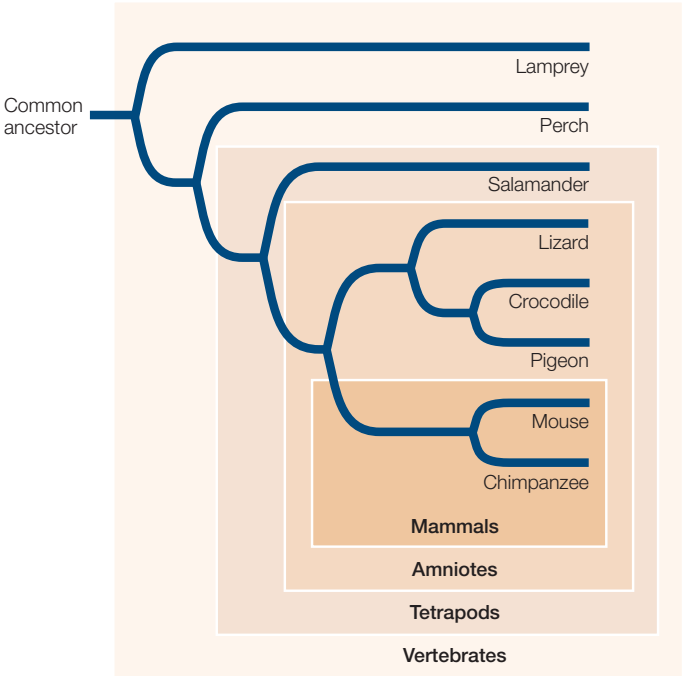


FIGURE 16.1 Clades Represent All the Descendants of a Common Ancestor All clades are subsets of larger clades, with all of life as the most inclusive taxon. In this example, the groups called mammals, amniotes, tetrapods, and vertebrates represent successively larger clades. Only a few species within each clade are represented on this tree.

classification. Biologists estimate that there are tens of millions of species on Earth. So far, however, only about 1.8 million species have been classified—that is, formally described and named. New species are being discovered all the time and phylogenetic analyses are constantly reviewed and revised, so our knowledge of the tree of life is far from complete. Yet knowledge of evolutionary relationships is essential for making comparisons in biology, so biologists build phylogenies for groups of interest as the need arises. The tree of life's evolutionary framework allows us to make many predictions about the behavior, ecology, physiology, genetics, and morphology of species that have not yet been studied in detail.

When biologists compare species, they observe traits that differ within the group of interest and try to understand when these traits evolved. In many cases, investigators are interested in how the evolution of a trait relates to environmental conditions or selective pressures. For instance, scientists have used phylogenetic analyses to discover changes in the genome of human immunodeficiency viruses that result in resistance to particular drug treatments. The association of a particular genetic change in HIV with a particular treatment provides a hypothesis about the evolution of resistance that can be tested experimentally.

Any features shared by two or more species that have been inherited from a common ancestor are said to be **homologous**. Homologous features may be any heritable traits, including DNA sequences, protein structures, anatomical structures, and

even some behavior patterns. For example, all living vertebrates have a vertebral column, as did the ancestral vertebrate. Therefore the vertebral column is judged to be homologous in all vertebrates.

Derived traits provide evidence of evolutionary relationships

In tracing the evolution of a character, biologists distinguish between ancestral and derived traits. Each character of an organism evolves from one condition, called the **ancestral trait**, to another condition, called the **derived trait**.

Derived traits that are shared among a group of organisms and are also viewed as evidence of the common ancestry of the group are called **synapomorphies** (*syn*, "shared"; *apo*, "derived"; *morph*, "form," referring to the "form" of a trait). Thus the vertebral column is considered a synapomorphy—a shared, derived trait—of the vertebrates. (The ancestral trait was an undivided supporting rod.)

Not all similar traits are evidence of relatedness. Similar traits in unrelated groups of organisms can develop for either of the following reasons:

- Superficially similar traits may evolve independently in different lineages, a phenomenon called **convergent evolution**. For example, although the *wing bones* of bats and birds are homologous, having been inherited from a common tetrapod ancestor, the *wings* of bats and birds are not homologous because they evolved independently from the forelimbs of different nonflying ancestors (FIGURE 16.2).

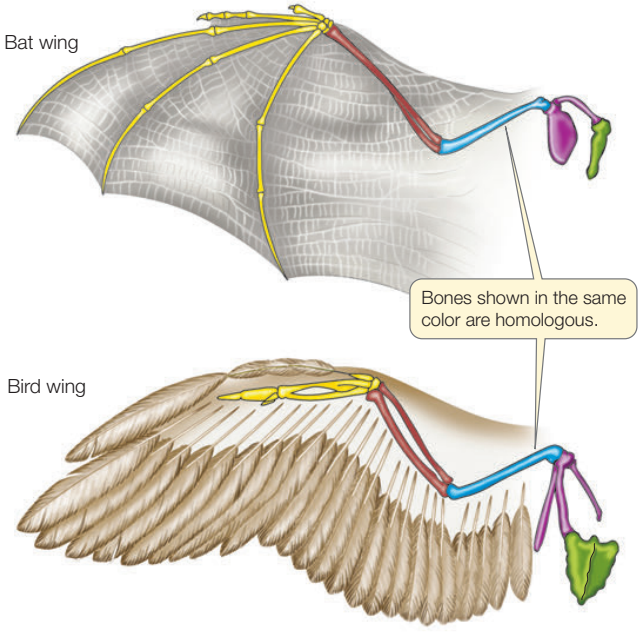


FIGURE 16.2 The Bones Are Homologous, the Wings Are Not
The supporting bone structures of both bat wings and bird wings are derived from a common tetrapod (four-limbed) ancestor and are thus homologous. However, the wings themselves—an adaptation for flight—evolved independently in the two groups.

- A character may revert from a derived state back to an ancestral state in an event called an **evolutionary reversal**. For example, the derived limbs of terrestrial tetrapods evolved from the ancestral fins of their aquatic ancestors. Then, within the mammals, the ancestors of modern cetaceans (whales and dolphins) returned to the ocean, and cetacean limbs evolved to once again resemble their ancestral state—fins. The superficial similarity of cetacean and fish fins does not suggest a close relationship between these groups. Instead, the similarity arises from evolutionary reversal.

Similar traits generated by convergent evolution and evolutionary reversals are called homoplastic traits or **homoplasies**.


Go to MEDIA CLIP 16.1
Morphing Arachnids
PoL2e.com/mc16.1

A particular trait may be ancestral or derived, depending on our point of reference. For example, all birds have feathers. We infer from this that feathers (which are highly modified scales) were present in the common ancestor of modern birds. Therefore we consider the presence of feathers to be an ancestral trait for any particular group of modern birds, such as the songbirds. However, feathers are not present in any other living animals. In reconstructing a phylogeny of all living vertebrates, the presence of feathers is a derived trait found only among birds, and thus is a synapomorphy of the birds.

CHECKpoint CONCEPT 16.1

- ✓ What biological processes can be represented in a phylogenetic tree?
- ✓ Why is it important to consider only homologous characters in constructing phylogenetic trees?
- ✓ What are some reasons that similar traits might arise independently in species that are only distantly related? Can you think of examples among familiar organisms?

Phylogenetic analyses of evolutionary history have become increasingly important to many types of biological research in recent years, and they are the basis for the comparative nature of biology. For the most part, however, evolutionary history cannot be observed directly. How, then, do biologists reconstruct the past?

CONCEPT 16.2 Phylogeny Can Be Reconstructed from Traits of Organisms

To illustrate how a phylogenetic tree is constructed, consider the eight vertebrate animals listed in TABLE 16.1: lamprey, perch, salamander, lizard, crocodile, pigeon, mouse, and chimpanzee. We will initially assume that any given derived trait arose only once during the evolution of these animals (that is,

TABLE 16.1 Eight Vertebrates and the Presence or Absence of Some Shared Derived Traits

	Derived trait							
Taxon	Jaws	Lungs	Claws or nails	Gizzard	Feathers	Fur	Mammary glands	Keratinous scales
Lamprey (outgroup)	–	–	–	–	–	–	–	–
Perch	+	–	–	–	–	–	–	–
Salamander	+	+	–	–	–	–	–	–
Lizard	+	+	+	–	–	–	–	+
Crocodile	+	+	+	+	–	–	–	+
Pigeon	+	+	+	+	+	–	–	+
Mouse	+	+	+	–	–	+	+	_
Chimpanzee	+	+	+	–	–	+	+	_

there has been no convergent evolution), and that no derived traits were lost from any of the descendant groups (there has been no evolutionary reversal). For simplicity, we have selected traits that are either present (+) or absent (–).

In a phylogenetic study, the group of organisms of primary interest is called the **ingroup**. As a point of reference, an ingroup is compared with an **outgroup**: a species or group that is closely related to the ingroup but is known to be phylogenetically outside it. In other words, the root of the tree is located between the ingroup and the outgroup. Any trait that is present in both the ingroup and the outgroup must have evolved before the origin of the ingroup and thus must be ancestral for the ingroup. In contrast, traits that are present in only some members of the ingroup must be derived traits within that ingroup. As we will see in Chapter 23, a group of jawless fishes called the lampreys is thought to have separated from the lineage leading to the other vertebrates before the jaw arose. Therefore we have included the lamprey as the outgroup for our analysis. Because derived traits are traits acquired by other members of the vertebrate lineage *after* they diverged from the outgroup, any trait that is present in both the lamprey and the other vertebrates is judged to be ancestral.

We begin by noting that the chimpanzee and mouse share two traits—mammary glands and fur—that are absent in both the outgroup and in the other species of the ingroup. Therefore we infer that mammary glands and fur are derived traits that evolved in a common ancestor of chimpanzees and mice after that lineage separated from the lineages leading to the other vertebrates. These characters are synapomorphies that unite chimpanzees and mice (as well as all other mammals, although we have not included other mammalian species in this example). By the same reasoning, we can infer that the other shared derived traits are synapomorphies for the various groups in which they are expressed. For instance, keratinous scales are a synapomorphy of the lizard, crocodile, and pigeon.

Table 16.1 also tells us that, among the animals in our ingroup, the pigeon has a unique trait: the presence of feathers. Feathers are a synapomorphy of birds and their extinct relatives. However, because we only have one bird in this example, the presence of feathers provides no clues concerning relationships among these eight species of vertebrates. However, gizzards are found in both birds and crocodiles, so this trait is evidence of a close relationship between birds and crocodilians.

By combining information about the various synapomorphies, we can construct a phylogenetic tree. We infer from our information that mice and chimpanzees—the only two animals that share fur and mammary glands—share a more recent common ancestor with each other than they do with pigeons and crocodiles. Otherwise we would need to assume that the ancestors of pigeons and crocodiles also had fur and mammary glands but subsequently lost them. There is no need to make these additional assumptions.

FIGURE 16.3 shows a phylogenetic tree for the vertebrates in Table 16.1, based on the shared derived traits we examined. This particular tree was easy to construct because it is based on a very small sample of traits, and the derived traits we examined evolved only once and were never lost after they appeared. Had we included a snake in the group, our analysis would not have been as straightforward. We would have needed to examine additional characters to determine that snakes evolved from a group of lizards that had limbs. In fact, the analysis of many characters shows that snakes evolved from burrowing lizards that became adapted to a subterranean existence.

Parsimony provides the simplest explanation for phylogenetic data

Typically, biologists construct phylogenetic trees using hundreds or thousands of traits. With larger data sets, we would expect to observe traits that have changed more than once, and thus would expect to see convergence and evolutionary reversal. How do we determine which traits are synapomorphies and which are homoplasies? One way is to invoke the principle of parsimony.

In its most general form, the **parsimony principle** states that the preferred explanation of observed data is the simplest

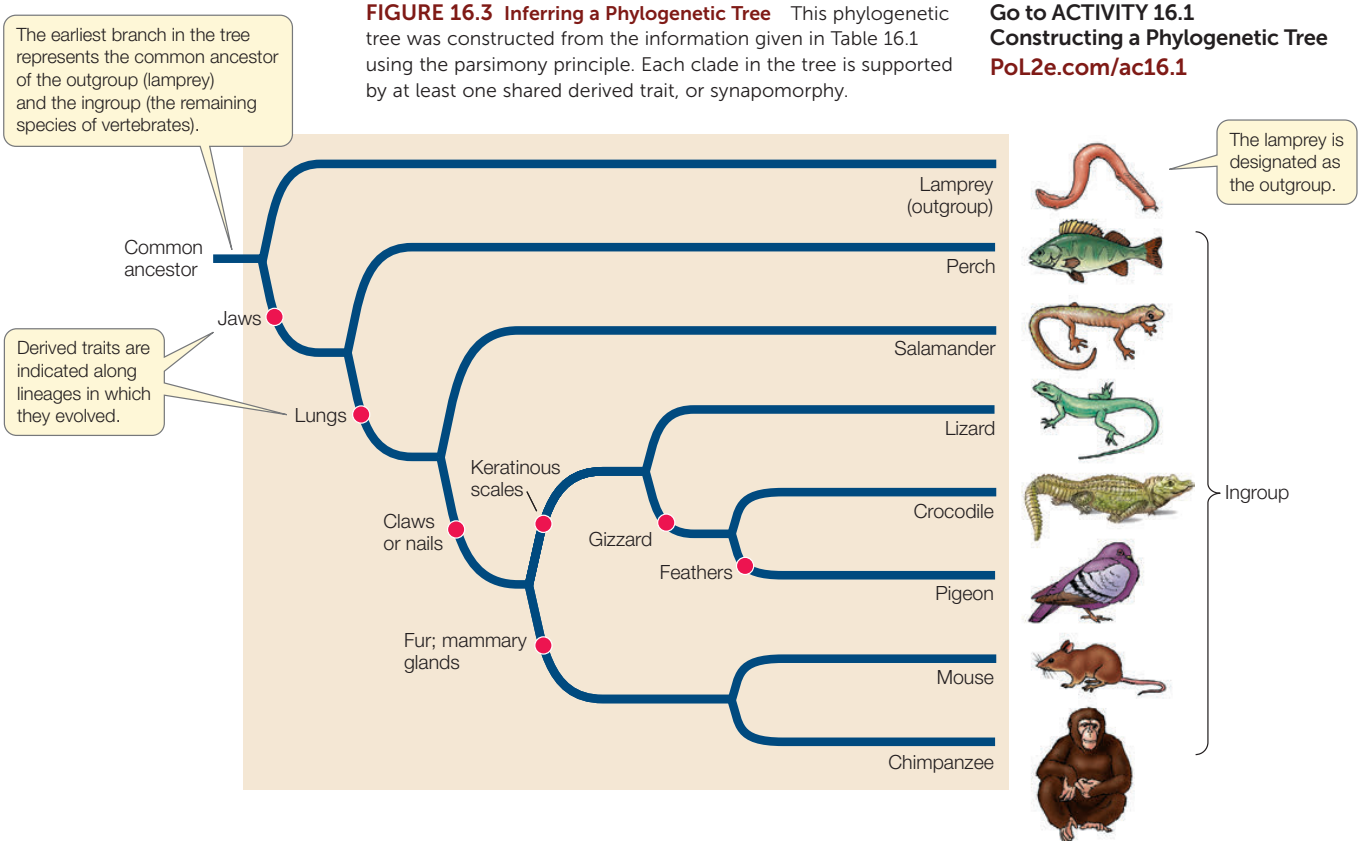


FIGURE 16.3 Inferring a Phylogenetic Tree This phylogenetic tree was constructed from the information given in Table 16.1 using the parsimony principle. Each clade in the tree is supported by at least one shared derived trait, or synapomorphy.

Go to ACTIVITY 16.1
Constructing a Phylogenetic Tree
PoL2e.com/ac16.1

APPLY THE CONCEPT

Phylogeny can be reconstructed from traits of organisms

The matrix below supplies data for seven land plants and an outgroup (an aquatic plant known as a stonewort). Each trait is scored as either present (+) or absent (–) in each of the plants. Use this data matrix to reconstruct the phylogeny of land plants and answer the questions. See Activity 16.1 for help with constructing a phylogenetic tree.

1. Which two of these taxa are most closely related?
2. Plants that produce seeds are known as seed plants. What is the sister group to the seed plants among these taxa?
3. Which two traits evolved along the same branch of your reconstructed phylogeny?
4. Are there any homoplasies in your reconstructed phylogeny?

	TRAIT						
TAXON	PROTECTED EMBRYOS	TRUE ROOTS	PERSISTENTLY GREEN SPOROPHYTE	VASCULAR CELLS	STOMATA	MEGAPHYLLS (TRUE LEAVES)	SEEDS
Stonewort (outgroup)	–	–	–	–	–	–	–
Liverwort	+	–	–	–	–	–	–
Pine tree	+	+	+	+	+	+	+
Bracken fern	+	+	+	+	+	+	–
Club moss	+	+	+	+	+	–	–
Sphagnum moss	+	–	–	–	+	–	–
Hornwort	+	–	+	–	+	–	–
Sunflower	+	+	+	+	+	+	+

explanation. Applying the principle of parsimony to the reconstruction of phylogenies entails minimizing the number of evolutionary changes that need to be assumed over all characters in all groups in the tree. In other words, the best hypothesis under the parsimony principle is one that requires the fewest homoplasies. This application of parsimony is a specific case of a general principle of reasoning called Occam's razor: the best explanation is the one that best fits the data while making the fewest assumptions. More complicated explanations are accepted only when the evidence requires them. Phylogenetic trees represent our best estimates about evolutionary relationships, given our current knowledge. They are continually modified as additional evidence becomes available.

Phylogenies are reconstructed from many sources of data

Naturalists have constructed various forms of phylogenetic trees for more than 150 years. In fact, the only figure in the first edition of *On the Origin of Species* was a phylogenetic tree. Tree construction has been revolutionized, however, by the advent of computer software that allows us to consider far more data and analyze many more traits than could ever before be processed. Combining these advances in methodology with the massive comparative data sets being generated through studies of genomes, biologists are learning details about the tree of life at a remarkable pace (see Appendix A: The Tree of Life).

Any trait that is genetically determined, and therefore heritable, can be used in a phylogenetic analysis. Evolutionary relationships can be revealed through studies of morphology, development, the fossil record, behavioral traits, and molecular traits such as DNA and protein sequences. Let's take a closer look at the types of data used in modern phylogenetic analyses.

Go to ANIMATED TUTORIAL 16.1
Phylogeny and Molecular Evolution Simulation
PoL2e.com/at16.1

MORPHOLOGY An important source of phylogenetic information is **morphology**: the presence, size, shape, and other attributes of body parts. Since living organisms have been observed, depicted, and studied for millennia, we have a wealth of recorded morphological data as well as extensive museum and herbarium collections of organisms whose traits can be measured. New technological tools, such as the electron microscope and computed tomography (CT) scans, enable systematists to examine and analyze the structures of organisms at much finer scales than was formerly possible.

Most species are described and known primarily by their morphology, and morphology still provides the most comprehensive data set available for many taxa. The morphological features that are important for phylogenetic analysis are often specific to a particular group. For example, the presence, development, shape, and size of various features of the skeletal system are important in vertebrate phylogeny, whereas floral structures are important for studying the relationships among flowering plants.

Morphological approaches to phylogenetic analysis have some limitations, however. Some taxa exhibit little morphological diversity, despite great species diversity. For example, the phylogeny of the leopard frogs of North and Central America would be difficult to infer from morphological differences alone, because the many species look very similar, despite important differences in their behavior and physiology. At the other extreme, few morphological traits can be compared across distantly related species (earthworms and mammals, for example). Furthermore, some morphological variation has an environmental (rather than a genetic) basis and so must be excluded from phylogenetic analyses. An accurate phylogenetic analysis often requires information beyond that supplied by morphology.

DEVELOPMENT Similarities in developmental patterns may reveal evolutionary relationships. Some organisms exhibit similarities in early developmental stages only. The larvae of marine creatures called sea squirts, for example, have a flexible gelatinous rod in the back—the notochord—that disappears as the larvae develop into adults. All vertebrate animals also have a notochord at some time during their development (**FIGURE 16.4**). This shared structure is one of the reasons for inferring that sea squirts are more closely related to vertebrates than would be suspected if only adult sea squirts were examined.

LINK

For more on the role of developmental processes in evolution, see **Concepts 14.4 and 14.5**

PALEONTOLOGY The fossil record is another important source of information on evolutionary history. Fossils show us where and when organisms lived in the past and give us an idea of what they looked like. Fossils provide important evidence that helps us distinguish ancestral from derived traits. The fossil record can also reveal when lineages diverged and began their independent evolutionary histories. Furthermore, in groups with few species that have survived to the present, information on extinct species is often critical to an understanding of the large divergences among the surviving species. The fossil record has limitations, however. Few or no fossils have been found for some groups, and the fossil record for many groups is fragmentary.

BEHAVIOR Some behavioral traits are culturally transmitted and others are genetically inherited. If a particular behavior is culturally transmitted, it may not accurately reflect evolutionary relationships (but may nonetheless reflect cultural connections). Many bird songs, for instance, are learned and may be inappropriate traits for phylogenetic analysis. Frog calls, however, are genetically determined and appear to be acceptable sources of information for reconstructing phylogenies.

MOLECULAR DATA All heritable variation is encoded in DNA, and so the complete genome of an organism contains an

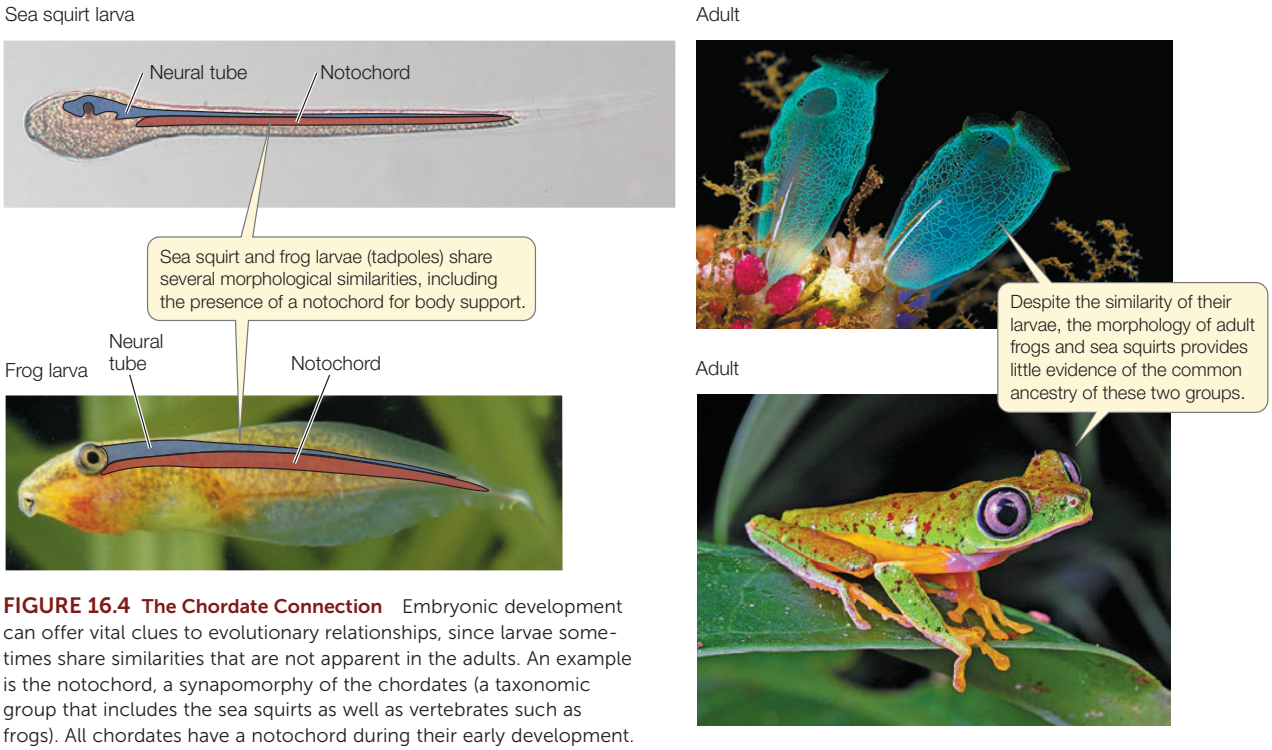


FIGURE 16.4 The Chordate Connection Embryonic development can offer vital clues to evolutionary relationships, since larvae sometimes share similarities that are not apparent in the adults. An example is the notochord, a synapomorphy of the chordates (a taxonomic group that includes the sea squirts as well as vertebrates such as frogs). All chordates have a notochord during their early development. The notochord is lost in adult sea squirts, whereas in adult frogs—as in all vertebrates—the vertebral column replaces the notochord as the body's support structure.

enormous set of traits (the individual nucleotide bases of DNA) that can be used in phylogenetic analyses. In recent years, DNA sequences have become among the most widely used sources of data for constructing phylogenetic trees. Comparisons of nucleotide sequences are not limited to the DNA in the cell nucleus. Eukaryotes have genes in their mitochondria as well as in their nuclei. Plant cells also have genes in their chloroplasts. The chloroplast genome (cpDNA), which is used extensively in phylogenetic studies of plants, has changed slowly over evolutionary time, so it is often used to study relatively ancient phylogenetic relationships. Most animal mitochondrial DNA (mtDNA) has changed more rapidly, so mitochondrial genes are used to study evolutionary relationships among closely related animal species (the mitochondrial genes of plants evolve more slowly). Many nuclear gene sequences are also commonly analyzed, and now that entire genomes have been sequenced from many species, they too are used to construct phylogenetic trees. Information on gene products (such as the amino acid sequences of proteins) is also widely used for phylogenetic analyses.

Mathematical models expand the power of phylogenetic reconstruction

As biologists began to use DNA sequences to infer phylogenies in the 1970s and 1980s, they developed explicit mathematical models describing how DNA sequences change over time. These models account for multiple changes at a given position in a DNA sequence. They also take into account different rates of change at different positions in a gene, at different positions in a codon, and among different nucleotides. For example, transitions (changes between two purines or between two pyrimidines) are usually more likely than are transversions (changes between a purine and pyrimidine).

Mathematical models can be used to compute how a tree might evolve given the observed data. A **maximum likelihood** method will identify the tree that most likely produced the observed data, given the assumed model of evolutionary change. Maximum likelihood methods can be used for any kind of characters, but they are most often used with molecular data, for which explicit mathematical models of evolutionary change are easier to develop. The principal advantages to maximum likelihood analyses are that they incorporate more information about evolutionary change than do parsimony methods, and they are easier to treat in a statistical framework. The principal disadvantages are that they are computationally intensive and require explicit models of evolutionary change (which may not be available for some kinds of character change).

The accuracy of phylogenetic methods can be tested

How can we test the accuracy of phylogenetic methods? After all, phylogenetic trees represent reconstructions of past events, and many of these events occurred before any humans were around. To address this issue, biologists have conducted experiments both in living organisms and with computer simulations to test the effectiveness and accuracy of phylogenetic methods.

In one experiment designed to test the accuracy of phylogenetic analysis, a single viral culture of bacteriophage T7 was used as a starting point, and lineages were allowed to evolve from this ancestral virus in the laboratory (**FIGURE 16.5**). The initial culture was split into two separate lineages, one of which

INVESTIGATION

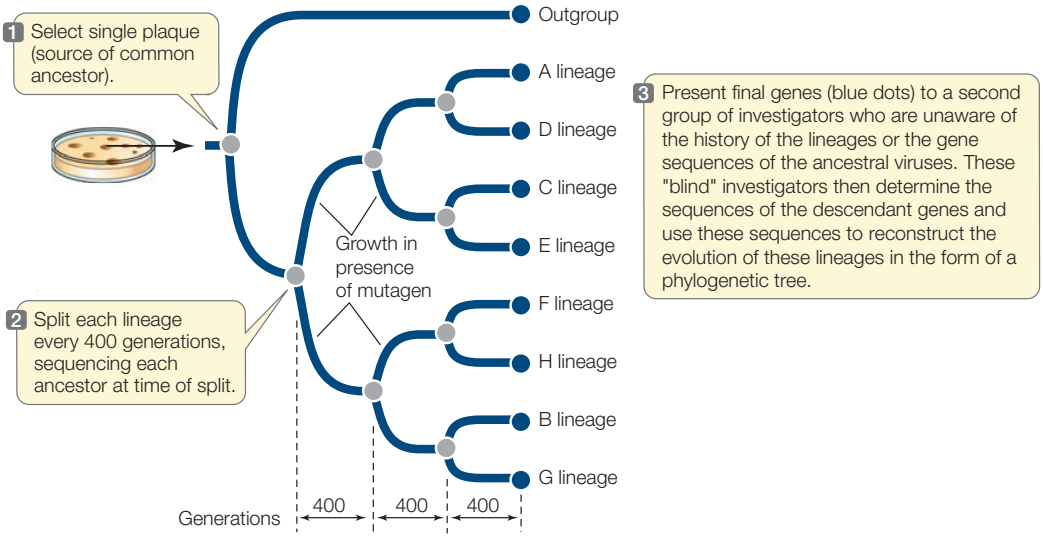
FIGURE 16.5 The Accuracy of Phylogenetic Analysis To test whether analysis of gene sequences can accurately reconstruct evolutionary history, we must have an unambiguously known phylogeny to compare against the reconstruction. Will the reconstruction match the observed phylogeny? See Animated Tutorial 16.1 for a simulation of this experiment.

HYPOTHESIS

A phylogenetic tree reconstructed from analysis of the DNA sequences of living organisms can accurately match the known evolutionary history of the organisms.

METHOD

In the laboratory, one group of investigators produced an experimental phylogeny of 9 viral lineages, enhancing the mutation rate to increase variation among the lineages.[a]



RESULTS

The true phylogeny and ancestral DNA sequences were accurately reconstructed solely from the DNA sequences of the viruses at the tips of the tree.

CONCLUSION

Phylogenetic analysis of DNA sequences can accurately reconstruct evolutionary history.

ANALYZE THE DATA

The full DNA sequences for the T7 strains in this experiment are thousands of nucleotides long. The nucleotides ("characters") at 23 DNA positions are given in the table.[b] See Activity 16.1 for help with constructing a phylogenetic tree.

	Character at position																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Outgroup	C	C	G	G	G	C	C	T	C	C	T	C	G	A	C	C	G	G	C	A	C	G	G
A	T	C	G	G	G	C	C	C	C	C	C	C	A	A	C	C	G	A	T	A	C	A	A
B	C	C	G	G	G	T	C	C	C	T	C	C	G	A	T	T	A	G	C	G	T	G	G
C	C	C	G	G	G	C	C	C	T	C	C	T	A	A	C	C	G	G	T	A	C	A	A
D	T	C	A	G	G	C	C	C	C	C	C	C	A	A	C	C	G	A	T	A	C	A	A
E	C	T	G	G	G	C	C	C	C	C	C	T	A	A	C	C	G	G	T	A	C	A	A
F	C	T	G	A	A	C	C	C	C	C	C	C	G	A	C	T	G	G	C	G	C	G	G
G	C	C	G	G	G	T	T	C	C	T	C	C	G	A	T	T	A	G	C	G	C	G	G
H	C	C	G	G	A	C	C	C	C	C	C	C	G	C	C	T	G	G	C	G	C	G	G

A. Construct a phylogenetic tree from these DNA positions using the parsimony method. Use the outgroup to root your tree. Assume that all changes among nucleotides are equally likely.

B. Using your tree, reconstruct the DNA sequences of the ancestral lineages.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]D. M. Hillis et al. 1992. *Science* 255: 589–295.
[b]J. J. Bull et al. 1993. *Evolution* 47: 993–1007.

became the ingroup for analysis and the other of which became the outgroup for rooting the tree. The lineages in the ingroup were split in two after every 400 generations, and samples of the virus were saved for analysis at each branching point. The lineages were allowed to evolve until there were eight lineages in the ingroup. Mutagens were added to the viral cultures to increase the mutation rate so that the amount of change and the degree of homoplasy would be typical of the organisms analyzed in average phylogenetic analyses. The investigators then sequenced samples from the end points of the eight lineages, as well as from the ancestors at the branching points. They then gave the sequences from the end points of the lineages to other investigators to analyze, without revealing the known history of the lineages or the sequences of the ancestral viruses.

After the phylogenetic analysis was completed, the investigators asked two questions. Did phylogenetic methods reconstruct the known history correctly? And were the sequences of the ancestral viruses reconstructed accurately? The answer in both cases was yes. The branching order of the lineages was reconstructed exactly as it had occurred, more than 98 percent of the nucleotide positions of the ancestral viruses were reconstructed correctly, and 100 percent of the amino acid changes in the viral proteins were reconstructed correctly.

Go to ANIMATED TUTORIAL 16.2
Using Phylogenetic Analysis to Reconstruct Evolutionary History
PoL2e.com/at16.2

The experiment shown in Figure 16.5 demonstrated that phylogenetic analysis was accurate under the conditions tested, but it did not examine all possible conditions. Other experimental studies have taken other factors into account, such as the sensitivity of phylogenetic analysis to parallel selection and highly variable rates of evolutionary change. In addition, computer simulations based on evolutionary models have been used extensively to study the effectiveness of phylogenetic analysis. These studies have also confirmed the accuracy of phylogenetic methods and have been used to refine those methods and extend them to new applications.

CHECKpoint CONCEPT 16.2

- ✓ How is the parsimony principle used in reconstructing evolutionary history?
- ✓ Why is it useful to consider the entire life cycle when reconstructing an organism's evolutionary history?
- ✓ What are some comparative advantages and disadvantages of morphological and molecular approaches for reconstructing phylogenetic trees?
- ✓ Contrast experimental and computer simulation approaches for testing the accuracy of phylogenetic reconstructions of evolutionary history. Can you think of some aspects of phylogenetic accuracy that might be more practical to test using computer simulation rather than experimental studies of viruses?

Why do biologists expend the time and effort necessary to reconstruct phylogenies? Information about the evolutionary relationships among organisms is a useful source of data for scientists investigating a wide variety of biological questions. Next we will describe how phylogenetic trees are used to answer questions about the past, and to predict and compare traits of organisms in the present.

CONCEPT 16.3 Phylogeny Makes Biology Comparative and Predictive

Once a phylogeny is reconstructed, what do we do with it? What beyond an understanding of evolutionary history does phylogeny offer us?

Phylogenies are important for reconstructing past events

Reconstructing past events is important for understanding many biological processes. In the case of zoonotic diseases (diseases caused by infectious organisms transmitted to humans from another animal host), it is important to understand when, where, and how the disease first entered a human population. Human immunodeficiency virus (HIV) is the cause of such a zoonotic disease, acquired immunodeficiency syndrome, or AIDS. Phylogenetic analyses have become important for studying the transmission of viruses such as HIV. Phylogenies are also important for understanding the present global diversity of HIV and for determining the virus's origins in human populations. A broader phylogenetic analysis of immunodeficiency viruses shows that humans acquired these viruses from two different hosts: HIV-1 from chimpanzees, and HIV-2 from sooty mangabeys (**FIGURE 16.6**).

HIV-1 is the common form of the virus in human populations in central Africa, where chimpanzees are hunted for food, and HIV-2 is the common form in human populations in western Africa, where sooty mangabeys are hunted for food. Thus it seems likely that these viruses entered human populations through hunters who cut themselves while skinning chimpanzees and sooty mangabeys. The global pandemic of AIDS occurred when these infections in local African populations rapidly spread through human populations around the world.

In recent years, phylogenetic analysis has become important in forensic investigations that involve viral transmission events. For example, phylogenetic analysis was critical for a criminal investigation of a physician who was accused of purposefully injecting blood from one of his HIV-positive patients into his former girlfriend in an attempt to kill her. The phylogenetic analysis revealed that the HIV strains present in the girlfriend were a subset of those present in the physician's patient (**FIGURE 16.7**). Other evidence was needed, of course, to connect the physician to this purposeful transmission event, but the phylogenetic analysis was important to confirm the viral transmission event from the patient to the victim.

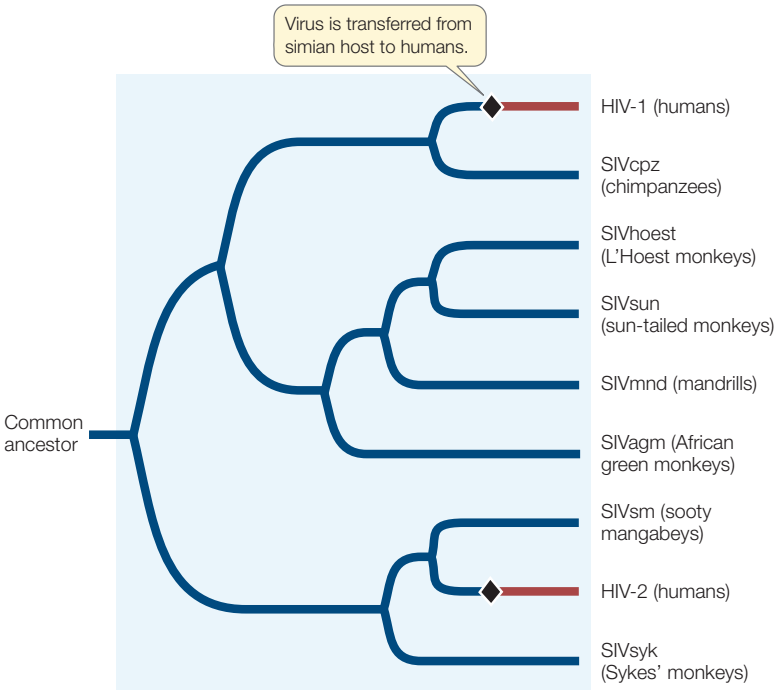


FIGURE 16.6 Phylogenetic Tree of Immunodeficiency Viruses
Immunodeficiency viruses have been transmitted to humans from two different simian hosts: HIV-1 from chimpanzees and HIV-2 from sooty mangabeys (the transmission events are marked by black diamonds). SIV stands for "simian immunodeficiency virus."

Phylogenies allow us to understand the evolution of complex traits

Male swordtails—a group of fishes in the genus *Xiphophorus*—have a long, colorful tail extension, and their reproductive success is closely associated with this appendage. Males with a long sword are more likely to mate successfully than are males with a short sword (an example of sexual selection; see Concept 15.2). Several explanations have been advanced for the evolution of this structure, including the hypothesis that the sword simply exploits a preexisting bias in the sensory system of the females. This sensory exploitation hypothesis suggests that female swordtails had a preference for males with long tails even before the tails evolved (perhaps because females assess the size of males by their total body length—including the tail—and prefer larger males).

To test the sensory exploitation hypothesis, phylogenetic analysis was used to identify the relatives of swordtails that had split most recently from their lineage before the evolution of swords. These closest relatives turned out to be fishes in the genus *Priapella*. Even though male *Priapella* do not normally have swords, when researchers attached artificial sword-like structures to the tails of male *Priapella*, female *Priapella* preferred those males. This result provided support for the hypothesis that female *Xiphophorus* had a preexisting sensory bias favoring tail extensions even before the trait evolved. Thus a long tail

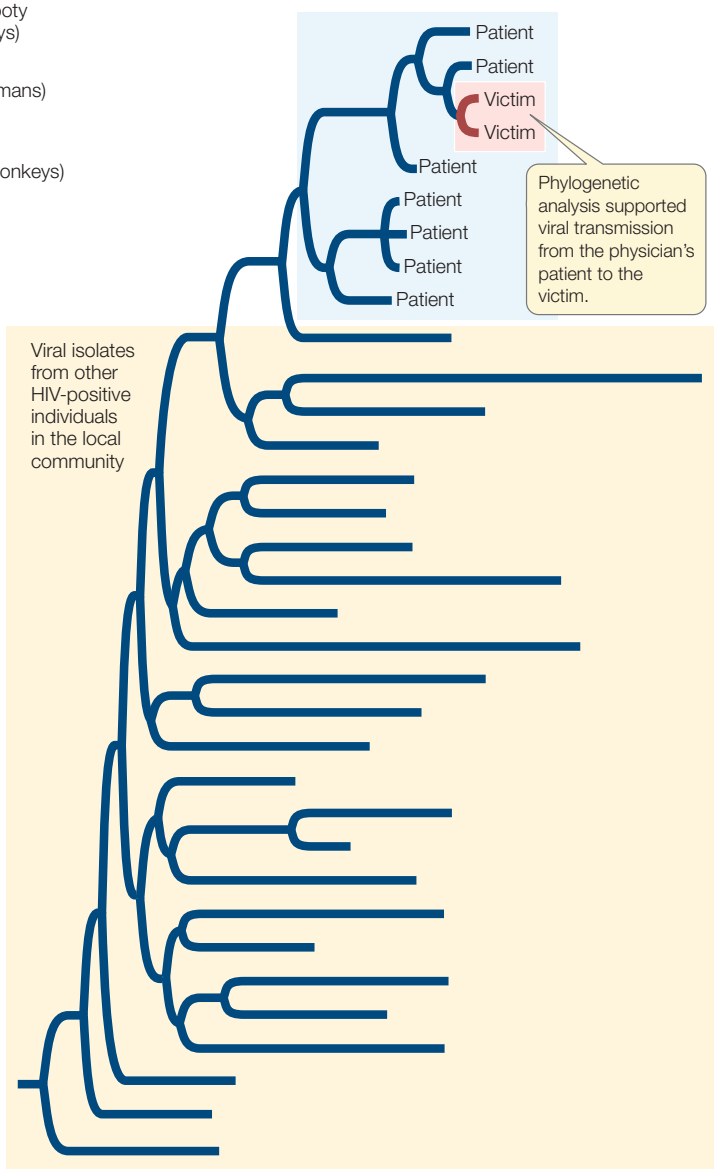


FIGURE 16.7 A Forensic Application of Phylogenetic Analysis
This phylogenetic analysis demonstrated that strains of HIV virus present in a victim (shown in red) were a phylogenetic subset of viruses isolated from a physician's patient (shown in blue). This analysis was part of the evidence used to show that the physician drew blood from his HIV-positive patient and injected it into the victim in an attempt to kill her. The physician was found guilty of attempted murder by the jury.

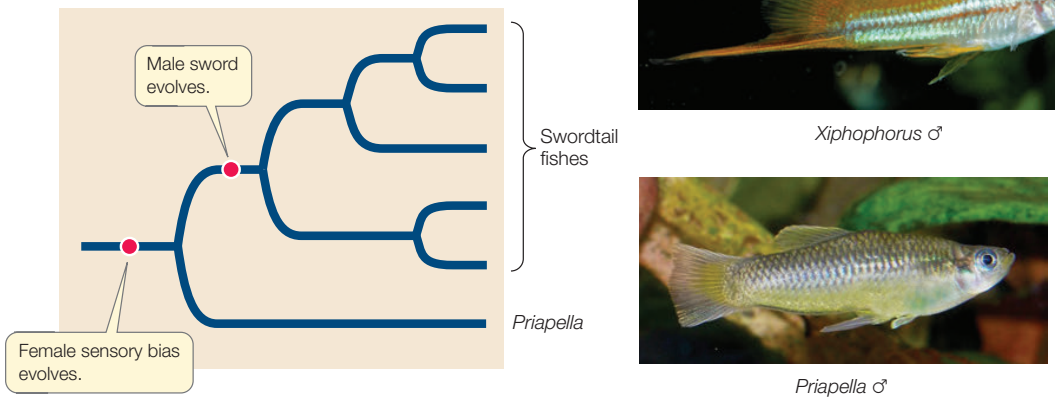


FIGURE 16.8 The Origin of a Sexually Selected Trait The long tail of male swordtail fishes (genus *Xiphophorus*) apparently evolved through sexual selection, with females mating preferentially with males with a longer "sword." Phylogenetic analysis reveals that the *Priapella* lineage split from the swordtails before the evolution of the sword. The independent finding that female *Priapella* prefer males with an artificial sword further supports the idea that this appendage evolved as a result of a preexisting preference in the females.

became a sexually selected trait because of the preexisting preference of the females (**FIGURE 16.8**).

Phylogenies can reveal convergent evolution

Like most animals, flowering plants (angiosperms) often reproduce by mating with another individual of the same species. But in many angiosperm species, the same individual produces both male and female gametes (contained within pollen and ovules, respectively). Self-incompatible species have mechanisms to prevent fertilization of the ovule by the individual's own pollen, and so must reproduce by outcrossing with another individual. Individuals of some species, however, regularly fertilize their ovules using their own pollen; they are self-fertilizing or selfing species, and their gametes are self-compatible.

LINK

Some mechanisms of self-incompatibility are discussed in **Concept 27.1**

The evolution of angiosperm fertilization mechanisms was examined in *Leptosiphon*, a genus in the phlox family that exhibits a diversity of mating systems and pollination mechanisms. The self-incompatible (outcrossing) species of *Leptosiphon* have long petals and are pollinated by long-tongued flies. In contrast, self-pollinating species have short petals and do not require insect pollinators to reproduce successfully. Using nuclear ribosomal DNA sequences, investigators reconstructed the phylogeny of this genus (**FIGURE 16.9**). They then determined whether each species was self-compatible by artificially pollinating flowers with the plant's own pollen or with pollen from other individuals and observing whether viable seeds formed.

The reconstructed phylogeny suggests that self-incompatibility is the ancestral state and that self-compatibility evolved three times within this group of *Leptosiphon*. The change to self-compatibility eliminated the plants' dependence on an outside pollinator and has been accompanied by the evolution of reduced petal size. Indeed, the striking morphological similarity of the flowers in the self-compatible groups once led to their being classified as members of a single species (*L. bicolor*). Phylogenetic analysis, however, shows them to be members of three distinct lineages. From this information we can infer that self-compatibility and its associated floral structure are convergent in the three independent lineages that had been called *L. bicolor*.

Ancestral states can be reconstructed

In addition to using phylogenetic methods to infer evolutionary relationships, biologists can use these techniques to reconstruct the morphology, behavior, or nucleotide and amino acid sequences of ancestral species (as was demonstrated for the ancestral sequence of bacteriophage T7 in Figure 16.5). In the opening of this chapter, we described how Mikhail Matz used phylogenetic analysis to reconstruct the sequence of changes in fluorescent proteins of corals to understand how red fluorescent proteins could be produced.

Reconstruction of ancient DNA sequences can also provide information about the biology of long-extinct organisms. For example, phylogenetic analysis was used to reconstruct an opsin protein in the ancestral archosaur (the most recent common ancestor of birds, dinosaurs, and crocodiles). Opsins are pigment proteins involved in vision; different opsins (with different amino acid sequences) are excited by different wavelengths of light. Knowledge of the opsin sequence in the ancestral archosaur would provide clues about the animal's visual capabilities and therefore about some of its probable behaviors. Investigators used phylogenetic analysis of opsin from living vertebrates to estimate the amino acid sequence of the pigment that existed in the ancestral archosaur. A protein with this same sequence was then constructed in the laboratory. The investigators tested the reconstructed opsin and found a significant shift toward the red end of the spectrum in the light sensitivity of this protein compared with that of most modern opsins. Modern species that exhibit similar sensitivity are adapted for nocturnal vision, so the investigators inferred that the ancestral

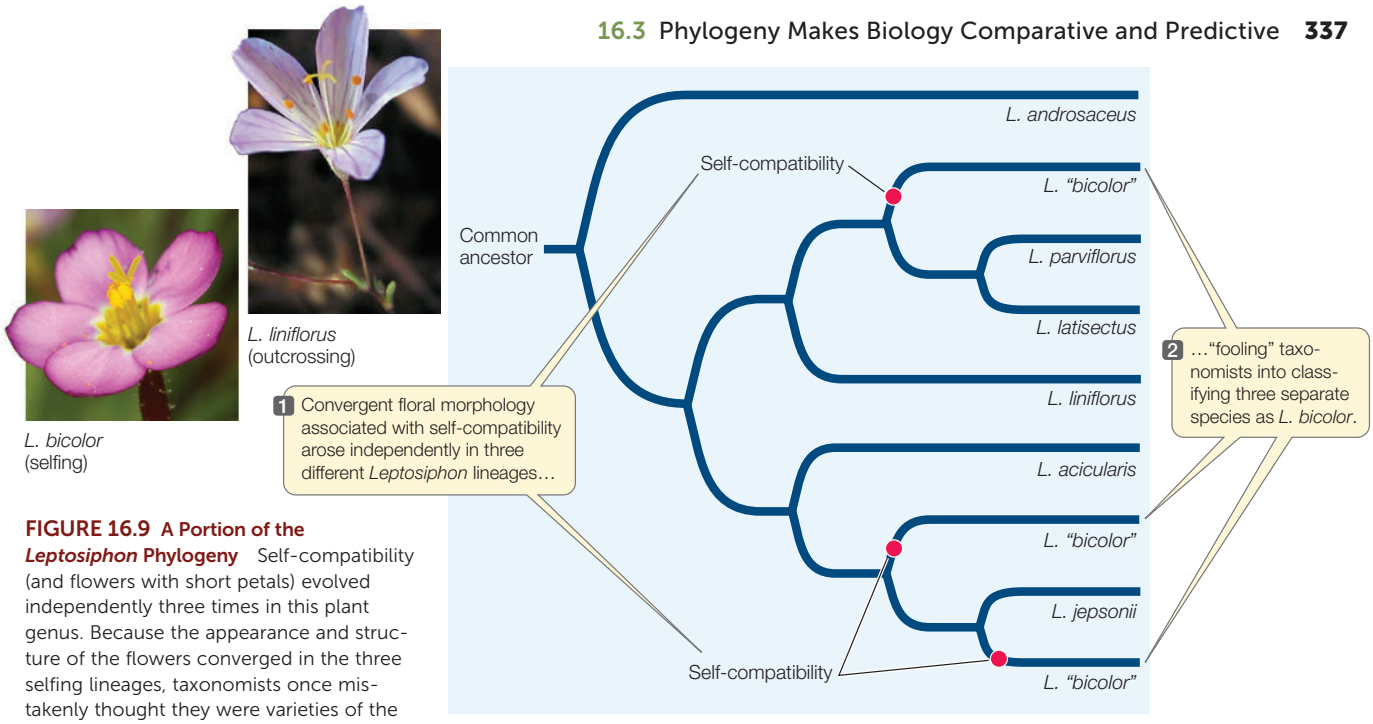


FIGURE 16.9 A Portion of the *Leptosiphon* Phylogeny Self-compatibility (and flowers with short petals) evolved independently three times in this plant genus. Because the appearance and structure of the flowers converged in the three selfing lineages, taxonomists once mistakenly thought they were varieties of the same species.

archosaur might have been active at night. Thus, reminiscent of the movie *Jurassic Park*, phylogenetic analyses are being used to reconstruct extinct species, one protein at a time.

Molecular clocks help date evolutionary events

For many applications, biologists want to know not only the order in which evolutionary lineages split but also the timing of those splits. In 1965, Emile Zuckerkandl and Linus Pauling hypothesized that rates of molecular change were constant enough that they could be used to predict evolutionary divergence times—an idea that has become known as the **molecular clock** hypothesis.

Of course, different genes evolve at different rates, and there are also differences in evolutionary rates among species related to differing generation times, environments, efficiencies of DNA repair systems, and other biological factors. Nonetheless, among closely related species, a given gene usually evolves at a reasonably constant rate. Therefore the protein encoded by the gene accumulates amino acid replacements at a relatively constant rate (**FIGURE 16.10**). A molecular clock uses the average rate at which a given gene or protein accumulates changes to gauge the time of divergence for a particular split in the phylogeny. Molecular clocks must be calibrated using independent data, such as the fossil record, known times of divergence, or biogeographic dates (e.g., the time of separations of continents). Using such calibrations, times of divergence have been estimated for many groups of species that have diverged over millions of years.

Molecular clocks are not only used to date ancient events; they are also used to study the timing of comparatively recent events. Most samples of HIV-1 have been collected from humans only since the early 1980s, although a few isolates from medical biopsies are available from as early as the 1950s. But biologists can use the observed changes in HIV-1 over the past several decades to project back to the common ancestor of all HIV-1 isolates, and estimate when HIV-1 first entered human populations from chimpanzees (**FIGURE 16.11**). This molecular clock was calibrated using the samples from the 1980s and 1990s, and then tested using the samples from the 1950s. As shown in Figure 16.11C, a sample from a 1959 biopsy is dated by molecular clock analysis at 1957 ± 10 years. Extrapolation back to

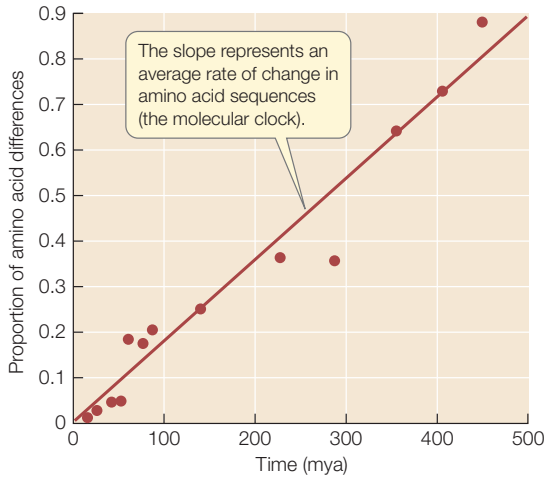


FIGURE 16.10 A Molecular Clock of the Protein Hemoglobin Amino acid replacements in hemoglobin have occurred at a relatively constant rate over nearly 500 million years of evolution. The graph shows the relationship between time of divergence and proportion of amino acid change for 13 pairs of vertebrate hemoglobin proteins. The average rate of change represents the molecular clock for hemoglobin in vertebrates.

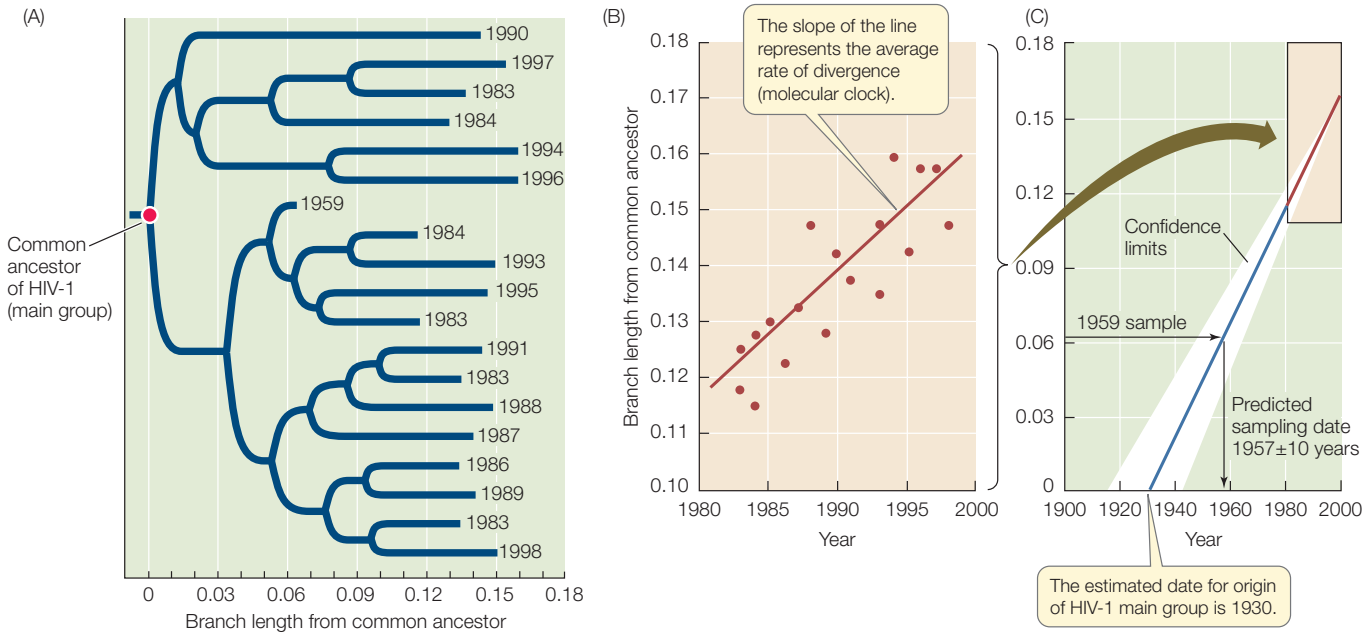


FIGURE 16.11 Dating the Origin of HIV-1 in Human Populations
(A) A phylogenetic analysis of the main group of HIV-1 viruses. The dates indicate the years in which samples were taken. (For clarity, only a small fraction of the samples that were examined in the original study are shown.) (B) A plot of year of isolation versus genetic divergence from the common ancestor provides an average rate of divergence, or a molecular clock. (C) The molecular clock is used to date a sample taken in 1959 (as a test of the clock) and the unknown date of origin of the HIV-1 main group (about 1930). Branch length from a common ancestor represents the average number of substitutions per nucleotide.

the common ancestor of the samples suggested a date of origin for this group of viruses of about 1930. Although AIDS was unknown to Western medicine until the 1980s, this analysis shows that HIV-1 was present (probably at a very low frequency) in human populations in Africa for at least a half-century before its emergence as a global pandemic. Biologists have used similar analyses to conclude that immunodeficiency viruses have been transmitted repeatedly into human populations from multiple primates for more than a century (see also Figure 16.6).

CHECKpoint CONCEPT 16.3

- ✓ How can phylogenetic trees help determine the number of times a particular trait evolved?
- ✓ How does the reconstruction of ancestral traits help biologists explain the biology of extinct species?
- ✓ What is the importance of adding a time dimension to phylogenetic trees, and how do biologists accomplish this?

All of life is connected through evolutionary history, and the relationships among organisms provide a natural basis for making biological comparisons. For these reasons, biologists use phylogenetic relationships as the basis for organizing life into a coherent classification system.

CONCEPT 16.4 Phylogeny Is the Basis of Biological Classification

The biological classification system in widespread use today is derived from that developed by the Swedish biologist Carolus Linnaeus in the mid-1700s. Linnaeus developed a system of **binomial nomenclature**. Linnaeus gave each species two names, one identifying the species itself and the other the group of closely related species, or **genus** (plural, *genera*), to which it belongs. Optionally, the name of the taxonomist who first proposed the species name may be added at the end. Thus *Homo sapiens* Linnaeus is the name of the modern human species. *Homo* is the genus, *sapiens* identifies the particular species in the genus *Homo*, and Linnaeus is the person who proposed the name *Homo sapiens*.

You can think of *Homo* as equivalent to your surname and *sapiens* as equivalent to your first name. The first letter of the genus name is capitalized, and the specific name is lowercase. Both of these formal designations are italicized. Rather than repeating the name of a genus when it is used several times in the same discussion, biologists often spell it out only once and abbreviate it with the initial letter thereafter (e.g., *D. melanogaster* rather than *Drosophila melanogaster*).

As we noted earlier, any group of organisms that is treated as a unit in a biological classification system, such as all species in the genus *Drosophila*, or all insects, or all arthropods, is called a taxon. In the Linnaean system, species and genera are further grouped into a hierarchical system of higher taxonomic

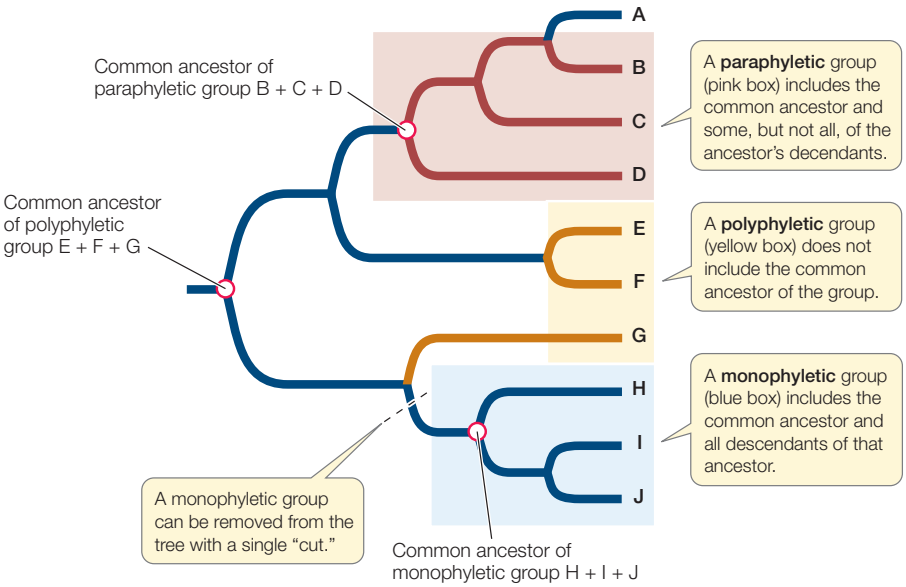


FIGURE 16.12 Monophyletic, Polyphyletic, and Paraphyletic Groups Monophyletic groups are the basis of biological taxa in modern classifications. Polyphyletic and paraphyletic groups do not accurately reflect evolutionary history.

Go to ACTIVITY 16.2 Types of Taxa
PoL2e.com/ac16.2

categories. The taxon above the genus in the Linnaean system is the **family**. The names of animal families end in the suffix "-idae." Thus Formicidae is the family that contains all ant species, and the family Hominidae contains humans and our recent fossil relatives, as well as our closest living relatives, the chimpanzees, gorillas, and orangutans. Family names are based on the name of a member genus; Formicidae is based on the genus *Formica*, and Hominidae is based on *Homo*. The same rules are used in classifying plants, except that the suffix "-aceae" is used for plant family names instead of "-idae." Thus Rosaceae is the family that includes the genus *Rosa* (roses) and its relatives.

In the Linnaean system, families are grouped into **orders**, orders into **classes**, and classes into **phyla** (singular *phylum*), and phyla into **kingdoms**. However, the ranking of taxa within Linnaean classification is subjective. Whether a particular taxon is considered, say, an order or a class is informative only with respect to the *relative* ranking of other related taxa. Although families are always grouped within orders, orders within classes, and so forth, there is nothing that makes a "family" in one group equivalent (in number of genera or in evolutionary age, for instance) to a "family" in another group.

Today the Linnaean terms above the genus level are used largely for convenience. Linnaeus recognized the overarching hierarchy of life, but he developed his system before evolutionary thought had become widespread. Biologists today recognize the tree of life as the basis for biological classification and often name taxa without placing them into the various Linnaean ranks. But regardless of whether they rank organisms into Linnaean categories or use unranked taxon names, modern biologists use evolutionary relationships as the basis for distinguishing, naming, and classifying biological groups.

Evolutionary history is the basis for modern biological classification

Today's biological classifications express the evolutionary relationships of organisms. Taxa are expected to be **monophyletic**, meaning that the taxon contains an ancestor and all descendants of that ancestor, and no other organisms. In other words, a monophyletic taxon is a historical group of related species, or a complete branch on the tree of life. As noted earlier, this is also the definition of a clade. A true monophyletic group can be removed from a phylogenetic tree by a single "cut" in the tree, as shown in FIGURE 16.12.

Note that there are many monophyletic groups on any phylogenetic tree, and that these groups are successively smaller subsets of larger monophyletic groups. This hierarchy of biological taxa, with all of life as the most inclusive taxon and many smaller taxa within larger taxa, down to the individual species, is the modern basis for biological classification.

Although biologists seek to describe and name only monophyletic taxa, the detailed phylogenetic information needed to do so is not always available. A group that does not include its common ancestor is **polyphyletic**. A group that does not include all the descendants of a common ancestor is referred to as **paraphyletic** (see Figure 16.12). Virtually all taxonomists now agree that polyphyletic and paraphyletic groups are inappropriate as taxonomic units because they do not correctly reflect evolutionary history. Some classifications still contain such groups because some organisms have not been evaluated phylogenetically. As mistakes in prior classifications are detected, taxonomic names are revised and polyphyletic and paraphyletic groups are eliminated from the classifications.

Several codes of biological nomenclature govern the use of scientific names

Several sets of explicit rules govern the use of scientific names. Biologists around the world follow these rules voluntarily to facilitate communication and dialogue. There may be dozens of common names for an organism in many different languages, and the same common name may refer to more than one species (FIGURE 16.13). The rules of biological nomenclature are designed so that there is only one correct scientific name for

FIGURE 16.13 Same Common Name, Not the Same Species
All three of these animals are known locally as anteaters. Unique scientific binomials allow biologists to communicate unambiguously about each species. (A) *Myrmecophaga tridactyla*, the giant anteater, searching for termites in Brazil. (B) *Tachyglossus aculeatus*, an echidna, is also called the spiny anteater. (C) *Orycteropus afer*, the aardvark, is also known as the Cape anteater.

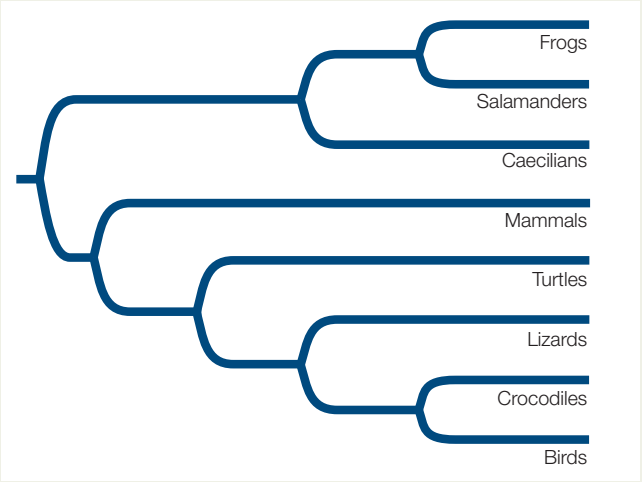
(A) *Myrmecophaga tridactyla*

(B) *Tachyglossus aculeatus*

(C) *Orycteropus afer*

APPLY THE CONCEPT

Phylogeny is the basis of biological classification



Classification One:

Named group	Included taxa
Amphibia	Frogs, salamanders, and caecilians
Mammalia	Mammals
Reptilia	Turtles, lizards, and crocodiles
Aves	Birds

Classification Two:

Named group	Included taxa
Amphibia	Frogs, salamanders, and caecilians
Mammalia	Mammals
Reptilia	Turtles, lizards, crocodiles, and birds

Classification Three:

Named group	Included taxa
Amphibia	Frogs, salamanders, and caecilians
Homothermia	Mammals and birds
Reptilia	Turtles, lizards, and crocodiles

Consider the above phylogeny and three possible classifications of the living taxa.

1. Which of these classifications contains a paraphyletic group?
2. Which of these classifications contains a polyphyletic group?
3. Which of these classifications is consistent with the goal of including only monophyletic groups in a biological classification?
4. Starting with the classification you named in Question 3, how many additional group names would you need to include all the clades shown in this phylogenetic tree?

any single recognized taxon, and (ideally) a given scientific name applies only to a single taxon (that is, each scientific name is unique). Sometimes the same species is named more than once (when more than one taxonomist has taken up the task). In these cases, the rules specify that the valid name is the first name that was proposed. If the same name is inadvertently given to two different species, then the species that was named second must be given a new name.

Because of the historical separation of the fields of zoology, botany (which originally included mycology, the study of fungi), and microbiology, different sets of taxonomic rules were developed for each of these groups. Yet another set of rules emerged later for classifying viruses. This separation of fields resulted in duplicated taxon names in groups governed by the different sets of rules. *Drosophila*, for example, is both a genus of fruit flies and a genus of fungi, and some species in both groups have identical names. Until recently these duplicated names caused little confusion, since traditionally biologists who studied fruit flies were unlikely to read the literature on fungi (and vice versa). Today, given the prevalence of large, universal biological databases (such as GenBank, which includes DNA sequences from across all life), it is increasingly important that each taxon have a unique and unambiguous name. Biologists are working on a universal code of nomenclature that can be applied to all organisms, so that every species will have a unique identifying name or registration number. This will assist efforts to build an online *Encyclopedia of Life* that links all the information for all the world's species.

CHECKpoint CONCEPT 16.4

- ✓ What is the difference between monophyletic, paraphyletic, and polyphyletic groups?
- ✓ Why do biologists prefer monophyletic groups in formal classifications?
- ✓ What advantages or disadvantages do you see to having separate sets of taxonomic rules for animals, plants, bacteria, and viruses?

Having described some of the mechanisms by which evolution occurs and how phylogenies can be used to study evolutionary relationships, we are now ready to consider the subject of speciation. Speciation is the process that leads to the splitting events (nodes) on the tree of life and eventually results in the millions of distinctive species that constitute Earth's biodiversity.

Q How are phylogenetic methods used to resurrect protein sequences from extinct organisms?

ANSWER Most genes and proteins of organisms that lived millions of years ago have decomposed in the fossil remains of these species. Nonetheless, sequences of many ancient genes and proteins can be reconstructed. Just as we can reconstruct the morphological features of a clade's ancestors, we can also reconstruct the ancestors' DNA and protein sequences—if we have enough information about the genomes of their descendants (Concept 16.3).

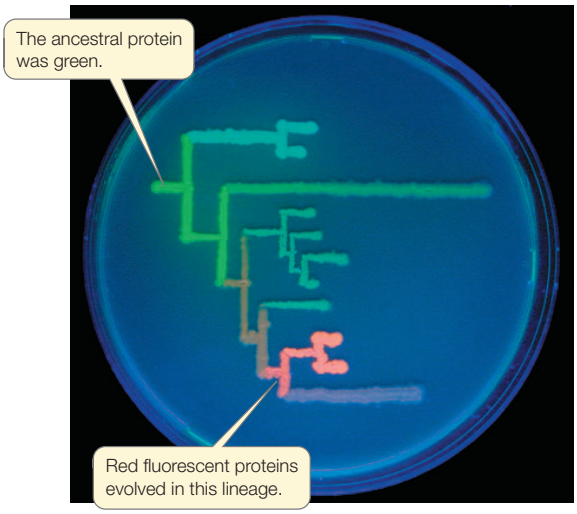


FIGURE 16.14 Evolution of Fluorescent Proteins of Corals Mikhail Matz and his colleagues used phylogenetic analysis to reconstruct the sequences of extinct fluorescent proteins that were present in the ancestors of modern corals. They then expressed these proteins in bacteria and plated the bacteria in the form of a phylogenetic tree to show how the colors evolved over time.

Biologists have now reconstructed gene sequences from many species that have been extinct for millions of years. Using this information, a laboratory can reconstruct real proteins that correspond to the proteins that were present in long-extinct species. This is how Mikhail Matz and his colleagues were able to resurrect fluorescent proteins from the extinct ancestors of modern corals and then visualize the colors produced by these proteins in the laboratory and recreate the probable evolutionary path of the different pigments (**FIGURE 16.14**).

Biologists have even used phylogenetic analysis to reconstruct some protein sequences that were present in the common ancestor of life (Concept 16.1). These hypothetical protein sequences can then be resurrected into actual proteins in the laboratory. When biologists measured the temperature optima for these resurrected proteins, they found that the proteins functioned best in the range of 55°C–65°C. This analysis is consistent with hypotheses that life evolved in a high-temperature environment.

To reconstruct protein sequences from species that have been extinct for millions or even billions of years, biologists use detailed mathematical models that take into account much of what we have learned of molecular evolution (Concept 16.2). These models incorporate information on rates of replacement among different amino acid residues in proteins, information on different substitution rates among nucleotides, and changes in the rate of molecular evolution among the major lineages of life.

SUMMARY

CONCEPT 16.1 All of Life Is Connected through Its Evolutionary History

- **Phylogeny** is the history of descent of organisms from their common ancestor. Groups of evolutionarily related species are represented as related branches in a **phylogenetic tree**.
- A group of species that consists of a common ancestor and all its evolutionary descendants is called a **clade**. Named clades and species are called **taxa**. Review Figure 16.1
- **Homologies** are similar traits that have been inherited from a common ancestor.
- A trait that is shared by two or more taxa and is derived through evolution from a common ancestral form is called a **synapomorphy**.
- Similar traits may occur among species that do not result from common ancestry. **Convergent evolution** and **evolutionary reversals** can give rise to such traits, which are called **homoplasies**. Review Figure 16.2

CONCEPT 16.2 Phylogeny Can Be Reconstructed from Traits of Organisms

- Phylogenetic trees can be inferred from synapomorphies and using the **parsimony principle**. Review Figure 16.3 and ACTIVITY 16.1
- Sources of phylogenetic information include morphology, patterns of development, the fossil record, behavioral traits, and molecular traits such as DNA and protein sequences. Review ANIMATED TUTORIAL 16.1
- Phylogenetic trees can be inferred with **maximum likelihood** methods, which calculate the probability that a particular tree will have generated the observed data. Review ANIMATED TUTORIAL 16.2

CONCEPT 16.3 Phylogeny Makes Biology Comparative and Predictive

- Phylogenetic trees are used to reconstruct past events. Review Figures 16.6 and 16.7
- Biologists can use phylogenetic trees to reconstruct ancestral states. Review Figure 16.8
- Phylogenetic trees are used to reveal convergent evolution. Review Figure 16.9
- Phylogenetic trees may include estimates of times of divergence of lineages determined by **molecular clock** analysis. Review Figures 16.10 and 16.11

CONCEPT 16.4 Phylogeny Is the Basis of Biological Classification

- Taxonomists organize biological diversity on the basis of evolutionary history.
- Taxa in modern classifications are expected to be clades, or **monophyletic** groups. **Paraphyletic** and **polyphyletic** groups are not considered appropriate taxonomic units. Review Figure 16.12 and ACTIVITY 16.2
- Several sets of rules govern the use of scientific names, with the goal of providing unique and universal names for biological taxa.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is16

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

17 Speciation

KEY CONCEPTS

This composite photograph shows several of the hundreds of species of diverse haplochromine cichlids that are endemic to Lake Malawi.

