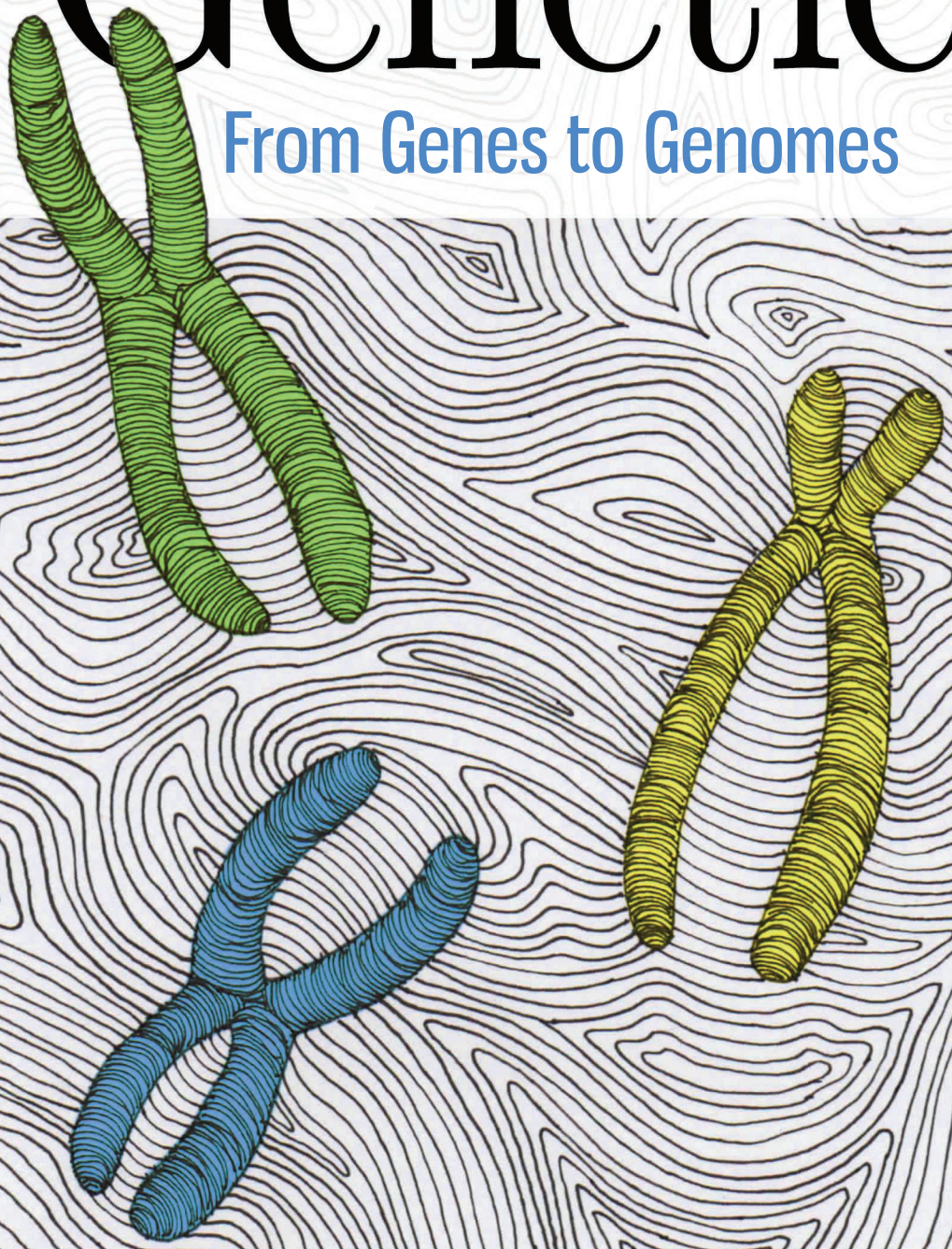
SIXTH EDITION

Genetics

From Genes to Genomes



HARTWELL ~ GOLDBERG ~ FISCHER ~ HOOD

Mc Graw Hill Education

SIXTH EDITION

Genetics

From Genes to Genomes

Leland H. Hartwell
FRED HUTCHISON CANCER CENTER

Michael L. Goldberg
CORNELL UNIVERSITY

Janice A. Fischer
UNIVERSITY OF TEXAS AT AUSTIN

Leroy Hood
THE INSTITUTE FOR SYSTEMS BIOLOGY

Mc Graw Hill Education

GENETICS: FROM GENES TO GENOMES, SIXTH EDITION

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 LWI 21 20 19 18 17

ISBN 978–1–259–70090–3
MHID 1–259–70090–9

Portfolio Manager: *Justin K.Wyatt, PhD*
Lead Product Developer: *Fran Simon*
Senior Marketing Manager: *Kelly Brown*
Senior Content Project Manager: *Vicki Krug*
Lead Content Project Manager: *Christina Nelson*
Buyer: *Jennifer Pickel*
Lead Designer: *David Hash*
Lead Content Licensing Specialist: *Lori Hancock*
Content Licensing Specialist: *Lori Slattery*
Cover Image: © *Sydney F. Vize*
Design Element: *Front and Back Matter DNA helix header, © MedicalRF.com*
Compositor: *Aptara, Inc*
Printer: *LSC Communications*

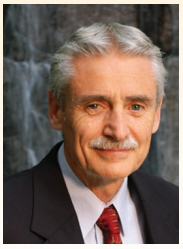
All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

Library of Congress Cataloging-in-Publication Data

Names: Hartwell, Leland, author.
Title: Genetics : from genes to genomes / Leland Hartwell, Michael Goldberg, Janice Fischer, Lee Hood, Charles F. Aquadro, Bruce Bejcek.
Description: Sixth edition. | New York, NY : McGraw-Hill Education, [2018] | Includes index.
Identifiers: LCCN 2017017175| ISBN 9781259700903 (alk. paper) | ISBN 1259700909 (alk. paper)
Subjects: LCSH: Genetics.
Classification: LCC QH430 .G458 2018 | DDC 576.5—dc23 LC record available at https://lccn.loc.gov/2017017175

mheducation.com/highered

About the Authors



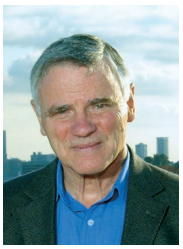
Dr. Leland Hartwell is President and Director of Seattle's Fred Hutchinson Cancer Research Center and Professor of Genome Sciences at the University of Washington.

Dr. Hartwell's primary research contributions were in identifying genes that control cell division in yeast, including those necessary for the division process as well as those necessary for the fidelity of genome reproduction. Subsequently, many of these same genes have been found to control cell division in humans and often to be the site of alteration in cancer cells.

Dr. Hartwell is a member of the National Academy of Sciences and has received the Albert Lasker Basic Medical Research Award, the Gairdner Foundation International Award, the Genetics Society Medal, and the 2001 Nobel Prize in Physiology or Medicine.

Dr. Michael Goldberg is a Professor at Cornell University, where he teaches introductory genetics and human genetics. He was an undergraduate at Yale University and received his Ph.D. in biochemistry from Stanford University. Dr. Goldberg performed postdoctoral research at the Biozentrum of the University of Basel (Switzerland) and at Harvard University, and he received an NIH Fogarty Senior International Fellowship for study at Imperial College (England) and fellowships from the Fondazione Cenci Bolognetti for sabbatical work at the University of Rome (Italy). His current research uses the tools of *Drosophila* genetics and the biochemical analysis of frog egg cell extracts to investigate the mechanisms that ensure proper cell cycle progression and chromosome segregation during mitosis and meiosis.

Dr. Janice Fischer is a Professor at The University of Texas at Austin, where she is an award-winning teacher of genetics and Director of the Biology Instructional Office. She received her Ph.D. in biochemistry and molecular biology from Harvard University, and did postdoctoral research at The University of California at Berkeley and The Whitehead Institute at MIT. In her research, Dr. Fischer used *Drosophila* first to determine how tissue-specific transcription works, and then to examine the roles of ubiquitin and endocytosis in cell signaling during development.



Dr. Lee Hood received an M.D. from the Johns Hopkins Medical School and a Ph.D. in biochemistry from the California Institute of Technology. His research interests include immunology, cancer biology, development, and the development of biological instrumentation (for example, the protein sequencer and the automated fluorescent DNA sequencer). His early research played a key role in unraveling the mysteries of antibody diversity. More recently he has pioneered systems approaches to biology and medicine.

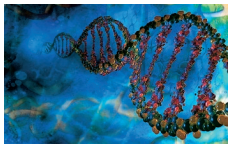
Dr. Hood has taught molecular evolution, immunology, molecular biology, genomics and biochemistry and has co-authored textbooks in biochemistry, molecular biology, and immunology, as well as *The Code of Codes*—a monograph about the Human Genome Project. He was one of the first advocates for the Human Genome Project and directed one of the federal genome centers that sequenced the human genome. Dr. Hood is currently the president (and co-founder) of the cross-disciplinary Institute for Systems Biology in Seattle, Washington.

Dr. Hood has received a variety of awards, including the Albert Lasker Award for Medical Research (1987), the Distinguished Service Award from the National Association of Teachers (1998) and the Lemelson/MIT Award for Invention (2003). He is the 2002 recipient of the Kyoto Prize in Advanced Biotechnology—an award recognizing his pioneering work in developing the protein and DNA synthesizers and sequencers that provide the technical foundation of modern biology. He is deeply involved in K–12 science education. His hobbies include running, mountain climbing, and reading.

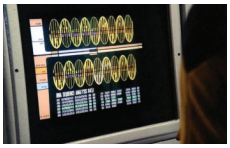
Brief Contents

Contents

PART II



PART III



PART IV

Preface

A Note from the Authors

The science of genetics is less than 150 years old, but its accomplishments within that short time have been astonishing. Gregor Mendel first described genes as abstract units of inheritance in 1865; his work was ignored and then rediscovered in 1900. Thomas Hunt Morgan and his students provided experimental verification of the idea that genes reside within chromosomes during the years 1910–1920. By 1944, Oswald Avery and his coworkers had established that genes are made of DNA. James Watson and Francis Crick published their pathbreaking structure of DNA in 1953. Remarkably, less than 50 years later (in 2001), an international consortium of investigators deciphered the sequence of the 3 billion nucleotides in the human genome. Twentieth century genetics made it possible to identify individual genes and to understand a great deal about their functions.

Today, scientists are able to access the enormous amounts of genetic data generated by the sequencing of many organisms' genomes. Analysis of these data will result in a deeper understanding of the complex molecular interactions within and among vast networks of genes, proteins, and other molecules that help bring organisms to life. Finding new methods and tools for analyzing these data will be a significant part of genetics in the twenty-first century.

Our sixth edition of *Genetics: From Genes to Genomes* emphasizes both the core concepts of genetics and the cutting-edge discoveries, modern tools, and analytic methods that will keep the science of genetics moving forward.

The authors of the sixth edition have worked together in revising every chapter in an effort not only to provide the most up-to-date information, but also to provide continuity and the clearest possible explanations of difficult concepts in one voice.

Our Focus—An Integrated Approach

Genetics: From Genes to Genomes represents a new approach to an undergraduate course in genetics. It reflects the way we, the authors, currently view the molecular basis of life.

We integrate:

- **Formal genetics:** the rules by which genes are transmitted.
- **Molecular genetics:** the structure of DNA and how it directs the structure of proteins.
- **Digital analysis and genomics:** recent technologies that allow a comprehensive analysis of the entire gene set and its expression in an organism.
- **Human genetics:** how genes contribute to health and diseases, including cancer.
- **The unity of life-forms:** the synthesis of information from many different organisms into coherent models.
- **Molecular evolution:** the molecular mechanisms by which biological systems, whole organisms, and populations have evolved and diverged.

The strength of this integrated approach is that students who complete the book will have a strong command of genetics as it is practiced today by both academic and corporate researchers. These scientists are rapidly changing our understanding of living organisms, including ourselves. Ultimately, this vital research may create the ability to replace or correct detrimental genes—those "inborn errors of metabolism," as researcher Archibald Garrod called them in 1923, as well as the later genetic alterations that lead to the many forms of cancer.

The Genetic Way of Thinking

Modern genetics is a molecular-level science, but an understanding of its origins and the discovery of its principles is a necessary context. To encourage a genetic way of thinking, we begin the book by reviewing Mendel's principles and the chromosomal basis of inheritance. From the outset, however, we aim to integrate organism-level genetics with fundamental molecular mechanisms.

Chapter 1 presents the foundation of this integration by summarizing the main biological themes we explore. In Chapter 2, we tie Mendel's studies of pea trait inheritance to the actions of enzymes that determine whether a pea is round or wrinkled, yellow or green, etc. In the same chapter, we point to the relatedness of the patterns of heredity in all organisms. Chapters 3–5 cover extensions to Mendel, the chromosome theory of inheritance, and the fundamentals of gene linkage and mapping. Starting in Chapter 6, we focus on the physical characteristics of DNA, on mutations, and on how DNA encodes, copies, and transmits biological information.

Beginning in Chapter 9, we move into the digital revolution in DNA analysis with a look at modern genetics techniques, including gene cloning, PCR, microarrays, and high-throughput genome sequencing. We explore how bioinformatics, an emergent analytical tool, can aid in discovery of genome features. This section concludes in Chapter 11 with case studies leading to the discovery of human disease genes.

The understanding of molecular and computer-based techniques carries into our discussion of chromosome specifics in Chapters 12–15, and also informs our analysis of gene regulation in Chapters 16 and 17. Chapter 18 describes the most recent technology that scientists can use to manipulate genomes at will – for research and practical purposes including gene therapy. Chapter 19 describes the use of genetic tools at the molecular level to uncover the complex interactions of eukaryotic development. In Chapter 20, we explain how our understanding of genetics and the development of molecular genetic technologies is enabling us to comprehend cancer and in some cases to cure it.

Chapters 21 and 22 cover population genetics, with a view of how molecular tools have provided information on species relatedness and on genome changes at the molecular level over time. In addition, we explain how bioinformatics can be combined with population genetics to understand inheritance of complex traits and to trace human ancestry.

Throughout our book, we present the scientific reasoning of some of the ingenious researchers of the field—from Mendel, to Watson and Crick, to the collaborators on the Human Genome Project. We hope student readers will see that genetics is not simply a set of data and facts, but also a human endeavor that relies on contributions from exceptional individuals.

Student-Friendly Features

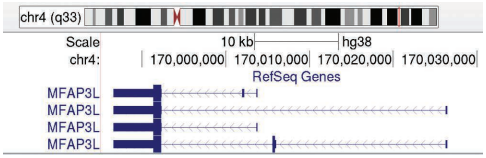
As digital components of the text become more and more crucial, we are very excited that Janice Fischer, a textbook author, is taking on a dual role as Digital Editor! Janice will ensure the important consistency between text and digital.

We have taken great pains to help the student make the leap to a deeper understanding of genetics. Numerous features of this book were developed with that goal in mind.

- ***One Voice Genetics:*** *Genes to Genomes* has a friendly, engaging reading style that helps students master the concepts throughout this book. The writing style provides the student with the focus and continuity required to make the book successful in the classroom.
- ***Visualizing Genetics*** The highly specialized art program developed for this book integrates photos and line art in a manner that provides the most engaging visual presentation of genetics available. Our Feature Figure illustrations break down complex processes into step-by-step illustrations that lead to greater student understanding. All illustrations are rendered with a consistent color theme—for example, all presentations of phosphate groups are the same color, as are all presentations of mRNA.
- ***Accessibility*** Our intention is to bring cutting-edge content to the student level. A number of more complex illustrations are revised and segmented to help the student follow the process. Legends have been streamlined to highlight only the most important ideas, and throughout the book, topics and examples have been chosen to focus on the most critical information.
- ***Problem Solving*** Developing strong problem-solving skills is vital for every genetics student. The authors have carefully created problem sets at the end of each chapter that allow students to improve upon their problem-solving ability.
- ***Solved Problems*** These cover topical material with complete answers provide insight into the step-by-step process of problem solving.
- ***Review Problems*** More than 700 questions involving a variety of levels of difficulty that develop excellent problem-solving skills. The problems are organized by chapter section and in order of increasing difficulty within each section for ease of use by instructors and students. The companion online *Study Guide and Solutions Manual,* completely revised for the 6th edition by Michael Goldberg and Janice Fischer, provides detailed analysis of strategies to solve all of the end-of-chapter problems.

SOLVED PROBLEMS

I. The following figure shows a screen shot from the UCSC Genome Browser, focusing on a region of the human genome encoding a gene called *MFAP3L.* (*Note: hg38* refers to version 38 of the human genome RefSeq.) If you do not remember how the browser represents the genome, refer to the key at the bottom of Fig. 10.3.



Source: University of California Genome Project, https://genome.ucsc.edu

a. Describe in approximate terms the genomic location of *MFAP3L.*

b. Is the gene transcribed in the direction from the centromere-to-telomere or from the telomere-to-centromere?

c. How many alternative splice forms of *MFAP3L* mRNA are indicated by the data?

b. The arrows within the introns of the gene show that the direction of transcription is from the telomere of *4q* toward the centromere of chromosome 4.

c. The data indicate four alternatively spliced forms of the mRNA. In the following parts, we list these as A to D from top to bottom.

d. The data suggest two promoters. One is roughly at position 170,037,000 and allows the transcription of a primary RNA alternatively spliced to produce mRNAs B and D. The other is roughly at position 170,013,000 and leads to the transcription of a primary RNA alternatively spliced to generate mRNAs A and C.

e. The data indicate that the *MFAP3* gene can encode two different but closely related proteins. mRNAs A, B, and C all encode the same protein; mRNA D a slightly larger protein that includes at its N terminus additional amino acids not found in the other protein. Otherwise these two proteins appear to be the same. The ORF that encodes the A B C protein form is about 880 bp long (a rough estimate); this corre-

Changes in the 6th Edition: A Chapter-by-Chapter Summary

The sixth edition has been revised and modernized significantly as compared with the fifth edition. We scrutinized the entire text and clarified the language wherever possible. In total, we created more than 50 new Figures and Tables, and revised more than 100 in addition. We also wrote more than 125 new end-of-chapter problems, and revised many other problems for clarity. The entire *Solutions Manual and Study Guide* was corrected and revised for clarity. We added several new Fast Forward, Genetics and Society, and Tools of Genetics Boxes on modern topics. Chapter 9 in the 5th edition was split into two separate chapters in the 6th edition: Chapter 9 (Digital Analysis of DNA) and Chapter 10 (Genome Annotation).

Along with the numerous text changes, the authors have also spent a great deal of time updating the test bank and question bank content to align more closely to the text. There will also be new video tutorials for difficult concepts in every chapter!

Every chapter of the sixth edition was improved significantly from the fifth edition. The most important changes in the sixth edition are summarized below:

Chapter 3 Extensions to Mendel's Laws
- Relationship between epistasis and complementation explained more clearly.
- Discussion of two-gene versus multifactorial inheritance now separated for clarity.
- Comprehensive Example about dog coat colors expanded to include molecular explanations for the various gene activities.

Chapter 4 The Chromosome Theory of Inheritance
- Figures and text altered to clarify that each chromatid has a centromere.
- New Fast Forward Box: *Visualizing X Chromosome Inactivation in Transgenic Mice*

Chapter 5 Linkage, Recombination, and the Mapping of Genes on Chromosomes
- New Fast Forward Box: *Mapping the Crossovers that Generate Individual Human Sperm*

Chapter 6 DNA Structure, Replication, and Recombination
- Improvements to diagrams of DSB repair model of recombination.
- New section about site-specific recombination.

Chapter 7 Anatomy and Function of Gene: Dissection Through Mutation
- Reorganized and clarified material to separate the discussion of DNA sequence alteration mechanisms from DNA repair mechanisms.

Chapter 9 Digital Analysis of DNA
- Improved depiction of plasmid cloning vectors.
- Renovated explanation of paired-end whole-genome shotgun sequencing.

Chapter 10 Genome Annotation
- Improved depiction of alternative RNA splicing.
- New illustration of consensus amino acid sequences in proteins.
- New material on the evolution of *de novo* genes.

Chapter 12 The Eukaryotic Chromosome
- New material on synthetic yeast chromosomes.

Chapter 15 Organellar Inheritance
- New Fast Forward Box about the Mitchondrial Eve concept.

Chapter 17 Gene Regulation in Eukaryotes
- New Tools of Genetics Box: *The Gal4/UAS$_G$ Binary Gene Expression System*
- New part of Epigenetics section: *Can Environmentally Acquired Traits Be Inherited?*
- New part of Regulation After Transcription section: Trans-*acting Proteins Regulate Translation*

Chapter 18 Manipulating the Genomes of Eukaryotes
- New part of Targeted Mutagenesis section: *CRISPR/Cas9 Allows Targeted Gene Editing in Any Organism*
- New Tools of Genetics Box: *How Bacteria Vaccinate Themselves Against Viral Infections with CRISPR/Cas9*
- New Genetics and Society Box: *Should We Alter the Genomes of Human Germ Lines?*

Chapter 19 The Genetic Analysis of Development
- Comprehensive Example of *Drosophila* body patterning revised to clarify that homeotic genes function within parasegments, and to clarify the concept of a morphogen.

Chapter 20 The Genetics of Cancer
- Clarified the fact that mutation drives cancer progression.
- Improved explanation of driver and passenger mutations.
- Increased coverage of tumor genome sequencing and the heterogeneity of mutations in different individuals with cancers in the same organ.

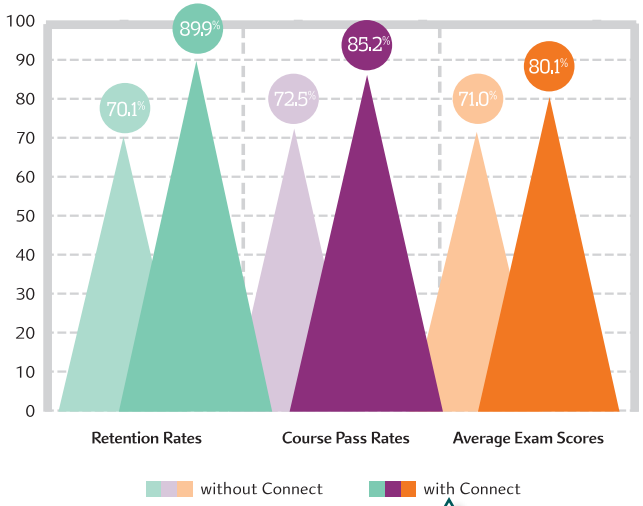
Chapter 22 The Genetics of Complex Traits
- Revised the section on heritability to clarify: lines of correlation and correlation coefficients; how to use different kinds of human twin studies to estimate the heritability of complex quantitative traits and complex discrete traits.
- New explanation of how to use the chi-square test for independence for GWAS.

McGraw-Hill Connect® is a highly reliable, easy-to-use homework and learning management solution that utilizes learning science and award-winning adaptive tools to improve student results.

Homework and Adaptive Learning

- Connect's assignments help students contextualize what they've learned through application, so they can better understand the material and think critically.
- Connect will create a personalized study path customized to individual student needs through SmartBook®.
- SmartBook helps students study more efficiently by delivering an interactive reading experience through adaptive highlighting and review.

Connect's Impact on Retention Rates, Pass Rates, and Average Exam Scores



Using **Connect** improves retention rates by **19.8%**, passing rates by **12.7%, and** exam scores by **9.1%.**

Over **7 billion questions** have been answered, making McGraw-Hill Education products more intelligent, reliable, and precise.

Quality Content and Learning Resources

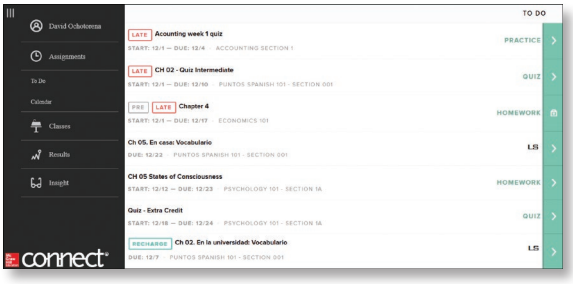
- Connect content is authored by the world's best subject matter experts, and is available to your class through a simple and intuitive interface.
- The Connect eBook makes it easy for students to access their reading material on smartphones and tablets. They can study on the go and don't need internet access to use the eBook as a reference, with full functionality.
- Multimedia content such as videos, simulations, and games drive student engagement and critical thinking skills.

73% of instructors who use **Connect** require it; instructor satisfaction **increases** by 28% when **Connect** is required.

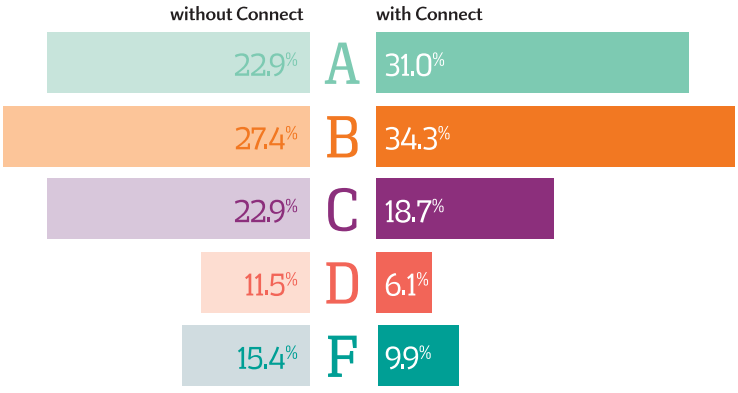
Robust Analytics and Reporting

- Connect Insight® generates easy-to-read reports on individual students, the class as a whole, and on specific assignments.
- The Connect Insight dashboard delivers data on performance, study behavior, and effort. Instructors can quickly identify students who struggle and focus on material that the class has yet to master.
- Connect automatically grades assignments and quizzes, providing easy-to-read reports on individual and class performance.

©Hero Images/Getty Images



Impact on Final Course Grade Distribution



More students earn **As** and **Bs** when they use **Connect**.

Trusted Service and Support

- Connect integrates with your LMS to provide single sign-on and automatic syncing of grades. Integration with Blackboard®, D2L®, and Canvas also provides automatic syncing of the course calendar and assignment-level linking.
- Connect offers comprehensive service, support, and training throughout every phase of your implementation.
- If you're looking for some guidance on how to use Connect, or want to learn tips and tricks from super users, you can find tutorials as you work. Our Digital Faculty Consultants and Student Ambassadors offer insight into how to achieve the results you want with Connect.

Guided Tour

Integrating Genetic Concepts

Genetics: From Genes to Genomes takes an integrated approach in its presentation of genetics, thereby giving students a strong command of genetics as it is practiced today by academic and corporate researchers. Principles are related throughout the text in examples, essays, case histories, and connections sections to make sure students fully understand the relationships between topics.

Chapter Outline

Every chapter opens with a brief outline of the chapter contents.

chapter outline

- 18.1 Creating Transgenic Organisms
- 18.2 Uses of Transgenic Organisms
- 18.3 Targeted Mutagenesis
- 18.4 Human Gene Therapy

Learning Objectives

Learning Objectives appear before each section, and are carefully written to clearly outline expectations.

essential concepts

- A wild-type transgene can be inserted into an embryo homozygous for a recessive mutant allele. If the normal phenotype is restored, then the transgene identifies the gene that was mutated.
- The creation of reporter constructs allows easy detection of when and in which tissues a gene is turned on or turned off in eukaryotes.
- Transgenic organisms produce medically important human proteins including insulin, blood clotting factors, and erythropoietin; transgenic crop plants can potentially make ingestible vaccines.
- GM soybeans are resistant to the weed killer glyphosate. Many crops, such as corn, soybean, canola, and cotton have been genetically modified to express Bt protein which discourages insect predation.
- Adding a transgene that carries a disease-causing, gain-of-function allele to a nonhuman animal model allows researchers to observe disease progression and to test possible therapeutic interventions.

18.2 Uses of Transgenic Organisms

learning objectives

1. Describe how transgenes can clarify which gene causes a mutant phenotype.
2. Summarize the use of transgene reporter constructs in gene expression studies.
3. Discuss examples of how transgenic organisms serve to produce proteins needed for human health.
4. List examples of GM organisms and discuss the pros and cons of their production.
5. Explain the use of transgenic animals to model gain-of-function genetic diseases in humans.

Essential Concepts

After each section, the most relevant points of content are now provided in concise, bulleted statements to reinforce crucial concepts and learning objectives for students.

WHAT'S NEXT

Manipulation of the genome is the basis for many of the experimental strategies we will describe in Chapter 19, where we discuss how genetic analysis has been a crucial tool in unraveling the biochemical pathways of development—the process by which a single-celled zygote becomes a complex multicellular organism. Transgenic technology is key to cloning the genes identified in mutant screens that are crucial for regulating development, and also to manipulating these genes in order to understand their precise functions in the organism.

What's Next

Each chapter closes with a What's Next section that serves as a bridge between the topics in the chapter just completed to those in the upcoming chapter or chapters. This spirals the learning and builds connections for students.