Not quite 2 million years ago, a tectonic split in the Great Rift Valley of East Africa led to the formation of Lake Malawi, which lies between the modern countries of Malawi, Tanzania, and Mozambique. A few fish species entered the new lake, including a type known as a haplochromine cichlid. Today the descendants of this early colonizer include nearly 1,000 species of haplochromine cichlids. All of them are endemic to Lake Malawi—they are found nowhere else. This vast array of cichlid species makes this the most diverse lake in the world in terms of its fish community. How did so many different species arise from a single ancestral species in less than 2 million years?

As we noted in Chapter 16, speciation is the process that produces the splits among lineages in the tree of life. Biologists have studied the history and timing of speciation events in Lake Malawi and have pieced together some of the processes that led to so many cichlid species. The earliest haplochromine cichlids to enter the new lake encountered diverse habitats in Lake Malawi, as some shores were rocky and others were sandy. Cichlid populations quickly adapted to these distinct habitat types. Fish in rocky habitats adapted to breeding and living in rocky conditions, and those in sandy habitats evolved specializations for life over sand. These changes resulted in an early speciation event.

Within each of these major habitat types, there were numerous opportunities for diet specialization. Various populations of cichlids became rock scrapers, bottom feeders, fish predators, scale biters, pelagic zooplankton eaters, and plant specialists. Each of these feeding specializations requires a different mouth morphology. The offspring of fish that bred with fish of similar morphology were more likely to survive than were fish with two very different parents. These differences in fitness led to the formation of many more new species, each adapted to a different feeding mode.

But still the Lake Malawi cichlids continued to diverge and form new species. Male cichlids compete for the attention of females through their bright body colors. Diversification of the body colors of males, and of the preferences of females for different body colors, led to many more new species of cichlids, each isolated from the other by their sexual preferences. Now biologists are studying the genomes of these Lake Malawi cichlids to understand the details of the genetic changes that have given rise to so many species over so little time.

Q Can biologists study the process of speciation in the laboratory?

You will find the answer to this question on page 356.

CONCEPT 17.1 Species Are Reproductively Isolated Lineages on the Tree of Life

Biological diversity does not vary in a smooth, incremental way. People have long recognized groups of similar organisms that mate with one another, and they have noticed that there are usually distinct morphological breaks between these groups. Groups of organisms that mate with one another are commonly called **species** (note that this is both the plural and singular form of the word). Species are the result of the process of **speciation**: the divergence of biological lineages and the emergence of reproductive isolation between lineages.

Although "species" is a useful and common term, its usage varies among biologists who are interested in different aspects of speciation. Different biologists think about species differently because they ask different questions: How can we recognize and identify species? How do new species arise? How do different species remain separate? Why do rates of speciation differ among groups of organisms? In answering these questions, biologists focus on different attributes of species, leading to several different ways of thinking about what species are and how they form. Most of the various **species concepts** proposed by biologists are simply different ways of approaching the question "What are species?" Let's compare three major classes of species concepts to contrast the way that biologists think about species.

We can recognize many species by their appearance

Someone who is knowledgeable about a group of organisms, such as birds or flowering plants, can usually distinguish the different species found in a particular area simply by looking at them. Standard field guides to birds, mammals, insects, and wildflowers are possible only because many species change little in appearance over large geographic distances (FIGURE 17.1A).

More than 250 years ago, Carolus Linnaeus developed the system of binomial nomenclature by which species are named today (see Concept 16.4). Linnaeus described and named thousands of species, but because he knew nothing about genetics or the mating behavior of the organisms he was naming, he classified them on the basis of their appearance alone. In other words, Linnaeus used a **morphological species concept**, a construct that assumes a species comprises individuals that look alike, and that individuals that do not look alike belong to different species. Although Linnaeus did not know it, the members of most of the groups he classified as species look alike because they share many alleles of genes that code for morphological features.

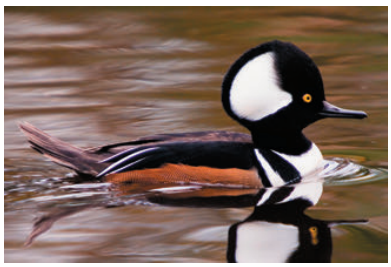
Using morphology to define species has limitations. Members of the same species do not always look alike. For example, males, females, and young individuals do not always resemble one another closely (FIGURE 17.1B). Furthermore, morphology is of little use in the case of cryptic species—instances in which two or more species are morphologically indistinguishable but do not interbreed (FIGURE 17.2). Biologists therefore cannot rely on appearance alone in determining whether individual organisms are members of the same or different species. Today biologists use several additional types of information—especially behavioral and genetic data—to differentiate species.

Reproductive isolation is key

The most important factor in the long-term isolation of sexually reproducing lineages from one another is the evolution of **reproductive isolation**, a state in which two groups of organisms can no longer exchange genes. If individuals of group "A" mate and reproduce only with one another, group "A" constitutes a distinct species within which genes recombine. In other words, group "A" is an independent evolutionary lineage—a separate branch on the tree of life.

Evolutionary biologist Ernst Mayr recognized the importance of reproductive isolation in maintaining species, and so he proposed the **biological species concept**: "*Species are groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups*." The phrase "actually or potentially" is an important element of this definition. "Actually" says that the individuals live in the same area and

(A)



Lophodytes cucullatus
Male, British Columbia

Lophodytes cucullatus
Male, New Mexico

(B)

Lophodytes cucullatus
Female

FIGURE 17.1 Members of the Same Species Look Alike—or Not (A) It is easy to identify these two male hooded mergansers as members of the same species, even though they were photographed thousands of miles apart in British Columbia and New Mexico, respectively. Despite their geographic separation, the two individuals are morphologically very similar. (B) Hooded mergansers are sexually dimorphic, which means the female's appearance is quite different from that of the male.

(A)

Hyla versicolor

(B)

Hyla chrysoscelis

FIGURE 17.2 Cryptic Species Look Alike but Do Not Interbreed These two species of gray treefrogs (*Hyla versicolor* and *H. chrysoscelis*) cannot be distinguished by their external morphology, but they do not interbreed even when they occupy the same geographic range. *H. versicolor* is a tetraploid species, whereas *H. chrysoscelis* is diploid. Although they look alike, the males have distinctive mating calls, and based on these calls, the females recognize and mate with males of their own species.

interbreed with one another. "Potentially" says that even though the individuals do not live in the same area, and therefore do not interbreed, other information suggests that they *would* do so if they were able to get together. This widely used species concept does not apply to organisms that reproduce asexually, and it is limited to a single point in evolutionary time.

The lineage approach takes a long-term view

Evolutionary biologists often think of species as branches on the tree of life. This idea can be termed a **lineage species concept**. In this framework for thinking about species, one species splits into two or more daughter species, which thereafter evolve as distinct lineages. A lineage concept allows biologists to consider species over evolutionary time.

A **lineage** is an ancestor–descendant series of populations followed over time. Each species has a history that starts with a speciation event by which one lineage on the tree is split into two, and ends either at extinction or at another speciation event, at which time the species produces two daughter species. The process of lineage splitting may be gradual, taking thousands of generations to complete. At the other extreme, an ancestral lineage may be split in two within a few generations (as happens with polyploidy, which we'll discuss in Concept 17.3). The gradual nature of some splitting events means that at a single point in time, the final outcome of the process may not be clear. In these cases, it may be difficult to predict whether the incipient species (lineages in an initial stage of development) will continue to diverge and become fully isolated from one another, or if they will merge again in the future.

The different species concepts are not mutually exclusive

Many named variants of these three major classes of species concepts exist. The various concepts are not entirely incompatible, however. Rather, they simply emphasize different aspects of species or speciation. The morphological species concept emphasizes the practical aspects of recognizing species, although it sometimes results in underestimation or overestimation of the actual number of species. Mayr's biological species concept emphasizes that reproductive isolation is what allows sexual species to evolve independently of one another. Lineage species concepts embrace the idea that sexual species are maintained by reproductive isolation, but extend the concept of a species as a lineage over evolutionary time. The species-as-lineage concept is also able to accommodate species that reproduce asexually.

Virtually all species exhibit some degree of genetic recombination among individuals, even if recombination events are relatively rare. Significant reproductive isolation between species is therefore necessary for lineages to remain distinct over evolutionary time. Furthermore, reproductive isolation is responsible for the morphological distinctiveness of most species, because mutations that result in morphological changes cannot spread between reproductively isolated species. Therefore no matter which species concept we emphasize, the evolution of reproductive isolation is important for understanding the origin of species.

CHECKpoint CONCEPT 17.1

- ✓ Why do different biologists emphasize different attributes of species in formulating species concepts?
- ✓ What makes reproductive isolation such an important component of each of the species concepts discussed here?
- ✓ Why is the biological species concept not applicable to asexually reproducing organisms? How do biologists consider the species status of asexual organisms?

Although Charles Darwin titled his groundbreaking book *On the Origin of Species*, in fact it included very little about speciation as we understand it today. Darwin devoted most of his attention to demonstrating that individual species are altered over time by natural selection. The following sections discuss the many aspects of speciation that biologists have learned about since Darwin's time.

CONCEPT 17.2 Speciation Is a Natural Consequence of Population Subdivision

Not all evolutionary changes result in new species. A single lineage may change over time without giving rise to a new species. Speciation requires the interruption of gene flow within a species whose members formerly exchanged genes. But if a genetic change prevents reproduction between individuals of a species, how can such a change spread through a species in the first place?

Incompatibilities between genes can produce reproductive isolation

If a new allele that causes reproductive incompatibility arises in a population, it cannot spread through the population because no other individuals will be reproductively compatible with the individual that carries the new allele. So how can one reproductively cohesive lineage ever split into two reproductively isolated species? Several early geneticists, including Theodosius Dobzhansky and Hermann Joseph Muller, developed a genetic model to explain this apparent conundrum (**FIGURE 17.3**).

The Dobzhansky–Muller model is quite simple. First, assume that a single ancestral population is subdivided into two daughter populations (by the formation of a new mountain range, for instance), which then evolve as independent lineages. In one of the descendant lineages, a new allele (*A*) arises and becomes fixed (see Figure 17.3). In the other population, another new allele (*B*) becomes fixed *at a different gene locus*. Neither new allele at either locus results in any loss of reproductive compatibility. However, the two new forms of these two different genes have never occurred together in the same individual or population. Recall that the products of many genes must work together in an organism. It is possible that the new protein forms encoded by the two new alleles will not be compatible with each other. If individuals from the two lineages come back together after these genetic changes, they may still be able to interbreed. However, the hybrid offspring may have a new combination of genes that is functionally inferior, or even lethal. This will not happen with all new combinations of genes, but over time, isolated lineages will accumulate many allele differences at many gene loci. Some combinations of these differentiated genes will not function well together in hybrids. Thus genetic incompatibility between the two isolated populations will develop over time.

Many empirical and experimental examples support the Dobzhansky–Muller model. This model works not only for pairs of individual genes but also for some kinds of chromosomal rearrangements. Bats of the genus *Rhogeessa*, for example, exhibit considerable variation in centric fusions of their chromosomes. The chromosomes of the various species contain the same basic chromosomal arms, but in some species two acrocentric (one-armed) chromosomes have fused at the centromere to form larger, metacentric (two-armed) chromosomes. A polymorphism in centric fusion causes few, if any, problems in meiosis because the respective chromosomes can still align and assort normally. Therefore a given centric fusion can become fixed in a lineage. However, if a *different* centric fusion becomes fixed in a second lineage, then hybrids between individuals of each lineage will not be able to produce normal gametes in meiosis (**FIGURE 17.4**). Most of the closely related species of *Rhogeessa* display different combinations of these centric fusions and are thereby reproductively isolated from one another.

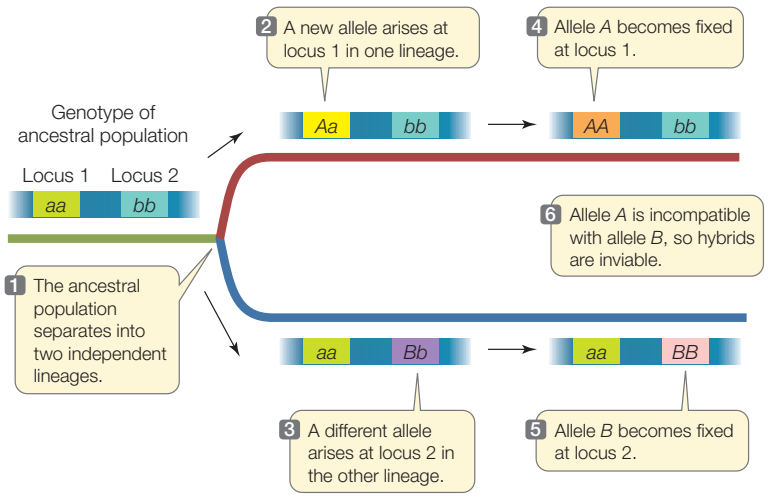


FIGURE 17.3 The Dobzhansky–Muller Model In this simple two-locus version of the model, two lineages from the same ancestral population become separated from each other and evolve independently. A new allele becomes fixed in each descendant lineage, but at two different genes. Neither of the new alleles is incompatible with the ancestral alleles, but the two new alleles in the two different genes are incompatible with each other. Thus the two descendant lineages are reproductively incompatible.

Go to ANIMATED TUTORIAL 17.1
Speciation Simulation
PoL2e.com/at17.1

Reproductive isolation develops with increasing genetic divergence

As pairs of species diverge genetically, they become increasingly reproductively isolated (**FIGURE 17.5**). Both the rate at which reproductive isolation develops and the mechanisms that produce it vary from group to group. Reproductive incompatibility has been shown to develop gradually in many groups of plants, animals, and fungi, reflecting the slow pace at which incompatible genes accumulate in each lineage. In some cases, complete reproductive isolation may take millions of years. In other cases (as with the chromosomal fusions of *Rhogeessa* described above), reproductive isolation can develop over just a few generations.

Partial reproductive isolation has evolved in strains of *Phlox drummondii* artificially isolated by humans. In 1835, Thomas Drummond, after whom this species of garden plant is named, collected seeds in Texas and distributed them to nurseries in Europe.

Rhogeessa tumida

FIGURE 17.4 Speciation by Centric Fusion In this chromosomal version of the Dobzhansky–Muller model, two independent centric fusions of one-armed chromosomes occur in two sister lineages of the bat genus *Rhogeessa*. Neither centric fusion by itself results in difficulties at meiosis, whether the fusion is found in just one or in both pairs of chromosomes. After fixation of the respective fusions in each lineage, however, F_1 hybrids between the two lineages are sterile, because the three different chromosomes involved in these centric fusions cannot pair normally at meiosis in hybrids. Most of the species in the bat genus *Rhogeessa* differ from one another by such centric fusions.

The European nurseries established more than 200 true-breeding strains of *P. drummondii* that differed in flower size, flower color, and plant growth form. The breeders did not select directly for reproductive incompatibility between strains, but in subsequent experiments in which strains were crossed and seed production was measured and compared, biologists found that reproductive compatibility between strains had been reduced by 14 to 50 percent, depending on the cross—even though the strains had been isolated from one another for less than two centuries.

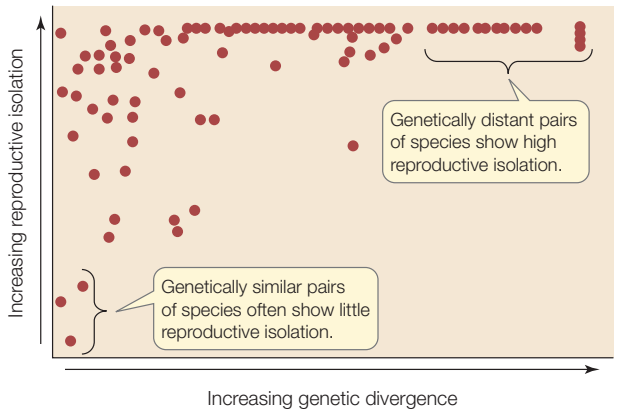


FIGURE 17.5 Reproductive Isolation Increases with Genetic Divergence Among pairs of *Drosophila* species, the more the species differ genetically, the greater their reproductive isolation from each other. Each dot represents a comparison of one species pair. Such positive relationships between genetic distance and reproductive isolation have been observed in many groups of plants, animals, and fungi.

CHECKpoint CONCEPT 17.2

- ✓ The Dobzhansky–Muller model suggests that divergence among alleles at *different* gene loci leads to genetic incompatibility between species. Why is genetic incompatibility between two alleles at the *same* locus considered less likely?
- ✓ Why do some combinations of chromosomal centric fusions cause problems in meiosis? Can you diagram what would happen at meiosis in a hybrid of the divergent lineages shown in Figure 17.4?
- ✓ Assume that the reproductive isolation seen in *Phlox* strains results from lethal combinations of incompatible alleles at several loci among the various strains. Given this assumption, why might the reproductive isolation seen among these strains be partial rather than complete?

We have now seen how splitting an ancestral population leads to genetic divergence in the two descendant lineages. Next we will consider ways in which the descendant lineages could have become separated in the first place.

CONCEPT 17.3 Speciation May Occur through Geographic Isolation or in Sympatry

Many scientists who study speciation have concentrated on geographic processes that result in the splitting of an ancestral species. Splitting the range of a species is one obvious way of achieving such a division, but it is not the only way.

Physical barriers give rise to allopatric speciation

Speciation that results when a population is divided by a physical barrier is known as **allopatric speciation** (Greek *allos*, "other"; *patria*, "homeland"). Allopatric speciation is thought to be the dominant mode of speciation in most groups of organisms. The physical barrier that divides the range of a species may be a body of water or a mountain range for terrestrial organisms, or dry land for aquatic organisms—in other words, any type of habitat that is inhospitable to the species. Such barriers can form when continents drift, sea levels rise and fall, glaciers advance and retreat, or climates change. The populations separated by such barriers are often, but not always, initially large. The lineages that descend from these founding populations evolve differences for a variety of reasons, including mutation, genetic drift, and adaptation to different environments in the two areas. As a result, many pairs of closely related **sister species**—species that are each other's closest relatives—may exist on opposite sides of the geographic barrier. An example of a physical geographical barrier that produced many pairs of sister species was the Pleistocene glaciation that isolated freshwater streams in the eastern highlands of the Appalachian Mountains from streams in the Ozark and Ouachita mountains about 20,000–80,000 years ago (FIGURE 17.6). This splitting event resulted in many parallel speciation events among isolated lineages of stream-dwelling organisms.

Allopatric speciation also may result when some members of a population cross an existing barrier and establish a new, isolated population. The 14 species of finches found in the Galápagos archipelago some 1,000 kilometers off the coast of Ecuador are the result of this process. Darwin's finches (as they are usually called, because Darwin was the first scientist to study them) arose in the Galápagos from a single South American finch species that colonized the islands. Today the Galápagos species differ strikingly from their closest mainland relative and from one another (FIGURE 17.7). The islands of the archipelago are sufficiently far apart that the finches move among them only infrequently. In addition, environmental conditions differ widely from island to island. Some islands are relatively flat and arid; others have forested mountain slopes. Over millions of years, finch lineages on the different islands have differentiated to the point that when occasional immigrants arrive from other islands, they either do not breed with the residents, or if they do, the resulting offspring do not survive as well as the offspring of established residents. The genetic distinctness of each finch species from the others and the genetic cohesiveness of the individual species are thus maintained.

Go to ANIMATED TUTORIAL 17.2
Founder Events and Allopatric Speciation
PoL2e.com/at17.2

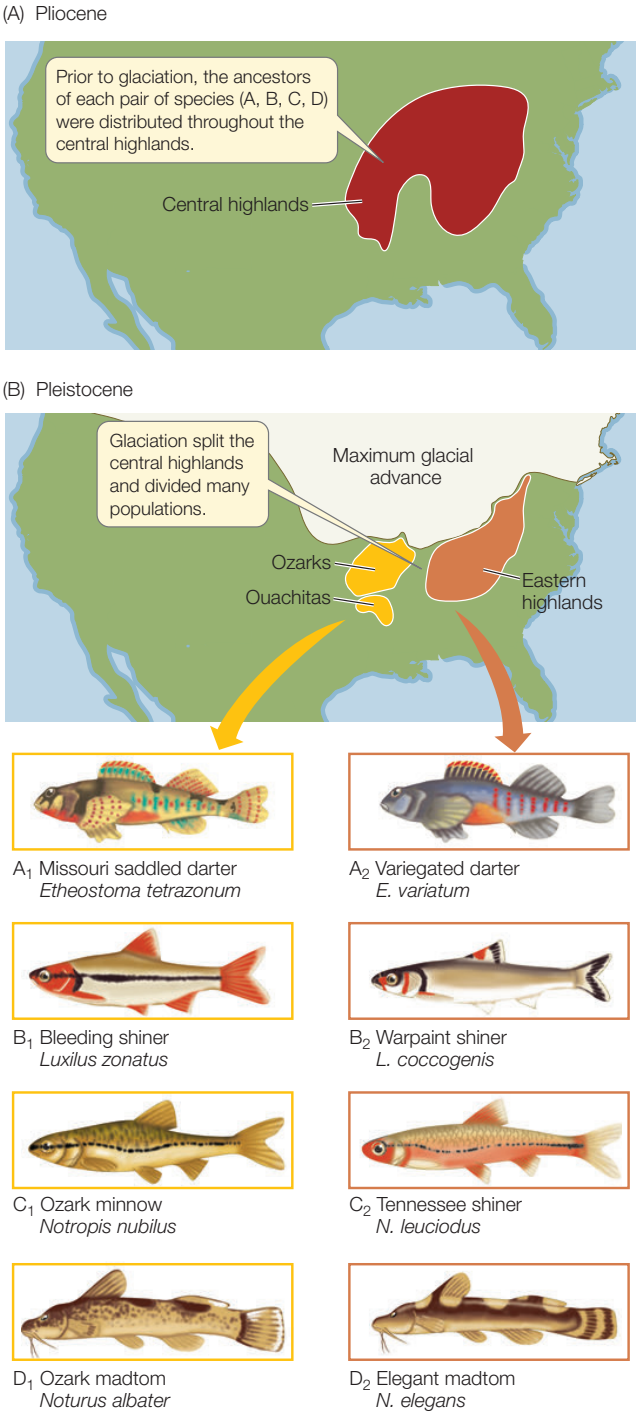


FIGURE 17.6 Allopatric Speciation Allopatric speciation may result when an ancestral population is divided into two separate populations by a physical barrier, and the lineages that descend from these populations then diverge. (A) Many species of freshwater stream fishes were distributed throughout the central highlands of North America in the Pliocene epoch (about 3–5 million years ago). (B) During the Pleistocene, glaciers advanced and isolated fish populations in the Ozark and Ouachita mountains to the west from fish populations in the highlands to the east. Numerous species diverged as a result of this separation, including the ancestors of the four species pairs shown here.

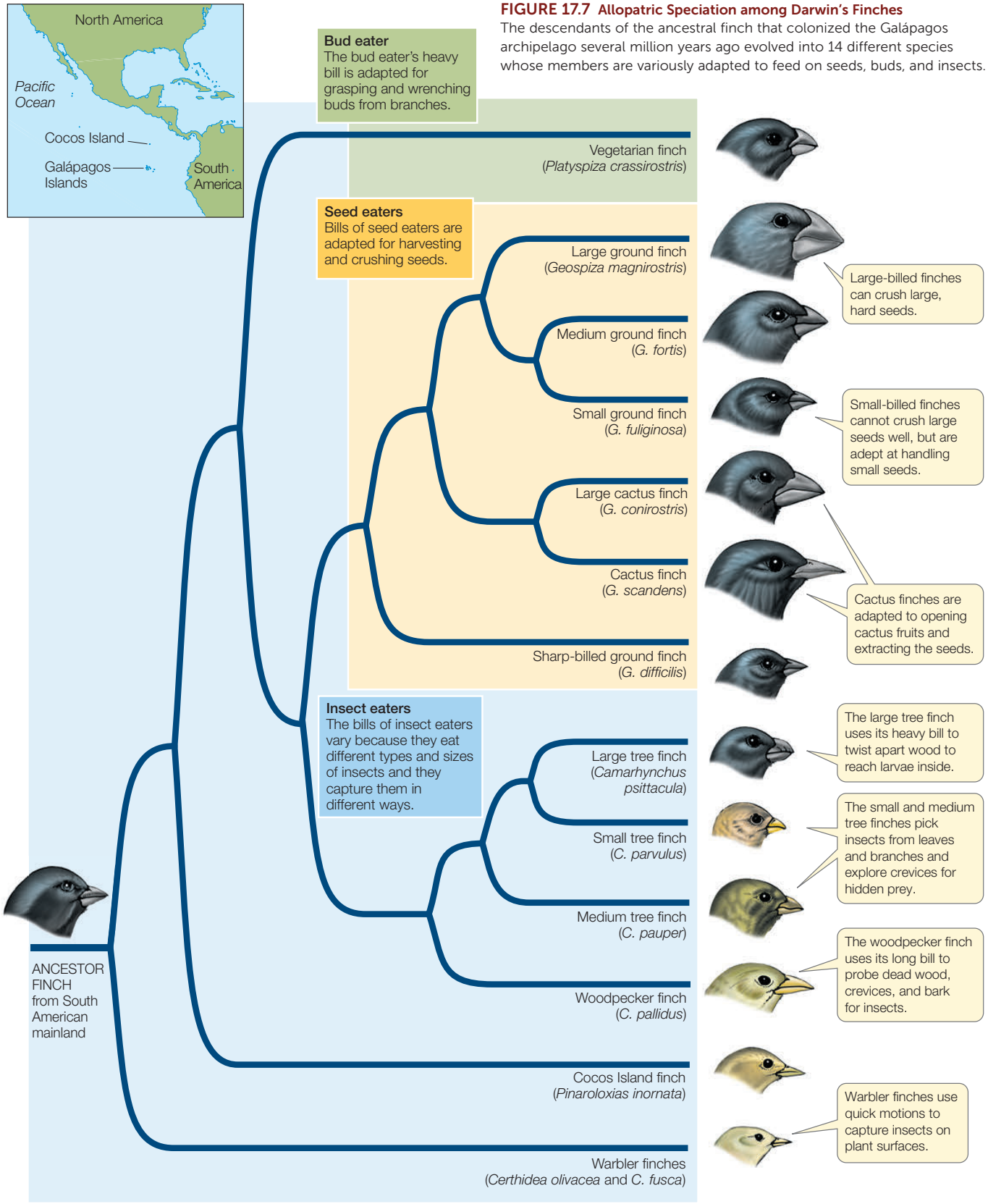


FIGURE 17.7 Allopatric Speciation among Darwin's Finches
The descendants of the ancestral finch that colonized the Galápagos archipelago several million years ago evolved into 14 different species whose members are variously adapted to feed on seeds, buds, and insects.

APPLY THE CONCEPT

Speciation may occur through geographic isolation

The different species of Darwin's finches shown in the phylogeny in Figure 17.7 have all evolved on islands of the Galápagos archipelago within the past 3 million years. Molecular clock analysis (see Concept 16.3) has been used to determine the dates of the various speciation events in that phylogeny. Geological techniques for dating rock samples (see Concept 18.1) have been used to determine the ages of the various Galápagos islands. The table shows the number of species of Darwin's finches and the number of islands that have existed in the archipelago at several times during the past 4 million years.

TIME (mya)	NUMBER OF ISLANDS	NUMBER OF FINCH SPECIES
0.25	18	14
0.50	18	9
0.75	9	7
1.00	6	5
2.00	4	3
3.00	4	1
4.00	3	0

1. Plot the number of species of Darwin's finches and the number of islands in the Galápagos archipelago (dependent variables) against time (independent variable).
2. Are the data consistent with the hypothesis that isolation of populations on newly formed islands is related to speciation in this group of birds? Why or why not?
3. If no more islands form in the Galápagos archipelago, do you think that speciation by geographic isolation will continue to occur among Darwin's finches? Why or why not? What additional data could you collect to test your hypothesis (without waiting to see if speciation occurs)?

Sympatric speciation occurs without physical barriers

Geographic isolation is usually required for speciation, but under some circumstances speciation can occur in the absence of a physical barrier. Speciation without physical isolation is called **sympatric speciation** (Greek *sym*, "together with"). But how can such speciation happen? Given that speciation is usually a gradual process, how can reproductive isolation develop when individuals have frequent opportunities to mate with one another?

Sympatric speciation may occur with some forms of disruptive selection (see Concept 15.4) in which individuals with certain genotypes have a preference for distinct microhabitats where mating takes place. The opening story of African cichlids was an example. Another example appears to be taking place in the apple maggot fly (*Rhagoletis pomonella*) of eastern North America. Until the mid-1800s, *Rhagoletis* flies courted, mated, and deposited their eggs only on hawthorn fruits. About 150 years ago, some flies began to lay their eggs on apples, which European immigrants had introduced into eastern North America. Apple trees are closely related to hawthorns, but the smell of the fruits differs, and the apple fruits appear earlier than those of hawthorns. Some early-emerging female *Rhagoletis* laid their eggs on apples, and over time, a genetic preference for the smell of apples evolved among early-emerging flies. When the offspring of these flies sought out apple trees for mating and egg deposition, they mated with other flies reared on apples, which shared the same preferences.

Today the two groups of *Rhagoletis pomonella* in the eastern U.S. appear to be on the way to becoming distinct species. One group mates and lays eggs primarily on hawthorn fruits, the other on apples. The incipient species are partially reproductively isolated because they mate primarily with individuals raised on the same fruit and because they emerge from their pupae at different times of the year. In addition, the apple-feeding flies now grow more rapidly on apples than they originally did. Sympatric speciation that arises from such host-plant specificity may be widespread among insects, many of which feed only on a single plant species.

The most common means of sympatric speciation, however, is **polyploidy**—the duplication of sets of chromosomes within individuals. Polyploidy can arise either from chromosome duplication in a single species (**autopolyploidy**) or from the combining of the chromosomes of two different species (**allopolyploidy**).

An autopolyploid individual originates when, for example, two accidentally unreduced diploid gametes (with two sets of chromosomes) combine to form a tetraploid individual (with four sets of chromosomes). Tetraploid and diploid individuals of the same species are reproductively isolated because their hybrid offspring are triploid. Even if these offspring survive, they are usually sterile; they cannot produce normal gametes because their chromosomes do not segregate evenly during meiosis. So a tetraploid individual cannot produce fertile offspring by mating with a diploid individual—but it *can* do so if it self-fertilizes or mates with another tetraploid. Thus polyploidy can result in complete reproductive isolation in two generations—an important exception to the general rule that speciation is a gradual process.

Allopolyploids may be produced when individuals of two different (but closely related) species interbreed. Such hybridization often disrupts normal meiosis, which can result in chromosomal doubling. Allopolyploids are often fertile because each of the chromosomes has a nearly identical partner with which to pair during meiosis.

Speciation by polyploidy has been particularly important in the evolution of plants, although it has contributed to speciation in animals as well (such as the example in Figure 17.2). Botanists estimate that about 70 percent of flowering plant species

and 95 percent of fern species are the result of recent polyploidization. Some of these species arose from hybridization between two species followed by chromosomal duplication and self-fertilization. Other species diverged from polyploid ancestors, so that the new species shared their ancestors' duplicated sets of chromosomes. New species arise by polyploidy more easily among plants than among animals because plants of many species can reproduce by self-fertilization. In addition, if polyploidy arises in several offspring of a single parent, the siblings can fertilize one another.

Go to ANIMATED TUTORIAL 17.3
Speciation Mechanisms
PoL2e.com/at17.3

CHECKpoint CONCEPT 17.3

- ✓ Explain how speciation via polyploidy can happen in only two generations.
- ✓ What are some obstacles to sympatric speciation?
- ✓ If allopatric speciation is the most prevalent mode of speciation, what do you predict about the geographic distributions of many closely related species? Does your answer differ for species that are sedentary versus species that are highly mobile?

Most populations separated by a physical barrier become reproductively isolated only slowly and gradually. If two incipient species once again come into contact with each other, what keeps them from merging back into a single species?

CONCEPT 17.4 Reproductive Isolation Is Reinforced When Diverging Species Come into Contact

As discussed in Concept 17.2, once a barrier to gene flow is established, reproductive isolation will begin to develop through genetic divergence. Over many generations, differences accumulate in the isolated lineages, reducing the probability that individuals from each lineage will mate successfully with one another when they come back into contact. In this way, reproductive isolation can evolve as a by-product of the genetic changes in the two diverging lineages.

Reproductive isolation may be incomplete when the incipient species come back into contact, however, in which case some hybridization will occur. If hybrid individuals are less fit than non-hybrids, selection favors parents that do not produce hybrid offspring. Under these conditions, selection results in strengthening, or **reinforcement**, of isolating mechanisms that prevent hybridization.

Mechanisms that prevent hybridization from occurring are called **prezygotic isolating mechanisms**. Mechanisms that reduce the fitness of hybrid offspring are called **postzygotic isolating mechanisms**. Postzygotic mechanisms result in selection against hybridization, which leads to the reinforcement of the prezygotic mechanisms.

Prezygotic isolating mechanisms prevent hybridization between species

Prezygotic isolating mechanisms, which come into play before fertilization, can prevent hybridization in several ways.

MECHANICAL ISOLATION Differences in the sizes and shapes of reproductive organs may prevent the union of gametes from different species. With animals, there may be a match between the shapes of the reproductive organs of males and females of the same species, so that reproduction between species with mismatched structures is not physically possible. In plants, mechanical isolation may involve a pollinator. For example, orchids of the genus *Ophrys* produce flowers that look and smell like the females of particular species of wasps (**FIGURE 17.8**). When a male wasp visits and attempts to mate with the flower (mistaking it for a female wasp of his species), his mating behavior results in the transfer of pollen to and from his body by appropriately configured anthers and stigmas on the flower. Insects that visit the flower but do not attempt to mate with it do not trigger the transfer of pollen between the insect and the flower.

TEMPORAL ISOLATION Many organisms have distinct mating seasons. If two closely related species breed at different times of the year (or different times of day), they may never have an opportunity to hybridize. For example, in sympatric populations of three closely related leopard frog species, each species breeds at a different time of year (**FIGURE 17.9**). Although there is some overlap in the breeding seasons, the opportunities for hybridization are minimal.

BEHAVIORAL ISOLATION Individuals may reject, or fail to recognize, individuals of other species as potential mating partners. For example, the mating calls of male frogs of related

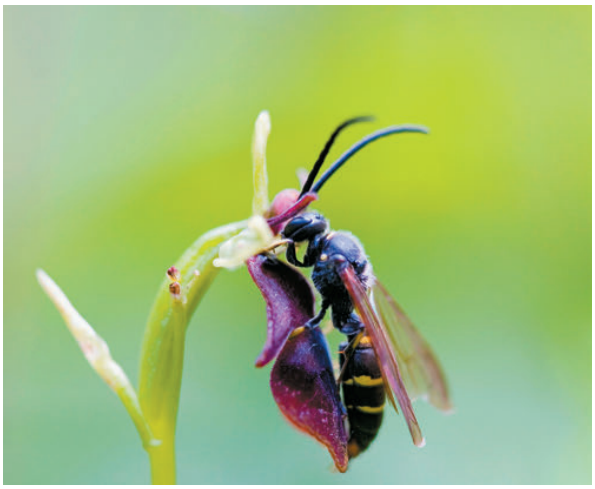


FIGURE 17.8 Mechanical Isolation through Mimicry Many orchid species maintain reproductive isolation by means of flowers that look and smell like females of a specific bee or wasp species. Shown here are a fly orchid (*Ophrys insectifera*) and its pollinator, a male wasp of the genus *Argogorytes*.

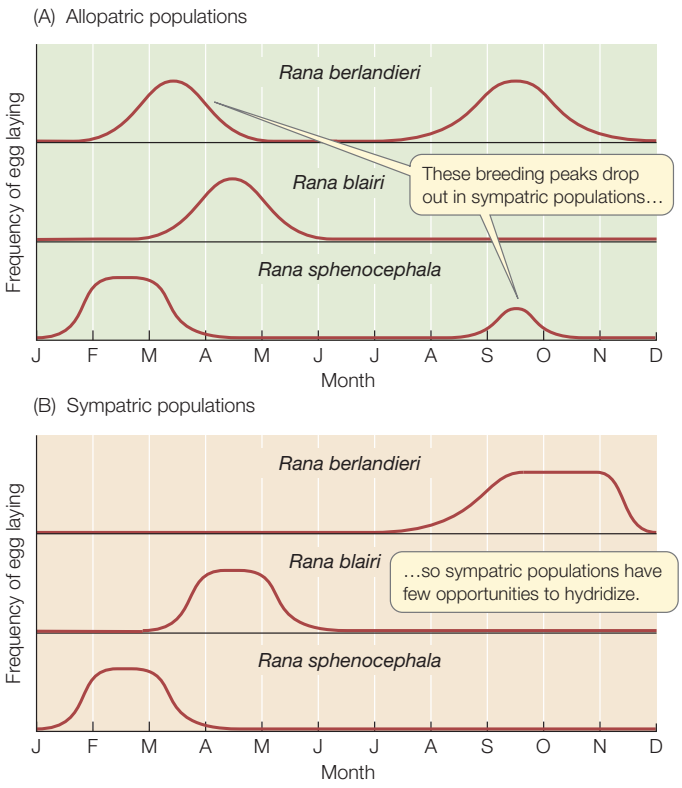


FIGURE 17.9 Temporal Isolation of Breeding Seasons (A) The peak breeding seasons of three species of *Rana* overlap when the species are physically separated (allopatry). (B) Where two or more species of *Rana* live together (sympatry), overlap between their peak breeding seasons is greatly reduced or eliminated. Selection against hybridization in areas of sympatry helps reinforce this prezygotic isolating mechanism.

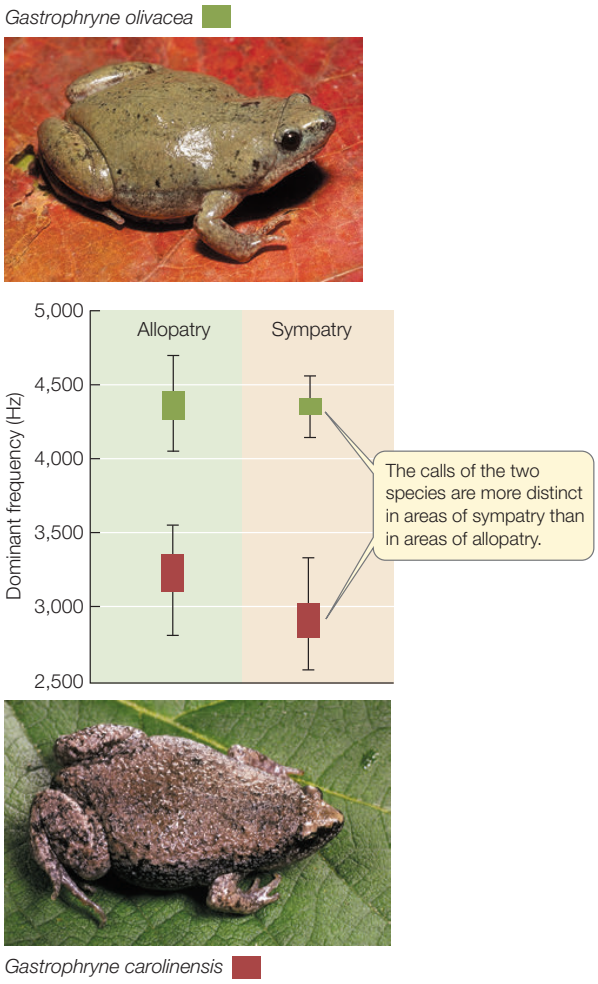


FIGURE 17.10 Behavioral Isolation in Mating Calls The males of most frog species produce species-specific calls. The calls of the two closely related frog species shown here differ in their dominant frequency (a high-frequency sound wave results in a high-pitched sound; a low frequency results in a low-pitched sound). Female frogs are attracted to the calls of males of their own species. The black lines indicate the range of individual calls; the upper and lower quartiles of calls are shown outside the colored boxes.

Go to MEDIA CLIP 17.1
Narrowmouth Toads Calling for Mates
PoL2e.com/mc17.1

species diverge quickly (**FIGURE 17.10**). Female frogs respond to mating calls from males of their own species but ignore the calls of other species, even closely related ones. The evolution of female preferences for certain male coloration patterns among the cichlids of Lake Malawi, described at the opening of this chapter, is another example of behavioral isolation.

Sometimes the mate choice of one species is mediated by the behavior of individuals of other species. For example, whether or not two plant species hybridize may depend on the food preferences of their pollinators. The floral traits of plants, including their color and shape, can enhance reproductive isolation either by influencing which pollinators are attracted to the flowers or by altering where pollen is deposited on the bodies of pollinators. A plant whose flowers are pendant (hanging downward; **FIGURE 17.11A**) will be pollinated by an animal with different physical characteristics than will a plant in which the flowers grow upright (**FIGURE 17.11B**). Because each pollinator prefers (and is adapted to) a different type of flower, the pollinators rarely transfer pollen from one plant species to the other.

Such isolation by pollinator behavior is seen in the mountains of California in two sympatric species of columbines (*Aquilegia*) that have diverged in flower color, structure, and orientation. *Aquilegia formosa* (**FIGURE 17.11C**) has pendant flowers with short spurs (spikelike, nectar-containing structures) and is pollinated by hummingbirds. *A. pubescens* (**FIGURE 17.11D**) has upright, lighter-colored flowers with long spurs and is pollinated by hawkmoths. The difference in pollinators means that these two species are effectively reproductively isolated even though they populate the same geographic range.

HABITAT ISOLATION When two closely related species evolve preferences for living or mating in different habitats, they may never come into contact during their respective mating periods. The *Rhagoletis* flies discussed in Concept 17.3 experienced such habitat isolation, as did the cichlid fishes that first adapted to rocky and sandy habitats upon entering Lake Malawi, as described at the opening of this chapter.

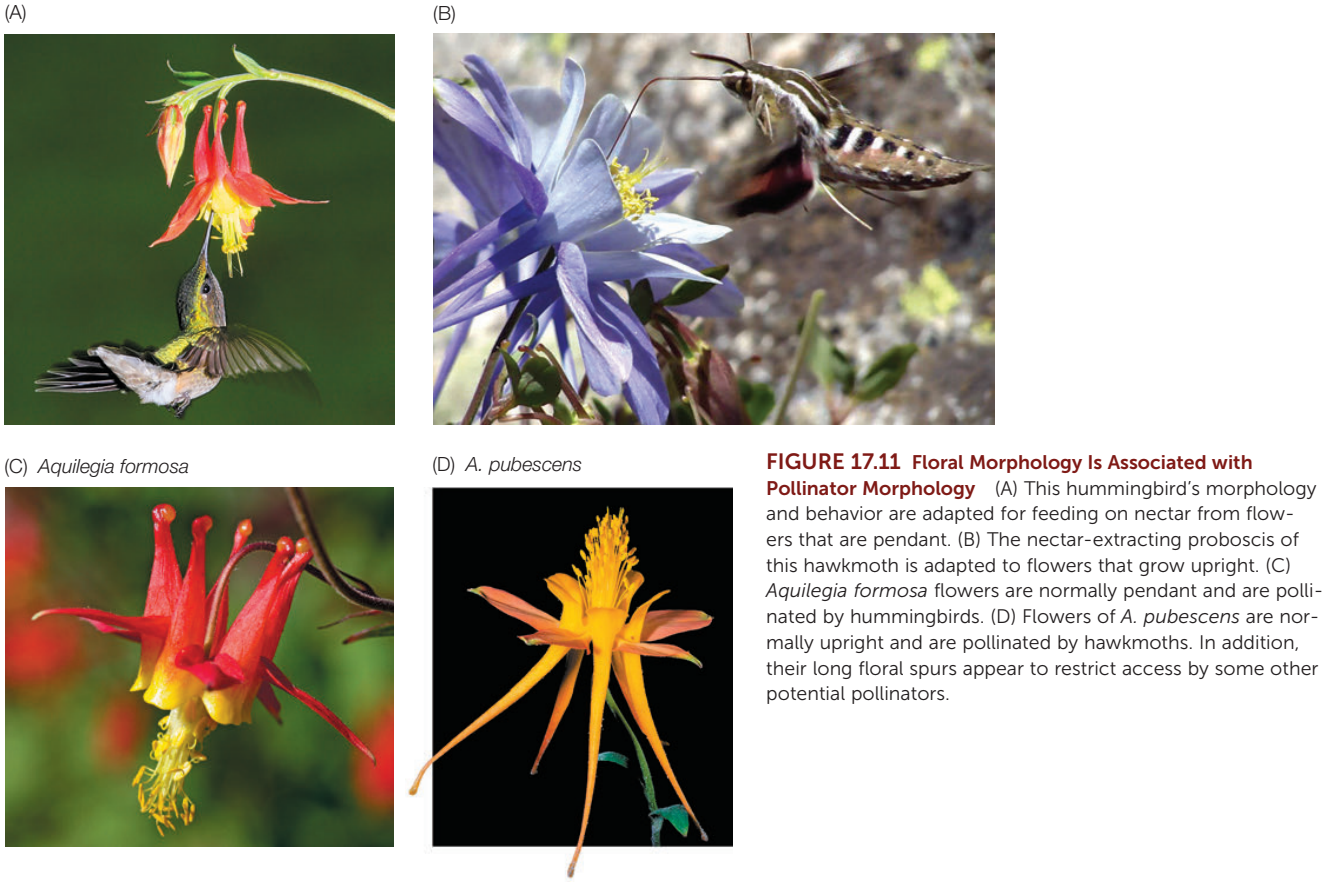


FIGURE 17.11 Floral Morphology Is Associated with Pollinator Morphology (A) This hummingbird's morphology and behavior are adapted for feeding on nectar from flowers that are pendant. (B) The nectar-extracting proboscis of this hawkmoth is adapted to flowers that grow upright. (C) *Aquilegia formosa* flowers are normally pendant and are pollinated by hummingbirds. (D) Flowers of *A. pubescens* are normally upright and are pollinated by hawkmoths. In addition, their long floral spurs appear to restrict access by some other potential pollinators.

LINK

Some plants and their pollinators become so tightly adapted to each other that they develop mutually dependent relationships, as described in **Concept 21.5**

GAMETIC ISOLATION The sperm of one species may not attach to the eggs of another species because the eggs do not release the appropriate attractive chemicals. Furthermore, the sperm may be unable to penetrate the egg because the two gametes are chemically incompatible. Thus even though the gametes of two species may come into contact, the gametes never fuse into a zygote.

Gametic isolation is extremely important for many aquatic species that spawn (release their gametes directly into the environment). For example, gametic isolation has been extensively studied in spawning sea urchins. A protein known as bindin is found in sea urchin sperm and functions in attaching ("binding") the sperm to eggs. All sea urchin species studied produce this egg-recognition protein, but the bindin gene sequence diverges so rapidly that it becomes species-specific. Since sperm can only attach to eggs of the same species, no interspecific hybridization occurs.

Postzygotic isolating mechanisms result in selection against hybridization

Genetic differences that accumulate between two diverging lineages may reduce the survival and reproductive rates of hybrid offspring in any of several ways:

- *Low hybrid zygote viability*. Hybrid zygotes may fail to mature normally, either dying during development or developing phenotypic abnormalities that prevent them from becoming reproductively capable adults.
- *Low hybrid adult viability*. Hybrid offspring may have lower survivorship than non-hybrid offspring.
- *Hybrid infertility*. Hybrids may mature into infertile adults. For example, the offspring of matings between horses and donkeys—mules—are sterile. Although otherwise healthy, mules usually produce no descendants.

Natural selection does not directly favor the evolution of postzygotic isolating mechanisms. But if hybrids are less fit, individuals that breed only within their own species will leave more surviving offspring than will individuals that interbreed with another species. Therefore individuals that can avoid interbreeding with members of other species will have a selective advantage, and any trait that contributes to such avoidance will be favored.

Donald Levin of the University of Texas has studied reinforcement of prezygotic isolating mechanisms in flowers of the genus *Phlox*. Levin noticed that individuals of *Phlox drummondii* in most of the range of the species in Texas usually have pink flowers. However, where *P. drummondii* is sympatric with its close relative the pink-flowered *P. cuspidata*, *P. drummondii* usually has red flowers. No other *Phlox* species has red flowers. Levin performed an experiment whose results showed that

reinforcement may explain why red flowers are favored where the two species are sympatric (FIGURE 17.12).

Likely cases of reinforcement are often detected by comparing sympatric and allopatric populations of potentially hybridizing species, as in the case of *Phlox*. If reinforcement is occurring, then sympatric populations of closely related species are expected to evolve more effective prezygotic reproductive barriers than do allopatric populations of the same species. As Figure 17.9 shows, the breeding seasons of sympatric populations of different leopard frog species overlap much less than do those of the corresponding allopatric populations. Similarly, the frequencies of the frog mating calls illustrated in Figure 17.10 are more divergent in sympatric populations than in allopatric populations. In both cases, there appears to have been natural selection against hybridization in areas of sympatry.

Hybrid zones may form if reproductive isolation is incomplete

Unless reproductive isolation is complete, closely related species may hybridize in areas where their ranges overlap, resulting in the formation of a **hybrid zone**. When a hybrid zone first forms, most hybrids are offspring of crosses between purebred individuals of the two species. However, subsequent generations will include a variety of individuals with varying proportions of their genes derived from the original two species. Thus hybrid zones often contain recombinant individuals resulting from many generations of hybridization.

Detailed genetic studies can tell us much about why narrow hybrid zones may persist for long periods between the ranges of two species. In Europe, the hybrid zone between two toad species of the genus *Bombina* has been studied intensively. The fire-bellied toad (*B. bombina*) lives in eastern Europe, whereas the closely related yellow-bellied toad (*B. variegata*) lives in western and southern Europe. The ranges of the two species overlap in a long but very narrow zone stretching 4,800 kilometers from eastern Germany to the Black Sea (FIGURE 17.13). Hybrids between the two species suffer from a range of defects, many of which are lethal. Those hybrids that survive often have skeletal abnormalities, such as misshapen mouths, ribs that

INVESTIGATION

FIGURE 17.12 Flower Color and Reproductive Isolation Most *Phlox drummondii* individuals have pink flowers, but in regions where the species is sympatric with *P. cuspidata*—which is always pink—most *P. drummondii* individuals have red flowers. Most pollinators preferentially visit flowers of one color or the other. In this experiment, Donald Levin explored whether flower color acts as a prezygotic isolating mechanism that lessens the chances of hybridization between the two species.[a]

HYPOTHESIS

Red-flowered *P. drummondii* are less likely to hybridize with *P. cuspidata* than are pink-flowered *P. drummondii*.

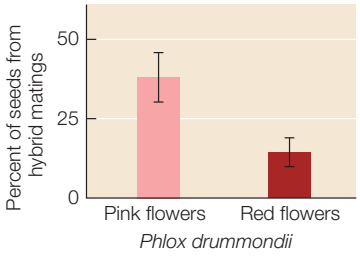
METHOD

1. Introduce equal numbers of red-flowered and pink-flowered *P. drummondii* individuals into an area where many pink-flowered *P. cuspidata* are growing.

2. After the flowering season ends, measure hybridization by assessing the genetic composition of the seeds produced by *P. drummondii* plants of both colors.

RESULTS

Of the seeds produced by pink-flowered *P. drummondii*, 38% were hybrids with *P. cuspidata*. Only 13% of the seeds produced by red-flowered *P. drummondii* were hybrids with *P. cuspidata*. Error bars show 95% confidence intervals (see Appendix B).



CONCLUSION

P. drummondii and *P. cuspidata* are less likely to hybridize if the flowers of the two species differ in color.

ANALYZE THE DATA

Data from Levin's experiment show that the frequency of hybridization between *Phlox drummondii* and *P. cuspidata* was strongly dependent on the former's flower color.

P. drummondii flower color	Number of progeny (seeds)		
	P. drummondii	Hybrid	Total
Red	181 (87%)	27 (13%)	208
Pink	86 (62%)	53 (38%)	139

A. The data reveal that red-flowered *P. drummondii* produced more seeds (208) than pink-flowered plants did (139), even though equal numbers of individuals of the two flower types were used. Does this difference influence your interpretation of Levin's results? Why or why not?

B. How would you improve the experimental design of this study? Should replicate or control sites be added? What kinds of additional test sites and conditions would you add?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]D. A. Levin. 1985. *Evolution* 39: 1275–1281.

APPLY THE CONCEPT

Reproductive isolation is reinforced when diverging species come into contact

As shown in Figure 17.9, the leopard frogs *Rana berlandieri* and *R. sphenocephala* usually have nonoverlapping breeding seasons in areas of sympatry, but where the species are allopatric, both species breed in both spring and fall. But when new ponds are created where the ranges of the two species come close together, frogs from previously allopatric populations may colonize the new ponds and hybridize during their overlapping breeding seasons.

LIFE STAGE	*R. BERLANDIERI*	*R. SPHENOCEPHALA*	F_1 HYBRIDS
Recently hatched tadpoles (spring, year 1)	155	125	238
Late-stage tadpoles (summer, year 1)	45	55	64
Newly metamorphosed froglets (fall, year 1)	32	42	15
Adult frogs (year 2)	10	15	1

Imagine you have collected and tabulated data on hybridization between these two frog species. You have sampled various life stages of frogs and their tadpoles for 2 years after an initial spring breeding season at a newly established pond. Use the data (above right) to answer the following questions.

1. Create four pie charts (one for each life stage) showing the percentage of each species and the percentage of hybrids at each stage.
2. What are some possible reasons for the differences in the percentages of hybrids found at each life stage? Suggest some postzygotic isolating mechanisms that are consistent with your data.
3. Over time, what changes might you expect in the breeding seasons of the two species at this particular pond, and why? How would future pie charts change if your predictions about breeding seasons are correct?

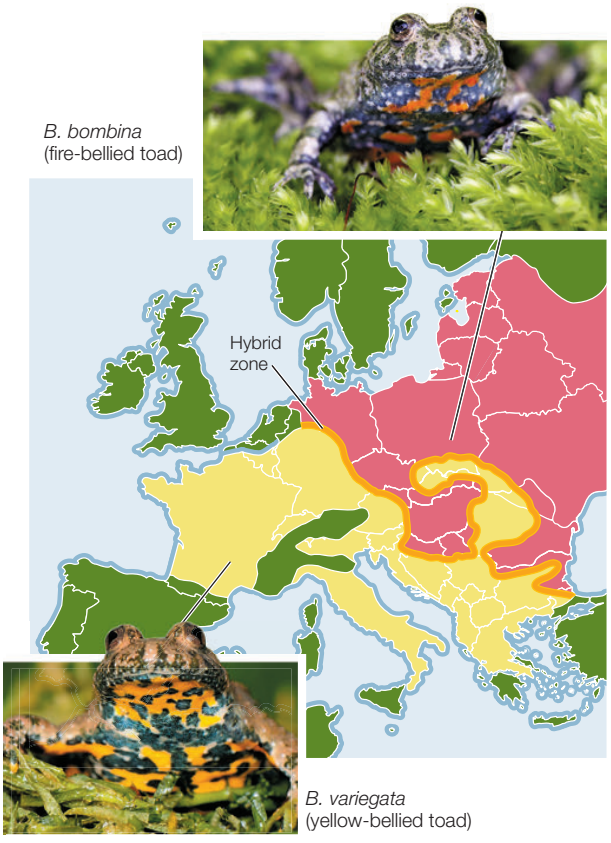


FIGURE 17.13 A Hybrid Zone The long, narrow zone where fire-bellied toads meet and hybridize with yellow-bellied toads has been stable for hundreds of years.

are fused to vertebrae, and a reduced number of vertebrae. By following the fates of thousands of toads from the hybrid zone, investigators found that a hybrid toad, on average, is only half as fit as a purebred individual of either species. The hybrid zone remains narrow because there is strong selection against hybrids and because adult toads do not move over long distances. The zone has persisted for hundreds of years, however, because individuals of both species continue to move short distances into it, continually replenishing the hybrid population.

CHECKpoint CONCEPT 17.4

- ✓ Why would reinforcement of prezygotic isolating mechanisms be expected even if postzygotic isolating mechanisms already exist?
- ✓ Why would you expect to find different types of prezygotic isolating mechanisms among different groups of organisms? Describe some specific examples.
- ✓ Why don't most narrow hybrid zones, such as the one between *Bombina bombina* and *B. variegata*, get wider over time?

The result of 4 billion of years of evolution has been many millions of species, each adapted to live in a particular environment and to use environmental resources in a particular way. In the next chapter we will describe how geological and biological processes interacted over the course of Earth's history to produce this vast biodiversity.

Q Can biologists study the process of speciation in the laboratory?

ANSWER Although speciation usually takes thousands or millions of years, and although it is typically studied in natural settings such as Lake Malawi, some aspects of speciation can be studied and observed in controlled laboratory experiments. Most such experiments use organisms with short generation times, in which evolution is expected to be relatively rapid.

William Rice and George Salt conducted an experiment in which fruit flies were allowed to choose food sources in different habitats, where mating also took place. The habitats were vials in different parts of an experimental cage (**FIGURE 17.14**). The vials differed in three environmental factors: (1) light; (2) the direction (up or down) in which the fruit flies had to move to reach food; and (3) the concentrations of two aromatic chemicals, ethanol and acetaldehyde. In just 35 generations, the two groups of flies that chose the most divergent habitats had become reproductively isolated from each other, having evolved distinct preferences for the different habitats.

The experiment by Rice and Salt demonstrated an example of habitat isolation evolving to function as a prezygotic isolating mechanism. Even though the different habitats were in the same cage, and individual fruit flies were capable of flying from one habitat to the other, habitat preferences were inherited by offspring from their parents, and populations from the two divergent habitats did not interbreed. Similar habitat selection is thought to have resulted in the early split in cichlids that preferred the rocky versus the sandy shores of Lake Malawi.

In controlled experiments like this one, biologists can observe many aspects of the process of speciation directly.

FIGURE 17.14 Evolution in the Laboratory For their experiments on the evolution of prezygotic isolating mechanisms in *Drosophila melanogaster*, Rice and Salt built an elaborate system of varying habitats contained within vials inside a large fly enclosure. Some groups of flies developed preferences for widely divergent habitats and became reproductively isolated within 35 generations.

SUMMARY

CONCEPT 17.1 Species Are Reproductively Isolated Lineages on the Tree of Life

- **Speciation** is the process by which one species splits into two or more daughter species, which thereafter evolve as distinct lineages.
- The **morphological species concept** distinguishes species on the basis of physical similarities; it often underestimates or overestimates the actual number of reproductively isolated species.
- The **biological species concept** distinguishes species on the basis of **reproductive isolation**.
- **Lineage species concepts**, which recognize independent evolutionary lineages as species, allow biologists to consider species over evolutionary time.

CONCEPT 17.2 Speciation Is a Natural Consequence of Population Subdivision

- Genetic divergence results from the interruption of gene flow within a population.
- The Dobzhansky–Muller model describes how reproductive isolation between two descendant lineages can develop through the accumulation of incompatible genes or chromosomal arrangements. **Review Figures 17.3 and 17.4 and ANIMATED TUTORIAL 17.1**
- Reproductive isolation increases with increasing genetic divergence between populations. **Review Figure 17.5**

CONCEPT 17.3 Speciation May Occur through Geographic Isolation or in Sympatry

- **Allopatric speciation**, which results when populations are separated by a physical barrier, is the dominant mode of speciation. This type of speciation may follow founder events, in which some members of a population cross a barrier and found a new, isolated population. **Review Figures 17.6 and 17.7 and ANIMATED TUTORIAL 17.2**
- **Sympatric speciation** results when two species diverge in the absence of geographic isolation. It can result from disruptive selection in two or more distinct microhabitats.
- Sympatric speciation can occur within two generations via **polyploidy**, an increase in the number of chromosomes sets. Polyploidy may arise from chromosome duplications within a species (**autopolyploidy**) or from hybridization that combines the chromosomes of two species (**allopolyploidy**). **Review ANIMATED TUTORIAL 17.3**

CONCEPT 17.4 Reproductive Isolation Is Reinforced When Diverging Species Come into Contact

- **Prezygotic isolating mechanisms** prevent hybridization; **postzygotic isolating mechanisms** reduce the fitness of hybrids.
- Postzygotic isolating mechanisms lead to **reinforcement** of prezygotic isolating mechanisms by natural selection. **Review Figures 17.9, 17.10, and 17.12**
- **Hybrid zones** may persist between species with incomplete reproductive isolation. **Review Figure 17.13**

See **ACTIVITY 17.1** for a concept review of this chapter.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is17

18 The History of Life on Earth

KEY CONCEPTS

Meganeuropsis permiana, shown in a reconstruction from fossils. Except for its size, this giant from the Permian period was similar to modern dragonflies (shown in the inset at the same scale).

Almost anyone who has spent time around freshwater ponds is familiar with dragonflies. Their bright colors and transparent wings stimulate our visual senses on bright summer afternoons as they fly about devouring mosquitoes, mating, and laying their eggs. The largest dragonflies alive today have wingspans that can be covered by a human hand. Three hundred million years ago, however, dragonflies such as *Meganeuropsis permiana* had wingspans of more than 70 centimeters—well over 2 feet, matching or exceeding the wingspans of many modern birds of prey. These dragonflies were the largest flying predators of their time.

No flying insects alive today are anywhere near this size. But during the Carboniferous and Permian geological periods, 350–250 million years ago, many groups of flying insects contained gigantic members. *Meganeuropsis* probably ate huge mayflies and other giant flying insects that shared their home in the Permian swamps. These enormous insects were themselves eaten by giant amphibians.

None of these insects or amphibians would be able to survive on Earth today. The oxygen concentrations in Earth's atmosphere were about 50 percent higher at that time compared to the present, and those high oxygen concentrations are thought to have been necessary to support giant insects and their huge amphibian predators.

Paleontologists have uncovered fossils of *Meganeuropsis permiana* in the rocks of Kansas. How do we know the age of these fossils, and how can we know how much oxygen that long-vanished atmosphere contained? The layering of the rocks allows us to tell their ages relative to one another, but it does not by itself indicate a given layer's absolute age.

One of the remarkable achievements of twentieth-century scientists was the development of sophisticated techniques that use the decay rates of various radioisotopes, the ratios of certain molecules in rocks and fossils, and changes in Earth's magnetic field to infer conditions and events in the remote past and to date them accurately. It is those methods that allow us to age the fossils of *Meganeuropsis* and to calculate the concentration of oxygen in Earth's atmosphere at the time.

Earth is about 4.5 billion years old, and life has existed on it for about 3.8 billion of those years. That means human civilizations have occupied Earth for less than 0.0003 percent of the history of life. Discovering what happened before humans were around is an ongoing and exciting area of science.

Q Can modern experiments test hypotheses about the evolutionary impact of ancient environmental changes?

You will find the answer to this question on page 374.

CONCEPT 18.1 Events in Earth's History Can Be Dated

Some evolutionary changes happen rapidly enough to be studied directly and manipulated experimentally. Plant and animal breeding by agriculturalists and evolution of resistance to pesticides are examples of rapid, short-term evolution that we saw in Chapter 15. Other evolutionary changes, such as the appearance of new species and evolutionary lineages, usually take place over much longer time scales.

To understand long-term patterns of evolutionary change, we must not only think in time scales spanning many millions of years, but also consider events and conditions very different from those we observe today. Earth of the distant past was so unlike our present Earth that it would seem like a foreign planet inhabited by strange organisms. The continents were not where they are now, and climates were sometimes dramatically different from those of today. We know this because much of Earth's history is recorded in its rocks.

We cannot tell the ages of rocks just by looking at them, but we can visually determine the ages of rocks *relative to one another*. The first person to formally recognize this fact was the seventeenth-century Danish physician Nicolaus Steno. Steno realized that in undisturbed **sedimentary rocks** (rocks formed by the accumulation of sediments), the oldest layers of rock, or **strata** (singular *stratum*), lie at the bottom, and successively higher strata are progressively younger.

Geologists subsequently combined Steno's insight with their observations of fossils contained in sedimentary rocks. They developed the following principles of **stratigraphy**:

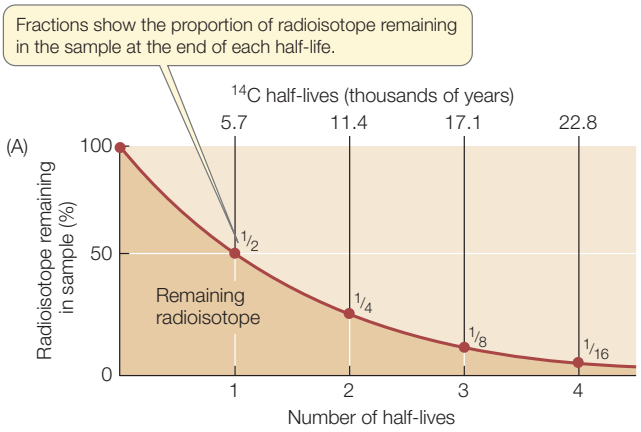
- Fossils of similar organisms are found in widely separated places on Earth.
- Certain fossils are always found in younger strata, and certain other fossils are always found in older strata.
- Organisms found in younger strata are more similar to modern organisms than are those found in older strata.

These patterns revealed much about the relative ages of sedimentary rocks and the fossils they contain, as well as patterns in the evolution of life. But the geologists still could not tell how old particular rocks were. A method of dating rocks did not become available until after radioactivity was discovered at the beginning of the twentieth century.

Radioisotopes provide a way to date rocks

Radioactive isotopes of atoms—radioisotopes—decay in a predictable pattern over long periods. Over a specific time interval, known as a **half-life**, half of the atoms in a radioisotope decay to become a different, stable (nonradioactive) isotope (**FIGURE 18.1A**). The use of this knowledge to date fossils and rocks is known as **radiometric dating**.

To use a radioisotope to date a past event, we must know or estimate the concentration of that isotope at the time of that event, and we must know the radioisotope's half-life. In the case of carbon-14, a radioisotope of carbon, the production of new carbon-14 (^{14}C) in the upper atmosphere—by the reaction of neutrons with nitrogen-14 (^{14}N, a stable isotope of nitrogen)—just balances the natural radioactive decay of ^{14}C into ^{14}N. Therefore the ratio of ^{14}C to the more common stable isotope of carbon, carbon-12 (^{12}C), is relatively constant in living organisms and in their environment. As soon as an organism dies, however, it ceases to exchange carbon compounds with its environment. Its decaying ^{14}C is no longer replenished, and the ratio of ^{14}C to ^{12}C in its remains decreases over time. Paleontologists can use the ratio of ^{14}C to ^{12}C in fossil material to date fossils that are less than 60,000 years old (and thus the sedimentary rocks that contain those fossils). If fossils are older than that, so little ^{14}C remains that the limits of detection using this particular isotope are reached.



(B)

Radioisotope	Decay product	Half-life (years)	Useful dating range (years)
Carbon-14 (^{14}C)	Nitrogen-14 (^{14}N)	5,700	100 – 60,000
Uranium-234 (^{234}U)	Thorium-230 (^{230}Th)	80,000	10,000 – 500,000
Uranium-235 (^{235}U)	Lead-207 (^{207}Pb)	704 million	200,000 – 4.5 billion
Potassium-40 (^{40}K)	Argon-40 (^{40}Ar)	1.3 billion	10 million – 4.5 billion

FIGURE 18.1 Radioactive Isotopes Allow Us to Date Ancient Rocks
The decay of radioactive isotopes into stable isotopes happens at a steady rate. A half-life is the time it takes for half of the remaining atoms to decay in this way. (A) The graph demonstrates the principle of half-life using carbon-14 (^{14}C) as an example. The half-life of ^{14}C is 5,700 years. (B) Examples of some radioisotopes with different characteristic half-lives that allow us to estimate the ages of many rocks.

Radiometric dating methods have been expanded and refined

Sedimentary rocks are formed from materials that existed for varying lengths of time before being weathered, fragmented, and transported, sometimes over long distances, to the site of their deposition. Therefore the radioisotopes in sedimentary rock do not contain reliable information about the date of its formation. Radiometric dating of rocks older than 60,000 years requires estimating radioisotope concentrations

APPLY THE CONCEPT

Events in Earth's history can be dated

Imagine you have been assigned the job of producing a geological map of volcanic rocks that were formed between 400 and 600 million years ago. You collect samples from ten sites (1–10 on the map at lower right). Then you can determine the ratio of ^{206}Pb to ^{238}U for each sample and use these ratios to estimate the ages of the rock samples, as given in the table below. Use the data to answer the following questions.

SITE	$^{206}PB/^{238}U$ RATIO	ESTIMATED AGE (mya)
1	0.076	474
2	0.077	479
3	0.069	431
4	0.081	505
5	0.076	474
6	0.070	435
7	0.089	550
8	0.080	500
9	0.079	495
10	0.077	479

1. Use the table shown here and Table 18.1 to assign each sample to a geological period.
2. Use these estimated ages and geological periods of the samples to mark rough boundaries between the geological periods among the sample locations on the map below.
3. If you wanted to refine the boundary between the Ordovician and Silurian on your map, which of three new sampling sites—*x*, *y*, or *z*—would you add to your analysis next?

in **igneous** rocks, which are formed when molten material cools. To date sedimentary strata, geologists search for places where volcanic ash or lava flows have intruded into the sedimentary rock.

A preliminary estimate of the age of an igneous rock determines which radioisotopes can be used to date it (**FIGURE 18.1B**). The decay of potassium-40 (which has a half-life of 1.3 billion years) to argon-40, for example, has been used to date many of the ancient events in the evolution of life. Fossils in the adjacent sedimentary rock that are similar to those in other rocks of known ages provide additional clues to the rock's age.

Scientists have used several methods to construct a geological time scale

Radiometric dating of rocks, combined with fossil analysis, is the most powerful method of determining geological age. But in places where sedimentary rocks do not contain suitable igneous intrusions and few fossils are present, paleontologists turn to other dating methods.

One method, known as **paleomagnetic dating**, relates the ages of rocks to patterns in Earth's magnetism, which change over time. Earth's magnetic poles move and occasionally reverse themselves. Both sedimentary and igneous rocks preserve a record of Earth's magnetic field at the time they were formed, and that record can be used to determine the ages of those rocks. Other dating methods use information about continental drift, information about sea level changes, and molecular clocks.

LINK

Molecular clocks are described in **Concept 16.3**

Using all of these methods, geologists developed a **geological time scale** (**TABLE 18.1**). They divided the broad history of life into four **eons**. The Hadean eon refers to the time on Earth before life evolved. The early history of life occurred in the Archean eon, which ended about the time that photosynthetic organisms first appeared on Earth. Prokaryotic life diversified rapidly in the Proterozoic eon, and the first eukaryotes in the fossil record date from this time. These three eons are sometimes referred to collectively as Precambrian time, or simply the **Precambrian**. The Precambrian lasted for approximately 3.8 billion years and thus accounts for the vast majority of geological time. It was in the Phanerozoic eon, however—a mere 542-million-year time span—that multicellular eukaryotes rapidly diversified. To emphasize the events of the Phanerozoic, Table 18.1 shows the subdivision of this eon into eras and periods. The boundaries between these divisions of time are based

TABLE 18.1 Earth's Geological History

Eon	Era	Period	Onset	Major physical changes on Earth
Phanerozoic (~0.5 billion years long)	Cenozoic	Quaternary (Q)	2.6 mya	Cold/dry climate; repeated glaciations
		Tertiary (T)	65.5 mya	Continents near current positions; climate cools
	Mesozoic	Cretaceous (K)	145.5 mya	Laurasian continents attached to one another; Gondwana begins to drift apart; meteorite strikes near current Yucatán Peninsula at end of period
		Jurassic (J)	201.6 mya	Two large continents form: Laurasia (north) and Gondwana (south); climate warm
		Triassic (Tr)	251.0 mya	Pangaea begins to drift apart; hot/humid climate
	Paleozoic	Permian (P)	299 mya	Extensive lowland swamps; O_2 levels 50% higher than present; by end of period continents aggregate to form Pangaea, and O_2 levels drop rapidly
		Carboniferous (C)	359 mya	Climate cools; marked latitudinal climate gradients
		Devonian (D)	416 mya	Continents collide at end of period; one or more giant meteorites probably strike Earth
		Silurian (S)	444 mya	Sea levels rise; two large land masses emerge; hot/humid climate
		Ordovician (O)	488 mya	Massive glaciation; sea level drops 50 meters
		Cambrian (Ꞓ)	542 mya	Atmospheric O_2 levels approach current levels
Proterozoic	Collectively called the Precambrian (~4 billion years long)		2.5 bya	Atmospheric O_2 levels increase from negligible to about 18%; "snowball Earth" from about 750 to 580 mya
Archean			3.8 bya	Earth accumulates more atmosphere (still almost no O_2); meteorite impacts greatly reduced
Hadean			4.5–4.6 bya	Formation of Earth; cooling of Earth's surface; atmosphere contains almost no free O_2; oceans form; Earth under almost continuous bombardment from meteorites

Note: mya, million years ago; bya, billion years ago.

largely on the striking differences geologists observe in the assemblages of fossil organisms contained in successive strata. This geological record of life reveals a remarkable story of a world in which the continents and biological communities are constantly changing.

CHECKpoint CONCEPT 18.1

- ✓ What observations about fossils suggested to geologists that they could be used to determine the relative ages of rocks?
- ✓ Why are radioisotopes not measured directly from sedimentary rocks to determine their ages?
- ✓ Given the problems of dating sedimentary rocks using radioisotopes, what other methods can geologists use to date sedimentary rocks?

As geologists began to develop accurate ways to age Earth, they began to understand that Earth is far older than anyone had previously understood. During its 4.5-billion-year history, Earth has undergone massive physical changes. These changes have influenced the evolution of life, and life, in its turn, has influenced Earth's physical environment.

CONCEPT 18.2 Changes in Earth's Physical Environment Have Affected the Evolution of Life

As we saw in the previous section, the Phanerozoic eon has been notable for the rapid diversification of multicellular eukaryotes. But the diversity of multicellular organisms has not simply increased steadily through time. New species have arisen, and species have gone extinct, throughout the history of life. Let's consider some of the physical changes on Earth that have resulted in such dramatic changes in life's diversity.

The continents have not always been where they are today

The globes and maps that adorn our walls, shelves, and books give an impression of a static Earth. It would be easy for us to assume that the continents have always been where they are. But we would be wrong. The idea that Earth's land masses

Major events in the history of life
Humans evolve; many large mammals become extinct
Diversification of birds, mammals, flowering plants, and insects
Dinosaurs continue to diversify; mass extinction at end of period (~76% of species lost)
Diverse dinosaurs; radiation of ray-finned fishes; first fossils of flowering plants
Early dinosaurs; first mammals; marine invertebrates diversify; mass extinction at end of period (~65% of species lost)
Reptiles diversify; giant amphibians and flying insects present; mass extinction at end of period (~96% of species lost)
Extensive fern/horsetail/giant club moss forests; first reptiles; insects diversify
Jawed fishes diversify; first insects and amphibians; mass extinction at end of period (~75% of marine species lost)
Jawless fishes diversify; first jawed fishes; plants and animals colonize land
Mass extinction at end of period (~75% of species lost)
Rapid diversification of multicellular animals; diverse photosynthetic protists
Origin of photosynthesis, multicellular organisms, and eukaryotes
Origin of life; prokaryotes flourish
Life not yet present

have changed their positions over the millennia, and that they continue to do so, was first put forth in 1912 by the German meteorologist and geophysicist Alfred Wegener. His idea, known as **continental drift**, was initially met with skepticism and resistance. By the 1960s, however, physical evidence and increased understanding of **plate tectonics**—the geophysics of the movement of major land masses—had convinced virtually all geologists of the reality of Wegener's vision. Plate tectonics provided the geological mechanism that explained Wegener's hypothesis of continental drift.

Earth's crust consists of several solid plates. Thick continental and thinner oceanic plates overlie a viscous, malleable layer of Earth's mantle. Heat produced by radioactive decay deep in Earth's core sets up large-scale convection currents in the mantle. New crust is formed as mantle material rises between diverging plates, pushing them apart.

Where oceanic plates and continental plates converge, the thinner oceanic plate is forced underneath the thicker continental plate, a process known as **subduction**. Subduction results in volcanism and mountain building on the continental boundary (FIGURE 18.2A). For example, in the Pacific Northwest of North America, a series of volcanoes formed the Cascade mountain range as the Juan de Fuca oceanic plate has been subducted beneath a portion of the continental North American Plate (FIGURE 18.2B). When two oceanic plates collide, one is also subducted below the other, producing a deep oceanic trench and associated volcanic activity.

When two thick continental plates collide, neither plate is subducted. Instead, the plates push up against one another, forming high mountain chains. The highest mountain chain in the world, the Himalayas, was formed this way when the Indian Plate collided with the Eurasian Plate. When continental plates diverge, new crust forms in the intervening spaces, resulting in deep clefts called rift valleys in which large freshwater lakes typically form. The Great Rift Valley lakes of eastern Africa, including Lake Malawi (discussed at the opening of Chapter 17), were formed in this way.

Many physical conditions on Earth have oscillated in response to plate tectonic processes. We now know that the movement of the plates has sometimes brought continents together and at other times has pushed them apart, as seen in the maps across the top of Figure 18.12. The positions and sizes of the continents influence oceanic circulation patterns, global climates, and sea levels. Sea levels are influenced directly by plate tectonic processes (which can influence the depth of ocean basins) and indirectly by oceanic circulation patterns, which affect patterns of glaciation. As climates cool, glaciers form and tie up water over land masses; as climates warm, glaciers melt and release water.

Some of these dramatic changes in Earth's physical parameters resulted in **mass extinctions**, during which a large proportion of the species living at the time disappeared. These mass extinctions are the cause of the striking differences in fossil assemblages that geologists used to divide the units of the geological time scale. After each mass extinction, the diversity of life rebounded, but recovery took millions of years.

Go to ANIMATED TUTORIAL 18.1
Evolution of the Continents
PoL2e.com/at18.1

Earth's climate has shifted between hot and cold conditions

Through much of its history, Earth's climate was considerably warmer than it is today, and temperatures decreased more gradually toward the poles. At other times, Earth was colder than it is today. Rapid drops in sea levels throughout the history of Earth have resulted mainly from increased global glaciation (FIGURE 18.3). Many of these drops in sea levels were accompanied by mass extinctions—particularly of marine organisms, which could not survive the disappearance of the shallow seas that covered vast areas of the continental shelves.

Earth's cold periods were separated by long periods of milder climates. Because we are living in one of the colder periods, it is difficult for us to imagine the mild climates that were found at high latitudes during much of the history of life. The Quaternary period has been marked by a series of glacial advances, interspersed with warmer interglacial intervals during which the glaciers retreated.

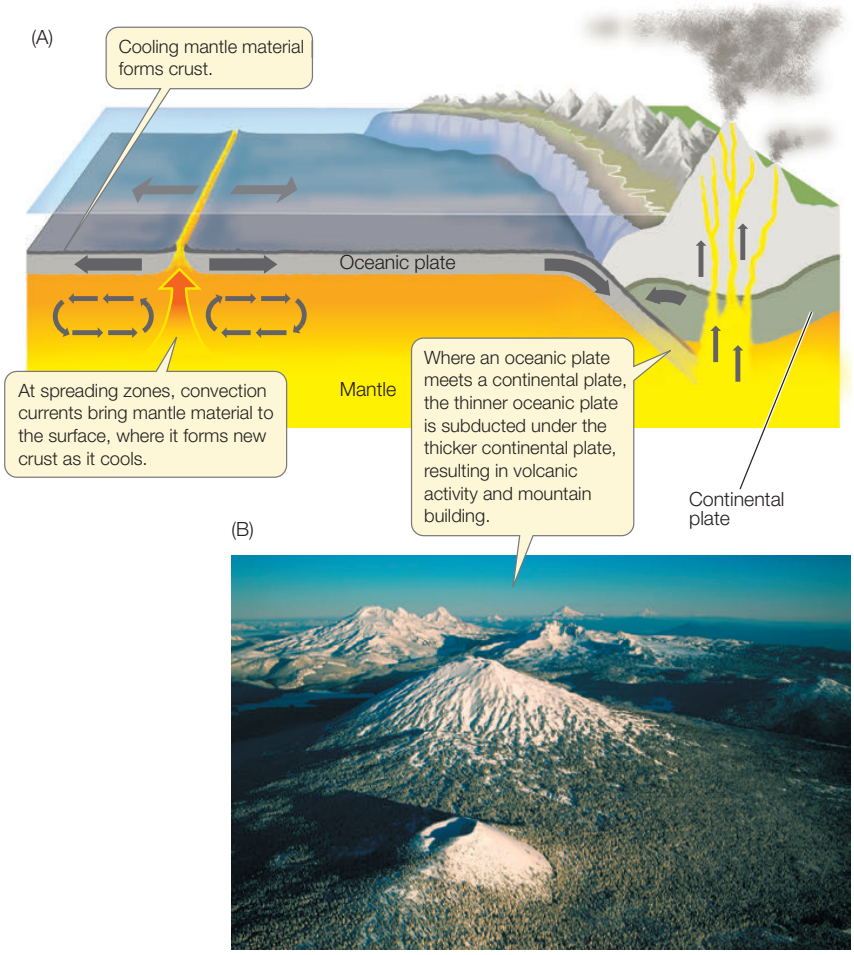


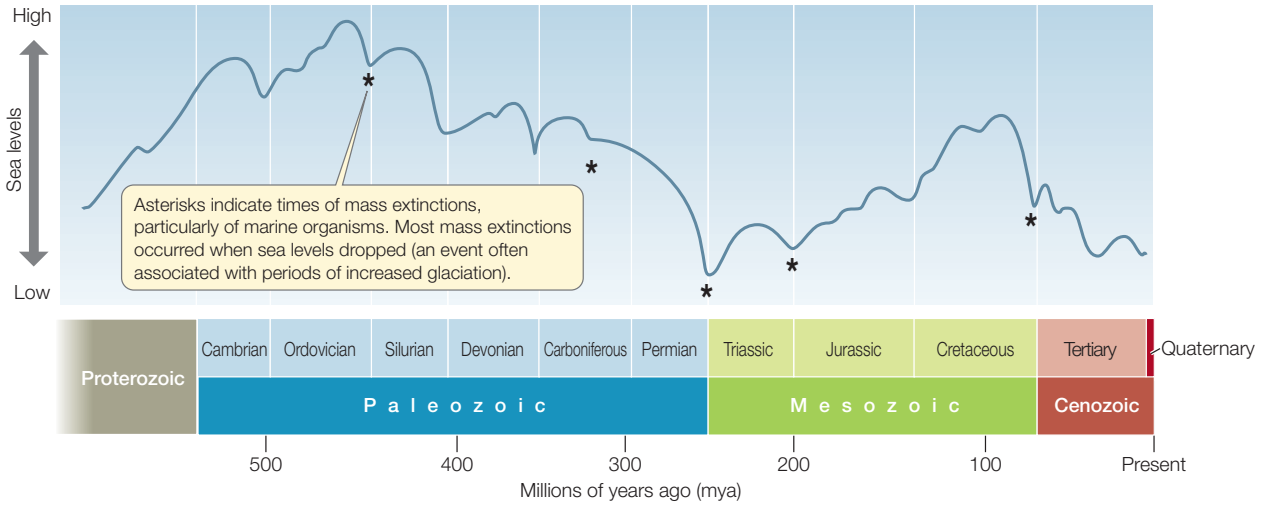
FIGURE 18.2 Plate Tectonics and Continental Drift (A) The heat of Earth's core generates convection currents in the viscous magma (the hot, molten rock of the mantle) underlying the oceanic and continental plates. Those currents push the continental plates, along with the land masses they carry, together or apart. Where plates collide, one may slide under the other, creating mountain ranges and often volcanoes. (B) The Cascade Range of the Pacific Northwest of North America is an example of a mountain range produced by subduction of an oceanic plate under a continental plate.

Go to MEDIA CLIP 18.1
Lava Flows and Magma Explosions
PoL2e.com/mc18.1

"Weather" refers to the daily events at a given location, such as individual storms and the high and low temperatures on a given day. "Climate" refers to long-term average expectations over the various seasons at a given location. Weather often changes rapidly, whereas climates typically change slowly. However, major climate shifts have taken place over periods as short as 5,000 to 10,000 years, primarily as a result of changes in Earth's orbit around the sun. A few climate shifts have been even more rapid. For example, during one Quaternary interglacial period, the ice-locked Antarctic Ocean became nearly ice-free in less than 100 years. Some climate changes have been so rapid that the extinctions caused by them appear to be nearly instantaneous in the fossil record. Such rapid changes are usually caused by sudden shifts in ocean currents.

We are currently living in a time of rapid climate change thought to be caused by a buildup of atmospheric CO_2, primarily from the burning of fossil fuels by human populations. We are reversing the energy transformations accrued in the burial and decomposition of organic material

FIGURE 18.3 Sea Levels Have Changed Repeatedly Rapid drops in sea levels are associated with periods of globally cooler temperatures and increased glaciation. Most mass extinctions of marine organisms have coincided with low sea levels.



that occurred (especially) in the Carboniferous, Permian, and Triassic, which gave rise to the fossil fuels we are using today. But we are burning these fuels over a few hundred years, rather than the many millions of years over which those deposits accumulated. The current rate of increase of atmospheric CO_2 is unprecedented in Earth's history. A doubling of the atmospheric CO_2 concentration—which may happen during the current century—is expected to increase the average temperature of Earth, change rainfall patterns, melt glaciers and ice caps, and raise sea levels.

LINK

The consequences of today's rapid climate changes are discussed in **Concept 45.5**

Volcanoes have occasionally changed the history of life

Most volcanic eruptions produce only local or short-lived effects, but a few large volcanic eruptions have had major consequences for life. When Krakatau (a volcanic island in the Sunda Strait off Indonesia) erupted in 1883, it ejected more than 25 cubic kilometers of ash and rock, as well as large quantities of sulfur dioxide gas (SO_2). The SO_2 was ejected into the stratosphere and carried by high-altitude winds around the planet. Its presence led to high concentrations of sulfurous acid (H_2SO_3) in high-altitude clouds, creating a "parasol effect" so that less sunlight reached Earth's surface. Global temperatures dropped by 1.2°C in the year following the eruption, and global weather patterns showed strong effects for another 5 years. More recently, the eruption of Mount Pinatubo in the Philippines in 1991 (**FIGURE 18.4**) temporarily reduced global temperatures by about 0.5°C.

Although these individual volcanoes had only relatively short-term effects on global temperatures, they suggest that the simultaneous eruption of many volcanoes could have a much stronger effect on Earth's climate. What would cause many volcanoes to erupt at the same time? The collision of continents during the Permian period, about 275 million years ago (mya), formed a single, gigantic land mass and caused massive volcanic eruptions as the continental plates overrode one another (see Figure 18.2). Emissions from these eruptions blocked considerable sunlight, contributing to the advance of glaciers and a consequent drop in sea levels (see Figure 18.3). Thus volcanoes were probably responsible, at least in part, for the greatest mass extinction in Earth's history.

Extraterrestrial events have triggered changes on Earth

At least 30 meteorites of sizes between tennis and soccer balls strike Earth each year. Collisions with larger meteorites or comets are rare, but such collisions have probably been responsible for several mass extinctions. Several types of evidence tell us about these collisions. Their craters, and the dramatically disfigured rocks that result from their impact, are found in many places. Geologists have discovered compounds in these rocks that contain helium and argon with isotope ratios characteristic

FIGURE 18.4 Volcanic Eruptions Can Cool Global Temperatures When Mount Pinatubo erupted in 1991, it increased the concentrations of sulfurous acid in high-altitude clouds, which temporarily lowered global temperatures by about 0.5°C.

of meteorites, which are very different from the ratios found elsewhere on Earth.

A meteorite caused or contributed to a mass extinction at the end of the Cretaceous period (about 65.5 mya). The first clue that a meteorite was responsible came from the abnormally high concentrations of the element iridium found in a thin layer separating rocks deposited during the Cretaceous from rocks deposited during the Tertiary (**FIGURE 18.5**). Iridium is abundant in some meteorites, but it is exceedingly rare on Earth's surface. When scientists then discovered a circular crater 180 km in diameter buried beneath the northern coast of the Yucatán Peninsula of Mexico, they constructed the following scenario. When it collided with Earth, the meteorite released energy equivalent to that of 100 million megatons of high explosives, creating great tsunamis. A massive plume of debris rose into the atmosphere, spread around Earth, and descended. The descending debris heated the atmosphere to several hundred degrees and ignited massive fires. It also blocked the sun, preventing plants from photosynthesizing. The settling debris formed the iridium-rich layer. About a billion tons of soot with a composition matching that of smoke from forest fires was also deposited. These events had devastating effects on biodiversity. Many fossil species (including non-avian dinosaurs) that are found in Cretaceous rocks are not found in the overlying Tertiary rocks.

Oxygen concentrations in Earth's atmosphere have changed over time

As the continents have moved over Earth's surface, the world has experienced other physical changes, including large

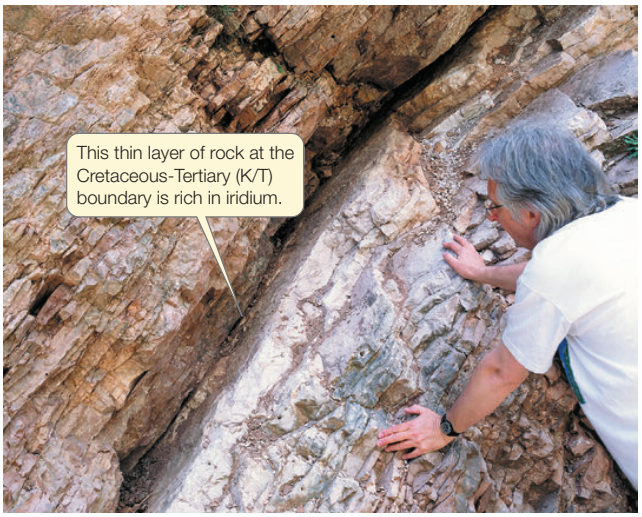


FIGURE 18.5 Evidence of a Meteorite Impact The white layers of rock are Cretaceous in age, whereas the layers at the upper left were deposited in the Tertiary. Between the two is a thin, dark layer of clay that contains large amounts of iridium, a metal common in some meteorites but rare on Earth. Its high concentration in this sediment layer, deposited about 65.5 million years ago, suggests the impact of a large meteorite at that time.

increases and decreases in atmospheric oxygen concentrations. The atmosphere of early Earth probably contained little or no free oxygen gas (O_2). The increase in atmospheric O_2 came in two big steps more than a billion years apart.

The first step occurred about 2.5 billion years ago (bya), when the ancestors of modern cyanobacteria evolved the ability to use water as the source of hydrogen ions for photosynthesis. By chemically splitting H_2O, these bacteria generated O_2 as a waste product. They also made electrons available for reducing CO_2 to form the carbohydrate end-products of photosynthesis (see Concepts 6.5 and 6.6). The O_2 they produced dissolved in water and reacted with dissolved iron. The reaction product then precipitated as iron oxide, which accumulated in alternating layers of red and dark rock known as banded iron formations (**FIGURE 18.6A**). These formations provide evidence for the earliest photosynthetic organisms. As photosynthetic organisms continued to release O_2, oxygen gas began to accumulate in the atmosphere.

The second step occurred about a billion years later, when a cyanobacteria-like ancestor became symbiotic within eukaryote cells, leading to the evolution of chloroplasts in photosynthetic plants and other eukaryotes. One group of O_2-generating cyanobacteria formed rocklike structures called stromatolites, which are abundantly preserved in the fossil record (**FIGURE 18.6B**).

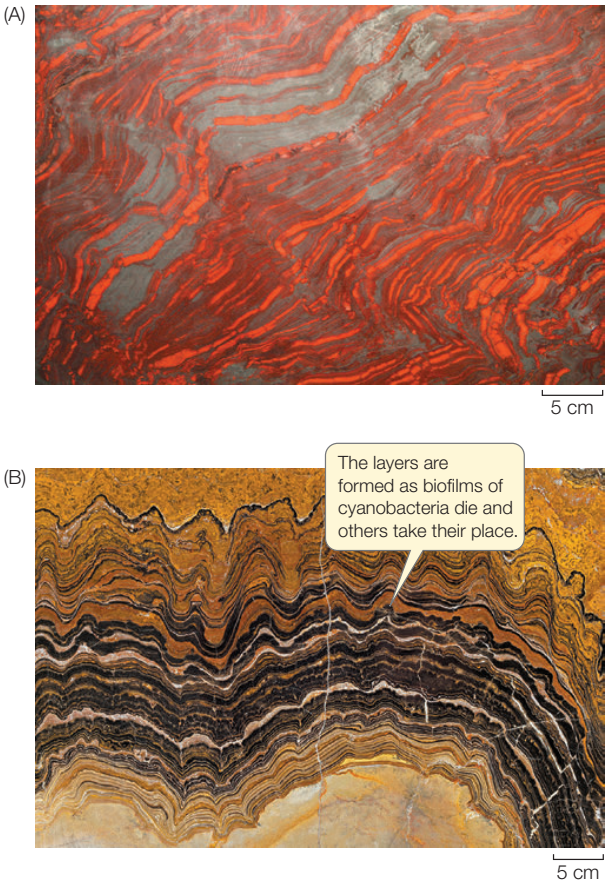
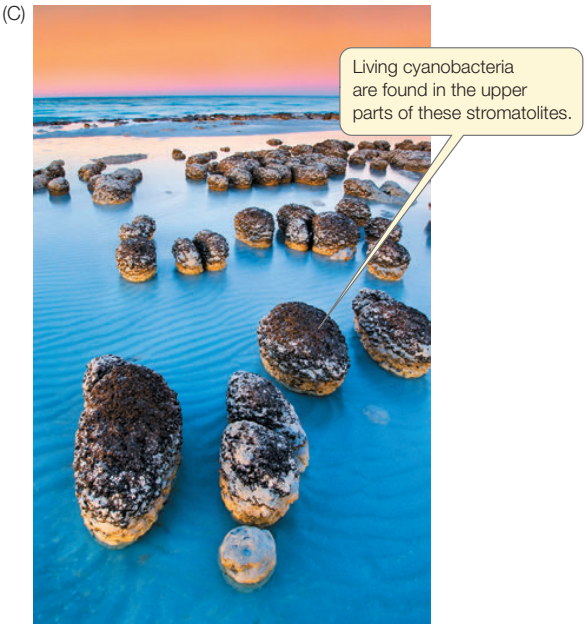


FIGURE 18.6 Evidence of Early Photosynthesis (A) The alternating red and dark layers in these "banded iron formations" resulted from a reaction between oxygen and dissolved iron, which produced iron oxide. The oxygen was produced by Earth's first photosynthetic organisms, beginning about 2.5 billion years ago. (B) A vertical section through a fossil stromatolite. Fossil stromatolites provide an early record of photosynthetic organisms. (C) These rocklike structures are living stromatolites that thrive in the very salty waters of Shark Bay in Western Australia.



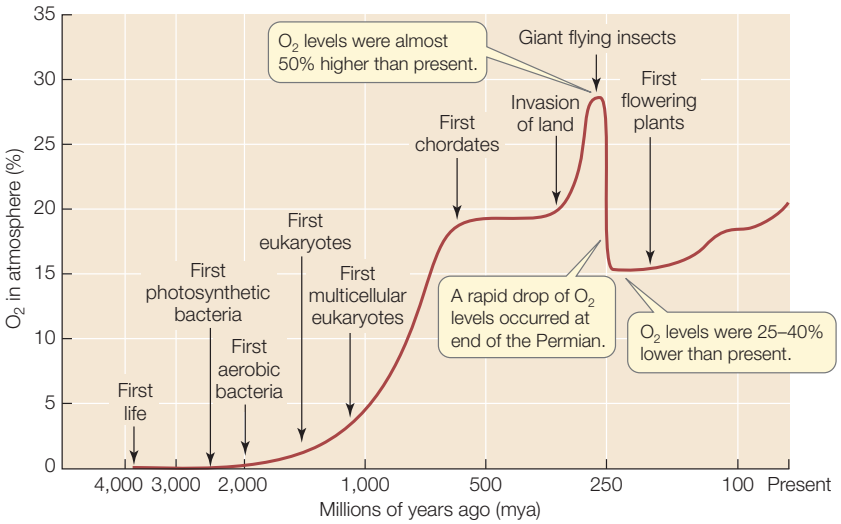


FIGURE 18.7 Atmospheric Oxygen Concentrations Have Changed Over Time Changes in atmospheric oxygen concentrations have strongly influenced, and been influenced by, the evolution of life. (Note that the horizontal axis of the graph is on a logarithmic scale.)

To this day, cyanobacteria still form stromatolites in a few very salty places (**FIGURE 18.6C**). Cyanobacteria liberated enough O_2 to open the way for the evolution of oxidation reactions as the energy source for the synthesis of ATP.

Thus the evolution of life irrevocably changed the physical nature of Earth. Those physical changes, in turn, influenced the evolution of life. When it first appeared in the atmosphere, O_2 was toxic to most of the anaerobic prokaryotes that inhabited Earth at the time. Over millennia, however, prokaryotes that evolved the ability to tolerate and use O_2 not only survived but gained the advantage. Aerobic metabolism proceeds more rapidly, and harvests energy more efficiently, than anaerobic metabolism. Organisms with aerobic metabolism replaced anaerobes in most of Earth's environments.

An atmosphere rich in O_2 also made possible larger and more complex organisms. Small single-celled aquatic organisms can obtain enough oxygen by simple diffusion even when dissolved oxygen concentrations in the water are very low. Larger single-celled organisms, however, have lower surface area-to-volume ratios. To obtain enough oxygen by simple diffusion, larger organisms must live in an environment with a relatively high oxygen concentration. Bacteria can thrive at 1 percent of the current oxygen concentration, but eukaryotic cells require levels that are at least 2–3 percent of the current concentration. For concentrations of dissolved oxygen in the oceans to have reached these levels, much higher atmospheric concentrations were needed.

Probably because it took many millions of years for Earth to develop an oxygenated atmosphere, only single-celled prokaryotes lived on Earth for more than 2 billion years. About 1.5 bya, atmospheric O_2 concentrations became high enough for larger eukaryotic cells to flourish (**FIGURE 18.7**). Further increases in atmospheric O_2 concentrations in the late Precambrian enabled several groups of multicellular organisms to evolve.

Oxygen concentrations increased again during the Carboniferous and Permian periods because of the evolution of large vascular plants. These plants lived in the expansive lowland swamps that existed at the time (see Table 18.1). Massive amounts of organic material were buried in these swamps as the plants died, leading to the formation of Earth's vast coal deposits. Because the buried organic material was not subject to oxidation as it decomposed, and because the living plants were producing large quantities of O_2, atmospheric O_2 increased to concentrations that have not been reached again in Earth's history (see Figure 18.7). As mentioned at the opening of this chapter, these high concentrations of atmospheric O_2 allowed the evolution of giant flying insects and amphibians that could not survive in today's atmosphere.

The drying of the lowland swamps at the end of the Permian reduced burial of organic matter as well as the production of O_2, so atmospheric O_2 concentrations dropped rapidly. Over the past 200 million years, with the diversification of flowering plants, O_2 concentrations have again increased, but not to the levels that characterized the Carboniferous and Permian periods.

Biologists have conducted experiments that demonstrate the changing selection pressures that can accompany changes in atmospheric O_2 concentrations. When fruit flies (*Drosophila*) are raised in hyperoxic conditions (i.e., with artificially increased atmospheric concentrations of O_2), they evolve larger body sizes in just a few generations (**FIGURE 18.8**). The present atmospheric O_2 concentrations appear to constrain body size in these flying insects, whereas increases in O_2 appear to relax those constraints. This experiment demonstrates that the stabilizing selection on body size at present O_2 concentrations can quickly switch to directional selection for a change in body size in response to a change in O_2 concentrations.

LINK

Stabilizing and directional selection are discussed and compared in **Concept 15.4**

INVESTIGATION

FIGURE 18.8 Atmospheric Oxygen Concentrations and Body Size in Insects C. Jaco Klok and his colleagues asked whether insects raised in hyperoxic conditions would evolve to be larger than their counterparts raised under today's atmospheric conditions. They raised strains of fruit flies (*Drosophila melanogaster*) under both conditions to test the effects of increased O_2 concentrations on the evolution of body size.[a]

HYPOTHESIS

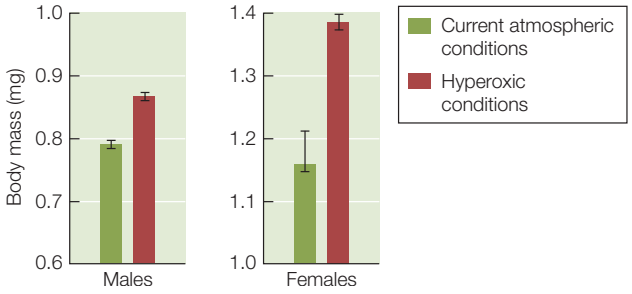
In hyperoxic conditions, increased partial pressure of oxygen results in evolution of increased body size in flying insects.

METHOD

1. Separate a population of fruit flies into multiple lines.
2. Raise half the lines in current atmospheric (control) conditions; raise the other lines in hyperoxic (experimental) conditions. Continue all lines for seven generations.
3. Raise the F_8 individuals of all lines under identical (current) atmospheric conditions.
4. Weigh 50 flies from each of the replicate lines and test for statistical differences in body weight.

RESULTS

The average body mass of F_8 individuals of both sexes raised under hyperoxic conditions was significantly ($P < 0.001$) greater than that of individuals in the control lines. Error bars show 95% confidence intervals for the mean (see Appendix B).



CONCLUSION

Increased O_2 concentrations led to evolution of larger body size in fruit flies, consistent with the trends seen among other flying insects in the fossil record.

ANALYZE THE DATA

The table shows the average body masses of the flies raised in hyperoxic conditions in the F_0 (i.e., before the first generation in hyperoxia), F_1, F_2, and F_8 generations.

	Average body mass (mg)	
Generation	Males	Females
F_0	0.732	1.179
F_1	0.847	1.189
F_2	0.848	1.254
F_8	0.878	1.392

A. Graph body mass versus generation for males and females.

B. Do the rates of evolution of larger body size appear to be constant throughout the experiment?

C. If you doubled the number of generations in the experiment, would you expect the increase in body mass seen under the hyperoxic conditions to double? Why or why not?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]C. J. Klok et al. 2009. *Journal of Evolutionary Biology* 22: 2496–2504.

CHECKpoint CONCEPT 18.2

- ✓ How have volcanic eruptions and meteorite strikes influenced the course of life's evolution?
- ✓ Explain why an occasional major winter blizzard is irrelevant to discussions of global climate warming.
- ✓ Research and explain the different climatic responses to the "parasol effect" of major volcanic eruptions and the "greenhouse effect" of long-term CO_2 emissions. What is the fundamental difference in the two effects?
- ✓ How have increases in atmospheric concentrations of O_2 affected the evolution of multicellular organisms?

The many dramatic physical events of Earth's history have influenced the nature and timing of evolutionary changes among Earth's living organisms. We will now look more closely at some of the major events that characterize the history of life on Earth.

APPLY THE CONCEPT

Changes in Earth's physical environment have affected the evolution of life

In the experiment shown in Figure 18.8, body mass of individuals in the experimental population of *Drosophila* increased (on average) about 2% per generation in the high-oxygen environment (although the rate of increase was not constant over the experiment). What rate of increase in body mass per generation would be sufficient to account for the giant dragonflies of the Permian?

1. Assume that the average rate of increase in dragonfly size during the Permian was much slower than the rate observed in the experiment in Figure 18.8. We'll assume that the actual rate of increase for dragonflies was just 0.01% per generation, rather than the 2% observed over a few generations for *Drosophila*. We'll assume further that dragonflies complete just one generation per year (as opposed to 40 or more generations for *Drosophila*). Starting with an average body mass of 1 gram, calculate the projected increase in body mass over 50,000 years.*
2. What percentage of the Permian period does 50,000 years represent? Use Table 18.1 for your calculation.
3. Given your calculations, do you think that increased oxygen concentrations during the Permian were sufficient to account for the evolution of giant dragonflies? Why or why not?

*This calculation is similar to computing compound interest for a savings account. Use the formula $W = S\,(1 + R)^N$, where W = the final mass, S = the starting mass (1 gram), R = the rate of increase per generation (0.0001 in this case), and N = the number of generations.

CONCEPT 18.3 Major Events in the Evolution of Life Can Be Read in the Fossil Record

How do we know about the physical changes in Earth's environment and their effects on the evolution of life? To reconstruct life's history, scientists rely heavily on the fossil record. As we have seen, geologists divided Earth's history into eons, eras, and periods based on their distinct fossil assemblages (see Table 18.1). Biologists refer to the assemblage of all organisms of all kinds living at a particular time or place as a **biota**. All of the plants living at a particular time or place are its **flora**, and all of the animals are its **fauna**.

About 300,000 species of fossil organisms have been described, and the number steadily grows. The number of named species, however, is only a tiny fraction of the species that have ever lived. We do not know how many species lived in the past, but we have ways of making reasonable estimates. Of the present-day biota, nearly 1.8 million species have been named. The actual number of living species is probably well over 10 million, and possibly much higher, because many species have not yet been discovered and described by biologists. So the number of described fossil species is only about 3 percent of the estimated minimum number of living species. Life has existed on Earth for about 3.8 billion years. Many species last only a few million years before undergoing speciation or going extinct. From this we know that Earth's biota must have turned over many times during geological history. So the total number of species that have lived over evolutionary time must vastly exceed the number living today. Why have only about 300,000 of these tens of millions of species been described from fossils to date?

Several processes contribute to the paucity of fossils

Only a tiny fraction of organisms ever become fossils, and only a tiny fraction of fossils are ever discovered by paleontologists. Most organisms live and die in oxygen-rich environments in which they quickly decompose. Organisms are not likely to become fossils unless they are transported by wind or water to sites that lack oxygen, where decomposition proceeds slowly or not at all. Furthermore, geological processes transform many rocks, destroying the fossils they contain, and many fossil-bearing rocks are deeply buried and inaccessible. Paleontologists have studied only a tiny fraction of the sites that contain fossils, although they find and describe many new ones every year.

The fossil record is most complete for marine animals that had hard skeletons (which resist decomposition). Among the nine major animal groups with hard-shelled members, approximately 200,000 species have been described from fossils—roughly twice the number of living marine species in these same groups. Paleontologists lean heavily on these groups in their interpretations of the evolution of life. Insects and spiders are also relatively well represented in the fossil record because they are numerically abundant and have hard exoskeletons (FIGURE 18.9). The fossil record, though incomplete, is good enough to document clearly the factual history of the evolution of life.

FIGURE 18.9 Insect Fossils Chunks of amber—fossilized tree resin—often contain insects that were preserved when they were trapped in the sticky resin.

By combining information about physical changes during Earth's history with evidence from the fossil record, scientists have composed portraits of what Earth and its inhabitants may have looked like at different times. We know in general where the continents were and how life changed over time, but many of the details are poorly known, especially for events in the more remote past.

Precambrian life was small and aquatic

Life first appeared on Earth about 3.8 bya (FIGURE 18.10). The fossil record of organisms that lived prior to the Phanerozoic is fragmentary, but it is good enough to establish that the total number of species and individuals increased dramatically in the late Precambrian.

For most of its history, life was confined to the oceans, and all organisms were small. For more than 3 billion years, all organisms lived in shallow seas. These seas slowly began to teem with microscopic prokaryotes. After the first eukaryotes appeared about 1.5 billion years ago, during the Proterozoic, unicellular eukaryotes and small multicellular animals fed on the microorganisms. Small floating organisms, known collectively as **plankton**, were strained from the water and eaten by slightly larger filter-feeding animals. Other animals ingested sediments on the seafloor and digested the remains of organisms within them. But it still took nearly a billion years before eukaryotes began to diversify rapidly into the many different morphological forms that we know today.

What limited the diversity of multicellular eukaryotes (in terms of their size and shape) for much of their early existence? It is likely that a combination of factors was responsible. We have already noted that O_2 levels increased throughout the Proterozoic, and it is likely that high atmospheric and dissolved O_2 concentrations were needed to support large multicellular organisms. In addition, geologic evidence points to a series of intensely cold periods during the late Proterozoic, which would have resulted in seas that were largely covered by ice and continents that were covered by glaciers. The "snowball

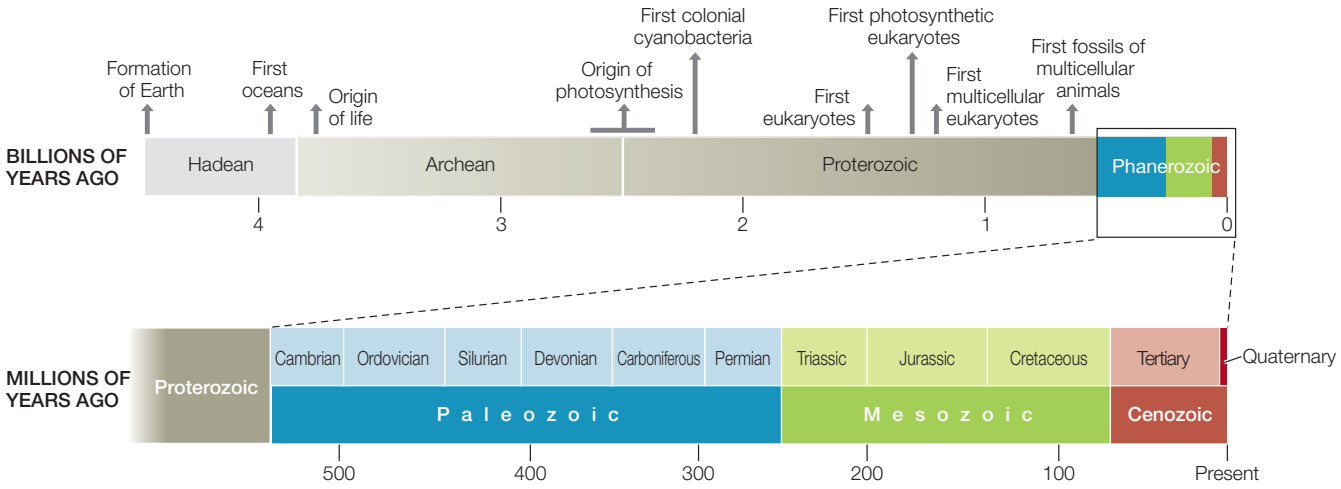


FIGURE 18.10 A Sense of Life's Time The top timeline shows the 4.5-billion-year history of Earth. Most of this history is accounted for by the Hadean, Archean, and Proterozoic eons, a 3.8-billion-year time span that saw the origin of life and the evolution of cells, photosynthesis, and multicellularity. The final 600 million years are expanded in the bottom timeline and detailed in Figure 18.12.

Earth" hypothesis suggests that cold conditions confined life to warm places such as hot springs, deep thermal vents, and perhaps a few equatorial oceans that avoided ice cover. The last of these Proterozoic glaciations ended about 580 million years ago, just before several major radiations of multicellular eukaryotes appear in the fossil record (**FIGURE 18.11**). Many of the multicellular organisms known from the late Proterozoic and early Phanerozoic were very different from any animals living today and may be members of groups that left no living descendants.

Life expanded rapidly during the Cambrian period

The Cambrian period (542–488 mya) marks the beginning of the Paleozoic, the first era of the Phanerozoic. The O_2 concentration in the Cambrian atmosphere was approaching its current level, and the glaciations of the late Proterozoic had ended nearly 40 million years earlier. A geologically rapid diversification of life took place that is often called the **Cambrian explosion**. This name is somewhat misleading, as the series of radiations it refers to actually began before the start of the Cambrian and continued for about 60 million years into the early Cambrian (see Figure 18.11). Nonetheless, 60 million years represents a relatively short amount of time, especially considering that the first eukaryotes had appeared about a billion (= 1,000 million) years earlier. Many of the major animal groups represented by species alive today first appeared during these evolutionary radiations. **FIGURE 18.12** provides an overview of the numerous

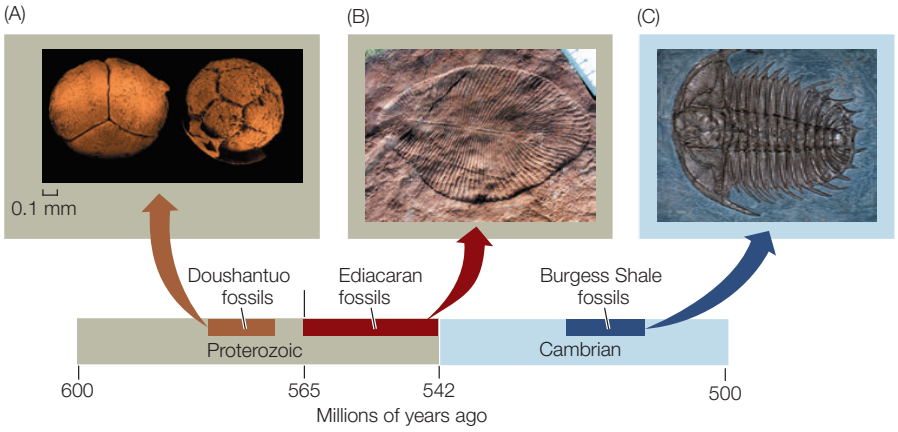
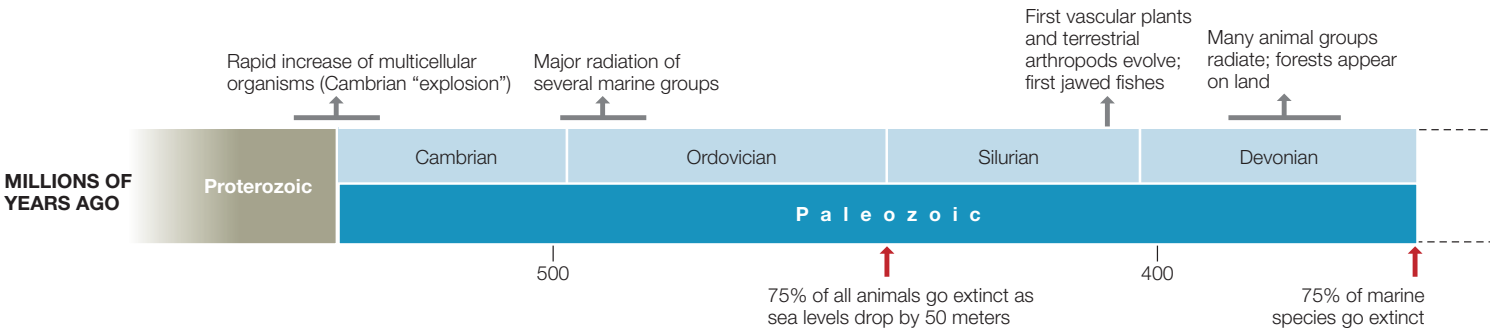
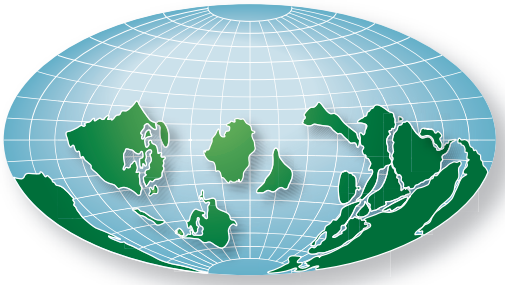


FIGURE 18.11 Diversification of Multicellular Organisms: The "Cambrian Explosion" Shortly after the end of Proterozoic glaciations (about 580 mya), several major radiations of multicellular organisms appear in the fossil record. (A) These microscopic fossils from the Doushantuo rock formation of China are the remains of tiny four- and eight-celled stages of multicellular organisms. (B) Unusual soft-bodied marine invertebrates, unlike any animals alive at present, characterize the fossilized fauna preserved at Ediacara is southern Australia. (C) By the early Phanerozoic, fossilized faunas such as those preserved in Canada's Burgess Shale include extinct representatives of some of the major animal groups alive today.

FIGURE 18.12 A Brief History of Multicellular Life on Earth The geologically rapid "explosion" of life before and during the Cambrian saw the rise of several animal groups that have representatives surviving today. The following three pages depict life's history through the Phanerozoic. The movements of the major continents during the past half-billion years are shown in the maps of Earth, and associated biotas for each time period are depicted. The artists' reconstructions are based on fossils such as those shown in the photographs.





Cambrian

Devonian

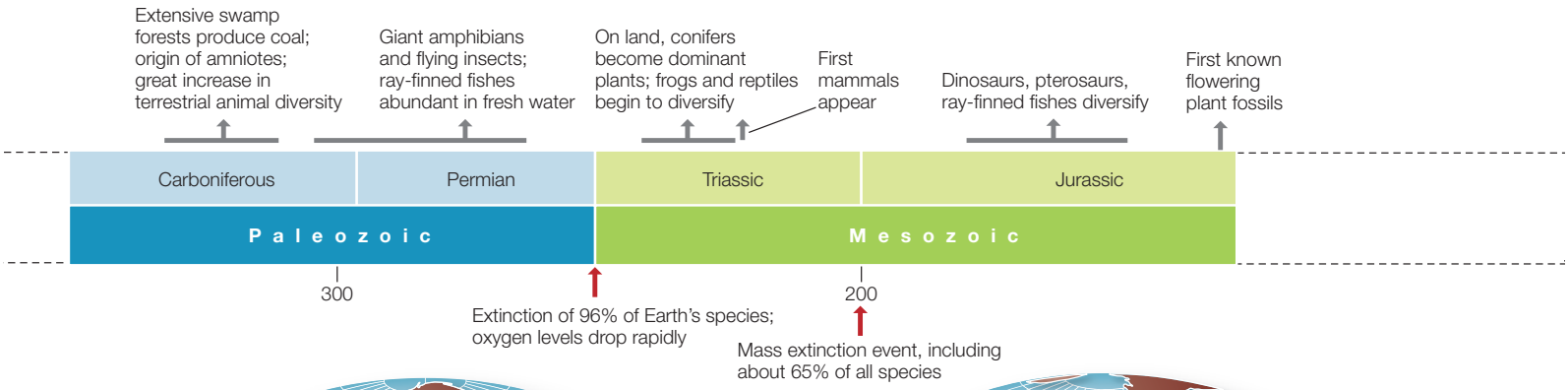
Marrella splendens

Ottoia sp.

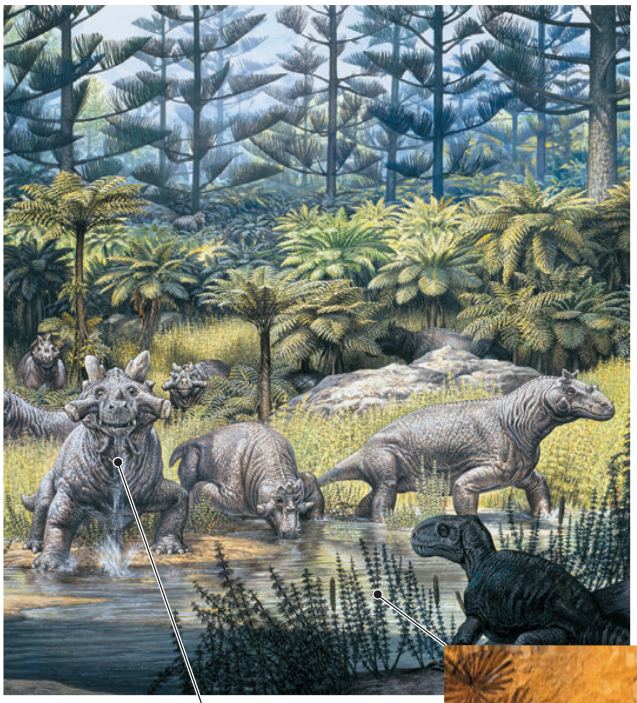
Anomalocaris canadensis (claw only)

Archaeopteris sp.

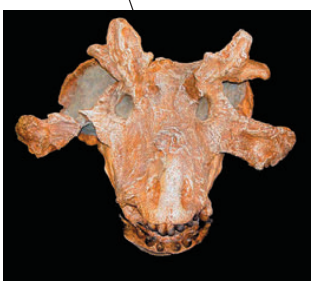
Eusthenopteron foordi



Permian



Jurassic

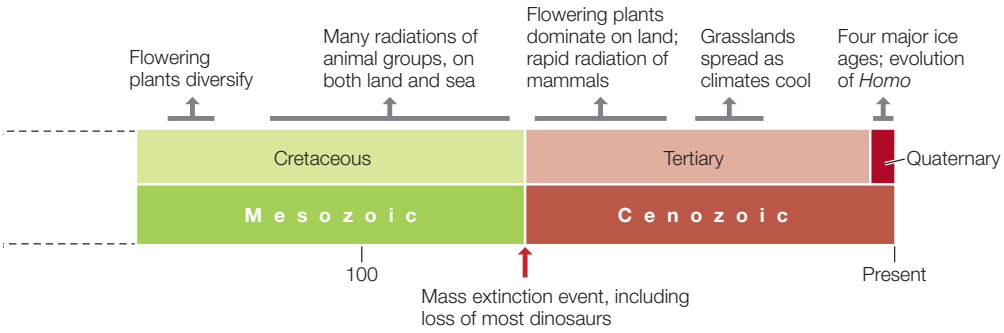


Estemmenosuchus sp.

Equisetum sp.

Ginko sp.

Europasaurus holgeri

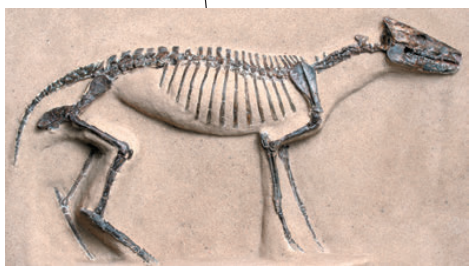


Cretaceous

Sapindopsis belviderensis (leaves)

Chasmosaurus belli

Tertiary



Hyracotherium leporinum

Plesiadapis fodinatus (jaw)

continental and biotic innovations that have characterized the Phanerozoic.

For the most part, fossils tell us only about the hard parts of organisms, but in some well-studied Cambrian fossil beds, the soft parts of many animals were preserved. Multicellular life was largely or completely aquatic during the Cambrian. If there was life on land at this time, it was probably restricted to microorganisms.

Many groups of organisms that arose during the Cambrian later diversified

Geologists divide the remainder of the Paleozoic era into the Ordovician, Silurian, Devonian, Carboniferous, and Permian periods. Each period is characterized by the diversification of specific groups of organisms. Mass extinctions marked the ends of the Ordovician, Devonian, and Permian.

THE ORDOVICIAN (488–444 MYA) During the Ordovician period, the continents, which were located primarily in the Southern Hemisphere, still lacked multicellular life. Evolutionary radiation of marine organisms was spectacular during the early Ordovician, especially among animals, such as brachiopods and mollusks, that lived on the seafloor and filtered small prey from the water. At the end of the Ordovician, as massive glaciers formed over the southern continents, sea levels dropped about 50 meters, and ocean temperatures dropped. About 75 percent of all animal species became extinct, probably because of these major environmental changes.

THE SILURIAN (444–416 MYA) During the Silurian period, the continents began to merge together. Marine life rebounded from the mass extinction at the end of the Ordovician. Animals able to swim in open water and feed above the ocean floor appeared for the first time. Fishes diversified as bony armor gave way to the less rigid scales of modern fishes, and the first jawed fishes and the first fishes with supporting rays in their fins appeared. The tropical sea was uninterrupted by land barriers, and most marine organisms were widely distributed. On land, the first vascular plants evolved late in the Silurian (about 420 mya). The first terrestrial arthropods—scorpions and millipedes—evolved at about the same time.

THE DEVONIAN (416–359 MYA) Rates of evolutionary change accelerated in many groups of organisms during the Devonian period. The major land masses continued to move slowly toward each other. In the oceans there were great evolutionary radiations of corals and of shelled, squidlike cephalopod mollusks.

Terrestrial communities changed dramatically during the Devonian. Club mosses, horsetails, and tree ferns became common, and some attained the size of large trees. Their roots accelerated the weathering of rocks, resulting in the development of the first forest soils. The first plants to produce seeds appeared in the Devonian. The earliest fossil centipedes, spiders, mites, and insects date to this period, as do the earliest terrestrial vertebrates.

A massive extinction of about 75 percent of all marine species marked the end of the Devonian. Paleontologists are uncertain about its cause, but two large meteorites collided with Earth at about that time and may have been responsible, or at least a contributing factor. The continued merging of the continents, with the corresponding reduction in the area of continental shelves, also may have contributed to this mass extinction.

THE CARBONIFEROUS (359–299 MYA) Large glaciers formed over high-latitude portions of the southern land masses during the Carboniferous period, but extensive swamp forests grew on the tropical continents. These forests were dominated by giant tree ferns and horsetails with small leaves. Their fossilized remains formed the coal we now mine for energy. In the seas, crinoids (a group of echinoderms, related to sea stars and sea urchins) reached their greatest diversity, forming "meadows" on the seafloor.

The diversity of terrestrial animals increased greatly during the Carboniferous. Snails, scorpions, centipedes, and insects were abundant and diverse. Insects evolved wings, becoming the first animals to fly. Flight gave herbivorous insects easy access to tall plants, and plant fossils from this period show evidence of chewing by insects (**FIGURE 18.13**). The terrestrial vertebrates split into two lineages. The amphibians became larger and better adapted to terrestrial existence, while the sister lineage led to the amniotes—vertebrates with well-protected eggs that can be laid in dry places.

THE PERMIAN (299–251 MYA) During the Permian period, the continents merged into a single supercontinent called **Pangaea**. Permian rocks contain representatives of many of the major groups of insects we know today. By the end of the period the amniotes had split into two lineages: the reptiles, and a second lineage that would lead to the mammals. Ray-finned fishes became common in the freshwaters of Pangaea.

Toward the end of the Permian, conditions for life deteriorated. Massive volcanic eruptions resulted in outpourings of lava that covered large areas of Earth. The ash and gases produced by the volcanoes blocked sunlight and cooled the climate. The death and decay of the massive Permian forests rapidly used up atmospheric oxygen, and the loss of photosynthetic

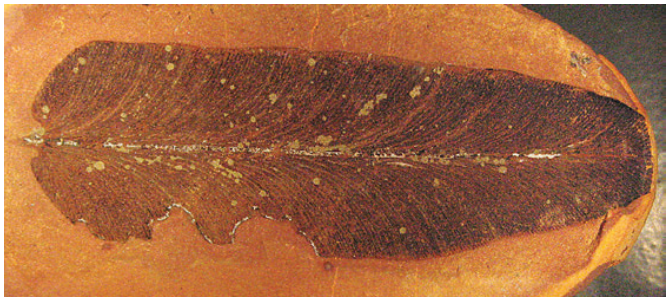


FIGURE 18.13 Evidence of Insect Diversification The margins of this fossil fern leaf from the Carboniferous have been chewed by insects.

organisms meant that relatively little new atmospheric oxygen was produced. In addition, much of Pangaea was located close to the South Pole by the end of the Permian. All of these factors combined to produce the most extensive continental glaciers since the "snowball Earth" times of the late Proterozoic. Atmospheric oxygen concentrations gradually dropped from about 30 percent to 15 percent. At such low concentrations, most animals would have been unable to survive at elevations above 500 meters, so about half of the land area would have been uninhabitable at the end of the Permian. The combination of these changes resulted in the most drastic mass extinction in Earth's history. Scientists estimate that about 96 percent of all multicellular species became extinct at the end of the Permian.

Geographic differentiation increased during the Mesozoic era

The few organisms that survived the Permian mass extinction found themselves in a relatively empty world at the start of the Mesozoic era (251 mya). As Pangaea slowly began to break apart in the Mesozoic, the biotas of the newly separated continents began to diverge. The oceans rose and once again flooded the continental shelves, forming huge, shallow inland seas. Atmospheric oxygen concentrations gradually rose. Life once again proliferated and diversified, but different groups of organisms came to the fore. The three groups of phytoplankton (floating photosynthetic organisms) that dominate today's oceans—dinoflagellates, coccolithophores, and diatoms—became ecologically important at this time, and their remains are the primary origin of the world's oil deposits. Seed-bearing plants replaced the trees that had ruled the Permian forests.

The Mesozoic era is divided into three periods: the Triassic, Jurassic, and Cretaceous. The Triassic and Cretaceous were terminated by mass extinctions, probably caused by meteorite impacts.

THE TRIASSIC (251–201.6 MYA) Pangaea remained largely intact through the Triassic. Many invertebrate groups diversified, and many burrowing animals evolved from groups living on the surfaces of seafloor sediments. On land, conifers and seed ferns were the dominant trees. The first frogs and turtles appeared. A great radiation of reptiles began, which eventually gave rise to crocodilians, dinosaurs, and birds. The first mammals appear. The end of the Triassic was marked by a mass extinction that eliminated about 65 percent of the species on Earth.

THE JURASSIC (201.6–145.5 MYA) Late in the Jurassic period, Pangaea became fully divided into two large continents: **Laurasia**, which drifted northward, and **Gondwana** in the south. Ray-finned fishes rapidly diversified in the oceans. The first lizards appeared, and flying reptiles (pterosaurs) evolved. Most of the large terrestrial predators and herbivores of the period were dinosaurs. Several groups of mammals made their first appearance, and the earliest known fossils of flowering plants are from late in this period.

THE CRETACEOUS (145.5–65.5 MYA) By the mid-Cretaceous period, Laurasia and Gondwana had largely broken apart into the continents we know today (although the Indian subcontinent was still separated from Asia). A continuous sea encircled the tropics. Sea levels were high, and Earth was warm and humid. Life proliferated both on land and in the oceans. Marine invertebrates increased in diversity. On land, the reptile radiation continued as dinosaurs diversified further and the first snakes appeared. Early in the Cretaceous, flowering plants began the radiation that led to their current dominance of the land. By the end of the period, many groups of mammals had appeared.

As described in Concept 18.2, another meteorite-caused mass extinction took place at the end of the Cretaceous. In the seas, many planktonic organisms and bottom-dwelling invertebrates became extinct. On land, almost all animals larger than about 25 kg in body weight became extinct. Many species of insects died out, perhaps because the growth of the plants they fed upon was greatly reduced following the impact. Some species in northern North America and Eurasia survived in areas that were not subjected to the devastating fires that engulfed most low-latitude regions.

Modern biotas evolved during the Cenozoic era

By the early Cenozoic era (65.5 mya), the continents were getting closer to their present positions, but the Indian subcontinent was still separated from Asia, and the Atlantic Ocean was much narrower. The Cenozoic was characterized by an extensive radiation of mammals, but other groups were also undergoing important changes.

Flowering plants diversified extensively and came to dominate world forests, except in the coolest regions, where the forests were composed primarily of gymnosperms. Mutations of two genes in one group of plants (the legumes) allowed them to use atmospheric nitrogen directly by forming symbioses with a few species of nitrogen-fixing bacteria. The evolution of this symbiosis was the first "green revolution" and dramatically increased the amount of nitrogen available for terrestrial plant growth. This symbiosis remains fundamental to the ecological base of life as we know it today.

LINK

The symbiosis between plants and nitrogen-fixing bacteria is covered in detail in **Concept 25.2**

The Cenozoic era is divided into the Tertiary and the Quaternary periods, which are commonly subdivided into **epochs** (**TABLE 18.2**).

THE TERTIARY (65.5–2.6 MYA) During the Tertiary period, the Indian subcontinent continued its northward drift. By about 55 mya it made initial contact with parts of southeastern Asia. By about 35 mya, the Indian Plate ran fully into the Eurasian Plate, and the Himalayas began to be pushed up as a result.

The early Tertiary was a hot and humid time, and the ranges of many plants shifted latitudinally. The tropics were probably too hot to support rainforest vegetation and instead were clothed in low-lying vegetation. In the middle of the Tertiary,

TABLE 18.2 Subdivisions of the Cenozoic Era

Period	Epoch	Onset (mya)
Quaternary	Holocene (Recent)	0.01 (~10,000 years ago)
	Pleistocene	2.6
Tertiary	Pliocene	5.3
	Miocene	23
	Oligocene	34
	Eocene	55.8
	Paleocene	65.5

however, Earth's climate became considerably cooler and drier. Many lineages of flowering plants evolved herbaceous (nonwoody) forms, and grasslands spread over much of Earth.

By the start of the Cenozoic era, invertebrate faunas had already come to resemble those of today. Frogs, snakes, lizards, birds, and mammals all underwent extensive radiations during the Tertiary. Three waves of mammals dispersed from Asia to North America across one of the several land bridges that have intermittently connected the two continents during the past 55 million years. Rodents, marsupials, primates, and hoofed mammals appeared in North America for the first time.

THE QUATERNARY (2.6 MYA TO PRESENT) We are living in the Quaternary period. It is subdivided into two epochs, the Pleistocene and the Holocene (the Holocene is also known as the Recent).

The Pleistocene was a time of drastic cooling and climate fluctuations. During 4 major and about 20 minor "ice ages," massive glaciers spread across the continents, and the ranges of animal and plant populations shifted toward the equator. The last of these glaciers retreated from temperate latitudes less than 15,000 years ago. Organisms are still adjusting to this change. Many high-latitude ecological communities have occupied their current locations for no more than a few thousand years.

It was during the Pleistocene that divergence within one group of mammals, the primates, resulted in the evolution of the hominoid lineage. Subsequent hominoid radiation eventually led to the species *Homo sapiens*—modern humans. Many large bird and mammal species became extinct in Australia and in the Americas when *H. sapiens* arrived on those continents about 45,000 and 15,000 years ago, respectively. Many paleontologists believe these extinctions were the result of hunting and other influences of *Homo sapiens*.

LINK

The evolution of modern humans and their close relatives during the Pleistocene is discussed in **Concept 23.7**

The tree of life is used to reconstruct evolutionary events

The fossil record reveals broad patterns in life's evolution. To reconstruct major events in the history of life, biologists also rely on the phylogenetic information in the tree of life. We can use phylogeny, in combination with the fossil record, to reconstruct the timing of such major events as the acquisition of mitochondria in the ancestral eukaryotic cell, the several independent origins of multicellular organisms, and the movement of life onto dry land. We can also follow major changes in the genomes of organisms, and we can even reconstruct many gene sequences of species that are long extinct.

LINK

Concept 16.3 describes how biologists reconstruct the gene sequences of extinct organisms

Changes in Earth's physical environment have clearly influenced the diversity of organisms we see on the planet today. To study the evolution of that diversity, biologists examine the evolutionary relationships among species. Deciphering phylogenetic relationships is an important step in understanding how life has diversified on Earth. The next part of this book will explore the major groups of life and the different solutions these groups have evolved to major challenges such as reproduction, energy acquisition, dispersal, and escape from predation.

CHECKpoint CONCEPT 18.3

- ✓ Why have so few of the organisms that have existed over Earth's history become fossilized?
- ✓ What do we mean by the "Cambrian explosion"? How long did it last, and in what sense was it an "explosion"?
- ✓ What are some of the ways in which continental drift has affected the evolution of life?

Q Can modern experiments test hypotheses about the evolutionary impact of ancient environmental changes?

ANSWER Several experiments have been conducted to test the link between O_2 concentrations and evolution of body size in flying insects. One of these is discussed in Figure 18.8. The results of these experiments are consistent with the evolution of larger body size in flying insects in hyperoxic (high-oxygen) environments.

Experiments have also been conducted under hypoxic (low-oxygen) conditions, as existed at the end of the Permian. These experiments suggest that the evolution of body size is constrained under hypoxic conditions, even under strong artificial selection for larger body size. These latter results are consistent with the extinction of many large flying insects at the end of the Permian as a result of rapidly decreasing O_2 concentrations. Giant flying insects simply could not have survived the lower O_2 concentrations that existed at that time. The mass extinction at the end of the Permian is the only known mass extinction that involved considerable loss of insect diversity.

SUMMARY

CONCEPT 18.1 Events in Earth's History Can Be Dated

- The relative ages of organisms can be determined by the dating of fossils and the **strata** of **sedimentary rocks** in which they are found.
- **Radiometric dating** techniques use a variety of radioisotopes with different **half-lives** to date events in the remote past. **Review Figure 18.1**
- Geologists divide the history of life into eons, eras, and periods, based on major differences in the fossil assemblages found in successive strata. **Review Table 18.1**

CONCEPT 18.2 Changes in Earth's Physical Environment Have Affected the Evolution of Life

- Earth's crust consists of solid plates that float on fluid magma. **Continental drift** is caused by convection currents in the magma, which move the plates and the continents that lie on top of them. **Review Figure 18.2 and ANIMATED TUTORIAL 18.1**
- Major physical events on Earth, such as continental collisions and volcanic eruptions, have affected Earth's climate, atmosphere, and sea levels. In addition, extraterrestrial events such as meteorite strikes have created sudden and dramatic environmental shifts. All of these changes affected the history of life. **Review Figure 18.3 and Table 18.1**
- Oxygen-generating cyanobacteria liberated enough O_2 to open the door to oxidation reactions in metabolic pathways. Aerobic prokaryotes were able to harvest more energy than anaerobic organisms and began to proliferate. Increases in atmospheric O_2 concentrations supported the evolution of large eukaryotic cells and, eventually, multicellular organisms. **Review Figures 18.7 and 18.8**

CONCEPT 18.3 Major Events in the Evolution of Life Can Be Read in the Fossil Record

- Paleontologists use fossils and evidence of geological changes to determine what Earth and its **biota** may have looked like at different times. **Review Figures 18.11 and 18.12**
- Before the Phanerozoic, life was almost completely confined to the oceans. Multicellular life diversified extensively during the **Cambrian explosion**, a prime example of an evolutionary radiation.
- The periods of the Paleozoic era were each characterized by the diversification of specific groups of organisms. During the Mesozoic era, distinct terrestrial biotas evolved on each continent.
- Five episodes of **mass extinction** punctuated the history of life in the Paleozoic and Mesozoic eras.
- The Cenozoic era is divided into the Tertiary and the Quaternary periods, which in turn are subdivided into **epochs**. This era saw the emergence of the modern biotas as mammals radiated extensively and flowering plants became dominant. **Review Table 18.2**
- The tree of life can be used to reconstruct the timing of evolutionary events.

See **ACTIVITY 18.1** for a concept review of this chapter.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is18

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

PART 4

Diversity

19

Bacteria, Archaea, and Viruses

KEY CONCEPTS

A satellite image reveals thousands of square kilometers of "milky seas" (arrow) in the Indian Ocean. This expanse of bioluminescence is produced by *Vibrio* bacteria.

On the night of January 25, 1995, the British merchant vessel *Lima* was off the coast of Somalia, near the Horn of Africa. This area is infamous for bands of pirates, so the crew was keeping a watchful eye on the seas. On the horizon, they spotted an eerie whitish glow. It was directly in their path, and there was no way to avoid it. Was the glow the result of some strange trick of piracy?

Within 15 minutes of first spotting the glow, the *Lima* was surrounded by glowing waters for as far as her crew could see. As the ship's log recorded, "it appeared as though the ship was sailing over a field of snow or gliding over the clouds." Fortunately for the crew, the glow had nothing to do with pirates.

For centuries, mariners in this part of the world had reported occasional "milky seas" in which the sea surface produced a strange glow at night, extending from horizon to horizon. Scientists to that point had never been able to confirm the reality or the cause of such phenomena. It was well established, however, that many organisms can emit light by bioluminescence—a complex, enzyme-catalyzed biochemical reaction that results in the emission of light but not heat.

What kind of organism could cause the vast expanse of bioluminescence observed by the *Lima*? Some marine organisms emit flashes of light when they are disturbed, but they could not produce the sustained and uniform glow seen in milky seas. The only organisms known to produce the quality of bioluminescence consistent with milky seas are prokaryotes, such as bacteria of the genus *Vibrio*. Using information supplied by the *Lima*, biologists scanned satellite images of the Indian Ocean for the specific light wavelengths emitted by *Vibrio*. The satellite images clearly showed thousands of square kilometers of *Vibrio*-produced milky seas.

Vibrio's bioluminescence requires a critical concentration of a specific chemical signal produced by the bacteria, so at low densities, free-living *Vibrio* populations do not glow. But as a colony establishes itself on phytoplankton, the bacteria's population density increases and concentrations of the luminescence signal build up. Eventually bacterial density (and concentration of the signal) becomes high enough for the huge colony to produce visible light. Such chemical-inducing action among bacterial cells is referred to as quorum sensing.

Q What adaptive advantage does bioluminescence provide to *Vibrio* bacteria?

You will find the answer to this question on page 399.

CONCEPT 19.1 Life Consists of Three Domains That Share a Common Ancestor

You may think that you have little in common with a bacterium. But all multicellular eukaryotes—including you—share many attributes with bacteria and archaea, together called **prokaryotes**. For example, all organisms, whether eukaryotes or prokaryotes,

- have cell membranes and ribosomes (see Chapter 4).
- have a common set of metabolic pathways, such as glycolysis (see Chapter 6).
- replicate DNA semiconservatively (see Chapter 9).
- use DNA as the genetic material to encode proteins, and use a similar genetic code to produce those proteins by transcription and translation (see Chapter 10).

These shared features support the conclusion that all living organisms are related. If life had multiple origins, there would be little reason to expect all organisms to use overwhelmingly similar genetic codes or to share structures as unique as ribosomes. Furthermore, similarities in the DNA sequences of genes that are shared by all organisms confirm the monophyly of life.

Despite these commonalities, major differences have also evolved across the diversity of life. Based on the deepest divisions of phylogenetic relationships, many biologists now recognize three **domains** (primary divisions) of life, two prokaryotic and one eukaryotic (FIGURE 19.1).

All prokaryotic organisms are unicellular, although they may form large coordinated colonies or biofilms consisting of many individuals. The domain Eukarya, by contrast, encompasses both unicellular and multicellular life forms. As was described in Chapter 4, prokaryotic cells differ from eukaryotic cells in some important ways:

- *Prokaryotic cells do not divide by mitosis.* Instead, after replicating their DNA, prokaryotic cells divide by their own method, binary fission (see Concept 7.2).
- *The organization of the genetic material differs.* The DNA of the prokaryotic cell is not organized within a membrane-enclosed nucleus. DNA molecules in prokaryotes are often circular. Many (but not all) prokaryotes have only one main chromosome and are effectively haploid, although many have additional smaller DNA molecules, called plasmids (see Concept 12.2).
- *Prokaryotes lack most of the membrane-enclosed cytoplasmic organelles—mitochondria, Golgi apparatus, and others—that are found in most eukaryotes.* However, the cytoplasm of a prokaryotic cell may contain a variety of infoldings of the cell membrane and photosynthetic membrane systems not found in eukaryotes.

Although the study and classification of eukaryotic organisms goes back centuries, much of our knowledge of the evolutionarily ancient prokaryotic domains is quite recent. Not until the final quarter of the twentieth century did advances in molecular genetics and biochemistry reach a point that enabled research that revealed deep-seated distinctions between the domains Bacteria and Archaea.

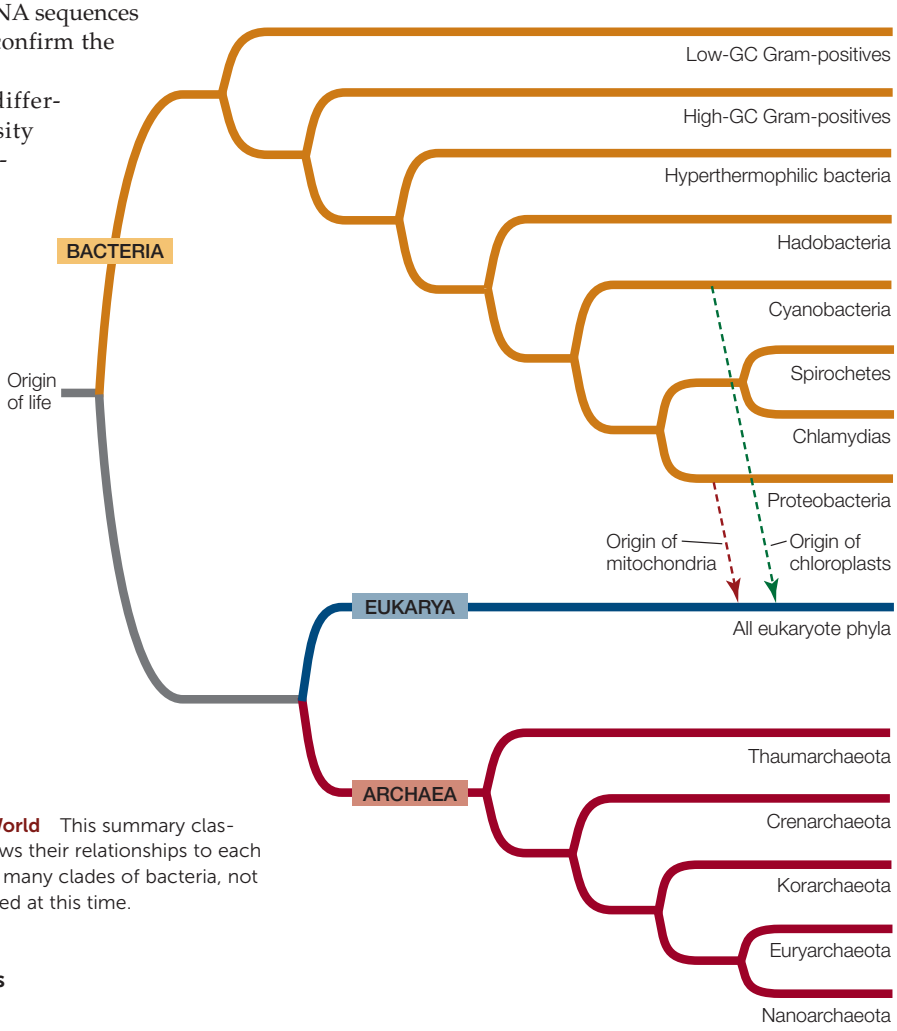


FIGURE 19.1 The Three Domains of the Living World This summary classification of the domains Bacteria and Archaea shows their relationships to each other and to Eukarya. The relationships among the many clades of bacteria, not all of which are listed here, are incompletely resolved at this time.

Go to ANIMATED TUTORIAL 19.1
The Evolution of the Three Domains
PoL2e.com/at19.1

TABLE 19.1 The Three Domains of Life on Earth

	Domain		
Characteristic	**Bacteria**	**Archaea**	**Eukarya**
Membrane-enclosed nucleus	Absent	Absent	**Present**
Membrane-enclosed organelles	Few	Absent	**Many**
Peptidoglycan in cell wall	**Present**	Absent	Absent
Membrane lipids	Ester-linked	**Ether-linked**	Ester-linked
	Unbranched	**Branched**	Unbranched
Ribosomes[a]	70S	70S	**80S**
Initiator tRNA	**Formylmethionine**	Methionine	Methionine
Operons	Yes	Yes	**Rare**
Plasmids	Yes	Yes	**Rare**
RNA polymerases	One	One[b]	**Three**
Ribosomes sensitive to chloramphenicol and streptomycin	**Yes**	No	No
Ribosomes sensitive to diphtheria toxin	**No**	Yes	Yes

[a]70S ribosomes are smaller than 80S ribosomes.
[b]Archaeal RNA polymerase is similar to eukaryotic polymerases.