New! Exciting Revised Content

Every chapter of the sixth edition has been revised and modernized significantly as compared with the fifth edition. More than 50 new Figures and Tables were created, and more than 100 were revised. More than 125 new end-of-chapter problems were written, and many more revised for clarity. The entire Solutions Manual and Study Guide was updated, corrected, and revised by Michael Goldberg and Janice Fischer. Several new Fast Forward, Genetics and Society, and Tools of Genetics Boxes covering modern topics were created. For breadth and clarity, Chapter 9 in the 5th edition was split into two separate chapters in the 6th edition: Chapter 9 (Digital Analysis of DNA) and Chapter 10 (Genome Annotation).

Figure 6.30 One site-specific recombination mechanism. The Cre and Flp enzymes discussed in the text function as shown. The *red* and *blue* target DNA sequences are identical to each other but are represented in different colors for clarity. These targets are embedded in different DNA molecules (*black* and *gray dots*). The subunits of the recombinase tetramer are *yellow ovals;* this enzyme catalyzes all steps of the reaction. *Black triangles* are sites where recombinase cleaves single-stranded DNA. Note that resolution of the Holliday junction intermediate involves cleavage of the *blue* and *red* DNA strands that were not cleaved initially.

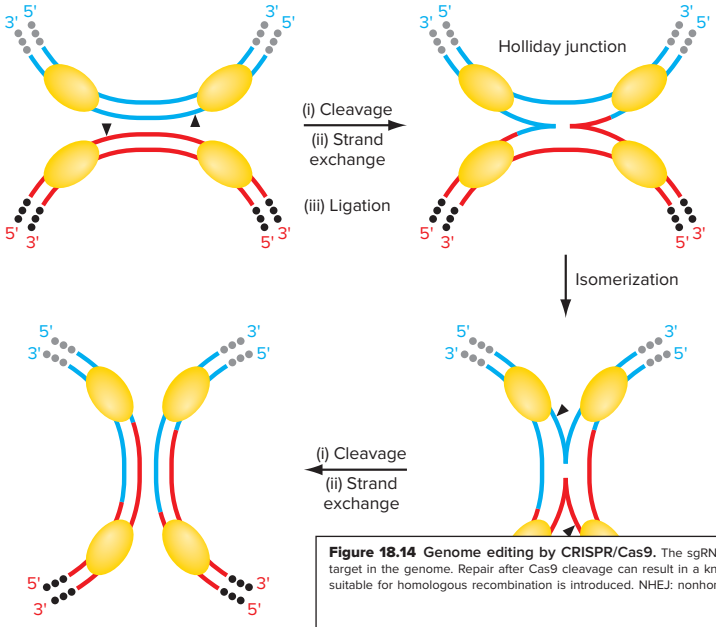
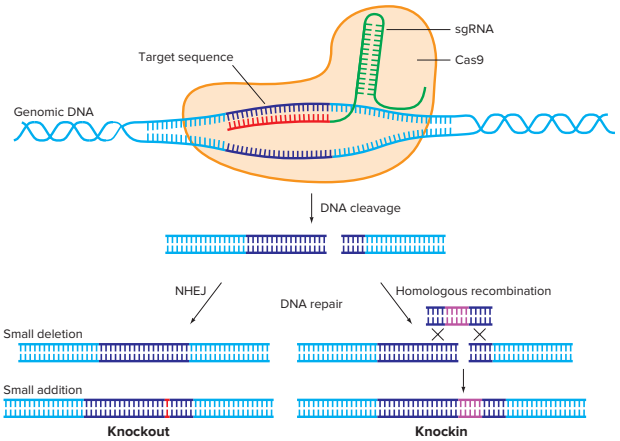


Figure 18.14 Genome editing by CRISPR/Cas9. The sgRNA sequence is designed to bring the Cas9 endonuclease to a specific target in the genome. Repair after Cas9 cleavage can result in a knockout or a knockin, depending on whether or not a DNA fragment suitable for homologous recombination is introduced. NHEJ: nonhomologous end-joining.



FAST FORWARD

Visualizing X Chromosome Inactivation in Transgenic Mice

Scientists have recently used molecular techniques and transgenic technology (similar to that described in the earlier Fast Forward Box *Transgenic Mice Prove That* SRY *Is the Maleness Factor*) to visualize the pattern of X chromosome inactivation in mice. The researchers generated XX mice containing two different *transgenes* (in this case, genes from a different species). One of these transgenes was a jellyfish gene that specifies green fluorescent protein (GFP); the other was a gene from red coral that makes red fluorescent protein (RFP) **(Fig. A).**

In the XX mice, the GFP gene is located on the X chromosome from the mother, and the RFP gene resides on the X chromosome from the father. Clonal patches of cells are either green or red depending on which X chromosome was turned into a Barr body in the original cell that established the patch **(Fig. B).**

Different XX mice display different green and red patchwork patterns, providing a clear demonstration of the random nature of X chromosome inactivation. The patchwork patterns reflect the cellular memory of which X chromosome was inactivated in the founder cell for each clonal patch. Geneticists currently use these transgenic mice to decipher the genetic details of how cells "remember" which X to inactivate after each cell division.

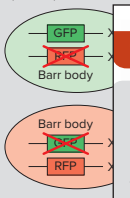


Figure A Cells of transgenic mice glow either green or red in response to X chromosome inactivation. The mouse carries a green (GFP) transgene inserted in the maternal X chromosome (X^M), and a red (RFP) transgene in the paternal X chromosome (X^P). Cells in which X^P is inactivated (*top*) glow green; cells glow red (*bottom*) when X^M is inactivated.

Figure B Heart cells of a transgenic mouse reveal a clonal patchwork of X inactivation. Patches of red or green cells represent cellular descendants of the founders that randomly inactivated one of their X chromosomes.

Fast Forward

This feature is one of the methods used to integrate the Mendelian principles introduced early in the content with the molecular content that will follow.

TOOLS OF GENETICS

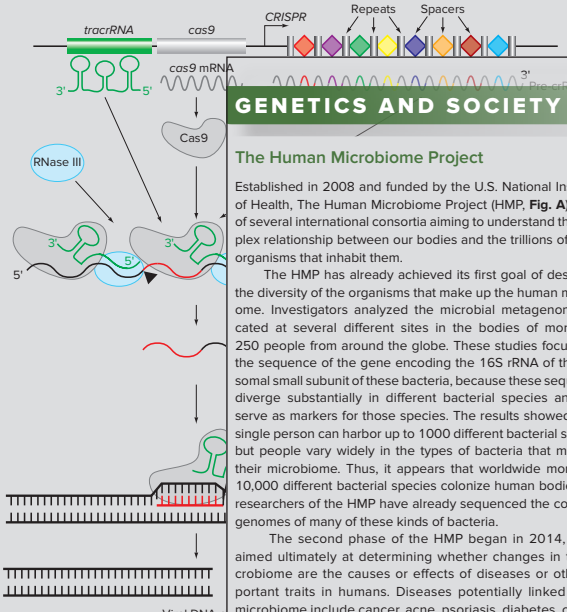
How Bacteria Vaccinate Themselves Against Viral Infections with CRISPR/Cas9

Researchers discovered clustered sequence repeats (CRISPRs) in bacterial genomes as early as 1987. When in 2005 some of these sequences were found to originate from bacteriophage genomes, several astute scientists speculated that CRISPRs might mediate a viral immunity system in bacteria. These ideas were largely ignored for several more years until the mechanism of resistance became clarified. And finally, in 2012–2013, the so-called *CRISPR craze* reached its full bloom when researchers including Feng Zhang, Jennifer Doudna, and Emmanuelle Charpentier developed methods to adapt this viral immunity system to engineer genomes in bacterial cells and in eukaryotic organisms.

At the *CRISPR* locus of bacterial genomes, short direct repeats are interrupted at regular intervals by unique spacer sequences (**Fig. A**). The spacer sequences are fragments of bacteriophage genomes captured by the host cell and integrated into the host genome by the action of two bacterially encoded Cas proteins (Cas1 and Cas2). The repeats within the *CRISPR* arrays are added by these endonucleolytic enzymes during the capture and integration process.

Viral immunity results from steps that begin with transcription of the *CRISPR* array into long RNA molecules called *pre-crRNAs* that are processed into short (24–48 nt) so-called *CRISPR RNAs* (*crRNAs*). In the bacterial species

Figure A The *CRISPR/Cas9* locus vaccinates bacteria against viruses.



Tools of Genetics Essays

Current readings explain various techniques and tools used by geneticists, including examples of applications in biology and medicine.

GENETICS AND SOCIETY

Crowd: © Image Source/Getty Images RF

The Human Microbiome Project

Established in 2008 and funded by the U.S. National Institutes of Health, The Human Microbiome Project (HMP, **Fig. A**) is one of several international consortia aiming to understand the complex relationship between our bodies and the trillions of microorganisms that inhabit them.

The HMP has already achieved its first goal of describing the diversity of the organisms that make up the human microbiome. Investigators analyzed the microbial metagenomes located at several different sites in the bodies of more than 250 people from around the globe. These studies focused on the sequence of the gene encoding the 16S rRNA of the ribosomal small subunit of these bacteria, because these sequences diverge substantially in different bacterial species and thus serve as markers for those species. The results showed that a single person can harbor up to 1000 different bacterial species, but people vary widely in the types of bacteria that make up their microbiome. Thus, it appears that worldwide more than 10,000 different bacterial species colonize human bodies. The researchers of the HMP have already sequenced the complete genomes of many of these kinds of bacteria.

The second phase of the HMP began in 2014, and is aimed ultimately at determining whether changes in the microbiome are the causes or effects of diseases or other important traits in humans. Diseases potentially linked to the microbiome include cancer, acne, psoriasis, diabetes, obesity, and inflammatory bowel disease; some investigators have suggested that the composition of microbiomes could also influence the mental health of their hosts. The first step in these studies will be to establish whether statistical correlations exist between specific kinds of microbial communities and disease states. As one example, one HMP phase II project currently underway is an analysis of vaginal host cells and microbes during pregnancy. Approximately 2000 pregnant women will be studied and their birth outcomes recorded. The goal of this project is to determine if changes in the microbiome correlate with premature birth or other complications of pregnancy.

Of course, the existence of any correlations found between microbiomes and disease does not prove cause or effect. But even if bacteria correlated with a disease state do not cause the disease, the existence of the correlation could be useful as a way to diagnose certain conditions. Nevertheless, the most exciting

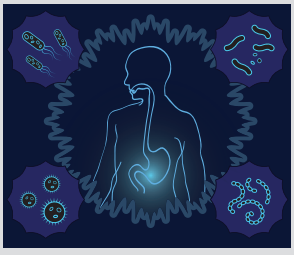


Figure A
© Anna Smirnova/Alamy RF

an effect? One method is to investigate in detail how the biological properties of the microbiome and the host might be changed by the interactions of bacteria and the humans they colonize. Thus, scientists will characterize whether and how the transcriptomes and proteomes of the bacteria and human cells are changed by bacterial colonization of human organs. These studies will further delve into *metabolomics* (characterizing metabolites in the human bloodstream).

A second and even more powerful method for establishing the cause and effect of microbiome changes is the use of *germ-free mice* raised in sterile environments. Surprisingly, germ-free mice can survive although they are not normal: they have altered immune systems, poor skin, and they need to eat more calories than do normal mice to maintain a normal body weight. Researchers can populate germ-free mice with a single bacterial species or a complex microbial community, and thus determine how microbiomes influence physiological states. Problem 8 at the end of this chapter will allow you to explore this approach by discussing an experiment recently performed with germ-free mice that asks if the microbiome plays a causal role in obesity.

If microbial communities indeed contribute to disease states in humans, then future treatments might aim to alter resident microbiomes. Thus, the flip side of the HMP is to in-

Genetics and Society Essays

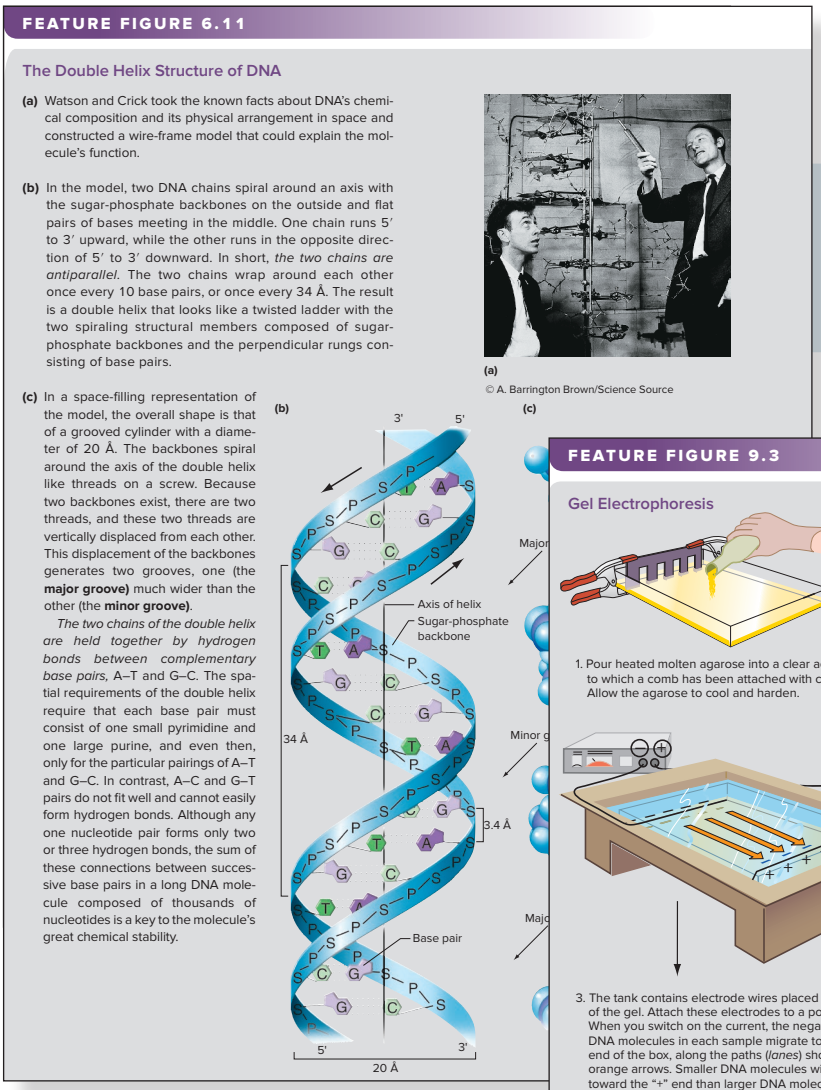
Dramatic essays explore the social and ethical issues created by the multiple applications of modern genetic research.

Visualizing Genetics

Full-color illustrations and photographs bring the printed word to life. These visual reinforcements support and further clarify the topics discussed throughout the text.

Feature Figures

Special multipage spreads integrate line art, photos, and text to summarize in detail important genetic concepts.



FEATURE FIGURE 6.11

The Double Helix Structure of DNA

(a) Watson and Crick took the known facts about DNA's chemical composition and its physical arrangement in space and constructed a wire-frame model that could explain the molecule's function.

(b) In the model, two DNA chains spiral around an axis with the sugar-phosphate backbones on the outside and flat pairs of bases meeting in the middle. One chain runs 5′ to 3′ upward, while the other runs in the opposite direction of 5′ to 3′ downward. In short, *the two chains are antiparallel*. The two chains wrap around each other once every 10 base pairs, or once every 34 Å. The result is a double helix that looks like a twisted ladder with the two spiraling structural members composed of sugar-phosphate backbones and the perpendicular rungs consisting of base pairs.

(c) In a space-filling representation of the model, the overall shape is that of a grooved cylinder with a diameter of 20 Å. The backbones spiral around the axis of the double helix like threads on a screw. Because two backbones exist, there are two threads, and these two threads are vertically displaced from each other. This displacement of the backbones generates two grooves, one (the **major groove)** much wider than the other (the **minor groove)**.

The two chains of the double helix are held together by hydrogen bonds between complementary base pairs, A–T and G–C. The spatial requirements of the double helix require that each base pair must consist of one small pyrimidine and one large purine, and even then, only for the particular pairings of A–T and G–C. In contrast, A–C and G–T pairs do not fit well and cannot easily form hydrogen bonds. Although any one nucleotide pair forms only two or three hydrogen bonds, the sum of these connections between successive base pairs in a long DNA molecule composed of thousands of nucleotides is a key to the molecule's great chemical stability.

(a)

© A. Barrington Brown/Science Source

(b) **(c)**

3′ 5′ Axis of helix Sugar-phosphate backbone 34 Å 3.4 Å Base pair 5′ 3′ 20 Å

FEATURE FIGURE 9.3

Gel Electrophoresis

1. Pour heated molten agarose into a clear acrylic plate to which a comb has been attached with clamps. Allow the agarose to cool and harden.

2. Remove the comb; shallow wells will be left in the gel. Remove the gel from the acrylic plate and transfer it to a tank containing a buffered solution. Use a micropipette to load a different DNA sample into each well of the gel. Each sample contains a blue dye to make it easier to see. One sample should contain DNA molecules of known length to serve as size markers.

3. The tank contains electrode wires placed along each end of the gel. Attach these electrodes to a power supply. When you switch on the current, the negatively charged DNA molecules in each sample migrate toward the "+" end of the box, along the paths (*lanes*) shown by the orange arrows. Smaller DNA molecules will move faster toward the "+" end than larger DNA molecules.

4. Remove the gel from the tank. Incubate in a solution containing ethidium bromide (which binds to DNA), then wash with water to remove excess dye from the gel.

5. Expose the gel to ultraviolet (UV) light. DNA molecules will fluoresce as orange bands because the ethidium bromide bound to the DNA absorbs UV photons and gives off photons in the visible red range. You can estimate the size of the DNA molecules in the unknown samples by comparing their migration in the gel with that of the size markers (*standards*) in the lane at the left.

Standard Human A B Lambda C D E Plasmid F G H Well Kb 20 10 7 5 4 3 2 1.5 1 0.7 0.5 0.4 0.3 0.2 Direction of electrophoresis

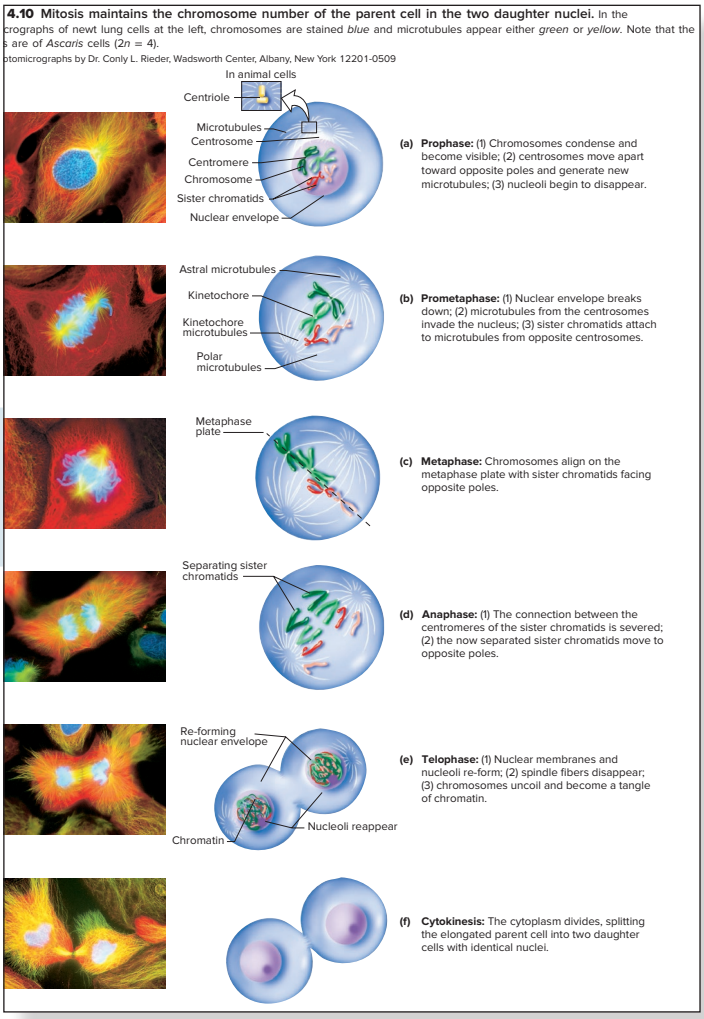
(5): © Lee Silver, Princeton University

Separating DNA molecules according to their size by agarose gel electrophoresis. To prepare an agarose gel with wells for samples, you pour the gel as shown in Step 1. You then transfer the gel to a tank containing a buffered solution with ions that allow current to flow, and load DNA samples in the wells (Step 2). You then connect the gel tank to a power supply and allow electrophoresis to run from 1 to 20 hours (depending on the DNA size and the voltage; Step 3). After incubating the gel with the fluorescent dye ethidium bromide (Step 4), you then expose the gel to UV light (Step 5). DNA molecules will appear as orange bands because they bind to the fluorescent dye.

Step 5 shows actual results from gel electrophoresis; because black-and-white film was used, DNA appears white rather than orange. The standard lane at left has DNA fragments of known sizes. Human genomic DNA was cut with *Eco*RI in lane A and with *Rsa*I in lane B. Smears containing hundreds of thousands of fragments are produced with an average size of about 4.1 kb for *Eco*RI and 256 bp for *Rsa*I. In C, D, and E, the chromosome of bacteriophage λ was cut with *Hin*dIII, *Eco*RI, and *Rsa*I, respectively. The sizes of the fragments in any one lane add up to 48.5 kb, the size of the viral genome. In F, G, and H plasmid DNA of total length 6.9 kb was cut with the same three enzymes. Note that the larger the genome analyzed, the more fragments are produced; moreover, the more bases in the restriction enzyme recognition site, the larger is the average size of the fragments produced.

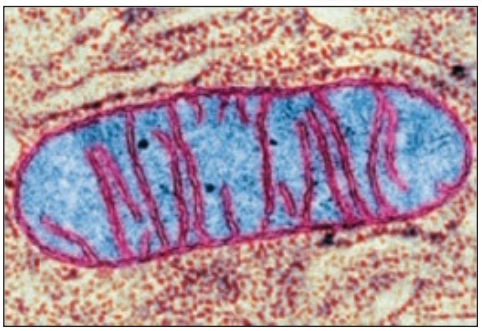
Process Figures

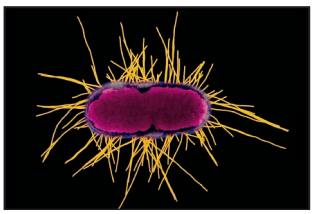
Step-by-step descriptions allow the student to walk through a compact summary of important details.

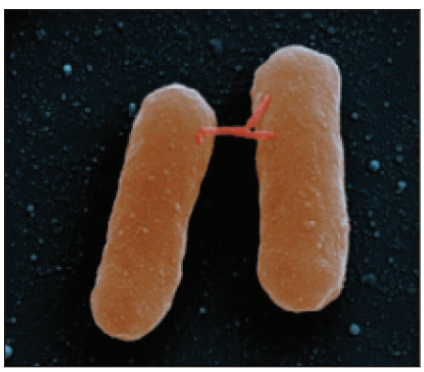


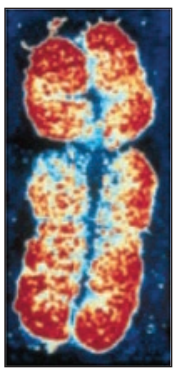
Micrographs

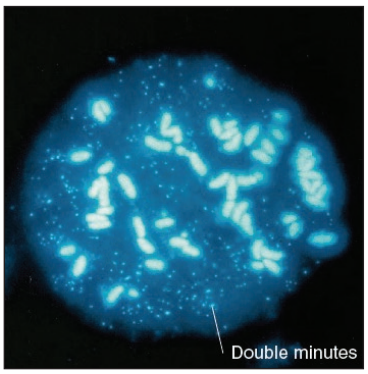
Stunning micrographs bring the genetics world to life.

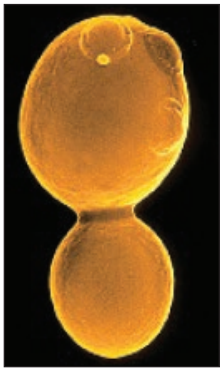


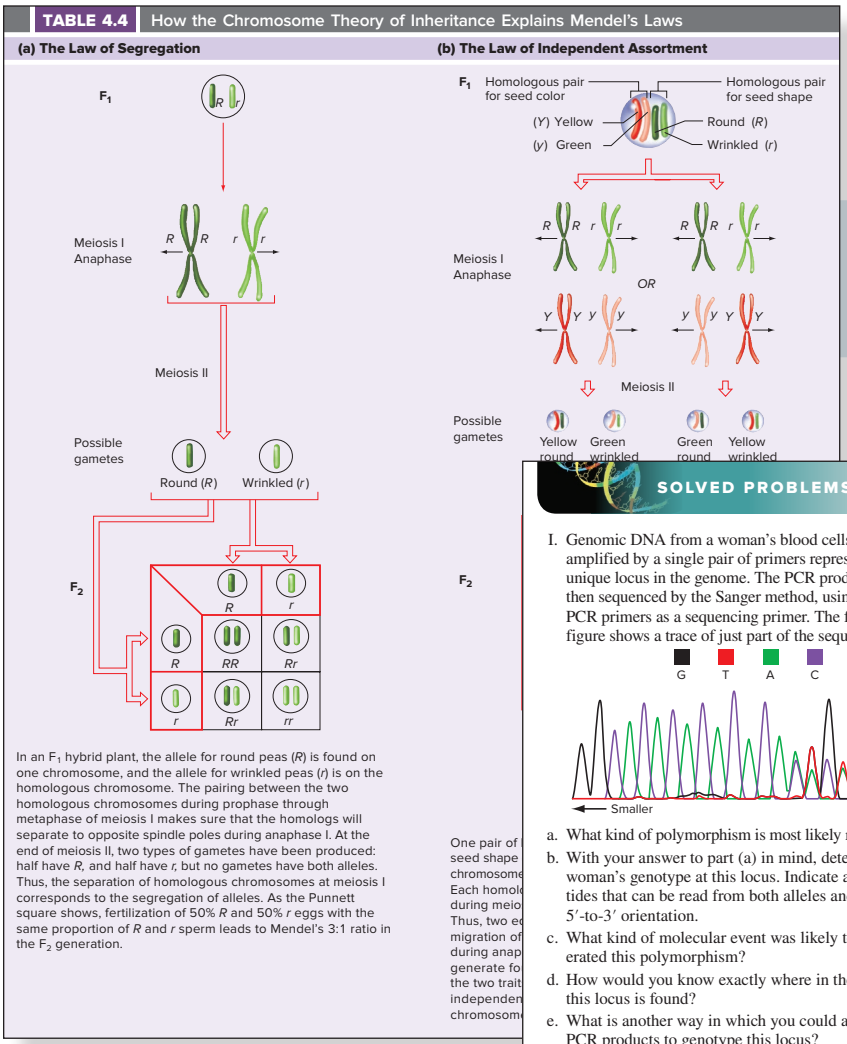






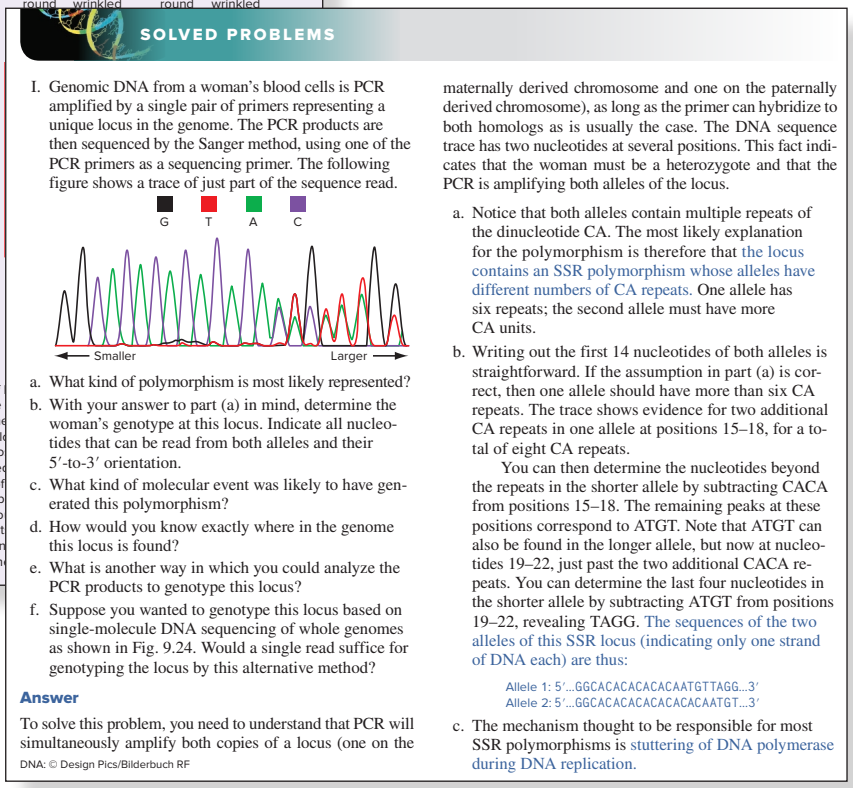






Comparative Figures

Comparison illustrations lay out the basic differences of often confusing principles.