The two prokaryotic domains differ in significant ways

A glance at **TABLE 19.1** will show you that there are major differences (most of which cannot be seen even under an electron microscope) between the two prokaryotic domains. In some ways archaea are more like eukaryotes; in other ways they are more like bacteria. (Note that we use lowercase when referring to the members of these domains and uppercase when referring to the domains themselves.) The basic unit of an archaeon (the term for a single archaeal organism) or bacterium (a single bacterial organism) is the prokaryotic cell. Each single-celled organism contains a full complement of genetic and protein-synthesizing systems, including DNA, RNA, and all the enzymes needed to transcribe and translate the genetic information into proteins. The prokaryotic cell also contains at least one system for generating the ATP it needs.

Genetic studies clearly indicate that all three domains had a single common ancestor. Across a major portion of their genome, eukaryotes share a more recent common ancestor with Archaea than they do with Bacteria (see Figure 19.1). However, the mitochondria of eukaryotes (as well as the chloroplasts of photosynthetic eukaryotes, such as plants) originated through endosymbiosis with a bacterium. Some biologists prefer to view the origin of eukaryotes as a fusion of two equal partners (one ancestor that was related to modern archaea, and another that was more closely related to modern bacteria). Others view the divergence of the early eukaryotes from the archaea as a separate and earlier event than the later endosymbioses. In either case, some eukaryote genes are most closely related to those of archaea, whereas others are most closely related to those of bacteria. The tree of life therefore contains some merging of lineages as well as the predominant divergence of lineages.

LINK

The origin of mitochondria and chloroplasts by endosymbiosis is described in **Concept 20.1**

Biologists estimate that the last common ancestor of the three domains lived about 3 billion years ago. We can deduce that it had DNA as its genetic material, and that its machinery for transcription and translation produced RNAs and proteins, respectively. This ancestor likely had a circular chromosome. Archaea, Bacteria, and Eukarya are all the products of billions of years of mutation, natural selection, and genetic drift, and they are all well adapted to their present-day environments. The earliest prokaryote fossils, which date back at least 3.5 billion years, indicate that there was considerable diversity among the prokaryotes even during those earliest days of life.

The small size of prokaryotes has hindered our study of their evolutionary relationships

Until about 300 years ago, nobody had even *seen* an individual prokaryote. Most prokaryotes remained invisible to humans until the invention of the first simple microscope. Prokaryotes are so small, however, that even the best light microscopes don't reveal much about them. It took advanced microscopic equipment and modern molecular techniques to open up the microbial world. (Microscopic organisms—both prokaryotes and eukaryotes—are often collectively referred to as "microbes.")

Before DNA sequencing became practical, taxonomists based prokaryote classification on observable characters such as shape, color, motility, nutritional requirements, and sensitivity to antibiotics. One of the characters most widely used to classify prokaryotes is the structure of their cell walls.

The cell walls of almost all bacteria contain **peptidoglycan**, a polymer that produces a meshlike structure around the cell. Peptidoglycan is a substance unique to bacteria. The absence of peptidoglycan from the cell walls of archaea is a key difference between the two prokaryotic domains. Peptidoglycan is also an excellent target for combating pathogenic (disease-causing) bacteria because it has no counterpart in eukaryotic cells. Antibiotics such as penicillin and ampicillin, as well as other agents that specifically interfere with the synthesis of

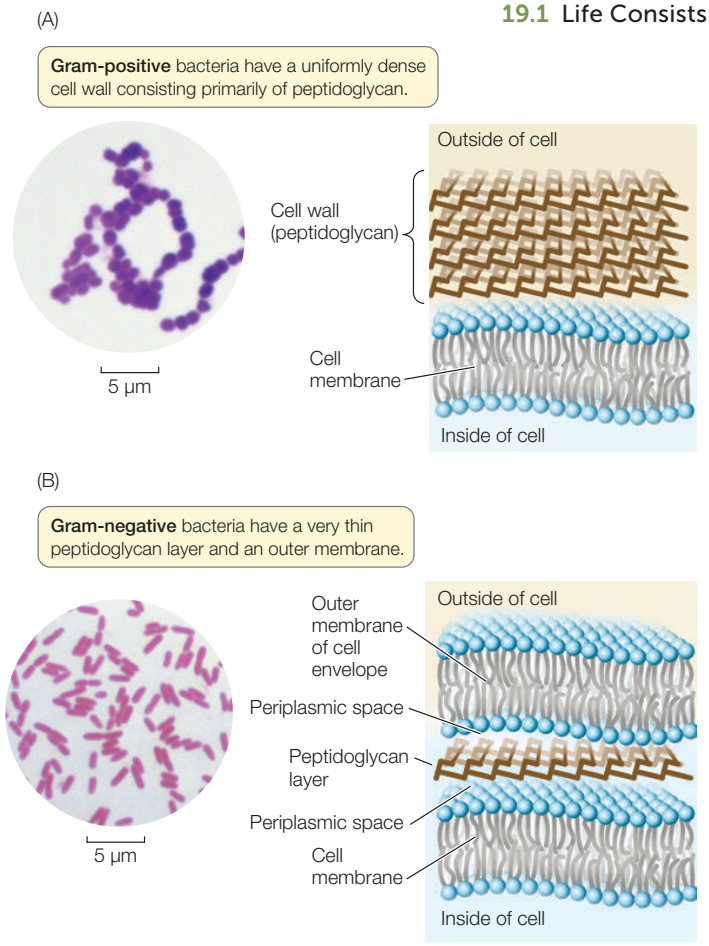


FIGURE 19.2 The Gram Stain and the Bacterial Cell Wall When treated with Gram stain, the cell walls of bacteria react in one of two ways. (A) Gram-positive bacteria have a thick peptidoglycan cell wall that retains the violet dye and appears deep blue or purple. (B) Gram-negative bacteria have a thin peptidoglycan layer that does not retain the violet dye, but picks up the counterstain and appears pink to red.

Go to ACTIVITY 19.1 Gram Stain and Bacteria
PoL2e.com/ac19.1

peptidoglycan-containing cell walls, tend to have little, if any, effect on the cells of humans and other eukaryotes.

The **Gram stain** is a laboratory technique that can be used to classify most types of bacteria into two distinct groups, Gram-positive and Gram-negative. A smear of bacterial cells on a microscope slide is soaked in a violet dye and treated with iodine. Next it is washed with alcohol and counterstained with a red dye (safranin). **Gram-positive bacteria** retain the violet dye and appear blue to purple (FIGURE 19.2A). The alcohol washes the violet stain out of **Gram-negative bacteria**, which then pick up the safranin counterstain and appear pink to red (FIGURE 19.2B). For most bacteria, the effect of the Gram stain is determined by the chemical structure of the cell wall:

- *A Gram-negative cell wall* usually has a thin peptidoglycan layer, which is surrounded by a second, outer membrane quite distinct in chemical makeup from the cell membrane (see Figure 19.2B). The space between the cell membrane and the outer membrane (known as the periplasmic space) contains proteins that are important in digesting some materials, transporting others, and detecting chemical gradients in the environment.
- *A Gram-positive cell wall* usually has about five times as much peptidoglycan as a Gram-negative cell wall. Its thick peptidoglycan layer is a meshwork that may serve some of the same purposes as the periplasmic space of the Gram-negative cell envelope (cell wall plus outer membrane).

Shape is another phenotypic characteristic that is useful for the basic identification of bacteria. Three major shapes are common—spheres, rods, and spiral forms (FIGURE 19.3). Many bacterial names are based on these shapes. A spherical bacterium is called a **coccus** (plural *cocci*). Cocci may live singly or may associate in two- or three-dimensional arrays as chains, plates, blocks, or clusters of cells. A rod-shaped bacterium is called a **bacillus** (plural *bacilli*). A spiral bacterium (shaped like a corkscrew) is called a **spirillum** (plural *spirilla*). Bacilli and spirilla may be single, form chains, or gather in regular clusters. Among the other bacterial shapes are long filaments and branched filaments.

Less is known about the shapes of archaea because many of these organisms have never been seen. Many archaea are known only from samples of DNA from the environment. However, the species whose morphologies are known include cocci, bacilli, and even triangular and square-shaped species. Some flattened species grow on surfaces, arranged like sheets of postage stamps.

The nucleotide sequences of prokaryotes reveal their evolutionary relationships

Analyses of the nucleotide sequences of ribosomal RNA (rRNA) genes provided the first comprehensive evidence of evolutionary relationships among prokaryotes. For several

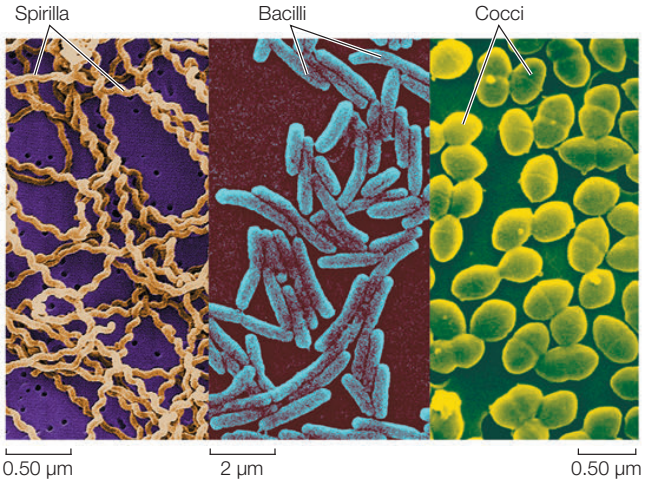


FIGURE 19.3 Bacterial Cell Shapes This composite colorized micrograph shows three common bacterial shapes. The spiral-shaped spirilla are *Leptospira* sp., a human pathogen. Rod-shaped cells are called bacilli; these *Legionella pneumophila* reside in the gut. Spherical cells are called cocci; these are a species of *Enterococcus*, also from the mammalian gut.

reasons, rRNA is particularly useful for phylogenetic studies of living organisms:

- rRNA was present in the common ancestor of all life and is therefore evolutionarily ancient.
- No free-living organism lacks rRNA, so rRNA genes can be compared throughout the tree of life.
- rRNA plays a critical role in translation in all organisms, so lateral transfer of rRNA genes among distantly related species is unlikely.
- rRNA has evolved slowly enough that gene sequences from even distantly related species can be aligned and analyzed.

Comparisons of rRNA genes from a great many organisms have revealed the probable phylogenetic relationships throughout the tree of life. Databases such as GenBank contain rRNA gene sequences from hundreds of thousands of species—more than any other type of gene sequence.

Although studies of rRNA genes reveal much about the evolutionary relationships of prokaryotes, they don't always reveal the entire evolutionary history of these organisms. In some groups of prokaryotes, analyses of multiple gene sequences have suggested several different phylogenetic patterns. How could such differences among different gene sequences arise? Studies of whole prokaryotic genomes have revealed that even distantly related prokaryotes sometimes exchange genetic material.

Lateral gene transfer can lead to discordant gene trees

As noted earlier, prokaryotes reproduce by binary fission. If we could follow these divisions back through evolutionary time, we would be tracing the path of the complete tree of life for bacteria and archaea. This underlying tree of relationships (represented in highly abbreviated form in Appendix A) is called the organismal (or species) tree. Because whole genomes are replicated during binary fission, we would expect phylogenetic trees constructed from most gene sequences (see Chapter 16) to reflect these same relationships.

Even though binary fission is an asexual process, there are other processes—including transformation, conjugation, and transduction—that allow the exchange of genetic information between some prokaryotes without reproduction. Thus prokaryotes can exchange and recombine their DNA with that of other individuals (this is sex in the genetic sense of the word), but this genetic exchange is not directly linked to reproduction, as it is in most eukaryotes.

LINK

Prokaryote exchange of genetic material by transformation, conjugation, and transduction is described in **Concept 8.4**

From early in evolution to the present day, some genes have been moving "sideways" from one prokaryote species to another, a phenomenon known as **lateral gene transfer**. Lateral gene transfers are well documented, especially among closely related species; some have been documented even across the domains of life.

Consider, for example, the genome of *Thermotoga maritima*, a bacterium that can survive extremely high temperatures. In comparing the 1,869 gene sequences of *T. maritima* against sequences encoding the same proteins in other species, investigators found that some of this bacterium's genes have their closest relationships not with the genes of other bacterial species, but with the genes of archaea that live in similar extreme environments.

When genes involved in lateral transfer events are sequenced and analyzed, the resulting individual gene trees will not match the organismal tree in every respect (**FIGURE 19.4**). The individual gene trees will vary because the history

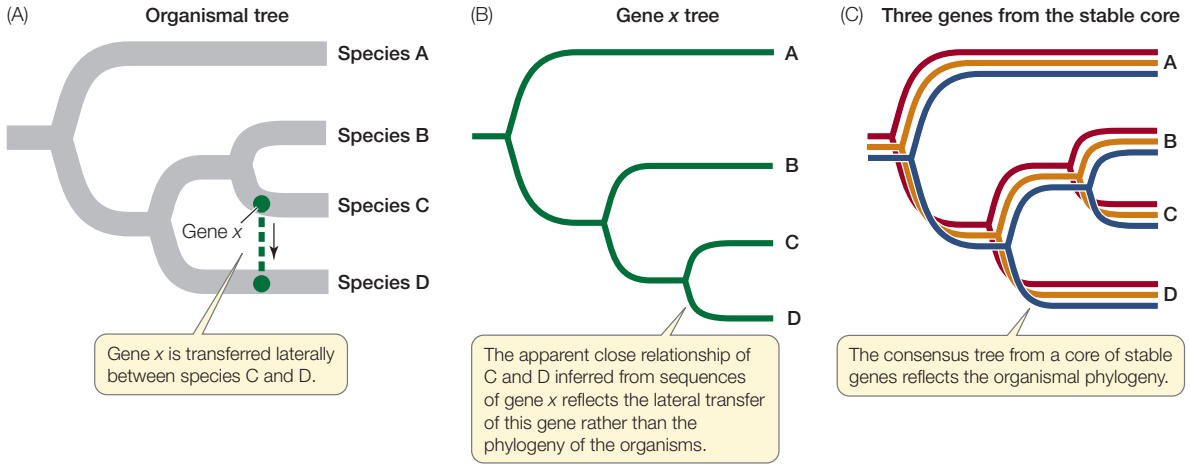


FIGURE 19.4 Lateral Gene Transfer Complicates Phylogenetic Relationships (A) The phylogeny of four hypothetical prokaryote species, two of which have been involved in a lateral transfer of gene *x*. (B) A tree based only on gene *x* shows the phylogeny of the laterally transferred gene, rather than the organismal phylogeny. (C) In many cases, a consensus tree based on a "stable core" of genes accurately reflects the organismal phylogeny.

of lateral transfer events is different for each gene. Biologists can reconstruct the underlying organismal phylogeny by comparing multiple genes (to produce a consensus tree) or by concentrating on genes that are unlikely to be involved in lateral gene transfer events. For example, genes that are involved in fundamental cellular processes (such as the rRNA genes discussed above) are unlikely to be replaced by the same genes from other species because functional, locally adapted copies of these genes are already present.

What kinds of genes are most likely to be involved in lateral gene transfer? Genes that result in a new adaptation that confers higher fitness on a recipient species are most likely to be transferred repeatedly among species. For example, genes that produce antibiotic resistance are often transferred among bacterial species on plasmids, especially under the strong selection pressure such as that imposed by modern antibiotic medications. Improper or overly frequent use of antibiotics can select for resistant strains of bacteria that are much harder to treat. This selection for antibiotic resistance explains why informed physicians have become more careful in prescribing antibiotics.

It is debatable whether lateral gene transfer has seriously complicated our attempts to resolve the tree of prokaryotic life. Recent work suggests that it has not. Lateral gene transfer rarely creates problems at higher taxonomic levels, even though it may complicate our understanding of the relationships among individual species. Some species clearly obtain some of their genes from otherwise distantly related species, so evolutionary histories of individual genes may differ within a single organism. But it is now possible to make nucleotide sequence comparisons involving entire genomes, and these studies are revealing a stable core of crucial genes that are uncomplicated by lateral gene transfer. Gene trees based on this stable core more accurately reveal the organismal phylogeny (see Figure 19.4). The problem remains, however, that only a very small proportion of the prokaryotic world has been described and studied.

The great majority of prokaryote species have never been studied

Most prokaryotes have defied all attempts to grow them in pure culture, causing biologists to wonder how many species, and possibly even clades, we might be missing. A window onto this problem was opened with the introduction of a new way of examining nucleic acid sequences. When biologists are unable to work with the whole genome of a single prokaryote species, they can instead examine individual genes collected from a random sample of the environment (see Figure 12.6).

Norman Pace of the University of Colorado isolated individual rRNA gene sequences from extracts of environmental samples such as soil and seawater. Comparing such sequences with previously known ones revealed that an extraordinary number of the sequences were new, implying that they came from previously unrecognized species. Biologists have described only about 10,000 species of bacteria and only a few hundred species of archaea (see Figure 1.4). The results of Pace's and similar studies suggest that there may be millions—perhaps hundreds of millions—of prokaryote species on Earth. Other biologists put the estimate much lower, arguing that the high dispersal ability of many bacterial species greatly reduces local endemism (i.e., the number of species restricted to a small geographic area). Only the magnitude of these estimates differs, however; all sides agree that we have just begun to uncover Earth's bacterial and archaeal diversity.

CHECKpoint CONCEPT 19.1

- ✓ Why were all prokaryotes once considered members of a single clade, and what findings led to the establishment of Bacteria and Archaea as separate domains?
- ✓ How did biologists classify bacteria before it became possible to determine nucleotide sequences?
- ✓ Why are nucleotide sequences of rRNA genes particularly useful for evolutionary studies?
- ✓ How does lateral gene transfer complicate evolutionary studies?

Despite the challenges of reconstructing the phylogeny of prokaryotes, taxonomists are beginning to establish evolutionary classification systems for these organisms. With a full understanding that new information requires periodic revisions in these classifications, we will next apply a current system of classification to organize our survey of prokaryote diversity.

CONCEPT 19.2 Prokaryote Diversity Reflects the Ancient Origins of Life

The prokaryotes were alone on Earth for a very long time, adapting to new environments and to changes in existing environments. They have survived to this day, in massive numbers and incredible diversity, and they are found everywhere. In numbers of individuals, prokaryotes are far more abundant than eukaryotes are. Individual bacteria and archaea in the oceans number more than 3×10^{28}—about a million times more than the number of stars in the universe. Closer to home, the individual bacteria living in your intestinal tract outnumber all the humans who have ever lived.

Given our still fragmentary knowledge of prokaryote diversity, it is not surprising that there are several different hypotheses about the relationships of the major groups of prokaryotes. In this book we use a widely accepted classification system that has considerable support from nucleotide sequence data. We will discuss eight major bacterial groups that have the broadest phylogenetic support and have received the most study: the low-GC Gram-positives, high-GC Gram-positives, hyperthermophilic bacteria, hadobacteria, cyanobacteria, spirochetes, chlamydias, and proteobacteria (see Figure 19.1). We will then describe the archaea, whose diversity is even less well studied than that of the bacteria.

The low-GC Gram-positives include the smallest cellular organisms

The **low-GC Gram-positives**, also known as Firmicutes, derive the first part of their name from the relatively low ratio of G-C to A-T nucleotide base pairs in their DNA. The second part of their name is less accurate: some of the low-GC Gram-positives are in fact Gram-negative, and some have no cell wall at all. Despite these differences, phylogenetic analyses of DNA sequences support the monophyly of this bacterial clade.

One group of low-GC Gram-positives can produce heat-resistant resting structures called **endospores** (FIGURE 19.5). When a key nutrient such as nitrogen or carbon becomes scarce, the bacterium replicates its DNA and encapsulates one copy, along with some of its cytoplasm, in a tough cell wall heavily thickened with peptidoglycan and surrounded by a spore coat. The parent cell then breaks down, releasing the endospore. Endospore production is not a reproductive process, as the endospore merely replaces the parent cell. The endospore, however, can survive harsh environmental conditions that would kill the parent cell, such as high or low temperatures or drought, because it is dormant—its normal metabolic activity is suspended. Later, if it encounters favorable conditions, the endospore becomes metabolically active and divides, forming new cells that are like the parent cells. Members of this endospore-forming group of low-GC Gram-positives include the many species of *Clostridium* and *Bacillus*. Some of their endospores can be reactivated after more than 1,000 years of dormancy. There are even credible claims of reactivation of *Bacillus* endospores that are millions of years old.

FIGURE 19.5 A Structure for Waiting Out Bad Times Some low-GC Gram-positive bacteria can replicate their DNA and encase it in a resistant endospore. In harsh conditions, the parent cell breaks down and the endospore survives in a dormant state until conditions improve.

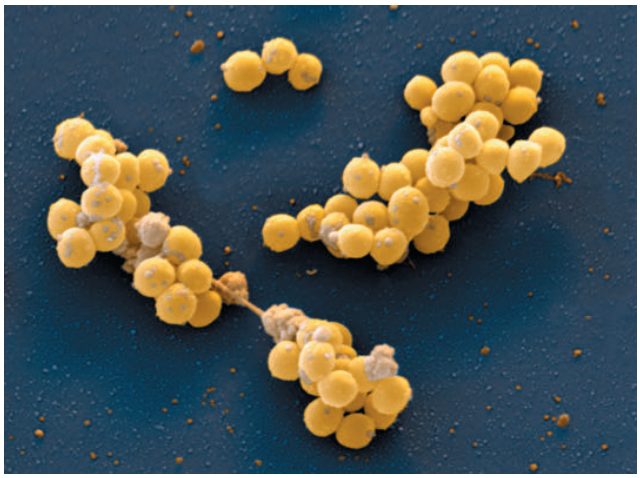


FIGURE 19.6 Staphylococci "Grape clusters" are the usual arrangement of these low-GC Gram-positive coccal bacteria, often the cause of skin or wound infections.

Endospores of *Bacillus anthracis* are the cause of anthrax. Anthrax is primarily a disease of grazing animals, but it can be fatal in humans if the infection starts in the respiratory system. When the endospores sense macrophages in mammalian blood, they reactivate and release toxins into the bloodstream. *Bacillus anthracis* has been used as a bioterrorism agent because large quantities of its endospores can be transported in a small space and then spread among human populations, where they may be inhaled or ingested.

Low-GC Gram-positives of the genus *Staphylococcus*—the **staphylococci** (FIGURE 19.6)—are abundant on the human body surface; they are responsible for boils and many other skin problems. *Staphylococcus aureus* is the best-known human pathogen in this genus; it is present in 20–40 percent of normal adults (and in 50–70 percent of hospitalized adults). In addition to skin diseases, *S. aureus* can cause respiratory, intestinal, and wound infections.

Another interesting group of low-GC Gram-positives, the **mycoplasmas**, lack cell walls, although some have a stiffening material outside the cell membrane. The mycoplasmas are among the smallest cellular organisms known (FIGURE 19.7). The smallest mycoplasmas capable of multiplying by binary fission have a diameter of about 0.2 µm. They are small in another crucial sense as well: they have less than half as much DNA as most other prokaryotes. It has been speculated that the DNA in a mycoplasma, which codes for fewer than 500 proteins, may be close to the minimum amount required to encode the essential properties of a living cell.

Some high-GC Gram-positives are valuable sources of antibiotics

High-GC Gram-positives, also known as actinobacteria, have a higher ratio of G-C to A-T nucleotide base pairs than do the low-GC Gram-positives. These bacteria develop an elaborately branched system of filaments (FIGURE 19.8) that resembles the

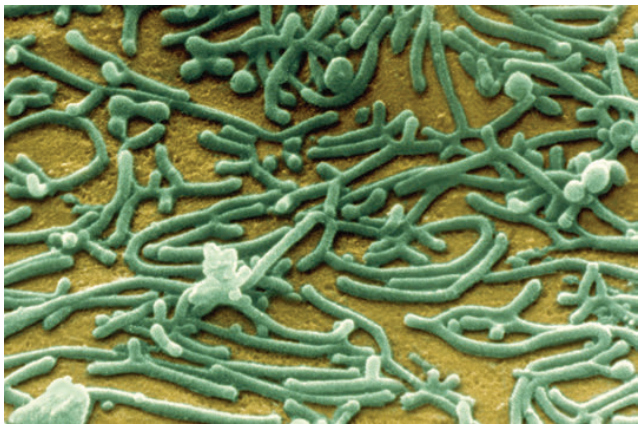


FIGURE 19.7 Tiny Cells With about one-fifth as much DNA as *E. coli*, mycoplasmas are among the smallest known bacteria.

filamentous growth habit of fungi, albeit at a reduced scale. Some high-GC Gram-positives reproduce by forming chains of spores at the tips of the filaments. In species that do not form spores, the branched, filamentous growth ceases and the structure breaks up into typical cocci or bacilli, which then reproduce by binary fission.

The high-GC Gram-positives include several medically important bacteria. *Mycobacterium tuberculosis* causes tuberculosis, which kills 3 million people each year. Genetic data suggest that this bacterium arose 3 million years ago in East Africa, making it the oldest known human bacterial pathogen. The genus *Streptomyces* produces streptomycin as well as hundreds of other antibiotics. We derive most of our antibiotics from members of the high-GC Gram-positives.

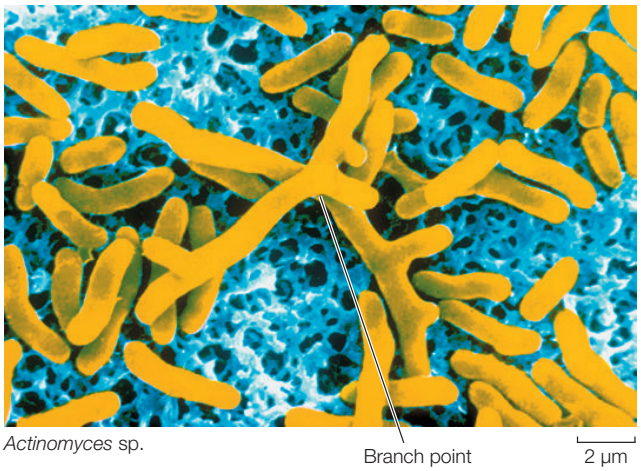


FIGURE 19.8 Actinobacteria Are High-GC Gram-Positives The tangled, branching filaments seen in this scanning electron micrograph are typical of this medically important bacterial group.

Hyperthermophilic bacteria live at very high temperatures

Several lineages of bacteria and archaea are **extremophiles**: they thrive under extreme conditions that would kill most other organisms. The **hyperthermophilic bacteria**, for example, are thermophiles (Greek, "heat-lovers"). Genera such as *Aquifex* live near volcanic vents and in hot springs, sometimes at temperatures near the boiling point of water. Some species of *Aquifex* need only hydrogen, oxygen, carbon dioxide, and mineral salts to live and grow. Species of the genus *Thermotoga* live deep underground in oil reservoirs, as well as in other high-temperature environments.

Biologists have hypothesized that high temperatures characterized the ancestral conditions for life on Earth, given that most environments on early Earth were much hotter than those of today. Reconstructions of ancestral bacterial genes have supported this hypothesis by showing that the ancestral sequences functioned best at elevated temperatures. The monophyly of the hyperthermophilic bacteria, however, is not well established.

Hadobacteria live in extreme environments

The **hadobacteria**, including such genera as *Deinococcus* and *Thermus*, are another group of thermophilic extremophiles. The group's name is derived from Hades, the ancient Greek name for the underworld. *Deinococcus* are resistant to radiation and can consume nuclear waste and other toxic materials. They can also survive extremes of cold as well as hot temperatures. Another member of this group, *Thermus aquaticus*, was the source of the thermally stable DNA polymerase that was critical for the development of the polymerase chain reaction. *Thermus aquaticus* was originally isolated from a hot spring, but it can be found wherever hot water occurs (including many residential hot water heaters).

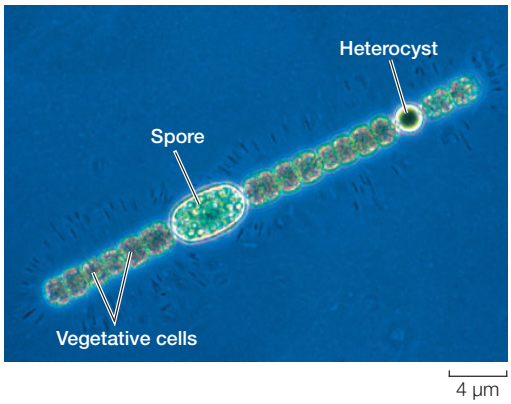
Cyanobacteria were the first photosynthesizers

Cyanobacteria, sometimes called blue-green bacteria because of their pigmentation, are photosynthetic bacteria that require only water, nitrogen gas, oxygen, a few mineral elements, light, and carbon dioxide to survive. They use chlorophyll *a* for photosynthesis and release oxygen gas (O_2); many species also fix nitrogen. Photosynthesis by these bacteria was the basis of the "green revolution" that transformed Earth's atmosphere (see Concept 18.3).

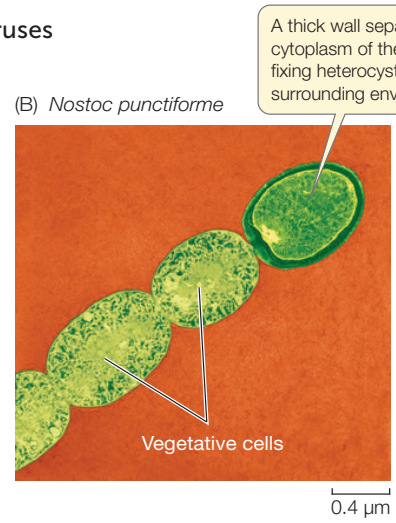
Cyanobacteria carry out the same type of photosynthesis that is characteristic of eukaryotic photosynthesizers. They contain elaborate and highly organized internal membrane systems called **photosynthetic lamellae**. As mentioned in Concept 19.1, the chloroplasts of photosynthetic eukaryotes are derived from an endosymbiotic cyanobacterium.

Cyanobacteria may live free as single cells or associate in multicellular colonies. Depending on the species and on growth conditions, these colonies may range from flat sheets one cell thick to filaments to spherical balls of cells. Some filamentous colonies of cyanobacteria differentiate into three specialized cell types: vegetative cells, spores, and heterocysts. **Vegetative cells** photosynthesize, **spores** are resting stages that

(A) *Anabaena* sp.



(B) *Nostoc punctiforme*



(C)

FIGURE 19.9 Cyanobacteria (A) Some cyanobacteria form filamentous colonies containing three cell types, including reproductive spores and photosynthesizing vegetative cells. (B) Heterocysts are specialized for nitrogen fixation and may serve as a breaking point when filaments reproduce. (C) This pond in Finland has experienced eutrophication: phosphorus and other nutrients generated by human activity have accumulated, feeding an immense green mat (commonly referred to as "pond scum") that is made up of several species of free-living cyanobacteria.

Go to MEDIA CLIP 19.1
Cyanobacteria
PoL2e.com/mc19.1

can survive harsh environmental conditions and eventually develop into new filaments, and **heterocysts** are cells specialized for nitrogen fixation (**FIGURE 19.9**). All of the known cyanobacteria with heterocysts fix nitrogen. Heterocysts also have a role in reproduction: when filaments break apart to reproduce, the heterocyst may serve as a breaking point.

Spirochetes move by means of axial filaments

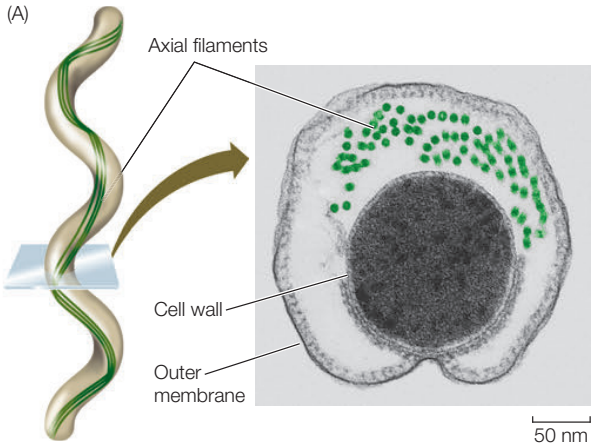
Spirochetes are Gram-negative, motile bacteria characterized by unique structures called axial filaments (**FIGURE 19.10A**), which are modified flagella running through the periplasmic space. The cell body is a long cylinder coiled into a spiral (**FIGURE 19.10B**). The axial filaments begin at either end of the cell and overlap in the middle. Motor proteins connect the axial filaments to the cell wall, enabling the corkscrew-like movement of the bacterium. Many spirochetes are parasites of humans; a few are pathogens, including those that cause syphilis and Lyme disease. Others live free in mud or water.

Chlamydias are extremely small parasites

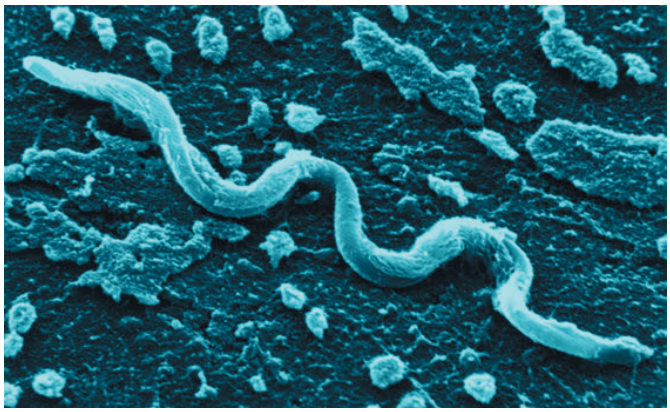
Chlamydias are among the smallest bacteria (0.2–1.5 µm in diameter). They can live only as parasites in the cells of other organisms. It was once believed that their obligate parasitism

FIGURE 19.10 Spirochetes Get Their Shape from Axial Filaments (A) A spirochete from the gut of a termite, seen in cross section, shows the axial filaments used to produce a corkscrew-like movement of these spiral prokaryotes. (B) This spirochete species causes syphilis in humans.

(A)



(B)



Cristispira sp.

2 µm

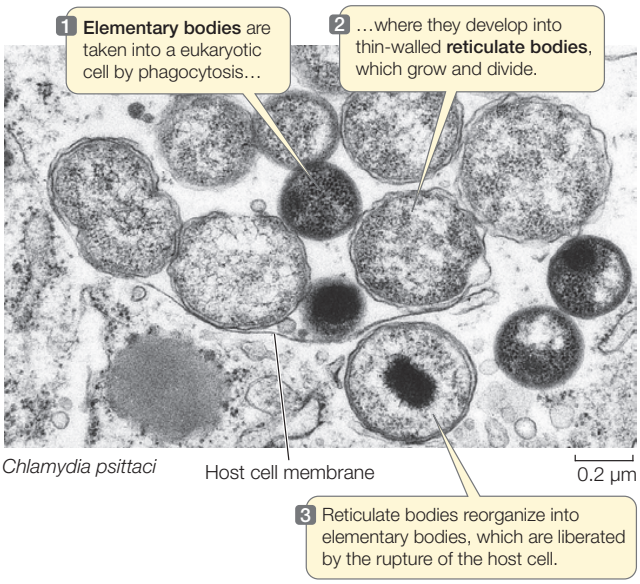


FIGURE 19.11 Chlamydias Change Form during Their Life Cycle
Elementary bodies and reticulate bodies are the two major phases of the chlamydia life cycle.

resulted from an inability to produce ATP—that chlamydias were "energy parasites." However, genome sequencing indicates that chlamydias have the genetic capacity to produce at least some ATP. They can augment this capacity by using an enzyme called a translocase, which allows them to take up ATP from the cytoplasm of their host in exchange for ADP from their own cells.

These tiny, Gram-negative cocci are unique among prokaryotes because of a complex life cycle that involves two different forms of cells, elementary bodies and reticulate bodies (**FIGURE 19.11**). Various strains of chlamydias cause eye infections (especially trachoma), sexually transmitted diseases, and some forms of pneumonia in humans.

The proteobacteria are a large and diverse group

By far the largest group of bacteria, in terms of number of described species, is the **proteobacteria**. The proteobacteria include many species of Gram-negative photoautotrophs that use light-driven reactions to metabolize sulfur, as well as dramatically diverse bacteria that bear no phenotypic resemblance to the photoautotrophic species. Genetic and morphological evidence indicates that the mitochondria of eukaryotes were derived from a proteobacterium by endosymbiosis.

Among the proteobacteria are some nitrogen-fixing genera, such as *Rhizobium*, and other bacteria that contribute to the global nitrogen and sulfur cycles. *Escherichia coli*, one of the most studied organisms on Earth, is a proteobacterium. So, too, are many of the most famous human pathogens, such as *Yersinia pestis* (which causes bubonic plague), *Vibrio cholerae* (cholera), and *Salmonella typhimurium* (gastrointestinal disease; **FIGURE 19.12**).

The bioluminescent *Vibrio* we discussed at the opening of this chapter are also members of this group. There are many potential applications of the genes that encode bioluminescent proteins in bacteria. Already, these genes are being inserted into the genomes of other species in which the resulting bioluminescence is used as a marker of gene expression. Futuristic proposals for making use of bioluminescence in bioengineered organisms include crop plants that glow when they become water-stressed and need to be irrigated, and glowing trees that could light highways at night in place of electric lights.

Although fungi cause most plant diseases, and viruses cause others, about 200 known plant diseases are of bacterial origin. Crown gall, with its characteristic tumors (**FIGURE 19.13**), is one of the most striking. The causal agent of crown gall is *Agrobacterium tumefaciens*, a proteobacterium that harbors a plasmid used in recombinant DNA studies as a vehicle for inserting genes into new plant hosts.

Gene sequencing enabled biologists to differentiate the domain Archaea

The separation of Archaea from Bacteria and Eukarya was originally based on phylogenetic relationships determined from sequences of rRNA genes. It was supported when biologists sequenced the first archaeal genome. That genome consisted of 1,738 genes, more than half of which were unlike any genes ever found in the other two domains.

Salmonella typhimurium

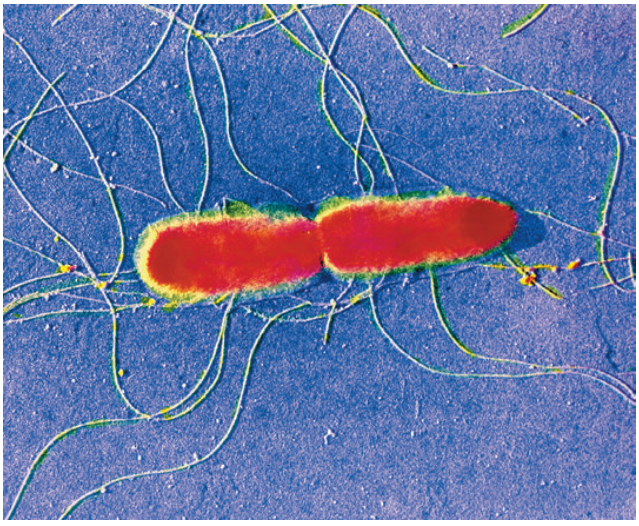


FIGURE 19.12 Proteobacteria Include Human Pathogens These conjugating cells of *Salmonella typhimurium* are exchanging genetic material (see Concept 8.4). Representing only one group of the highly diverse proteobacteria, this pathogen causes a wide range of gastrointestinal illnesses in humans.

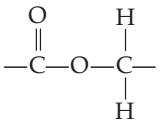
Go to MEDIA CLIP 19.2
A Swarm of *Salmonella*
PoL2e.com/mc19.2

FIGURE 19.13 Crown Gall Crown gall, a type of tumor shown here growing on the trunk of a white oak, is caused by the proteobacterium *Agrobacterium tumefaciens*.

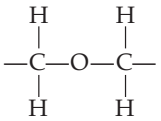
Archaea are well known for living in extreme habitats such as those with high salinity (salt content), low oxygen concentrations, high temperatures, or high or low pH (**FIGURE 19.14**). Many archaea are not extremophiles, however, but live in moderate habitats—they are common in soil, for example. Perhaps the largest numbers of archaea live in the ocean depths.

One current classification scheme divides Archaea into two principal groups, **Crenarchaeota** and **Euryarchaeota**. Less is known about three more recently discovered groups, **Korarchaeota**, **Nanoarchaeota**, and **Thaumarchaeota**. In fact, we know relatively little about the phylogeny of archaea, in part because the study of these prokaryotes is still in its early stages.

Two characteristics shared by all archaea are the absence of peptidoglycan in their cell walls and the presence of lipids of distinctive composition in their cell membranes (see Table 19.1). The unusual lipids in the membranes of archaea are found in all archaea and in no bacteria or eukaryotes. Most lipids in bacterial and eukaryotic membranes contain unbranched long-chain fatty acids connected to glycerol molecules by **ester linkages**:



In contrast, some lipids in archaeal membranes contain long-chain hydrocarbons connected to glycerol molecules by **ether linkages**:



INVESTIGATION

FIGURE 19.14 What Is the Highest Temperature Compatible with Life? Can any organism thrive at temperatures above 120°C? This is the temperature used for sterilization, known to destroy all previously described organisms. Kazem Kashefi and Derek Lovley isolated an unidentified prokaryote from water samples taken near a hydrothermal vent and found that it survived and even multiplied at 121°C. The organism was dubbed "Strain 121," and its gene sequencing results indicate that it is an archaeal species.[a]

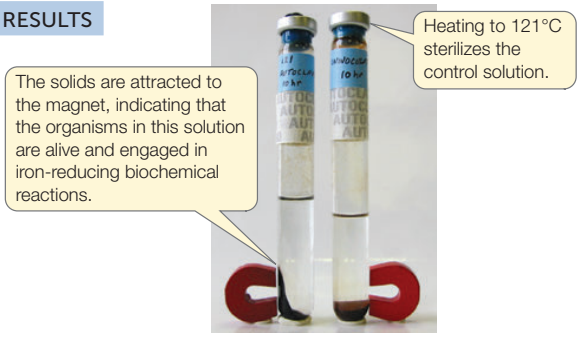
HYPOTHESIS

Some prokaryotes can survive at temperatures above the 120°C threshold of sterilization.

METHOD

1. Seal samples of unidentified, iron-reducing, thermal vent prokaryotes in tubes with a medium containing Fe^{3+} as an electron acceptor. Control tubes contain Fe^{3+} but no organisms.
2. Hold both tubes in a sterilizer at 121°C for 10 hours. If the iron-reducing organisms are metabolically active, they will reduce the Fe^{3+} to Fe^{2+} (as magnetite, which can be detected with a magnet).

RESULTS



CONCLUSION

This thermal vent organism (Strain 121) can survive at temperatures above the previously defined sterilization limit.

ANALYZE THE DATA

After Strain 121 was isolated, its growth was examined at various temperatures. The table shows generation time (the time between cell divisions) at nine temperatures.

Temperature (°C)	Generation time (hr)
85	10
90	4
95	3
100	2.5
105	2
110	4
115	6
120	20
130	No growth, but cells not killed

A. Make a graph showing generation time as a function of temperature.

B. Which temperature appears to be closest to the optimum for growth of Strain 121?

C. Note that no growth occurred at 130°C, but that the cells were not killed. How would you demonstrate that these cells were still alive?

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]K. Kashefi and D. R. Lovley. 2003. *Science* 301: 934.

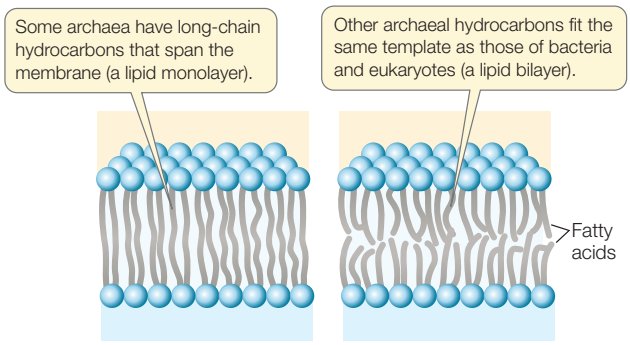


FIGURE 19.15 Membrane Architecture in Archaea The long-chain hydrocarbons of many archaeal lipids have glycerol molecules at both ends, so that the membranes they form consist of a lipid monolayer. In contrast, the membranes of other archaea, bacteria, and eukaryotes consist of a lipid bilayer.

These ether linkages are a synapomorphy (a shared, derived trait viewed as evidence of common ancestry) of archaea. In addition, the long-chain hydrocarbons of archaea are branched. One class of archaeal lipids, with hydrocarbon chains 40 carbon atoms in length, contains glycerol at *both* ends of the hydrocarbons (**FIGURE 19.15**). These lipids form a lipid monolayer structure that is unique to archaea. They still fit into a biological membrane because they are twice as long as the typical lipids in the bilayers of other membranes. Lipid monolayers and bilayers are both found among the archaea. The effects, if any, of these structural features on membrane performance are unknown. In spite of this striking difference in their lipids, the membranes of all three domains have similar overall structures, dimensions, and functions.

Most crenarchaeotes live in hot or acidic places

Most known crenarchaeotes are either thermophilic, acidophilic (acid loving), or both. Members of the genus *Sulfolobus* live in hot sulfur springs at temperatures of 70°C–75°C. They become metabolically inactive at 55°C (131°F). Hot sulfur springs are also extremely acidic. *Sulfolobus* grows best in the range of pH 2–pH 3, but some members of this genus readily tolerate pH values as low as 0.9. Most acidophilic thermophiles maintain an internal pH of 5.5–7 (close to neutral) in spite of their acidic environment. These and other thermophiles thrive where few other organisms can even survive (**FIGURE 19.16**).

Euryarchaeotes are found in surprising places

Some species of Euryarchaeota are **methanogens**: they produce methane (CH_4) by reducing carbon dioxide, and this is the key step in their energy metabolism. All of the methanogens are obligate anaerobes (see Concept 19.3). Comparison of their rRNA gene sequences has revealed a close evolutionary relationship among these methanogenic species, which were previously assigned to several different groups of bacteria.

Methanogenic euryarchaeotes release approximately 2 billion tons of methane gas into Earth's atmosphere each year, accounting for 80–90 percent of the methane that enters the atmosphere, including that produced in some mammalian digestive systems. Approximately a third of this methane comes from methanogens living in the guts of ruminants such as cattle, sheep, and deer, and another large fraction comes from methanogens living in the guts of termites and cockroaches. Methane is increasing in Earth's atmosphere by about 1 percent per year and contributes to the greenhouse effect (see Concept 45.4). Part of that increase is due to increases in cattle and rice farming and the methanogens associated with both.

FIGURE 19.16 Some Crenarchaeotes Like It Hot Thermophilic archaea can thrive in the intense heat of volcanic hot sulfur springs such as those in Wyoming's Yellowstone National Park.

Another group of euryarchaeotes, the **extreme halophiles** (salt lovers), lives exclusively in very salty environments. Because they contain pink carotenoid pigments, these archaea are sometimes easy to see (**FIGURE 19.17**). Extreme halophiles grow in the Dead Sea and in brines of all types. The reddish pink spots that can occur on pickled fish are colonies of halophilic archaea. Few other organisms can live in the saltiest of

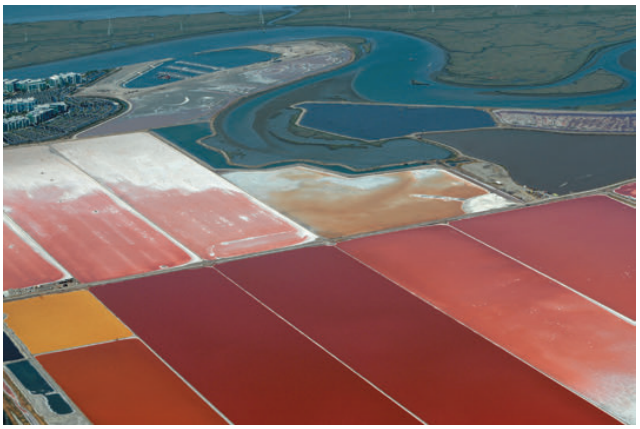


FIGURE 19.17 Extreme Halophiles Highly saline environments such as these commercial seawater evaporating ponds in San Francisco Bay are home to extreme halophiles. The archaea are easily visible here because of the rich red coloration from their carotenoid pigments.

the homes that the extreme halophiles occupy. In these environments, most organisms would "dry" to death, losing too much water to the hypertonic environment. Extreme halophiles have been found in lakes with pH values as high as 11.5—the most alkaline environment inhabited by living organisms, and almost as alkaline as household ammonia.

Some of the extreme halophiles have a unique system for trapping light energy and using it to form ATP—without using any form of chlorophyll—when oxygen is in short supply. They use the pigment retinal (also found in the vertebrate eye) combined with a protein to form a light-absorbing molecule called microbial rhodopsin.

Another member of the Euryarchaeota, *Thermoplasma*, has no cell wall. It is thermophilic and acidophilic, its metabolism is aerobic, and it lives in coal deposits. Its genome of 1,100,000 base pairs is among the smallest (along with the mycoplasmas) of any free-living organism, although some parasitic organisms have even smaller genomes.

Several lineages of Archaea are poorly known

Most known archaea are crenarchaeotes or euryarchaeotes, but studies of extreme environments have identified several small lineages that are not closely related to either of these major groups. For example, the korarchaeotes and thaumarchaeotes are known only from DNA isolated directly from hot environments. Neither group has been successfully grown in pure culture. The thaumarchaeotes oxidize ammonia and may play an important role in the nitrogen cycle.

Another distinctive archaeal lineage has been discovered at a deep-sea hydrothermal vent off the coast of Iceland. It is the first representative of the group christened Nanoarchaeota because of their minute size (Greek *nanos*, "dwarf"). This organism is a parasite that lives on cells of *Ignicoccus*, a crenarchaeote. Because of their association, the two species can be grown together in culture (FIGURE 19.18).

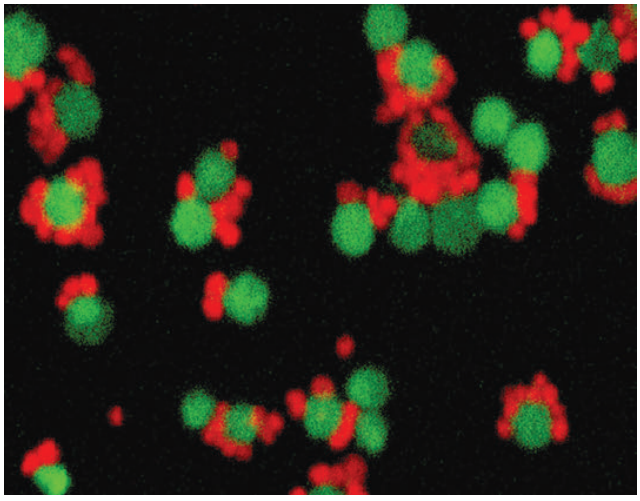


FIGURE 19.18 A Nanoarchaeote Growing in Mixed Culture with a Crenarchaeote *Nanoarchaeum equitans* (red), discovered living near deep-sea hydrothermal vents, is the only representative of the nanoarchaeote group so far identified. This tiny organism lives attached to cells of the crenarchaeote *Ignicoccus* (green). For this confocal laser micrograph, the two species were visually differentiated by fluorescent dye "tags" that are specific to their distinct gene sequences.

CHECKpoint CONCEPT 19.2

- ✓ How does the diversity of environments occupied by prokaryotes compare with the diversity of environments occupied by multicellular organisms you may be familiar with?
- ✓ Explain why Gram staining is of limited use in understanding the evolutionary relationships of bacteria.
- ✓ What makes the membranes of archaea unique?
- ✓ Given that all species of life have evolved for the same amount of time since their common origin, how would you respond to someone who characterizes prokaryotes as "primitive"?

Prokaryotes are found almost everywhere on Earth and live in a wide variety of ecosystems. Next we will examine the contributions of prokaryotes to the functioning of those ecosystems.

CONCEPT 19.3 Ecological Communities Depend on Prokaryotes

Prokaryotic cells and their associations do not usually live in isolation. Rather, they live in communities of many different species, often including microscopic eukaryotes. Whereas some microbial communities are harmful to humans, others provide important services. They help us digest our food, break down municipal waste, and recycle organic matter and chemical elements in the environment.

Many prokaryotes form complex communities

Some microbial communities form layers in sediments, and others form clumps a meter or more in diameter. Many microbial communities tend to form dense **biofilms**. Upon contacting a solid surface, the cells bind to that surface and secrete a sticky, gel-like polysaccharide matrix that traps other cells (FIGURE 19.19). Once a biofilm forms, the cells become more difficult to kill.

Biofilms are found in many places, and in some of those places they cause problems for humans. The material on our teeth that we call dental plaque is a biofilm. Pathogenic bacteria are difficult for the immune system—and modern medicine—to combat once they form a biofilm, which may be impermeable to antibiotics. Worse, some drugs stimulate the bacteria in a biofilm to lay down more matrix, making the film even more impermeable. Biofilms may form on just about any available surface, including contact lenses and artificial joint replacements. They foul metal pipes and cause corrosion, a major problem in steam-driven electricity generation plants. Fossil stromatolites—large, rocky structures made up

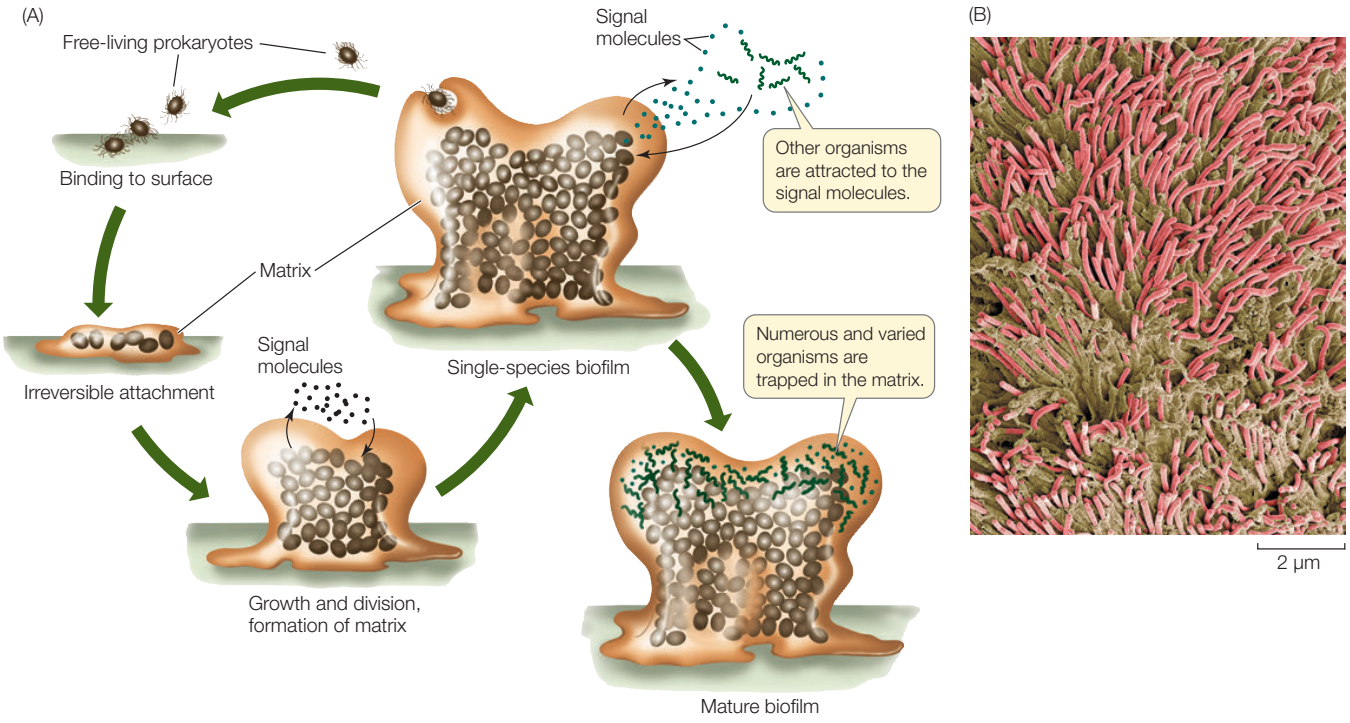


FIGURE 19.19 Forming a Biofilm (A) Free-living prokaryotes readily attach themselves to surfaces and form films that are stabilized and protected by a surrounding matrix. Once the population is large enough, the developing biofilm can send chemical signals that attract other microorganisms. (B) Scanning electron microscopy reveals a biofilm of dental plaque. The bacteria (red) are embedded in a matrix consisting of proteins from both bacterial secretions and saliva.

of alternating layers of fossilized biofilm and calcium carbonate—are among the oldest remnants of life on Earth (see Figure 18.6B).

Some biologists are studying the chemical signals that prokaryotes use to communicate with one another and that trigger density-linked activities such as biofilm formation. We saw one example of this type of communication—called **quorum sensing**—in the chapter-opening discussion of bioluminescent *Vibrio*. In the case of health-threatening bacteria, researchers hope to find ways to block the quorum-sensing signals that lead to the production of the matrix polysaccharides, thus preventing pathological biofilms from forming.

Microbiomes are critical to the health of many eukaryotes

Although only a few bacterial species are pathogens, popular notions of bacteria as "germs" and fear of the consequences of infection cause many people to assume that most bacteria are harmful. Increasingly, however, biologists are discovering that the health of humans (as well as that of other eukaryotes) depends in large part on the health of our **microbiomes**: the communities of bacteria and archaea that inhabit our bodies.

Every surface of your body is covered with diverse communities of bacteria. The Human Microbiome Project has identified more than 10,000 species living in and on humans. Inside your body, your digestive system teems with bacteria (**FIGURE 19.20**). If you count up all the cells in a human body, only about 10 percent of them are human cells. The rest are microbes—mostly bacteria, along with some archaea and microscopic eukaryotes. When these communities are disrupted, they must be restored before the body can function normally.

Biologists are discovering that many complex health problems are linked to the disruption of our microbiomes. These diverse microbial communities affect the expression of our genes and play a critical role in the development and maintenance of a healthy immune system. When our microbiomes contain an appropriate community of beneficial species, our bodies function normally. But these communities are strongly affected by our life experiences, by the food we eat, by the medicines we take, and by our exposure to various environmental toxins. The recent rapid increase in the rate of autoimmune diseases in humans—diseases in which the immune system begins to attack the body—has been linked to the changing diversity and composition of our microbiomes.

The early acquisition of an appropriate microbiome is critical for lifelong health. Normally, a human infant acquires much of its microbiome at birth, from the microbiome in its mother's vagina. Other components of the microbiome are also acquired from the mother, especially through breast feeding. Recent studies have shown that babies born by cesarean section, as well as babies that are bottle-fed on artificial milk formula, typically acquire microbes from a wider variety of sources. Many of the bacteria acquired in this way are not well suited for human health. Biologists have discovered that the incidence of many autoimmune diseases is higher in people who were born

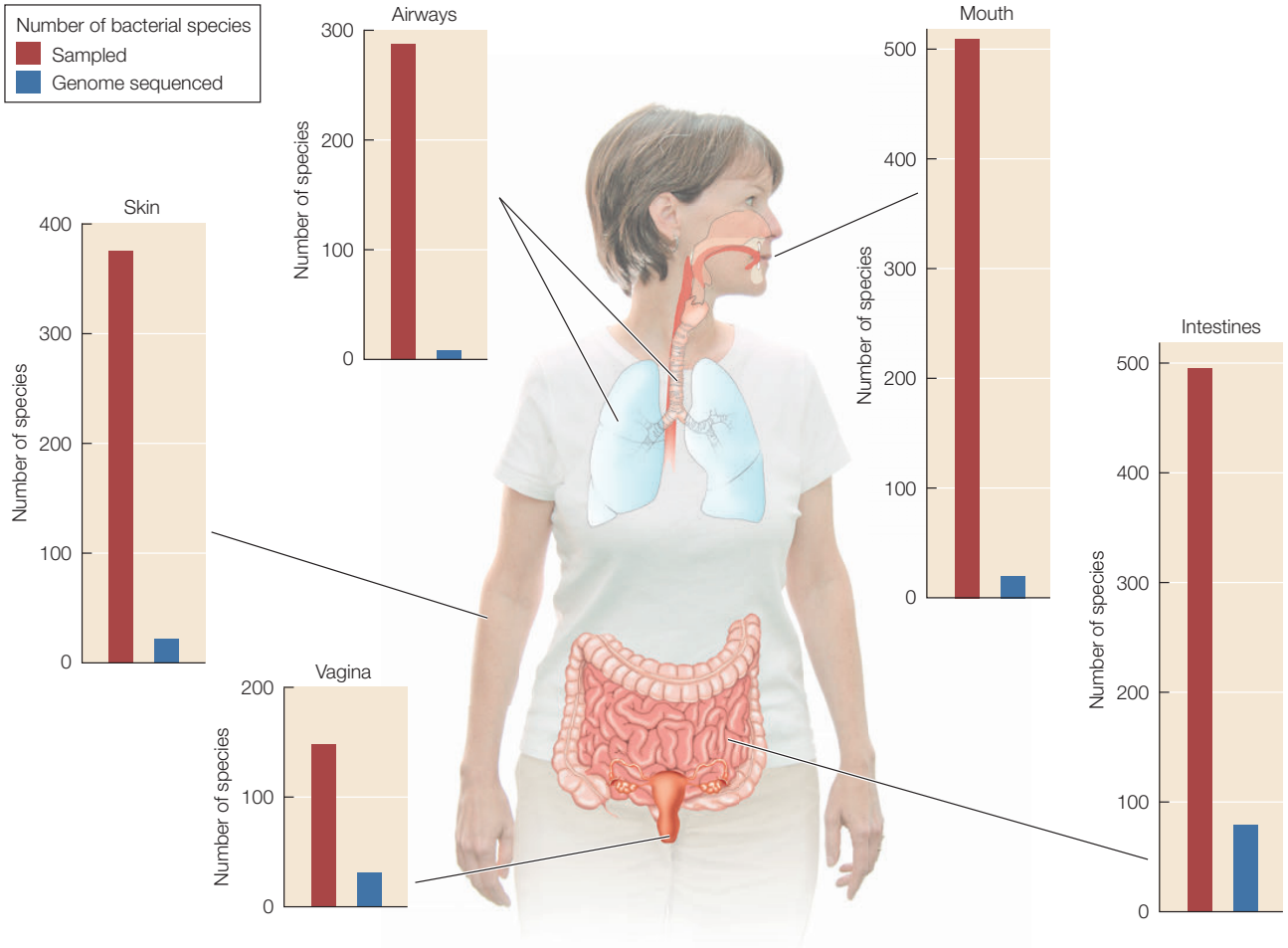


FIGURE 19.20 The Human Body's Microbiome Is Critical to the Maintenance of Health Surveys of the human microbiome have shown that this diverse community includes thousands of microbial species that are adapted to grow in or on various parts of the human body. Although it has become clear that the composition of this microbiome is closely associated with many aspects of human health, most of the component species are poorly characterized and remain largely unstudied by biologists. What has become clear is that, although the "subcommunities" in different parts of the body share similarities, each is a site-specific assemblage of many distinctive species.

by cesarean section and in those who were fed on formula as infants, compared with individuals who were born vaginally and breast-fed as infants. The difference appears to be related to the composition of the individual's original microbiome.

Our microbiomes may be related to many other health concerns as well. For example, physicians have long noted a connection between autism and gastrointestinal disorders in humans. In 2012, microbiologists discovered that children with autism have high levels of a genus of bacteria known as *Sutterella* adherent to their intestinal walls. These bacteria are absent or rare in children without autism. It is not yet known if these bacteria are causing the gastrointestinal problems, or if they are merely symptoms of the gastrointestinal disorders, but there is growing evidence that intestinal microbiomes are distinct in children with autism.

Humans require some of the metabolic products—especially vitamins B_{12} and K—produced by the microbiome living in the large intestine. Communities of bacteria line our intestines with a dense biofilm that is in intimate contact with the mucosal lining of the gut. This biofilm facilitates nutrient transfer from the intestine into the body, functioning like a specialized "tissue" that is essential to our health. This biofilm has a complex ecology that scientists have just begun to explore in detail—including the possibility that the species composition of an individual's gut microbiome may contribute to obesity (or the resistance to it).

Animals also harbor a variety of bacteria and archaea in their digestive tracts, many of which play important roles in digestion. Cattle depend on prokaryotes to break down plant material. Like most animals, cattle cannot produce cellulase, the enzyme needed to start the digestion of the cellulose that makes up the bulk of their plant food. However, bacteria living in a special section of the gut, called the rumen, produce enough cellulase to process the daily diet for the cattle (see Concept 30.2).

Marshall and Warren set out to satisfy Koch's postulates:

Test 1

The microorganism must be present in every case of the disease.

Results: Biopsies from the stomachs of many patients revealed that the bacterium was always present if the stomach was inflamed or ulcerated.

Test 2

The microorganism must be cultured from a sick host.

Results: The bacterium was isolated from biopsy material and eventually grown in culture media in the laboratory.

Test 3

The isolated and cultured bacteria must be able to induce the disease.

Results: Marshall was examined and found to be free of bacteria and inflammation in his stomach. After drinking a pure culture of the bacterium, he developed stomach inflammation (gastritis).

Test 4

The bacteria must be recoverable from newly infected individuals.

Results: Biopsy of Marshall's stomach 2 weeks after he ingested the bacteria revealed the presence of the bacterium, now christened *Helicobacter pylori*, in the inflamed tissue.

Conclusion

Antibiotic treatment eliminated the bacteria and the inflammation in Marshall. The experiment was repeated on healthy volunteers, and many patients with gastric ulcers were cured with antibiotics.Thus Marshall and Warren demonstrated that the stomach inflammation leading to ulcers is caused by *H. pylori* infections in the stomach.

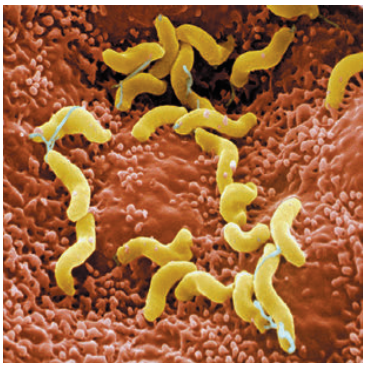


FIGURE 19.21 Satisfying Koch's Postulates Robin Warren and Barry Marshall of the University of Western Australia won the 2005 Nobel Prize in Medicine for showing that ulcers are caused not by the action of stomach acid but by infection with the bacterium *Helicobacter pylori*.

A small minority of bacteria are pathogens

The late nineteenth century was a productive era in the history of medicine—a time when bacteriologists, chemists, and physicians proved that many diseases are caused by microbial agents. During this time, the German physician Robert Koch laid down a set of four rules for establishing that a particular microorganism causes a particular disease:

1. The microorganism is always found in individuals with the disease.
2. The microorganism can be taken from the host and grown in pure culture.
3. A sample of the culture produces the same disease when injected into a new, healthy host.
4. The newly infected host yields a new, pure culture of microorganisms identical to those obtained in the second step.

These rules, called **Koch's postulates**, were important tools in a time when it was not widely understood that microorganisms cause disease. Although modern medical science has more powerful diagnostic tools, Koch's postulates remain useful. For example, physicians were taken aback in the 1990s when stomach ulcers—long accepted and treated as the result of excess stomach acid—were shown by Koch's postulates to be caused by the bacterium *Helicobacter pylori* (**FIGURE 19.21**).

For an organism to be a successful pathogen, it must:

- arrive at the body surface of a potential host;
- enter the host's body;
- evade the host's defenses;
- reproduce inside the host; and
- infect a new host.

Failure to complete any of these steps ends the reproductive career of a pathogenic organism. Yet in spite of the many defenses available to potential hosts (see Chapter 39), some bacteria are

APPLY THE CONCEPT

A small minority of bacteria are pathogens

Imagine you are in charge of maintaining a trout hatchery. Some trout are exhibiting loss of tissue at the tips of their fins, and you suspect a bacterial infection. You isolate and culture two species of bacteria from a trout with affected fins.

1. How would you satisfy the first of Koch's postulates?
2. Imagine that bacterium 1 satisfies the first of Koch's postulates, but bacterium 2 does not. What presence or absence of data from the cultures from many samples of infected fish would lead you to this conclusion?
3. You already know that bacterium 1 satisfies the second of Koch's postulates. You decide to conduct a test of the third postulate. To your surprise, the test animals are all healthy and show no sign of disease. What are some possible explanations? How would you test your hypotheses?

very successful pathogens. Pathogenic bacteria are often surprisingly difficult to combat, even with today's arsenal of antibiotics. One source of this difficulty is their ability to form biofilms.

For the host, the consequences of a bacterial infection depend on several factors. One is the **invasiveness** of the pathogen: its ability to multiply in the host's body. Another is its **toxigenicity**: its ability to produce toxins (chemical substances that are harmful to the host's tissues). *Corynebacterium diphtheriae*, the agent that causes diphtheria, has low invasiveness and multiplies only in the throat, but its toxigenicity is so great that the entire body is affected. In contrast, *Bacillus anthracis*, which causes anthrax, has low toxigenicity but is so invasive that the entire bloodstream ultimately teems with the bacteria.

There are two general types of bacterial toxins: endotoxins and exotoxins. **Endotoxins** are released when certain Gram-negative bacteria grow or lyse (burst). Endotoxins are lipopolysaccharides (complexes consisting of a polysaccharide and a lipid component) that form part of the outer bacterial membrane. Endotoxins are rarely fatal to the host; they normally cause fever, vomiting, and diarrhea. Among the endotoxin producers are some strains of the proteobacteria *Salmonella* and *Escherichia*.

Exotoxins are soluble proteins released by living, multiplying bacteria. They are highly toxic—sometimes fatal—to the host. Human diseases induced by bacterial exotoxins include tetanus (*Clostridium tetani*), cholera (*Vibrio cholerae*), and bubonic plague (*Yersinia pestis*). Anthrax is caused by three exotoxins produced by *Bacillus anthracis*. Botulism is caused by exotoxins produced by *Clostridium botulinum*, which are among the most poisonous ever discovered. The lethal dose for humans of one exotoxin of *C. botulinum* is about one-millionth of a gram. Nonetheless, much smaller doses of this exotoxin, marketed under various trade names (e.g., Botox), are used to treat muscle spasms and for cosmetic purposes (temporary wrinkle reduction in the skin).

Prokaryotes have amazingly diverse metabolic pathways

Bacteria and archaea outdo the eukaryotes in terms of metabolic diversity. Although they are much more diverse in size and shape, eukaryotes draw on fewer metabolic mechanisms for their energy needs. In fact, much of the eukaryotes' energy metabolism is carried out in organelles—mitochondria and chloroplasts—that are endosymbiotic descendants of bacteria. The long evolutionary history of bacteria and archaea, during which they have had time to explore a wide variety of habitats, has led to the extraordinary diversity of their metabolic "lifestyles"—their use or nonuse of oxygen, their energy sources, their sources of carbon atoms, and the materials they release as waste products.

ANAEROBIC VERSUS AEROBIC METABOLISM Some prokaryotes can live only by anaerobic metabolism because oxygen is poisonous to them. These oxygen-sensitive organisms are called **obligate anaerobes**. Other prokaryotes can shift their metabolism between anaerobic and aerobic modes and are thus called **facultative anaerobes**. Many facultative anaerobes alternate between anaerobic metabolism (such as fermentation) and cellular respiration as conditions dictate. **Aerotolerant anaerobes** cannot conduct aerobic cellular respiration, but they are not damaged by oxygen when it is present. By definition, an anaerobe does not use oxygen as an electron acceptor for its respiration.

At the other extreme from the obligate anaerobes, some prokaryotes are **obligate aerobes**, unable to survive for extended periods in the *absence* of oxygen. They require oxygen for cellular respiration.

TABLE 19.2 How Organisms Obtain Their Energy and Carbon

Nutritional category	Energy source	Carbon source
Photoautotrophs (some bacteria, some eukaryotes)	Light	Carbon dioxide
Photoheterotrophs (some bacteria)	Light	Organic compounds
Chemoautotrophs (some bacteria, many archaea)	Inorganic substances	Carbon dioxide
Chemoheterotrophs (found in all three domains)	Usually organic compounds; sometimes inorganic substances	Organic compounds

NUTRITIONAL CATEGORIES All living organisms face the same nutritional challenges: they must synthesize energy-rich compounds such as ATP to power their life-sustaining metabolic reactions, and they must obtain carbon atoms to build their own organic molecules. Biologists recognize four broad nutritional categories of organisms: photoautotrophs, photoheterotrophs, chemoautotrophs, and chemoheterotrophs. Prokaryotes are represented in all four groups (TABLE 19.2).

Photoautotrophs perform photosynthesis. They use light as their energy source and carbon dioxide (CO_2) as their carbon source. The cyanobacteria, like green plants and other photosynthetic eukaryotes, use chlorophyll *a* as their key photosynthetic pigment and produce oxygen gas (O_2) as a by-product of noncyclic electron transport.