Solving Genetics Problems

The best way for students to assess and increase their understanding of genetics is to practice through problems. Found at the end of each chapter, problem sets assist students in evaluating their grasp of key concepts and allow them to apply what they have learned to real-life issues.

Review Problems

Problems are organized by chapter section and in order of increasing difficulty to help students develop strong problem-solving skills. The answers to select problems can be found in the back of this text.

Solved Problems

Solved problems offer step-by-step guidance needed to understand the problem-solving process.

Acknowledgements

The creation of a project of this scope is never solely the work of the authors. We are grateful to our colleagues who answered our numerous questions, or took the time to share with us their suggestions for improvement of the previous edition. Their willingness to share their expertise and expectations was a tremendous help to us.

- Charles Aquadro, *Cornell University*
- Daniel Barbash, *Cornell University*
- Johann Eberhart, *University of Texas at Austin*
- Tom Fox, *Cornell University*
- Kathy Gardner, *University of Pittsburgh*
- Larry Gilbert, *University of Texas at Austin*
- Nancy Hollingsworth, *Stony Brook University*
- Mark Kirkpatrick, *University of Texas at Austin*
- Alan Lloyd, *University of Texas at Austin*
- Paul Macdonald, *University of Texas at Austin*
- Kyle Miller, *University of Texas at Austin*
- Debra Nero, *Cornell University*
- Howard Ochman, *University of Texas at Austin*
- Kristin Patterson, *University of Texas at Austin*
- Inder Saxena, *University of Texas at Austin*

Janice Fischer and Michael Goldberg would also like to thank their Genetics students at The University of Texas at Austin and Cornell University for their amazing questions. Many of their ideas have influenced the 6th edition.

A special thank-you to Kevin Campbell for his extensive feedback on this sixth edition. We would also like to thank the highly skilled publishing professionals at McGraw-Hill who guided the development and production of the sixth edition of ***Genetics: From Genes to Genomes:*** Justin Wyatt and Michelle Vogler for their support; Mandy Clark for her organizational skills and tireless work to tie up all loose ends; and Vicki Krug and the entire production team for their careful attention to detail and ability to move the schedule along.

chapter 1

Genetics: The Study of Biological Information

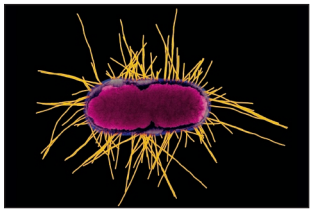
Information can be stored in many ways, including the patterns of letters and words in books and the sequence of nucleotides in DNA molecules.

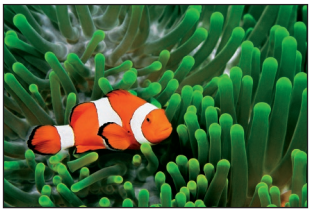
chapter outline

GENETICS, THE SCIENCE of heredity, is at its core the study of biological information. All living organisms—from single-celled bacteria and protozoa to multicellular plants and animals—must store and use vast quantities of information to develop, survive, and reproduce in their environments (**Fig. 1.1**). Geneticists examine how organisms use biological information during their lifetimes and pass it on to their progeny.

This book introduces you to the field of genetics as currently practiced in the early twenty-first century. Several broad themes recur throughout this presentation. First, we know that biological information is encoded in *DNA*, and that the *proteins* responsible for an organism's many functions are built from this code. Second, we have found that all living forms are related at the molecular level. With the aid of high-speed computers and other technologies, we can now study *genomes* at the level of DNA sequence. These new methods have revealed that genomes have a modular construction that has allowed rapid evolution of complexity. Finally, our focus in this book is on human genetics and the application of genetic discoveries to human problems.

Figure 1.1 The biological information in DNA generates an enormous diversity of living organisms.


(a) Bacteria


(b) Clown fish


(c) Lion


(d) Oak tree


(e) Poppies


(f) Hummingbird


(g) Red-eyed tree frog


(h) Humans

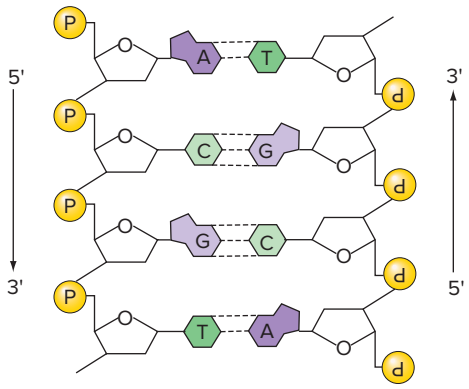
1.1 DNA: Life's Fundamental Information Molecule

learning objectives

1. Relate the structure of DNA to its function.
2. Differentiate between a chromosome, DNA, a gene, a base pair, and a protein.

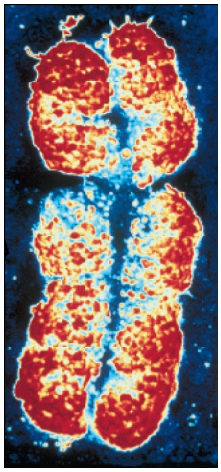
The process of **evolution**—the change in traits of groups of organisms over time—has taken close to 4 billion years to generate the amazing mechanisms for storing, replicating, expressing, and diversifying biological information seen in organisms now inhabiting the earth. The linear **DNA** molecule stores biological information in units known as **nucleotides.** Within each DNA molecule, the sequence of the four letters of the DNA alphabet—G, C, A, and T—specify which proteins an organism will make as well as when and where protein synthesis will occur. The letters refer to the **bases**—guanine, cytosine, adenine, and thymine—that are components of the nucleotide building blocks of DNA. The DNA molecule itself is a double strand of nucleotides carrying complementary G–C or A–T base pairs (**Fig. 1.2**). These **complementary base pairs** bind together through hydrogen bonds. The molecular **complementarity** of double-stranded DNA is its most important property and the key to understanding how DNA functions.

Figure 1.2 Complementary base pairs are a key feature of the DNA molecule. A single strand of DNA is composed of nucleotide subunits each consisting of a deoxyribose sugar (*white pentagons*), a phosphate (*yellow circles*), and one of four nitrogenous bases—adenine, thymine, cytosine, or guanine (designated as *lavender* or *green* A*s*, *T*s, *C*s, or *G*s). Hydrogen bonds (*dotted lines*) enable A to associate tightly with T, and C to associate tightly with G. Thus the two strands are complementary to each other. The arrows labeled 5′ to 3′ show that the strands have opposite orientations.



Although the DNA molecule is three-dimensional, most of its information is one-dimensional and digital. The information is one-dimensional because it is encoded as a specific sequence of letters along the length of the molecule. It is digital because each unit of information—one of the four letters of the DNA alphabet—is discrete. Because genetic information is digital, it can be stored as readily in a computer memory as in a DNA molecule. Indeed, the

Figure 1.3 A human chromosome. Each chromosome contains hundreds to thousands of genes.



combined power of DNA sequencers, computers, and DNA synthesizers makes it possible to store, interpret, replicate, and transmit genetic information electronically from one place to another anywhere on the planet.

The DNA regions that encode proteins are called **genes.** Just as the limited number of letters in a written alphabet places no restrictions on the stories one can tell, so too the limited number of letters in the genetic code alphabet places no restrictions on the kinds of proteins and thus the kinds of organisms genetic information can define.

Within the cells of an organism, DNA molecules carrying the genes are assembled into **chromosomes:** organized structures containing DNA and proteins that package and manage the storage, duplication, expression, and evolution of DNA (**Fig. 1.3**). The DNA within the entire collection of chromosomes in each cell of an organism is its **genome.** Human cells, for example, contain 24 distinct kinds of chromosomes carrying approximately 3×10^9 base pairs and roughly 27,000 genes. The amount of information that can be encoded in this size genome is equivalent to 6 million pages of text containing 250 words per page, with each letter corresponding to one *base pair*.

To appreciate the long journey from a finite amount of genetic information easily storable on a computer disk to the production of a human being, we next must examine proteins, the primary molecules that determine how complex systems of cells, tissues, and organisms function.

essential concepts

- *DNA*, a double-stranded macromolecule composed of four nucleotides, is the repository of genetic information.
- DNA is organized into *chromosomes* (of 24 different types in humans) that collectively constitute an organism's *genome*.
- The human genome contains about 27,000 genes, most of which encode *proteins*.

1.2 Proteins: The Functional Molecules of Life Processes

learning objectives

1. Compare the chemical structures of DNA and proteins.
2. Differentiate between the functions of DNA and the functions of proteins.

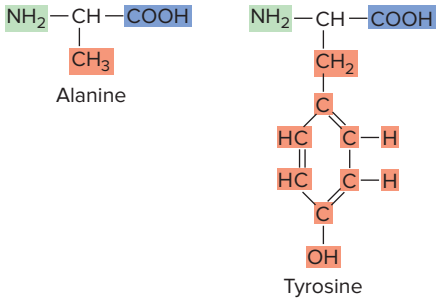
Although no single characteristic distinguishes living organisms from inanimate matter, you would have little trouble deciding which entities in a group of objects are alive. Over time, these living organisms, governed by the laws of physics and chemistry as well as a genetic program, would be able to reproduce themselves. Most of the organisms would also have an elaborate and complicated structure that would change over time—sometimes drastically, as when an insect larva metamorphoses into an adult. Yet another characteristic of life is the ability to move. Animals swim, fly, walk, or run, while plants grow toward or away from light. Still another characteristic is the capacity to adapt selectively to the environment. Finally, a key characteristic of living organisms is the ability to use sources of energy and matter to grow—that is, the ability to convert foreign material into their own body parts. The chemical and physical reactions that carry out these conversions are known as *metabolism.*

Most properties of living organisms arise ultimately from the class of molecules known as **proteins**—large polymers composed of hundreds to thousands of **amino acid** subunits strung together in long chains. Each chain folds into a specific three-dimensional conformation dictated by the sequence of its amino acids (**Fig. 1.4**). Most proteins are composed of 20 different amino acids. The information in the DNA of genes dictates, via a *genetic code,* the order of amino acids in a protein molecule.

You can think of proteins as constructed from a set of 20 different kinds of snap beads distinguished by color and shape. If you were to arrange the beads in any order, make strings of a thousand beads each, and then fold or twist the chains into shapes dictated by the order of their beads, you would be able to make a nearly infinite number of different three-dimensional shapes. The astonishing diversity of three-dimensional protein structures generates the extraordinary diversity of protein functions that is the basis

Figure 1.4 Proteins are polymers of amino acids that fold in three dimensions. **(a)** Structural formulas for two amino acids: alanine and tyrosine. All amino acids have a basic amino group (—NH_2; *green*) at one end and an acidic carboxyl group (—COOH; *blue*) at the other. The specific side chain (*red*) determines the amino acid's chemical properties. **(b)** The amino acid sequences of two different human proteins: the ß chain of hemoglobin (*green*), and the enzyme lactate dehydrogenase (*purple*). **(c)** The different amino acid sequences of these proteins dictate different three-dimensional shapes. The specific sequence of amino acids in a chain determines the precise three-dimensional shape of the protein.

(a)



(b)

Hemoglobin β chain

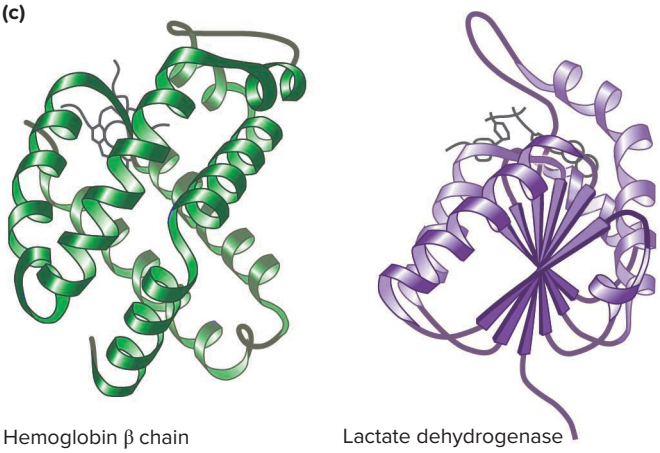
MVHLTPEEKSAVTALWGKVNVDEVGEALGRLLVVYPWTQRLFESFGDLFTPDAVMGNPKVKAHG
KKVLGAFSDGPAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGKEFTPPVQAA
YQKVVAGVANALAHKYH

Lactate dehydrogenase

MATIKSELIKNFAEEEAIHHNKISIVGTGSVGVACAISILLKGLSDELVLVDVDEGKLKGETMDL
QHGSPFMKMPNIVSSKDYLVTANSNLVIITAGARQKKGETRLDLVQRNVSIFKLMIPNITQYSPH
CKLLIVTNPVDILTYVAWKLSGFPKNRVIGSGCNLDSARFRYFIGQRLGIHSESCHGLILGEHGD
SSVPVWSGVNIAGVPLKDLNPDIGTDKDPEQWENVHKKVISSGYEMVKMKGYTSWGISLSVADLT
ESILKNLRRVHPVSTLSKGLYGINEDIFLSVPCILGENGITDLIKVKLTLEEEACLQKSAETLWEIQKELKL

A = Ala	= alanine	G = Gly	= glycine	M = Met	= methionine	S = Ser	= serine
C = Cys	= cysteine	H = His	= histidine	N = Asn	= asparagine	T = Thr	= threonine
D = Asp	= aspartic acid	I = Ile	= isoleucine	P = Pro	= proline	V = Val	= valine
E = Glu	= glutamic acid	K = Lys	= lysine	Q = Gln	= glutamine	W = Trp	= tryptophan
F = Phe	= phenylalanine	L = Leu	= leucine	R = Arg	= arginine	Y = Tyr	= tyrosine

(c)



Hemoglobin β chain — Lactate dehydrogenase

of each organism's complex and adaptive behavior (Fig. 1.4b and c). The structure and shape of the hemoglobin protein, for example, allow it to transport oxygen in the bloodstream and release it to the tissues. In contrast, lactate dehydrogenase is an enzyme that converts lactate to pyruvate, an important step in producing cellular energy. Most of the properties associated with life emerge from the constellation of protein molecules that an organism synthesizes according to instructions contained in its DNA.

essential concepts

- *Proteins* are responsible for most biological functions of cells and organisms.
- A protein is a macromolecule consisting of *amino acids* linked in a linear sequence.
- The sequences of amino acids in proteins are encoded by *genes* within the DNA.

1.3 Molecular Similarities of All Life-Forms

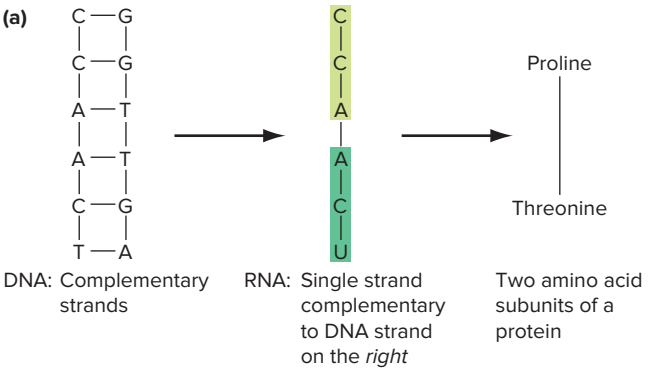
learning objective

1. Summarize the molecular evidence for the common origin of living organisms.

The evolution of biological information is a fascinating story spanning the 4 billion years of earth's history. Many biologists think that **RNA** was the first information-processing molecule to appear. Very similar to DNA, RNA molecules are also composed of four subunits: the bases G, C, A, and U (for uracil, which replaces the T of DNA). Like DNA, RNA has the capacity to store, replicate, and express information; like proteins, RNA can fold in three dimensions to produce molecules capable of catalyzing the chemistry of life. In fact, you will learn that the ultimate function of some genes is to encode RNA molecules instead of proteins. RNA molecules, however, are intrinsically unstable. Thus, it is probable that the more stable DNA took over the linear information storage and replication functions of RNA, while proteins, with their far greater capacity for diversity, preempted in large part the functions derived from RNA's three-dimensional folding. With this division of labor, RNA became primarily an intermediary in converting the information in DNA into the sequence of amino acids in protein (**Fig. 1.5a**). The separation that placed information storage in DNA and biological function mainly in proteins was so successful that all known organisms alive today descend from the first organisms that happened upon this molecular specialization.

The evidence for the common origin of all living forms is present in their DNA sequences. All living organisms use essentially the same **genetic code** in which various triplet groupings of the four letters of the DNA and RNA alphabets encode the 20 letters of the amino acid alphabet (**Fig. 1.5b**).

Figure 1.5 RNA is an intermediary in the conversion of DNA information into protein via the genetic code. (a) The linear bases of DNA are copied through molecular complementarity into the linear bases of RNA. The bases of RNA are read three at a time (that is, as triplets) to encode the amino acid subunits of proteins. **(b)** The genetic code dictionary specifies the relationship between RNA triplets and the amino acid subunits of proteins.



(b)

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

Figure 1.6 Comparisons of gene products in different species provide evidence for the relatedness of living organisms. This chart shows the amino acid sequence for equivalent portions of the cytochrome c protein in six species: *Saccharomyces cerevisiae* (yeast), *Arabidopsis thaliana* (a weedlike flowering plant), *Caenorhabditis elegans* (a nematode), *Drosophila melanogaster* (the fruit fly), *Mus musculus* (the house mouse), and *Homo sapiens* (humans). Consult Fig. 1.4b for the key to amino acid names.

* Indicates identical and . indicates similar

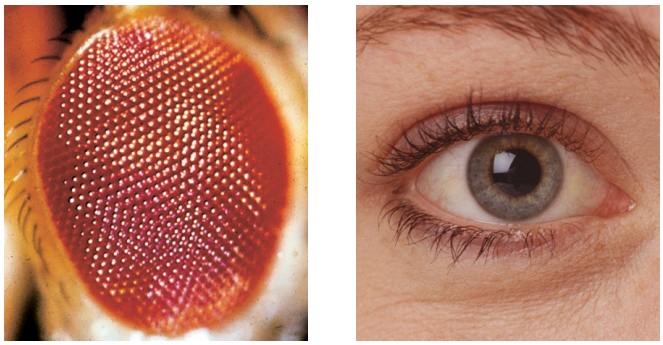
The relatedness of all living organisms is also evident from comparisons of genes with similar functions in very different organisms. A striking similarity exists between the genes for many corresponding proteins in bacteria, yeast, plants, worms, flies, mice, and humans. For example, most of the amino acids in the cytochrome c proteins of diverse species are identical to each other (**Fig. 1.6**), indicating that these proteins all derived from a common ancestral protein. It is also important to note that some amino acids in these various cytochrome c proteins are different. The reason is that different **mutations,** that is, changes in nucleotide pairs, can occur when genes are passed from one generation of an organism to the next. The accumulation of these mutations in genomes is the main driving force of evolution.

Despite the occurrence of mutations that alter DNA and thus protein sequences, it is often possible to place a gene from one organism into the genome of a very different organism and see it function normally in the new environment. Human genes that help regulate cell division, for example, can replace related genes in yeast and enable the yeast cells to function normally.

One of the most striking examples of relatedness at this level of biological information was uncovered in studies of eye development. Both insects and vertebrates (including humans) have eyes, but they are of very different types (**Fig. 1.7**). Biologists had long assumed that the evolution of eyes occurred independently, and in many evolution textbooks, eyes were used as an example of *convergent evolution,* in which structurally unrelated but functionally analogous organs emerge in different species as a result of natural selection. Studies of a gene called *Pax6* have turned this view upside down.

Mutations in the *Pax6* gene lead to a failure of eye development in both people and mice, and molecular studies have suggested that *Pax6* might play a central role in the initiation of eye development in all vertebrates. Remarkably, when the human *Pax6* gene is expressed in cells along the surface of the fruit fly body, it induces numerous little eyes to develop there. It turns out that fruit flies also have a gene specifying

Figure 1.7 The eyes of insects and humans have a common ancestor. (a) A fly eye and **(b)** a human eye.



(a) (b)

a protein whose amino acid sequence is distantly but clearly related to that of the protein specified by human *Pax6*; and furthermore, certain mutations in the fly gene result in animals lacking eyes. Taken together, these results demonstrate that during 600 million years of *divergent evolution,* an ancestral gene that served as the main control switch for initiating eye development accumulated different mutations in the lineages leading to people and fruit flies, but the gene still serves the same function in both species.

The usefulness of the relatedness and unity at all levels of biological information cannot be overstated. It means that in many cases, the experimental manipulation of organisms known as *model organisms* can shed light on gene functions in humans. If genes similar to human genes function in simple model organisms such as fruit flies or bacteria, scientists can determine gene function and regulation in these experimentally manipulable organisms and bring these insights to an understanding of the human organism.

essential concepts

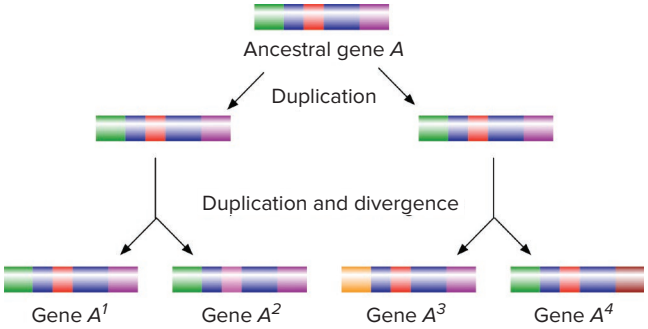
- Living organisms exhibit marked similarities at the molecular level in the ways they use DNA and RNA to make proteins.
- Certain genes have persisted throughout the evolution of widely divergent species.

1.4 The Modular Construction of Genomes

learning objectives

1. Describe mechanisms by which new genes could arise.
2. Explain how regulation of gene expression can alter gene function.

Figure 1.8 How genes arise by duplication and divergence. Ancestral gene *A* contains exons (*green, red,* and *purple*) separated by introns in *blue*. Gene *A* is duplicated to create two copies that are originally identical, but mutations in either or both (*other colors*) cause the copies to diverge. Additional rounds of duplication and divergence create a family of related genes.

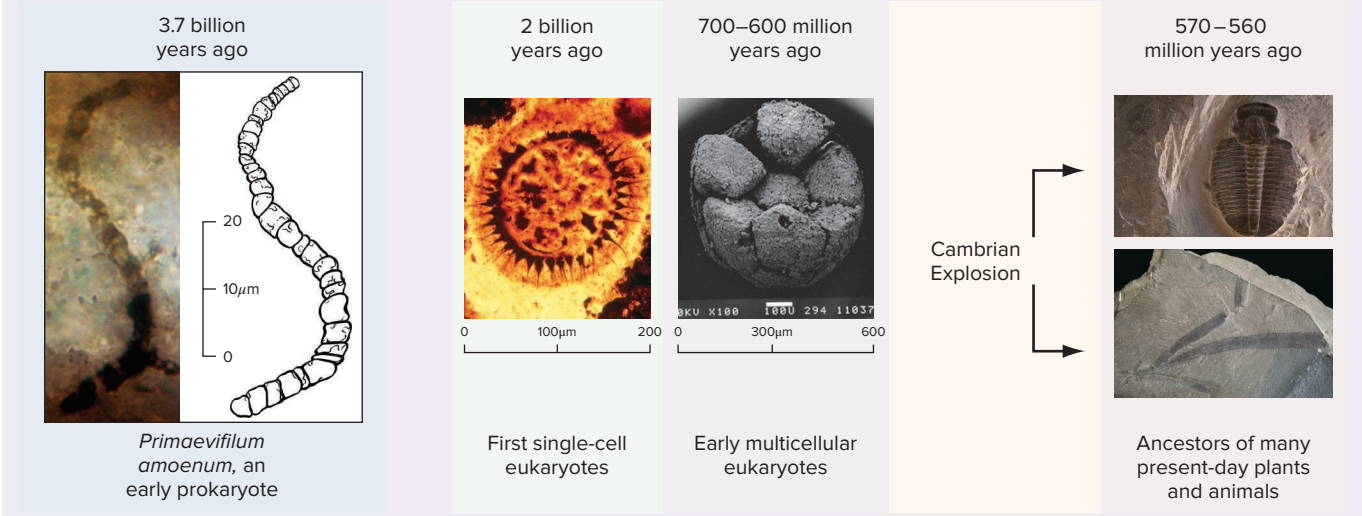


We have seen that roughly 27,000 genes direct human growth and development. How did such complexity arise? Recent technical advances have enabled researchers to complete structural analyses of the entire genome of many organisms. The information obtained reveals that **gene families** have arisen by duplication of a primordial gene; after duplication, mutations may cause the two copies to diverge from each other (**Fig. 1.8**). In both humans and chimpanzees, for example, four different genes produce different rhodopsin proteins that are expressed in photoreceptors of distinct retinal cells. Each of these proteins functions in a slightly different way so that four kinds of retinal cells respond to light of different wavelengths and intensities, resulting in color vision. The four rhodopsin genes arose from a single primordial gene by several duplications followed by slight divergences in structure.

Duplication followed by divergence underlies the evolution of new genes with new functions. This principle appears to have been built into the genome structure of all multicellular organisms. The protein-coding region of most genes is subdivided into as many as 10 or more small pieces (called *exons*), separated by DNA that does not code for protein (called *introns*) as shown in Fig. 1.8. This modular construction facilitates the rearrangement of different modules from different genes to create new combinations during evolution. It is likely that this process of modular reassortment facilitated the rapid diversification of living forms about 570 million years ago (see Fig. 1.8).

The tremendous advantage of the duplication and divergence of existing pieces of genetic information is evident in the history of life's evolution (**Table 1.1**). *Prokaryotic* cells such as bacteria, which do not have a membrane-bounded nucleus, evolved about 3.7 billion years ago; *eukaryotic* cells such as algae, which have a membrane-bounded nucleus, emerged around 2 billion years ago; and multicellular eukaryotic organisms appeared 700–600 million years ago. Then, about 570 million years

TABLE 1.1 Fossil Evidence for Some Major Stages in the Evolution of Life



(prokaryote): © J.W. Schopf; (eukaryotes): © Prof. Andrew Knoll; (trilobite): © Brand X Pictures/PunchStock RF; (sponge): © Alan Sirulnikoff/Science Source

ago, within the relatively short evolutionary time of roughly 20–50 million years known as the *Cambrian explosion*, the multicellular life-forms diverged into a bewildering array of organisms, including primitive vertebrates.

Figure 1.9 Two-winged and four-winged flies. Geneticists converted a contemporary normal two-winged fly to a four-winged insect resembling the fly's evolutionary antecedent. They accomplished this by mutating a key element in the fly's regulatory network. Note the club-shaped halteres (*arrows*) behind the wings of the fly at the top.

(both): © Edward Lewis, California Institute of Technology

A fascinating question is: How could the multicellular forms achieve such enormous diversity in only 20–50 million years? The answer lies, in part, in the hierarchic organization of the information encoded in chromosomes. Exons are arranged into genes; genes duplicate and diverge to generate gene families; and gene families sometimes rapidly expand to **gene superfamilies** containing hundreds of related genes. In both mouse and human adults, for example, the immune system is encoded by a gene superfamily composed of hundreds of closely related but slightly divergent genes. With the emergence of each successively larger informational unit, evolution gains the ability to duplicate increasingly complex informational modules through single genetic events.

Probably even more important for the evolution of complexity is the rapid change of regulatory networks that specify how genes behave (that is, when, where, and to what degree they are expressed) during development. For example, the two-winged fly evolved from a four-winged ancestor not because of changes in gene-encoded structural proteins, but rather because of a rewiring of the regulatory network, which converted one pair of wings into two small balancing organs known as *halteres* (**Fig. 1.9**).

essential concepts

- *Gene duplication* followed by the *divergence* of copies is one explanation for how new functions evolve.
- The *reshuffling of exons* in eukaryotes provides another mechanism for the rapid diversification of genomes.
- Changes in DNA that affect *gene regulation*—where, when, and to what degree genes are expressed—also generate evolutionary change.

1.5 Modern Genetic Techniques

learning objectives

1. Explain how advances in technology have accelerated the analysis of genomes.
2. Compare the knowledge obtained from genetic dissection and from genome sequencing.
3. Discuss how genome sequence information can be used to treat or cure diseases.

The complexity of living systems has developed over 4 billion years from the continuous amplification and refinement of genetic information. The simplest bacterial cells contain about 1000 genes that interact in complex networks. Yeast cells, the simplest eukaryotic cells, contain about 5800 genes. Nematodes (roundworms) contain about 20,000 genes, and fruit flies contain roughly 13,000 genes. Humans have approximately 27,000 genes; surprisingly, the flowering plant *Arabidopsis* has as many and the zebrafish *D. rerio* has even more (**Fig. 1.10**). Each of these organisms has provided valuable insights into aspects of biology that are conserved among all organisms as well as other phenomena that are species-specific.