There are other photoautotrophs among the bacteria, but these organisms use bacteriochlorophyll as their key photosynthetic pigment, and they do not produce O_2. Instead, some of these photosynthesizers produce particles of pure sulfur, because hydrogen sulfide (H_2S), rather than H_2O, is their electron donor for photophosphorylation. Many proteobacteria fit into this category. Bacteriochlorophyll molecules absorb light of longer wavelengths than the chlorophyll molecules used by other photosynthesizing organisms. As a result, bacteria using this pigment can grow in water under fairly dense layers of algae, using light of wavelengths that are not absorbed by the algae (FIGURE 19.22).

Photoheterotrophs use light as their energy source but must obtain their carbon atoms from organic compounds made by other organisms. Their "food" consists of organic compounds

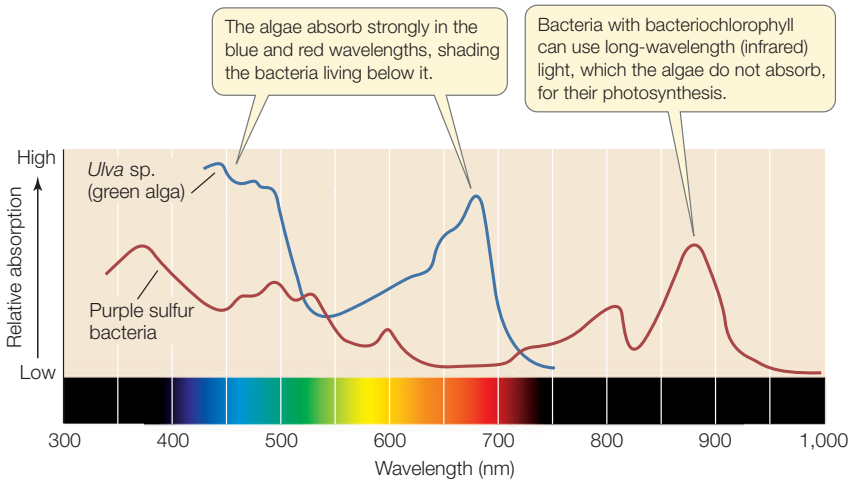


FIGURE 19.22 Bacteriochlorophyll Absorbs Long-Wavelength Light The green alga *Ulva* contains chlorophyll, which absorbs no light of wavelengths longer than 750 nm. Purple sulfur bacteria, which contain bacteriochlorophyll, can conduct photosynthesis using longer wavelengths. As a result, these bacteria can grow under layers of algae.

such as carbohydrates, fatty acids, and alcohols. For example, compounds released from plant roots (as in rice paddies) or from decomposing photosynthetic bacteria in hot springs are taken up by photoheterotrophs and metabolized to form building blocks for other compounds. Sunlight provides the energy necessary for metabolism through photophosphorylation.

Chemoautotrophs obtain their energy by oxidizing inorganic substances, and they use some of that energy to fix carbon. Some chemoautotrophs use reactions identical to those of the typical photosynthetic cycle, but others use alternative pathways for carbon fixation. Some bacteria oxidize ammonia or nitrite ions to form nitrate ions. Others oxidize hydrogen gas, hydrogen sulfide, sulfur, and other materials. Many archaea are chemoautotrophs.

Finally, **chemoheterotrophs** obtain both energy and carbon atoms from one or more complex organic compounds that have been synthesized by other organisms. Most known bacteria and archaea are chemoheterotrophs—as are all animals and fungi and many protists.

Although most chemoheterotrophs rely on the breakdown of organic compounds for energy, some chemoheterotrophic prokaryotes obtain their energy by breaking down inorganic substances. Organisms that obtain energy from oxidizing inorganic substances (both chemoautotrophs as well as some chemoheterotrophs) are also known as lithotrophs (Greek, "rock consumers").

Prokaryotes play important roles in element cycling

The metabolic diversity of the prokaryotes makes them key players in the cycles that keep elements moving through ecosystems. Many prokaryotes are decomposers: organisms that metabolize organic compounds from dead organic material and return the products to the environment as inorganic substances. Prokaryotes, along with fungi, return tremendous quantities of carbon to the atmosphere as carbon dioxide, thus carrying out a key step in the carbon cycle.

The key metabolic reactions of many prokaryotes involve nitrogen or sulfur. For example, some bacteria carry out respiratory electron transport without using oxygen as an electron acceptor. These organisms use oxidized inorganic ions such as nitrate, nitrite, or sulfate as electron acceptors. Examples include the **denitrifiers**, which release nitrogen to the atmosphere as nitrogen gas (N_2). These normally aerobic bacteria, mostly species of the genera *Bacillus* and *Pseudomonas*, use nitrate (NO_3^-) as an electron acceptor in place of oxygen if they are kept under anaerobic conditions:

$$2\,NO_3^- + 10\,e^- + 12\,H^+ \rightarrow N_2 + 6\,H_2O$$

Denitrifiers play a key role in the cycling of nitrogen through ecosystems. Without denitrifiers, which convert nitrate ions back into nitrogen gas, all forms of nitrogen would leach from the soil and end up in lakes and oceans, making life on land much more difficult.

Nitrogen fixers convert atmospheric nitrogen gas into a chemical form (ammonia) that is usable by the nitrogen fixers themselves as well as by other organisms:

$$N_2 + 6\,H \rightarrow 2\,NH_3$$

All organisms require nitrogen in order to build proteins, nucleic acids, and other important compounds. Nitrogen fixation is thus vital to life as we know it. This all-important biochemical process is carried out by a wide variety of archaea and bacteria (including cyanobacteria) but by no other organisms, so we depend on these prokaryotes for our very existence.

LINK

For descriptions of the role of nitrogen in plant nutrition and of the global nitrogen cycle, see **Concepts 25.2 and 45.3**

Ammonia is oxidized to nitrate in soil and in seawater by chemoautotrophic bacteria called **nitrifiers**. Bacteria of two genera, *Nitrosomonas* and *Nitrosococcus*, convert ammonia (NH_3) to nitrite ions (NO_2^-), and *Nitrobacter* oxidize nitrite to nitrate (NO_3^-), the form of nitrogen most easily used by many plants. What do the nitrifiers get out of these reactions? Their metabolism is powered by the energy released by the oxidation

of ammonia or nitrite. For example, by passing the electrons from nitrite through an electron transport system, *Nitrobacter* can make ATP and, using some of this ATP, can also make NADH. With this ATP and NADH, the bacterium can convert CO_2 and H_2O into glucose.

We have already seen the importance of the cyanobacteria in the cycling of oxygen: in ancient times, the oxygen generated by their photosynthesis converted Earth's atmosphere from an anaerobic to an aerobic environment. Other prokaryotes—both bacteria and archaea—contribute to the cycling of sulfur. Deep-sea hydrothermal vent ecosystems depend on chemoautotrophic prokaryotes that are incorporated into large communities of crabs, mollusks, and giant worms, all living at a depth of 2,500 meters—below any hint of sunlight. These bacteria obtain energy by oxidizing hydrogen sulfide and other substances released in the near-boiling water flowing from volcanic vents in the ocean floor.

CHECKpoint CONCEPT 19.3

- ✓ How do biofilms form, and why are they of special interest to researchers?
- ✓ What are three ways that modern lifestyles may be disrupting human microbiomes in ways that affect our health?
- ✓ How is nitrogen metabolism in the prokaryotes vital to other organisms? Given the roles of bacteria in the nitrogen cycle, how would you answer people who consider all bacteria to be "germs" and dangerous?

Before moving on to discuss the diversity of eukaryotic life, it is appropriate to consider another category of life that includes some pathogens: the viruses. Although they are not cellular, viruses are numerically among the most abundant forms of life on Earth. Their effects on other organisms are enormous. Where did viruses come from, and how do they fit into the tree of life? Biologists are still working to answer these questions.

CONCEPT 19.4 Viruses Have Evolved Many Times

Some biologists do not think of viruses as living organisms, primarily because they are not cellular and must depend on cellular organisms for basic life functions such as replication and metabolism. But viruses are derived from the cells of living organisms. They use the same essential forms of genetic information storage and transmission as do cellular organisms. Viruses infect all cellular forms of life—bacteria, archaea, and eukaryotes. They replicate, mutate, evolve, and interact with other organisms, often causing serious diseases in their hosts. Finally, viruses clearly evolve independently of other organisms, so it is almost impossible not to treat them as a part of life.

Viruses are abundant in many environments. In some freshwater and marine ecosystems, they can occur at densities of up to 10 million viruses per milliliter of water. Biologists estimate that there are approximately 10^{31} individual virus particles on Earth—about 1,000 times the number of cellular organisms on the planet. Viruses have an enormous effect on the ecology of the oceans. Every day, about one-half of the bacteria in the oceans are killed by viruses. Huge marine blooms of bacteria, such as the *Vibrio* bloom that produced the milky seas described at the opening of this chapter, do not last for long because viral blooms soon follow the initial bacterial bloom. As the viruses increase, they begin to kill bacteria faster than the bacteria can reproduce.

Although viruses are everywhere and play an important role in many ecosystems, many aspects of their ecology and evolution are still poorly known. For example, several factors make virus phylogeny difficult to resolve. The tiny size of many virus genomes restricts the phylogenetic analyses that can be conducted to relate viruses to cellular organisms. Even viruses with large genomes often contain many genes that are difficult to align with the genomes of cellular organisms. Their rapid mutation rate, which results in rapid evolution of virus genomes, tends to cloud evolutionary relationships over long periods. There are no known fossil viruses (viruses are too small and delicate to fossilize), so the paleontological record offers no clues to virus origins. Finally, viruses are highly diverse (**FIGURE 19.23**). Several lines of evidence support the hypothesis that viruses have evolved repeatedly within each of the major groups of life. The difficulty in resolving deep evolutionary relationships of viruses makes a phylogeny-based classification difficult. Instead, viruses are placed in one of several functionally similar groups on the basis of the structure of their genomes (for example, whether the genomes are composed of RNA or DNA, and are double- or single-stranded). Most of these defined groups are not thought to represent monophyletic taxa, however.

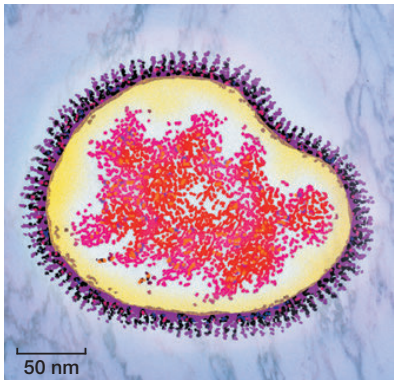
Many RNA viruses probably represent escaped genomic components

Although viruses are now obligate parasites of cellular species, many viruses may once have been cellular components involved in basic cellular functions—that is, they may be "escaped" components of cellular life that now evolve independently of their hosts.

NEGATIVE-SENSE SINGLE-STRANDED RNA VIRUSES A case in point is a class of viruses whose genome is composed of single-stranded **negative-sense RNA**: RNA that is the complement of the mRNA needed for protein translation. Many of these negative-sense single-stranded RNA viruses have only a few genes, including one for an RNA-dependent RNA polymerase that allows them to make mRNA from their negative-sense RNA genome. Modern cellular organisms cannot generate mRNA in this manner (at least in the absence of viral infections), but scientists speculate that single-stranded RNA genomes may have been common in the distant past, before DNA became the primary molecule for genetic information storage.

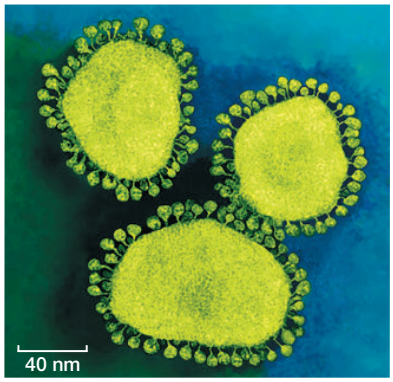
A self-replicating RNA polymerase gene that began to replicate independently of a cellular genome could conceivably acquire a few additional protein-coding genes through recombination with its host's DNA. If one or more of these genes were to foster the development of a protein coat, the virus

(A)



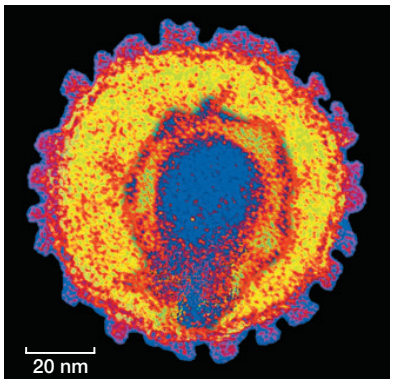
A negative-sense single-stranded RNA virus: The influenza A virus. This virus is responsible for seasonal influenza epidemics in humans. Surface view.

(B)



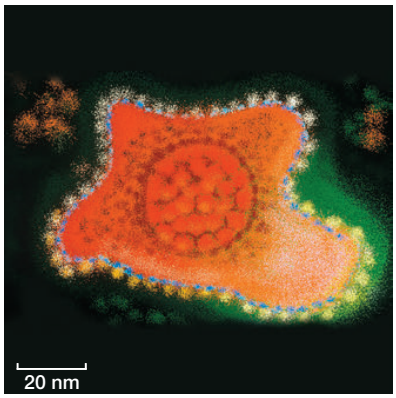
A positive-sense single-stranded RNA virus: Coronavirus of a type thought to be responsible for severe acute respiratory syndrome (SARS). Surface view.

(C)



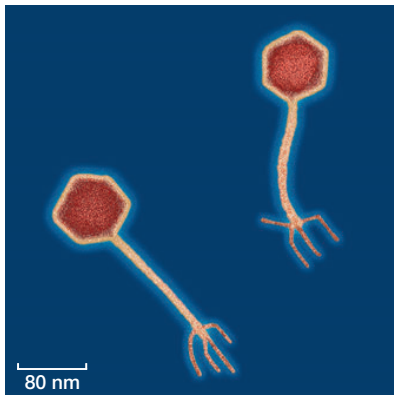
An RNA retrovirus: One of the human immunodeficiency viruses (HIV) that causes AIDS. Cutaway view.

(D)



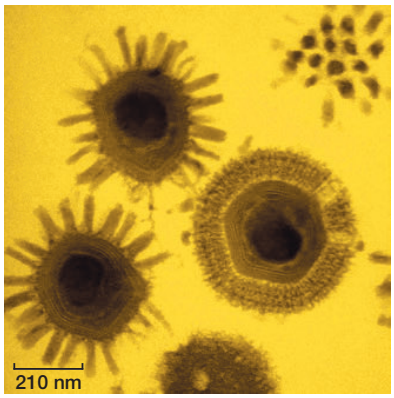
A double-stranded DNA virus: One of the many herpes viruses (Herpesviridae). Different herpes viruses are responsible for many human infections, including chicken pox, shingles, cold sores, and genital herpes (HSV1/2). Surface view.

(E)



A double-stranded DNA virus: Bacteriophage T4. Viruses that infect bacteria are referred to as bacteriophage (or simply phage). T4 attaches leglike fibers to the outside of its host cell and injects its DNA into the cytoplasm through its "tail" (pink structure in this rendition).

(F)



A double-stranded DNA mimivirus: This *Acanthamoeba polyphaga* mimivirus (APMV) has a genome larger than some prokaryote genomes. Cutaway view.

FIGURE 19.23 Viruses Are Diverse Relatively small genomes and rapid evolutionary rates make it difficult to reconstruct phylogenetic relationships among some classes of viruses. Instead, viruses are classified largely by general characteristics of their genomes. The images here are derived from cryoelectron micrographs.

might then survive outside the host and infect new hosts. It is believed that this scenario has been repeated many times independently across the tree of life, given that many of the negative-sense single-stranded RNA viruses that infect organisms from bacteria to humans are not closely related to one another. In other words, negative-sense single-stranded RNA viruses do not represent a distinct taxonomic group, but rather exemplify a particular process of cellular escape that probably happened many different times.

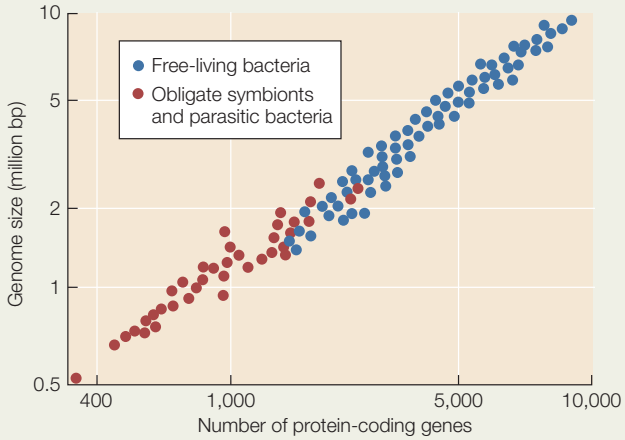
Familiar examples of negative-sense single-stranded RNA viruses include the viruses that cause measles, mumps, rabies, and influenza (see Figure 19.23A).

POSITIVE-SENSE SINGLE-STRANDED RNA VIRUSES The genome of another type of single-stranded RNA virus is composed of positive-sense RNA. Positive-sense genomes are already set for translation; no replication of the genome to form a complement strand is needed before protein translation can take place. Positive-sense single-stranded RNA viruses (see Figure 19.23B) are the most abundant and diverse class of viruses. Most of the viruses that cause diseases in crop plants are members of this group. These viruses kill patches of cells in the leaves or stems of plants, leaving live cells amid a patchwork of discolored dead tissue (giving them the name of mosaic or mottle viruses; **FIGURE 19.24**). Other viruses in this group infect bacteria, fungi, and animals. Human diseases caused by positive-sense single-stranded RNA viruses include polio, hepatitis C, and the common cold. As is true of the other functionally defined groups of viruses, these viruses appear to have evolved multiple times across the tree of life from different groups of cellular ancestors.

APPLY THE CONCEPT

Viruses have evolved many times

When Gram staining revealed the first mimivirus (discovered living inside an amoeba), it was mistakenly identified as a parasitic Gram-positive bacterium. It was soon discovered that this species was not cellular, however, and it was reclassified as a virus. The graph below shows the genome size and number of protein-coding genes for various species of bacteria. Use the graph to answer the following questions.



1. What is a likely explanation for the generally smaller genome size and smaller number of protein-coding genes in parasitic bacteria than in most free-living species? Can you form a hypothesis about which species of parasitic bacteria are likely to have existed for the longest time as parasites?
2. The genome of *Acanthamoeba polyphaga mimivirus* (the first discovered mimivirus) is 1,181,404 base pairs, encompassing 911 protein-coding genes. Plot the relevant data for this mimivirus on the graph. How does your result fit with the hypothesis that the mimivirus evolved from a parasitic bacterium?
3. Earlier in this chapter we described a minute, parasitic group of archaea, the Nanoarchaeota (see Figure 19.18). The genome of one species, *Nanoarchaeum equitans*, is 490,885 base pairs and contains 536 protein-coding genes. Plot the relevant data for this species on the graph. How does the genome of this parasitic archaea species differ from the genomes of the parasitic bacteria? From the genome of the mimivirus?

RNA RETROVIRUSES The RNA retroviruses are best known as the group that includes the human immunodeficiency viruses (HIV; see Figure 19.23C). Like the previous two categories of viruses, RNA retroviruses have genomes composed of single-stranded RNA and probably evolved as escaped cellular components.

Retroviruses are so named because they regenerate themselves by reverse transcription. When the retrovirus enters the nucleus of its vertebrate host, viral reverse transcriptase produces complementary DNA (cDNA) from the viral RNA genome and then replicates the single-stranded cDNA to produce double-stranded DNA. Another virally encoded enzyme called integrase catalyzes the integration of the new piece of double-stranded viral DNA into the host's genome. The viral genome is then replicated along with the host cell's DNA. The integrated retroviral DNA is known as a **provirus**.

FIGURE 19.24 Mosaic Viruses Are a Problem for Agriculture
Mosaic, or "mottle," viruses are the most diverse class of viruses. This leaf is from an apple tree infected with mosaic virus.

Retroviruses are only known to infect vertebrates, although genomic elements that resemble portions of these viruses are a component of the genomes of a wide variety of organisms, including bacteria, plants, and many animals. Several retroviruses are associated with the development of various forms of cancer, as cells infected with these viruses are more likely to undergo uncontrolled replication.

As retroviruses become incorporated into the genomes of their hosts, many become nonfunctional copies that are no longer expressed as functional viruses. These sequences may provide a record of ancient viral infections that plagued our ancestors. Humans, for example, carry about 100,000 fragments of **endogenous retroviruses** in our genome. These fragments make up about 8 percent of our DNA—a considerably larger fraction of our genome than the fraction that comprises all our protein-coding genes (about 1.2 percent of our genome).

DOUBLE-STRANDED RNA VIRUSES Double-stranded RNA viruses may have evolved repeatedly from single-stranded RNA ancestors—or perhaps vice versa. These viruses, which

are not closely related to one another, infect organisms from throughout the tree of life. Many plant diseases are caused by double-stranded RNA viruses. Other viruses of this type cause many cases of infant diarrhea in humans.

Some DNA viruses may have evolved from reduced cellular organisms

Another class of viruses is composed of viruses that have a double-stranded DNA genome (see Figure 19.23D–F). This group is also almost certainly polyphyletic (with many independent origins). Many of the common phage that infect bacteria are double-stranded DNA viruses, as are the viruses that cause smallpox and herpes in humans.

Some biologists think that at least some of the DNA viruses may represent highly reduced parasitic organisms that have lost their cellular structure as well as their ability to survive as free-living species. For example, the mimiviruses (see Figure 19.23F) have a genome in excess of a million base pairs of DNA that encode more than 900 proteins. This genome is similar in size to the genomes of many parasitic bacteria and about twice as large as the genome of the smallest bacteria. The recently discovered pandoraviruses have even larger genomes (up to about 2.5 million base pairs)—larger than the smallest genomes of eukaryotes. Phylogenetic analyses of DNA viruses suggest that they have evolved repeatedly from cellular organisms, possibly including major branches of life that are now extinct. Furthermore, recombination among different viruses may have allowed the exchange of various genetic modules, further complicating the history and origins of these viruses.

Viruses can be used to fight bacterial infections

Although some viruses cause devastating diseases, other viruses have been used to fight disease. Most bacterial diseases are treated today with antibiotics. But antibiotics were first discovered in the 1930s, and they were not widely used to treat bacterial diseases until the 1940s. So antibiotics were not yet available during World War I, when bacterial infections plagued the battlefields. Battlefield wounds were often infected by bacteria, and in the absence of antibiotics, these infections often led to the loss of limbs and lives. While trying to find a way to combat this problem, a physician named Felix d'Herelle discovered the first evidence of viruses that attack bacteria. He named these viruses bacteriophage, or "eaters of bacteria." Herelle extracted bacteriophage from the stool of infected patients. He then used these extracts to treat patients with deadly bacterial infections, including dysentery, cholera, and bubonic plague. This practice became known as **phage therapy**. After the war, phage therapy was widely used among the general public to treat bacterial infections of the skin and intestines.

Go to MEDIA CLIP 19.3
Bacteriophages Attack *E. coli*
PoL2e.com/mc19.3

Phage therapy was mostly replaced by the use of antibiotics in the 1930s and 1940s as physicians grew concerned about treating patients with live viruses. Phage therapy continued to be used in the Soviet Union but largely disappeared from Western medical practice. Today, however, many antibiotics are losing their effectiveness as bacterial pathogens evolve resistance to these drugs. Phage therapy is once again an active area of research, and it is likely that bacteriophage will become increasingly important as weapons against bacterial diseases. One advantage that bacteriophage may have over antibiotics is that, like bacteria, bacteriophage can evolve. As bacteria evolve resistance to a strain of bacteriophage, biologists can select for new strains of bacteriophage that retain their effectiveness against the pathogens. In this way, biologists are using their understanding of evolution to combat the problem of antibiotic-resistant bacteria.

CHECKpoint CONCEPT 19.4

- ✓ Why is it difficult to place viruses precisely within the tree of life?
- ✓ What are the two main hypotheses of virus origins?
- ✓ How can viruses be used to treat some human diseases?

It appears that the enormous diversity of viruses is, at least in part, a result of their multiple origins from many different cellular organisms. It may be best to view viruses as spinoffs from the various branches on the tree of life—sometimes evolving independently of cellular genomes, sometimes recombining with them. One way to think of viruses is as the "bark" on the tree of life: certainly an important component all across the tree, but not quite like the main branches.

Q What adaptive advantage does bioluminescence provide to *Vibrio* bacteria?

ANSWER Although marine *Vibrio* can live independently, they thrive inside the guts of fish and other marine animals. Inside a fish, *Vibrio* cells attach themselves to food particles, including phytoplankton, and are eventually expelled into the ocean as waste. How can they get back into another fish gut? The bioluminescent glow produced by a dense colony of free-living *Vibrio* growing on phytoplankton attracts fish, which consume the phytoplankton and thus ingest the bacteria—which gets the bacteria into a new host fish.

SUMMARY

CONCEPT 19.1 Life Consists of Three Domains That Share a Common Ancestor

- Two of life's three **domains**, Bacteria and Archaea, are **prokaryotic**. They are distinguished from Eukarya in several ways, including their lack of a nucleus and of membrane-enclosed organelles. **Review Table 19.1**
- Eukaryotes are related to both Archaea and Bacteria and appear to have formed through endosymbiosis between members of these two lineages. The last common ancestor of all three domains probably lived about 3 billion years ago. **Review Figure 19.1 and ANIMATED TUTORIAL 19.1**
- Early attempts to classify prokaryotes were hampered by the organisms' small size and the difficulties of growing them in pure culture.
- The cell walls of almost all bacteria contain **peptidoglycan**, whereas the cell walls of archaea and eukaryotes lack peptidoglycan.
- Bacteria can be differentiated into two groups by the **Gram stain**. **Gram-negative bacteria** have a periplasmic space between the cell membrane and a distinct outer membrane. **Gram-positive bacteria** have a thick cell wall containing about five times as much peptidoglycan as a Gram-negative wall. **Review Figure 19.2 and ACTIVITY 19.1**
- The three most common bacterial morphologies are **spirilla** (spirals), **bacilli** (rods), and **cocci** (spheres). **Review Figure 19.3**
- The cells of some bacteria aggregate, forming multicellular colonies.
- Phylogenetic classification of prokaryotes is now based principally on the nucleotide sequences of rRNA and other genes involved in fundamental cellular processes.
- Prokaryotes reproduce asexually by binary fission, but many can exchange genetic material. Reproduction and genetic exchange are not directly linked in prokaryotes.
- Although **lateral gene transfer** has occurred throughout prokaryotic evolutionary history, elucidation of prokaryote phylogeny is still possible. **Review Figure 19.4**

CONCEPT 19.2 Prokaryote Diversity Reflects the Ancient Origins of Life

- Prokaryotes are the most numerous organisms on Earth, but only a small fraction of prokaryote diversity has been characterized to date.
- Some **high-GC Gram-positives** produce important antibiotics.
- The **low-GC Gram-positives** include the **mycoplasmas**, which are among the smallest cellular organisms ever discovered.
- The photosynthetic **cyanobacteria** release oxygen into the atmosphere. Cyanobacteria may live free as single cells or associate in multicellular colonies. **Review Figure 19.9**
- **Spirochetes** have unique structures called axial filaments that allow them to move in a corkscrew-like manner. **Review Figure 19.10**
- The **proteobacteria** embrace the largest number of known species of bacteria. Smaller groups include the **hyperthermophilic bacteria**, **hadobacteria**, and **chlamydias**.
- The best-studied groups of Archaea are **Crenarchaeota** and **Euryarchaeota**.
- Many archaea are **extremophiles**. Most known crenarchaeotes are thermophilic, acidophilic, or both. Some euryarchaeotes are **methanogens**; **extreme halophiles** are also found among the euryarchaeotes. **Review Figure 19.14**
- **Ether linkages** in the branched long-chain hydrocarbons of the lipids that make up their cell membranes are a synapomorphy of archaea. **Review Figure 19.15**

CONCEPT 19.3 Ecological Communities Depend on Prokaryotes

- Prokaryotes form complex communities, of which **biofilms** are one example. **Review Figure 19.19**
- Communities of bacteria can communicate information about their density using chemical signals in a process known as **quorum sensing**.
- **Microbiomes** are the communities of prokaryotes that live in and on the bodies of multicellular organisms. These communities are often important to the health of the hosts, and changes to the microbiome may lead to serious health consequences. **Review Figure 19.20**
- Prokaryotes inhabit the guts of many animals (including humans) and help them digest food.
- **Koch's postulates** establish the criteria by which an organism may be classified as a pathogen. Relatively few bacteria—and no archaea—are known to be pathogens. **Review Figure 19.21**
- Prokaryote metabolism is very diverse. Some prokaryotes are anaerobic, others are aerobic, and still others can shift between these modes.
- Prokaryotes fall into four broad nutritional categories: **photoautotrophs**, **photoheterotrophs**, **chemoautotrophs**, and **chemoheterotrophs**. **Review Table 19.2**
- Prokaryotes play key roles in the cycling of elements such as nitrogen, oxygen, sulfur, and carbon. One such role is as decomposers of dead organisms.
- Some prokaryotes metabolize sulfur or nitrogen. **Nitrogen fixers** convert nitrogen gas into a form that organisms can metabolize. **Nitrifiers** convert that nitrogen into forms that can be used by plants, and **denitrifiers** return nitrogen gas to the atmosphere.

CONCEPT 19.4 Viruses Have Evolved Many Times

- Viruses have evolved many times from many different groups of cellular organisms. They do not represent a single taxonomic group.
- Some viruses are probably derived from escaped genetic elements of cellular species; others are thought to have evolved as highly reduced parasites.
- Viruses are categorized by the nature of their genomes.
- Viruses can be used to fight bacterial infections in a process known as **phage therapy**.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is19

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources..

20

The Origin and Diversification of Eukaryotes

KEY CONCEPTS

Blooms of dinoflagellates can cause toxic red tides, such as this one in Puget Sound in the U.S. state of Washington.

In summer 2005, a devastating red tide crippled the shellfish industry along the Atlantic coast of North America from Canada to Massachusetts. This red tide was produced by a bloom of dinoflagellates of the genus *Alexandrium*. These protists produce a powerful toxin that accumulates in clams, mussels, and oysters. A person who eats a mollusk contaminated with the toxin can experience a syndrome known as paralytic shellfish poisoning. Many people were sickened by eating mollusks that were harvested before the problem was diagnosed, and losses to the shellfish industry in 2005 were estimated at $50 million.

Several species of dinoflagellates produce toxic red tides in many parts of the world. Along the Gulf of Mexico, red tides caused by dinoflagellates of the genus *Karenia* produce a neurotoxin that affects the central nervous systems of fish, which become paralyzed and cannot respire effectively. Huge numbers of dead fish wash up on Gulf Coast beaches during a *Karenia* red tide. In addition, wave action can produce aerosols of the *Karenia* toxin, and these aerosols often cause asthma-like symptoms in humans on shore.

After the losses that resulted from the 2005 red tide, biologists at the Woods Hole Oceanographic Institution (WHOI) on Cape Cod began to monitor and model dinoflagellate populations off the New England coast. If biologists could accurately forecast future blooms, people in the area could be made aware of the problem in advance and adjust the shellfish harvest (and their eating habits) accordingly.

Biologists from WHOI monitored counts of dinoflagellates in the water and in seafloor sediments. They also monitored river runoff, water currents, water temperature and salinity, winds, and tides. An additional environmental factor was the "nor'easter" storms common along the New England coast. By correlating their measurements of these environmental factors with dinoflagellate counts, biologists produced a model that predicted growth of dinoflagellate populations.

In spring 2008, the WHOI team determined that all the factors were in place to produce another red tide like the one of 2005—if a nor'easter occurred to blow the dinoflagellates toward the coast. A nor'easter did occur at just the wrong time, and another red tide materialized in summer 2008, just as predicted. But this time, people were warned. Shellfish harvesters adjusted their harvest, and many fewer people were harmed by eating toxic mollusks.

Q Red tides are harmful, but can dinoflagellates also be beneficial to marine ecosystems?

You will find the answer to this question on page 418.

CONCEPT 20.1 Eukaryotes Acquired Features from Both Archaea and Bacteria

We easily recognize trees, mushrooms, and insects as plants, fungi, and animals, respectively. But there is a dazzling assortment of other eukaryotic organisms—mostly microscopic—that do not fit into these three groups. Eukaryotes that are not plants, animals, or fungi have traditionally been called **protists**. But phylogenetic analyses reveal that many of the groups we commonly refer to as protists are not, in fact, closely related. Thus the term "protist" does not describe a formal taxonomic group, but is a convenience term for "all the eukaryotes that are not plants, animals, or fungi."

The unique characteristics of the eukaryotic cell lead scientists to conclude that the eukaryotes are monophyletic, and that a single eukaryotic ancestor diversified into the many different protist lineages as well as giving rise to the plants, fungi, and animals. As we saw in Concept 19.1, eukaryotes are generally thought to be more closely related to archaea than to bacteria. The mitochondria and chloroplasts of eukaryotes, however, are clearly derived from bacterial lineages (see Figure 19.1).

Traditionally, biologists have hypothesized that the *split* of Eukarya from Archaea was followed by the endosymbioses with bacterial lineages that led to the origin of mitochondria and chloroplasts. Some biologists prefer to view the origin of eukaryotes as the *fusion* of lineages from the two prokaryote groups. This difference is largely a semantic one that hinges on the subjective point at which we deem the eukaryote lineage to have become definitively "eukaryotic." In either case, we can make some reasonable inferences about the events that led to the evolution of a new cell type, bearing in mind that the environment underwent an enormous change—from low to high availability of free atmospheric oxygen—during the course of these events.

LINK

You can compare eukaryotic and prokaryotic cells by reviewing **Concepts 4.2 and 4.3**

The modern eukaryotic cell arose in several steps

Several events were important in the origin of the modern eukaryotic cell (**FIGURE 20.1**):

- The origin of a flexible cell surface

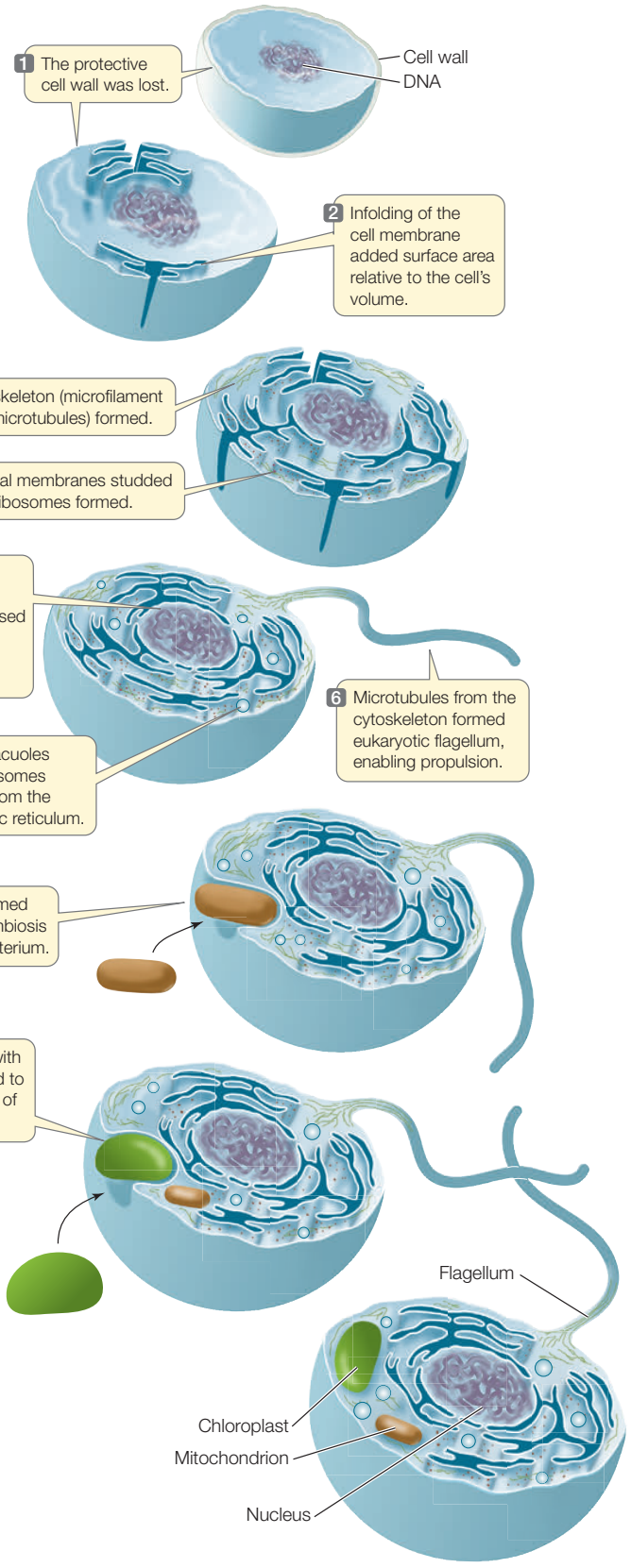


FIGURE 20.1 Evolution of the Eukaryotic Cell The loss of a rigid cell wall allowed the cell membrane to fold inward and create more surface area, which facilitated the evolution of larger cells. As cells grew larger, cytoskeletal complexity increased, and the cell became increasing compartmentalized. Endosymbioses involving bacteria gave rise to mitochondria and (in photosynthetic eukaryotes) to chloroplasts.

- The origin of a cytoskeleton
- The origin of a nuclear envelope, which enclosed a genome organized into chromosomes
- The appearance of digestive vacuoles
- The acquisition of certain organelles via endosymbiosis

FLEXIBLE CELL SURFACE We presume that ancient prokaryotic organisms, like most present-day prokaryotic cells, had firm cell walls. The first step toward the eukaryotic condition was the loss of the cell wall by an ancestral prokaryotic cell. This wall-less condition occurs in some present-day prokaryotes.

Consider the possibilities open to a flexible cell without a firm cell wall, starting with cell size. As a cell grows larger, its surface area-to-volume ratio decreases (see Figure 4.2). Unless the surface area can be increased, the cell volume will reach an upper limit. If the cell's surface is flexible, however, it can fold inward and become more elaborate, creating more surface area for gas and nutrient exchange. With a surface flexible enough to allow infolding, the cell can exchange materials with its environment rapidly enough to sustain a larger volume and more rapid metabolism (see Figure 20.1, steps 1–2). Furthermore, a flexible surface can pinch off bits of the environment, bringing them into the cell by endocytosis. These infoldings of the cell surface, which also exist in some modern prokaryotes, were important for the evolution of large eukaryotic cells.

CHANGES IN CELL STRUCTURE AND FUNCTION Other early steps that were important for the evolution of the eukaryotic cell involved increased compartmentalization and complexity of the cell (see Figure 20.1, steps 3–7):

- The development of a more complex cytoskeleton
- The formation of ribosome-studded internal membranes, some of which surrounded the DNA
- The enclosure of the cell's DNA in a nucleus
- The formation of a flagellum from microtubules of the cytoskelton
- The evolution of digestive vacuoles

Until a few years ago, biologists thought that cytoskeletons were restricted to eukaryotes. Improved imaging technology and molecular analyses have now revealed homologs of many cytoskeletal proteins in prokaryotes, so simple cytoskeletons evolved before the origin of eukaryotes. The cytoskeleton of a eukaryote, however, is much more developed and complex than that of a prokaryote. This greater development of microfilaments and microtubules supports the eukaryotic cell and allows it to manage changes in shape, to distribute daughter chromosomes, and to move materials from one part of its larger cell to other parts. In addition, the presence of microtubules in the cytoskeleton allowed some cells to develop the characteristic eukaryotic flagellum.

The DNA of a prokaryotic cell is attached to a site on its cell membrane. If that region of the cell membrane were to fold into the cell, the first step would be taken toward the evolution of a nucleus, a primary feature of the eukaryotic cell. The nuclear envelope appeared early in the eukaryote lineage. The

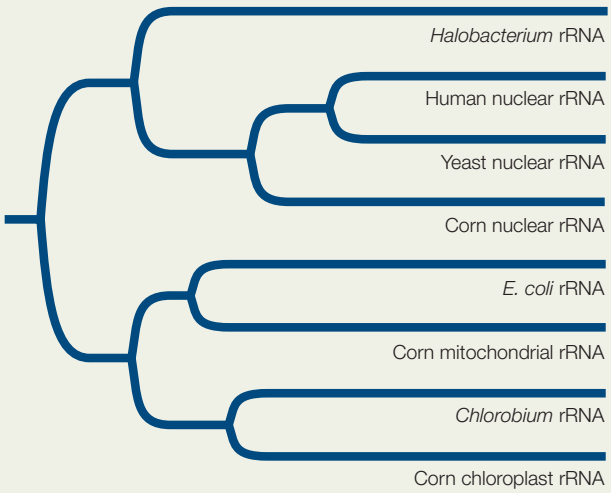
APPLY THE CONCEPT

Eukaryotes acquired features from both archaea and bacteria

Ribosomal RNA (rRNA) genes are present in the nuclear genome of eukaryotes. There are also rRNA genes in the genomes of mitochondria and chloroplasts. Therefore, photosynthetic eukaryotes have three different sets of rRNA genes, which encode the structural RNA of separate ribosomes in the nucleus, mitochondria, and chloroplasts, respectively. Translation of each genome takes place on its own set of ribosomes.

The gene tree shows the evolutionary relationships of rRNA gene sequences isolated from the nuclear genomes of humans, yeast, and corn; from an archaeon (*Halobacterium*), a proteobacterium (*E. coli*), and a cyanobacterium (*Chlorobium*); and from the mitochondrial and chloroplast genomes of corn. Use the gene tree to answer the following questions.

1. Why aren't the three rRNA genes of corn one another's closest relatives?
2. How would you explain the closer relationship of the mitochondrial rRNA gene of corn to the rRNA gene of *E. coli* than to the nuclear rRNA genes of other eukaryotes? Can you explain the relationship of the rRNA gene from the chloroplast of corn to the rRNA gene of the cyanobacterium?



3. If you were to sequence the rRNA genes from human and yeast mitochondrial genomes, where would you expect these two sequences to fit on the gene tree?

next step was probably phagocytosis—the ability to engulf and digest other cells.

ENDOSYMBIOSIS At the same time the processes outlined above were taking place, cyanobacteria were generating O_2 as a product of photosynthesis. The increasing concentrations of O_2 in the oceans, and eventually in the atmosphere, had disastrous consequences for most organisms of the time, which were unable to tolerate the newly oxidizing environment. But some prokaryotes evolved strategies to use the increasing oxygen, and—fortunately for us—so did some of the new phagocytic eukaryotes.

At about this time, endosymbioses began to play a role in eukaryote evolution (see Figure 20.1, steps 8–9). The theory of endosymbiosis proposes that certain organelles are the descendants of prokaryotes engulfed, but not digested, by ancient eukaryotic cells. One crucial event in the history of eukaryotes was the incorporation of a proteobacterium that evolved into the mitochondrion. Initially the new organelle's primary function was probably to detoxify O_2 by reducing it to water. Later this reduction became coupled with the formation of ATP in cellular respiration. Upon completion of this step, the essential modern eukaryotic cell was complete.

LINK

You may wish to review the reactions of cellular respiration in **Concept 6.2**

Photosynthetic eukaryotes are the result of yet another endosymbiotic step: the incorporation of a prokaryote related to today's cyanobacteria. This endosymbiont evolved into the modern chloroplast.

Chloroplasts have been transferred among eukaryotes several times

Eukaryotes in several different groups possess chloroplasts, and groups with chloroplasts appear in several distantly related eukaryote clades. Some of these groups differ in the photosynthetic pigments their chloroplasts contain. And not all chloroplasts are limited to a pair of surrounding membranes—in some photosynthetic eukaryotes, they are surrounded by three or more membranes. We now view these observations as evidence of a remarkable series of endosymbioses. This conclusion is supported by extensive evidence from electron microscopy and nucleic acid sequence comparisons.

All chloroplasts trace their ancestry back to the engulfment of one cyanobacterium by a larger eukaryotic cell. This event, the step that first gave rise to the photosynthetic eukaryotes, is known as **primary endosymbiosis** (**FIGURE 20.2A**). The cyanobacterium, a Gram-negative bacterium, had both an inner and an outer membrane (see Figure 19.2B). Thus the original chloroplasts had two surrounding membranes: the inner and outer membranes of the cyanobacterium. Remnants of the peptidoglycan-containing cell wall of the bacterium are present in the form of a bit of peptidoglycan between the chloroplast membranes of glaucophytes, the first eukaryote group to branch off following primary endosymbiosis (as we will see in Chapter 21). Primary

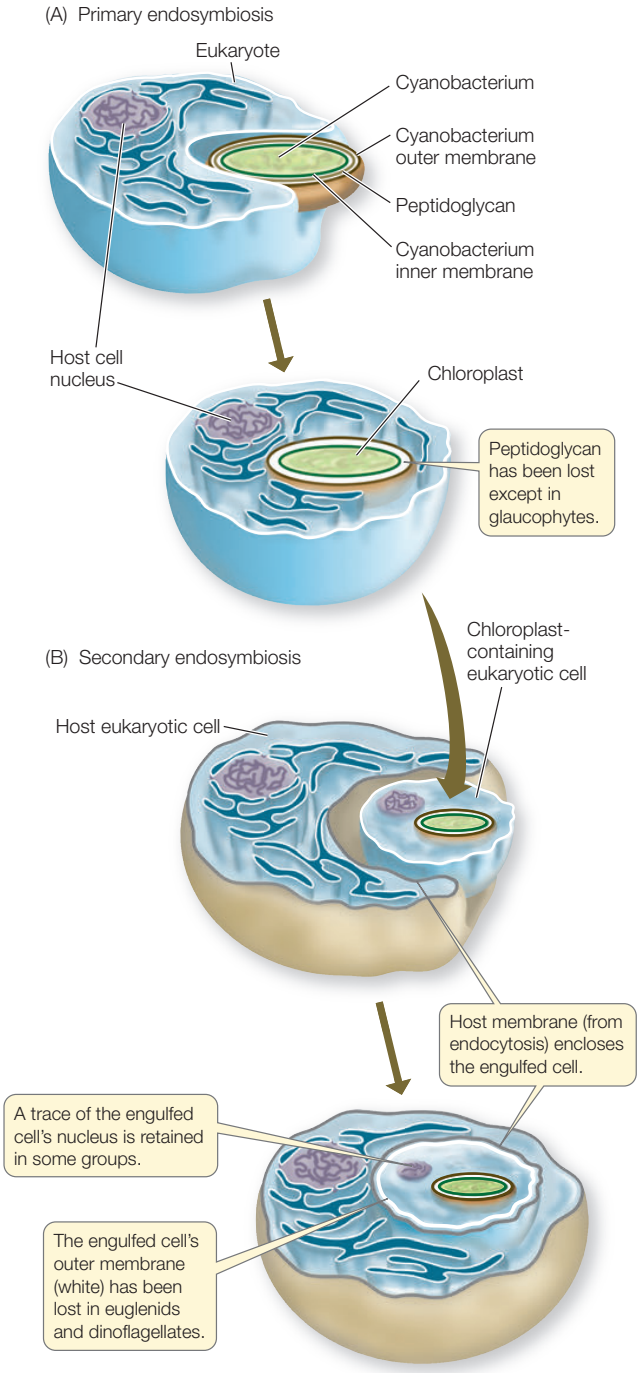


FIGURE 20.2 Endosymbiotic Events in the Evolution of Chloroplasts (A) A single instance of primary endosymbiosis ultimately gave rise to all of today's chloroplasts. (B) Secondary endosymbiosis—the uptake and retention of a chloroplast-containing cell by another eukaryotic cell—took place several times, independently.

Go to ANIMATED TUTORIAL 20.1
Family Tree of Chloroplasts
PoL2e.com/at20.1

endosymbiosis also gave rise to the chloroplasts of the red algae, green algae, and land plants. The red algal chloroplast retains certain pigments of the original cyanobacterial endosymbiont that are absent in green algal chloroplasts.

Almost all remaining photosynthetic eukaryotes are the result of additional rounds of endosymbiosis. For example, the photosynthetic euglenids derived their chloroplasts from **secondary endosymbiosis** (FIGURE 20.2B). Their ancestor took up a unicellular green alga, retaining its chloroplast and eventually losing the rest of the constituents of the alga. This history explains why the photosynthetic euglenids have the same photosynthetic pigments as the green algae and land plants. It also accounts for the third membrane of the euglenid chloroplast, which is derived from the euglenid's cell membrane (as a result of endocytosis). An additional round—**tertiary endosymbiosis**—occurred when a dinoflagellate apparently lost its chloroplast and took up another protist that had acquired its chloroplast through secondary endosymbiosis.

CHECKpoint CONCEPT 20.1

- ✓ Why was the development of a flexible cell surface a key event for eukaryote evolution?
- ✓ What do you consider the most critical events that led to the evolution of the eukaryotic cell? Why?
- ✓ Explain how increased availability of atmospheric oxygen (O_2) could have affected the evolution of the eukaryotic cell.

The features that eukaryotes gained from archaea and bacteria have allowed them to exploit many different environments. This led to the evolution of great diversity among eukaryotes, beginning with a radiation that started in the Precambrian.

CONCEPT 20.2 Major Lineages of Eukaryotes Diversified in the Precambrian

Most eukaryotes can be classified in one of eight major clades that began to diversify about 1.5 billion years ago: alveolates, excavates, stramenopiles, plants, rhizarians, amoebozoans, fungi, and animals (FIGURE 20.3). Plants, fungi, and animals each have close protist relatives (such as the choanoflagellate relatives of animals), which we will discuss along with those major multicellular eukaryote groups in Chapters 21–23.

Each of the five major groups of protist eukaryotes covered in this chapter consists of organisms with enormously diverse body forms and nutritional lifestyles. Some protists are motile, whereas others do not move. Some protists are photosynthetic, whereas others are heterotrophic. Most protists are unicellular, but some are multicellular. Most protists are microscopic, but a few are huge (giant kelps, for example, can grow to half the length of a football field). We refer to the unicellular species of protists as **microbial eukaryotes**, but keep in mind that there are large, multicellular protists as well.

Multicellularity has arisen dozens of times across the evolutionary history of eukaryotes. Four of the origins of multicellularity resulted in large organisms that are familiar to most

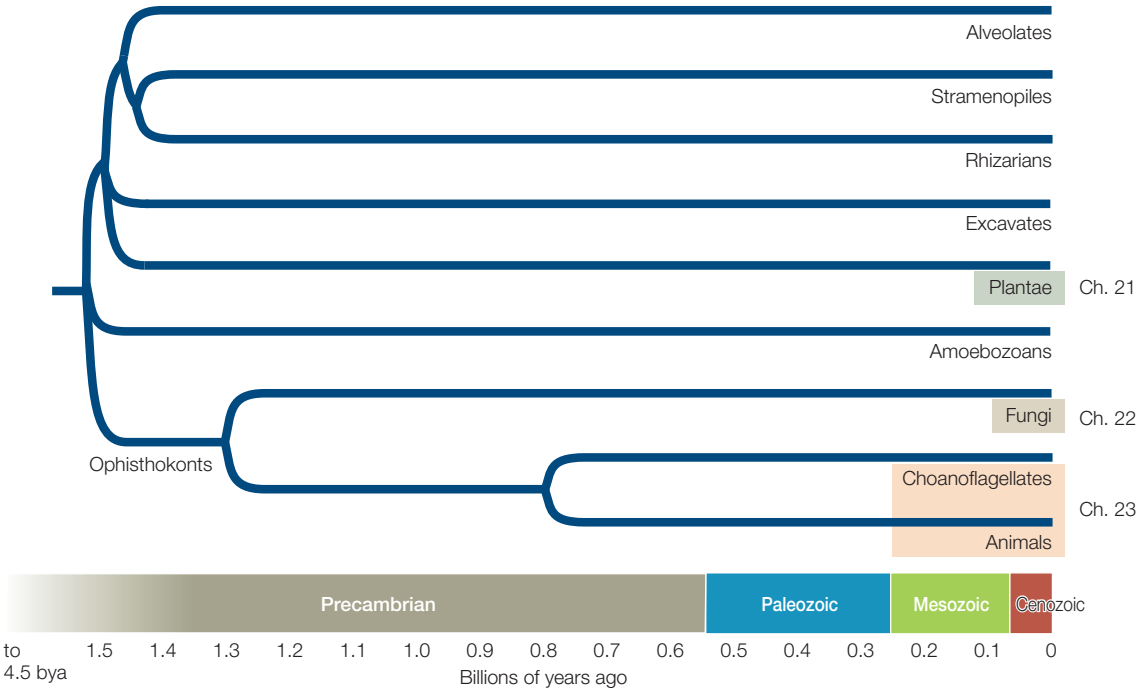


FIGURE 20.3 Precambrian Divergence of Major Eukaryote Groups
A phylogenetic tree shows one current hypothesis and estimated time line for the origin of the major groups of eukaryotes. The rapid divergence of major lineages between 1.5 and 1.4 billion years ago makes reconstruction of their precise relationships difficult. The major multicellular groups (tinted boxes) will be covered in subsequent chapters.

people: plants, animals, fungi, and brown algae (the last are a group of stramenopiles). In addition, there are dozens of smaller and less familiar groups among the eukaryotes that include multicellular species. Recent experimental studies have shown that artificial selection for multicellularity can produce repeated, convergent evolution of multicellular forms over just a few months in some normally unicellular eukaryotic species. In addition, many unicellular species retain individual identities but nonetheless associate in large multicellular colonies. There is a near-continuum between fully integrated, multicellular organisms on the one hand and loosely integrated multicellular colonies of cells on the other. Biologists do not always agree on where to draw the line between the two.

Biologists used to classify protists largely on the basis of their life histories and reproductive features (see Concept 20.3). In recent years, however, electron microscopy and gene sequencing have revealed many new patterns of evolutionary relatedness among these groups. Analyses of slowly evolving gene sequences are making it possible to explore evolutionary relationships among eukaryotes in ever greater detail and with greater confidence. Nonetheless, some substantial areas of uncertainty remain, and lateral gene transfer may complicate efforts to reconstruct the evolutionary history of protists (as was also true for prokaryotes; see Concept 19.1). Today we recognize great diversity among the many distantly related protist clades.

Alveolates have sacs under their cell membranes

Alveolates are so named because they possess sacs, called alveoli, just beneath their cell membranes, which may play a role in supporting the cell surface. All alveolates are unicellular, and most are photosynthetic, but they are diverse in body form. The groups considered in detail here are the dinoflagellates, apicomplexans, and ciliates.

DINOFLAGELLATES Most **dinoflagellates** are marine and photosynthetic; they are important primary producers of organic matter in the oceans. Although fewer species of dinoflagellates live in fresh water, individuals can be abundant in freshwater environments. The dinoflagellates are of great ecological, evolutionary, and morphological interest. A distinctive mixture of photosynthetic and accessory pigments gives their chloroplasts a golden brown color. Some dinoflagellate species cause red tides, as discussed at the start of this chapter. Other species are photosynthetic endosymbionts that live within the cells of other organisms, including invertebrate animals (such as corals; see Figure 20.19) and other marine protists (see Figure 20.11A). Some are nonphotosynthetic and live as parasites within various marine organisms.

Dinoflagellates have a distinctive appearance. They generally have two flagella, one in an equatorial groove around the cell, the other starting near the same point as the first and passing down a longitudinal groove before extending into the surrounding medium (**FIGURE 20.4**). Some dinoflagellates can take on different forms, including amoeboid ones, depending on environmental conditions. It has been claimed that the dinoflagellate *Pfiesteria piscicida* can occur in at least two dozen distinct forms, although this claim is highly controversial. In any case, this remarkable dinoflagellate is harmful to fish and can, when present in great numbers, both stun and feed on them.

Amphidiniopsis kofoidii

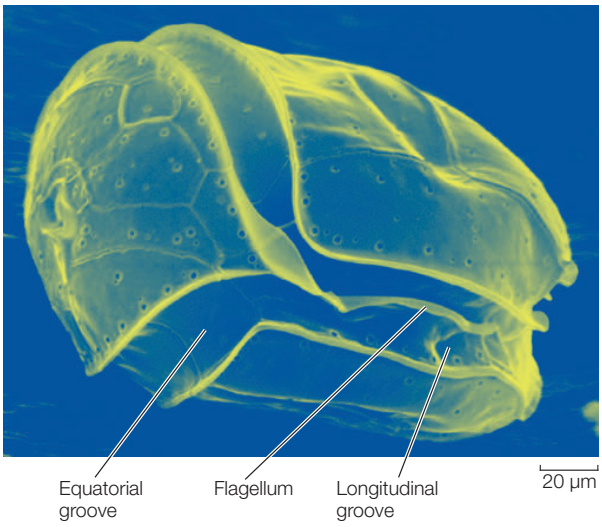


FIGURE 20.4 A Dinoflagellate The presence of two flagella is characteristic of many dinoflagellates, although only one of the two flagella is visible in this photograph. One flagellum originates within the equatorial groove and provides forward thrust and spin to the organism. The second flagellum originates in the longitudinal groove and acts like the rudder of a boat.

Go to MEDIA CLIP 20.1
A Dinoflagellate Shows Off Its Flagellum
PoL2e.com/mc20.1

APICOMPLEXANS The exclusively parasitic **apicomplexans** derive their name from the apical complex, a mass of organelles contained in the apical end (the tip) of the cell. These organelles help the apicomplexan invade its host's tissues. For example, the apical complex enables *Plasmodium*, the causative agent of malaria, to enter its target cells in the human body after transmission by a mosquito.

Like many obligate parasites, apicomplexans have elaborate life cycles featuring asexual and sexual reproduction by a series of very dissimilar life stages. In many species, these life stages are associated with two different types of host organisms, as is the case with *Plasmodium*. Another apicomplexan, *Toxoplasma*, alternates between cats and rats to complete its life cycle. A rat infected with *Toxoplasma* loses its fear of cats, which makes it more likely to be eaten by, and thus transfer the parasite to, a cat.

CILIATES The **ciliates** are named for their numerous hairlike cilia, which are shorter than, but otherwise identical to, eukaryotic flagella. The ciliates are much more complex in body form than are most other unicellular eukaryotes (**FIGURE 20.5**). Their definitive characteristic is the possession of two types of nuclei (whose roles we will describe in Concept 20.3 when we discuss protist reproduction). Almost all ciliates

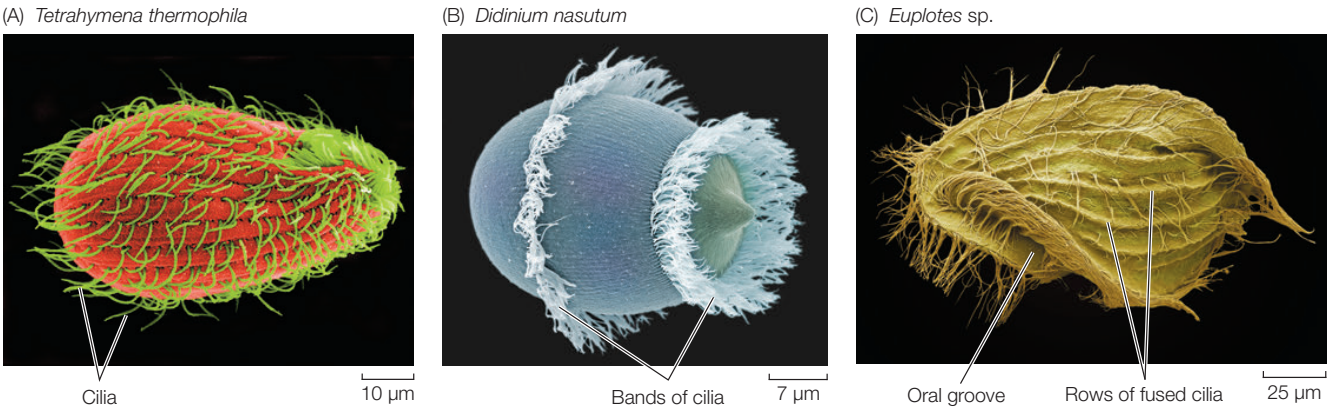


FIGURE 20.5 Diversity among the Ciliates (A) *Tetrahymena thermophila* is used as a model organism for research on gene expression as well as the structure and function of microtubule arrays. (B) The barrel-shaped *Didinium nasutum* feeds on other ciliates, including *Paramecium*. Their cilia occur in two separate bands. (C) Some of the cilia in *Euplotes* fuse into flat sheets that direct food particles into an oral groove.

are heterotrophic, although a few contain photosynthetic endosymbionts.

Paramecium, a frequently studied ciliate genus, exemplifies the complex structure and behavior of ciliates (**FIGURE 20.6**). The slipper-shaped cell is covered by an elaborate pellicle, a structure composed principally of an outer membrane and an inner layer of closely packed, membrane-enclosed sacs (the alveoli) that surround the bases of the cilia. Defensive organelles called trichocysts are also present in the pellicle. In response to a threat, a microscopic explosion expels the trichocysts in a few milliseconds, and they emerge as sharp darts, driven forward at the tip of a long, expanding filament.

The cilia provide *Paramecium* with a form of locomotion that is generally more precise than locomotion by flagella or pseudopods. A *Paramecium* can coordinate the beating of its cilia to propel itself either forward or backward in a spiraling manner. It can also back off swiftly when it encounters a barrier or a negative stimulus. The coordination of ciliary beating is probably the result of a differential distribution of ion channels in the cell membrane near the two ends of the cell.

Organisms living in fresh water are hypertonic to their environment. Many freshwater protists, including *Paramecium*, address this problem by means of specialized **contractile vacuoles** that excrete the excess water the organisms constantly take in by osmosis. The excess water collects in the contractile vacuoles, which then contract and expel the water from the cell.

Paramecium and many other protists engulf solid food by endocytosis, forming a **digestive vacuole** within which the food is digested. Smaller vesicles containing digested food pinch away from the digestive vacuole and enter the cytoplasm. These tiny vesicles provide a large surface area across which the products of digestion can be absorbed by the rest of the cell.

Go to ANIMATED TUTORIAL 20.2
Digestive Vacuoles
PoL2e.com/at20.2

LINK

The role of a pathogen's cell surface recognition molecules in the mammalian immune response is covered in **Chapter 39**

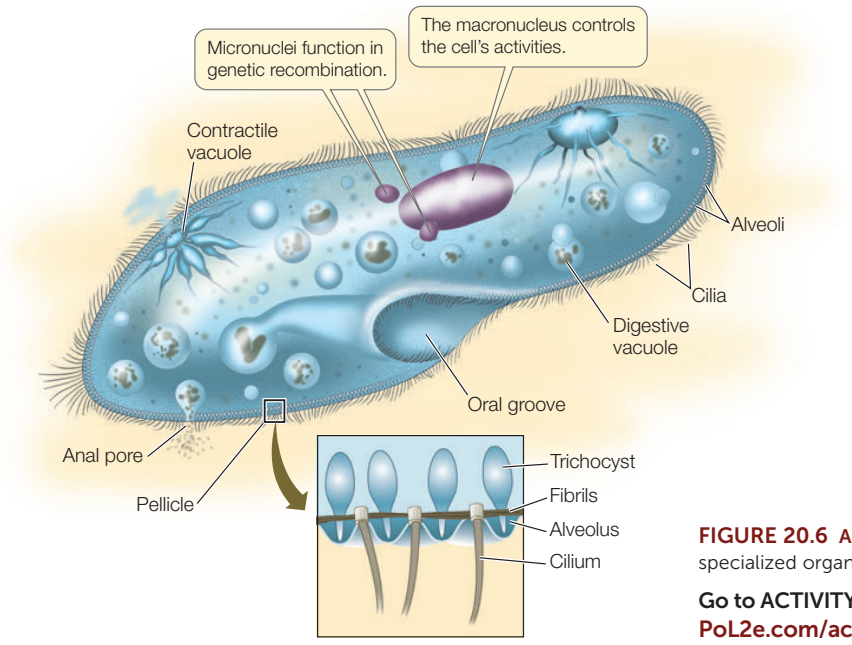


FIGURE 20.6 Anatomy of a *Paramecium* A *Paramecium*, with its many specialized organelles, exemplifies the complex body form of ciliates.

Go to ACTIVITY 20.1 Anatomy of a *Paramecium*
PoL2e.com/ac20.1

Stramenopiles typically have two unequal flagella, one with hairs

A morphological synapomorphy of most **stramenopiles** is the possession of rows of tubular hairs on the longer of their two flagella. Some stramenopiles lack flagella, but they are descended from ancestors that possessed flagella. The stramenopiles include the diatoms and the brown algae, which are photosynthetic, and the oomycetes, which are not.

DIATOMS All of the **diatoms** are unicellular, although some species associate in filaments. Many have sufficient carotenoids in their chloroplasts to give them a yellow or brownish color. All of them synthesize carbohydrates and oils as photosynthetic storage products. Diatoms lack flagella except in male gametes.

Architectural magnificence on a microscopic scale is the hallmark of the diatoms. Almost all diatoms deposit silica (hydrated silicon dioxide) in their cell walls. The cell wall of a diatom is constructed in two pieces, with the top overlapping the bottom like the top of a petri dish. The silica-impregnated walls have intricate patterns unique to each species (FIGURE 20.7). Despite their remarkable morphological diversity, all diatoms are symmetrical—either bilaterally (with "right" and "left" halves) or radially (with the type of symmetry possessed by a circle).

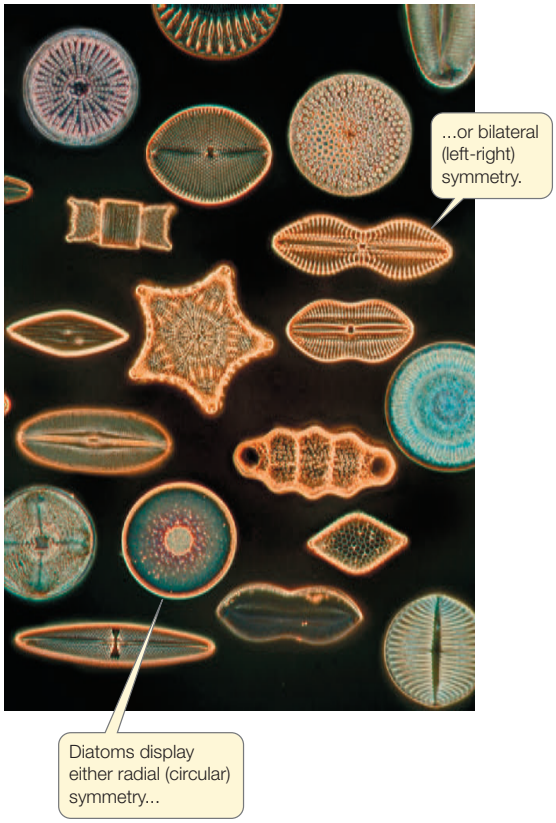


FIGURE 20.7 Diatom Diversity This dark-field micrograph illustrates the variety of species-specific forms found among the diatoms.

Go to MEDIA CLIP 20.2
Diatoms in Action
PoL2e.com/mc20.2

Diatoms reproduce both sexually and asexually. Asexual reproduction by binary fission is somewhat constrained by the stiff cell wall. Both the top and bottom of the "petri dish" become tops of new "dishes" without changing appreciably in size. As a result, the new cell made from the former bottom is smaller than the parent cell. If this process continued indefinitely, one cell line would simply vanish, but sexual reproduction largely solves this potential problem. Gametes are formed, shed their cell walls, and fuse. The resulting zygote then grows substantially in size before a new cell wall is laid down.

Diatoms are found in all the oceans and are frequently present in great numbers. They are major photosynthetic producers in coastal waters and are among the dominant organisms in the dense "blooms" of phytoplankton that occasionally appear in the open ocean (see Concept 20.4). Diatoms are also common in fresh water and even occur on the wet surfaces of terrestrial mosses.

BROWN ALGAE The **brown algae** obtain their namesake color from the carotenoid fucoxanthin, which is abundant in their chloroplasts. The combination of this yellow-orange pigment with the green of chlorophylls *a* and *c* yields a brownish tinge. All brown algae are multicellular, and some are extremely large. Giant kelps, such as those of the genus *Macrocystis*, may be up to 60 meters long.

Go to MEDIA CLIP 20.3
A Kelp Forest
PoL2e.com/mc20.3

The brown algae are almost exclusively marine. They are composed either of branched filaments (FIGURE 20.8A) or of leaflike growths (FIGURE 20.8B). Some float in the open ocean. The most famous example is the genus *Sargassum*, which forms dense mats in the Sargasso Sea in the mid-Atlantic. Most brown algae, however, attach themselves to rocks near the shore. A few thrive only where they are regularly exposed to heavy surf. All of the attached forms develop a specialized structure, called a holdfast, that literally glues them to the rocks. The "glue" of the holdfast is alginic acid, a gummy polymer of sugar acids found in the walls of many brown algal cells. In addition to its function in holdfasts, it cements algal cells and filaments together. It is harvested and used by humans as an emulsifier in ice cream, cosmetics, and other products.

OOMYCETES The **oomycetes** consist in large part of the water molds and their terrestrial relatives, such as the downy mildews. Water molds are filamentous and stationary. They are **absorptive heterotrophs**—that is, they secrete enzymes that digest large food molecules into smaller molecules that they can absorb. They are all aquatic and **saprobic**—meaning they feed on dead organic matter. If you have seen a whitish, cottony mold growing on dead fish or dead insects in water, it was probably a water mold of the common genus *Saprolegnia* (FIGURE 20.9).

(A) *Himanthalia elongata*

(B) *Postelsia palmaeformis*

FIGURE 20.8 Brown Algae (A) This seaweed illustrates the filamentous growth form of the brown algae. (B) Sea palms exemplify the leaflike growth form of brown algae. Sea palms and many other brown algal species are "glued" to the rocks by tough, branched structures called holdfasts that can withstand the pounding of the surf.

Some other oomycetes, such as the downy mildews, are terrestrial. Although most of the terrestrial oomycetes are harmless or helpful decomposers of dead matter, a few are plant parasites that attack crops such as avocados, grapes, and potatoes.

Oomycetes were once classified as fungi. However, we now know that their similarity to fungi is only superficial, and that the oomycetes are more distantly related to the fungi than are many other eukaryote groups, including humans (see Figure 20.3). For example, the cell walls of oomycetes are typically made of cellulose, whereas those of fungi are made of chitin.

Saprolegnia sp.

FIGURE 20.9 An Oomycete The filaments of a water mold radiate from the carcass of a beetle.

Rhizarians typically have long, thin pseudopods

The three primary groups of **rhizarians**—cercozoans, foraminiferans, and radiolarians—are unicellular and mostly aquatic. These organisms typically have long, thin pseudopods that contrast with the broader, lobelike pseudopods of the more familiar amoebozoans. The rhizarians have contributed to ocean sediments, some of which have become terrestrial features in the course of geological history.

CERCOZOANS The **cercozoans** are a diverse group with many forms and habitats. Some are aquatic, and others live in soil. One group of cercozoans possesses chloroplasts derived from a green alga by secondary endosymbiosis, and those chloroplasts contain a trace of the alga's nucleus.

FORAMINIFERANS Some **foraminiferans** secrete external shells of calcium carbonate (**FIGURE 20.10**). These shells have accumulated over time to produce much of the world's limestone. Foraminiferans live as plankton (drifing with the curent) or on the seafloor. Living foraminiferans have been found 10,896 meters down in the western Pacific's Challenger Deep—the deepest point in the world's oceans. At that depth, however, they cannot secrete normal shells because the surrounding water is too poor in calcium carbonate.

In living planktonic foraminiferans, long, threadlike, branched pseudopods extend through numerous microscopic openings in the shell and interconnect to create a sticky network, which the foraminiferans use to catch smaller

Elphidium sp.

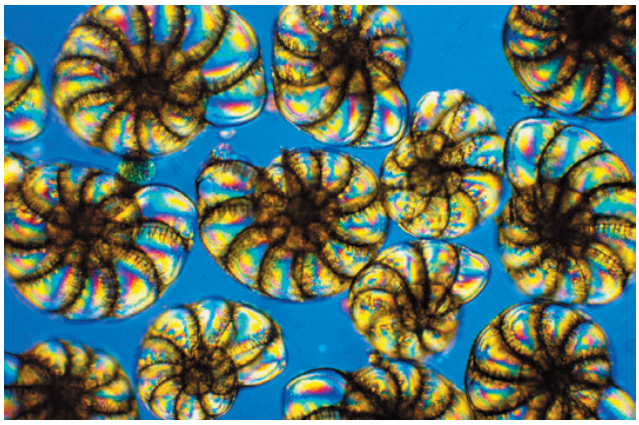


FIGURE 20.10 Building Blocks of Limestone Some foraminiferans secrete calcium carbonate to form shells. The shells of different species have distinctive shapes. Over millions of years, the shells of foraminiferans have accumulated to form limestone deposits.

plankton. In some foraminiferan species, the pseudopods provide locomotion.

RADIOLARIANS **Radiolarians** are recognizable by their thin, stiff pseudopods, which are reinforced by microtubules (**FIGURE 20.11A**). These pseudopods greatly increase the surface area of the cell, and they help the cell stay afloat in its marine environment.