Genetic Dissection of Model Organisms Reveals the Working of Biological Processes

Model organisms including bacteria, yeast, nematodes, fruit flies, *Arabidopsis*, zebrafish, and mice are extremely valuable to researchers, who can use these organisms to analyze the complexity of a genome piece by piece. The logic used in *genetic dissection* is quite simple: inactivate a gene in a model organism and observe the consequences. For example, the loss of a gene for visual pigment produces fruit flies with white eyes instead of eyes of the normal red color. One can thus conclude that the protein product of this gene plays a key role in eye pigmentation. From their study of model organisms, geneticists are amassing a detailed picture of the complexity of living systems.

Whole-Genome Sequencing Can Identify Mutant Genes that Cause Disease

A complementary way to study an organism's genetic complexity is to look not just at one gene at a time, but rather at the genome as a whole. The new tools of **genomics,** particularly high-throughput DNA sequencers, have the capacity to analyze all the genes of any living thing. In fact, the complete nucleotide sequences of representative genomes of the model species listed above, as well as of humans, have all been determined.

The first draft of the human genome sequence announced by the Human Genome Project in 2001 cost $3 billion and took over 10 years to produce. Since then, rapid advances in genome sequencing technology have made it possible in 2016 to determine the genome sequence of an individual in just a few days for about $1000. Alongside the advances in DNA sequencing technology have been the development of computer algorithms to analyze the sequence data and the establishment of online databases that catalog the differences in individual genome sequences.

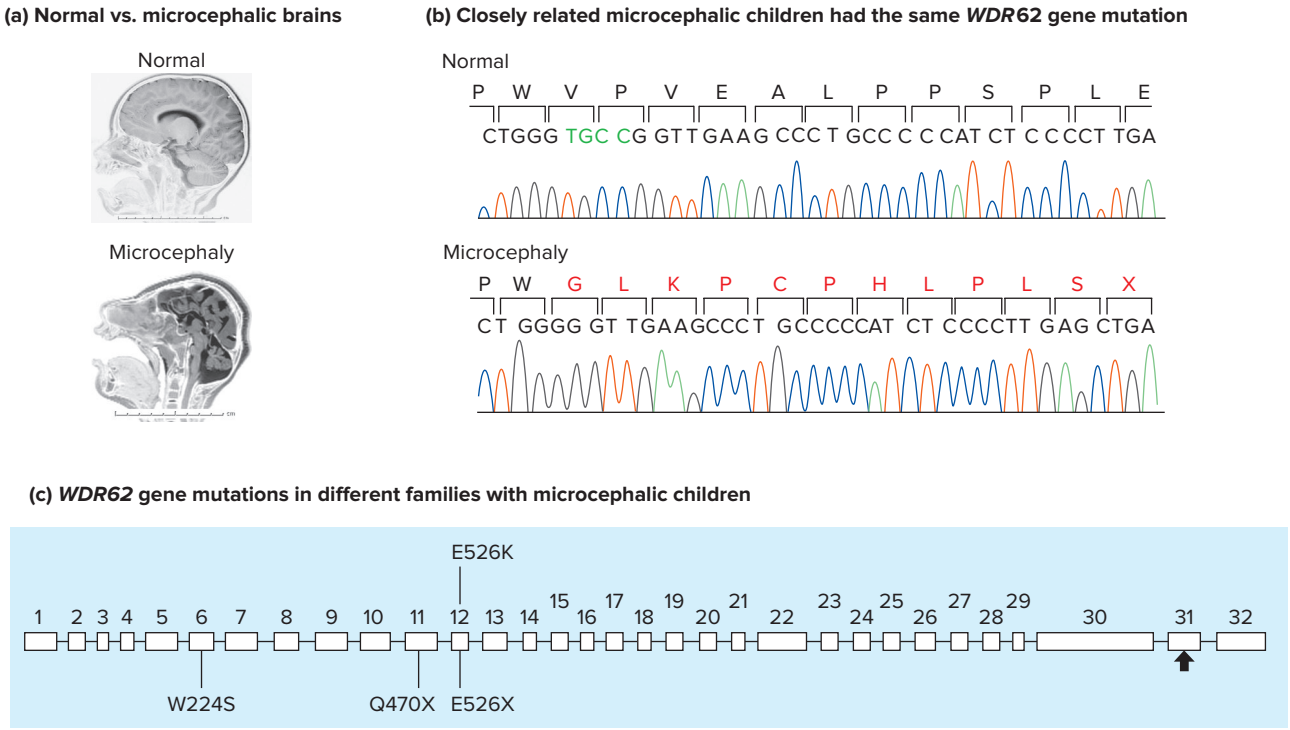
No example better illustrates the power of genome sequencing technology than its use in the identification of gene mutations that cause human genetic diseases. For diseases that result from mutation of a single gene, the gene responsible often may be identified by determining the genome sequence of just a few people or sometimes even that of a single individual.

In the case shown in **Fig. 1.11a,** geneticists analyzed whole genome sequences to find a gene mutation underlying a rare brain malformation disease called *microcephaly.* The

Figure 1.10 Seven model organisms whose genomes were sequenced as part of the Human Genome Project. The chart indicates genome size in millions of base pairs, or megabases (Mb). The bottom row shows the approximate number of genes for each organism. (*E. coli*): © David M. Phillips/Science Source; (*S. cerevisiae*): © CMSP/Getty Images; (*C. elegans*): © Sinclair Stammers/Science Source; (*A. thaliana*): Source: Courtesy USDA/Peggy Greb, photographer; (*D. melanogaster*): © Hemis.fr/SuperStock; (*D. rerio*): © A Hartl/Blickwinkel/agefotostock; (*M. musculus*): © imageBROKER/SuperStock RF

Organism	*E. coli*	*S. cerevisiae*	*C. elegans*	*A. thaliana*	*D. melanogaster*	*D. rerio*	*M. musculus*
Genome size	4.6 Mb	12 Mb	100 Mb	125 Mb	130 Mb	1500 Mb	2700 Mb
Number of protein-coding genes (approximate)	4300	5800	20,000	27,000	13,000	36,000	25,000

Figure 1.11 A causal gene for microcephaly identified by genome sequencing. **(a)** Magnetic resonance images of normal and microcephalic brains. **(b)** Sequence analysis of normal and mutant copies of the *WDR62* gene. The mutation is a deletion of the four nucleotides TGCC (*green*) that causes a major change in the amino acid sequence of the protein product of the gene. The letters above each triplet sequence identify the encoded amino acid. **(c)** Five different mutations in the *WDR62* gene in five different families are shown. Four of the mutations affect the identity of a single amino acid in the protein encoded by the gene. For example, W224S means that the 224th amino acid is normally W (tryptophan) but is changed to S (serine) by the mutation. The *arrow* indicates the position of the TGCC deletion mutation shown in (b).



inheritance pattern observed for microcephaly indicated that it is a so-called *recessive* disease, meaning that diseased people inherit two mutant gene copies, one from each normal parent. The parents have one normal gene copy and one mutant copy, explaining why the parents' brains do not have this malformation. Sequencing and analysis of the genomes of two children with microcephaly from the same family identified a single rare gene mutation present in both siblings—a deletion of four base pairs within a gene called *WDR62* (**Fig. 1.11b**). Each parent was found to have one normal copy of *WDR62* and one copy with the four base pair deletion (Fig. 1.11b). Subsequently, the researchers found that different families with microcephalic children harbored different mutations in the same gene (**Fig. 1.11c**), thus confirming *WDR62* as a causative gene for microcephaly.

Gene Therapy May Help Cure Genetic Diseases

By enabling rapid disease gene discovery, the sequencing of whole genomes is revolutionizing medicine. Knowledge of disease genes can inform parents whether their children might suffer from a severely debilitating disease like microcephaly, allowing the parents to consider ways in which to avoid or to prepare for such an outcome. Moreover, identification of a disease gene provides information about the protein encoded by that gene that can sometimes guide the design of effective therapeutics to treat the disease. This strategy has not yet been useful in the case of microcephaly, but it has already been of tremendous value in developing drugs to combat other genetic diseases, including certain kinds of cancer.

Dramatic progress within the last few years offers hope that medical scientists may eventually be able to treat genetic diseases by modifying the genomes of the *somatic* (body) *cells* affected by the disease syndrome. One method under development is called **gene therapy;** here, scientists introduce normal copies of genes into human cells, where they can be expressed and compensate for their mutant, nonfunctional counterparts in the genome. An alternative, very recent gene therapy approach is *genome editing,* in which researchers change the base-pair sequence of a mutant gene to that of its normal counterpart. Gene therapy and genome editing have been used in model organisms such as mice to restore proper gene function and sometimes reverse the disease process, but the application of these techniques to human conditions is still only in very early stages.

essential concepts

- Scientists analyze mutations in model organisms to understand the molecular basis of many biological processes.
- Automated sequencing and computer analysis have made possible the rapid determination of DNA sequences in genomes, allowing researchers to pinpoint genes responsible for genetic diseases.
- Knowledge of disease genes helps parents make informed reproductive decisions, allows pharmaceutical companies to design effective drugs, and in the future will enable medical researchers to manipulate somatic cell genomes to reverse disease effects.

1.6 Human Genetics and Society

learning objectives

1. Describe the types of information that can be obtained from an individual's genome sequence.
2. Discuss the social issues that arise from the availability of personal genome sequences.

Over the next 25 years, geneticists will identify hundreds of genes with variations that predispose people to many types of disease: cardiovascular, cancerous, immunological, mental, metabolic. Some mutations will always cause disease, as in the microcephaly example just discussed; others will only predispose to disease. For example, a change in a specific single DNA base pair in the *β-globin* gene will nearly always cause sickle-cell anemia, a painful, life-threatening condition that leads to severe anemia. By contrast, a mutation in the *breast cancer 1 (BRCA1)* gene increases the risk of breast cancer to between 40% and 80% in a woman carrying one copy of the mutation. This conditional state arises because the *BRCA1* gene interacts with environmental factors that affect the probability of activating the cancerous condition, and because various forms of other genes can modify the effect of the *BRCA1* gene mutation.

Physicians may be able to use DNA diagnostics—a collection of techniques for characterizing genes—to analyze an individual's DNA for genes that predispose to some diseases. With this genetic profile, doctors may be able to write out a predictive health history based on probabilities for some medical conditions.

Many Social Issues Need to Be Addressed

Although biological information is similar to other types of information from a strictly technical point of view, it is as different as can be in its meaning and impact on individual human beings and on human society as a whole. The difference lies in the personal nature of the unique genetic profile carried by each person from birth. Within the genome of each individual is complex information that provides greater or lower susceptibility or resistance to many diseases, as well as greater or lesser potential for the expression of many physiological and neurological attributes that distinguish people from each other. Until now, almost all of this information has remained hidden away. But if research continues at its present pace, in less than a decade it will become possible to understand many aspects of a person's genetic make-up, and with this information will come the power to make some limited predictions about future possibilities and risks.

As you will see in many of the Genetics and Society boxes throughout this book, society can use genetic information not only to help people but also to restrict their lives (for example, by denying insurance or employment). We believe that just as our society respects an individual's right to privacy in other realms, it should also respect the privacy of an individual's genetic profile and work against all types of discrimination. Indeed, in 2008 the United States federal government passed the Genetic Information Nondiscrimination Act prohibiting insurance companies and employers from discrimination on the basis of genetic tests.

Another issue raised by the potential for detailed genetic profiles is the interpretation or misinterpretation of that information. Without accurate interpretation, the information becomes useless at best and harmful at worst. Proper interpretation of genetic information requires some understanding of statistical concepts such as risk and probability. As an example, women found to have the *BRCA1* mutation described above need to weigh the possible risks and benefits of preventive treatments such as mastectomy against the statistical probability that they would suffer otherwise from breast cancer. To help people understand these concepts, widespread education in this area will be essential.

To many people, the most frightening potential of the new genetics is the development of gene therapy technologies that can alter or add to the genes present within the *germ line* (reproductive cell precursors) of human embryos. In contrast with the previously described use of such *genetic engineering* to treat diseases in non-inheritable somatic cells, germ-line cells that have been manipulated in these ways can be transmitted between generations and thus have the potential to influence the evolution of our species.

Some people caution that developing the power to alter our own genomes is a step we should not take, arguing that if genetic information and technology are misused, as they certainly have been in the past, the consequences could be horrific. Attempts to use genetic information for social purposes were prevalent in the early twentieth century, leading to enforced sterilization of individuals thought to be inferior, to laws that prohibited interracial marriage, and to

laws prohibiting immigration of certain ethnic groups. The basis of these actions was not scientific and has been thoroughly discredited.

Others agree that we must not repeat the mistakes of the past, but they warn that if the new technologies could help children and adults lead healthier, happier lives, we cannot outlaw their application. Most agree that the biological revolution we are living through will have a greater impact on human society than any technological revolution of the past and that education and public debate are the key to preparing for the consequences of this revolution.

The focus on human genetics in this book looks forward into the new era of biology and genetic analysis. These new possibilities raise serious moral and ethical issues that will demand wisdom and humility. It is in the hope of educating young people for the moral and ethical challenges awaiting the next generation that we write this book.

essential concepts

- Genome sequences may identify not only genes that always cause disease but also genes that predispose an individual to a disease.
- As a society, we must ensure that genetic knowledge is properly interpreted and that individuals' privacy is protected.

WHAT'S NEXT

Genetics, the study of biological information, is also the study of the DNA and RNA molecules that store, replicate, transmit, express, and evolve information for the construction of proteins. At the molecular level, all living things are closely related, and as a result, observations of model organisms as different as yeast and mice can provide insights into general biological principles as well as human biology.

Remarkably, more than 75 years before the discovery of DNA, Gregor Mendel, an Augustinian monk, delineated the basic laws of gene transmission with no knowledge of the molecular basis of heredity. He accomplished this by following simple traits, such as flower or seed color, through several generations of the pea plant (*Pisum sativum*). We now know that his findings apply to all sexually reproducing organisms. Chapter 2 describes Mendel's studies and insights, which became the foundation of the field of genetics.

PROBLEMS

Vocabulary

1. Choose the phrase from the right column that best fits the term in the left column.

a. complementarity	1. a linear polymer of amino acids that folds into a particular shape
b. nucleotide	2. part of a gene that does not contain protein coding information
c. chromosomes	3. a polymer of nucleotides that is an intermediary in the synthesis of proteins from instructions in DNA
d. protein	4. G–C and A–T base pairing in DNA through hydrogen bonds
e. genome	5. alteration of DNA sequence
f. gene	6. part of a gene that can contain protein coding information
g. uracil	7. DNA/protein structures that contain genes
h. exon	8. DNA information for a single function, such as production of a protein
i. intron	9. the entirety of an organism's hereditary information
j. DNA	10. a double-stranded polymer of nucleotides that stores the inherited blueprint of an organism
k. RNA	11. subunit of the DNA macromolecule
l. mutation	12. the only one of the four bases in RNA that is not in DNA

Section 1.1

2. If one strand of a DNA molecule has the base sequence 5′-AGCATTAAGCT-3′, what is the base sequence of the other, complementary strand?
3. The size of one copy of the human genome is approximately 3 billion base pairs, and it contains about 27,000 genes organized into 23 chromosomes.
   a. Human chromosomes vary in size. What would you predict is the size of the average chromosome?
   b. Assuming that genes are spread evenly among chromosomes, how many genes does an average human chromosome contain?
   c. About half of the DNA in chromosomes contains genes. How large (in base pairs) is an average human gene?

Section 1.2

4. Indicate whether each of the following words or phrases applies to proteins, DNA, or both.
   a. a macromolecule composed of a string of subunits
   b. double-stranded
   c. four different subunits
   d. 20 different subunits
   e. composed of amino acids
   f. composed of nucleotides
   g. contains a code used to generate other macromolecules
   h. performs chemical reactions
5. a. How many different DNA strands composed of 100 nucleotides could possibly exist?
   b. How many different proteins composed of 100 amino acids could possibly exist?

Section 1.3

6. RNA shares with proteins the ability to fold into complex three-dimensional shapes. As a result, RNA molecules can, like protein molecules, catalyze biochemical reactions (that is, both kinds of molecules can act as *enzymes*, or biological catalysts). These statements are not true of DNA. Why can some RNA molecules act as enzymes whereas DNA molecules cannot? (*Hint:* Most RNA molecules consist of a single strand of nucleotides while most DNA molecules are double helixes made of two strands of nucleotides.)
7. The human protein lactate dehydrogenase shown in Fig. 1.4 has 332 amino acids. What is the smallest possible combined size of the parts of the gene that specify this protein using the genetic code?
8. a. Are the triplets in the genetic code table shown in Fig. 1.5b written as DNA or RNA?
   b. Two amino acids are each specified only by a single triplet. Identify these two amino acids and the corresponding triplets.
   c. If you know the sequence of amino acids in a protein, what does the genetic code table allow you to infer about the sequence of base pairs in the gene that specifies that protein?
9. Why do scientists think that all forms of life on earth have a common origin?
10. Why would a geneticist study a yeast cell or a fruit fly or a mouse in order to understand human genes and human biology?
11. How can a scientist tell if a protein present in bacteria and a fruit fly protein have a common origin? How can a scientist determine whether a protein with a common origin in bacteria and a fruit fly function in a common pathway?
12. Figure 1.6 shows the amino acid sequences of parts of the cytochrome c proteins from several different organisms. Some of these amino acids are highlighted in *dark orange,* some in *light orange,* and some are not highlighted at all. Which of these three classes of amino acids is likely to be most important for the biochemical function of cytochrome c proteins?

Section 1.4

13. Why do scientists think that new genes arise by duplication of an original gene and divergence by mutation?
14. Explain how the exon/intron structure of genes contributes to the generation of new gene functions during evolution.
15. Mutations in genes that change their pattern of expression (the time and cell type in which the gene product is produced) are thought to be a major factor in the evolution of different organisms. Would you expect the same protein to work in the same way (for example, to perform the same kind of enzymatic reaction) in two different types of cells (for example, cells in the retina of the eye and muscle cells)? Is it possible that the same protein might function in different biochemical pathways in eye cells and muscle cells even if the protein's basic mechanism always remains the same?

Section 1.5

16. A single zebrafish gene function was inactivated completely by mutation, and a zebrafish with this mutation had none of its normal horizontal stripes. For each of the following statements, indicate whether the

statement is certainly true, certainly untrue, or if there is insufficient information to decide.

a. The normal gene function is required for the viability of the zebrafish.

b. The normal gene function is required for the formation of stripes.

c. The normal gene function is required to make the pigment deposited in the stripes.

d. The gene is required in zebrafish only for stripe formation.

17. Different mutations in the *WDR62* gene that inactivate gene function were found in the genomes of many different people with microcephaly. This information provided strong support for the idea that the *WDR62* gene mutation causes microcephaly.

a. The human genome sequence identified *WDR62* as one of the approximately 27,000 genes in the human genome. What information about the function of *WDR62* do you think was learned originally from the DNA sequence of the normal human genome?

b. What additional information was provided by identification of *WDR62* as the microcephaly disease gene?

c. The mouse genome contains a gene similar to human *WDR62*. Experiments in mice have shown that the mouse *WDR62* gene is expressed in the brain. Technology is now available that allows scientists to generate mice in which the two normal copies of the *WDR62* gene are replaced with mutant copies of the gene that are nonfunctional. Why would a scientist want to generate such *WDR62* mutant mice?

18. Researchers have successfully used gene therapy to ameliorate some human genetic diseases by adding a normal gene copy to cells whose genomes originally had only nonfunctional mutant copies of that gene. For example, a form of blindness due to the lack of a single protein called RPE65 has been reversed by introduction of a normal *RPE65* gene to cells of the retina of adults.

a. The success of this gene therapy approach provides us with clues about the role of the RPE65 protein in the retina. Do you think that RPE65 is needed for the proper development of the human eye?

b. Can you see a potential difficulty in applying this gene therapy approach for diseases like microcephaly?

Section 1.6

19. By the time this book is published, it will likely be possible for you to obtain the sequence of your genome at nominal cost. Do you want this information? Explain the factors that affected your decision. (You may not be able to answer this question until you are finished reading this book.)

chapter 2

Mendel's Principles of Heredity

Although Mendel's laws can predict the probability that an individual will have a certain genetic makeup, the chance meeting of particular male and female gametes determines an individual's actual genetic fate.

chapter outline

A QUICK GLANCE at an extended family portrait is likely to reveal children who resemble one parent or the other, or who look like a combination of the two (**Fig. 2.1**). Some children, however, look unlike any of the assembled relatives and more like a great, great grandparent. What causes the similarities and differences of appearance and the skipping of generations?

The answers lie in our **genes,** the basic units of biological information, and in **heredity,** the way genes transmit physiological, anatomical, and behavioral traits from parents to offspring. Each of us starts out as a single fertilized egg cell that develops, by division and differentiation, into a mature adult made up of 10^{14} (a hundred trillion) specialized cells capable of carrying out all the body's functions and controlling our outward appearance. Genes, passed from one generation to the next, underlie the formation of every heritable trait. Such traits are as diverse as the presence of a cleft in your chin, the tendency to lose hair as you age, your hair, eye, and skin color, and even your susceptibility to certain cancers. All such traits run in families in predictable patterns that impose some possibilities and exclude others.

Genetics, the science of heredity, pursues a precise explanation of the biological structures and mechanisms that determine inheritance. In some instances, the relationship between gene and trait is remarkably simple. A single change in a single gene, for example, results in sickle-cell anemia, a disease in which the hemoglobin molecule found in red blood cells is defective. In other instances, the correlations between genes and traits are bewilderingly complex. An example is the genetic basis of facial features, in which many genes determine a large number of molecules that interact to generate the combination we recognize as a friend's face.

Gregor Mendel (1822–1884; **Fig. 2.2**), a stocky, bespectacled Augustinian monk and expert plant breeder, discovered the basic principles of genetics in the mid-nineteenth century. He published his findings in 1866, just seven years after Darwin's *On the Origin of Species* appeared in print. Mendel lived and worked in Brünn, Austria (now Brno in the Czech Republic), where he examined the inheritance of clear-cut alternative traits in pea plants, such as purple versus white flowers or yellow versus green seeds. In so

Figure 2.1 A family portrait. The extended family shown here includes members of four generations.

Figure 2.2 Gregor Mendel. Photographed around 1862 holding one of his experimental plants.

Figure 2.3 Like begets like and unlike. A Labrador retriever with her litter of pups.

doing, he inferred genetic laws that allowed him to make verifiable predictions about which traits would appear, disappear, and then reappear, and in which generations.

Mendel's laws are based on the hypothesis that observable traits are determined by independent units of inheritance not visible to the naked eye. We now call these units *genes*. The concept of the gene continues to change as research deepens and refines our understanding. Today, a gene is recognized as a region of DNA that encodes a specific protein or a particular type of RNA. In the beginning, however, it was an abstraction—an imagined particle with no physical features, the function of which was to control a visible trait by an unknown mechanism.

We begin our study of genetics with a detailed look at what Mendel's laws are and how they were discovered. In subsequent chapters, we discuss logical extensions to these laws and describe how Mendel's successors grounded the abstract concept of hereditary units (genes) in an actual biological molecule (DNA).

Four general themes emerge from our detailed discussion of Mendel's work. The first is that variation, as expressed in alternative forms of a trait, is widespread in nature. This genetic diversity provides the raw material for the continuously evolving variety of life we see around us. Second, observable variation is essential for following genes from one generation to the next. Third, variation is not distributed solely by chance; rather, it is inherited according to genetic laws that explain why like begets both like and unlike. Dogs beget other dogs—but hundreds of breeds of dogs are known. Even within a breed, such as Labrador retrievers, genetic variation exists: Two black dogs could have a litter of black, chocolate (brown), and yellow puppies (**Fig. 2.3**). Mendel's insights help explain why this is so. Fourth, the laws Mendel discovered about heredity apply equally well to all sexually reproducing organisms, from protozoans to peas to dogs to people.

2.1 The Puzzle of Inheritance

learning objectives

1. Relate how Mendel's experimental approach is similar to the process of modern scientific inquiry.
2. Describe how Mendel cross-fertilized and self-fertilized pea plants.
3. Explain the importance of Mendel's inclusion of reciprocal crosses within his controlled breeding program of pea plants.
4. Predict the type of progeny produced by Mendel's crosses between pure-breeding plants with discrete, antagonistic traits, such as purple versus white flowers.

Several steps lead to an understanding of genetic phenomena: the careful observation over time of groups of organisms, such as human families, herds of cattle, or fields of

corn or tomatoes; the rigorous analysis of systematically recorded information gleaned from these observations; and the development of a theoretical framework that can explain the origins and relationships of these phenomena. In the mid-nineteenth century, Gregor Mendel became the first person to combine data collection, analysis, and theory in a successful pursuit of the true basis of heredity. For many thousands of years before that, the only genetic practice was the selective breeding of domesticated plants and animals, with no guarantee of what a particular mating would produce.

Artificial Selection Was the First Applied Genetic Technique

A rudimentary use of genetics was the driving force behind a key transition in human civilization, allowing hunters and gatherers to settle in villages and survive as shepherds and farmers. Even before recorded history, people practiced applied genetics as they domesticated plants and animals for their own uses. From a large litter of semitamed wolves, for example, they sent the savage and the misbehaving to the stew pot while sparing the alert sentries and friendly companions for longer life and eventual mating. As a result of this **artificial selection**—purposeful control over mating by choice of parents for the next generation—the domestic dog (*Canis lupus familiaris*) slowly arose from ancestral wolves (*Canis lupus*). The oldest bones identified indisputably as dog (and not wolf) are a skull excavated from a 20,000-year-old Alaskan settlement. Many millennia of evolution guided by artificial selection have produced massive Great Danes and minuscule Chihuahuas as well as hundreds of other modern breeds of dog. By 10,000 years ago, people had begun to use this same kind of genetic manipulation to develop economically valuable herds of reindeer, sheep, goats, pigs, and cattle that produced life-sustaining meat, milk, hides, and wools.

Farmers also carried out artificial selection of plants, storing seed from the hardiest and tastiest individuals for the next planting, eventually obtaining strains that grew better, produced more, and were easier to cultivate and harvest. In this way, scrawny, weed-like plants gradually, with human guidance, turned into rice, wheat, barley, lentils, and dates in Asia; corn, squash, tomatoes, potatoes, and peppers in the Americas; yams, peanuts, and gourds in Africa. Later, plant breeders recognized male and female organs in plants and carried out artificial pollination. An Assyrian frieze carved in the ninth century B.C., pictured in **Fig. 2.4,** is the oldest known visual record of this kind of genetic experiment. It depicts priests brushing the flowers of female date palms with selected male pollen. By this method of artificial selection, early practical geneticists produced several hundred varieties of dates, each differing in specific observable qualities, such as the fruit's size, color, or taste. A 1929 botanical survey of three oases in Egypt turned up 400 varieties of date-bearing palms, twentieth-century evidence of the natural and artificially generated variation among these trees.

Figure 2.4 The earliest known record of applied genetics. In this 2800-year-old Assyrian relief from the Northwest Palace of Assurnasirpal II (883–859 B.C.), priests wearing bird masks artificially pollinate flowers of female date palms.

Desirable Traits Sometimes Disappear and Reappear

In 1822, the year of Mendel's birth, what people in Austria understood about the basic principles of heredity was not much different from what the people of ancient Assyria had known. By the nineteenth century, plant and animal breeders had created many strains in which offspring often carried a prized parental trait. Using such strains, breeders could produce plants or animals with desired characteristics for food and fiber, but they could not always predict why a valued trait would sometimes disappear and then reappear in only some offspring.

For example, selective breeding practices had resulted in valuable flocks of merino sheep producing large quantities of soft, fine wool, but at the 1837 annual meeting of the Moravian Sheep Breeders Society, one breeder's dilemma epitomized the state of the art. He possessed an outstanding ram that would be priceless "if its advantages are inherited by its offspring," but "if they are not inherited, then it is

worth no more than the cost of wool, meat, and skin." Which would it be? According to the meeting's recorded minutes, current breeding practices offered no definite answers. In his concluding remarks at this sheep-breeders meeting, the Abbot Cyril Napp pointed to a possible way out. He proposed that breeders could improve their ability to predict what traits would appear in the offspring by finding the answers to three basic questions: What is inherited? How is it inherited? What is the role of chance in heredity?

This quandary is where matters stood in 1843 when 21-year-old Gregor Mendel entered the monastery in Brünn, presided over by the same Abbot Napp. Although Mendel was a monk trained in theology, he was not a rank amateur in science. The province of Moravia, in which Brünn was located, was a center of learning and scientific activity. Mendel was able to acquire a copy of Darwin's *On the Origin of Species* shortly after it was translated into German in 1863. Abbot Napp, recognizing Mendel's intellectual abilities, sent him to the University of Vienna—all expenses paid—where he prescribed his own course of study. Mendel's choices were an unusual mix: physics, mathematics, chemistry, botany, paleontology, and plant physiology. Christian Doppler, discoverer of the Doppler effect, was one of his teachers. The cross-pollination of ideas from several disciplines would play a significant role in Mendel's discoveries. One year after he returned to Brünn, he began his series of seminal genetic experiments. **Figure 2.5** shows where Mendel worked and the microscope he used.

Figure 2.5 Mendel's garden and microscope. **(a)** Gregor Mendel's garden was part of his monastery's property in Brünn. **(b)** Mendel used this microscope to examine plant reproductive organs and to pursue his interests in natural history.

(a)

(b)

Mendel Devised a New Experimental Approach

Before Mendel, many misconceptions clouded people's thinking about heredity. Two of the prevailing errors were particularly misleading. The first was that one parent contributes most to an offspring's inherited features; Nicolaas Hartsoeker, one of the earliest microscopists, contended in 1694 that it was the male, by way of a fully formed *homunculus* inside the sperm (**Fig. 2.6**). Another deceptive notion was the concept of *blended inheritance*, the idea that parental traits become mixed and forever changed in the offspring, as when blue and yellow pigments merge to green on a painter's palette. The theory of blending may have grown out of a natural tendency for parents to see a combination of their own traits in their offspring. While blending could account for children who look like a combination of their parents, it could not explain obvious differences between biological brothers and sisters nor the persistence of variation within extended families.

The experiments Mendel devised would lay these myths to rest by providing precise, verifiable answers to the three questions Abbot Napp had raised almost 15 years earlier: What is inherited? How is it inherited? What is the

Figure 2.6 The homunculus: A misconception. Well into the nineteenth century, many prominent microscopists believed they saw a fully formed, miniature fetus crouched within the head of a sperm.

role of chance in heredity? A key component of Mendel's breakthrough was the way he set up his experiments.

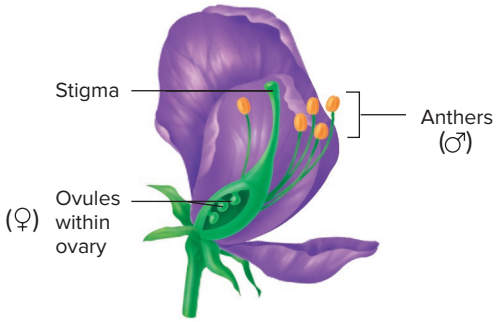
What did Mendel do differently from those who preceded him? First, he chose the garden pea *(Pisum sativum)* as his experimental organism (**Figs. 2.7a** and **b**). Peas grew well in Brünn, and with male and female organs in the same flower, they were normally self-fertilizing. In **self-fertilization** (or **selfing**), both egg and pollen come from the same plant. The particular anatomy of pea flowers, however, makes it easy to prevent self-fertilization and instead to **cross-fertilize** (or *cross*) two individuals by brushing pollen from one plant onto a female organ of another plant, as illustrated in **Fig. 2.7c.** Peas offered yet another advantage. For each successive generation, Mendel could obtain large numbers of individuals within a relatively short growing season. By comparison, if he had worked with sheep, each mating would have generated only a few offspring and the time between generations would have been several years.

Second, Mendel examined the inheritance of clear-cut alternative forms of particular traits—purple versus white flowers, yellow versus green peas. Using such either-or traits, he could distinguish and trace unambiguously the transmission of one or the other observed characteristic, because no intermediate forms existed. (The opposite of these so-called **discrete traits** are **continuous traits,** such as height and skin color in humans. Continuous traits show many intermediate forms.)

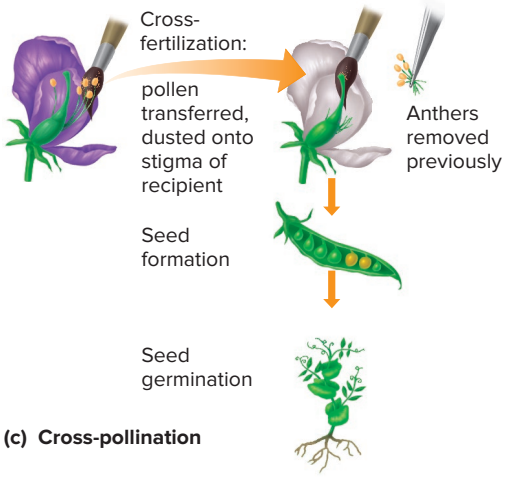
Third, Mendel collected and perpetuated lines of peas that bred true. Matings within such **pure-breeding** (or **true-breeding) lines** produce offspring carrying specific parental traits that remain constant from generation to generation. These lines are also called *inbred* because they have been mated only to each other for many generations. Mendel observed his pure-breeding lines for up to eight generations. Plants with white flowers always produced offspring with white flowers; plants with purple flowers produced only offspring with purple flowers. Mendel called constant but mutually exclusive, alternative traits,

Figure 2.7 Mendel's experimental organism: The garden pea. (a) Pea plants with white flowers. **(b)** Pollen is produced in the anthers. Mature pollen lands on the stigma, which is connected to the ovary (which becomes the pea pod). After landing, the pollen grows a tube that extends through the stigma to one of the ovules (immature seeds), allowing fertilization to take place. **(c)** To prevent self-fertilization, breeders remove the anthers from the female parents (here, the white flower) before the plant produces mature pollen. Pollen is then transferred with a paintbrush from the anthers of the male parent (here, the purple flower) to the stigma of the female parent. Each fertilized ovule becomes an individual pea (mature seed) that can grow into a new pea plant. All of the peas produced from one flower are encased in the same pea pod, but these peas form from different pollen grains and ovules.

(a) ***Pisum sativum***



(b) Pea flower anatomy



(c) Cross-pollination

such as purple versus white flowers or yellow versus green seeds *antagonistic pairs*, and he settled on seven such pairs for his study (**Fig. 2.8**). In his experiments, Mendel not only perpetuated pure-breeding stocks for each member of a pair, but he also cross-fertilized pairs of plants to produce **hybrids,** offspring of genetically dissimilar parents, for each pair of antagonistic traits. Figure 2.8 shows the appearance of the hybrids he studied.

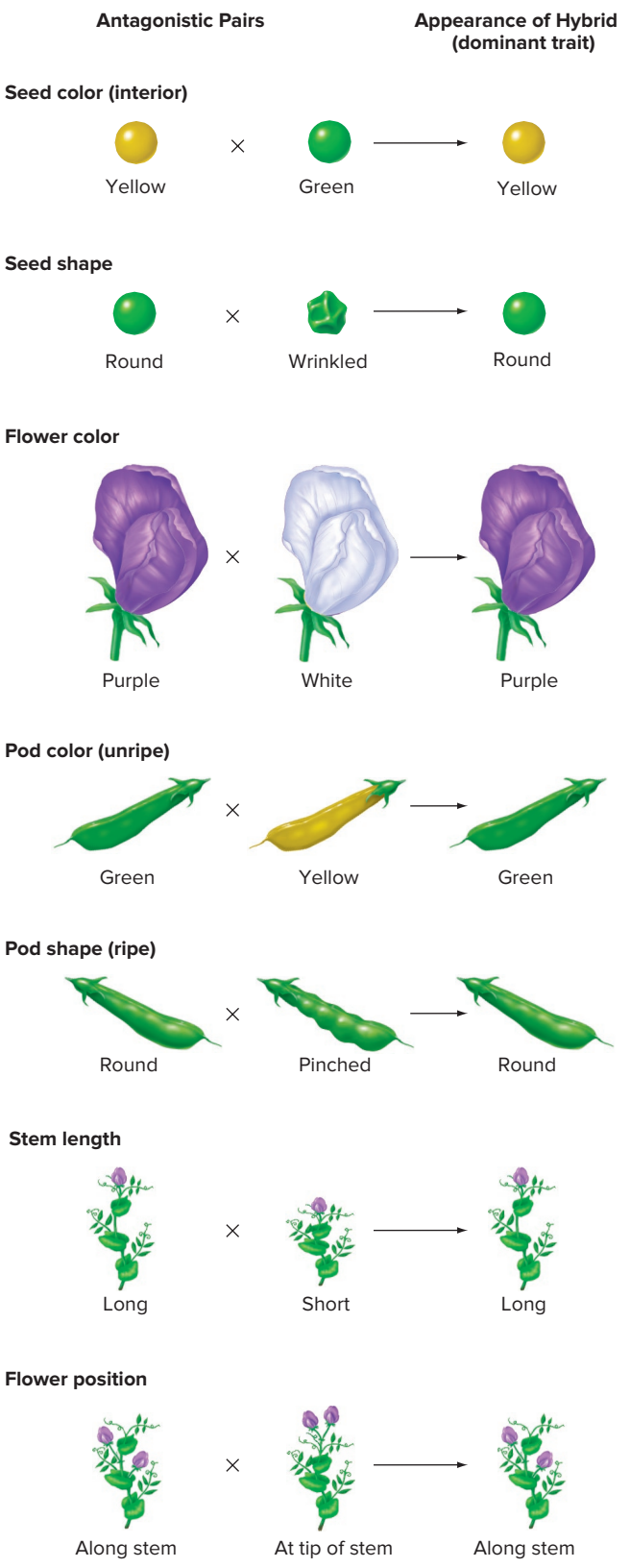
Fourth, being an expert plant breeder, Mendel carefully controlled his matings, going to great lengths to ensure that the progeny he observed really resulted from the specific fertilizations he intended. Thus Mendel painstakingly prevented the intrusion of any foreign pollen and assured self- or cross-pollination as the experiment demanded. Not only did this allow him to carry out controlled breedings of selected traits, but he could also make **reciprocal crosses.** In such crosses, he reversed the traits of the male and female parents, thus controlling whether a particular trait was transmitted via the egg cell within the ovule or via a sperm cell within the pollen. For example, he could use pollen from a purple flower to fertilize the eggs of a white flower and also use pollen from a white flower to fertilize the eggs of a purple flower. Because the progeny of these reciprocal crosses were similar, Mendel demonstrated that the two parents contribute equally to inheritance. "It is immaterial to the form of the hybrid," he wrote, "which of the parental types was the seed or pollen plant."

Fifth, Mendel worked with large numbers of plants, counted all offspring, subjected his findings to numerical analysis, and then compared his results with predictions based on his models. He was the first person to study inheritance in this manner, and no doubt his background in physics and mathematics contributed to this quantitative approach. Mendel's careful numerical analysis revealed patterns of transmission that reflected basic laws of heredity.

Finally, Mendel was a brilliant practical experimentalist. When comparing tall and short plants, for example, he made sure that the short ones were out of the shade of the tall ones so their growth would not be stunted. Eventually he focused on certain traits of the pea seeds themselves, such as their color or shape, rather than on traits of the plants arising from the seeds. In this way, he could observe many more individuals from the limited space of the monastery garden, and he could evaluate the results of a cross in a single growing season.

In short, Mendel purposely set up a simplified black-and-white experimental system and then figured out how it worked. He did not look at the vast number of variables that determine the development of a prize ram nor at the origin of differences between species. Rather, he looked at discrete traits that came in two mutually exclusive forms and asked questions that could be answered by observation and computation.

Figure 2.8 The mating of parents with antagonistic traits produces hybrids. Note that each of the hybrids for the seven antagonistic traits studied by Mendel resembles only one of the parents. The parental trait that shows up in the hybrid is known as the *dominant trait.*



essential concepts

- People practiced *artificial selection* of crop plants and domestic animals for thousands of years without understanding how heredity works.
- Mendel established *pure-breeding* lines of peas in which a specific characteristic would remain constant from one generation to the next.
- When Mendel crossed pure-breeding lines with alternative traits, the *hybrid* progeny always had the characteristics of one parent.
- In Mendel's experiments, the hybrid progeny produced by reciprocal cross-fertilizations had the same characteristics; it did not matter which parent was male and which was female.

2.2 Genetic Analysis According to Mendel

learning objectives

1. Explain Mendel's law of segregation and how it predicts the 3:1 dominant-to-recessive phenotypic ratio among the F_2 generation of a monohybrid cross.
2. Distinguish between a monohybrid cross and a testcross.
3. Explain Mendel's law of independent assortment and how the 9:3:3:1 phenotypic ratio among the F_2 of a dihybrid cross provides evidence for this law.
4. Interpret phenotypic ratios of progeny to infer how particular traits are inherited.
5. Predict the genotypic and phenotypic ratios among progeny of complex multihybrid crosses using simple rules of probability.
6. Cite the most common molecular explanations for dominant and recessive alleles.

In early 1865 at the age of 43, Gregor Mendel presented a paper entitled *Experiments on Plant Hybrids* before the Natural Science Society of Brünn. Despite its modest heading, this was a scientific paper of uncommon clarity and simplicity that summarized a decade of original observations and experiments. In it Mendel describes in detail the transmission of visible characteristics in pea plants, defines unseen but logically deduced units (genes) that determine when and how often these visible traits appear, and analyzes the behavior of genes in simple mathematical terms to reveal previously unsuspected principles of heredity.

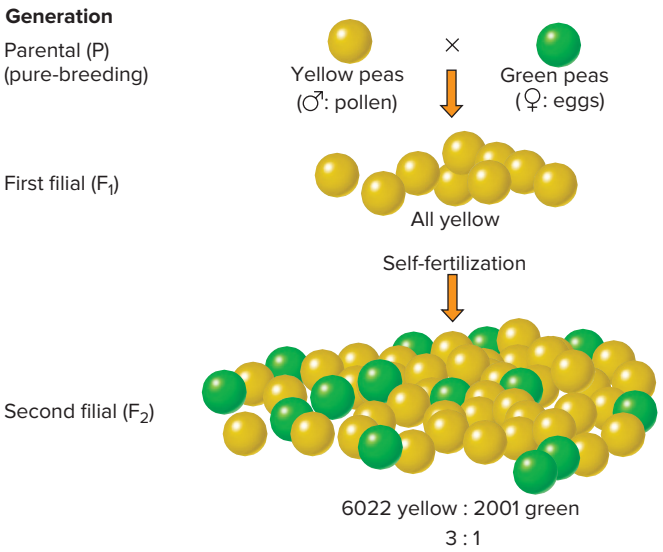
Published the following year, the paper would eventually become the cornerstone of modern genetics. Its stated purpose was to see whether there is a "generally applicable law governing the formation and development of hybrids." Let us examine its insights.

Monohybrid Crosses Reveal the Law of Segregation

Once Mendel had isolated pure-breeding lines for several sets of characteristics, he carried out a series of matings between individuals that differed in only one trait, such as seed color or stem length. In each cross, one parent carries one form of the trait, and the other parent carries an alternative form of the same trait. **Figure 2.9** illustrates one such mating. Early in the spring of 1854, for example, Mendel planted pure-breeding green peas and pure-breeding yellow peas and allowed them to grow into the **parental (P) generation.** Later that spring when the plants had flowered, he dusted the female stigma of green-pea plant flowers with pollen from yellow-pea plants. He also performed the reciprocal cross, dusting yellow-pea plant stigmas with green-pea pollen. In the fall, when he collected and separately analyzed the progeny peas of these reciprocal crosses, he found that in both cases, the peas were all yellow.

These yellow peas, progeny of the P generation, were the beginning of what we now call the **first filial (F_1) generation.** To learn whether the green trait had disappeared entirely or remained intact but hidden in these F_1 yellow peas, Mendel planted them to obtain mature F_1 plants that he allowed to self-fertilize. Such experiments involving hybrids for a single trait are often called **monohybrid crosses.** He then harvested and counted the peas of the resulting **second filial (F_2) generation,** progeny of the F_1 generation. Among the progeny of one series of F_1 self-fertilizations, there were 6022 yellow and 2001 green F_2 peas, an almost

Figure 2.9 Analyzing a monohybrid cross. Cross-pollination of pure-breeding parental plants produces F_1 hybrids, all of which resemble one of the parents. Self-pollination of F_1 plants gives rise to an F_2 generation with a 3:1 ratio of individuals resembling the two original parental types. For simplicity, we do not show the plants that produce the peas or that grow from the planted peas.



perfect ratio of 3 yellow : 1 green. F_1 plants derived from the reciprocal of the original cross produced a similar ratio of yellow to green F_2 progeny.

Reappearance of the recessive trait

The presence of green peas in the F_2 generation was irrefutable evidence that blending had not occurred. If it had, the information necessary to make green peas would have been lost irretrievably in the F_1 hybrids. Instead, the information remained intact and was able to direct the formation of 2001 green peas actually harvested from the second filial generation. These green peas were indistinguishable from their green grandparents.

Mendel concluded that two types of yellow peas must exist: those that breed true like the yellow peas of the P generation, and those that can yield some green offspring like the yellow F_1 hybrids. This second type somehow contains latent information for green peas. He called the trait that appeared in all the F_1 hybrids—in this case, yellow seeds—**dominant** (see Fig. 2.8) and the antagonistic green-pea trait that remained hidden in the F_1 hybrids but reappeared in the F_2 generation **recessive.** But how did he explain the 3:1 ratio of yellow to green F_2 peas?

Genes: Discrete units of inheritance

To account for his observations, Mendel proposed that for each trait, every plant carries two copies of a unit of inheritance, receiving one from its maternal parent and the other from the paternal parent. Today, we call these units of inheritance **genes.** Each unit determines the appearance of a specific characteristic. The pea plants in Mendel's collection had two copies of a gene for seed color, two copies of another for seed shape, two copies of a third for stem length, and so forth.

Mendel further proposed that each gene comes in alternative forms, and combinations of these alternative forms determine the contrasting characteristics he was studying. Today we call the alternative forms of a single gene **alleles.** The gene for pea color, for example, has yellow and green alleles; the gene for pea shape has round and wrinkled alleles. In Mendel's monohybrid crosses, one allele of each gene was dominant, the other recessive. In the P generation, one parent carried two dominant alleles for the trait under consideration; the other parent, two recessive alleles. The F_1 generation hybrids carried one dominant and one recessive allele for the trait. Individuals having two different alleles for a single trait are **monohybrids.**

The law of segregation

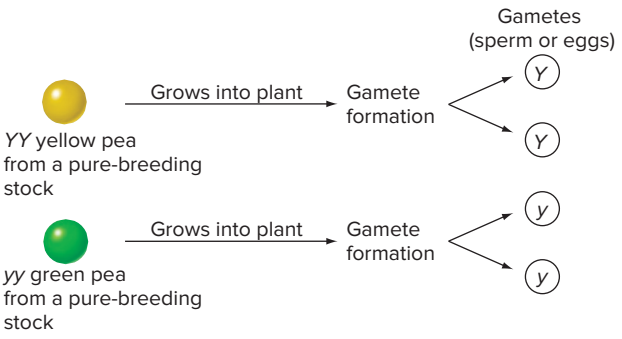
If a plant has two copies of every gene, how does it pass only one copy of each to its progeny? And how do the offspring then end up with two copies of these same genes, one from each parent? Mendel drew on his background in plant physiology and answered these questions in terms of the two biological mechanisms behind reproduction: gamete formation and the random union of gametes at fertilization.

Gametes are the specialized cells—eggs within the ovules of the female parent and sperm cells within the pollen grains—that carry genes between generations. Mendel imagined that during the formation of eggs and sperm, the two copies of each gene in the parent separate (or *segregate*) so that each gamete receives only one allele for each trait (**Fig. 2.10a**). Thus, each egg and each sperm receives only one allele for pea color (either yellow or green).

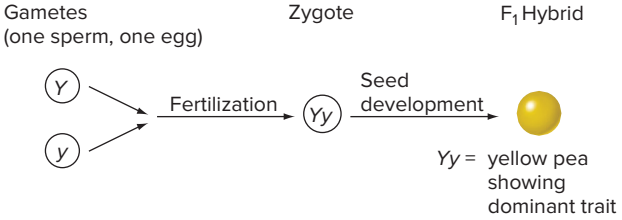
At fertilization, a sperm with one or the other allele unites at random with an egg carrying one or the other allele, restoring the two copies of the gene for each trait in the fertilized egg, or **zygote** (**Fig. 2.10b**). If the sperm carries yellow and the egg green, the result will be a hybrid yellow pea like the F_1 monohybrids that resulted when pure-breeding parents of opposite types mated. If the yellow-carrying sperm unites with a yellow-carrying egg, the result will be a yellow pea that grows into a pure-breeding plant like those of the P generation that produced only yellow peas.

Figure 2.10 The law of segregation. **(a)** The two identical alleles of pure-breeding plants separate (segregate) during gamete formation. As a result, each sperm or egg carries only one of each pair of parental alleles. **(b)** Cross-pollination and fertilization between pure-breeding parents with antagonistic traits result in F_1 hybrid zygotes with two different alleles. For the seed color gene, a *Yy* hybrid zygote will develop into a yellow pea.

(a) The two alleles for each trait separate during gamete formation.



(b) Two gametes, one from each parent, unite at random at fertilization.



Y = yellow-determining allele of pea color gene
y = green-determining allele of pea color gene

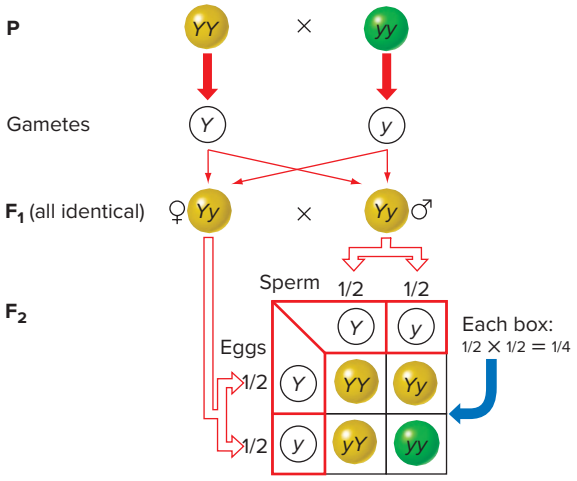
And finally, if sperm carrying the allele for green peas fertilizes a green-carrying egg, the progeny will be a pure-breeding green pea.

Mendel's **law of segregation** encapsulates this general principle of heredity: *The two alleles for each trait separate (segregate) during gamete formation, and then unite at random, one from each parent, at fertilization.* Throughout this book, the term **segregation** refers to such *equal segregation* in which one allele, and only one allele, of each gene goes to each gamete. Note that the law of segregation makes a clear distinction between the **somatic cells** (*body cells*) of an organism, which have two copies of each gene, and the **gametes,** which bear only a single copy of each gene.

The Punnett square

Figure 2.11 shows a simple way of visualizing the results of the segregation and random union of alleles during gamete formation and fertilization. Mendel invented a system of symbols that allowed him to analyze all his crosses in the same way. He designated dominant alleles with a capital *A*, *B*, or *C* and recessive ones with a lowercase *a, b*, or *c*. Modern geneticists have adopted this convention for naming genes in peas and many other organisms, but they often choose a symbol with some reference to the trait in question—a *Y* for yellow or an *R* for round. Throughout this book, we present gene symbols in italics. In Fig. 2.11, we denote the dominant yellow allele with a capital *Y* and the recessive green allele with a lowercase *y*. The pure-breeding plants of the parental generation are either *YY* (yellow peas) or *yy* (green peas). The *YY* parent can produce only *Y* gametes, the *yy* parent only *y* gametes. You can see in Fig. 2.11 why every cross between *YY* and *yy* produces exactly the same result—a *Yy* hybrid—no matter which parent (male or female) contributes which particular allele.

Figure 2.11 The Punnett square: Visual summary of a cross. This Punnett square illustrates the combinations that can arise when an F_1 hybrid undergoes gamete formation and self-fertilization. The F_2 generation should have a 3:1 ratio of yellow to green peas.



Next, to visualize what happens when the *Yy* hybrids self-fertilize, we set up a **Punnett square** (named after the British mathematician Reginald Punnett, who introduced it in 1906; Fig. 2.11). The square provides a simple and convenient method for tracking the kinds of gametes produced as well as all the possible combinations that might occur at fertilization. As the Punnett square shows in the first column and the first row, each hybrid produces two kinds of gametes, *Y* and *y*, in a ratio of 1:1. Thus, half the sperm and half the eggs carry *Y*, the other half of each gamete type carries *y*.

Each box in the Punnett square in Fig. 2.11 containing a colored pea represents one possible fertilization event. At fertilization, 1/4 of the progeny will be *YY*, 1/4 *Yy*, 1/4 *yY*, and 1/4 *yy*. Because the gametic source of an allele (egg or sperm) for the traits Mendel studied had no influence on the allele's effect, *Yy* and *yY* are equivalent. This means that 1/2 of the progeny are yellow *Yy* hybrids, 1/4 *YY* true-breeding yellows, and 1/4 true-breeding *yy* greens. The diagram illustrates how the segregation of alleles during gamete formation and the random union of egg and sperm at fertilization can produce the 3:1 ratio of yellow to green that Mendel observed in the F_2 generation.

Mendel's Results Reflect Basic Rules of Probability

Though you may not have realized it, the Punnett square illustrates two simple rules of probability—the *product rule* and the *sum rule*—that are central to the analysis of genetic crosses. These rules predict the likelihood that a particular combination of events will occur.

The product rule

The **product rule** states that the probability of two or more *independent events* occurring together is the *product* of the probabilities that each event will occur by itself. With independent events:

Probability of event 1 *and* event 2 =

Probability of event 1 × probability of event 2.

Consecutive coin tosses are obviously independent events; a heads in one toss neither increases nor decreases the probability of a heads in the next toss. If you toss two coins at the same time, the results are also independent events. A heads for one coin neither increases nor decreases the probability of a heads for the other coin. Thus, the probability of a given combination is the product of their independent probabilities. For example, the probability that both coins will turn up heads is:

$$1/2 \times 1/2 = 1/4.$$

Similarly, the formation of egg and sperm are independent events; in a hybrid plant, the probability is 1/2 that a given

gamete will carry *Y* and 1/2 that it will carry *y*. Because fertilization happens at random, the probability that a particular combination of maternal and paternal alleles will occur simultaneously in the same zygote is the product of the independent probabilities of these alleles being packaged in egg and sperm. Thus, to find the chance of a *Y* egg (formed as the result of one event) uniting with a *Y* sperm (the result of an independent event), you simply multiply $1/2 \times 1/2$ to get 1/4. This is the same fraction of *YY* progeny seen in the Punnett square of Fig. 2.11, which demonstrates that the Punnett square is simply another way of depicting the product rule. It is important to realize that each box in the Punnett square represents an equally likely outcome of the cross (an equally likely fertilization event) *only because* each of the two types of sperm and eggs (*Y* and *y*) are produced at equal frequencies.

The sum rule

While we can describe the moment of random fertilization as the simultaneous occurrence of two independent events, we can also say that two different fertilization events are mutually exclusive. For instance, if *Y* combines with *Y,* it cannot also combine with *y* in the same zygote. A second rule of probability, the **sum rule,** states that the probability of either of two such *mutually exclusive events* occurring is the *sum* of their individual probabilities. With mutually exclusive events:

Probability of event 1 *or* event 2 =

Probability of event 1 + probability of event 2.

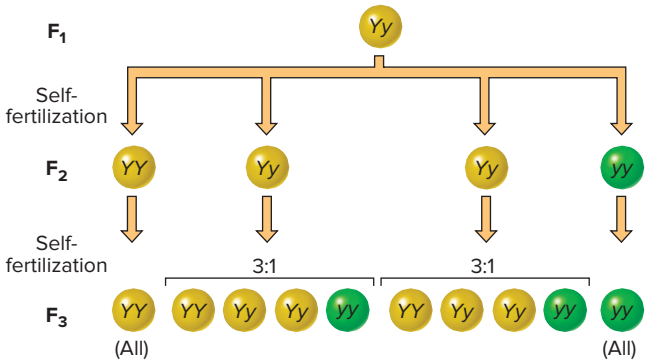
To find the likelihood that an offspring of a *Yy* hybrid self-fertilization will be a hybrid like the parents, you add 1/4 (the probability of maternal *Y* uniting with paternal *y*) and 1/4 (the probability of the mutually exclusive event where paternal *Y* unites with maternal *y*) to get 1/2, again the same result as in the Punnett square.

In another use of the sum rule, you could predict the ratio of yellow to green F_2 progeny. The fraction of F_2 peas that will be yellow is the sum of 1/4 (the event producing *YY*) plus 1/4 (the mutually exclusive event generating *Yy*) plus 1/4 (the mutually exclusive event producing *yY*) to get 3/4. The remaining 1/4 of the F_2 progeny will be green. So the yellow-to-green ratio is 3/4 to 1/4, or more simply, 3:1.

Further Crosses Verify the Law of Segregation

The law of segregation was a hypothesis that explained the data from simple crosses involving monohybrid peas, but Mendel needed to perform additional experiments to check its validity. Mendel's hypothesis, summarized in Fig. 2.11, made the testable prediction that the F_2 should have two kinds of yellow peas (*YY* and *Yy*) but only one kind of green pea (*yy*). In addition, his hypothesis predicted that the *YY* and *Yy* yellow peas in the F_2 should be present in a ratio of $1YY : 2Yy$.