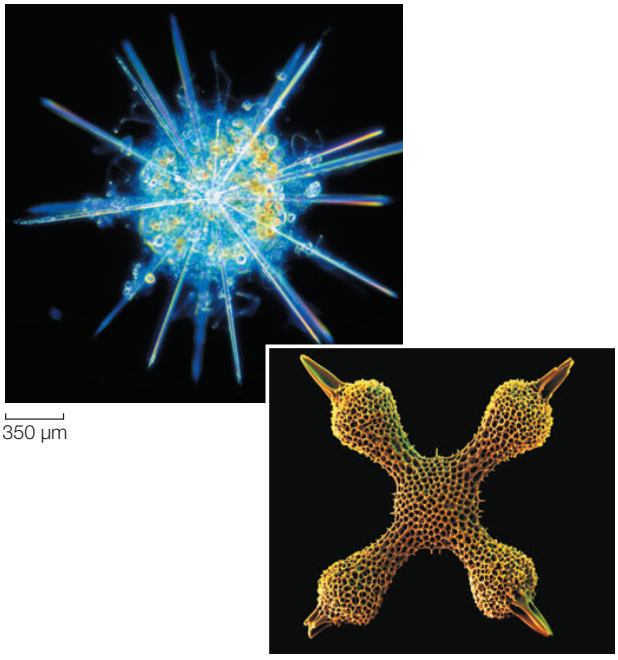
Radiolarians also are immediately recognizable by their distinctive radial symmetry. Almost all radiolarian species secrete glassy endoskeletons (internal skeletons). The skeletons of the different species are as varied as snowflakes, and many have elaborate geometric designs (**FIGURE 20.11B**). A few radiolarians are among the largest of the unicellular eukaryotes, measuring several millimeters in diameter.

Excavates began to diversify about 1.5 billion years ago

The **excavates** include a number of diverse groups that began to split from one another soon after the origin of eukaryotes. Several groups of excavates lack mitochondria, an absence that once led to the view that these groups might represent early-diverging eukaryotes that diversified before the evolution of mitochondria. However, the discovery of nuclear genes normally associated with mitochondria in these organisms suggests that the absence of mitochondria is a derived condition. In other words, ancestors of these excavate groups probably possessed mitochondria that were lost or reduced over the course of evolution. The existence of these organisms today shows that eukaryotic life is possible without mitochondria.

DIPLOMONADS AND PARABASALIDS The **diplomonads** and **parabasalids** are unicellular and lack mitochondria. *Giardia lamblia*, a diplomonad, is a familiar parasite that contaminates water supplies and causes the intestinal disease giardiasis. This tiny organism contains two nuclei bounded by

(A)



(B) *Crucella* sp.

50 µm

FIGURE 20.11 Radiolarians Exhibit Distinctive Pseudopods and Radial Symmetry (A) The radiolarians are distinguished by their thin, stiff pseudopods and by their radial symmetry. The pigmentation seen at the center of this radiolarian's glassy endoskeleton is imparted by endosymbiotic dinoflagellates. (B) The endoskeleton secreted by a radiolarian.

nuclear envelopes, and it has a cytoskeleton and multiple flagella (**FIGURE 20.12A**).

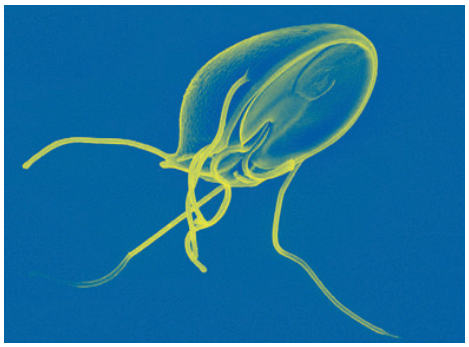
In addition to flagella and a cytoskeleton, the parabasalids have undulating membranes that also contribute to the cell's locomotion. *Trichomonas vaginalis* (**FIGURE 20.12B**) is a parabasalid responsible for a sexually transmitted disease (trichomoniasis) in humans. Infection of the male urethra, where it may occur without symptoms, is less common than infection of the vagina.

HETEROLOBOSEANS The constantly changing amoeboid body form appears in several protist groups that are only distantly related to one another. The body forms of **heterolobose-ans**, for example, resemble those of loboseans, an amoebozoan group that is not at all closely related to heterolobose-ans (see p. 412). Amoebas of the free-living heterolobosean genus *Naegleria* usually have a two-stage life cycle in which one stage has amoeboid cells and the other flagellated cells. Some amoebas of this genus can enter the human body and cause a fatal disease of the nervous system.

EUGLENIDS AND KINETOPLASTIDS The **euglenids** and **kinetoplastids** together constitute a clade of unicellular excavates with flagella. Their mitochondria contain distinctive, disc-shaped cristae, and their flagella contain a crystalline rod not found in other organisms. They reproduce primarily asexually by binary fission.

The flagella of euglenids arise from a pocket at the anterior end of the cell. Spiraling strips of proteins under the

(A) *Giardia muris*



(B) *Trichomonas vaginalis*

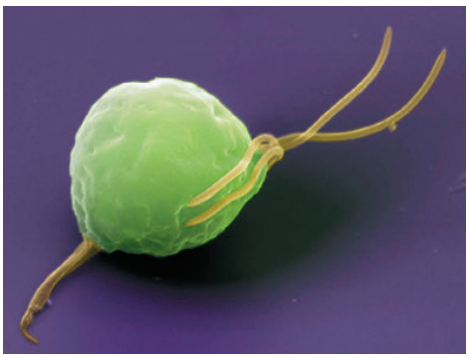


FIGURE 20.12 Some Excavate Groups Lack Mitochondria
(A) *Giardia*, a diplomonad, has flagella and two nuclei. (B) *Trichomonas*, a parabasalid, has flagella and undulating membranes. Neither of these organisms possesses mitochondria.

cell membrane control the cell's shape. Some euglenids are photosynthetic.

FIGURE 20.13 depicts a cell of the genus *Euglena*. Like most other euglenids, this common freshwater organism has a complex cell structure. It propels itself through the water with the longer of its two flagella, which may also serve as an anchor to hold the organism in place. The second flagellum is often rudimentary.

Go to MEDIA CLIP 20.4
Euglenids
PoL2e.com/mc20.4

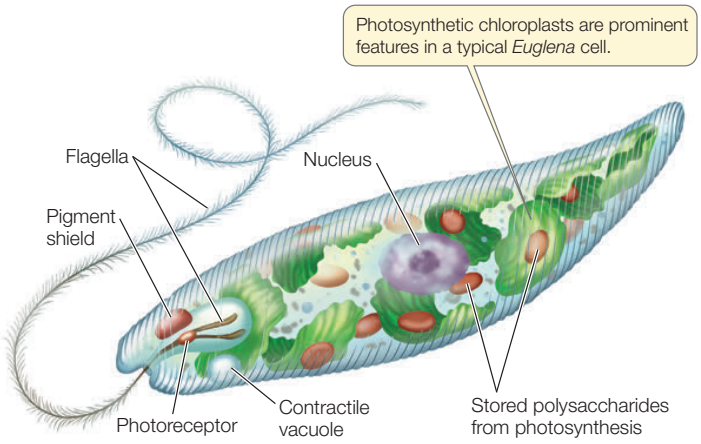


FIGURE 20.13 A Photosynthetic Euglenid In the *Euglena* species illustrated in this drawing, the second flagellum is rudimentary. Note that the primary flagellum originates at the anterior of the organism and trails toward its posterior.

The euglenids have diverse nutritional requirements. Many species are always heterotrophic. Other species are fully autotrophic in sunlight, using chloroplasts to synthesize organic compounds through photosynthesis. When kept in the dark, these euglenids lose their photosynthetic pigment and begin to feed exclusively on dissolved organic material in the water around them. Such a "bleached" *Euglena* resynthesizes its photosynthetic pigment when it is returned to the light and becomes autotrophic again. But *Euglena* cells treated with certain antibiotics or mutagens lose their photosynthetic pigment completely; neither they nor their descendants are ever autotrophs again. However, those descendants function well as heterotrophs.

The kinetoplastids are unicellular parasites with two flagella and a single, large mitochondrion. That mitochondrion contains a kinetoplast, a unique structure housing multiple circular DNA molecules and associated proteins. Some of these DNA molecules encode "guide proteins" that edit mRNA within the mitochondrion.

The kinetoplastids include several medically important species of pathogenic trypanosomes (**TABLE 20.1**). Some of these

TABLE 20.1 A Comparison of Three Kinetoplastid Trypanosomes

	Trypanosoma brucei	*Trypanosoma cruzi*	*Leishmania major*
Human disease	Sleeping sickness	Chagas' disease	Leishmaniasis
Insect vector	Tsetse fly	Assassin bug	Sand fly
Vaccine or effective cure	None	None	None
Strategy for survival	Changes surface recognition molecules frequently	Causes changes in surface recognition molecules on host cell	Reduces effectiveness of macrophage hosts
Site in human body	Bloodstream; attacks nerve tissue in final stages	Enters cells, especially muscle cells	Enters cells, primarily macrophages
Approximate number of deaths per year	50,000	43,000	60,000

organisms are able to change their cell surface recognition molecules frequently, allowing them to evade our best attempts to kill them and thus eradicate the diseases they cause.

Amoebozoans use lobe-shaped pseudopods for locomotion

Amoebozoans appear to have diverged from other eukaryotes about 1.5 billion years ago (see Figure 20.3). It is not yet clear if they are more closely related to opisthokonts (including fungi and animals) or to other major groups of eukaryotes.

The lobe-shaped pseudopods of amoebozoans (**FIGURE 20.14**) are a hallmark of the amoeboid body form. Amoebozoan pseudopods differ in form and function from the slender pseudopods of rhizarians. We consider three amoebozoan groups here: the loboseans and two groups known as slime molds.

LOBOSEANS **Loboseans** are small amoebozoans that feed on other small organisms and particles of organic matter by phagocytosis, engulfing them with pseudopods. Many loboseans are adapted for life on the bottoms of lakes, ponds, and other bodies of water. Their creeping locomotion and their manner of engulfing food particles fit them for life close to a relatively rich supply of sedentary organisms or organic particles. Most loboseans exist as predators, parasites, or scavengers. Members of one group of loboseans, the testate amoebas, live inside shells. Some of these amoebas produce casings by gluing sand grains together (**FIGURE 20.15**). Other testate amoebas have shells secreted by the organism itself.

Amoeba proteus

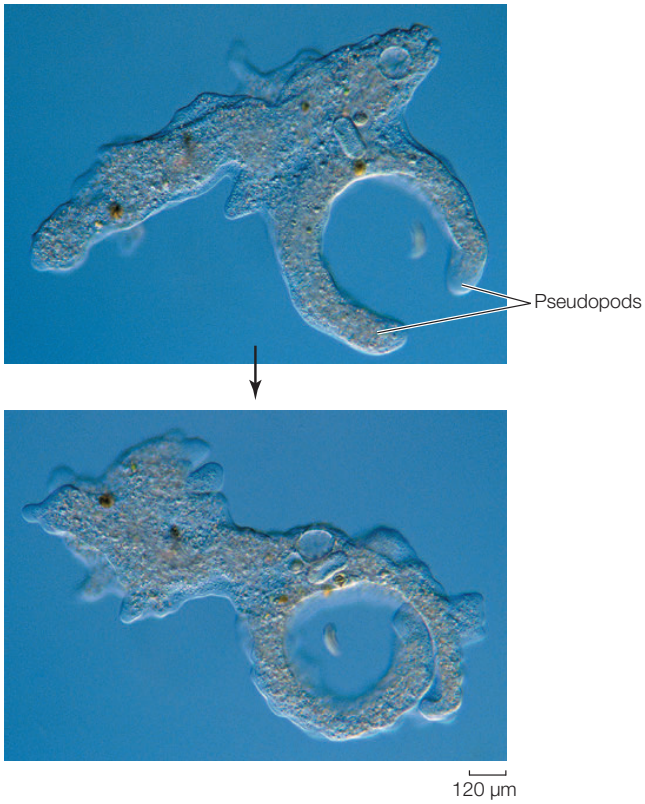


FIGURE 20.14 An Amoeba in Motion The flowing pseudopods of this "chaos amoeba" are constantly changing shape as it moves and feeds.

Go to MEDIA CLIP 20.5
Amoeboid Movement
PoL2e.com/mc20.5

Nebela collaris

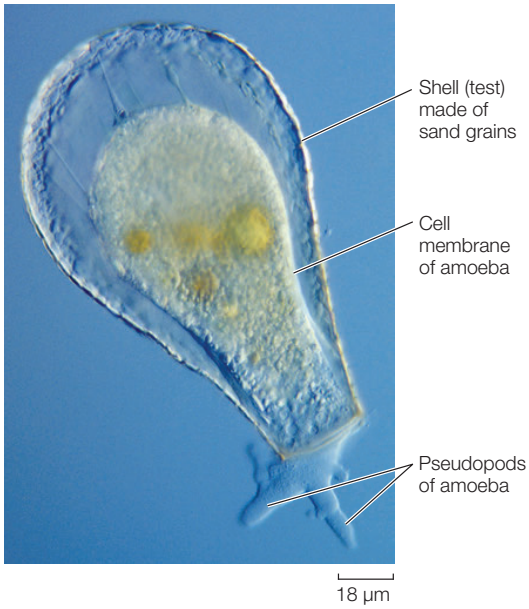
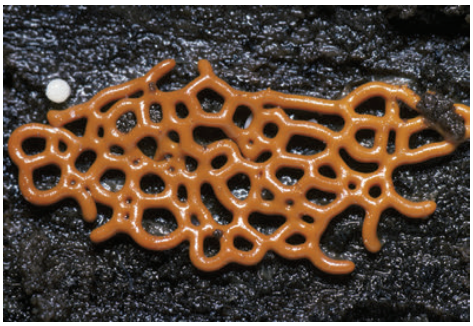


FIGURE 20.15 Life in a Glass House This testate amoeba has built a lightbulb-shaped shell, or test, by gluing sand grains together. Its pseudopods extend through the single aperture in the test.

PLASMODIAL SLIME MOLDS If the nucleus of an amoeba began rapid mitotic division, accompanied by a tremendous increase in cytoplasm and organelles but no cytokinesis (division of the cytoplasm), the resulting organism would resemble the multinucleate mass of a **plasmodial slime mold**. During its vegetative (feeding, nonreproductive) stage, a plasmodial slime mold is a wall-less mass of cytoplasm with numerous diploid nuclei. This mass streams very slowly over its substrate in a remarkable network of strands called a plasmodium (**FIGURE 20.16A**). The plasmodium of such a slime mold is an example of a **coenocyte**: many nuclei enclosed in a single cell membrane. The outer cytoplasm of the plasmodium (closest to the environment) is normally less fluid than the interior cytoplasm and thus provides some structural rigidity.

Plasmodial slime molds provide a dramatic example of movement by **cytoplasmic streaming**. The outer cytoplasmic region of the plasmodium becomes more fluid in places, and cytoplasm rushes into those areas, stretching the plasmodium. This streaming somehow reverses its direction every few minutes as cytoplasm rushes into a new area and drains away from an older one, moving the plasmodium over its substrate. Sometimes an entire wave of plasmodium moves across a surface, leaving strands behind. Microfilaments and a contractile protein interact to produce the streaming movement. As it

(A)



30 mm

(B)

1.5 mm

FIGURE 20.16 Plasmodial Slime Molds (A) The plasmodial form of the slime mold *Hemitrichia serpula* covers rocks, decaying logs, and other objects as it engulfs bacteria and other food items; it is also responsible for its common name of "pretzel mold." (B) Fruiting structures of *Hemitrichia*.

Go to MEDIA CLIP 20.6
Plasmodial Slime Mold Growth
PoL2e.com/mc20.6

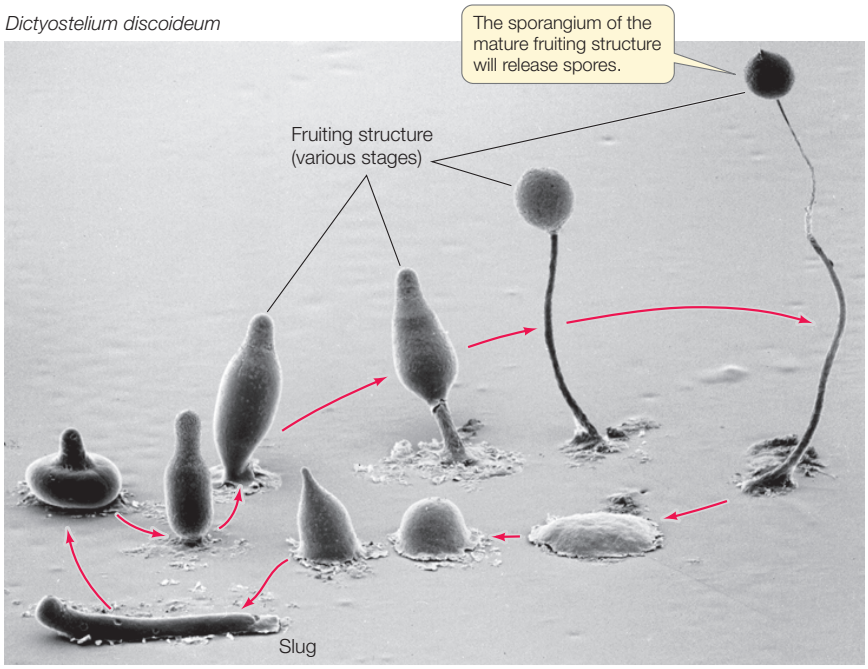
moves, the plasmodium engulfs food particles by endocytosis—predominantly bacteria, yeasts, spores of fungi, and other small organisms as well as decaying animal and plant remains.

A plasmodial slime mold can grow almost indefinitely in its plasmodial stage as long as the food supply is adequate and other conditions, such as moisture and pH, are favorable. If conditions become unfavorable, however, one of two things can happen. First, the plasmodium can form an irregular mass of hardened cell-like components. This resting structure rapidly becomes a plasmodium again when favorable conditions are restored.

Alternatively, the plasmodium can transform itself into spore-bearing fruiting structures (**FIGURE 20.16B**). These stalked or branched structures rise from heaped masses of plasmodium. They derive their rigidity from walls that form and thicken between their nuclei. The diploid nuclei of the plasmodium divide by meiosis as the fruiting structure develops. One or more knobs, called sporangia, develop on the end of the stalk. Within a sporangium, haploid nuclei become surrounded by walls to form spores. Eventually, as the fruiting structure dries, it sheds its spores.

The spores germinate into wall-less, haploid cells called swarm cells, which can either divide mitotically to produce more haploid swarm cells or function as gametes. Swarm cells can live as separate individual cells that move by means of flagella or pseudopods, or they can become walled and resistant resting cysts when conditions are unfavorable. When conditions improve again, the cysts release swarm cells. Two swarm cells can also fuse to form a diploid zygote, which divides by mitosis (but without a wall forming between the nuclei) and thus forms a new, coenocytic plasmodium.

Dictyostelium discoideum



0.25 mm

CELLULAR SLIME MOLDS Whereas the plasmodium is the basic vegetative unit of the plasmodial slime molds, an amoeboid cell is the vegetative unit of the **cellular slime molds** (**FIGURE 20.17**). Cells called myxamoebas, which have single haploid nuclei, engulf bacteria and other food particles by endocytosis and reproduce by mitosis and binary fission. This simple life cycle stage, consisting of swarms of independent, isolated cells, can persist indefinitely as long as food and moisture are available.

FIGURE 20.17 A Cellular Slime Mold This composite micrograph shows the life cycle of the slime mold *Dictyostelium*.

Go to MEDIA CLIP 20.7
Cellular Slime Mold Aggregation
PoL2e.com/mc20.7

When conditions become unfavorable, the cellular slime molds aggregate and form fruiting structures, as do their plasmodial counterparts. The individual myxamoebas aggregate into a mass called a **slug** or **pseudoplasmodium**. Unlike the true plasmodium of the plasmodial slime molds, however, this structure is not simply a giant sheet of cytoplasm with many nuclei. Instead, the individual myxamoebas retain their cell membranes, and therefore their identity.

A slug may migrate over a substrate for several hours before becoming motionless and reorganizing to construct a delicate, stalked fruiting structure. Cells at the top of the fruiting structure develop into thick-walled spores, which are eventually released. Later, under favorable conditions, the spores germinate, releasing myxamoebas.

The cycle from myxamoebas through slug and spores to new myxamoebas is asexual. Cellular slime molds also have a sexual cycle, in which two myxamoebas fuse. The product of this fusion develops into a spherical structure that ultimately germinates, releasing new haploid myxamoebas.

CHECKpoint CONCEPT 20.2

- ✓ Explain why the term "protists" does not refer to a formal taxonomic group.
- ✓ Contrast the major distinctive features of alveolates, excavates, stramenopiles, rhizarians, and amoebozoans. Which of these groups is most closely related to fungi and animals? What morphological evidence supports this relationship?
- ✓ The fossil record of eukaryotes in the Precambrian is poor compared with those of the Cambrian and later geological periods, even though eukaryotes were diversifying for the last billion years of the Precambrian. Can you think of some possible reasons for the better fossil record beginning in the Cambrian?

The ancient origins of major eukaryote lineages, and the adaptation of these lineages to a wide variety of lifestyles and environments, resulted in enormous protist diversity. It is not surprising, then, that reproductive modes among protists are also highly diverse.

CONCEPT 20.3 Protists Reproduce Sexually and Asexually

Although most protists engage in both asexual and sexual reproduction, sexual reproduction has yet to be observed in some groups. In some protists, as in all prokaryotes, the acts of sex and reproduction are not directly linked.

Several asexual reproductive processes have been observed among the protists:

- The equal splitting of one cell into two by mitosis followed by cytokinesis
- The splitting of one cell into multiple (i.e., more than two) cells
- The outgrowth of a new cell from the surface of an old one (known as **budding**)
- The formation of specialized cells (spores) that are capable of developing into new individuals (known as **sporulation**)

Asexual reproduction results in offspring that are genetically nearly identical to their parents (they only differ by new mutations that may arise during DNA replication). Such asexually reproduced groups of nearly identical organisms are known as **clonal lineages**.

Sexual reproduction among the protists takes various forms. In some protists, as in animals, the gametes are the only haploid cells. In others, the zygote is the only diploid cell. In still others, both diploid and haploid cells undergo mitosis, giving rise to alternating multicellular diploid and haploid life stages.

LINK

Asexual and sexual reproduction are described in **Concept 7.1**

Some protists have reproduction without sex and sex without reproduction

As we noted in Concept 20.2, members of the genus *Paramecium* are ciliates, which commonly have two types of nuclei in a single cell (one macronucleus and from one to several micronuclei; see Figure 20.6). The micronuclei, which are typical eukaryotic nuclei, are essential for genetic recombination. Each macronucleus contains many copies of the genetic information, packaged in units containing very few genes each. The macronuclear DNA is transcribed and translated to regulate the life of the cell.

When paramecia reproduce asexually, all of the nuclei are copied before the cell divides. Paramecia also have an elaborate sexual behavior called **conjugation**, in which two individuals line up tightly against each other and fuse in the oral groove region of the body. Nuclear material is extensively reorganized and exchanged over the next several hours (**FIGURE 20.18**). Each cell ends up with two haploid micronuclei, one of its own and one from the other cell, which fuse to form a new diploid micronucleus. A new macronucleus develops from the micronucleus through a series of dramatic chromosomal rearrangements. The exchange of nuclei is fully reciprocal: each of the two paramecia gives and receives an equal amount of DNA. The two organisms then separate and go their own ways, each equipped with new combinations of alleles.

Conjugation in *Paramecium* is a sexual process of genetic recombination, but it is not a reproductive process. Two cells begin the process, and two cells are there at the end, so no new cells are created. As a rule, each asexual clone of paramecia must periodically conjugate. Experiments have shown that if some species are not permitted to conjugate, the clones can live through no more than approximately 350 cell divisions before they die out.

Some protist life cycles feature alternation of generations

Alternation of generations is a type of life cycle found in many multicellular protists, all land plants, and some fungi. A

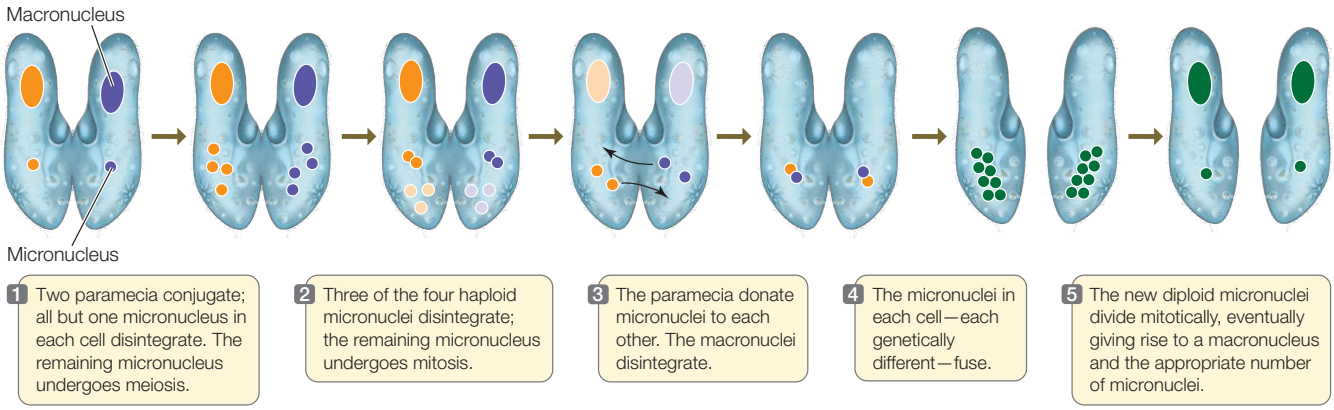


FIGURE 20.18 Paramecia Achieve Genetic Recombination by Conjugating The exchange of micronuclei by two conjugating *Paramecium* individuals results in genetic recombination. After conjugation, the cells separate and continue their lives as two individuals.

multicellular, diploid, spore-producing organism gives rise to a multicellular, haploid, gamete-producing organism. When two haploid gametes fuse, a diploid organism is produced. The haploid organism, the diploid organism, or both may also reproduce asexually.

The two alternating (spore-producing and gamete-producing) generations differ genetically (one has diploid cells, the other haploid cells), but they may or may not differ morphologically. In **heteromorphic** alternation of generations, the two generations differ morphologically, but in **isomorphic** alternation of generations, they do not. Examples of both heteromorphic and isomorphic alternation of generations are found among the brown algae.

The gamete-producing generation does not produce gametes by meiosis because the gamete-producing organism is already haploid. Instead, specialized cells of the diploid spore-producing organism, called **sporocytes**, divide meiotically to produce four haploid spores. The spores may eventually germinate and divide mitotically to produce the multicellular haploid generation, which then produces gametes by mitosis and cytokinesis.

Gametes, unlike spores, can produce new organisms only by fusing with other gametes. The fusion of two gametes produces a diploid zygote, which then undergoes mitotic divisions to produce a diploid organism. The diploid organism's sporocytes then undergo meiosis and produce haploid spores, starting the cycle anew.

CHECKpoint CONCEPT 20.3

- ✓ Why is conjugation between paramecia considered a sexual process but not a reproductive process?
- ✓ Why do you think paramecia that are not allowed to conjugate begin to die out after about 350 rounds of asexual reproduction?
- ✓ Although most diploid animals have haploid stages (for example, eggs and sperm), their life cycles are not considered alternation of generations. Why not?

CONCEPT 20.4 Protists Are Critical Components of Many Ecosystems

Some protists are food for marine animals, while others poison those animals. Some are packaged as nutritional supplements for humans, and some are human pathogens. The remains of some form the sands of many modern beaches, and others are a major source of the oil that sometimes fouls those beaches.

Phytoplankton are primary producers

A single protist clade, the diatoms, performs about one-fifth of all photosynthetic carbon fixation on Earth—about the same amount as all of Earth's rainforests. These spectacular unicellular organisms (see Figure 20.7) are the predominant component of the phytoplankton, but the phytoplankton include many other protists that contribute heavily to global photosynthesis. Like green plants on land, these "floating photosynthesizers" are the gateway for energy from the sun into the rest of the living world. In other words, they are **primary producers**. These autotrophs are eaten by heterotrophs, including animals and many other protists. Those consumers are, in turn, eaten by other consumers. Most aquatic heterotrophs (with the exception of some species in the deep sea) depend on photosynthesis performed by phytoplankton.

LINK

Much of the energy contained in autotrophs is unavailable to the heterotrophs that eat them; see **Concept 44.3** for a discussion of energy flow through communities

Some microbial eukaryotes are deadly

Some microbial eukaryotes are pathogens that cause serious diseases in humans and other vertebrates. The best-known pathogenic protists are members of the genus *Plasmodium*, a highly specialized group of apicomplexans that spend part of their life cycle as parasites in human red blood cells, where they are the cause of malaria. In terms of the number of people affected, malaria is one of the world's three most serious infectious diseases: it infects more than 350 million people, and kills more than 1 million people, each year. On average, about two people die from malaria every minute of every day—most of

APPLY THE CONCEPT

Protists are critical components of many ecosystems

In most temperate regions of the oceans, there is a spring bloom of phytoplankton. Although the red tide blooms described in this chapter's opening story are harmful, phytoplankton blooms can also be beneficial for marine communities. In fact, many species of marine life depend on these blooms for their survival.

The dates of spring phytoplankton blooms near the coast of Nova Scotia, Canada, were determined by examining remote satellite images. The table to the right presents these dates as deviations from the mean date of the spring bloom in this region.[a] The table also gives the survival index for larval haddock (an important commercial fish) for the year after each bloom. The survival index is the ratio of the mass of juvenile fish to the mass of mature fish; higher values indicate better survival of larval fish.

1. Plot the survival index of larval haddock against the deviation in the date of the spring phytoplankton bloom. Calculate a correlation coefficient for their relationship (see Appendix B).
2. Formulate one or more hypotheses to explain your results. Keep in mind that larval haddock include phytoplankton in their diet, and that phytoplankton blooms also provide some cover in which larval fish can hide from potential predators.

YEAR	DEVIATION IN BLOOM DATE* (DAYS)	SURVIVAL INDEX
1	+5	1.9
2	+11	2.2
3	–15	6.8
4	+5	1.9
5	–4	4.9
6	–20	10.3
7	+6	2.1
8	+14	1.9

*Negative values indicate blooms occurring earlier than the mean date; positive values indicate later blooms.

[a]T. Platt et al. 2003. *Nature* 423: 398–399.

them in sub-Saharan Africa, although malaria occurs in more than 100 countries.

Mosquitoes of the genus *Anopheles* transmit *Plasmodium* to humans. The parasites enter the human circulatory system when an infected female *Anopheles* mosquito penetrates the skin in search of blood. The parasites find their way to cells in the liver and the lymphatic system, change their form, multiply, and re-enter the bloodstream, where they invade red blood cells.

The parasites multiply inside the red blood cells, which then burst, releasing new swarms of parasites. These episodes of bursting red blood cells coincide with the primary symptoms of malaria, which include fever, shivering, vomiting, joint pains, and convulsions. If another *Anopheles* bites the victim, the mosquito takes in *Plasmodium* cells along with blood. Some of the ingested cells develop into gametes that unite in the mosquito, forming zygotes. The zygotes lodge in the mosquito's gut, divide several times, and move into its salivary glands, from which they can be passed on to another human host. Thus *Plasmodium* is an extracellular parasite in the mosquito vector and an intracellular parasite in the human host.

Go to ANIMATED TUTORIAL 20.3
Life Cycle of the Malarial Parasite
PoL2e.com/at20.3

Plasmodium has proved to be a singularly difficult pathogen to attack. The complex *Plasmodium* life cycle is best broken by the removal of stagnant water, in which mosquitoes breed. Using insecticides to reduce the *Anopheles* population can also be effective, but the benefits must be weighed against the ecological, economic, and health risks posed by the insecticides themselves.

Even some of the phytoplankton that are such important primary producers can be deadly, as described in this chapter's opening story. Some diatoms and dinoflagellates reproduce in enormous numbers when environmental conditions are favorable for their growth. In the resulting red tides, the concentration of dinoflagellates may reach 60 million per liter of ocean water and produce potent nerve toxins that harm or kill many vertebrates, especially fish.

Some microbial eukaryotes are endosymbionts

Endosymbiosis is common among the microbial eukaryotes, many of which live within the cells of animals. Many radiolarians harbor photosynthetic endosymbionts (see Figure 20.11A). As a result, these radiolarians, which are not photosynthetic themselves, appear greenish or golden, depending on the type of endosymbiont they contain. This arrangement is often mutually beneficial: the radiolarian can make use of the carbon compounds produced by its photosynthetic guest, and the guest may in turn make use of metabolites made by the host or receive physical protection. In some cases, the guest is exploited for its photosynthetic products while receiving little or no benefit itself.

Dinoflagellates are also common endosymbionts and can be found in both animals and other protists. Most, but not all, dinoflagellate endosymbionts are photosynthetic. Some dinoflagellates live endosymbiotically in the cells of corals, contributing the products of their photosynthesis to the partnership. Their importance to the corals is demonstrated when the dinoflagellates die or are expelled by the corals as a result of changing environmental conditions such as rising water temperatures or increased water turbidity. This phenomenon is known as **coral bleaching**. Unless the corals can acquire new endosymbionts, they are ultimately damaged or destroyed as a result of their reduced food supply (**FIGURE 20.19**).

INVESTIGATION

FIGURE 20.19 Can Corals Reacquire Dinoflagellate Endosymbionts Lost to Bleaching? Some corals lose their chief nutritional source when their photosynthetic endosymbionts die, often as a result of changing environmental conditions. This experiment by Cynthia Lewis and Mary Alice Coffroth investigated the ability of corals to acquire new endosymbionts after bleaching.[a]

HYPOTHESIS

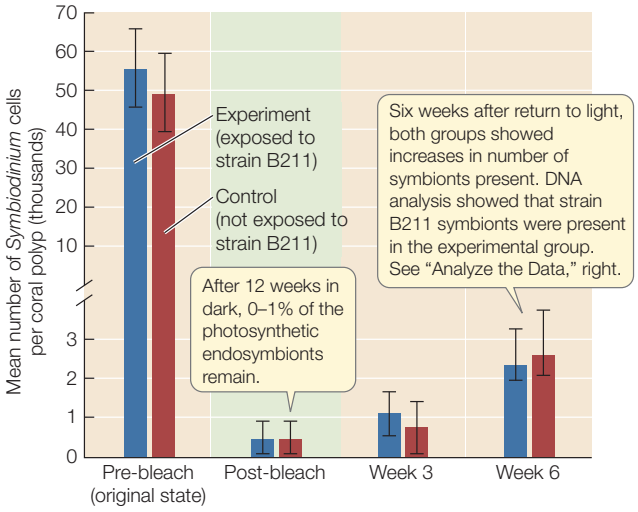
Bleached corals can acquire new photosynthetic endosymbionts from their environment.

METHOD

1. Count numbers of *Symbiodinium*, a photosynthetic dinoflagellate, living symbiotically in samples of a coral (*Briareum* sp.).
2. Stimulate bleaching by maintaining all *Briareum* colonies in darkness for 12 weeks.
3. After 12 weeks of darkness, count numbers of *Symbiodinium* in the coral samples; then return all colonies to light.
4. In some of the bleached colonies (the experimental group), introduce *Symbiodinium* strain B211—dinoflagellates that contain a unique molecular marker. A control group of bleached colonies is not exposed to strain B211. Maintain both groups in the light for 6 weeks.

RESULTS

Error bars indicate 95% confidence intervals for the mean (see Appendix B).



CONCLUSION

Corals can acquire new strains of endosymbionts from their environment following bleaching.

ANALYZE THE DATA

These data—the results of DNA analysis of the *Symbiodinium* endosymbionts—reveal that many of the experimental colonies took up strain B211 from their environment. The control colonies recovered their native *Symbiodinium*, except in colonies in which endosymbionts were completely lost. Use these data to answer the questions below.

	Symbiodinium strain present (% of colonies)			
	Non-B211	B211	None*	Colony died
Experimental colonies (strain B211 added)				
Pre-bleach	100	0	0	0
Post-bleach	58	0	42	0
Week 3	0	92	0	8
Week 6	8	58	8	25
Control colonies (no strain B211)				
Pre-bleach	100	0	0	0
Post-bleach	67	0	33	0
Week 3	67	0	33	0
Week 6	67	0	17	17

*Colonies remained alive but no *Symbiodinium* were detected.

A. Are new strains of *Symbiodinium* taken up only by coral colonies that have lost all their original endosymbionts?

B. Does the acquisition of a new *Symbiodinium* strain always result in survival of a recovering *Briareum* colony?

C. In week 3, only strain B211 was detected in the experimental colonies, but in week 6, non-B211 *Symbiodinium* were detected in 8% of the experimental colonies. Can you suggest an explanation for this observation?

Pre-bleach

Post-bleach

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]C. L. Lewis and M. A. Coffroth. 2004. *Science* 304: 1490–1492.

We rely on the remains of ancient marine protists

Diatoms are lovely to look at, but their importance to us goes far beyond aesthetics, and even beyond their role as primary producers. Diatoms store oil as an energy reserve and to keep themselves afloat at the correct depth in the ocean. Over millions of years, diatoms have died and sunk to the ocean floor, where they have been compressed by sediments and undergone chemical changes. In this way, they have become a major source of petroleum and natural gas, two of our most important energy supplies and political concerns.

Because the silica-containing cell walls of dead diatoms resist decomposition, some sedimentary rocks are composed almost entirely of diatom skeletons that sank to the seafloor over time. Diatomaceous earth, which is obtained from such rocks, has many industrial uses, such as insulation, filtration, and metal polishing. It has also been used as an "Earth-friendly" insecticide that clogs the tracheae (breathing structures) of insects.

Other ancient marine protists have also contributed to today's world. Some foraminiferans, as we have seen, secrete shells of calcium carbonate. After they reproduce (by mitosis and cytokinesis), the daughter cells abandon the parent shell and make new shells of their own. The discarded shells of ancient foraminiferans make up extensive limestone deposits in various parts of the world, forming a layer hundreds to thousands of meters deep over millions of square kilometers of ocean bottom. Foraminiferan shells also make up much of the sand of some beaches. A single gram of such sand may contain as many as 50,000 foraminiferan shells and shell fragments.

The shells of individual foraminiferans are easily preserved as fossils in marine sediments. Each geological period has a distinctive assemblage of foraminiferan species. Because the shells of foraminiferan species have distinctive shapes (see Figure 20.10), and because they are so abundant, the remains of foraminiferans are especially valuable in classifying and dating sedimentary rocks. In addition, analyses of the chemical makeup of foraminiferan shells can be used to estimate the global temperatures prevalent at the time when the shells were formed.

CHECKpoint CONCEPT 20.4

- ✓ What is the role of female *Anopheles* mosquitoes in the transmission of malaria?
- ✓ Explain the role of dinoflagellates in the two very different phenomena of coral bleaching and red tides.
- ✓ What are some of the ways in which diatoms are important to human society?

The next three chapters will explore the three major groups of multicellular eukaryotes, along with the protist ancestors from which they arose. Chapter 21 will describe the origin and diversification of the plants, Chapter 22 will present the fungi, and Chapter 23 will describe the animals.

FIGURE 20.20 Light Up the Sea Bioluminescent dinoflagellates flash as an outrigger disturbs the ocean surface off the island of Bali.

Q Red tides are harmful, but can dinoflagellates also be beneficial to marine ecosystems?

ANSWER Not all dinoflagellate blooms produce problems for other species. Dinoflagellates are important components of many ecosystems, as we have seen throughout this chapter. Corals and many other species depend on symbiotic dinoflagellates for photosynthesis (see Figure 20.19). In addition, as photosynthetic organisms, free-living planktonic dinoflagellates are among the most important primary producers in aquatic food webs. They are a major component of the phytoplankton and provide an important food source for many species (see Concept 20.4). Photosynthetic dinoflagellates also produce much of the atmospheric oxygen that most animals need to survive.

Some dinoflagellates produce a beautiful bioluminescence (**FIGURE 20.20**). Unlike the bioluminescent bacteria discussed at the start of Chapter 19, however, dinoflagellate bioluminescence is not steady. These protists produce flashes of light when disturbed, as people who swim in the ocean at night in certain regions often observe.

SUMMARY

CONCEPT 20.1 Eukaryotes Acquired Features from Both Archaea and Bacteria

- Early events in the evolution of the eukaryotic cell probably included the loss of the firm cell wall and infolding of the cell membrane. Such infolding probably led to segregation of the genetic material in a membrane-enclosed nucleus. **Review Figure 20.1**
- Some organelles were acquired by endosymbiosis. Mitochondria evolved by endosymbiosis with a proteobacterium.
- **Primary endosymbiosis** of a eukaryote and a cyanobacterium gave rise to the first chloroplasts. **Secondary endosymbiosis** and **tertiary endosymbiosis** between chloroplast-containing eukaryotes and other eukaryotes gave rise to the distinctive chloroplasts of euglenids, dinoflagellates, and other groups. **Review Figure 20.2 and ANIMATED TUTORIAL 20.1**

CONCEPT 20.2 Major Lineages of Eukaryotes Diversified in the Precambrian

- Most eukaryotes can be placed in one of eight major clades that originated in the Precambrian: alveolates, excavates, stramenopiles, rhizarians, amoebozoans, Plantae, fungi, and animals. The first five of these clades are collectively referred to as **protists**. **Review Figure 20.3**
- The term "protist" does not describe a formal taxonomic group, but rather is shorthand for "all eukaryotes that are not plants, animals, or fungi." Most, but not all, protists are unicellular.
- **Alveolates** are unicellular organisms with sacs (alveoli) beneath their cell membranes. Alveolate clades include the marine **dinoflagellates**, the parasitic **apicomplexans**, and the diverse, highly motile **ciliates**. **Review Figure 20.6, ACTIVITY 20.1, and ANIMATED TUTORIAL 20.2**
- **Stramenopiles** typically have two flagella of unequal length, the longer one bearing rows of tubular hairs. Among the stramenopiles are the unicellular **diatoms**, the multicellular **brown algae**, and the nonphotosynthetic **oomycetes**, many of which are **saprobic**.
- **Rhizarians** are unicellular and aquatic. They include the **foraminiferans**, whose shells have contributed to great limestone deposits; the **radiolarians**, which have thin, stiff pseudopods and glassy endoskeletons; and the **cercozoans**, which take many forms and live in diverse habitats.
- The **excavates** include a wide variety of symbiotic as well as free-living species. The **diplomonads** and **parabasalids** lack mitochondria, having apparently lost them during the course of their evolution. **Heteroloboseans** are amoebas with a two-stage life cycle. **Euglenids** are often photosynthetic and have anterior flagella and spiraling strips of protein that support their cell surface. The **kinetoplastids**, which include several human pathogens, have a single, large mitochondrion. **Review Figure 20.13**
- The **amoebozoans** move by means of lobe-shaped pseudopods. A **lobosean** consists of a single amoeboid cell. **Plasmodial slime molds** are amoebozoans whose vegetative stage is a **coenocyte** that moves by cytoplasmic streaming. In **cellular slime molds**, the individual cells maintain their identity at all times but aggregate to form fruiting structures. **Review Figure 20.17**

CONCEPT 20.3 Protists Reproduce Sexually and Asexually

- Asexual reproduction gives rise to **clonal lineages** of organisms.
- **Conjugation** in *Paramecium* is a sexual process but not a reproductive one. **Review Figure 20.18**
- **Alternation of generations**, which includes a multicellular diploid phase and a multicellular haploid phase, is a feature of the life cycle of many multicellular protists (as well as of some fungi and all land plants). The alternating generations may be **heteromorphic** or **isomorphic**.
- In alternation of generations, specialized cells of the diploid organism, called **sporocytes**, divide meiotically to produce haploid spores. Spores give rise to the multicellular, haploid, gamete-producing generation through mitosis. Gametes fuse and give rise to the diploid organism.

CONCEPT 20.4 Protists Are Critical Components of Many Ecosystems

- The diatoms are responsible for about one-fifth of the photosynthetic carbon fixation on Earth. They and other members of the phytoplankton are the **primary producers** in the marine environment.
- Endosymbiotic relationships are common among microbial protists and typically benefit both the endosymbionts and their protist or animal partners. **Review Figure 20.19**
- Some protists are pathogens of humans and other vertebrates. **Review ANIMATED TUTORIAL 20.3**
- Ancient diatoms are the major source of today's petroleum and natural gas deposits.

Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is20

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

21

The Evolution of Plants

KEY CONCEPTS

On examining a specimen of the orchid *Angraecum sesquipedale*, Charles Darwin predicted the existence of a pollinator with an exceptionally long proboscis. This pollinator, the sphinx moth *Xanthopan morgani*, was not discovered until after Darwin's death.

In the early 1860s, while the United States was entangled in its tragic civil war, much of middle- and upper-class England was caught up in an orchid frenzy. Amateur plant breeders and professional botanists alike were enchanted with raising the beautiful flowers. After *On the Origin of Species* appeared in 1859, Charles Darwin wrote his next book on this group of plants, publishing *Fertilisation of Orchids* in 1862.

There are more than 25,000 species of orchids, which makes them one of the most diverse plant groups. Darwin wanted to know why orchids had experienced such rapid diversification and was particularly impressed with the role that insect pollinators might have played in this process. He wanted examples to demonstrate the power of natural selection; he found such examples in abundance among the orchids.

Orchids show an impressive variety of specialized pollination mechanisms, many of which demonstrate that they have coevolved with their pollinators. For example, Darwin observed a South American orchid of the genus *Catasetum* shooting a packet of pollen at an insect that landed on its flower. When he was shown *Angraecum sesquipedale*, an orchid from Madagascar with a nectar tube over a foot long, Darwin hypothesized that there must be a moth with a proboscis of unprecedented length that fed from and pollinated that flower. Many people scoffed at his vision, but the moth he described was eventually discovered—21 years after his death.

In 1836 the explorer Robert Schomburgk shook the botanical world with a report that he had seen flowers described as belonging to three different genera of orchids—*Catasetum*, *Monachanthus*, and *Myanthus*—growing together on a single plant. The English botanist John Lindley remarked that this observation would "shake to the foundation all our ideas of the stability of genera and species." Orchid enthusiasts were befuddled by their efforts to grow specimens of *Myanthus*, only to have them flower with the more common blooms of *Catasetum*. Darwin knew that he needed to find the explanation for these odd observations, for otherwise he would have to conclude that individual plants were able to change their specific identity, something that did not fit with his explanations of the evolution of diversity.

Q What was Darwin's explanation for the three distinct flowers growing on a single orchid plant?

You will find the answer to this question on page 447.

CONCEPT 21.1 Primary Endosymbiosis Produced the First Photosynthetic Eukaryotes

More than a billion years ago, when a cyanobacterium was first engulfed by an early eukaryote, the history of life was altered radically. The chloroplasts that resulted from primary endosymbiosis of this cyanobacterium (see Figure 20.2) were obviously important for the evolution of plants and other photosynthetic eukaryotes, but they were also critical to the evolution of all life on land. Until photosynthetic plants were able to move onto land, there was very little on land to support multicellular animals or fungi, and almost all life was restricted to the oceans and fresh waters.

Primary endosymbiosis is a shared derived trait—a synapomorphy—of the group known as **Plantae** (FIGURE 21.1). Although *Plantae* is Latin for "plants," in everyday language—and throughout this book—the unmodified common name "plants" is usually used to refer only to the land plants. However, the first several clades to branch off the tree of life after primary endosymbiosis are all aquatic. Most aquatic photosynthetic eukaryotes (other than those secondarily derived from land plants) are known by the common name **algae**. This name, however, is just a convenient way to refer to all such groups, which are not all closely related.

Several distinct clades of algae were among the first photosynthetic eukaryotes

The ancestor of Plantae was unicellular and may have been similar in general form to the modern **glaucophytes** (FIGURE 21.2A). These microscopic freshwater algae are thought to be the sister group to the rest of Plantae (see Figure 21.1A). The chloroplast of glaucophytes is unique in containing a small amount of peptidoglycan between its inner and outer membranes—the same arrangement found in cyanobacteria. Peptidoglycan has been lost from the remaining photosynthetic eukaryotes.

In contrast to the glaucophytes, almost all **red algae** are multicellular (FIGURE 21.2B). Their characteristic color is a result of the accessory photosynthetic pigment **phycoerythrin**, which is found in relatively large amounts in the chloroplasts of many red algae. In addition to phycoerythrin, red algal chloroplasts contain chlorophyll *a* as well as several other accessory pigments.

The red algae include species that grow in the shallowest tide pools as well as the photosynthesizers found deepest in the ocean (as deep as 260 meters if nutrient conditions are right and the water is clear enough to permit light to penetrate). A few red algae inhabit fresh water. Most grow attached to a substrate by a structure known as a holdfast.

Despite their name, red algae don't always appear red in color. The ratio of two pigments—phycoerythrin (red) and chlorophyll *a* (green)—depends largely on the intensity of light that reaches the alga. In deep water, where light is dim, algae accumulate large amounts of phycoerythrin (which absorbs light at short wavelengths) and appear red. But many species growing near the surface contain a higher concentration of chlorophyll *a* and thus appear bright green.

The remaining algal groups in Plantae are the various "green algae." Like land plants, the green algae contain both chlorophylls *a* and *b* and store their reserve of photosynthetic products as starch in chloroplasts. All the groups that share these features are commonly called **green plants** because both of their photosynthetic pigments are green.

Three important clades of green algae are the **chlorophytes**, **coleochaetophytes**, and **stoneworts** (FIGURE 21.3). The chlorophytes are the sister group of all the other green plants, which are collectively called **streptophytes**. Among the streptophytes, the coleochaetophytes and stoneworts retain their eggs in the parental organism, as do land plants. They also share other cellular features with the land plants (see Figure 21.1A). Of these two groups, the stoneworts are thought to be the closest relatives of the land plants, based largely on similarities of their genes. The stoneworts also exhibit the branched apical growth that is typical of many land plants (see Figure 21.3C).

There are ten major groups of land plants

One of the key synapomorphies of the **land plants** is development from an embryo that is protected by tissues of the parent plant. For this reason, land plants are sometimes called **embryophytes** (*phyton*, "plant"). The green plants, the streptophytes, and the land plants each have been called "the plant kingdom" by different authorities, and others take an even broader view to include red algae and glaucophytes as "plants." To avoid confusion in this chapter, we will use modifying terms ("land plants" or "green plants," for example) to refer to the various clades of Plantae shown in Figure 21.1.

The land plants that exist today fall naturally into ten major clades (TABLE 21.1). Members of seven of those clades possess well-developed vascular systems that transport materials throughout the plant body. We call these seven groups, collectively, the **vascular plants**, or **tracheophytes**, because they all possess fluid-conducting cells called **tracheids**. The remaining three clades (liverworts, hornworts, and mosses) lack tracheids and are referred to collectively as **nonvascular land plants**. Note, however, that the three groups of nonvascular land plants do *not* form a clade (unlike the vascular plants, which *are* a clade).

CHECKpoint CONCEPT 21.1

- ✓ Explain the different possible uses of the term "plant."
- ✓ What are some of the key differences between glaucophytes, red algae, and the various clades of green algae?
- ✓ What evidence supports the phylogenetic relationship between land plants and green algae? Why doesn't the name "algae" designate a formal taxonomic group?

The green algal ancestors of the land plants lived at the margins of ponds or marshes, ringing them with a mat of dense green. It was from such a marginal habitat, which was sometimes wet and sometimes dry, that early plants made the transition onto land.

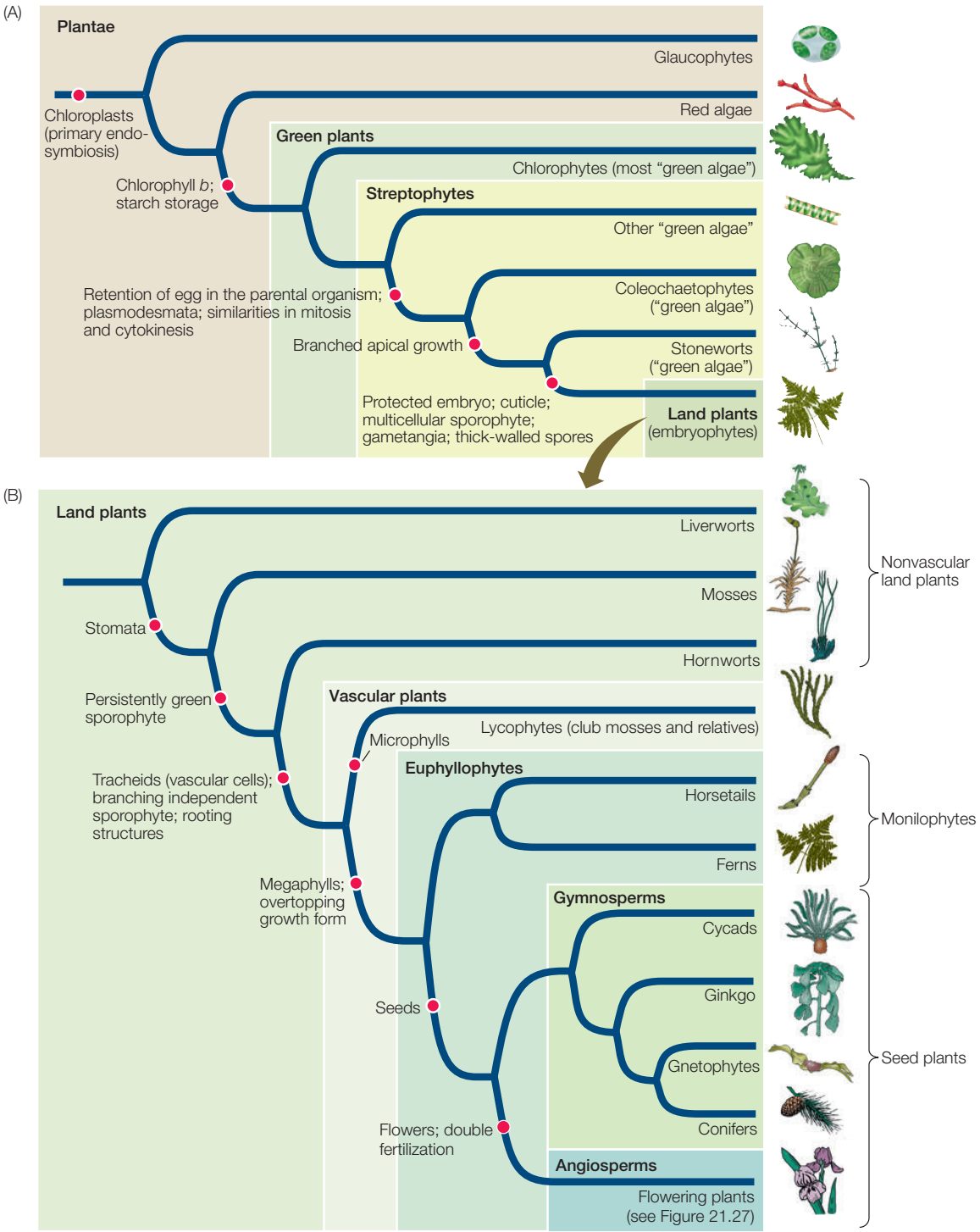
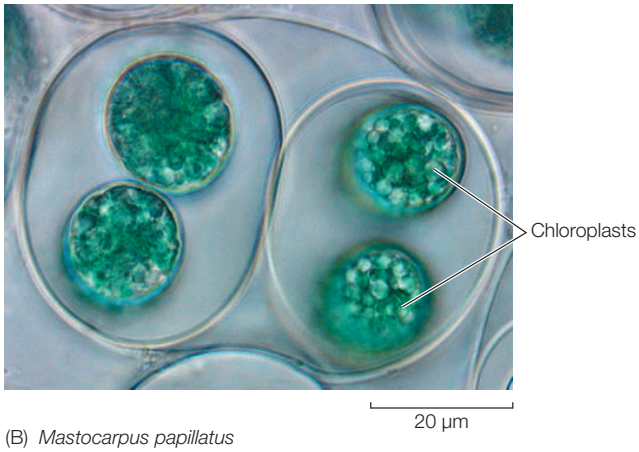


FIGURE 21.1 The Evolution of Plants (A) In its broadest definition, the term "plant" includes the glaucophytes, red algae, and green plants—all the groups descended from a common ancestor with chloroplasts. (B) Some biologists restrict the term "plant" to the green plants (those with chlorophyll *b*) or, even more narrowly, to the land plants. Three key characteristics that emerged during the evolution of land plants—protected embryos, vascular tissues, and seeds—led to their success in the terrestrial environment. (See Analyze the Data 21.1 at LaunchPad.)

(A) *Glaucocystis* sp.



(B) *Mastocarpus papillatus*

FIGURE 21.2 Early Branching Groups of Plantae (A) The large chloroplasts of unicellular glaucophytes differ from chloroplasts of other Plantae in retaining a layer of peptidoglycan. This feature is thought to have been retained from the endosymbiotic cyanobacteria that gave rise to the chloroplasts of Plantae. This photograph shows a colony of two individuals, each with two chloroplasts. (B) The photosynthetic pigment phycoerythrin gives this red alga its rich hue, showing up vividly against a background of sand.

CONCEPT 21.2 Key Adaptations Permitted Plants to Colonize Land

How did the land plants arise? To address this question, we can compare land plants with their closest relatives among the green algae. The features that differ between the two groups include the adaptations that allowed the first land plants to survive in the terrestrial environment.

Adaptations to life on land distinguish land plants from green algae

Land plants first appeared in the terrestrial environment between 450 and 500 million years ago. How did they survive in an environment that differed so dramatically from the aquatic environment of their ancestors? While the water essential for

(A) *Ulva latuca*

(B) *Coleochaete* sp.

(C) *Chara vulgaris*

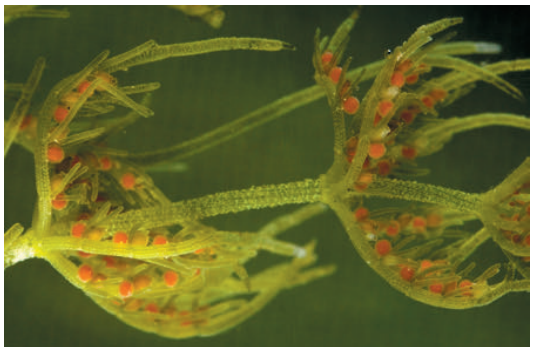


FIGURE 21.3 "Green Algae" Consist of Several Distantly Related Groups (A) Sea lettuce, a chlorophyte, grows in ocean tidewaters. (B) The coleochaetophytes are thought to be the sister group of stoneworts plus land plants. (C) The land plants probably evolved from a common ancestor shared with stoneworts, which display the branching pattern we otherwise associate with land plants. A stonewort in the common freshwater genus *Chara* is shown here. The orange structures are male sex organs.

Go to MEDIA CLIP 21.1
Reproductive Structures of *Chara*
PoL2e.com/mc21.1

TABLE 21.1 Classification of Land Plants

Group	Common name	Characteristics
NONVASCULAR LAND PLANTS		
Hepatophyta	Liverworts	No filamentous stage; gametophyte flat
Anthocerophyta	Hornworts	Embedded archegonia; sporophyte grows basally (from the ground)
Bryophyta	Mosses	Filamentous stage; sporophyte grows apically (from the tip)
VASCULAR PLANTS		
Lycopodiophyta	Lycophytes: Club mosses and allies	Microphylls in spirals; sporangia in leaf axils
Monilophyta	Horsetails, ferns	Megaphylls with a leaf gap (a space in the stem from which the leaf emerges)
SEED PLANTS		
Gymnosperms		
Cycadophyta	Cycads	Compound leaves; swimming sperm; seeds on modified leaves
Ginkgophyta	Ginkgo	Deciduous; fan-shaped leaves; swimming sperm
Gnetophyta	Gnetophytes	Vessels in vascular tissue; opposite, simple leaves
Coniferophyta	Conifers	Seeds in cones; needlelike or scalelike leaves
Angiosperms	Flowering plants	Endosperm; carpels; gametophytes much reduced; seeds within fruit

life is everywhere in the aquatic environment, water is difficult to obtain and retain in the terrestrial environment.

No longer bathed in fluid, organisms on land faced potentially lethal desiccation (drying). Large terrestrial organisms had to develop ways to transport water to body parts distant from the source of the water. And whereas water provides aquatic organisms with support against gravity, a plant living on land must either have some other support system or sprawl unsupported on the ground. A land plant must also use different mechanisms for dispersing its gametes and progeny than its aquatic relatives, which can simply release them into the water.

Survival on land was facilitated by the evolution among plants of numerous adaptations, including:

- The *cuticle*, a coating of waxy lipids that retards water loss
- *Stomata*, small closable openings in leaves and stems that are used to regulate gas exchange and water loss
- *Gametangia*, multicellular organs that enclose plant gametes and prevent them from drying out
- *Embryos*, young plants contained within a protective structure
- Certain *pigments* that afford protection against the mutagenic ultraviolet radiation that bathes the terrestrial environment
- Thick *spore walls* containing a *polymer* that protects the spores from desiccation and resists decay
- A *mutually beneficial association with fungi* (mycorrhizae) that promotes nutrient uptake from the soil

The cuticle may be the most important—and the earliest—of these features. Composed of several unique waxy lipids that coat the leaves and stems of land plants, the cuticle has several functions, the most obvious and important of which is to keep water from evaporating from the plant body.

As ancient plants colonized the land, they not only adapted to the terrestrial environment, they also modified it by contributing to the formation of soil. Acids secreted by plants help break down rock, and the organic compounds produced by the breakdown of dead plants contribute nutrients to the soil. Such effects are repeated today as plants grow in new areas.

Life cycles of land plants feature alternation of generations

A universal feature of the life cycles of land plants is alternation of generations. Recall from Concept 20.3 the two hallmarks of alternation of generations:

- The life cycle includes both a multicellular diploid stage and a multicellular haploid stage.
- Gametes are produced by mitosis, not by meiosis. Meiosis produces **spores** that develop into multicellular haploid organisms.

If we begin looking at the land plant life cycle at the single-cell stage—the diploid zygote—then the first phase of the cycle is the formation, by mitosis and cytokinesis, of a multicellular **embryo**, which eventually grows into a mature diploid plant. This multicellular diploid plant is called the **sporophyte** ("spore plant").

Cells contained within specialized reproductive organs of the sporophyte, called **sporangia** (singular *sporangium*), undergo meiosis to produce haploid, unicellular spores. By mitosis and cytokinesis, a spore develops into a haploid plant. This multicellular haploid plant, called the **gametophyte** ("gamete plant"), produces haploid gametes by mitosis. The fusion of two gametes (fertilization) forms a single diploid cell—the zygote—and the cycle is repeated (FIGURE 21.4).

The sporophyte generation extends from the zygote through the adult multicellular diploid plant and sporangium formation. In contrast, the gametophyte generation extends from the

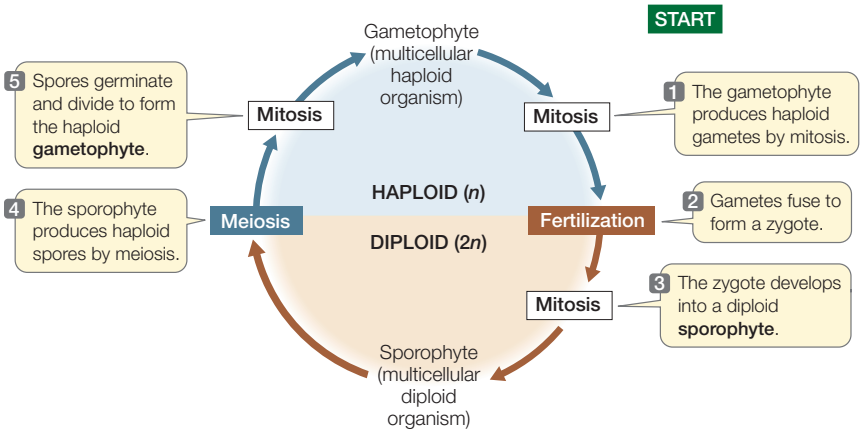


FIGURE 21.4 Alternation of Generations in Land Plants A multicellular diploid sporophyte generation that produces spores by meiosis alternates with a multicellular haploid gametophyte generation that produces gametes by mitosis.

spore through the adult multicellular haploid plant to the gametes. The transitions between the generations are accomplished by fertilization and by meiosis. In all land plants, the sporophyte and the gametophyte differ genetically: the sporophyte has diploid cells, and the gametophyte has haploid cells.

There is a trend toward reduction of the gametophyte generation in plant evolution. In the nonvascular land plants, the gametophyte is larger, longer-lived, and more self-sufficient than the sporophyte. In those groups that appeared later in plant evolution, however, the sporophyte is the larger, more conspicuous, longer-lived, and more self-sufficient generation.

Nonvascular land plants live where water is readily available

The living species of nonvascular land plants are the liverworts, mosses, and hornworts. These three groups are thought to be similar in many ways to the earliest land plants. Most of these plants grow in dense mats, usually in moist habitats. Even the largest of these species are only about half a meter tall, and most are only a few centimeters tall or long. Why have they not evolved to be taller? The probable answer is that they lack an efficient vascular system for transporting water and minerals from the soil to distant parts of the plant body.

The nonvascular land plants lack the true leaves, stems, and roots that characterize the vascular plants, although they have structures analogous to each. Their growth form allows water to move through the mats of plants by capillary action. They have leaflike structures that readily catch and hold any water that splashes onto them. They are small enough that minerals can be distributed throughout their bodies by diffusion. As in all land plants, layers of maternal tissue protect their embryos from desiccation. Nonvascular land plants also have a cuticle, although it is often very thin (or even absent in some species) and thus is not highly effective in retarding water loss.

Most nonvascular land plants live on the soil or on vascular plants, but some grow on bare rock, on dead and fallen tree trunks, and even on buildings. Their ability to grow on such marginal surfaces results from a mutualistic association with fungi. The earliest association of land plants with fungi dates back at least 460 million years. This mutualism probably facilitated the absorption of water and minerals, especially phosphorus, from the first soils.

LINK

Land plants of many groups have mutualistic associations with fungi, as described in **Concept 22.2**

Nonvascular land plants are widely distributed over six continents and even exist (albeit very locally) on the coast of the seventh, Antarctica. Most are terrestrial. Although a few species live in fresh water, these aquatic species are descended from terrestrial ones. None live in the oceans.

LIVERWORTS There are about 9,000 species of **liverworts**. Most liverworts have green, leaflike gametophytes that lie close to or flat on the ground (**FIGURE 21.5A**). The simplest liverworts are flat plates of cells a centimeter or so long with structures that produce sperm or eggs on their upper surfaces and rootlike filaments on their lower surfaces. The sporophyte remains attached to the larger gametophyte and rarely exceeds a few millimeters in length. Most liverworts can reproduce asexually (through simple division of the gametophyte) as well as sexually.

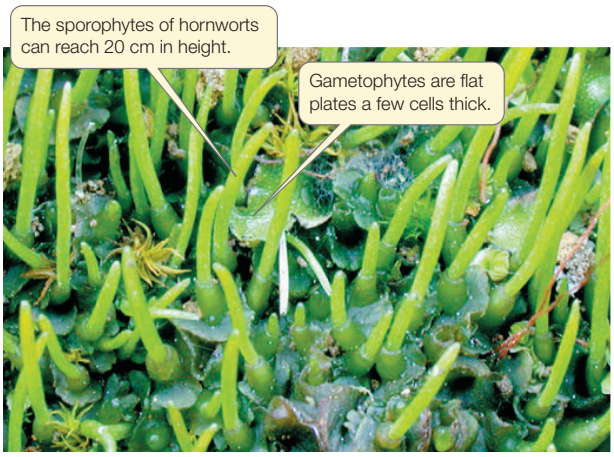
Go to MEDIA CLIP 21.2
Liverwort Life Cycle
PoL2e.com/mc21.2

MOSSES The most familiar of the nonvascular land plants are the **mosses** (**FIGURE 21.5B**). These hardy little plants, of which there are about 15,000 species, are found in almost every terrestrial environment. They are often found on damp, cool ground, where they form thick mats.

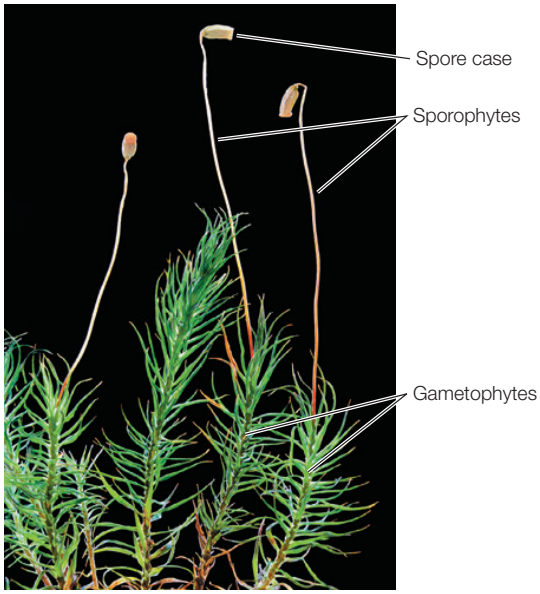
The mosses are the sister lineage to the vascular plants plus the hornworts (see Figure 21.1B). They share with those lineages an advance over the liverworts in their adaptation to life on land: they have stomata, which are important for both water retention and gas exchange.

Some moss gametophytes are so large that they cannot transport enough water throughout their bodies solely by diffusion. Gametophytes and sporophytes of many mosses

(A) *Marchantia polymorpha*



(C) *Anthoceros* sp.



(B) *Polytrichum commune*

FIGURE 21.5 Diversity among Nonvascular Land Plants (A) Gametophytes of a liverwort. (B) The sporophytes and gametophytes are easily distinguished in this moss. (C) The sporophytes of many hornworts resemble little horns.

contain a type of cell called a hydroid, which dies and leaves a tiny channel through which water can travel. The hydroid is functionally similar to the tracheid, the characteristic water-conducting cell of vascular plants, but it lacks the lignin and the cell-wall structure that characterize tracheids. The possession of hydroids shows that the term "nonvascular land plant" is somewhat misleading when applied to mosses. Despite their simple systems of internal transport, however, the mosses are not considered vascular plants because they lack tracheids and other vascular tissues.

HORNWORTS The approximately 100 species of **hornworts** are so named because their sporophytes often look like little horns (**FIGURE 21.5C**). Hornworts have two characteristics that distinguish them from liverworts and mosses. First, the cells of hornworts each contain a single large, platelike chloroplast, whereas the cells of the other two groups contain numerous small, lens-shaped chloroplasts. Second, of the sporophytes in all three groups, those of the hornworts come closest to being capable of growth without a set limit. Liverwort and moss sporophytes have a stalk that stops growing as the spore-producing structure matures, so elongation of the sporophyte is strictly limited. The hornwort sporophyte, however, has no stalk, and it is persistently green (a trait shared with vascular plants). A basal region of the sporangium remains capable of indefinite cell division, continuously producing new spore-bearing tissue above. The sporophytes of some hornworts growing in mild and continuously moist conditions can become as tall as 20 centimeters. Eventually, however, the sporophyte's growth is limited by the lack of a transport system.

Hornworts have evolved a symbiotic relationship that promotes their growth by providing them with greater access to nitrogen, which is often a limiting resource. Hornworts have internal cavities filled with mucilage, and the cavities are often populated by symbiotic cyanobacteria that convert atmospheric nitrogen gas into a form of nitrogen usable by their host plant.

The sporophytes of nonvascular land plants are dependent on the gametophytes

In the nonvascular land plants, the conspicuous green structure visible to the naked eye is the gametophyte. The gametophyte is photosynthetic and is therefore nutritionally independent. The sporophyte may or may not be photosynthetic, but it is always nutritionally dependent on the gametophyte and remains permanently attached to it.

FIGURE 21.6 illustrates a moss life cycle that is typical of the life cycles of nonvascular land plants. A sporophyte produces unicellular haploid spores as products of meiosis within a sporangium. When a spore germinates, it gives rise to a multicellular haploid gametophyte whose cells contain chloroplasts and are thus photosynthetic. Eventually gametes form within specialized

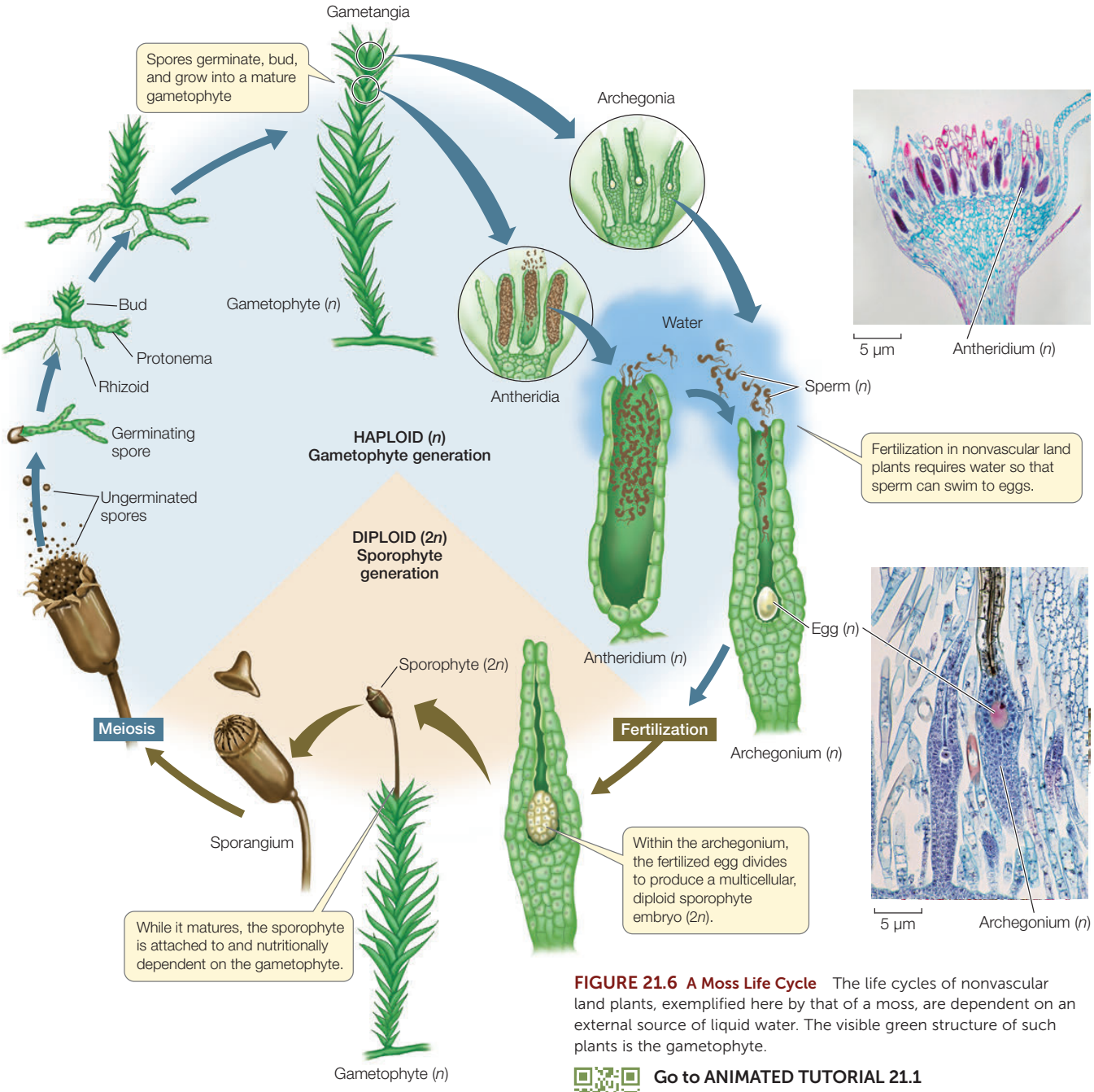


FIGURE 21.6 A Moss Life Cycle The life cycles of nonvascular land plants, exemplified here by that of a moss, are dependent on an external source of liquid water. The visible green structure of such plants is the gametophyte.