Figure 2.12 Yellow F_2 peas are of two types: Pure breeding and hybrid. The distribution of a pair of contrasting alleles (*Y* and *y*) after two generations of self-fertilization. The homozygous individuals of each generation breed true, whereas the hybrids do not.



To verify these expectations, Mendel allowed self-fertilization of all the plants in the F_2 generation and counted the types of F_3 progeny (**Fig. 2.12**). He found that the plants that developed from F_2 green peas all produced only green peas in the F_3, and when the resulting F_3 plants self-fertilized, the next generation (the F_4) also produced green peas (*not shown*). This is what we (and Mendel) would expect of pure-breeding *yy* lines carrying two copies of the recessive allele. The yellow peas were a different story. When Mendel allowed 518 F_2 plants that developed from yellow peas to self-fertilize, he observed that 166, roughly 1/3 of the total, were pure-breeding yellow through several generations, but the other 352 (2/3 of the total yellow F_2 plants) were hybrids because they gave rise to yellow and green F_3 peas in a ratio of 3:1. Therefore, as Mendel's theory anticipated, the ratio of *YY* to *Yy* among the 518 F_2 yellow pea plants was indeed 1:2.

It took Mendel years to conduct such rigorous experiments on seven pairs of pea traits, but in the end, he was able to conclude that the segregation of dominant and recessive alleles during gamete formation and their random union at fertilization could indeed explain the 3:1 ratios he observed whenever he allowed hybrids to self-fertilize. His results, however, raised yet another question, one of some importance to future plant and animal breeders. Plants showing a dominant trait, such as yellow peas, can be either pure-breeding (*YY*) or hybrid (*Yy*). How can you distinguish one from the other? For self-fertilizing plants, the answer is to observe the appearance of the next generation. But how would you distinguish pure-breeding from hybrid individuals in species that do not self-fertilize?

Figure 2.13 Genotype versus phenotype in homozygotes and heterozygotes. The relationship between genotype and phenotype with a pair of contrasting alleles where one allele (*Y*) shows complete dominance over the other (*y*).

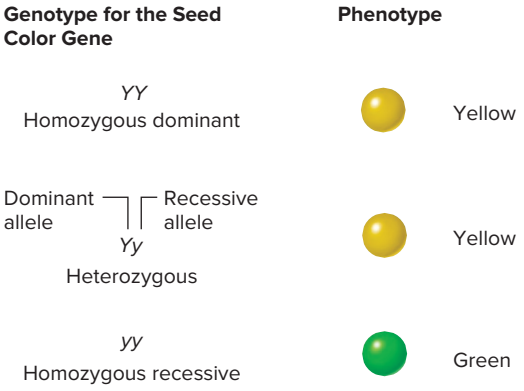
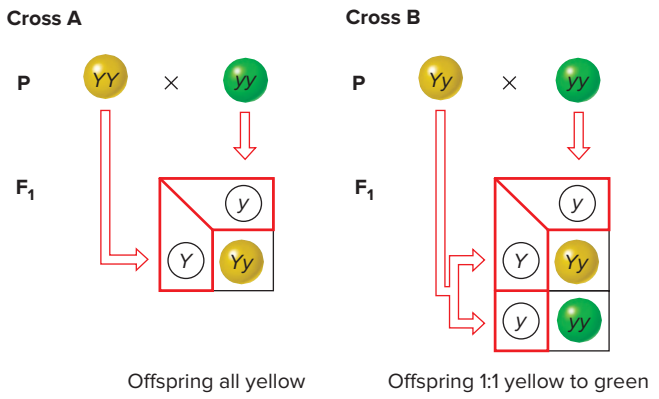


Figure 2.14 How a testcross reveals genotype. An individual of unknown genotype, but dominant phenotype, is crossed with a homozygous recessive. If the unknown genotype is homozygous, all progeny will exhibit the dominant phenotype (*cross A*). If the unknown genotype is heterozygous, half the progeny will exhibit the dominant trait, half the recessive trait (*cross B*).



Testcrosses: A way to establish genotype

Before describing Mendel's answer, we need to define a few more terms. An observable characteristic, such as yellow or green pea seeds, is a **phenotype,** while the actual pair of alleles present in an individual is its **genotype.** A *YY* or a *yy* genotype is called **homozygous,** because the two copies of the gene that determine the particular trait in question are the same. In contrast, a genotype with two different alleles for a trait is **heterozygous;** in other words, it is a hybrid for that trait (**Fig. 2.13**). An individual with a homozygous genotype is a **homozygote;** one with a heterozygous genotype is a **heterozygote.**

Note that the phenotype of a heterozygote (that is, of a hybrid) defines which allele is dominant: Because *Yy* peas are yellow, the yellow allele *Y* is dominant to the *y* allele for green. If you know the genotype and the dominance relation of the alleles, you can accurately predict the phenotype. The reverse is not true, however, because some phenotypes can derive from more than one genotype. For example, the phenotype of yellow peas can result from either the *YY* or the *Yy* genotype.

With these distinctions in mind, we can look at the method Mendel devised for deciphering the unknown genotype. We'll call it *Y*–, responsible for a dominant phenotype; the dash represents the unknown second allele, either *Y* or *y*. This method, called the **testcross,** is a mating in which an individual showing the dominant phenotype, for instance, a *Y*– plant grown from a yellow pea, is crossed with an individual expressing the recessive phenotype, in this case a *yy* plant grown from a green pea. As the Punnett squares in **Fig. 2.14** illustrate, if the dominant phenotype in question derives from a homozygous *YY* genotype, all the offspring of the testcross will show the dominant yellow phenotype. But if the dominant parent of unknown genotype is a heterozygous hybrid (*Yy*), half of the progeny are expected to be yellow peas, and the other half green. In this way, the testcross establishes the genotype behind a dominant phenotype, resolving any uncertainty.

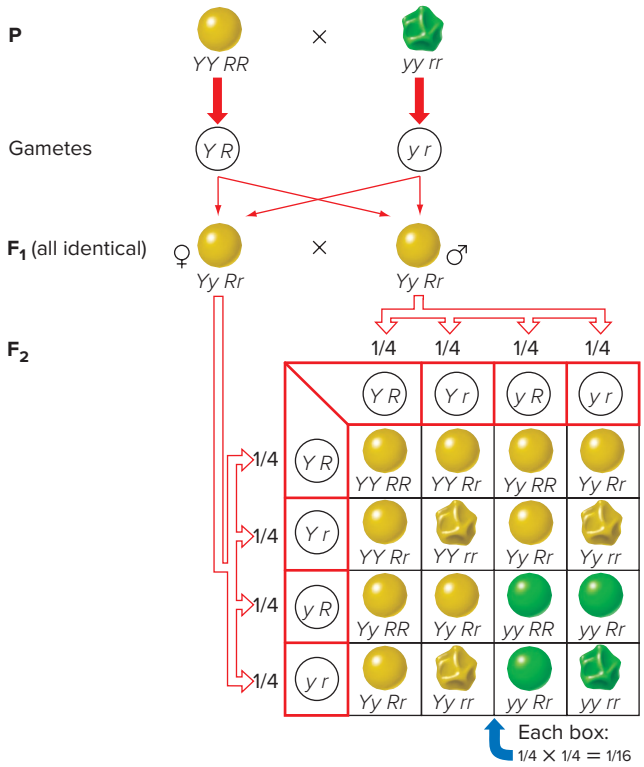
As we mentioned earlier, Mendel deliberately simplified the problem of heredity, focusing on traits that come in only two forms. He was able to replicate his basic monohybrid findings with corn, beans, and four-o'clocks (plants with tubular, white or bright red flowers). As it turns out, his concept of the gene and his law of segregation can be generalized to almost all sexually reproducing organisms.

Dihybrid Crosses Reveal the Law of Independent Assortment

Having determined from monohybrid crosses that genes are inherited according to the law of segregation, Mendel turned his attention to the simultaneous inheritance of two or more apparently unrelated traits in peas. He asked how two pairs of alleles would segregate in a **dihybrid** individual, that is, in a plant that is heterozygous for two genes at the same time.

To construct such a dihybrid, Mendel mated true-breeding plants grown from yellow round peas (*YY RR*) with true-breeding plants grown from green wrinkled peas (*yy rr*). From this cross he obtained a dihybrid F_1 generation (*Yy Rr*) showing the two dominant phenotypes, yellow and round (**Fig. 2.15**). He then allowed these F_1 dihybrids to self-fertilize to produce the F_2 generation. Mendel could not predict the outcome of this mating. Would all the F_2 progeny be **parental types** that looked like either the original yellow round parent or the green wrinkled parent? Or would some new combinations of phenotypes occur that were not seen in the parental lines, such as yellow wrinkled or green round peas? New phenotypic combinations like these are called **recombinant types.**

Figure 2.15 A dihybrid cross produces parental types and recombinant types. In this dihybrid cross, pure-breeding parents (P) produce a genetically uniform generation of F_1 dihybrids. Self-pollination or cross-pollination of the F_1 plants yields the characteristic F_2 phenotypic ratio of 9:3:3:1.



Type	Genotype	Phenotype	Number	Phenotypic Ratio
Parental	*Y– R–*	yellow round	315	9/16
Recombinant	*yy R–*	green round	108	3/16
Recombinant	*Y– rr*	yellow wrinkled	101	3/16
Parental	*yy rr*	green wrinkled	32	1/16

Ratio of yellow (dominant) to green (recessive) = 12:4 or 3:1

Ratio of round (dominant) to wrinkled (recessive) = 12:4 or 3:1

When Mendel counted the F_2 generation of one experiment, he found 315 yellow round peas, 108 green round, 101 yellow wrinkled, and 32 green wrinkled. Both yellow wrinkled and green round recombinant phenotypes did, in fact, appear, providing evidence that some shuffling of the alleles of different genes had taken place.

The law of independent assortment

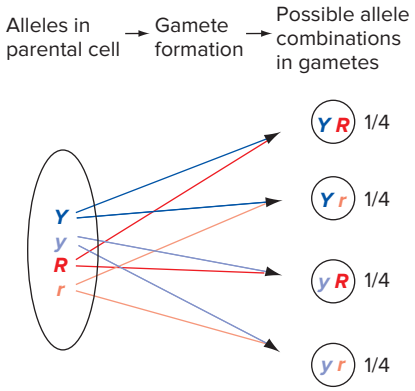
From the observed ratios, Mendel inferred the biological mechanism of that shuffling—the **independent assortment** of gene pairs during gamete formation. Because the genes for pea color and for pea shape assort independently, the allele for pea shape in a gamete carrying *Y* could with equal likelihood be either *R* or *r*. Thus, the presence of a particular allele of one gene, say, the dominant *Y* for pea color, provides no information whatsoever about the allele of the second gene. Each dihybrid of the F_1 generation can therefore make four kinds of gametes: *Y R, Y r, y R*, and *y r*. In a large number of gametes, the four kinds will appear in an almost perfect ratio of 1:1:1:1, or put another way, roughly 1/4 of the eggs and 1/4 of the sperm will contain each of the four possible combinations of alleles. That "the different kinds of germinal cells [eggs or sperm] of a hybrid are produced on the average in equal numbers" was yet another one of Mendel's incisive insights.

At fertilization then, in a mating of dihybrids, 4 different kinds of eggs can each combine with any 1 of 4 different kinds of sperm, producing a total of 16 possible zygotes. Once again, a Punnett square is a convenient way to visualize the process (Fig. 2.15). Using the same kind of logic previously applied to the Punnett square for monohybrid crosses (review Fig. 2.11), each of the 16 boxes with colored peas in the Punnett square for the dihybrid cross in Fig. 2.15 represents an equally likely fertilization event. Again, each box is an equally likely outcome *only because* each of the different gamete types is produced at equal frequency in each parent. Therefore, using the product rule, the frequency of the progeny type in each box is $1/4 \times 1/4 = 1/16$.

If you look at the square in Fig. 2.15, you will see that some of the 16 potential allelic combinations are identical. In fact, only nine different genotypes exist—*YY RR, YY Rr, Yy RR, Yy Rr, yy RR, yy Rr, YY rr, Yy rr*, and *yy rr*—because the source of the alleles (egg or sperm) does not make any difference. If you look at the combinations of traits determined by the nine genotypes, you will see only four phenotypes—yellow round, green round, yellow wrinkled, and green wrinkled—in a ratio of 9:3:3:1. If, however, you look only at pea color or only at pea shape, you can see that each trait is inherited in the 3:1 ratio predicted by Mendel's law of segregation. In the Punnett square, there are 12 yellow for every 4 green and 12 round for every 4 wrinkled. In other words, the ratio of each dominant trait (yellow or round) to its antagonistic recessive trait (green or wrinkled) is 12:4, or 3:1. This means that the inheritance of the gene for pea color is unaffected by the inheritance of the gene for pea shape, and *vice versa.*

The preceding analysis became the basis of Mendel's second general genetic principle, the **law of independent assortment:** *During gamete formation, different pairs of alleles segregate independently of each other* (**Fig. 2.16**). The independence of their segregation and the subsequent random union of gametes at fertilization determine the phenotypes observed. Using the product rule for assessing the probability of independent events, you can see mathematically how the 9:3:3:1 phenotypic ratio observed in a dihybrid cross derives from two separate 3:1 phenotypic ratios. If the two sets of

Figure 2.16 The law of independent assortment. In a dihybrid cross, each pair of alleles assorts independently during gamete formation. In the gametes, *Y* is equally likely to be found with *R* or *r* (that is, *Y R* = *Y r*); the same is true for *y* (that is, *y R* = *y r*). As a result, all four possible types of gametes (*Y R*, *Y r*, *y R*, and *y r*) are produced in equal frequency among a large population.



alleles assort independently, the yellow-to-green ratio in the F_2 generation will be 3/4 : 1/4, and likewise, the round-to-wrinkled ratio will be 3/4 : 1/4. To find the probability that two independent events such as yellow and round will occur simultaneously in the same plant, you multiply as follows:

Probability of yellow round = 3/4 × 3/4 = 9/16

Probability of green round = 1/4 × 3/4 = 3/16

Probability of yellow wrinkled = 3/4 × 1/4 = 3/16

Probability of green wrinkled = 1/4 × 1/4 = 1/16

Thus, in a population of F_2 plants, there will be a 9:3:3:1 phenotypic ratio of yellow round to green round to yellow wrinkled to green wrinkled.

Branched-line diagrams

A convenient way to keep track of the probabilities of each potential outcome in a genetic cross is to construct a **branched-line diagram** (**Fig. 2.17**), which shows all the possible genotypes or phenotypes for each gene in a sequence of columns. In Fig. 2.17, the first column shows the two possible pea color phenotypes; the second column demonstrates that each pea color can occur with either of two pea shapes. Again, the 9:3:3:1 ratio of phenotypes is apparent. You will see later that branched-line diagrams are more convenient than Punnett squares for predicting the outcomes of crosses involving more than two genes.

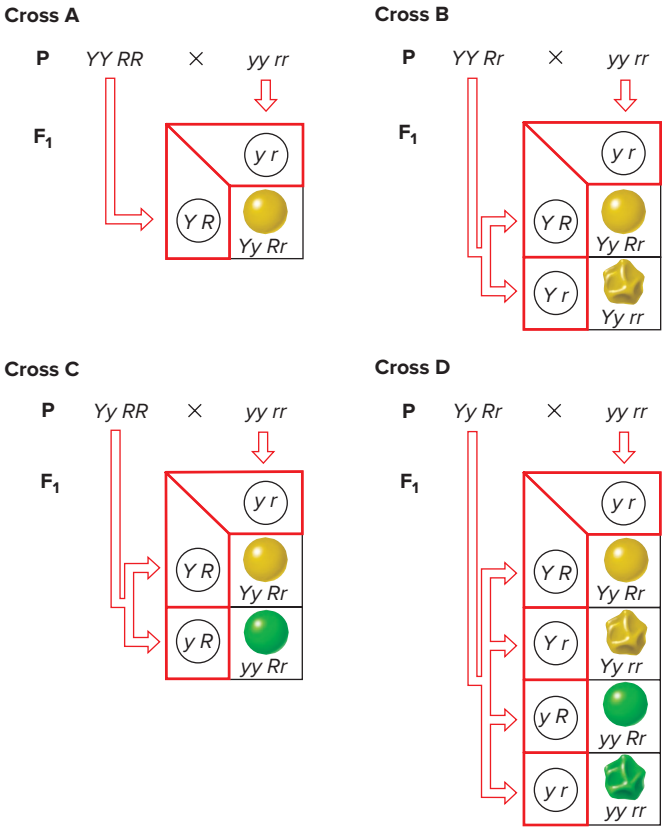
Figure 2.17 Following crosses with branched-line diagrams. A branched-line diagram, which uses a series of columns to track every gene in a cross, provides an organized overview of all possible outcomes. This branched-line diagram of a dihybrid cross generates the same phenotypic ratios as the Punnett square in Fig. 2.15, showing that the two methods are equivalent.

Testcrosses with dihybrids

An understanding of dihybrid crosses has many applications. Suppose, for example, that you work for a nursery that has three pure-breeding strains: yellow wrinkled, green round, and green wrinkled. Your assignment is to grow pure-breeding plants guaranteed to produce yellow round peas. How would you proceed?

One answer is to cross your two pure-breeding strains (*YY rr* × *yy RR*) to generate a dihybrid (*Yy Rr*). Then self-cross the dihybrid and plant only the yellow round peas. Only one out of nine of such progeny—those grown from peas with a *YY RR* genotype—will be appropriate for your uses. To find these plants, you could subject each yellow round candidate to a testcross for genotype with a green wrinkled (*yy rr*) plant, as illustrated in **Fig. 2.18**. If the

Figure 2.18 Testcrosses with dihybrids. Testcrosses involving two pairs of independently assorting alleles yield different, predictable results depending on the tested individual's genotype for the two genes in question.



testcross yields all yellow round offspring (testcross A), you can sell your test plant, because you know it is homozygous for both pea color and pea shape. If your testcross yields 1/2 yellow round and 1/2 yellow wrinkled (testcross B), or 1/2 yellow round and 1/2 green round (testcross C), you know that the candidate plant in question is genetically homozygous for one trait and heterozygous for the other and must therefore be discarded. Finally, if the testcross yields 1/4 yellow round, 1/4 yellow wrinkled, 1/4 green round, and 1/4 green wrinkled (testcross D), you know that the plant is a heterozygote for both the pea color and the pea shape genes.

Geneticists Use Mendel's Laws to Calculate Probabilities and Make Predictions

Mendel performed several sets of dihybrid crosses and also carried out **multihybrid crosses:** matings between the F_1 progeny of pure-breeding parents that differed in three or more unrelated traits. In all of these experiments, he observed numbers and ratios very close to what he expected on the basis of his two general biological principles: The alleles of a gene segregate during the formation of egg or sperm, and the alleles of different genes assort independently of each other. Mendel's laws of inheritance, in conjunction with the mathematical rules of probability, provide geneticists with powerful tools for predicting and interpreting the results of genetic crosses. But as with all tools, they have their limitations. We examine here both the power and the limitations of Mendelian analysis.

First, the power: Using simple Mendelian analysis, it is possible to make accurate predictions about the offspring of extremely complex crosses. Suppose you want to predict the occurrence of one specific genotype in a cross involving several independently assorting genes. For example, if hybrids that are heterozygous for four traits are allowed to self-fertilize—*Aa Bb Cc Dd* × *Aa Bb Cc Dd*—what proportion of their progeny will have the genotype *AA bb Cc Dd?* You could set up a Punnett square to answer the question. Because for each trait there are two different alleles, the number of different eggs or sperm is found by raising 2 to the power of the number of differing traits (2^n, where n is the number of traits). By this calculation, each hybrid parent in this cross with 4 traits would make $2^4 = 16$ different kinds of gametes. The Punnett square depicting such a cross would thus contain 256 boxes (16 × 16).

Setting up such a square may be fine if you live in a monastery with a bit of time on your hands, but not if you're taking a 1-hour exam. It would be much simpler to analyze the problem by breaking down the multihybrid cross into four independently assorting monohybrid crosses. Remember that the genotypic ratios of each monohybrid cross are 1 homozygote for the dominant allele, to 2 heterozygotes, to 1 homozygote for the recessive allele = 1/4 : 2/4 : 1/4. Thus, you can find the probability of *AA bb Cc Dd* by multiplying the probability of each independent event: *AA* (1/4 of the progeny produced by *Aa* × *Aa*); *bb* (1/4); *Cc* (2/4); *Dd* (2/4):

$$1/4 \times 1/4 \times 2/4 \times 2/4 = 4/256 = 1/64.$$

The Punnett square approach would provide the same answer, but it would require much more time.

If instead of a specific genotype, you want to predict the probability of a certain phenotype, you can again use the product rule as long as you know the phenotypic ratios produced by each pair of alleles in the cross. For example, if in the multihybrid cross of *Aa Bb Cc Dd* × *Aa Bb Cc Dd*, you want to know how many offspring will show the dominant A trait (genotype *AA* or *Aa* = 1/4 + 2/4, or 3/4), the recessive b trait (genotype *bb* = 1/4), the dominant C trait (genotype *CC* or *Cc* = 3/4), and the dominant D trait (genotype *DD* or *Dd* = 3/4), you simply multiply:

$$3/4 \times 1/4 \times 3/4 \times 3/4 = 27/256.$$

In this way, the rules of probability make it possible to predict the outcome of very complex crosses.

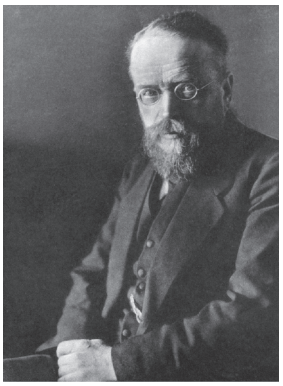
You can see from these examples that particular problems in genetics are amenable to particular modes of analysis. As a rule of thumb, Punnett squares are excellent for visualizing simple crosses involving a few genes, but they become unwieldy in the dissection of more complicated matings. Direct calculations of probabilities, such as those in the two preceding problems, are useful when you want to know the chances of one or a few outcomes of complex crosses. If, however, you want to know all the outcomes of a multihybrid cross, a branched-line diagram is the best way to go as it will keep track of all the possibilities in an organized fashion.

Now, the limitations of Mendelian analysis: Like Mendel, if you were to breed pea plants or corn or any other organism, you would most likely observe some deviation from the ratios you expected in each generation. What can account for such variation? One element is chance, as witnessed in the common coin toss experiment. With each throw, the probability of the coin coming up heads is equal to the likelihood it will come up tails. But if you toss a coin 10 times, you may get 30% (3) heads and 70% (7) tails, or *vice versa.* If you toss it 100 times, you are more likely to get a result closer to the expected 50% heads and 50% tails. The larger the number of trials, the lower the probability that chance significantly skews the data. The statistical benefit is one reason Mendel worked with large numbers of pea plants.

Mendel's laws, in fact, have great predictive power for populations of organisms, but they do not tell us what will happen in any one individual. With a garden full of self-fertilizing monohybrid pea plants, for example, you can expect that 3/4 of the F_2 progeny will show the dominant phenotype and 1/4 the recessive, but you cannot predict the phenotype of any particular F_2 plant. In Chapter 5, we discuss mathematical methods for assessing whether the

Figure 2.19 The science of genetics begins with the rediscovery of Mendel. Working independently near the beginning of the twentieth century, Correns, de Vries, and von Tschermak each came to the same conclusions as those Mendel summarized in his laws. (a, c): © SPL/Science Source; (b): © INTERFOTO/Alamy; (d): © ullstein bild/Getty Images

(a) Gregor Mendel



(b) Carl Correns

(c) Hugo de Vries

(d) Erich von Tschermak

chance variation observed in a sample of individuals within a population is compatible with a genetic hypothesis.

Mendel's Genius Was Unappreciated Before 1900

Mendel's insights into the workings of heredity were a breakthrough of monumental proportions. By counting and analyzing data from hundreds of pea plant crosses, he inferred the existence of genes—independent units that determine the observable patterns of inheritance for particular traits. His work explained the reappearance of hidden traits, disproved the idea of blended inheritance, and showed that mother and father make an equal genetic contribution to the next generation. The model of heredity that he formulated was so specific that he could test predictions based on it by observation and experiment.

With the exception of Abbot Napp, none of Mendel's contemporaries appreciated the importance of his research. Mendel did not teach at a prestigious university and was not well known outside Brünn. Even in Brünn, members of the Natural Science Society were disappointed when he presented *Experiments on Plant Hybrids* to them. They wanted to view and discuss intriguing mutants and lovely flowers, so they did not appreciate his numerical analyses. Mendel, it seems, was far ahead of his time. Sadly, despite written requests from Mendel that others try to replicate his studies, no one repeated his experiments. Several citations of his paper between 1866 and 1900 referred to his expertise as a plant breeder but made no mention of his laws. Moreover, at the time Mendel presented his work, no one had yet seen the structures within cells, the *chromosomes,* that actually carry the genes. That would happen only in the next few decades (as described in Chapter 4). If scientists had been able to see these structures, they might have more readily accepted Mendel's ideas, because the chromosomes are actual physical structures that behave exactly as Mendel predicted.

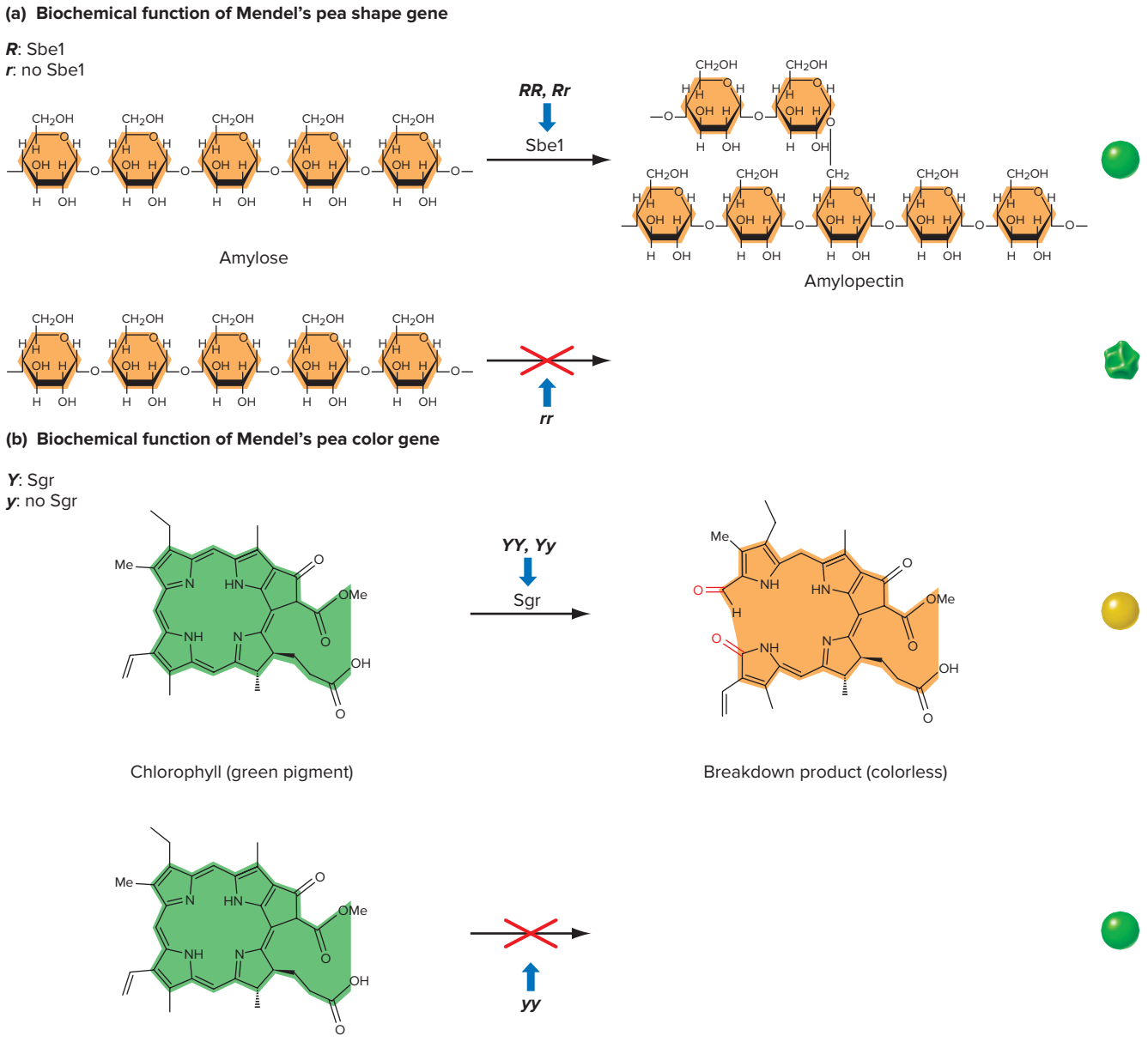
Mendel's work might have had an important influence on early debates about evolution if it had been more widely appreciated. Charles Darwin (1809–1882), who was unfamiliar with Mendel's work, was plagued in his later years by criticism that his explanations for the persistence of variation in organisms were insufficient. Darwin considered such variation a cornerstone of his theory of evolution, maintaining that natural selection would favor particular variants in a given population in a given environment. If the selected combinations of variant traits were passed on to subsequent generations, this transmission of variation would propel evolution. He could not, however, say how that transmission might occur. Had Darwin been aware of Mendel's ideas, he might not have been backed into such an uncomfortable corner.