Go to ANIMATED TUTORIAL 21.1
Life Cycle of a Moss
PoL2e.com/at21.1

Go to MEDIA CLIP 21.3
Bryophyte Reproduction
PoL2e.com/mc21.3

sex organs, called the **gametangia**. The **archegonium** is a multicellular, flask-shaped female sex organ that produces a single egg. The **antheridium** is a male sex organ in which sperm, each bearing two flagella, are produced in large numbers. Both archegonia and antheridia are produced on the same individual, so each individual has both male and female reproductive structures. Adjacent individuals often fertilize one another's gametes, however, which helps maintain genetic diversity in the population.

Once released from the antheridium, the sperm must swim or be splashed by raindrops to a nearby archegonium on the same or a neighboring plant—a constraint that reflects the aquatic origins of the nonvascular land plants' ancestors. The sperm are aided on their journey by chemical attractants released by the egg or the archegonium. Before sperm can enter the archegonium, however, certain cells in the neck of the archegonium must break down, leaving a water-filled canal through which the sperm can swim to complete their journey. Notice that *all of these events require liquid water*.

Once sperm arrive at an egg, the nucleus of a sperm fuses with the egg nucleus to form a diploid zygote. Mitotic divisions of the zygote produce a multicellular, diploid sporophyte embryo. The sporophyte matures and produces a sporangium, within which meiotic divisions produce spores and thus the next gametophyte generation.

CHECKpoint CONCEPT 21.2

- ✓ Describe several adaptations of plants to the terrestrial environment, and describe the distribution of those adaptations in liverworts, mosses, and hornworts.
- ✓ Explain what is meant by alternation of generations.
- ✓ What aspects of the reproductive cycle of nonvascular land plants appear to have been retained from their aquatic ancestors?

Further adaptations to the terrestrial environment appeared as plants continued to evolve. One of the most important of these later adaptations was the appearance of vascular tissues—the characteristic that defines the vascular plants.

CONCEPT 21.3 Vascular Tissues Led to Rapid Diversification of Land Plants

The first plants possessing vascular tissues did not arise until tens of millions of years after the earliest plants had colonized the land. But once vascular tissues arose, their ability to transport water and food throughout the plant body allowed vascular plants to spread to new terrestrial environments and to diversify rapidly.

Vascular tissues transport water and dissolved materials

Vascular plants differ from the other land plants in crucial ways, one of which is the possession of a well-developed vascular system consisting of tissues that are specialized for the transport of materials from one part of the plant to another. One type of vascular tissue, the **xylem**, conducts water and minerals from the soil to aerial parts of the plant. Because some of its cell walls contain a stiffening substance called lignin, xylem also provides support against gravity in the terrestrial environment. The other type of vascular tissue, the **phloem**, conducts the products of photosynthesis from sites where they are produced or released to sites where they are used or stored.

LINK

The vascular tissues of plants are described in detail in Chapter 24

Although the vascular plants are an extraordinarily large and diverse group, a particular event was critical to their evolution. Sometime during the Paleozoic era, probably in the mid-Silurian (430 mya), a new cell type—the tracheid—evolved in sporophytes of the earliest vascular plants. The tracheid is the principal water-conducting element of the xylem in all vascular plants except the angiosperms (flowering plants) and one small group of gymnosperms—and tracheids persist even in these groups, along with a more specialized and efficient system derived from them.

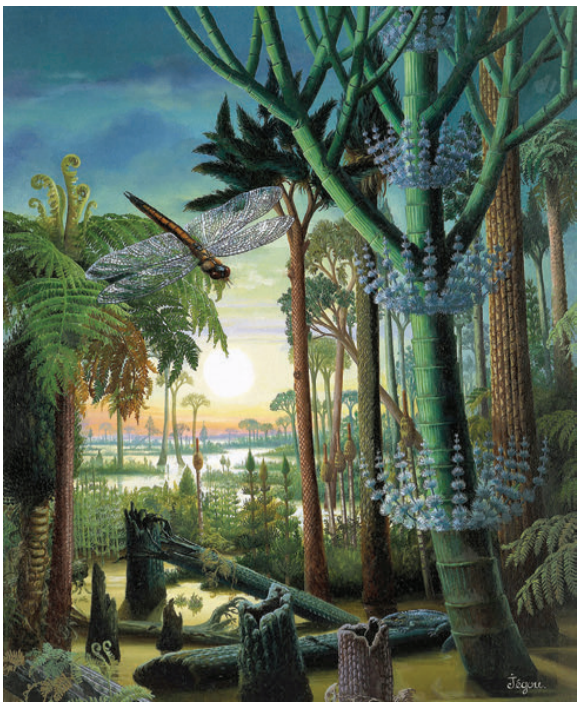


FIGURE 21.7 Reconstruction of an Ancient Forest Forests of the Carboniferous period were characterized by abundant vascular plants such as club mosses, ferns, and horsetails, some of which reached heights of 40 meters. Huge flying insects (see p. 357) thrived in these forests, which are the source of modern coal deposits.

The evolution of tracheids set the stage for the complete and permanent invasion of land by plants. First, these cells provided a pathway for the transport of water and mineral nutrients from a source of supply to regions of need in the plant body. And second, the cell walls of tracheids, stiffened by lignin, provided rigid structural support. This support is a crucial factor in a terrestrial environment because it allows plants to grow upward and thus compete for sunlight. A taller plant can intercept more direct sunlight (and thus conduct photosynthesis more readily) than a shorter plant, which may be shaded by the taller one. Increased height also improves the dispersal of spores.

The vascular plants featured another evolutionary novelty: a branching, independent sporophyte. A branching sporophyte body can produce more spores than an unbranched body, and it can develop in complex ways. The sporophyte of a vascular plant is nutritionally independent of the gametophyte at maturity. Among the vascular plants, the sporophyte is the large and obvious plant that one normally pays attention to in nature, in contrast to the relatively small, dependent sporophytes typical of most nonvascular land plants.

The diversification of vascular plants made land more suitable for animals

The initial absence of herbivores (plant-eating animals) on land helped make the first vascular plants successful. By the late

(A) *Lycopodium annotinum*

(B) *Equisetum pratense*

(C) *Marsilea mutica*

(D) *Alsophilia spinulosa*

FIGURE 21.8 Lycophytes and Monilophytes (A) A strobilus is visible at the tip of this club moss. Club mosses have microphylls arranged spirally on their stems. (B) Horsetails have a distinctive growth pattern in which the stem grows in segments above each whorl of leaves. These are fertile shoots with sporangia-bearing structures at the apex. (C) The leaves of a species of water fern. (D) A tree fern on a mountain in India.

Silurian period (about 425 mya), the proliferation of land plants made the terrestrial environment more hospitable to animals. Arthropods, vertebrates, and other animals moved onto land only after vascular plants became established there.

Trees of various kinds appeared in the Devonian period and dominated the landscape of the Carboniferous period (359–299 mya). Forests of lycophytes (club mosses) up to 40 meters tall, along with horsetails and tree ferns, flourished in the tropical swamps of what would become North America and Europe (**FIGURE 21.7**). Plant parts from those forests sank into the swamps and were gradually covered by layers of sediment. Over millions of years, as the buried plant material was subjected to intense pressure and elevated temperatures, it was transformed into coal. Today that coal provides more than half of our electricity. The world's coal deposits, although huge, are not infinite, and humans are burning coal deposits at a far faster rate than they were produced.

In the subsequent Permian period, when the continents came together to form Pangaea, the continental interior became warmer and drier. The 200-million-year reign of the lycophyte–fern forests came to an end as they were replaced by forests of early gymnosperms.

The earliest vascular plants lacked roots

The earliest known vascular plants belonged to a now-extinct group called the **rhyniophytes**. The rhyniophytes were one of a very few types of vascular plants in the Silurian period. The landscape at that time probably consisted mostly of bare ground, with stands of rhyniophytes in low-lying moist areas. Early versions of the structural features of all the other vascular plant groups appeared in the rhyniophytes of that time. These shared features strengthen the case for the origin of all vascular plants from a common nonvascular land plant ancestor.

Rhyniophytes did not have roots. Like most modern ferns and lycophytes, they were apparently anchored in the soil by horizontal portions of stem, called **rhizomes**, which bore water-absorbing unicellular filaments called **rhizoids**. These plants also bore aerial branches, and sporangia—homologous to the sporangia of mosses—were found at the tips of those branches. Their branching pattern was **dichotomous**, meaning the apex (tip) of the shoot divided to produce two equivalent new branches, each pair diverging at approximately the same angle from the original stem.

The lycophytes are sister to the other vascular plants

The club mosses and their relatives, the spike mosses and quillworts, are collectively called **lycophytes**. The lycophytes are the sister group to the remaining vascular plants (see Figure 21.1B). There are relatively few (just over 1,200) surviving species of lycophytes.

The lycophytes have true roots that branch dichotomously. The arrangement of vascular tissue in their stems is simpler than that in other vascular plants. They bear simple leaflike structures called **microphylls**, which are arranged spirally on the stem. Growth in lycophytes comes entirely from apical cell division. Branching in the stems, which is also dichotomous, occurs by division of an apical cluster of dividing cells.

The sporangia of many club mosses are aggregated in cone-like structures called **strobili** (singular *strobilus*; **FIGURE 21.8A**). The strobilus of a club moss is a cluster of spore-bearing leaves inserted on an axis (a linear supporting structure). Other club mosses lack strobili and bear their sporangia on (or adjacent to) the upper surfaces of specialized leaves.

Horsetails and ferns constitute a clade

The horsetails and ferns were once thought to be only distantly related. As a result of genomic analyses, we now know that

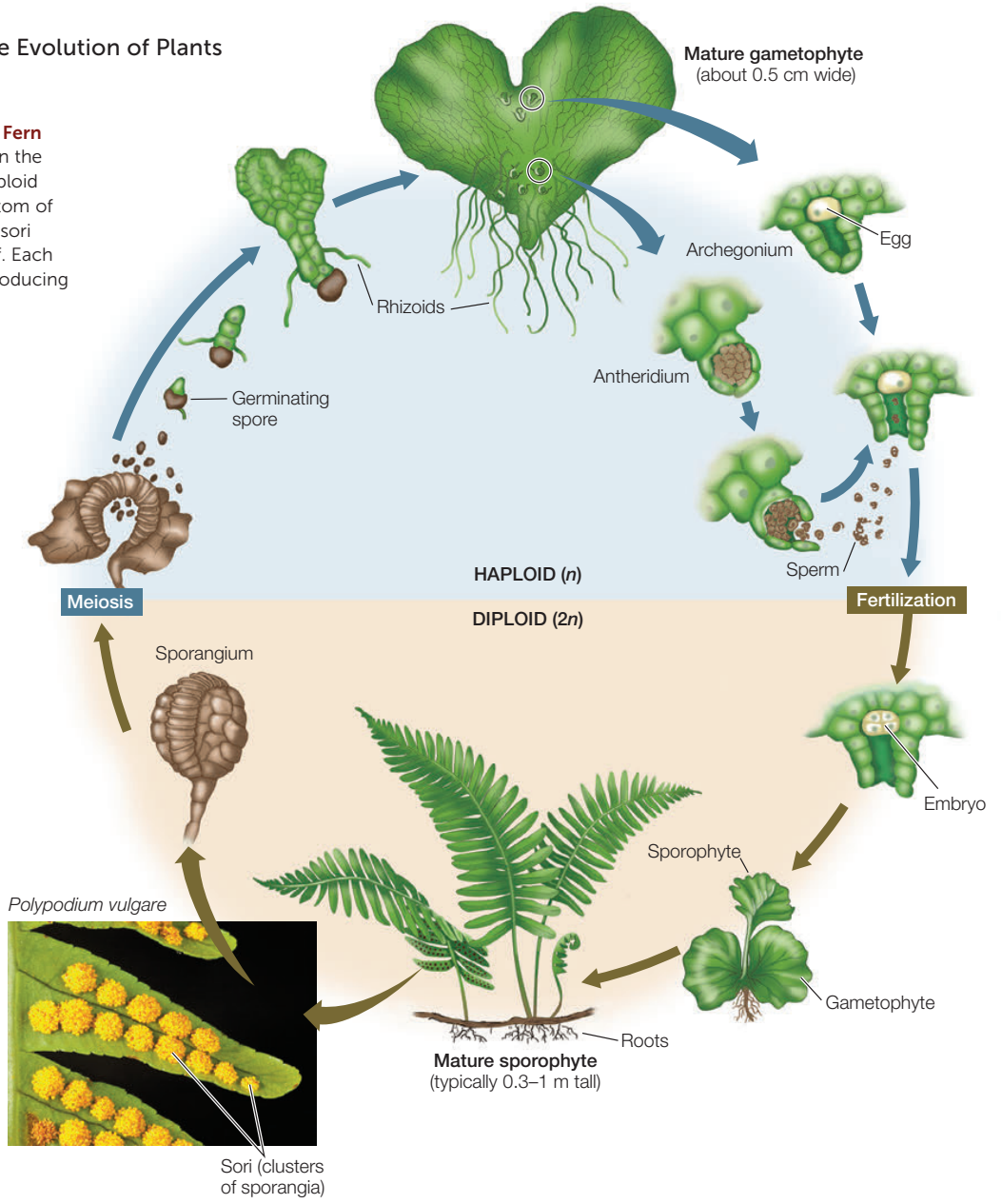


FIGURE 21.9 Life Cycle of a Fern
The most conspicuous stage in the fern life cycle is the mature diploid sporophyte, shown at the bottom of this diagram. The inset shows sori on the underside of a fern leaf. Each sorus contains many spore-producing sporangia.

Go to ACTIVITY 21.1
The Fern Life Cycle
PoL2e.com/ac21.1

they form a clade: the **monilophytes**. In the monilophytes—as in the seed plants, to which they are the sister group (see Figure 21.1B)—there is differentiation between a main stem and side branches. This pattern contrasts with the dichotomous branching characteristic of the lycophytes and rhyniophytes, in which each split gives rise to two branches of similar size.

Today there are only about 15 species of **horsetails**, all in the genus *Equisetum*. The horsetails have reduced true leaves that form in distinct whorls (circles) around the stem (**FIGURE 21.8B**). Horsetails are sometimes called "scouring rushes" because rough silica deposits found in their cell walls once made them useful for cleaning. They have true roots that branch irregularly. Horsetails have a large sporophyte and a small gametophyte, each independent of the other.

The first **ferns** appeared during the Devonian period. Today this group comprises more than 12,000 species. Analyses of gene sequences indicate that a few species traditionally allied with ferns may in fact be more closely related to horsetails than to ferns. Nonetheless, the majority of ferns form a monophyletic group.

Although most ferns are terrestrial, a few species live in shallow fresh water (**FIGURE 21.8C**). Terrestrial ferns are characterized by large leaves with branching vascular strands (**FIGURE 21.8D**). Some fern leaves become climbing organs and may grow to be as long as 30 meters.

In the alternating generations of a fern, the gametophyte is small, delicate, and short-lived, but the sporophyte can be very large and can sometimes survive for hundreds of years (**FIGURE 21.9**). Ferns require liquid water for the transport of the male gametes to the female gametes, so most ferns inhabit shaded, moist woodlands and swamps. The sporangia of ferns typically are borne on a stalk in clusters called **sori** (singular

sorus). The sori are found on the undersurfaces of the leaves, sometimes covering the entire undersurface and sometimes located at the edges.

The vascular plants branched out

Several features that were new to the vascular plants evolved in lycophytes and monilophytes. Roots probably had their evolutionary origins as a branch, either of a rhizome or of the aboveground portion of a stem. That branch presumably penetrated the soil and branched further. The underground portion could anchor the plant firmly, and even in this primitive condition, it could absorb water and minerals.

The microphylls of lycophytes were probably the first leaflike structures to evolve among the vascular plants. Microphylls are usually small and only rarely have more than a single vascular strand, at least in existing species. Some biologists believe that microphylls had their evolutionary origins as sterile sporangia (FIGURE 21.10A). The principal characteristic of a microphyll is a vascular strand that departs from the vascular system of the stem in such a way that the structure of the stem's vascular system is scarcely disturbed. This pattern was evident even in the lycophyte trees of the Carboniferous period, many of which had microphylls many centimeters long.

The monilophytes and seed plants constitute a clade called the **euphyllophytes** (*eu*, "true"; *phyllos*, "leaf"). An important synapomorphy of the euphyllophytes is **overtopping**, a growth pattern in which one branch differentiates from and grows beyond the others (FIGURE 21.10B). Overtopping would have given these plants an advantage in the competition for light, enabling them to shade their dichotomously branching competitors. The overtopping growth of the euphyllophytes also allowed a new type of leaflike structure to evolve. This larger, more complex leaf is called a **megaphyll**. The megaphyll is thought to have arisen from the flattening of a portion of a branching stem system that exhibited overtopping growth. This change was followed by the development of photosynthetic tissue between the members of overtopped groups of branches, which had the advantage of increasing the photosynthetic surface area of those branches.

Heterospory appeared among the vascular plants

Among the earliest vascular plants, the gametophyte and the sporophyte were independent of one another, and both were photosynthetic. The spores produced by the sporophyte were of a single type and developed into a single type of gametophyte that bore both female and male reproductive organs. This is still true in some surviving lineages of vascular plants, such as horsetails and many ferns. Such plants, which bear a single type of spore, are said to be **homosporous** (FIGURE 21.11A).

A system with two distinct types of spores evolved several times among the monilophytes and in the ancestor of seed plants. Plants of this type are said to be **heterosporous** (FIGURE 21.11B). In heterospory, one type of spore—the **megaspore**—develops into a specifically female gametophyte (a **megagametophyte**) that produces only eggs. The other type, the **microspore**, is smaller and develops into a male gametophyte (a

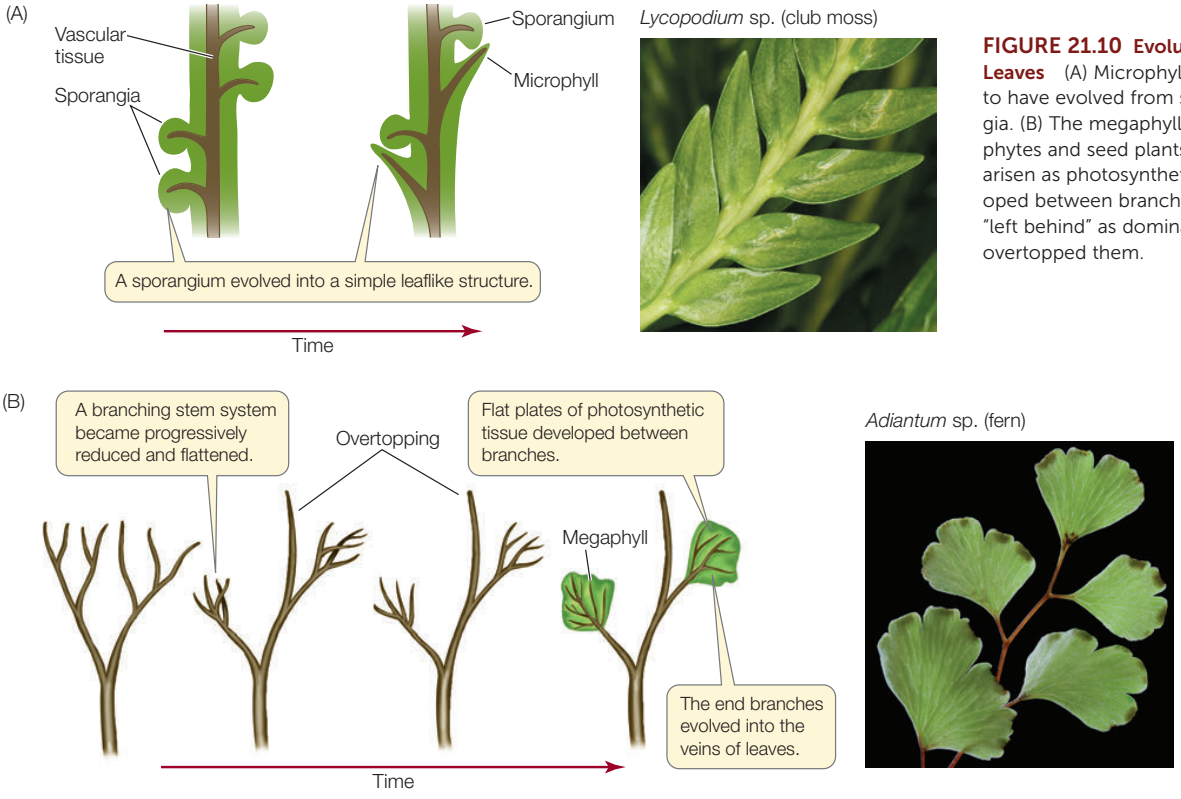


FIGURE 21.10 Evolution of Leaves (A) Microphylls are thought to have evolved from sterile sporangia. (B) The megaphylls of monilophytes and seed plants may have arisen as photosynthetic tissue developed between branch pairs that were "left behind" as dominant branches overtopped them.

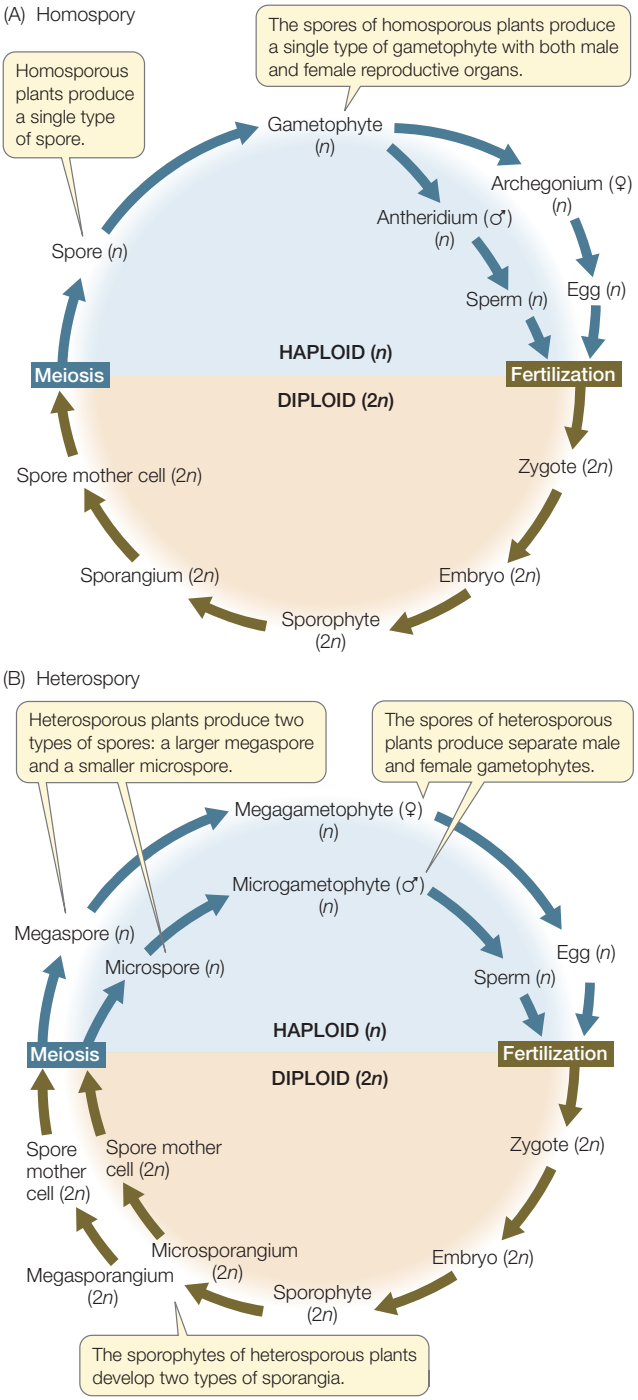


FIGURE 21.11 Homospory and Heterospory (A) Homosporous plants bear a single type of spore. Each gametophyte has two types of sex organs, antheridia (male) and archegonia (female). (B) Heterosporous plants bear two types of spores that develop into distinctly male and female gametophytes.

Go to ACTIVITY 21.2 Homospory
PoL2e.com/ac21.2

Go to ACTIVITY 21.3 Heterospory
PoL2e.com/ac21.3

microgametophyte) that produces only sperm. The sporophyte produces megaspores in small numbers in **megasporangia** and microspores in large numbers in **microsporangia**. Heterospory affects not only the spores and the gametophytes, but also the sporophyte plant itself, which must develop two types of sporangia.

The fact that heterospory evolved repeatedly suggests that it affords selective advantages. What advantages does heterospory provide? Heterospory allows the production of many small microspores, which are easily transported from plant to plant. Heterospory also results in the production of a few large megaspores in large megasporangia, which can provide nutrition and protection for the developing embryo. With this division, plants can increase the opportunities for long-distance cross-fertilization, and at the same time increase the chances for survival of the developing embryo.

CHECKpoint CONCEPT 21.3

- ✓ How do the vascular tissues xylem and phloem serve the vascular plants?
- ✓ Describe the evolution and distribution of different kinds of leaves and roots among the vascular plants.
- ✓ Explain the concept of heterospory. How does heterospory provide selective advantages over homospory?

All of the vascular plant groups we have discussed thus far disperse by means of spores. The embryos of these seedless vascular plants develop directly into sporophytes, which either survive or die, depending on environmental conditions. The spores of some seedless plants may remain dormant and viable for long periods, but the embryos of seedless plants are relatively unprotected (see Figure 21.9). Greater protection of the embryo evolved in the seed plants.

CONCEPT 21.4 Seeds Protect Plant Embryos

The **seed plants**—the gymnosperms and the angiosperms—provide a secure and lasting dormant stage for the embryo, known as a seed. In seeds, embryos may safely lie dormant, in some cases for many years (even centuries), until conditions are right for germination. How is this protection provided?

Features of the seed plant life cycle protect gametes and embryos

In Concept 21.2 we described a trend in plant evolution: the sporophyte became less dependent on the gametophyte, which became smaller in relation to the sporophyte. This trend continued with the seed plants, whose gametophyte generation is reduced even further than it is in the ferns (**FIGURE 21.12**). The haploid gametophyte develops partly or entirely while attached to and nutritionally dependent on the diploid sporophyte.

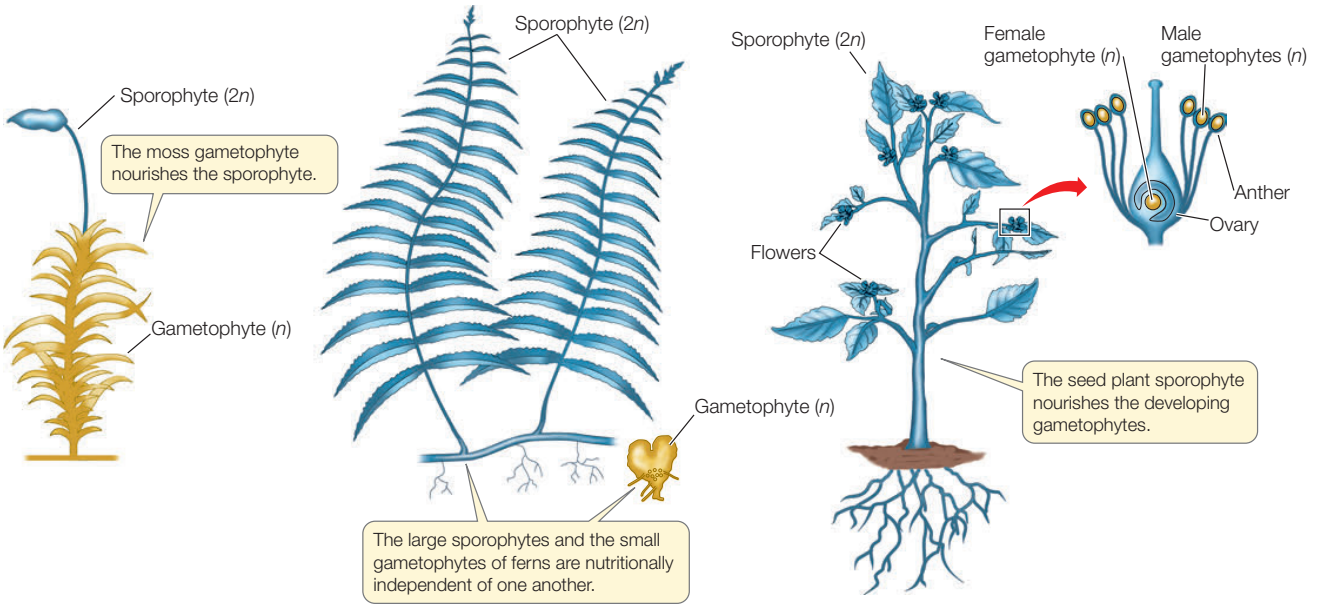


FIGURE 21.12 The Relationship between Sporophyte and Gametophyte In the course of plant evolution, the gametophyte (brown) has been reduced and the sporophyte (blue) has become more prominent.

Among the seed plants, only the earliest diverging groups of gymnosperms (including modern cycads and ginkgos) have swimming sperm. Even in these groups, sperm is transferred via pollen grains, so fertilization does not require liquid water outside the plant body. The evolution of pollen, along with the advent of seeds, gave seed plants the opportunity to colonize drier areas and spread over the terrestrial environment.

Seed plants are heterosporous (see Figure 21.11B)—that is, they produce two types of spores, one that becomes a microgametophyte (male gametophyte) and one that becomes a megagametophyte (female gametophyte). They form separate microsporangia and megasporangia on structures that are grouped on short stems, such as the stamens and carpels of an angiosperm flower.

Within the microsporangium, the meiotic products are microspores. Within its spore wall, a microspore divides mitotically one or a few times to form a multicellular male gametophyte called a **pollen grain**. Pollen grains are released from the microsporangium to be distributed by wind or by an animal pollinator (**FIGURE 21.13**). The spore wall that surrounds the pollen grain contains a substance called sporopollenin, the most chemically resistant biological compound known. Sporopollenin protects the pollen grain against dehydration and chemical damage—another advantage in terms of survival in the terrestrial environment.

In contrast to the microspores, the megaspores of seed plants are not shed. Instead they develop into female gametophytes within the megasporangia. These megagametophytes are dependent on the sporophyte for food and water.

In most seed plant species, only one of the meiotic products in a megasporangium survives. The surviving haploid nucleus divides mitotically, and the resulting cells divide again to produce a multicellular female gametophyte. The megasporangium is surrounded by sterile sporophytic structures, which form an **integument** that protects the megasporangium

FIGURE 21.13 Pollen Grains Pollen grains are the male gametophytes of seed plants. The pollen of the hazel tree is dispersed by the wind. The grains may land near female gametophytes of the same or other hazel plants.

Go to MEDIA CLIP 21.4
Pollen Transfer by Wind
PoL2e.com/mc21.4

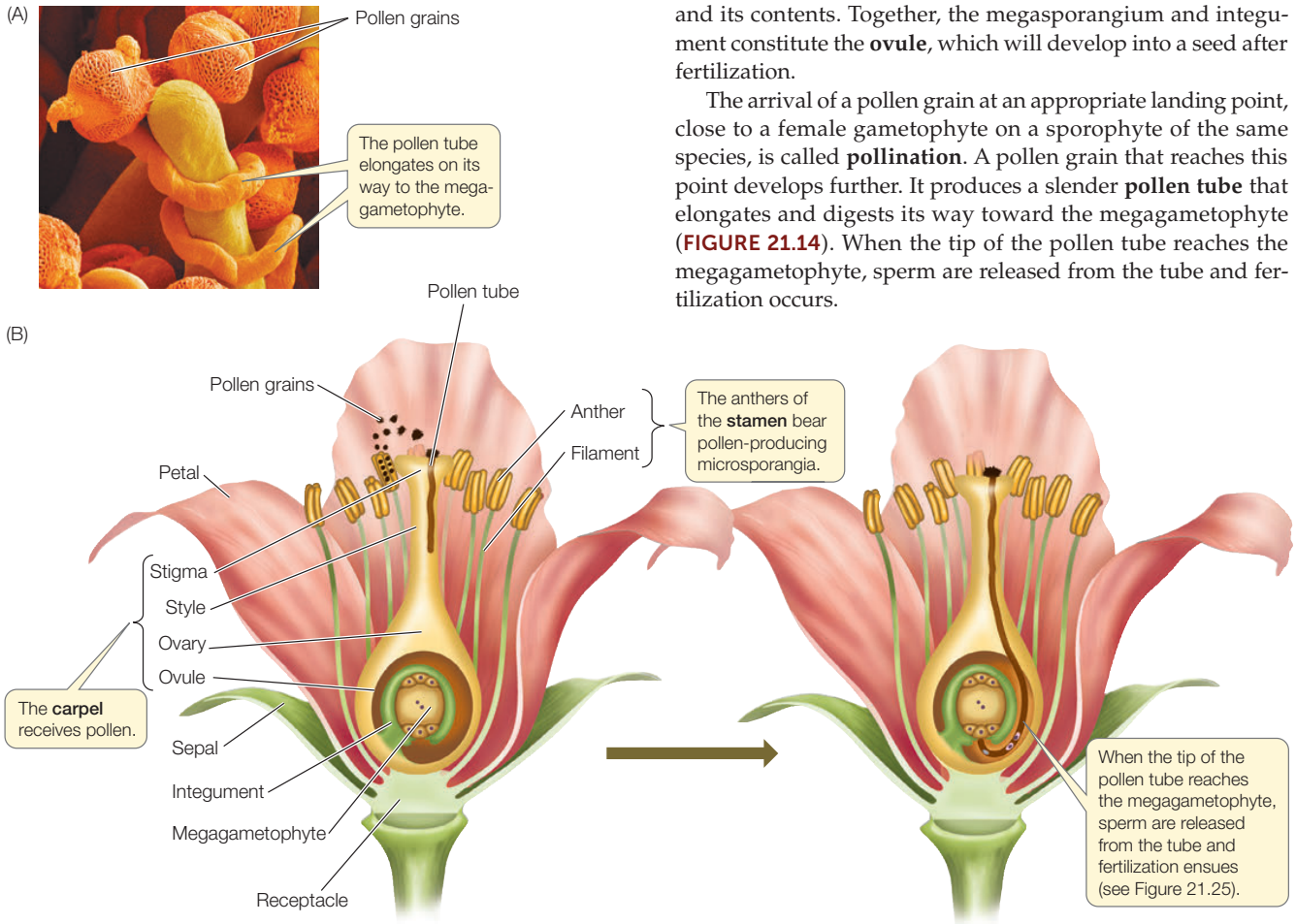


FIGURE 21.14 Pollination in Seed Plants In all seed plants, a pollen tube grows from the pollen grain to the megagametophyte, where sperm are released. (A) Scanning electron micrograph of a pollen tube growing in a prairie gentian flower. (B) The process of pollination is diagrammed for a generalized angiosperm flower.

and its contents. Together, the megasporangium and integument constitute the **ovule**, which will develop into a seed after fertilization.

The arrival of a pollen grain at an appropriate landing point, close to a female gametophyte on a sporophyte of the same species, is called **pollination**. A pollen grain that reaches this point develops further. It produces a slender **pollen tube** that elongates and digests its way toward the megagametophyte (**FIGURE 21.14**). When the tip of the pollen tube reaches the megagametophyte, sperm are released from the tube and fertilization occurs.

APPLY THE CONCEPT

Seeds protect plant embryos

In 1879 W. J. Beal began an experiment that he could not hope to finish in his lifetime. He prepared 20 lots of seeds for long-term storage. Each lot consisted of 50 seeds from each of 23 species. He mixed each lot of seeds with sand and placed the mixture in an uncapped bottle, then buried all the bottles on a sandy knoll. At regular intervals over the next century, several different biologists have excavated a bottle and checked the viability of its contents. The table shows the number of germinating seeds (of the original 50) from three species in years 50–100 of this ongoing experiment.

	YEARS AFTER BURIAL					
SPECIES	50	60	70	80	90	100
Oenothera biennis (Evening primrose)	19	12	7	5	0	0
Rumex crispus (Curly dock)	26	2	7	1	0	0
Verbascum blattaria (Moth mullein)	31	34	37	35	10	21

1. Calculate the percentage of surviving seeds for these three species in years 50–100 and graph seed survivorship as a function of time buried.
2. No seeds of the first two species were viable after 90 years of the experiment. Assume 100% seed viability at the start of the experiment (year 0), and predict from your graph the approximate year when you think the last of the *Verbascum blattaria* seeds will germinate.
3. What factors do you think might influence the differences in long-term seed viability among the species?

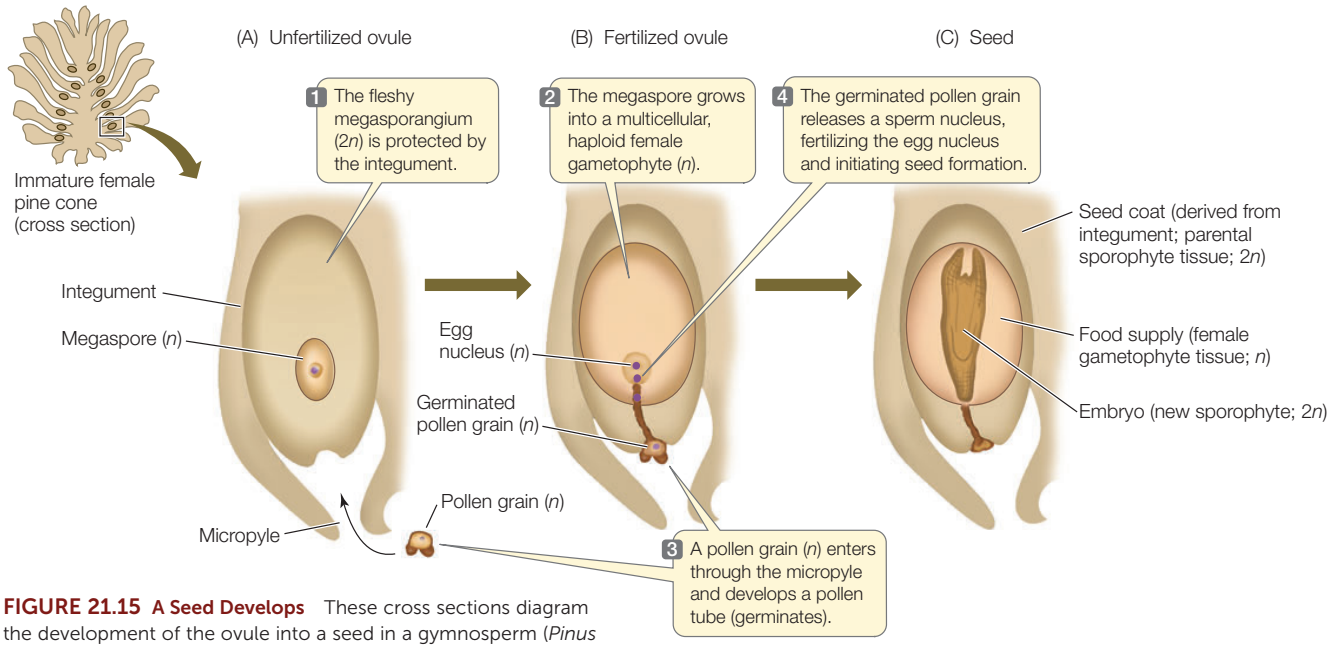


FIGURE 21.15 A Seed Develops These cross sections diagram the development of the ovule into a seed in a gymnosperm (*Pinus* sp.). Angiosperm seed development has differences (e.g., angiosperm integuments have two layers rather than one, and the angiosperm embryo is nourished by specialized tissue called endosperm) but follows the same general steps. (A) The haploid megaspore is nourished by tissues of the parental sporophyte (the diploid megasporangium). (B) The mature megaspore is fertilized by a pollen grain that penetrates the integument, germinates (grows a pollen tube), and releases a sperm nucleus. (C) Fertilization initiates production of a seed. A mature seed contains three generations: a diploid embryo (the new sporophyte), which is surrounded by haploid female gametophyte tissue that supplies nutrition, which is in turn surrounded by the seed coat (diploid parental sporophyte tissue).

The resulting diploid zygote divides repeatedly, forming an embryonic sporophyte. After a period of embryonic development, growth is temporarily suspended (the embryo enters a dormant stage). The end product at this stage is the multicellular seed.

The seed is a complex, well-protected package

A **seed** contains tissues from three generations (FIGURE 21.15). A seed coat develops from the integument—the tissues of the diploid sporophyte parent that surround the megasporangium. Within the megasporangium is haploid tissue from the female gametophyte, which contains a supply of nutrients for the developing embryo. (This tissue is fairly extensive in most gymnosperm seeds. In angiosperm seeds it is greatly reduced, and nutrition for the embryo is supplied instead by a tissue called endosperm.) In the center of the seed is the third generation, the embryo of the new diploid sporophyte.

The seed is a well-protected resting stage. The seeds of some species may remain dormant but stay viable (capable of growth and development) for many years, germinating only when conditions are favorable for the growth of the sporophyte. During the dormant stage, the seed coat protects the embryo from excessive drying and may also protect it against potential predators that would otherwise consume the embryo and its nutrient reserves. Many seeds have structural adaptations or are contained in fruits that promote their dispersal by wind or, more often, by animals. When the young sporophyte resumes growth, it draws on the food reserves in the seed. The possession of seeds is a major reason for the enormous evolutionary success of the seed plants, which are the dominant life forms of most modern terrestrial floras.

A change in stem anatomy enabled seed plants to grow to great heights

The earliest fossil seed plants have been found in late Devonian rocks. These plants had extensively thickened woody stems, which developed through the proliferation of xylem. This type of growth, which increases the diameter of stems and roots in some modern seed plants, is called **secondary growth**, and its product is called secondary xylem, or wood.

The younger portion of the wood produced by secondary growth is well adapted for water transport, but older wood becomes clogged with resins or other materials. Although no longer functional in transport, the older wood continues to provide support for the plant. This support allows woody plants to grow taller than other plants around them and thus capture more light for photosynthesis.

Not all seed plants are woody. In the course of seed plant evolution, many groups lost the woody growth habit. Nonetheless, other advantageous attributes helped them become established in an astonishing variety of places.

Gymnosperms have naked seeds

The two major groups of living seed plants are the **gymnosperms** (such as pines and cycads) and the **angiosperms** (flowering plants; see Figure 21.1B). We'll discuss the flowering plants in Concept 21.5 and examine the gymnosperms here.

The gymnosperms are seed plants that do not form flowers or fruits. Gymnosperms (which means "naked-seeded")

(A) *Encephalartos* sp.

(B) *Ginkgo biloba*

(C) *Welwitschia mirabilis*

(D) *Picea glauca* and *P. mariana*

FIGURE 21.16 Diversity among the Gymnosperms (A) Many cycads have growth forms that resemble both ferns and palms, although cycads are not closely related to either group. (B) The characteristic broad leaves of the maidenhair tree. (C) The straplike leaves of *Welwitschia*, a gnetophyte, grow throughout the life of the plant, breaking and splitting as they grow. (D) Conifers dominate many types of landscapes in the Northern Hemisphere. Shown here are white and black spruce in Canada.

are so named because their ovules and seeds, unlike those of angiosperms, are not protected by ovary or fruit tissue. Although there are probably fewer than 1,200 living species of gymnosperms, these plants are second only to the angiosperms in their dominance of the terrestrial environment. The gymnosperms can be divided into four major groups:

- **Cycads** are palmlike plants of the tropics and subtropics (**FIGURE 21.16A**). Of the present-day gymnosperms, the cycads are probably the earliest-diverging clade. There are about 300 species, some of which grow as tall as 20 meters. The tissues of many species are highly toxic to humans if ingested.
- **Ginkgos**, common during the Mesozoic era, are represented today by a single genus and species: *Ginkgo biloba*, the maidenhair tree (**FIGURE 21.16B**). There are both male (microsporangiate) and female (megasporangiate) maidenhair trees. The difference is determined by X and Y sex chromosomes, as in humans. Few other plants have distinct sex chromosomes.
- **Gnetophytes** number about 90 species in three very different genera, which share certain characteristics analogous to ones found in the angiosperms. One of the gnetophytes is *Welwitschia* (**FIGURE 21.16C**), a long-lived desert plant with just two permanent straplike leaves, which split into many pieces that sprawl across the sand.
- **Conifers** are by far the most abundant of the gymnosperms. There are about 700 species of these cone-bearing plants, including the pines, spruce, and redwoods (**FIGURE 21.16D**).

With the exception of the gnetophytes, the living gymnosperm groups have only tracheids as water-conducting and support cells within the xylem. Most gymnosperms lack the vessel elements and fibers (cells specialized for water conduction and support, respectively) that are found in angiosperms. While the gymnosperm water-transport and support system may thus seem somewhat less efficient than that of the angiosperms, it serves some of the largest trees known. The coastal redwoods of California are the tallest gymnosperms, with some individuals growing to well over 100 meters tall.

During the Permian, as environments became warmer and dryer, the conifers and cycads flourished. Gymnosperm forests changed over time as the gymnosperm groups evolved. Gymnosperms dominated the Mesozoic era, during which the continents drifted apart and large dinosaurs lived. Gymnosperms were the principal trees in all forests until about 65 million years ago, and even today conifers are the dominant trees in many forests, especially at high latitudes and altitudes. The oldest living single organism on Earth today is a gymnosperm

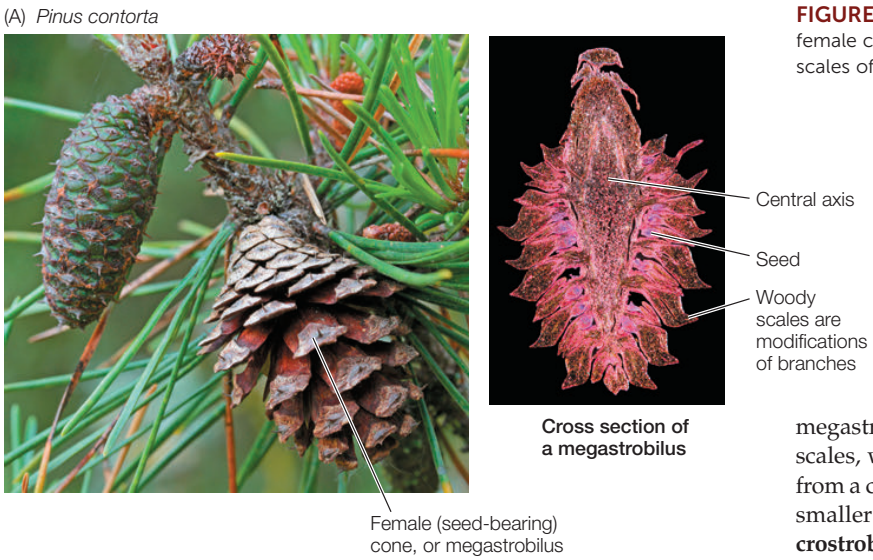


FIGURE 21.17 Female and Male Cones (A) The scales of female cones (megastrobili) are modified branches. (B) The scales of male cones (microstrobili) are modified leaves.

in California—a bristlecone pine that germinated about 4,800 years ago, at about the time the ancient Egyptians were starting to develop writing.

Conifers have cones and lack swimming sperm

The great Douglas fir and cedar forests found in the northwestern United States and the massive boreal forests of pine, fir, and spruce of the northern regions of Eurasia and North America, as well as on the upper slopes of mountain ranges elsewhere, rank among the great forests of the world. All these trees belong to one group of gymnosperms: the conifers, or cone-bearers.

Male and female **cones** contain the reproductive structures of conifers. The female (seed-bearing) cone is known as a **megastrobilus** (plural *megastrobili*). An example of a familiar megastrobilus is the woody cone of pine trees. The seeds in a megastrobilus are protected by a tight cluster of woody scales, which are modifications of branches extending from a central axis (FIGURE 21.17A). The typically much smaller male (pollen-bearing) cone is known as a **microstrobilus**. The microstrobilus is typically herbaceous rather than woody, as its scales are composed of modified leaves, beneath which are the pollen-bearing microsporangia (FIGURE 21.17B).

The life cycle of a pine illustrates reproduction in gymnosperms (FIGURE 21.18). The production of male gametophytes in the form of pollen grains frees the plant completely from its dependence on liquid water for fertilization. Wind, rather than water, assists conifer pollen grains in their first stage of travel from the microstrobilus to the female gametophyte inside a cone. A pollen tube provides the sperm with the means for the last stage of travel by elongating through maternal sporophytic tissue. When the pollen tube reaches the female gametophyte, it releases two sperm, one of which degenerates after the other unites with an egg. Union of sperm and egg results in a zygote. Mitotic divisions and further development of the zygote result in an embryo.

The megasporangium, in which the female gametophyte will form, is enclosed in a layer of sporophytic tissue—the integument—that will eventually develop into the seed coat that protects the embryo. The integument, the megasporangium inside it, and the tissue attaching it to the maternal sporophyte constitute the ovule. The pollen grain enters through a small opening in the integument at the tip of the ovule, the **micropyle**.

Most conifer ovules (which will develop into seeds after fertilization) are borne exposed on the upper surfaces of the scales of the cone (megastrobilus). The only protection of the ovules comes from the scales, which are tightly pressed against one another within the cone. Some pines, such as the lodgepole pine, have such tightly closed cones that only fire suffices to split them open and release the seeds. These species are said to be fire-adapted, and fire is essential to their reproduction.

About half of all conifer species have soft, fleshy modifications of cones that envelop their seeds. Some of these are fleshy, fruitlike cones, as in junipers. Others are fruitlike

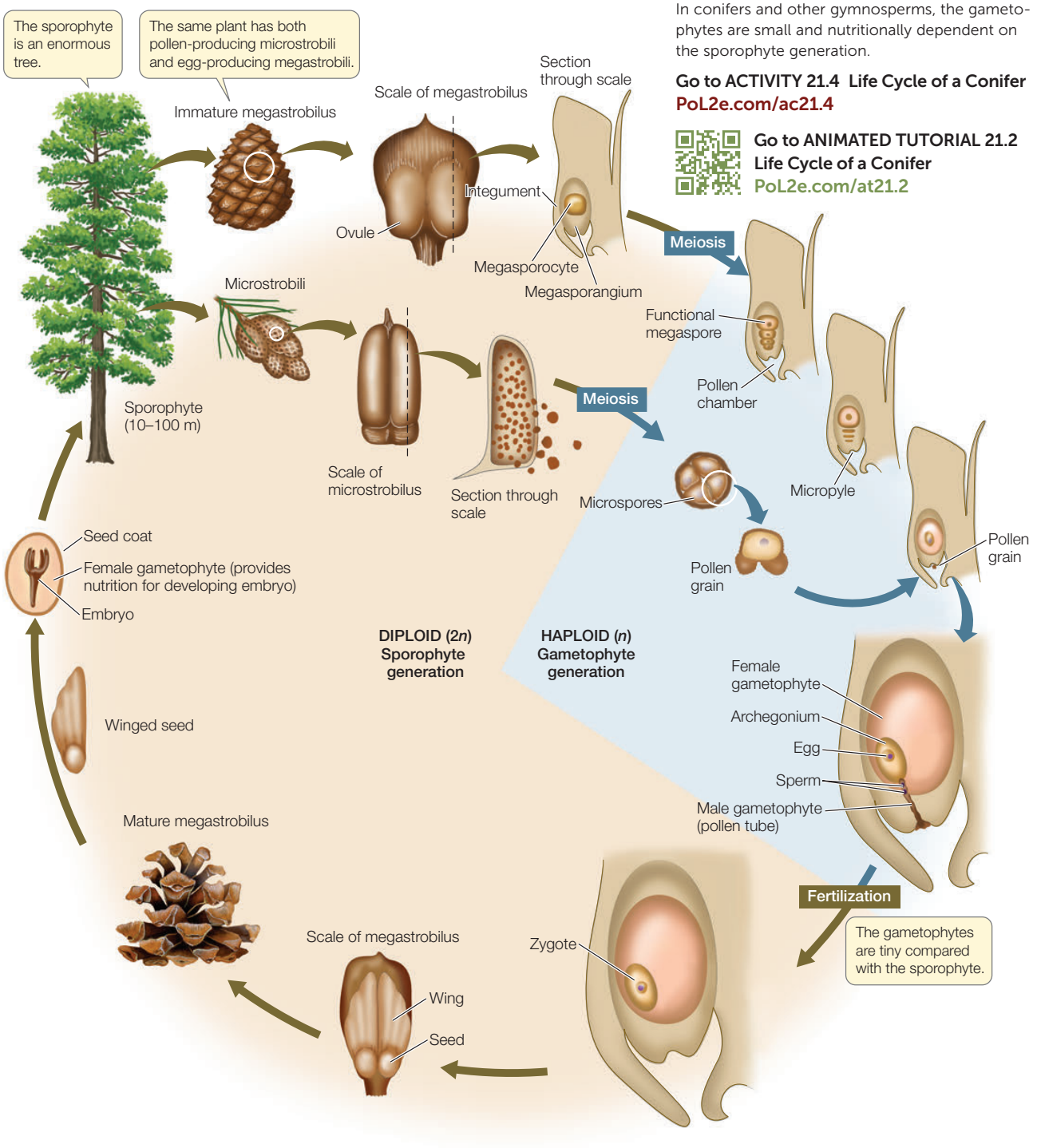


FIGURE 21.18 The Life Cycle of a Pine Tree
In conifers and other gymnosperms, the gametophytes are small and nutritionally dependent on the sporophyte generation.

Go to ACTIVITY 21.4 Life Cycle of a Conifer
PoL2e.com/ac21.4

Go to ANIMATED TUTORIAL 21.2 Life Cycle of a Conifer
PoL2e.com/at21.2

extensions of the seeds, as in yews. These tissues, although often mistaken for "berries," are not true fruits. As we will see when we discuss angiosperms, true fruits are a plant's ripened ovaries, which are absent in gymnosperms. Nonetheless, the fleshy tissues that surround many conifer seeds serve a similar purpose as that of the fruits of flowering plants, acting as an enticement for seed-dispersing animals. Animals eat these fleshy tissues and disperse the seeds in their feces, often depositing the seeds considerable distances from the parent plant.

CHECKpoint CONCEPT 21.4

- ✓ Distinguish between the roles of the megagametophyte and the pollen grain.
- ✓ Explain the importance of pollen in freeing seed plants from dependence on liquid water.
- ✓ What are some of the advantages afforded by seeds? By wood?
- ✓ Explain how fire can be necessary for the survival of some plant species.

CONCEPT 21.5 Flowers and Fruits Increase the Reproductive Success of Angiosperms

The most obvious feature defining the angiosperms is the **flower**, which is their sexual structure. Production of **fruits** is also a shared derived trait of angiosperms. After fertilization, the ovary of a flower (together with the seeds it contains) develops into a fruit that protects the seeds and can promote seed dispersal. As we will see, both flowers and fruits afford major reproductive advantages to angiosperms.

Angiosperms have many shared derived traits

The name *angiosperm* ("enclosed seed") is derived from another distinctive trait of flowering plants that is related to the formation of fruits: the ovules and seeds are enclosed in a modified leaf called a **carpel**. Besides protecting the ovules and seeds, the carpel often interacts with incoming pollen to prevent self-fertilization, thus favoring cross-fertilization and increasing genetic diversity.

The female gametophyte of angiosperms is even more reduced than that of gymnosperms, usually consisting of only seven cells. Thus the angiosperms represent the current extreme of the trend we have traced throughout the evolution of vascular plants: the sporophyte generation becomes larger and more independent of the gametophyte, while the gametophyte becomes smaller and more dependent on the sporophyte.

The xylem of most angiosperms is distinguished by the presence of specialized water-transporting cells called **vessel elements**. These cells are larger in diameter than tracheids and connect with one another without obstruction, allowing easy water movement. A second distinctive cell type in angiosperm xylem is the **fiber**, which plays an important role in supporting the plant body. Angiosperm phloem possesses another unique cell type, called a companion cell. Like the gymnosperms, woody angiosperms show secondary growth, increasing in diameter by producing secondary xylem and secondary phloem.

A more comprehensive list of angiosperm synapomorphies, then, includes the following (some of these traits will be discussed later in this chapter):

- Flowers
- Fruits
- Ovules and seeds enclosed in a carpel
- Highly reduced gametophytes
- Germination of pollen on a stigma
- Double fertilization
- Endosperm (nutritive tissue for the embryo)
- Phloem with companion cells

The majority of these traits bear directly on angiosperm reproduction, which is a large factor in the success of this dominant plant group.

The sexual structures of angiosperms are flowers

Flowers come in an astonishing variety of forms—just think of some of the flowers you recognize. Flowers may be single, or they may be grouped together to form an **inflorescence**. Different families of flowering plants have characteristic types of inflorescences, such as the compound umbels of the carrot family (FIGURE 21.19A), the heads of the aster family (FIGURE 21.19B), and the spikes of many grasses (FIGURE 21.19C).

If you examine any familiar flower, you will notice that the outer parts look somewhat like leaves. In fact, all the parts of a flower *are* modified leaves. The diagram in Figure 21.14B represents a generalized flower (for which there is no exact counterpart in nature). The structures bearing microsporangia are called **stamens**. Each stamen is composed of a **filament** bearing an **anther** that contains the pollen-producing microsporangia. The structures bearing megasporangia are called carpels. The swollen base of the carpel, containing one or more ovules (each containing a megasporangium surrounded by two protective integuments), is called the **ovary**. The stalk at the top of the carpel is the **style**, and the terminal surface that receives pollen grains is the **stigma**. Two or more fused carpels, or a single carpel if only one is present, are also called a **pistil**.

Go to ACTIVITY 21.5 Flower Morphology
PoL2e.com/ac21.5

In addition, many flowers have specialized sterile (non-spore-bearing) leaves. The inner ones are called **petals** (collectively, the **corolla**) and the outer ones **sepals** (collectively, the calyx). The corolla and calyx can be quite showy and often play roles in attracting animal pollinators to the flower. The calyx more commonly protects the immature flower in bud. From base to apex, these floral organs—sepals, petals, stamens, and carpels—are usually positioned in circular arrangements or whorls and attached to a central stalk.

The generalized flower in Figure 21.14B has functional megasporangia and microsporangia. Such flowers are referred to as **perfect** (or hermaphroditic). Many angiosperms produce two types of flowers, one with only megasporangia and the other with only microsporangia. Consequently, either the stamens or the carpels are nonfunctional or absent in a given flower, and the flower is referred to as **imperfect**.

Species such as corn or birch, in which both megasporangiate (female) and microsporangiate (male) flowers occur on the same plant, are said to be **monoecious** ("one-housed"—but, it must be added, one house with separate rooms). Complete

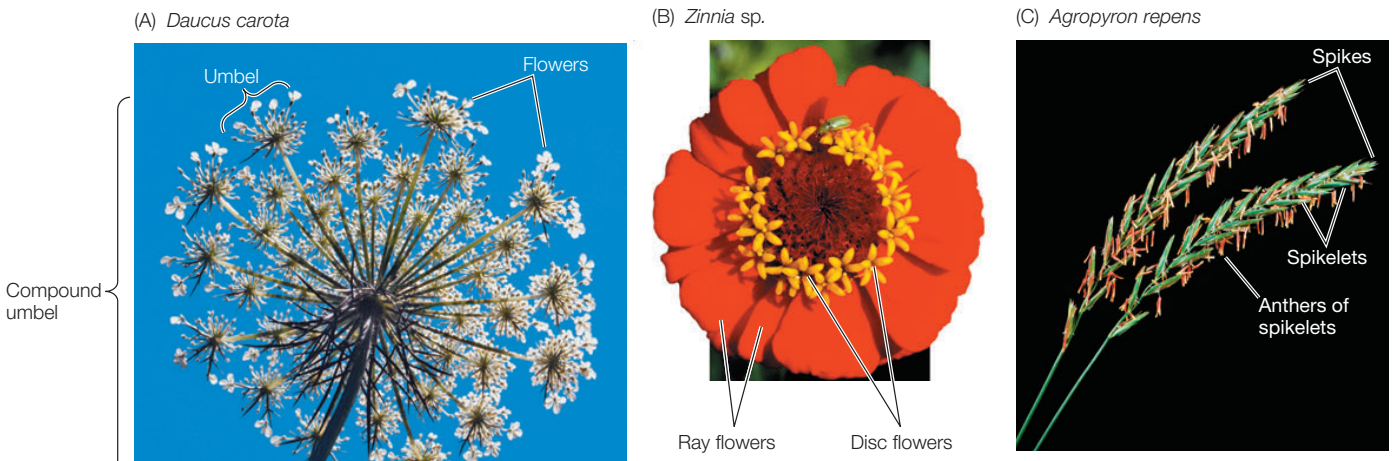


FIGURE 21.19 Inflorescences (A) The inflorescence of Queen Anne's lace, a member of the carrot family, is a compound umbel. Each umbel bears flowers on stalks that arise from a common center. (B) Zinnias are members of the aster family; their inflorescence is a head. Within the head, each of the long, petal-like structures is a ray flower; the central portion of the head consists of dozens to hundreds of disc flowers. (C) Some grasses, such as quack grass, have inflorescences called spikes, which are composed of many individual flowers, or spikelets.

separation of imperfect flowers occurs in some other angiosperm species, such as willows and date palms; in these species, an individual plant produces either flowers with stamens or flowers with carpels, but never both. Such species are said to be **dioecious** ("two-housed").

Flower structure has evolved over time

The flowers of the earliest-diverging clades of angiosperms have a large and variable number of tepals (undifferentiated sepals and petals), carpels, and stamens (**FIGURE 21.20A**). Evolutionary change within the angiosperms has included some striking modifications of this early condition: reductions in the number of each type of floral organ to a fixed number, differentiation of petals from sepals, and changes in symmetry from radial (as in a lily or magnolia) to bilateral (as in a sweet pea or orchid), often accompanied by an extensive fusion of parts (**FIGURE 21.20B**).

According to one hypothesis, the first carpels to evolve were leaves with marginal sporangia, folded but incompletely closed. Early in angiosperm evolution, the carpels fused with one another, forming a single, multichambered ovary (**FIGURE 21.21A**). In some flowers, the other floral organs are attached at the top of the ovary, rather than at the bottom as in Figure 21.14B. The stamens of the most ancient flowers may have been leaflike (**FIGURE 21.21B**), with little resemblance to the stamens of the generalized flower in Figure 21.14B.

A perfect flower represents a compromise of sorts. On the one hand, by attracting a pollinating bird or insect, the plant is attending to both its female and male functions with a single flower type, whereas plants with imperfect flowers must create that attraction twice—once for each type of flower. On the other hand, the perfect flower can favor self-pollination, which is usually disadvantageous. Another potential problem is that the female and male functions might interfere with each other—for example, the stigma might be so placed as to make it difficult for pollinators to reach the anthers, thus reducing the export of pollen to other flowers.

Might there be a way around these problems? One solution is seen in the bush monkeyflower (*Mimulus aurantiacus*), which is pollinated by hummingbirds. Its flower has a stigma that

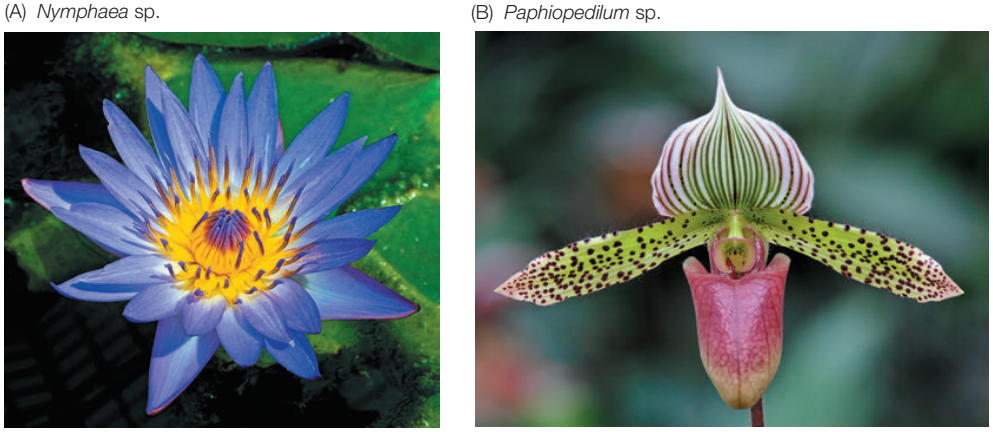


FIGURE 21.20 Flower Form and Evolution (A) A water lily shows the major features of early flowers: it is radially symmetrical, and the individual tepals, stamens, and carpels are separate, numerous, and attached at their bases. (B) Orchids such as this lady slipper have a bilaterally symmetrical structure that evolved much later than radial flower symmetry.

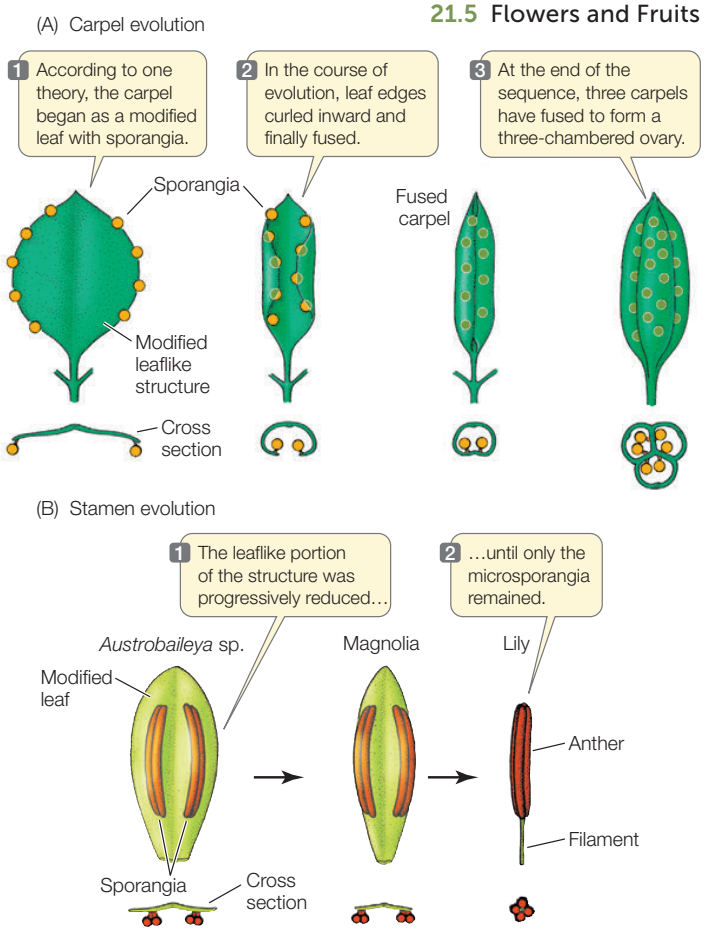


FIGURE 21.21 Carpels and Stamens Evolved from Leaflike Structures (A) Possible stages in the evolution of a carpel from a more leaflike structure. (B) The stamens of three modern plants show three possible stages in the evolution of that organ. (It is *not* implied that these species evolved from one another; their structures simply illustrate the possible stages.)

initially serves as a screen, hiding the anthers (FIGURE 21.22). Once a hummingbird touches the stigma, one of the stigma's two lobes is retracted, so that subsequent hummingbird visitors pick up pollen from the previously screened anthers. Thus the first bird to visit the flower transfers pollen from another plant to the stigma, eventually leading to fertilization. Later visitors pick up pollen from the now-accessible anthers, fulfilling the flower's male function. FIGURE 21.23 describes the experiment that revealed the function of this mechanism.

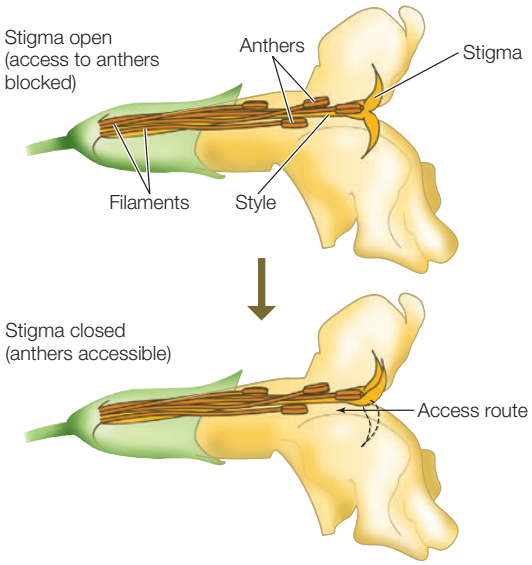


FIGURE 21.22 An Unusual Way to Prevent Self-Fertilization Both long stamens and long styles facilitate cross-pollination, but if these male and female structures are too close to each other, the likelihood of (disadvantageous) self-pollination increases. In *Mimulus aurantiacus*, the stigma is initially open, blocking access to the anthers. A hummingbird's touch as it deposits pollen on the stigma causes one lobe of the stigma to retract, creating a path to the anthers and allowing pollen dispersal.

Angiosperms have coevolved with animals

Whereas most gymnosperms are pollinated by wind, most angiosperms are pollinated by animals. The many different mutualistic pollination relationships between plants and animals are vital to both parties. We mentioned coevolution between insects and orchids at the opening of this chapter, but we'll consider a few additional aspects of plant–pollinator coevolution here.

Many flowers entice animals to visit them by providing food rewards. Pollen grains themselves sometimes serve as food for animals. In addition, some flowers produce a sugary fluid called nectar as a pollinator attractant, and some of these flowers have specialized structures to store and distribute it, as we saw at the opening of this chapter. In the process of visiting flowers to obtain nectar or pollen, animals often carry pollen from one flower to another or from one plant to another. Thus, in their quest for food, the animals contribute to the genetic diversity of the plant population. Insects, especially bees, are among the most important pollinators. Other major pollinators include some species of birds and bats.

Go to MEDIA CLIP 21.5
Pollen Transfer by a Bat
PoL2e.com/mc21.5

For more than 150 million years, angiosperms and their animal pollinators have coevolved in the terrestrial environment. The animals have affected the evolution of the plants, and the plants have affected the evolution of the animals. Flower structure has become incredibly diverse under these selection pressures. Some of the products of coevolution are highly specific. For example, the flowers of some yucca species are pollinated by only one species of yucca moth, and that moth may exclusively pollinate just one species of yucca. Such specific relationships provide plants with a reliable mechanism for transferring pollen only to members of their own species.

Most plant–pollinator interactions are much less specific. In most cases, many different animal species pollinate the same plant species, and the same animal species pollinates many different plant species. However, even these less

INVESTIGATION

FIGURE 21.23 The Effect of Stigma Retraction in Monkeyflowers
Elizabeth Fetscher's experiments showed that the unusual stigma retraction response to pollination in monkeyflowers (illustrated in Figure 21.22) enhances the dispersal of pollen to other flowers.[a]

HYPOTHESIS

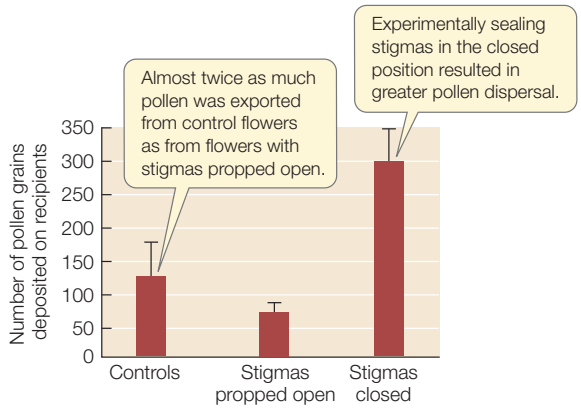
The stigma-retraction response in *M. aurantiacus* increases the likelihood than an individual flower's pollen will be exported to another flower once pollen from another flower has been deposited on its stigma.

METHOD

1. Set up three groups of monkeyflower arrays. Each array consists of one pollen-donor flower and multiple pollen-recipient flowers (with the anthers removed to prevent pollen donation).
2. In control arrays, the stigma of the pollen donor is allowed to function normally.
3. In one set of experimental arrays, the stigma of the pollen donor is permanently propped open.
4. In a second set of experimental arrays, the stigma of the pollen donor is artificially sealed closed.
5. Allow hummingbirds to visit the arrays, then count the pollen grains transferred from each donor flower to the recipient flowers in the same array.

RESULTS

Error bars show 1 standard error of the mean; the three groups are significantly different at $P < 0.05$ (see Appendix B).



CONCLUSION

The stigma-retraction response enhances the male function of the flower (dispersal of pollen) once the female function (receipt of pollen) has been performed.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]A. E. Fetscher. 2001. *Proceedings of the Royal Society B* 268: 525–529.

specific interactions have developed some specialization. Bird-pollinated flowers are often red and odorless. Many insect-pollinated flowers have characteristic odors, and bee-pollinated flowers may have conspicuous markings, called

Taraxacum officinale

FIGURE 21.24 See Like a Bee To normal human vision (above), the petals of a dandelion appear solid yellow. Ultraviolet photography reveals patterns that attract bees to the central region, where pollen and nectar are located.

nectar guides, that are conspicuous only to animals, such as bees, that can see colors in the ultraviolet region of the spectrum (**FIGURE 21.24**).

The angiosperm life cycle produces diploid zygotes nourished by triploid endosperms

Like all seed plants, angiosperms are heterosporous. As we have seen, their ovules are contained within carpels rather than being exposed on the surfaces of scales, as in most gymnosperms. The male gametophytes, as in the gymnosperms, are pollen grains.

Pollination in the angiosperms consists of the arrival of a microgametophyte—a pollen grain—on a receptive surface in a flower (the stigma). As in the gymnosperms, pollination is the first in a series of events that results in the formation of a seed. The next event is the growth of a pollen tube extending to the megagametophyte. The third event is a fertilization process that, in detail, is unique to the angiosperms (**FIGURE 21.25**).

In nearly all angiosperms, *two* male gametes, contained in a single microgametophyte, participate in fertilization. The nucleus of one sperm combines with that of the egg to produce a diploid zygote, the first cell of the sporophyte generation. In most angiosperms, the other sperm nucleus combines with two other haploid nuclei of the female gametophyte to form a cell with a *triploid* ($3n$) nucleus. That cell, in turn, gives rise to triploid tissue, the **endosperm**, which nourishes the embryonic sporophyte

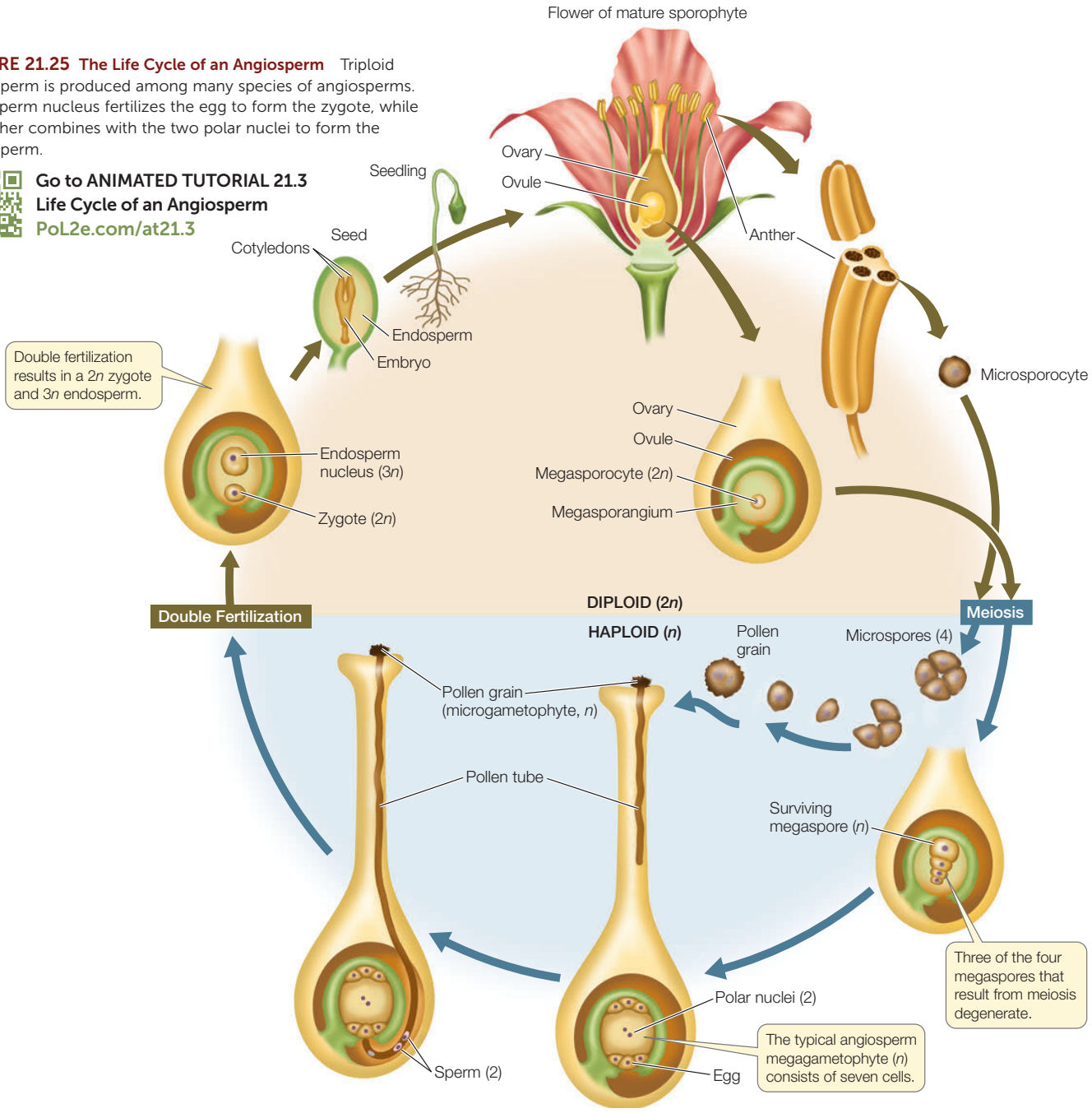


FIGURE 21.25 The Life Cycle of an Angiosperm Triploid endosperm is produced among many species of angiosperms. One sperm nucleus fertilizes the egg to form the zygote, while the other combines with the two polar nuclei to form the endosperm.

Go to ANIMATED TUTORIAL 21.3
Life Cycle of an Angiosperm
PoL2e.com/at21.3

during its early development. This process, in which two fertilization events take place, is known as **double fertilization**.

As Figure 21.25 shows, the zygote develops into an embryo, which consists of an embryonic axis (the "backbone" that will become a stem and a root) and one or two **cotyledons**, or "seed leaves." The cotyledons have different fates in different plants. In many, they serve as absorptive organs that take up and digest the endosperm. In others, they enlarge and become photosynthetic when the seed germinates. Often they play both roles.

The ovule develops into a seed containing the products of the double fertilization that characterizes angiosperms: a diploid zygote and a triploid endosperm (see Figure 21.25). The endosperm serves as storage tissue for starch or lipids, proteins, and other substances that will be needed by the developing embryo.

Fruits aid angiosperm seed dispersal

Fruits typically aid in seed dispersal. Fruits may attach to or be eaten by an animal. The animal is then likely to move, after which the seeds may fall off or be defecated. Fruits are not necessarily fleshy. Fruits can also be hard and woody, or small and have modified structures that allow the seeds to be dispersed by wind or water.

A fruit may consist of only the mature ovary and its seeds, or it may include other parts of the flower or structures associated with it. A **simple fruit** is one that develops from a single

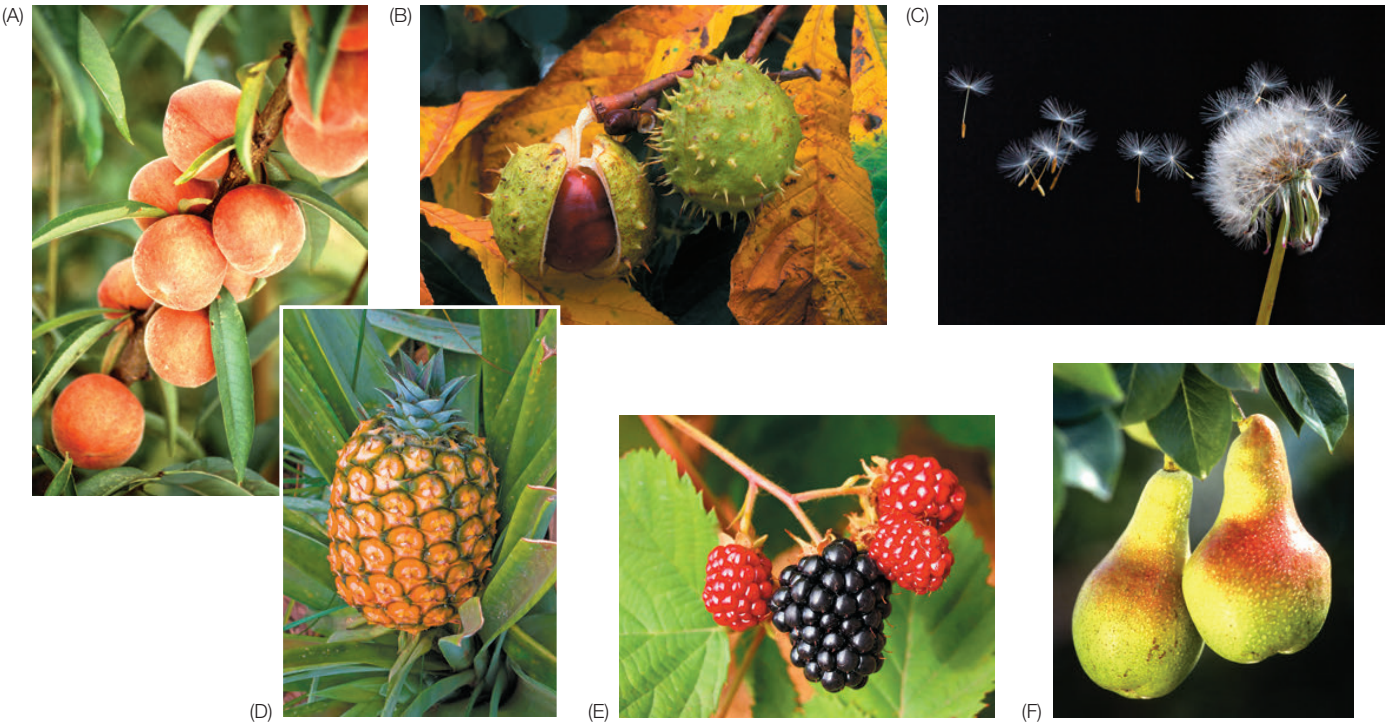


FIGURE 21.26 Fruits Come in Many Forms (A) The single seeds inside the simple fruits of peaches are dispersed by animals. (B) Each horse chestnut seed is covered by a spiny husk. (C) The highly reduced simple fruits of dandelions are dispersed by wind. (D) A multiple fruit, the pineapple (*Ananas comosus*), has become one of the most economically significant fruit crops of the tropics. (E) An aggregate fruit (blackberry). (F) An accessory fruit (pear).

carpel or several fused carpels, such as a plum or peach. A raspberry is an example of an **aggregate fruit**—one that develops from several separate carpels of a single flower. Pineapples and figs are examples of **multiple fruits**, formed from a cluster of flowers (an inflorescence). Fruits derived from parts in addition to the carpel and seeds are called **accessory fruits**—examples are apples, pears, and strawberries (FIGURE 21.26).

APPLY THE CONCEPT

Fruits increase the reproductive success of angiosperms

Many fleshy fruits attract animals, which eat the fruit and then disperse the seeds in their feces. If all other factors are equal, large seeds have a better chance of producing a successful seedling than small seeds. So why isn't there selection for larger seeds in the fruits of all plants?

In one study in Peru, the feces of the spider monkey *Ateles paniscus* were found to contain seeds from 71 species of plants. After eating fruit, the monkeys usually travel some distance before defecating, thus dispersing any undigested seeds.

If monkey feces are left undisturbed on the forest floor, rodents eat and destroy the vast majority of the seeds in the feces. To germinate successfully, the seeds in spider monkey feces need to be buried by dung beetles, which makes the discovery and destruction of seeds by rodents much less likely.

Ellen Andresen hypothesized that dung beetles were more likely to remove larger than smaller seeds from spider monkey dung before burying the dung. She added plastic beads of various diameters to spider monkey dung (to simulate seeds) and measured the percentage of beads buried with the dung by the beetles. Use her data to answer the questions at right.[a]

Bead diameter (mm)	2	4	6	8	10	12
Percentage buried	100	76	52	39	20	4

[a]Andresen, E. (1999). *Biotropica* 31(1): 145–158.

1. Plot bead size (the independent variable) versus percentage of beads buried by dung beetles (the dependent variable).
2. Calculate a regression line for the relationship shown in your graph (see Appendix B). Approximately what percentage of beads with a diameter of 5 mm would you predict would be buried by the beetles? What about beads 14 mm in diameter?
3. What other factors besides size might influence the probability of seed burial by dung beetles? Can you design an experiment to test your hypotheses?
4. Describe how changes in the population sizes of spider monkeys, rodents, and dung beetles would affect the reproductive success of various plant species.

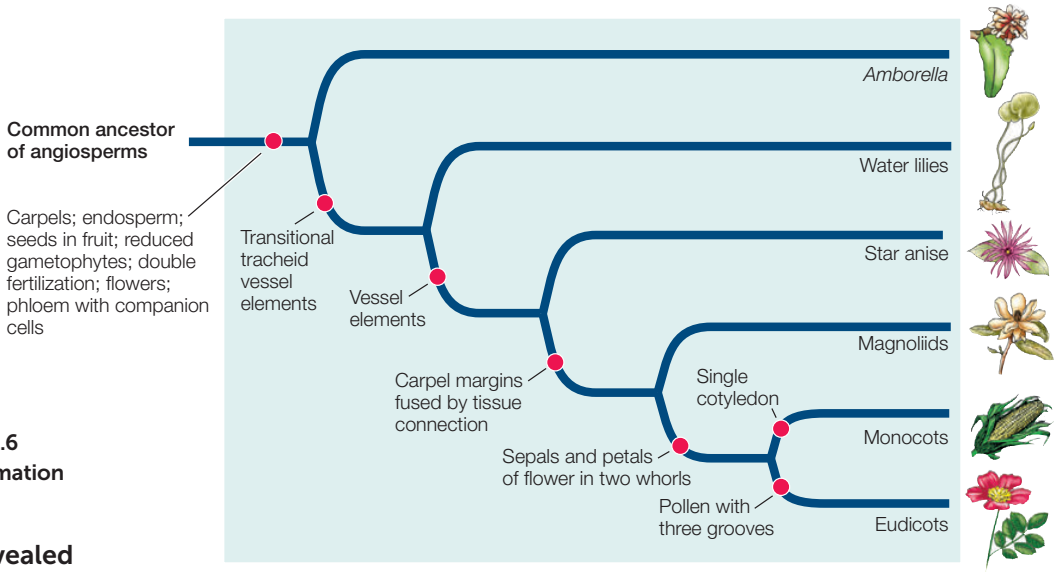


FIGURE 21.27 Evolutionary Relationships among the Angiosperms Recent analyses of many angiosperm genes have clarified the relationships among the major groups.

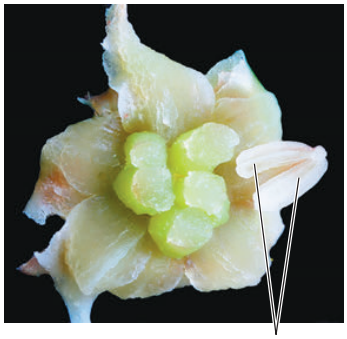
Go to MEDIA CLIP 21.6
Flower and Fruit Formation
PoL2e.com/mc21.6

Recent analyses have revealed the phylogenetic relationships of angiosperms

FIGURE 21.27 shows the relationships among the major angiosperm clades. Recent molecular and morphological analyses have supported the hypothesis that the sister group of remaining flowering plants is a single species of the genus *Amborella* (see Figure 21.28A). This woody shrub, with cream-colored flowers, lives only on New Caledonia, an island in the South Pacific. Other early branching angiosperm groups include the water lilies, star anise and its relatives, and the **magnoliids** (**FIGURE 21.28**). The magnoliids include many familiar and useful plants, such as avocados, cinnamon, black pepper, and magnolias.

FIGURE 21.28 Monocots and Eudicots Are Not the Only Surviving Angiosperms (A) *Amborella*, a shrub, is sister to the remaining extant angiosperms. Notice the sterile stamens on this female flower, which may serve to lure insects that are searching for pollen. (B) The water lily clade was the next to diverge after *Amborella*. (C) Star anise and its relatives belong to another early-diverging angiosperm clade. (D,E) The largest clade other than the monocots and eudicots is the magnoliid complex, which includes magnolias and the group known as "Dutchman's pipe."

(A) *Amborella trichopoda*



(B) *Victoria cruziana*

(C) *Illicium floridanum*

(D) *Magnolia* sp.

(E) *Aristolochia littoralis*

(A) *Tulipa* sp. and *Narcissus* sp.

(B) *Triticum aestivum*

(C) *Posidonia* sp.

(D) *Phoenix dactylifera*

FIGURE 21.29 Monocots (A) Monocots include many popular garden flowers such as these tulips (pink and white), and daffodils (yellow). (B) Monocot grasses such as wheat feed the world; sugarcane, rice, and maize (corn) are also grasses. (C) Seagrasses form "meadows" in the shallow, sunlit waters of the world's oceans. (D) Palms are among the few monocot trees. Date palms like these are a major food source in some areas of the world.

The two largest clades—the **monocots** and the **eudicots**—include the great majority of angiosperm species. The monocots are so called because they have a single embryonic cotyledon, whereas the eudicots have two.

Representatives of the two largest angiosperm clades are everywhere. The monocots (**FIGURE 21.29**) include grasses, cattails, lilies, orchids, and palms. The eudicots (**FIGURE 21.30**) include the vast majority of familiar seed plants, including most herbs (i.e., nonwoody plants), vines, trees, and shrubs. Among the eudicots are such diverse plants as oaks, willows, beans, snapdragons, roses, and sunflowers.

CHECKpoint CONCEPT 21.5

- ✓ Explain the difference between pollination and fertilization.
- ✓ Give some examples of how animals have affected the evolution of the angiosperms.
- ✓ What are the respective roles of the two sperm in double fertilization?
- ✓ What are the different functions of flowers, fruits, and seeds?

Once life moved onto land, it was plants that shaped the terrestrial environment. Terrestrial ecosystems could not function without the foods and habitats provided by plants. Plants produce oxygen and remove carbon dioxide from the atmosphere. They play important roles in forming soils and renewing their fertility. Plant roots help hold soil in place, providing protection against erosion by wind and water. Plants also moderate local

(A) *Malus* sp.

(B) *Banksia coccinea*

(C) *Rafflesia arnoldii*

(D) *Echinopsis calochlora*

FIGURE 21.30 Eudicots (A) Eudicots include many trees, such as this crabapple tree. (B) Scarlet Banksia is a species of an Australian genus of eudicots that attracts a wide diversity of pollinators by producing large quantities of nectar. (C) The largest flower in the world is *Rafflesia arnoldii*, found in the rainforests of Indonesia. This plant is a root parasite of tropical vines and has lost its leaf, stem, and even root structures. The flower is the only part of the plant that can be easily seen. (D) Cacti comprise a large group of eudicots, with about 1,500 species in the Americas. Many bear large flowers for a brief period of each year.

climates in various ways, such as by increasing humidity, providing shade, and blocking wind. All of these ecosystem services permit a great diversity of fungi and animals to exist on land.

Q What was Darwin's explanation for the three distinct flowers growing on a single orchid plant?

ANSWER After obtaining specimens of the plant in question and dissecting its flowers, Darwin was able to demonstrate that the orchid was a single species (*Catasetum macrocarpum*) that bore three distinct types of flowers: megasporangiate (female), microsporangiate (male), and perfect (hermaphroditic). The three types of flowers were remarkable in their morphological differences, which had misled botanists into describing the different flower types as species in different genera. Most plants were either male (specimens identified as *Catasetum*) or female (specimens identified as *Monachanthus*), but some individuals that bore predominately male or female flowers also produced perfect flowers (specimens identified as *Myanthus*).

The case of *C. macrocarpum* demonstrates that some plants blur the lines between strict dioecy (male and female flowers in separate individuals), monoecy (male and female flowers in the same individual), and perfect flowers (flowers with both male and female parts). The flowers on a *C. macrocarpum* are either male or female at any one time, except that plants can bear some perfect flowers as well.

Why do the male and female flowers of *C. macrocarpum* look so different? Part of the explanation is their different roles in pollination. Recall Darwin's observation of a *Catasetum* flower shooting a packet of pollen at an insect that landed on its flower. The pollinia (pollen packets) and associated structures in male flowers of *Catasetum* are coiled like springs and are released suddenly when disturbed by an insect. This release forcefully propels the pollinia precisely into position on the

back of the insect. The insect pollinator of *C. macrocarpum* is a specific bee species, the males of which are attracted to the odor of the flowers. The flowers produce no nectar reward, but the male bee does gather the chemical that produces the scent. The bee then moves on to another flower. When the bee visits a female flower on a different *C. macrocarpum* individual (again attracted by the same scent), no such "loaded spring" awaits. Instead, the morphology of the female flower enhances the removal of the pollinia from the insect's body. In this way, floral morphology makes cross-fertilization more likely and reduces the chances of self-pollination.

Orchids were important in forming Darwin's ideas about the mechanisms of evolution, for they showed that even aspects of the coloration and form of flowers evolved in response to natural selection. This conclusion ran counter to the thinking of the day. For example, Thomas Huxley, one of Darwin's earliest and strongest supporters, doubted that the beauty of color in plants and animals could be explained on the basis of their importance to function. Darwin showed that the beauty of flowers is indeed connected to their reproductive success and is a key element in explaining the great diversity of plants.

SUMMARY

CONCEPT 21.1 Primary Endosymbiosis Produced the First Photosynthetic Eukaryotes

- Primary endosymbiosis gave rise to chloroplasts and the subsequent diversification of the **Plantae**. The descendants of the first photosynthetic eukaryote include **glaucophytes**, **red algae**, several groups of green algae, and **land plants**, all of which contain chlorophyll *a*. Review Figure 21.1
- **Streptophytes** include the land plants and two groups of green algae. **Green plants**, which include the streptophytes and the remaining green algae, are characterized by the presence of chlorophyll *b* (in addition to chlorophyll *a*). Review Figure 21.1
- Land plants, also known as **embryophytes**, arose from an aquatic green algal ancestor related to today's **stoneworts**. Land plants develop from embryos that are protected by parental tissue. Review Figure 21.1 and Table 21.1

CONCEPT 21.2 Key Adaptations Permitted Plants to Colonize Land

- The acquisition of a cuticle, stomata, gametangia, a protected embryo, protective pigments, thick spore walls with a protective polymer, and a mutualistic association with a fungus were all adaptations of land plants to terrestrial life.
- All land plant life cycles feature alternation of generations, in which a multicellular diploid **sporophyte** alternates with a multicellular haploid **gametophyte**. Review Figure 21.4
- The **nonvascular land plants** comprise the **liverworts**, **hornworts**, and **mosses**. These groups lack specialized vascular tissues for the conduction of water or nutrients through the plant body.
- The life cycles of nonvascular land plants depend on liquid water. The sporophyte is usually smaller than the gametophyte and depends on it for water and nutrition.
- In many land plants, spores form in structures called **sporangia** and gametes form in structures called **gametangia**. Female and male gametangia are, respectively, an **archegonium** and an **antheridium**. Review Figure 21.6 and ANIMATED TUTORIAL 21.1

CONCEPT 21.3 Vascular Tissues Led to Rapid Diversification of Land Plants

- The **vascular plants** have a vascular system consisting of xylem and phloem that conducts water, minerals, and products of photosynthesis through the plant body. The vascular system includes cells called **tracheids**.
- The **rhyniophytes**, the earliest known vascular plants, are known to us only in fossil form. They lacked true roots and leaves but apparently possessed **rhizomes** and **rhizoids**.
- Among living vascular plant groups, the **lycophytes** (club mosses and relatives) have only small, simple leaflike structures (**microphylls**). True leaves (**megaphylls**) are found in **monilophytes** (which include **horsetails** and **ferns**). The monilophytes and the seed plants are collectively called **euphyllophytes**. Review ACTIVITY 21.1
- Roots may have evolved either from rhizomes or from stems. Microphylls probably evolved from sterile sporangia, and megaphylls may have resulted from the flattening and reduction of a portion of a stem system with **overtopping** growth. Review Figure 21.10
- The earliest-diverging groups of vascular plants are **homosporous**, but **heterospory**—the production of distinct **megaspores** and **microspores**—has evolved several times. Megaspores develop into female **megagametophytes**; microspores develop into male **microgametophytes**. Review Figure 21.11 and ACTIVITIES 21.2 and 21.3

CONCEPT 21.4 Seeds Protect Plant Embryos

- All **seed plants** are heterosporous, and their gametophytes are much smaller than (and dependent on) their sporophytes. Review Figure 21.12
- Seed plants do not require liquid water for fertilization. **Pollen grains**, the microgametophytes of seed plants, are carried to a megagametophyte by wind or by animals. Following **pollination**, a **pollen tube** emerges from the pollen grain and elongates to deliver gametes to the megagametophyte. Review Figure 21.14
- An **ovule** consists of the seed plant megagametophyte and the **integument** of sporophytic tissue that protects it. The ovule develops into a **seed**. Review Figure 21.15
- Seeds are well protected, and they are often capable of long periods of dormancy, germinating when conditions are favorable.
- The two major groups of living seed plants are the **gymnosperms** and **angiosperms**. Review Figure 21.1B
- The gymnosperms produce ovules and seeds that are not protected by ovary or fruit tissues. The major gymnosperm groups are the **cycads**, **ginkgos**, **gnetophytes**, and **conifers**. Review Figure 21.16
- The megaspores of conifers are produced in woody **cones** called **megastrobili**; the microspores are produced in herbaceous cones called **microstrobili**. Pollen reaches the megagametophyte by way of the **micropyle**, an opening in the integument of the ovule. Review Figures 21.17 and 21.18, ACTIVITY 21.4, and ANIMATED TUTORIAL 21.2

SUMMARY *(continued)*

CONCEPT 21.5 Flowers and Fruits Increase the Reproductive Success of Angiosperms

- **Flowers** and **fruits** are unique to the angiosperms, distinguishing them from the gymnosperms.
- The xylem of angiosperms is more complex than that of the gymnosperms. It contains two specialized cell types: **vessel elements**, which function in water transport, and **fibers**, which play an important role in structural support.
- The floral organs, from the base to the apex of the flower, are the **sepals**, **petals**, **stamens**, and one or more **carpels**. Stamens bear microsporangia in **anthers**. The carpel includes an **ovary** containing ovules and a receptive surface called the **stigma**. **Review Figure 21.14B and ACTIVITY 21.5**
- The structure of flowers has evolved over time. A flower with both megasporangia and microsporangia is referred to as **perfect**; a flower with only one or the other is **imperfect**. Some plants with perfect flowers have adaptations to prevent self-fertilization. **Review Figures 21.22 and 21.23**
- A **monoecious** species has megasporangiate and microsporangiate flowers on the same plant. A **dioecious** species is one in which megasporangiate and microsporangiate flowers occur on different plants.
- Flowers may be pollinated by wind or by animals. Many angiosperms have coevolved with their animal pollinators.
- Nearly all angiosperms exhibit **double fertilization**, resulting in the production of a diploid zygote and an **endosperm** (which is triploid in most species). **Review Figure 21.25 and ANIMATED TUTORIAL 21.3**
- The oldest evolutionary split among the angiosperms is between the clade represented by the single species in the genus *Amborella* and all the remaining flowering plants. **Review Figure 21.27**
- The **magnoliids** are the sister group to the monocots and eudicots. The most species-rich angiosperm clades are the **monocots** and the **eudicots**.

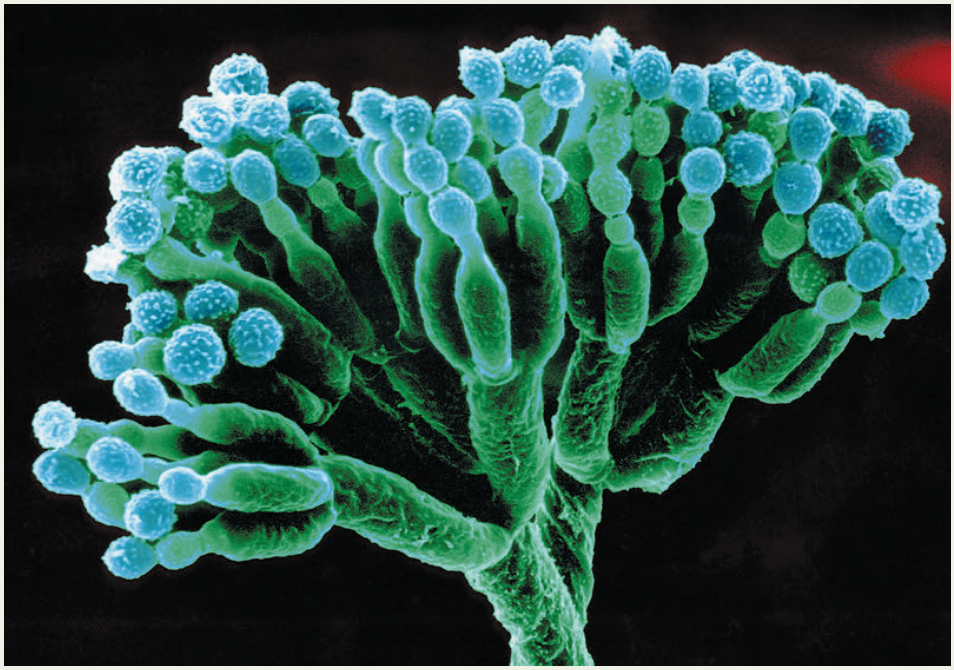
Go to the Interactive Summary to review key figures, Animated Tutorials, and Activities
PoL2e.com/is21

Go to LaunchPad at **macmillanhighered.com/launchpad** for additional resources, including LearningCurve Quizzes, Flashcards, and many other study and review resources.

22

The Evolution and Diversity of Fungi

KEY CONCEPTS



All species of the fungus *Penicillium* are recognizable by their dense spore-bearing structures (see Figure 22.18).

Alexander Fleming was already a famous scientist in 1928, but his laboratory was often a mess. That year he was studying the properties of *Staphylococcus* bacteria, the agents of dangerous staph infections. In August he took a long vacation with his family. When he returned in early September, he found that some of his petri dishes of *Staphylococcus* had become infested with a fungus that killed many of the bacteria.

Many scientists would have sighed at the loss, thrown out the petri dishes, and started new cultures of bacteria. But when Fleming looked at the dishes, he saw something exciting. Around each colony of fungi was a ring within which all the bacteria were dead.

Fleming hypothesized that the bacteria-free rings around the fungal colonies were produced by a substance excreted from the fungi, which he initially called "mould juice." He identified the fungi as members of the genus *Penicillium* and eventually named the antibacterial substance produced by these fungi penicillin. Fleming published his discovery in 1929, but initially the finding received very little attention.

Over the next decade, Fleming produced small quantities of penicillin for testing as an antibacterial agent. Some of the tests showed promise, but many were inconclusive, and eventually Fleming gave up on the research. But his tests had shown enough promise to attract the attention of several chemists, who worked out the practical problems of producing a stable form of the substance. Clinical trials of this stable form of penicillin were extremely successful, and by 1945 it was being produced and distributed as an antibiotic on a large scale. That same year, Fleming and two of the chemists, Howard Florey and Ernst Chain, won the Nobel Prize in Medicine for their work on penicillin.

The development of penicillin was one of the most important achievements in modern medicine. Until the introduction of modern antibiotics, the most widespread agents of human death included bacterial infections such as gangrene, tuberculosis, and syphilis. Penicillin proved to be highly effective in curing such infections, and its success led to the creation of the modern pharmaceutical industry. Soon many additional antibiotic compounds were isolated from other fungi or synthesized in the laboratory, leading to a "golden age" of human health.

Q Have antibiotics derived from fungi eliminated the danger of bacterial diseases in human populations?

You will find the answer to this question on page 467.

CONCEPT 22.1 Fungi Live by Absorptive Heterotrophy

Fungi are organisms that digest their food outside their bodies. They secrete digestive enzymes to break down large food molecules in the environment, then absorb the breakdown products through their cell membranes in a process known as **absorptive heterotrophy**. This mode of nutrition is successful in a wide variety of environments. Many fungi are **saprobes**, meaning they absorb nutrients from dead organic matter. Others are parasites, absorbing nutrients from living hosts. Still others are mutualists, living in intimate associations with other organisms that benefit both partners.

Modern fungi are believed to have evolved from a unicellular protist ancestor that had a flagellum. The probable common ancestor of the animals was also a flagellated protist much like the living choanoflagellates (see Figure 23.2). Current evidence, including the sequences of many genes, suggests that the fungi, choanoflagellates, and animals share a common ancestor not shared by other eukaryotes. These three lineages are often grouped together as the **opisthokonts** (FIGURE 22.1). A synapomorphy (shared derived trait) of the opisthokonts is a flagellum that, if present, is posterior, as in animal sperm. The flagella of all other eukaryotes are anterior.

Synapomorphies that distinguish the fungi as a group among the opisthokonts include absorptive heterotrophy and the presence of **chitin**, a structural polysaccharide, in their cell walls. The fungi represent one of four independent evolutionary origins of large multicellular organisms (plants, brown algae, and animals are the other three).

Unicellular yeasts absorb nutrients directly

Most fungi are multicellular, but single-celled species are found in most fungal groups. Unicellular, free-living fungi are referred to as **yeasts** (FIGURE 22.2). Some fungi that have yeast life stages also have multicellular life stages. Thus the term "yeast" does not refer to a single taxonomic group, but rather to a lifestyle that has evolved multiple times. Yeasts live in liquid or moist environments and absorb nutrients directly across their cell surfaces.

The ease with which many yeasts can be cultured, combined with their rapid growth rates, has made them ideal model organisms for study in the laboratory. They present many of the same advantages to laboratory investigators as do many bacteria, but because they are eukaryotes, their genome structures and cells are much more like those of humans and other eukaryotes than are those of bacteria.

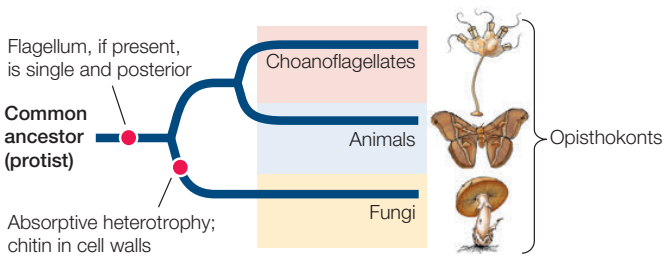


FIGURE 22.1 Fungi in Evolutionary Context Absorptive heterotrophy and the presence of chitin in their cell walls distinguish the fungi from other opisthokonts.

Saccharomyces cerevisiae

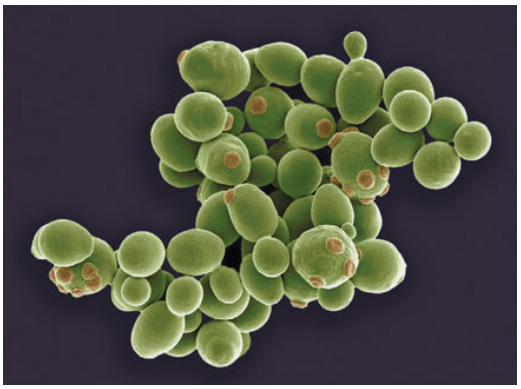


FIGURE 22.2 Yeasts Unicellular, free-living fungi are known as yeasts. Many yeasts reproduce by budding—mitosis followed by asymmetrical cell division—as those shown here are doing.

LINK

You can read about other model organisms in **Concept 12.3**

Multicellular fungi use hyphae to absorb nutrients

The body of a multicellular fungus is called a **mycelium** (plural *mycelia*). A mycelium is composed of a mass of individual tubular filaments called **hyphae** (singular *hypha*; FIGURE 22.3A), in which absorption of nutrients takes place. The cell walls of the hyphae are greatly strengthened by microscopic strands of a complex polysaccharide called chitin. In some species of fungi, the hyphae are subdivided into cell-like compartments by incomplete cross-walls called **septa** (singular *septum*). These subdivided hyphae are referred to as **septate hyphae**. Septa do not completely close off compartments in the hyphae. Pores at the centers of the septa allow organelles—sometimes even nuclei—to move in a controlled way between compartments (FIGURE 22.3B). In other species of fungi, the hyphae lack septa but may contain hundreds of nuclei. These multinucleate, undivided hyphae are referred to as **coenocytic**. The coenocytic condition results from repeated nuclear divisions without cytokinesis.

Certain modified hyphae, called **rhizoids**, anchor some fungi to their substrate (i.e., the dead organism or other matter on which they are feeding). These rhizoids are not homologous to the rhizoids of plants, and they are not specialized to absorb nutrients and water.

Fungi can grow very rapidly when conditions are favorable. In some species, the total hyphal growth of a fungal mycelium (not the growth of an individual hypha) may exceed 1 kilometer a day! The hyphae may be widely dispersed to forage for nutrients over a large area, or they may clump together in

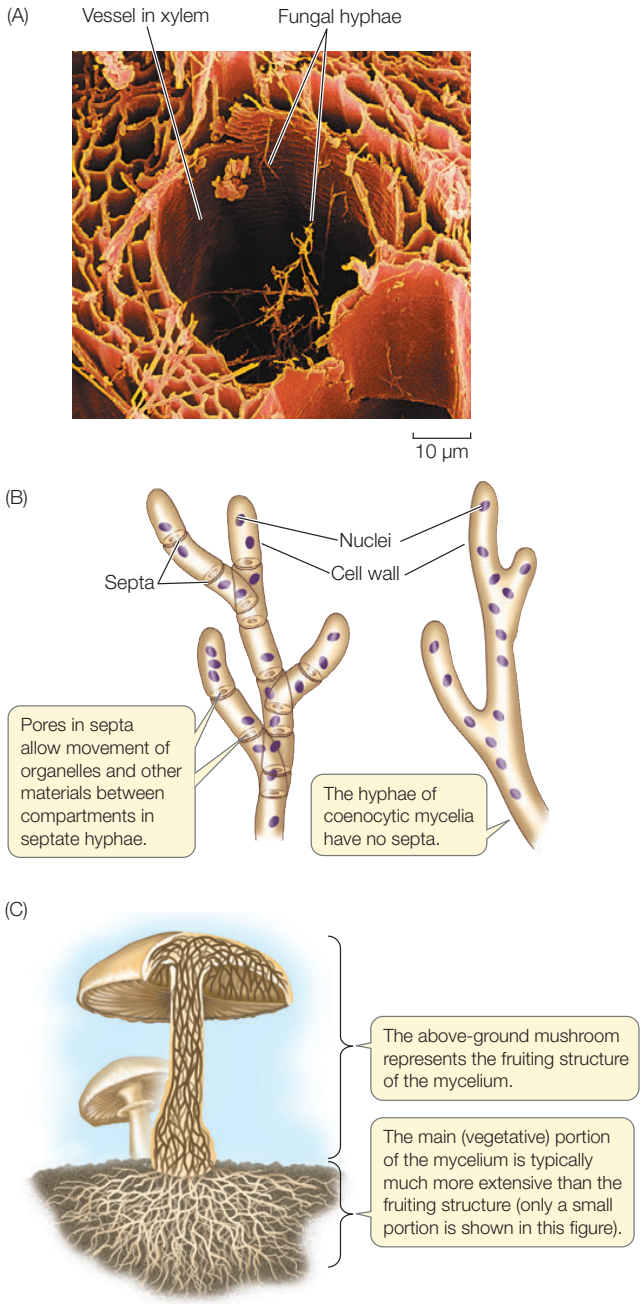


FIGURE 22.3 Mycelia Are Made Up of Hyphae (A) The minute individual hyphae of fungal mycelia can penetrate small spaces. In this artificially colored micrograph, hyphae (yellow structures) of a dry-rot fungus are penetrating the xylem tissues of a log. (B) The hyphae of septate fungal species are divided into organelle-containing compartments by porous septa, whereas coenocytic hyphae have no septa. (C) The fruiting structure of a club fungus is short-lived, but the filamentous, nutrient-absorbing mycelium can be long-lived and cover a large area.

a cottony mass to exploit a rich nutrient source. The familiar mushrooms you may notice growing in moist areas are spore-producing fruiting structures (FIGURE 22.3C). In the fungal species that produce these structures, the mycelial mass is often far larger than the visible mushroom. The mycelium of one individual fungus discovered in Oregon covers almost 900 hectares underground and weighs considerably more than a blue whale (the largest animal). Aboveground, this individual is evident only as isolated clumps of mushrooms.

Fungi are in intimate contact with their environment

The filamentous hyphae of a fungus give it a unique relationship with its physical environment. The fungal mycelium has an enormous surface area-to-volume ratio compared with that of most large multicellular organisms. This large ratio is a marvelous adaptation for absorptive heterotrophy. Throughout the mycelium (except in fruiting structures), all of the hyphae are very close to their food source.

The downside of the large surface area-to-volume ratio of the mycelium is its tendency to lose water rapidly in a dry environment. Thus fungi are most common in moist environments. You have probably observed the tendency of molds, toadstools, and other fungi to appear in damp places.

Another characteristic of some fungi is a tolerance for highly hypertonic environments (those with a solute concentration higher than their own; see Concept 5.2). Many fungi are more resilient than bacteria in hypertonic surroundings. Jelly in the refrigerator, for example, will not become a growth medium for bacteria because it is too hypertonic to those organisms, but it may eventually harbor mold colonies. Mold in the refrigerator illustrates yet another trait of many fungi: tolerance of temperature extremes. Many fungi grow in temperatures as low as –6°C, and some tolerate temperatures higher than 50°C.

CHECKpoint CONCEPT 22.1

- ✓ Describe the relationship between fungal structure and absorptive heterotrophy.
- ✓ What are the advantages and disadvantages to multicellular fungi of the large surface-to-volume ratio of the mycelium?

Fungi are important components of healthy ecosystems. They interact with other organisms in many ways, some of which are harmful and some beneficial to those other organisms.

CONCEPT 22.2 Fungi Can Be Saprobic, Parasitic, Predatory, or Mutualistic

Without the fungi, our planet would be very different. Picture Earth with only a few stunted plants and watery environments choked with the remains of dead organisms. Fungi do much of Earth's garbage disposal. Fungi not only help clean up the landscape and form soil, but also play a key role in recycling mineral nutrients. Furthermore, the colonization of the terrestrial environment was made possible in large part by associations that fungi formed with land plants and other organisms.

Saprobic fungi are critical to the planetary carbon cycle

Saprobic fungi, along with bacteria, are the major decomposers on Earth, contributing to the decay of nonliving organic matter and thus to recycling of the elements used by living things. In forests, for example, the mycelia of fungi absorb nutrients from fallen trees, thus decomposing their wood. Fungi are the principal decomposers of cellulose and lignin, the main components of plant cell walls (most bacteria cannot break down these materials). Other fungi produce enzymes that decompose keratin and thus break down animal structures such as hair and nails.

Were it not for the fungal decomposers, Earth's carbon cycle would fail. Great quantities of carbon atoms would remain trapped forever on forest floors and elsewhere. Instead, those carbon atoms are returned to the atmosphere in the form of CO_2 by fungal respiration, where they are again available for photosynthesis by plants.

LINK

Earth's carbon cycle is described in **Concept 45.3**

There was a time in Earth's history when populations of saprobic fungi declined dramatically. Vast tropical swamps existed during the Carboniferous period, as we saw in Chapter 21. When plants in these swamps died, they began to form peat. Peat formation led to acidification of the swamps. That acidity, in turn, drastically reduced the fungal population. The result? With the decomposers largely absent, large quantities of peat remained on the swamp floor and over time were converted into coal.

In contrast to their decline during the Carboniferous, fungi did very well at the end of the Permian, a quarter of a billion years ago, when the aggregation of continents produced volcanic eruptions that triggered a global mass extinction (see Chapter 18). The fossil record shows that even as 96 percent of all multicellular species became extinct, fungi flourished—demonstrating both their hardiness and their role in recycling the elements in dead plants and animals.

Simple sugars and the breakdown products of complex polysaccharides are the favored source of carbon for saprobic fungi. Most fungi obtain nitrogen from proteins or the products of protein breakdown. Many fungi can use nitrate (NO_3^-) or ammonium (NH_4^+) ions as their sole source of nitrogen. No known fungus can get its nitrogen directly from inorganic nitrogen gas, however, as can some bacteria and plant–bacteria associations (that is, fungi cannot fix nitrogen; see Concept 19.3).

What happens when a fungus faces a dwindling food supply? A common strategy is to reproduce rapidly and abundantly. When conditions are good, fungi produce great quantities of spores, but the rate of spore production is commonly even higher when nutrient supplies go down. The spores may then remain dormant until conditions improve, or they may be dispersed to areas where nutrient supplies are higher.

Not only are fungal spores abundant in number, but they are extremely tiny and easily spread by wind or water (**FIGURE 22.4**). These attributes virtually ensure that they will be scattered over great distances, and that at least some of them will find conditions suitable for growth. The air we breathe contains as many as 10,000 fungal spores per cubic meter. No wonder we find fungi just about everywhere.

Lycoperdon perlatum

FIGURE 22.4 Spores Galore Puffballs (a type of club fungus) disperse trillions of spores in great bursts. Few of the spores travel very far, however; some 99 percent of them fall within 100 meters of the parent puffball.

Go to MEDIA CLIP 22.1
Fungal Decomposers
PoL2e.com/mc22.1

Some fungi engage in parasitic or predatory interactions

Whereas saprobic fungi obtain their energy, carbon, and nitrogen directly from dead organic matter, other species of fungi obtain their nutrition from parasitic—and even predatory—interactions.

PARASITIC FUNGI Mycologists (biologists who study fungi) distinguish between two classes of parasitic fungi based on their degree of dependence on their host. **Facultative parasites** can grow on living organisms but can also grow independently (including on artificial media). **Obligate parasites** can grow only on a living host. The fact that their growth depends on a living host shows that obligate parasites have specialized nutritional requirements.

Plants and insects are the most common hosts of parasitic fungi. The filamentous structure of fungal hyphae is especially well suited to a life of absorbing nutrients from living plants. The slender hyphae of a parasitic fungus can invade a plant through

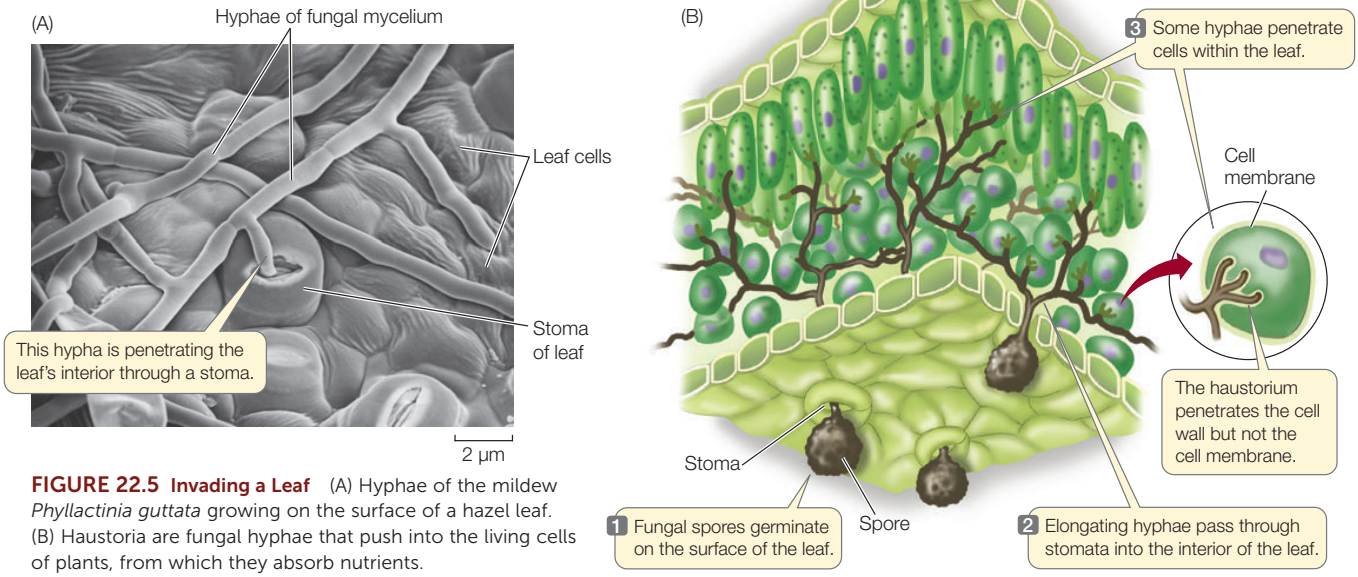


FIGURE 22.5 Invading a Leaf (A) Hyphae of the mildew *Phyllactinia guttata* growing on the surface of a hazel leaf. (B) Haustoria are fungal hyphae that push into the living cells of plants, from which they absorb nutrients.

stomata, through wounds, or in some cases, by direct penetration of epidermal cell walls (**FIGURE 22.5A**). Once inside the plant, the hyphae branch out to expand the mycelium. Some hyphae produce **haustoria** (singular *haustorium*), branching projections that push through cell walls into living plant cells, absorbing the nutrients within those cells. The haustoria do not break through the cell membranes inside the cell walls, but instead push pockets into the membranes, so that the cell membrane fits them like a glove (**FIGURE 22.5B**). Fruiting structures may form, either within the plant body or on its surface.

Some parasitic fungi live in a close physical (symbiotic) relationship with their host that is usually not lethal to the plant. Others are pathogenic, sickening or even killing the host from which they derive nutrition.

Go to MEDIA CLIP 22.2
Mind-Control Killer Fungi
PoL2e.com/mc22.2

PATHOGENIC FUNGI Although most human diseases are caused by bacteria or viruses, fungal pathogens are a major cause of death among people with compromised immune systems. Many people with AIDS die of fungal diseases, such as the pneumonia caused by *Pneumocystis jirovecii*. Even *Candida albicans* and certain other yeasts that are normally a part of a healthy microbiome can cause severe diseases, such as esophagitis (which impairs swallowing), in individuals with AIDS and in individuals taking immunosuppressive drugs. Various fungi cause other, less threatening human diseases, such as ringworm and athlete's foot. Our limited understanding of the basic biology of these fungi still hampers our ability to treat the diseases they cause. As a result, fungal diseases are a growing international health problem.

The worldwide decline of amphibian species has been linked to the spread of a chytrid fungus, *Batrachochytrium dendrobatidis* (or *Bd* for simplicity). In some areas of the world where this fungus has been present for millennia, amphibian populations appear to be tolerant of *Bd* and it does not cause widespread die-offs. But in other areas, such as western North America and Australia, introductions of *Bd* are implicated in the loss or decline of many frog and salamander populations, or the extinction of entire species. Some strains of *Bd* are native to southern Africa, where they do not appear to cause widespread amphibian declines. Biologists have hypothesized that the spread of *Bd* around the world may have been initiated in the 1930s with exports of the African clawed frog (*Xenopus laevis*), which was once widely used in human pregnancy tests. Recent studies by Erica Rosenblum and her colleagues have shown that Africa may indeed have been one source of the introductions, but that strains of *Bd* appear to have been introduced multiple times from several geographic areas (**FIGURE 22.6**).

Fungi are by far the most important plant pathogens, causing crop losses amounting to billions of dollars. Bacteria and viruses are less important than fungi as plant pathogens. Major fungal diseases of crop plants include black stem rust of wheat and other diseases of wheat, corn, and oats. The agent of black stem rust is *Puccinia graminis*, which has a complicated life cycle that involves two plant hosts (wheat and barberry). In an epidemic in 1935, *P. graminis* was responsible for the loss of about one-fourth of the wheat crop in Canada and the United States.

Although many pathogenic fungi cause problems when they attack agricultural crops, pathogenic fungi of the genus *Fusarium* can benefit agriculture by killing certain weed species, such as witchweed, a serious pest of cereal crops. Another strain of *Fusarium* has been proposed as a tool in the war against cocaine. The fungus could be applied to kill coca plants, the source of the drug. The use of *Fusarium* to kill coca plants is highly controversial, however, as the fungus might infect neighboring plant species in addition to the coca plants.

PREDATORY FUNGI Some fungi have adaptations that enable them to function as active predators, trapping nearby microscopic protists or animals. The most common predatory

INVESTIGATION

FIGURE 22.6 What is the Origin of Amphibian-Killing Chytrids in Western North America? An investigation by Erica Rosenblum and colleagues of *Batrachochytrium dendrobatidis* samples from around the world suggests that this pathogenic fungus was introduced to western North America several times, from different geographic localities.[a]

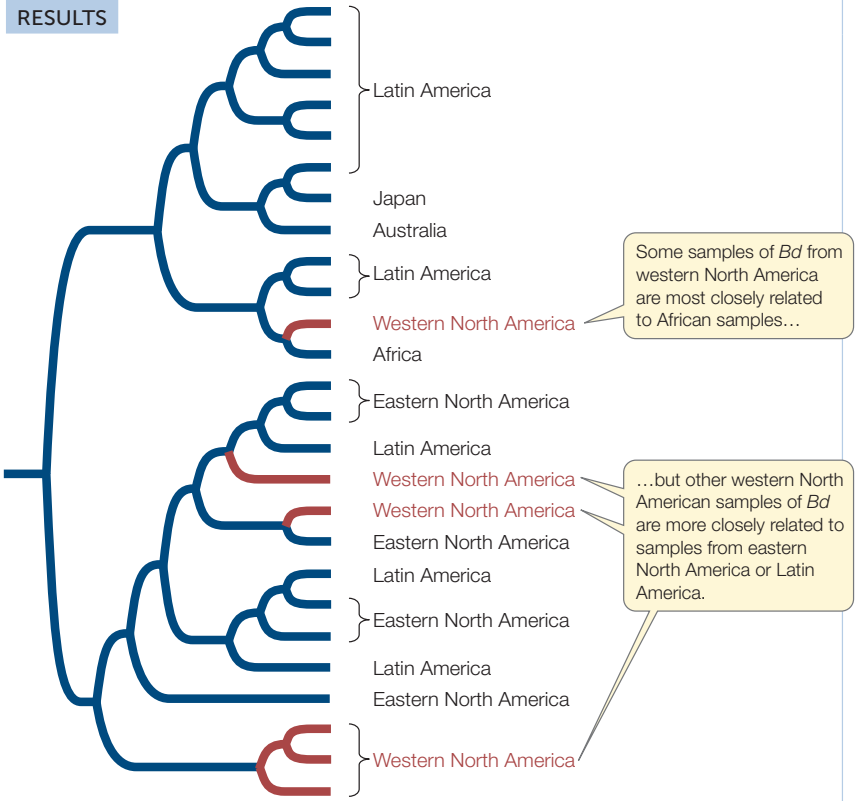
HYPOTHESIS

Pathogenic forms of *Bd* that kill many amphibians were introduced into western North America from Africa, through importation of African clawed frogs beginning in the 1930s.

METHOD

1. Collect samples of *Bd* from throughout the world.
2. Sequence the genomes of the sampled chytrids.
3. Construct a phylogenetic tree by comparing the DNA sequences of the different samples of *Bd*.
4. Assess whether or not the resulting tree is consistent with the introduction of *Bd* from Africa to western North America.

RESULTS



CONCLUSION

Pathogenic strains of *Bd* appear to have been introduced to western North America from several different sources. At least one strain of *Bd* is consistent with an introduction from Africa, but other strains were more likely introduced from eastern North America or Latin America.

Go to **LaunchPad** for discussion and relevant links for all **INVESTIGATION** figures.

[a]E. B. Rosenblum et al. 2013. *Proceedings of the National Academy of Sciences USA* 110: 9385–9390.

strategy seen in fungi is to secrete sticky substances from the hyphae so that passing organisms stick to them. The hyphae then quickly invade the trapped prey, growing and branching within it, spreading through its body, absorbing nutrients, and eventually killing it.

A more dramatic adaptation for predation is the constricting ring formed by some species of soil fungi (**FIGURE 22.7**). When nematodes (tiny roundworms) are present in the soil, these fungi form three-celled rings with a diameter that just fits a nematode. A nematode crawling through one of these rings stimulates the fungus, causing the cells of the ring to swell and trap the worm. Fungal hyphae quickly invade and digest the unlucky victim.

Mutualistic fungi engage in relationships beneficial to both partners

Certain kinds of relationships between fungi and other organisms have nutritional consequences for both partners. Two relationships of this type are highly specific and are **symbiotic** (the partners live in close, permanent contact with each other) as well as **mutualistic** (the relationship benefits both partners).

LINK

The obligate mutualistic relationship between leaf-cutter ants and the fungus they cultivate is necessary for the survival of each partner; see the opener of **Chapter 43 and Concepts 43.3 and 43.4**

LICHENS A **lichen** is not a single organism, but rather a symbiosis between two radically different species: a fungus and a photosynthetic microorganism. Together the organisms that constitute a lichen can survive some of the harshest environments on Earth (although they are sensitive to poor air quality; see Concept 22.4). The biota of Antarctica, for example, features more than 100 times as many species of lichens as of vascular plants. Relatively little experimental work has focused on lichens, perhaps because they grow so slowly—typically less than 1 centimeter in a year.

There are nearly 20,000 described "species" of lichens, each of which is assigned the name of its fungal component. These fungal components may constitute as many as 20 percent of all fungal species. Most of them are sac fungi (Ascomycota). Some of them are able to grow independently without a photosynthetic partner, but most have never been observed in nature other than in a lichen association. The photosynthetic component of a lichen is most often a unicellular green alga, but it can be a cyanobacterium, or may even include both.

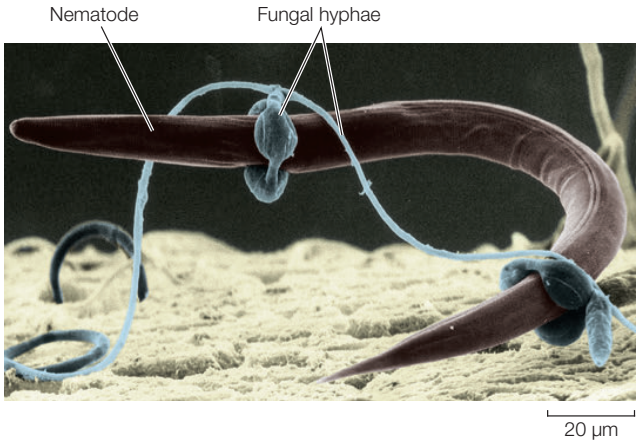


FIGURE 22.7 Fungus as Predator A nematode is trapped by hyphal rings of the soil-dwelling fungus *Arthrobotrys dactyloides*.

Lichens are found in all sorts of exposed habitats: on tree bark, on open soil, and on bare rock. Reindeer moss (not a moss at all, but the lichen *Cladonia subtenuis*) covers vast areas in Arctic, sub-Arctic, and boreal regions, where it is an important part of the diets of reindeer and other large mammals.

The body forms of lichens fall into three principal categories: **Crustose** (crustlike) lichens look like colored powder dusted over their substrate (**FIGURE 22.8A**). **Foliose** (leafy) lichens are more loosely attached and have large leaflike forms, whereas **fruticose** (shrubby) lichens are highly branched and grow upward like little shrubs, or hang in long strands from tree branches or rocks (**FIGURE 22.8B**).

A cross section of a typical foliose lichen reveals a tight upper region of fungal hyphae, a layer of photosynthetic cyanobacteria or algae, a looser hyphal layer, and finally hyphal rhizoids that attach the entire structure to its substrate (**FIGURE 22.9**). The meshwork of fungal hyphae takes up some mineral nutrients needed by the photosynthetic cells and also holds water tenaciously, providing a suitably moist environment. The fungi obtain fixed carbon from the photosynthetic products of the algal or cyanobacterial cells.

Within the lichen, the fungal hyphae are tightly pressed against the algal or cyanobacterial cells and sometimes even invade them without breaching the cell membrane (similar to the haustoria in parasitic fungi; see Figure 22.5). The photosynthetic cells not only survive these intrusions but continue to grow. Algal cells in a lichen "leak" photosynthetic products at a greater rate than do similar cells growing on their own, and photosynthetic cells taken from lichens grow more rapidly on their own than when associated with a fungus. On the basis of these observations, we could consider lichen fungi to be parasitic on their photosynthetic partners. In many places where lichens grow, however, the photosynthetic cells could not grow at all on their own.

Lichens can reproduce simply by fragmentation of the vegetative body (the **thallus**) or by means of specialized structures called **soredia** (singular *soredium*). Soredia consist of one or a few photosynthetic cells bound by fungal hyphae. They become detached from the lichen, are dispersed by air currents,

FIGURE 22.8 Lichen Body Forms Lichens fall into three principal categories—crustose, fruticose, or foliose—based on their body form. (A) A crustose lichen growing on the surface of an exposed rock. (B) Fruticose lichens appear "shrubby," whereas foliose lichens look "leafy."

and upon arriving at a favorable location, develop into a new lichen thallus. Alternatively, the fungal partner may go through its sexual reproductive cycle, producing haploid spores. When these spores are discharged, however, they disperse alone, unaccompanied by the photosynthetic partner.

Lichens are often the first colonists on new areas of bare rock. They get most of the mineral nutrients they need from the air and rainwater, augmented by minerals absorbed from dust. A lichen begins to grow shortly after a rain, as it begins to dry. As it grows, the lichen acidifies its environment slightly, and this acidity contributes to the slow breakdown of rocks, an early step in soil formation. With further drying, the lichen's photosynthesis ceases. The water content of the lichen may drop to less than 10 percent of its dry weight, at which point it becomes highly insensitive to extremes of temperature.

MYCORRHIZAE Many vascular plants depend on a symbiotic association with fungi. This ancient association between

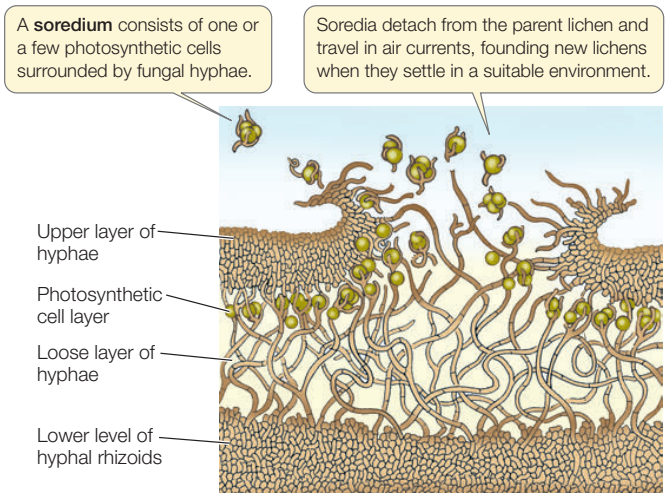


FIGURE 22.9 Lichen Anatomy Cross section showing the layers of a foliose lichen and the release of soredia.

plants and fungi was critical to the successful exploitation of the terrestrial environment by plants. Unassisted, the root hairs of many plants often do not take up enough water or minerals to sustain growth. However, the roots of such plants usually become infected with fungi, forming an association called a **mycorrhiza** (plural *mycorrhizae*). There are two types of mycorrhizae, distinguished by whether or not the fungal hyphae penetrate the plant cell walls.

In **ectomycorrhizae**, the fungus wraps around the root, and its mass is often as great as that of the root itself (**FIGURE 22.10A**). The fungal hyphae wrap around individual cells in the root but do not penetrate the cell walls. An extensive web of hyphae penetrates the soil in the area around the root, so that up to 25 percent of the soil volume near the root may be fungal hyphae. The hyphae attached to the root increase the surface area for the absorption of water and minerals, and the mass of hyphae in the soil acts like a sponge to hold water in the neighborhood of the root. Infected roots are short, swollen, and club-shaped, and they lack root hairs.

The fungal hyphae of **arbuscular mycorrhizae** enter the root and penetrate the cell walls of the root cells, forming arbuscular (treelike) structures inside the cell wall but outside the cell membrane. These structures, like the haustoria of parasitic fungi and the contact regions of fungal hyphae and photosynthetic cells in lichens, become the primary site of exchange between plant and fungus (**FIGURE 22.10B**). As in the ectomycorrhizae, the fungus forms a vast web of hyphae leading from the root surface into the surrounding soil.

LINK

Concept 25.2 contains a detailed description of the role of arbuscular mycorrhizae in plant nutrition

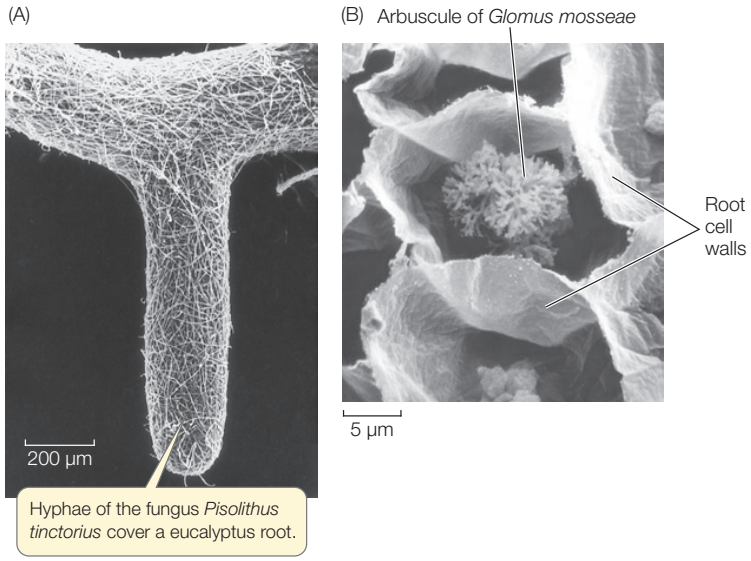


FIGURE 22.10 Mycorrhizal Associations (A) Ectomycorrhizal fungi wrap themselves around a plant root, increasing the area available for absorption of water and minerals. (B) Hyphae of arbuscular mycorrhizal fungi infect the root internally and penetrate the root cell walls, branching within the cells and forming a treelike structure, the arbuscule. (For purposes of this scanning electron micrograph, the cell cytoplasm was removed to better visualize the arbuscule.)

The mycorrhizal association is important to both partners. The fungus obtains needed organic compounds, such as sugars and amino acids, from the plant. In return, the fungus, because of its very high surface area-to-volume ratio and its ability to penetrate the fine structure of the soil, greatly increases the plant's ability to absorb water and minerals (especially phosphorus). The fungus may also provide the plant with certain growth hormones and may protect it against attack by disease-causing microorganisms.

Plants that have active arbuscular mycorrhizae typically are a deeper green and may resist drought and temperature extremes better than plants of the same species that have little mycorrhizal development. Attempts to introduce some plant species to new areas have failed until a bit of soil from the native area (presumably containing the fungus necessary to establish mycorrhizae) was provided. Trees without ectomycorrhizae do not grow well in the absence of abundant nutrients and water, so the health of our forests depends on the presence of ectomycorrhizal fungi. Many agricultural crops require inoculation of seeds with appropriate mycorrhizal fungi prior to planting. Without these fungi, the plants are unlikely to grow well, or in some cases at all. Certain plants that live in nitrogen-poor habitats, such as cranberry bushes and orchids, invariably have mycorrhizae. Orchid seeds will not germinate in nature unless they are already infected by the fungus that will form their mycorrhizae. Plants that lack chlorophyll always have mycorrhizae, which they often share with the roots of green, photosynthetic plants. In effect, these plants without chlorophyll are feeding on nearby green plants, using the fungus as a bridge.

Endophytic fungi protect some plants from pathogens, herbivores, and stress

In a tropical rainforest, 10,000 or more fungal spores may land on a single leaf each day. Some are plant pathogens, some do not affect the plant at all, and some invade the plant in a beneficial way. Fungi that live within aboveground parts of plants without causing obvious deleterious symptoms are called **endophytic fungi**. Recent research has shown that endophytic fungi are abundant in plants in all terrestrial environments.

Among the grasses, individual plants with endophytic fungi are more resistant to pathogens and to insect and mammalian herbivores than are plants lacking endophytes. The fungi produce alkaloids (nitrogen-containing compounds) that are toxic to animals. The alkaloids do not harm the host plant. In fact, some plants produce alkaloids (such as nicotine) themselves. The fungal alkaloids also increase the ability of grasses to resist stress of various types, including drought (water shortage) and salty soils. Such resistance is useful in agriculture.

The role, if any, of endophytic fungi in most broad-leaved plants is unclear. They may convey protection against pathogens, or they may simply occupy space within leaves without conferring any benefit, but also without doing harm. The benefit, in fact, might be all for the fungus.

CHECKpoint CONCEPT 22.2

- ✓ What is the role of fungi in Earth's carbon cycle?
- ✓ Describe the nature and benefits of the lichen association.
- ✓ Why do plants grow better when infected with mycorrhizal fungi?

Before molecular techniques clarified the phylogenetic relationships of fungi, one criterion used for assigning fungi to taxonomic groups was the nature of their life cycles—including the types of fruiting bodies they produce. The next section will take a closer look at life cycles in the six major groups of fungi.

CONCEPT 22.3 Major Groups of Fungi Differ in Their Life Cycles

Major fungal groups were originally defined by their structures and processes for sexual reproduction and also, to a lesser extent, by other morphological differences. Although fungal life cycles are even more diverse than was once realized, specific types of life cycles generally distinguish the six major groups of fungi: microsporidia, chytrids, zygospore fungi (Zygomycota), arbuscular mycorrhizal fungi (Glomeromycota), sac fungi (Ascomycota), and club fungi (Basidiomycota). FIGURE 22.11 diagrams the evolutionary relationships of these groups as they are understood today.

The chytrids and the zygospore fungi may not represent monophyletic groups, as they each consist of several distantly related lineages that retain some ancestral features. The clades

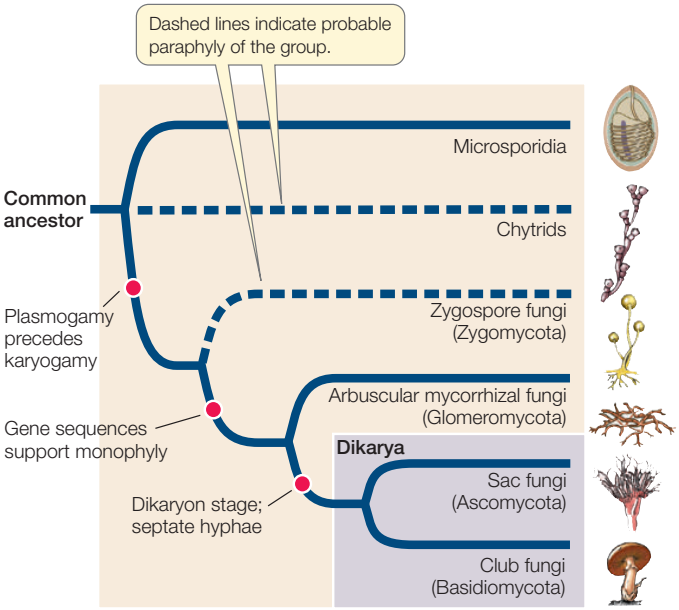


FIGURE 22.11 A Phylogeny of the Fungi Microsporidia are reduced, parasitic fungi whose relationships among the fungi are uncertain. They may be the sister group of most other fungi or more closely related to particular groups of chytrids or zygospore fungi. The dashed lines indicate that chytrids and zygospore fungi are thought to be paraphyletic; the relationships of the lineages within these two informal groups (see Table 22.1) are not yet well resolved. The sac fungi and club fungi together form the clade Dikarya.

Go to ACTIVITY 22.1 Fungal Phylogeny
PoL2e.com/ac22.1

that are thought to be monophyletic within these two informal groupings are listed in TABLE 22.1. Recent evidence from DNA analyses has established the placement of the microsporidia among the fungi, the likely paraphyly of the chytrids and the zygospore fungi (in other words, members of these groups are probably not each other's closest relatives), the independence of arbuscular mycorrhizal fungi from the other fungal groups, and the monophyly of sac fungi and club fungi.

Fungi reproduce both sexually and asexually

Both asexual and sexual reproduction occur among the fungi (FIGURE 22.12). Asexual reproduction takes several forms:

- The production of (usually) haploid spores within sporangia
- The production of haploid spores (not enclosed in sporangia) at the tips of hyphae; such spores are called **conidia** (Greek *konis*, "dust")
- Cell division by unicellular fungi—either a relatively equal division of one cell into two (fission) or an asymmetrical division in which a smaller daughter cell is produced (budding)
- Simple breakage of the mycelium

Sexual reproduction is rare (or even unknown) in some groups of fungi but common in others. Sexual reproduction may not