For 34 years, Mendel's laws lay dormant—untested, unconfirmed, and unapplied. Then in 1900, 16 years after Mendel's death, Carl Correns, Hugo de Vries, and Erich von Tschermak independently rediscovered and acknowledged his work (**Fig. 2.19**). The scientific community had finally caught up with Mendel. Within a decade, investigators had coined many of the modern terms we have been using: phenotype, genotype, homozygote, heterozygote, gene, and genetics, the label given to the twentieth-century science of heredity. Mendel's paper provided the new discipline's foundation. His principles and analytic techniques endure today, guiding geneticists and evolutionary biologists in their studies of genetic variation.

The Influence of Molecules on Phenotype Determines Whether Alleles are Dominant or Recessive

We now know that genes specify the proteins (and RNAs) that cells produce and that dictate cellular structure and function. Recently, two genes were identified that are likely to correspond to Mendel's genes for seed shape and seed

Figure 2.20 Molecular explanations of Mendel's pea shape and pea color genes. The *R* allele of the pea shape gene specifies the enzyme Sbe1, which converts unbranched starch (amylose) to branched starch (amylopectin). The *r* allele does not produce Sbe1. The buildup of unbranched starch in *rr* peas ultimately causes seed wrinkling. The *Y* allele of the pea color gene specifies the enzyme Sgr, which functions in a pathway to break down chlorophyll during pea maturation, resulting in yellow peas. The *y* allele does not produce Sgr. Chlorophyll is not broken down in *yy* peas, and they remain green.



color. The pea shape gene specifies an enzyme known as Sbe1 (for Starch-branching enzyme 1). Sbe1 catalyzes the conversion of amylose, an unbranched linear molecule of starch, to amylopectin, a starch molecule composed of several branching chains (**Fig. 2.20a**). The dominant *R* allele of the pea shape gene determines a normal, functioning Sbe1 enzyme. In contrast, the recessive allele *r* specifies no Sbe1 enzyme. As a result, *RR* homozygotes contain a high proportion of branched starch molecules, which allows the peas to maintain a rounded shape. In homozygous recessive *rr* peas, sucrose builds up because less of it is converted into starch. The excess sucrose modifies the osmotic pressure, causing water to enter the young seeds. As the seeds mature, they lose the water, shrink, and wrinkle. The single dominant allele in *Rr* heterozygotes apparently specifies enough of the normal Sbe1 enzyme to prevent wrinkling.

The pea color gene determines an enzyme called Sgr (for Stay green). Sgr performs one step in the pathway leading to the breakdown of the green pigment chlorophyll, a process that occurs naturally in peas as they mature (**Fig. 2.20b**). The dominant *Y* allele specifies Sgr, and the

recessive *y* allele does not. Homozygous *YY* or heterozygous *Yy* peas are yellow because they each have enough Sgr to break down all the chlorophyll. Homozygous *yy* peas stay green because they lack the Sgr enzyme, and the chlorophyll remains.

Two general principles emerge from these molecular discoveries. First, a specific gene determines a specific protein (in these cases an enzyme). The activity of the protein may affect the phenotype of the pea plant in any number of ways, depending on the biochemical pathway in which it functions. Second, a pattern can be seen in both of these examples: The dominant allele determines a normally functioning protein, while the recessive allele does not specify a functional protein. You will see in Chapter 3 that, although it is certainly not always the case, the molecular explanation described here is the most common reason why one allele is dominant to another (recessive) allele. Genes likely to be those Mendel described for stem length and flower color have also been identified recently. In both cases, the dominant allele encodes a normally functioning protein, and the recessive allele specifies either no protein or a less functional version of the normal protein.

essential concepts

- Discrete units called *genes* control the appearance of inherited traits; genes come in alternative forms called *alleles*.
- A sexually reproducing organism's body cells contain two *alleles* for every gene. These alleles may be the same (in a *homozygote*) or different (in a *heterozygote*).
- *Genotype* refers to the alleles an individual possesses; *phenotype* refers to the traits the individual exhibits.
- The *dominant* allele controls the phenotype of a trait in heterozygotes; the other allele in the heterozygote is *recessive*. In monohybrid crosses, the dominant and recessive phenotypes will appear in the progeny in a ratio of 3:1.
- Alleles segregate during the formation of gametes, which thus contain only one allele of each gene. Male and female gametes unite at random at fertilization. These two processes correspond to Mendel's *law of segregation*.
- The segregation of alleles of any one gene is independent of the segregation of the alleles of other genes. This principle is Mendel's *law of independent assortment*. According to this law, crosses between *Aa Bb* dihybrids will generate progeny with a phenotypic ratio of 9 (*A– B–*) : 3 (*A– bb*) : 3 (*aa B–*) : 1 (*aa bb*).
- Most often, the dominant allele of a gene specifies a functional product (a protein), while the recessive allele determines either a less functional or nonfunctional version of the protein, or no protein at all.

2.3 Mendelian Inheritance in Humans

learning objectives

1. Analyze human pedigrees to determine whether a genetic disease exhibits recessive or dominant inheritance.
2. Explain why Huntington disease is inherited as a dominant allele while cystic fibrosis is caused by a recessive allele.

Although many human traits clearly run in families, most do not show a simple Mendelian pattern of inheritance. Suppose, for example, that you have brown eyes, but both your parents' eyes appear to be blue. Because blue is normally considered recessive to brown, does this mean that you are adopted or that your father isn't really your father? Not necessarily, because eye color is influenced by more than one gene.

Like eye color, most common and obvious human phenotypes arise from the interaction of many genes. In contrast, single-gene traits in people usually involve an abnormality that is disabling or life-threatening. Examples are the progressive neurological damage of Huntington disease and the clogged lungs and potential respiratory failure of cystic fibrosis. A defective allele of a single gene gives rise to Huntington disease; defective alleles of a different gene are responsible for cystic fibrosis. **Table 2.1** lists some of the roughly 6000 such single-gene, or Mendelian, traits known in humans as of 2016. As you will see, the allele that causes Huntington disease is dominant and the normal (nondisease) allele of this gene is recessive. The opposite is true for cystic fibrosis—the disease-causing allele is recessive and the normal (nondisease) allele is dominant.

Pedigrees Aid the Study of Hereditary Traits in Human Families

Determining a genetic defect's pattern of transmission is not always an easy task because people make slippery genetic subjects. Their generation time is long, and the families they produce are relatively small, which makes statistical analysis difficult. Humans do not base their choice of mates on purely genetic considerations. Thus, no pure-breeding lines exist and no controlled matings are possible. Furthermore, people rarely produce a true F_2 generation (like the one in which Mendel observed the 3:1 ratios from which he derived his rules) because brothers and sisters almost never mate.

Geneticists circumvent these difficulties by working with a large number of families or with several generations

TABLE 2.1 Some of the Most Common Single-Gene Traits in Humans

Disease	Effect	Incidence of Disease
Caused by a Recessive Allele		
Thalassemia (chromosome 16 or 11)	Reduced amounts of hemoglobin; anemia, bone and spleen enlargement	1/10 in parts of Italy
Sickle-cell anemia (chromosome 11)	Abnormal hemoglobin; sickle-shaped red cells, anemia, blocked circulation; increased resistance to malaria	1/625 African-Americans
Cystic fibrosis (chromosome 7)	Defective cell membrane protein; excessive mucus production; digestive and respiratory failure	1/2000 Caucasians
Tay-Sachs disease (chromosome 15)	Missing enzyme; buildup of fatty deposit in brain that disrupts mental development	1/3000 Eastern European Jews
Phenylketonuria (PKU) (chromosome 12)	Missing enzyme; mental deficiency	1/10,000 Caucasians
Caused by a Dominant Allele		
Hypercholesterolemia (chromosome 19)	Missing protein that removes cholesterol from the blood; heart attack by age 50	1/122 French Canadians
Huntington disease (chromosome 4)	Abnormal Huntingtin protein; progressive mental and neurological damage; neurologic disorders by ages 40–70	1/25,000 Caucasians

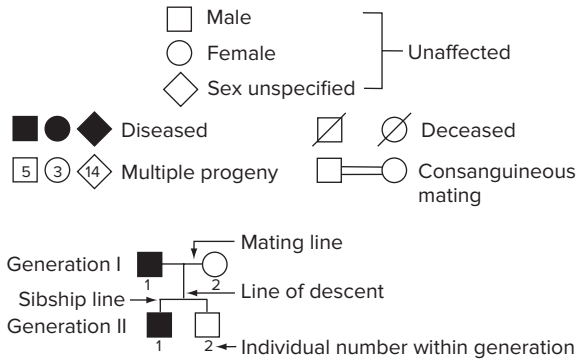
of a very large family. In this way, scientists can study the large numbers of genetically related individuals needed to establish the inheritance patterns of specific traits. A family history, known as a **pedigree,** is an orderly diagram of a family's relevant genetic features, extending back to at least both sets of grandparents and preferably through as many additional generations as possible. From systematic pedigree analysis in the light of Mendel's laws, geneticists can tell if a trait is determined by alternative alleles of a single gene and whether a single-gene trait is dominant or recessive. Because Mendel's principles are so simple and straightforward, a little logic can go a long way in explaining how traits are inherited in humans.

Figure 2.21 shows how to interpret a family pedigree diagram. Squares (□) represent males, circles (○) are females, diamonds (◊) indicate that the sex is unspecified. Family members affected by the trait in question are indicated by a filled-in symbol (for example, ■). A single horizontal line connecting a male and a female (□—○) represents a mating; a double connecting line (□═○) designates a **consanguineous mating,** that is, a mating between relatives; and a horizontal line above a series of symbols (○ □ ○) indicates the children of the same parents (a *sibship*) arranged and numbered from left to right in order of their birth. Roman numerals to the left or right of the diagram indicate the generations.

To reach a conclusion about the mode of inheritance of a family trait, human geneticists must use a pedigree that supplies sufficient information. For example, researchers could not determine whether the allele causing the disease depicted at the bottom of Fig. 2.21 is dominant or recessive solely on the basis of the simple pedigree shown. The data are consistent with both possibilities. If the trait is dominant, then the father and the affected son are heterozygotes, while the mother and the unaffected son are homozygotes for the recessive normal allele. If instead the trait is recessive, the father and affected son are homozygotes for the recessive disease-causing allele, while the mother and the unaffected son are heterozygotes.

Several kinds of additional information could help resolve this uncertainty. Human geneticists would particularly want to know the frequency at which the trait in question is found in the population from which the family came. *If the trait is rare in the population, then the allele giving rise to the trait should also be rare, and the most likely hypothesis*

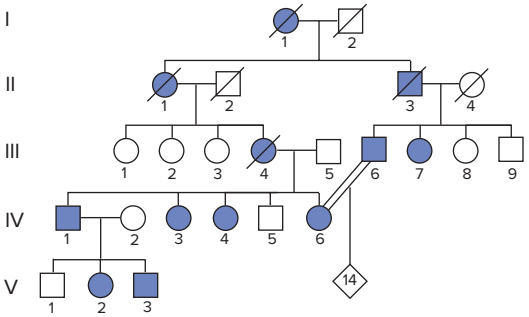
Figure 2.21 Symbols used in pedigree analysis. In the simple pedigree at the bottom, I-1 is the father, I-2 is the mother, and II-1 and II-2 are their sons. The father and the first son are both affected by the disease trait.



would require that the fewest genetically unrelated people carry the allele. Only the father in Fig. 2.21 would need to have a dominant disease-causing allele, but both parents would need to carry a recessive disease-causing allele (the father two copies and the mother one). However, even the information that the trait is rare does not allow us to draw the firm conclusion that it is inherited in a dominant fashion. The pedigree in the figure is so limited that we cannot be sure the two parents are themselves unrelated. As we discuss later in more detail, related parents might have both received the same rare recessive allele from their common ancestor. This example illustrates why human geneticists try to collect family histories that cover several generations.

We now look at more extensive pedigrees for the dominant trait of Huntington disease and for the recessive condition of cystic fibrosis. The patterns by which these traits appear in the pedigrees provide important clues that can indicate modes of inheritance and allow geneticists to assign genotypes to family members.

Figure 2.22 **Huntington disease: A rare dominant trait.** All individuals represented by filled-in symbols are heterozygotes (except I-1, who could instead have been homozygous for the dominant *HD* disease allele); all individuals represented by open symbols are homozygotes for the recessive HD^+ normal allele. Among the 14 children of the consanguineous mating, DNA testing shows that some are *HD HD,* some are *HD* HD^+, and some are HD^+ HD^+. The diamond designation masks personal details to protect confidentiality.



A Vertical Pattern of Inheritance Indicates a Rare Dominant Trait

Huntington disease is named for George Huntington, the New York physician who first described its course. This illness usually shows up in middle age and slowly destroys its victims both mentally and physically. Symptoms include intellectual deterioration, severe depression, and jerky, irregular movements, all caused by the progressive death of nerve cells. If one parent develops the symptoms, his or her children usually have a 50% probability of suffering from the disease, provided they live to adulthood. Because symptoms are not present at birth and manifest themselves only later in life, Huntington disease is known as a **late-onset genetic trait.**

How would you proceed in assigning genotypes to the individuals in the Huntington disease pedigree depicted in **Fig. 2.22**? First, you would need to find out if the disease-producing allele is dominant or recessive. Several clues suggest that Huntington disease is transmitted by a dominant allele of a single gene. Everyone who develops the disease has at least one parent who shows the trait, and in several generations, approximately half of the offspring are affected. The pattern of affected individuals is thus *vertical:* If you trace back through the ancestors of any affected individual, you would see at least one affected person in each generation, giving a continuous line of family members with the disease. When a disease is rare in the population as a whole, a vertical pattern is strong evidence that a dominant allele causes the trait; the alternative would require that many unrelated people carry a rare recessive allele. (A recessive trait that is extremely common might also show up in every generation; we examine this possibility in Problem 40 at the end of this chapter.)

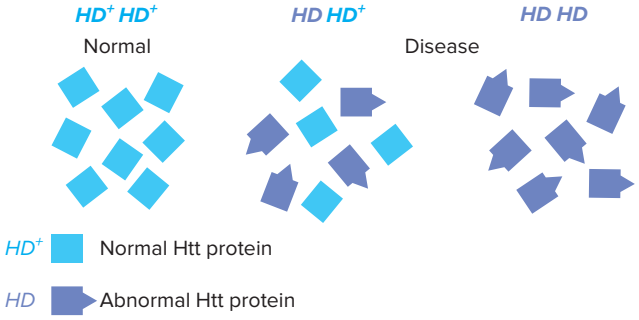
In tracking a dominant allele through a pedigree, you can view every mating between an affected and an unaffected partner as analogous to a testcross. If some of the offspring do not have Huntington disease, you know the parent showing the trait is a heterozygote. As an exercise, you should check your own genotype assignments against the answers in the caption to Fig. 2.22.

Notice also in the legend to Fig. 2.22 that human geneticists use different symbols than Mendel's for alleles of genes. In human genotypes, all alleles are written in uppercase. If the allele specifies a normally functioning gene product, the allele symbol has a superscript +. Alleles that specify no gene product or abnormal gene products sometimes have no superscript at all, as in the Fig. 2.22 legend, but in other cases they have a superscript other than + that signifies a particular type of abnormal allele. (See the Appendix *Guidelines for Gene Nomenclature* for further discussion of genetic notation.)

Like Mendel's pea genes, the gene that causes Huntington disease has been identified and studied at the molecular level. In fact, in 1988 this was the first human disease gene identified molecularly using methods that will be described in Chapter 11. The protein product of the Huntington disease gene, called Huntingtin or Htt, is needed for the proper physiology of nerve cells, but the protein's precise role in these cells is not yet understood. The dominant disease allele (*HD*) specifies a defective Htt protein that over time damages nerve cells (**Fig. 2.23**).

The disease allele is dominant to the normal allele because the presence of the normal Htt protein in heterozygotes does not prevent the abnormal protein from damaging the cells. It is important to note that this explanation for the Huntington disease allele is only one of many different

Figure 2.23 Why the allele for Huntington disease is dominant. People who are *HD HD* or *HD HD*$^+$ exhibit Huntington disease because the *HD* allele produces an abnormal Htt protein that damages nerve cells. Homozygotes for the normal allele (*HD*$^+$ *HD*$^+$) produce only normal Htt protein and do not have the disease. The disease allele (*HD*) is dominant because even when the normal protein is present—in *HD HD*$^+$ heterozygotes—the abnormal protein damages nerve cells. *HD HD* homozygosity is possible because the abnormal Htt protein retains some function of the normal protein.



molecular mechanisms that may result in a disease allele that is dominant to the normal allele of a particular gene.

No effective treatment yet exists for Huntington disease, and because of its late onset, there was until the 1980s no way for children of a Huntington parent to know before middle age—usually until well after their own childbearing years—whether they carried the Huntington disease allele (*HD*). Most people with the disease allele are *HD HD*$^+$ heterozygotes, so their children would have a 50% probability of inheriting *HD* and, before they are diagnosed, a 25% probability of passing the defective allele on to one of their children.

In the mid-1980s, with new knowledge of the gene, molecular geneticists developed a DNA test that determines whether an individual carries the *HD* allele. (This test will be explained in detail in Chapter 11.) Because of the lack of effective treatment for the disease, some young adults whose parents died of Huntington disease prefer not to be tested so that they will not learn their own fate prematurely. However, other at-risk individuals employ the test for the *HD* allele to guide their decisions about having children. If someone whose parent had Huntington disease does not have *HD*, he or she has no chance of developing the disease or of transmitting it to offspring. If the test shows the presence of *HD*, the at-risk person and his or her partner might choose to conceive a child using *in vitro* fertilization (IVF) technology (described in Chapter 11) that allows for genotyping of early-stage embryos. Using IVF, only embryos lacking the *HD* disease allele would be introduced into the mother's womb.

The Genetics and Society Box *Developing Guidelines for Genetic Screening* discusses significant social and ethical issues raised by information obtained from family pedigrees and molecular tests.

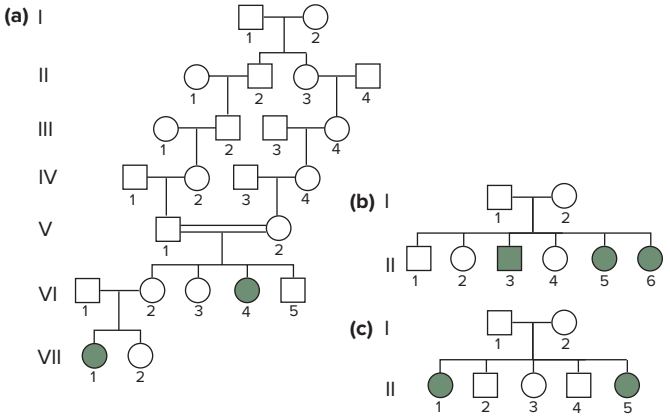
A Horizontal Pattern of Inheritance Indicates a Rare Recessive Trait

Unlike Huntington disease, most confirmed single-gene diseases in humans are caused by recessive alleles. One reason is that, with the exception of late-onset traits, deleterious dominant alleles are unlikely to be transmitted to the next generation. For example, if people affected with Huntington disease all died by the age of 10, the disease would disappear from the population. In contrast, individuals can carry one allele for a recessive disease without ever being affected by any symptoms.

Figure 2.24 shows three pedigrees for cystic fibrosis, the most commonly inherited recessive disease among Caucasian children in the United States. A double dose of the recessive *CF* allele (meaning the absence of a *CF*$^+$ allele) causes a fatal disorder in which the lungs, pancreas, and other organs become clogged with a thick, viscous mucus that can interfere with breathing and digestion. One in every 2000 white Americans is born with cystic fibrosis, and only 10% of them survive into their 30s.

Note two salient features of the cystic fibrosis pedigrees. First, the family pattern of people showing the trait is often *horizontal:* The parents, grandparents, and great-grandparents of children born with cystic fibrosis do not themselves manifest the disease, while several brothers and sisters in a single generation may. A horizontal pedigree pattern is a strong indication that the trait is recessive. The unaffected parents are heterozygous **carriers:** They bear a dominant

Figure 2.24 Cystic fibrosis: A recessive condition. In **(a)**, the two affected individuals (VI-4 and VII-1) are *CF CF*; that is, homozygotes for the recessive disease allele. Their unaffected parents must be carriers, so V-1, V-2, VI-1, and VI-2 must all be *CF CF*$^+$. Individuals II-2, II-3, III-2, III-4, IV-2, and IV-4 are probably also carriers. We cannot determine which of the founders (I-1 or I-2) was a carrier, so we designate their genotypes as *CF*$^+$–. Because the *CF* allele is relatively rare, it is likely that II-1, II-4, III-1, III-3, IV-1, and IV-3 are *CF*$^+$*CF*$^+$ homozygotes. The genotype of the remaining unaffected people (VI-3, VI-5, and VII-2) is uncertain (*CF*$^+$–). (**b** and **c**) These two families demonstrate horizontal patterns of inheritance. Without further information, the unaffected children in each pedigree must be regarded as having a *CF*$^+$– genotype.



GENETICS AND SOCIETY

Crowd: © Image Source/Getty Images RF

Developing Guidelines for Genetic Screening

In the early 1970s, the United States launched a national screening program for carriers of sickle-cell anemia, a recessive genetic disease that afflicts roughly 1 in 600 African-Americans. The disease is caused by a particular allele, called $Hb\beta^S$, of the β-globin gene; the dominant normal allele is $Hb\beta^A$. The protein determined by the β-globin gene is one component of the oxygen-carrying hemoglobin molecule. $Hb\beta^S$ $Hb\beta^S$ homozygotes have sickle-shaped red blood cells; these patients suffer a decrease in oxygen supply, tire easily, and often develop heart failure from stress on the circulatory system.

The national screening program for sickle-cell anemia was based on a simple test of hemoglobin mobility: Normal and *sickling* hemoglobins move at different rates in a gel. People who participated in the screening program and found they were carriers could use the test results to make informed reproductive decisions. A healthy man, for example, who learned he was a carrier (that is, that he was a $Hb\beta^S$ $Hb\beta^A$ heterozygote) would not have to worry about having an affected child if his mate was a noncarrier.

The original sickle-cell screening program, based on detection of the abnormal hemoglobin protein, was not an unqualified success, largely because of insufficient educational follow-through. Many who learned they were carriers mistakenly thought they had the disease. Moreover, employers and insurance companies that obtained access to the information denied jobs or health insurance to some heterozygotes for no acceptable reason. Problems of public relations and education thus made a reliable screening test into a source of dissent and alienation.

Today, at-risk families may be screened for a growing number of genetic disorders, thanks to the ability to evaluate genotypes directly. The need to establish guidelines for genetic screening thus becomes more and more pressing. Several related questions reveal the complexity of the issue.

- *Why carry out genetic screening at all?* The first reason for screening is to obtain information that will benefit individuals. For example, if you learn at an early age that you have a genetic predisposition to heart disease, you can change your lifestyle to improve your chances of staying healthy. You can also use the results from genetic screening to make informed reproductive decisions.
- The second reason for genetic screening, which often conflicts with the first, is to benefit groups within society. Insurance companies and employers, for example, would like to know who is at risk for various genetic conditions.
- *Should I be screened if a test is available*? For most inherited diseases, no cures presently exist. The psychological burden of anticipating a fatal late-onset disease for which there is no treatment could be devastating, and therefore some people might decide not to be tested. Others may object to testing for religious reasons, or because of confidentiality concerns.
- *If a screening program is established, who should be tested?* The answer depends on what the test is trying to accomplish as well as on its expense. Ultimately, the cost of a procedure must be weighed against the usefulness of the data it provides. In the United States, for example, only one-tenth as many African-Americans as Caucasians are affected by cystic fibrosis, and Asians almost never exhibit the disease. Should all racial groups be tested for cystic fibrosis, or only Caucasians?
- *Should private employers and insurance companies be allowed to test their clients and employees?* Some employers advocate genetic screening to reduce the incidence of occupational disease, arguing that they can use genetic test results to make sure employees are not assigned to environments that might cause them harm. Critics of this position say that screening violates workers' rights, including the right to privacy, increases racial and ethnic discrimination in the workplace, and provides insurers with an excuse to deny coverage. In 2008, President George W. Bush signed into law the Genetic Information Nondiscrimination Act, which prohibits insurance companies and employers in the United States from discriminating on the basis of information derived from genetic tests.
- *Finally, how should people be educated about the meaning of test results?* In one small-community screening program, people identified as carriers of the recessive, life-threatening blood disorder known as β-thalassemia were ostracized; as a result, carriers ended up marrying one another, only making medical matters worse. By contrast, in Ferrara, Italy, where 30 new cases of β-thalassemia had been reported every year, extensive screening combined with education was so successful that the 1980s passed with only a few new cases of the disease.

Given all of these considerations, what kind of guidelines would you like to see established to ensure that genetic screening reaches the right people at the right time, and that information gained from such screening is used for the right purposes?

normal allele (CF^+) that masks the effects of the recessive abnormal one. An estimated 12 million Americans are carriers of a recessive *CF* allele. **Table 2.2** summarizes some of the clues found in pedigrees that can help you decide whether a trait is caused by a dominant or a recessive allele.

The second salient feature of the cystic fibrosis pedigrees is that many of the couples who produce afflicted children are blood relatives; that is, their mating is consanguineous (as indicated by the double line). In Fig. 2.24a, the consanguineous mating in generation V is between third cousins. Of course, children with cystic fibrosis can also have unrelated carrier parents, but because relatives share genes, their offspring have a much greater than average chance of receiving two copies of a rare allele. Whether or

TABLE 2.2 How to Recognize Dominant and Recessive Traits in Pedigrees

Dominant Traits
1. Affected children always have at least one affected parent.
2. Dominant traits show a *vertical pattern* of inheritance: The trait shows up in every generation.
3. Two affected parents can produce unaffected children, if both parents are heterozygotes.
Recessive Traits
1. Affected individuals can be the children of two unaffected carriers, particularly as a result of consanguineous matings.
2. All the children of two affected parents should be affected.
3. *Rare* recessive traits show a *horizontal pattern* of inheritance: The trait first appears among several members of one generation and is not seen in earlier generations.
4. Recessive traits may show a vertical pattern of inheritance if the trait is extremely common in the population.

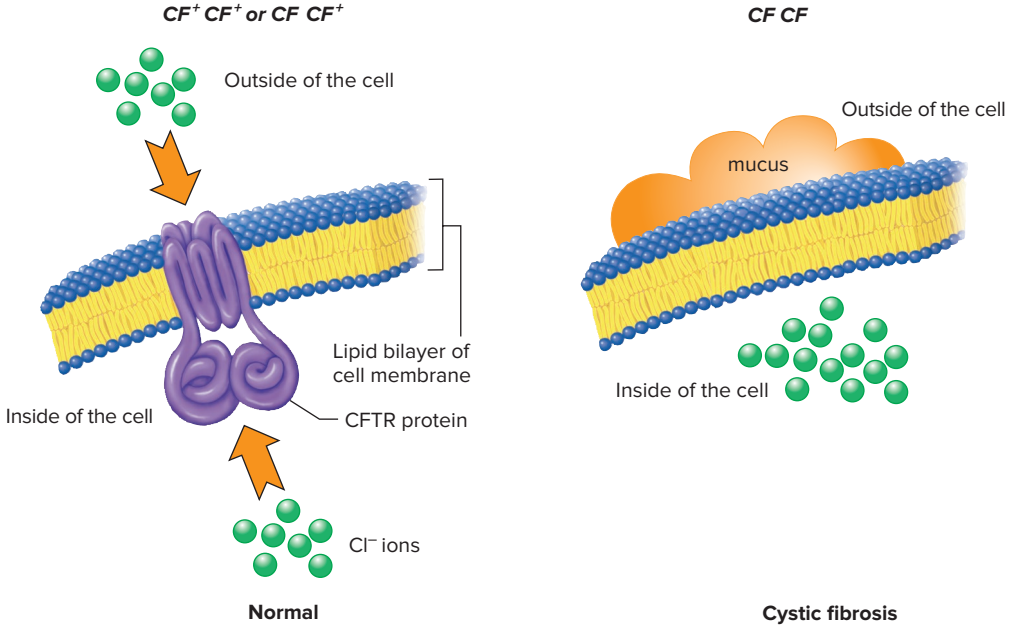
not they are related, carrier parents are both heterozygotes. Thus among their offspring, the proportion of unaffected to affected children is expected to be 3:1. To look at it another way, the chances are that 1 out of 4 children of two heterozygous carriers will be homozygous cystic fibrosis sufferers.

You can gauge your understanding of this inheritance pattern by assigning a genotype to each person in Fig. 2.24 and then checking your answers against the caption. Note that for several individuals, such as the generation I individuals in part (a) of the figure, it is impossible to assign a full genotype. We know that one of these people must be the carrier who supplied the original *CF* allele, but we do not know if it was the male or the female. As with an ambiguous dominant phenotype in peas, the unknown second allele is indicated by a dash (–).

In Fig. 2.24a, a mating between the unrelated carriers VI-1 and VI-2 produced a child with cystic fibrosis. How likely is such a marriage between unrelated carriers for a recessive genetic condition? The answer depends on the gene in question and the particular population into which a person is born. As Table 2.1 shows, the incidence of genetic diseases (and thus the frequency of their carriers) varies markedly among populations. Such variation reflects the distinct genetic histories of different groups. The area of genetics that analyzes differences among groups of individuals is called *population genetics*, a subject we cover in detail in Chapter 21. Notice that in Fig. 2.24a, several unrelated, unaffected people, such as II-1 and II-4, married into the family under consideration. Although it is highly probable that these individuals are homozygotes for the normal allele of the gene (CF^+CF^+), a small chance (whose magnitude depends on the population) exists that any one of them could be a carrier of the disease.

Genetic researchers identified the cystic fibrosis gene in 1989, soon after the Huntington disease gene was identified. The normal, dominant CF^+ allele makes a protein called cystic fibrosis transmembrane conductance regulator (CFTR). CFTR protein forms a channel in the cell membranes that controls the flow of chloride ions through lung cells. Recessive *CF* disease alleles either produce no CFTR or produce nonfunctional or less functional versions of the protein (**Fig. 2.25**). Because of osmosis, water flows into lung cells without CFTR, while a thick, dehydrated

Figure 2.25 Why the allele for cystic fibrosis is recessive. The CFTR protein regulates the passage of chloride ions (*green spheres*) through the cell membrane. People who are homozygous for a cystic fibrosis disease allele (*CF CF*) have the disease because recessive disease alleles either specify no CFTR protein as shown, or encode abnormal CFTR proteins that do not function at all or do not function as well as the normal protein (*not shown*). Disease alleles (*CF*) are recessive because *CF* CF^+ heterozygotes produce CFTR from the normal (CF^+) allele, and this amount of CFTR is sufficient for normal lung function.



mucus builds up outside the cells. Thus, *CF CF* homozygotes have no functional CFTR (or not enough of this protein) and exhibit cystic fibrosis. Gene therapy—insertion of a normal CF^+ gene into lung cells of patients—has been tried to ameliorate the disease's debilitating symptoms, but so far without success.

Despite the failure to date of gene therapy, identification of the gene responsible for cystic fibrosis has very recently led to effective treatments for the disease in patients with particular mutant alleles. For example, in 2015 the U.S. Food and Drug Administration approved a drug cocktail called Orkambi® that helps the particular defective form of CFTR specified by one of these alleles to function properly. Varied approaches to the treatment of cystic fibrosis and other inherited diseases will be discussed later in the book.

essential concepts

- In a *vertical* pattern of transmission, a trait that appears in an affected individual also appears in at least one parent, one of the affected parent's parents, and so on. If a trait is rare, a pedigree with a vertical pattern usually indicates that the disease-causing allele is dominant.
- In a *horizontal* pattern of transmission, a trait that appears in an affected individual may not appear in any ancestors, but it may appear in some of the person's siblings. A pedigree with a horizontal pattern usually indicates a rare recessive disease-causing allele. Affected individuals are often products of consanguineous mating.
- Various kinds of biochemical events may explain why some disease alleles are dominant. In the case of Huntington disease, the disease-causing *HD* allele specifies an abnormal, deleterious version of the protein produced by the normal, recessive allele.
- Recessive disease alleles, like the *CF* alleles that cause cystic fibrosis, usually specify either no protein or less-functional versions of the protein that the normal, dominant allele produces.

WHAT'S NEXT

Mendel answered the three basic questions about heredity as follows: To *What is inherited?* he replied, "alleles of genes." To *How is it inherited?* he responded, "according to the principles of segregation and independent assortment." And to *What is the role of chance in heredity?* he said, "for each individual, inheritance is determined by chance, but within a population, this chance operates in a context of strictly defined probabilities."

Within a decade of the 1900 rediscovery of Mendel's work, numerous breeding studies had shown that Mendel's laws hold true not only for seven pairs of antagonistic characteristics in peas, but also for many diverse traits in a wide variety of sexually reproducing plant and animal species. Some of these same breeding studies, however, raised a challenge to the new genetics. For certain traits in certain species, the studies uncovered unanticipated phenotypic ratios, or the results included F_1 and F_2 progeny with novel phenotypes that resembled those of neither pure-breeding parent.

These phenomena could not be explained by Mendel's hypothesis that for each gene, two alternative alleles, one completely dominant, the other recessive, determine a single trait. We now know that most common traits, including skin color, eye color, and height in humans, are determined by interactions between two or more genes. We also know that within a given population, more than two alleles may be present for some of those genes. Chapter 3 shows how the genetic analysis of such complex traits, that is, traits produced by complex interactions between genes and between genes and the environment, extended rather than contradicted Mendel's laws of inheritance.

SOLVED PROBLEMS

Solving Genetics Problems

The best way to evaluate and increase your understanding of the material in the chapter is to apply your knowledge in solving genetics problems. Genetics word problems are like puzzles. Take them in slowly—don't be overwhelmed by the whole problem. Identify useful facts given in the problem, and use the facts to deduce additional information. Use genetic principles and logic to work toward the

solutions. The more problems you do, the easier they become. In solving problems, you will not only solidify your understanding of genetic concepts, but you will also develop basic analytical skills that are applicable in many disciplines.

Note that some of the problems at the end of each chapter are designed to introduce supplementary but important concepts that expand on the information in the text. You can nonetheless answer such problems using logical inferences from your reading.

Solving genetics problems requires much more than simply plugging numbers into formulas. Each problem is unique and requires thoughtful evaluation of the information given and the question being asked. The following are general guidelines you can follow in approaching these word problems:

a. Read through the problem once to get some sense of the concepts involved.
b. Go back through the problem, noting all the information supplied to you. For example, genotypes or phenotypes of offspring or parents may be given to you or implied in the problem. Represent the known information in a symbolic format—assign symbols for alleles; use these symbols to indicate genotypes; make a diagram of the crosses including genotypes and phenotypes given or implied. Be sure that you do not assign different letters of the alphabet to two alleles of the same gene, as this can cause confusion. Also, be careful to discriminate clearly between the upper- and lowercases of letters, such as *C*(*c*) or *S*(*s*).
c. Now, reassess the question and work toward the solution using the information given. Make sure you answer the question being asked!
d. When you finish the problem, check to see that the answer makes sense. You can often check solutions by working backwards; that is, see if you can reconstruct the data from your answer.
e. After you have completed a question and checked your answer, spend a minute to think about which major concepts were involved in the solution. This is a critical step for improving your understanding of genetics.

For each chapter, the logic involved in solving two or three types of problems is described in detail.

I. In cats, white patches are caused by the dominant allele *P,* while *pp* individuals are solid-colored. Short hair is caused by a dominant allele *S,* while *ss* cats have long hair. A long-haired cat with patches whose mother was solid-colored and short-haired mates with a short-haired, solid-colored cat whose mother was long-haired and solid-colored. What kinds of kittens can arise from this mating, and in what proportions?

Answer

The solution to this problem requires an understanding of dominance/recessiveness, gamete formation, and the independent assortment of alleles of two genes in a cross.

First make a representation of the known information:

Mothers:	solid, short-haired		solid, long-haired
Cross:	cat 1		cat 2
	patches, long-haired	×	solid, short-haired

What genotypes can you assign? Any cat showing a recessive phenotype must be homozygous for the recessive allele. Therefore the long-haired cats are *ss;* solid cats are *pp.* Cat 1 is long-haired, so it must be homozygous for the recessive allele (*ss*). This cat has the dominant phenotype of patches and could be either *PP* or *Pp,* but because the mother was *pp* and could only contribute a *p* allele in her gametes, cat 1 must be *Pp.* Cat 1's full genotype is *Pp ss.* Similarly, cat 2 is solid-colored, so it must be homozygous for the recessive allele (*pp*). Because this cat is short-haired, it could have either the *SS* or *Ss* genotype. Its mother was long-haired (*ss*) and could only contribute an *s* allele in her gamete, so cat 2 must be heterozygous *Ss.* The full genotype is *pp Ss.*

The cross is therefore between *Pp ss* (cat 1) and *pp Ss* (cat 2). To determine the types of kittens, first establish the types of gametes that can be produced by each cat and then set up a Punnett square to determine the genotypes of the offspring. Cat 1 (*Pp ss*) produces *Ps* and *ps* gametes in equal proportions. Cat 2 (*pp Ss*) produces *pS* and *ps* gametes in equal proportions. Four types of kittens can result from this mating with equal probability: *Pp Ss* (patches, short-haired), *Pp ss* (patches, long-haired), *pp Ss* (solid, short-haired), and *pp ss* (solid, long-haired).

		Cat 1	
		P s	*p s*
Cat 2	*p S*	*Pp Ss*	*pp Ss*
	p s	*Pp ss*	*pp ss*

The following table demonstrates that you could also work through this problem using the product rule of probability instead of a Punnett square. The principles are the same: Gametes produced in equal amounts by either parent are combined at random.

Cat 1 gamete		Cat 2 gamete	Progeny
1/2 *P s*	×	1/2 *p S*	1/4 *Pp Ss* patches, short-haired
1/2 *P s*	×	1/2 *p s*	1/4 *Pp ss* patches, long-haired
1/2 *P s*	×	1/2 *p S*	1/4 *pp Ss* solid-colored, short-haired
1/2 *P s*	×	1/2 *p s*	1/4 *pp ss* solid-colored, long-haired

II. In tomatoes, red fruit is dominant to yellow fruit, and purple stems are dominant to green stems. The progeny from one mating consisted of 305 red fruit, purple stem plants; 328 red fruit, green stem plants; 110 yellow fruit, purple stem plants; and 97 yellow fruit, green stem plants. What were the genotypes of the parents in this cross?

Answer

This problem requires an understanding of independent assortment in a dihybrid cross as well as the ratios predicted from monohybrid crosses.

Designate the alleles:

R = red, r = yellow

P = purple stems, p = green stems

In genetics problems, the ratios of offspring can indicate the genotype of parents. You will usually need to total the number of progeny and approximate the ratio of offspring in each of the different classes. For this problem, in which the inheritance of two traits is given, consider each trait independently. For red fruit, there are 305 + 328 = 633 red-fruited plants out of a total of 840 plants. This value (633/840) is close to 3/4. About 1/4 of the plants have yellow fruit (110 + 97 = 207/840). From Mendel's work, you know that a 3:1 phenotypic ratio results from crosses between plants that are hybrid (heterozygous) for one gene. Therefore, the genotype for fruit color of each parent must have been *Rr*.

For stem color, 305 + 110 or 415/840 plants had purple stems. About half had purple stems, and the other half (328 + 97) had green stems. A 1:1 phenotypic ratio occurs when a heterozygote is mated to a homozygous recessive (as in a testcross). The parents' genotypes must have been *Pp* and *pp* for stem color.

The complete genotype of the parent plants in this cross was *Rr Pp* × *Rr pp*.

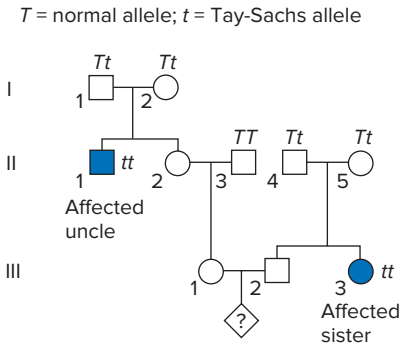
III. Tay-Sachs is a recessive lethal disease in which there is neurological deterioration early in life. This disease is rare in the population overall but is found at relatively high frequency in Ashkenazi Jews from Eastern Europe. A woman whose maternal uncle had the disease is trying to determine the probability that she and her husband could have an affected child. Her father does not come from a high-risk population. Her husband's sister died of the disease at an early age.

a. Draw the pedigree of the individuals described. Include the genotypes where possible.

b. Determine the probability that the couple's first child will be affected.

Answer

This problem requires an understanding of dominance/recessiveness and probability. First diagram the pedigree, and then assign as many genotypes as possible using the following allele designations:



The genotypes of the two affected individuals, the woman's uncle (II-1) and the husband's sister (III-3), are *tt*. Because the uncle was affected, both of his parents must have been heterozygous.

Similarly, as the husband's sister (III-3) is affected, both of her parents (II-4 and II-5) must be heterozygotes. Finally, because individual II-3 is not from a high-risk population, the most likely assumption is that he is *TT*.

You next need to determine the chance that a child of III-1 and III-2 (that is, individual IV-1) would have Tay-Sachs (*tt*). For that to be possible, both III-1 and III-2 must be *Tt* given that neither is *tt*. For III-1 to be *Tt*, II-2 must be *Tt*. Calculating the chance that II-2 is *Tt* is a bit tricky. At first, it appears that the chance is 1/2 that the daughter of two heterozygous (*Tt*) parents would be *Tt*: the expected progeny ratio is 1 *TT* : 2 *Tt* : 1 *tt*. However, in this case you have additional information to consider: II-2 is unaffected and thus the genotype *tt* is ruled out. That leaves 1 *TT* : 2 *Tt*, or a 2/3 chance that II-2 is *Tt*. If so, the chance that II-2 would transmit the *t* allele to III-1 is 1/2. Thus, the probability that III-1 is *Tt* is 2/3 × 1/2 = 1/3. This fact implies that II-2 could be either *TT* (probability = 1/3) or *Tt* (probability = 2/3). If II-2 is *Tt,* the chance that she would transmit the *t* allele to III-1 is 1/2. Thus, the probability that III-1 is *Tt* is 2/3 × 1/2 = 1/3.

What is the chance that III-2 is *Tt*? Both of his parents are heterozygous, and he is unaffected. Thus, using similar logic, the likelihood that III-2 is *Tt* is 2/3.

The probability that both III-1 and III-2 are *Tt* is 1/3 × 2/3 = 2/9. The chance that the child of two *Tt* parents would be *tt* is 1/4. Thus, the overall likelihood that IV-1, the child of III-1 and III-2, would have Tay-Sachs is 2/9 × 1/4 = 1/18.

PROBLEMS

Vocabulary

1. For each of the terms in the left column, choose the best matching phrase in the right column.

a. phenotype	1. having two identical alleles of a given gene
b. alleles	2. the allele expressed in the phenotype of the heterozygote
c. independent assortment	3. alternate forms of a gene
d. gametes	4. observable characteristic
e. gene	5. a cross between individuals both heterozygous for two genes
f. segregation	6. alleles of one gene separate into gametes randomly with respect to alleles of other genes
g. heterozygote	7. reproductive cells containing only one copy of each gene
h. dominant	8. the allele that does not contribute to the phenotype of the heterozygote
i. F_1	9. the cross of an individual of ambiguous genotype with a homozygous recessive individual
j. testcross	10. an individual with two different alleles of a gene
k. genotype	11. the heritable entity that determines a characteristic
l. recessive	12. the alleles an individual has
m. dihybrid cross	13. the separation of the two alleles of a gene into different gametes
n. homozygote	14. offspring of the P generation

Section 2.1

2. During the millennia in which selective breeding was practiced, why did breeders fail to uncover the principle that traits are governed by discrete units of inheritance (that is, by genes)?
3. Describe the characteristics of the garden pea that made it a good organism for Mendel's analysis of the basic principles of inheritance. Evaluate how easy or difficult it would be to make a similar study of inheritance in humans by considering the same attributes you described for the pea.

Section 2.2

4. An albino corn snake is crossed with a normal-colored corn snake. The offspring are all normal-colored. When these first-generation progeny snakes are crossed among themselves, they produce 32 normal-colored snakes and 10 albino snakes.
   a. How do you know that only a single gene is responsible for the color differences between these snakes?
   b. Which of these phenotypes is controlled by the dominant allele?
   c. A normal-colored female snake is involved in a testcross. This cross produces 10 normal-colored and 11 albino offspring. What are the genotypes of the parents and the offspring?
5. Two short-haired cats mate and produce six short-haired and two long-haired kittens. What does this information suggest about how hair length is inherited?
6. Piebald spotting is a condition found in humans in which there are patches of skin that lack pigmentation. The condition results from the inability of pigment-producing cells to migrate properly during development. Two adults with piebald spotting have one child who has this trait and a second child with normal skin pigmentation.
   a. Is the piebald spotting trait dominant or recessive? What information led you to this answer?
   b. What are the genotypes of the parents?
7. As a *Drosophila* research geneticist, you keep stocks of flies of specific genotypes. You have a fly that has normal wings (dominant phenotype). Flies with short wings are homozygous for a recessive allele of the wing-length gene. You need to know if this fly with normal wings is pure-breeding or heterozygous for the wing-length trait. What cross would you do to determine the genotype, and what results would you expect for each possible genotype?
8. A mutant cucumber plant has flowers that fail to open when mature. Crosses can be done with this plant by manually opening and pollinating the flowers with pollen from another plant. When closed × open crosses were done, all the F_1 progeny were open. The F_2 plants were 145 open and 59 closed. A cross of closed × F_1 gave 81 open and 77 closed. How is the closed trait inherited? What evidence led you to your conclusion?
9. In a particular population of mice, certain individuals display a phenotype called *short tail,* which is inherited as a dominant trait. Some individuals display a recessive trait called *dilute,* which affects coat color. Which of these traits would be easier to eliminate from the population by selective breeding? Why?
10. In humans, a dimple in the chin is a dominant characteristic controlled by a single gene.
    a. A man who does not have a chin dimple has children with a woman with a chin dimple whose

mother lacked the dimple. What proportion of their children would be expected to have a chin dimple?

b. A man with a chin dimple and a woman who lacks the dimple produce a child who lacks a dimple. What is the man's genotype?

c. A man with a chin dimple and a nondimpled woman produce eight children, all having the chin dimple. Can you be certain of the man's genotype? Why or why not? What genotype is more likely, and why?

11. Some inbred strains of the weedy plant *Arabidopsis thaliana* flower early in the growing season, but other strains flower at later times. Four different *Arabdiposis* plants (1–4) were crossed, and the resulting progeny were tabulated as follows:

Mating	Progeny
1 × 2	77 late : 81 early
1 × 3	134 late
1 × 4	93 late : 32 early
2 × 3	111 late
2 × 4	65 late : 61 early
3 × 4	126 late

a. Explain the genetic basis for the difference in flowering time. How do you know that among this group of plants, the flowering time trait is influenced by the action of a single gene? Which allele is dominant and which recessive?

b. Ascribe genotypes to the four plants.

c. What kinds of progeny would you expect if you allowed plants 1–4 to self-fertilize, and in what ratios?

12. Among Native Americans, two types of earwax (cerumen) are seen, dry and sticky. A geneticist studied the inheritance of this trait by observing the types of offspring produced by different kinds of matings. He observed the following numbers:

		Offspring	
Parents	**Number of mating pairs**	**Sticky**	**Dry**
Sticky × sticky	10	32	6
Sticky × dry	8	21	9
Dry × dry	12	0	42

a. How is earwax type inherited?

b. Why are no 3:1 or 1:1 ratios present in the data shown in the chart?

13. Imagine you have just purchased a black stallion of unknown genotype. You mate him to a red mare, and she delivers twin foals, one red and one black. Can you tell from these results how color is inherited, assuming that alternative alleles of a single gene are involved? What crosses could you do to determine how color is inherited?

14. If you roll a die (singular of dice), what is the probability you will roll: (a) a 6? (b) an even number? (c) a number divisible by 3? (d) If you roll a pair of dice, what is the probability that you will roll two 6s? (e) an even number on one and an odd number on the other? (f) matching numbers? (g) two numbers both over 4?

15. In a standard deck of playing cards, four suits exist (red suits = hearts and diamonds, black suits = spades and clubs). Each suit has 13 cards: Ace (A), 2, 3, 4, 5, 6, 7, 8, 9, 10, and the face cards Jack (J), Queen (Q), and King (K). In a single draw, what is the probability that you will draw a face card? A red card? A red face card?

16. How many genetically different eggs could be formed by women with the following genotypes?

a. *Aa bb CC DD*

b. *AA Bb Cc dd*

c. *Aa Bb cc Dd*

d. *Aa Bb Cc Dd*

17. What is the probability of producing a child that will phenotypically resemble either one of the two parents in the following four crosses? How many phenotypically different kinds of progeny could potentially result from each of the four crosses?

a. *Aa Bb Cc Dd* × *aa bb cc dd*

b. *aa bb cc dd* × *AA BB CC DD*

c. *Aa Bb Cc Dd* × *Aa Bb Cc Dd*

d. *aa bb cc dd* × *aa bb cc dd*

18. A mouse sperm of genotype *a B C D E* fertilizes an egg of genotype *a b c D e*. What are all the possibilities for the genotypes of (a) the zygote and (b) a sperm or egg produced by the mouse that develops from this fertilization?

19. Your friend is pregnant with triplets. She thinks that it is equally likely that she will be the mother of 3 sons, 3 daughters, 2 sons and 1 daughter, or 1 son and 2 daughters. Is she correct? Explain. (Assume that each of the triplets is from a separate fertilization, and that boys and girls are equally likely.)

20. Galactosemia is a recessive human disease that is treatable by restricting lactose and glucose in the diet. Susan Smithers and her husband are both heterozygous for the galactosemia gene.

a. Susan is pregnant with twins. If she has fraternal (nonidentical) twins, what is the probability both of the twins will be girls who have galactosemia?

b. If the twins are identical, what is the probability that both will be girls and have galactosemia?

For parts (c–g), assume that none of the children is a twin.

c. If Susan and her husband have four children, what is the probability that none of the four will have galactosemia?

d. If the couple has four children, what is the probability that at least one child will have galactosemia?

e. If the couple has four children, what is the probability that the first two will have galactosemia and the second two will not?

f. If the couple has three children, what is the probability that two of the children will have galactosemia and one will not, regardless of order?

g. If the couple has four children with galactosemia, what is the probability that their next child will have galactosemia?

21. Albinism is a condition in which pigmentation is lacking. In humans, the result is white hair, nonpigmented skin, and pink eyes. The trait in humans is caused by a recessive allele. Two normal parents have an albino child. What are the parents' genotypes? What is the probability that the next child will be albino?

22. A cross between two pea plants, both of which grew from yellow round seeds, gave the following numbers of seeds: 156 yellow round and 54 yellow wrinkled. What are the genotypes of the parent plants? (Yellow and round are dominant traits.)

23. A third-grader decided to breed guinea pigs for her school science project. She went to a pet store and bought a male with smooth black fur and a female with rough white fur. She wanted to study the inheritance of those features and was sorry to see that the first litter of eight contained only rough black animals. To her disappointment, the second litter from those same parents contained seven rough black animals. Soon the first litter had begun to produce F_2 offspring, and they showed a variety of coat types. Before long, the child had 125 F_2 guinea pigs. Eight of them had smooth white coats, 25 had smooth black coats, 23 were rough and white, and 69 were rough and black.

a. How are the coat color and texture characteristics inherited? What evidence supports your conclusions?

b. What phenotypes and proportions of offspring should the girl expect if she mates one of the smooth white F_2 females to an F_1 male?

24. The self-fertilization of an F_1 pea plant produced from a parent plant homozygous for yellow and wrinkled seeds and a parent homozygous for green and round seeds resulted in a pod containing seven F_2 peas. (Yellow and round are dominant.) What is the probability that all seven peas in the pod are yellow and round?

25. The achoo syndrome (sneezing in response to bright light) and trembling chin (triggered by anxiety) are both dominant traits in humans.

a. What is the probability that the first child of parents who are heterozygous for both the achoo gene and trembling chin will have achoo syndrome but lack the trembling chin?

b. What is the probability that the first child will have neither achoo syndrome nor trembling chin?

26. A pea plant from a pure-breeding strain that is tall, has green pods, and has purple flowers that are terminal is crossed to a plant from a pure-breeding strain that is dwarf, has yellow pods, and has white flowers that are axial. The F_1 plants are all tall and have purple axial flowers as well as green pods.

a. What phenotypes do you expect to see in the F_2?

b. What phenotypes and ratios would you predict in the progeny from crossing an F_1 plant to the dwarf parent?

27. The following table shows the results of different matings between jimsonweed plants that had either purple or white flowers and spiny or smooth pods. Determine the dominant allele for the two traits and indicate the genotypes of the parents for each of the crosses.

Parents	Offspring			
	Purple Spiny	**White Spiny**	**Purple Smooth**	**White Smooth**
a. purple spiny × purple spiny	94	32	28	11
b. purple spiny × purple smooth	40	0	38	0
c. purple spiny × white spiny	34	30	0	0
d. purple spiny × white spiny	89	92	31	27
e. purple smooth × purple smooth	0	0	36	11
f. white spiny × white spiny	0	45	0	16

28. A pea plant heterozygous for plant height, pod shape, and flower color was selfed. The progeny consisted of 272 tall, inflated pods, purple flowers; 92 tall, inflated, white flowers; 88 tall, flat pods, purple; 93 dwarf, inflated, purple; 35 tall, flat, white; 31 dwarf, inflated, white; 29 dwarf, flat, purple; 11 dwarf, flat, white. Which alleles are dominant in this cross?

29. In the fruit fly *Drosophila melanogaster,* the following genes and mutations are known:

Wing size: recessive allele for tiny wings *t;* dominant allele for normal wings *T.*

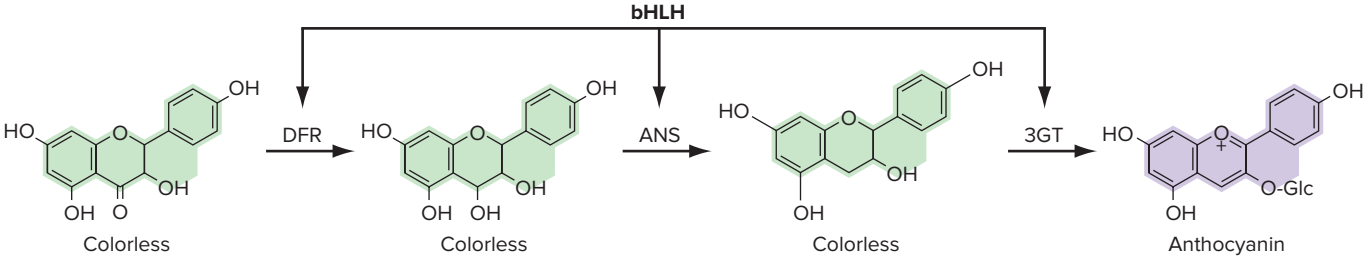
Eye shape: recessive allele for narrow eyes *n;* dominant allele for normal (oval) eyes *N.*

For each of the four following crosses, give the genotypes of each of the parents.

	Male			Female		
	Wings	**Eyes**		**Wings**	**Eyes**	**Offspring**
1	tiny	oval	×	tiny	oval	78 tiny wings, oval eyes 24 tiny wings, narrow eyes
2	normal	narrow	×	tiny	oval	45 normal wings, oval eyes 40 normal wings, narrow eyes 38 tiny wings, oval eyes 44 tiny wings, narrow eyes
3	normal	narrow	×	normal	oval	35 normal wings, oval eyes 29 normal wings, narrow eyes 10 tiny wings, oval eyes 11 tiny wings, narrow eyes
4	normal	narrow	×	normal	oval	62 normal wings, oval eyes 19 tiny wings, oval eyes

30. Based on the information you discovered in the previous problem, answer the following:
    a. A female fruit fly with genotype *Tt nn* is mated to a male of genotype *Tt Nn*. What is the probability that any one of their offspring will have normal phenotypes for both characters?
    b. What phenotypes would you expect among the offspring of this cross? If you obtained 200 progeny, how many of each phenotypic class would you expect?
31. Considering the yellow and green pea color phenotypes studied by Gregor Mendel:
    a. What is the biochemical function of the protein that is specified by the gene responsible for the pea color phenotype?
    b. *A null allele* of a gene is an allele that does not specify any of the biochemical function that the gene normally provides. Of the two alleles *Y* and *y*, which is more likely to be a null allele?
    c. In terms of the underlying biochemistry, why is the *Y* allele dominant to the *y* allele?
    d. Why are peas that are *yy* homozygotes green?
    e. The amount of the protein specified by a gene is roughly proportional to the number of functional copies of the gene carried by a cell or individual. What do the phenotypes of *YY* homozygotes, *Yy* heterozygotes, and *yy* homozygotes tell us about the amount of the Sgr enzyme (the product of the pea color gene) needed to produce a yellow color?
    f. The Sgr enzyme is not needed for the survival of a pea plant, but the genomes of organisms contain many so-called *essential genes* needed for an individual's survival. For such genes, heterozygotes for the normal allele and the null allele survive, but individuals homozygous for the null allele die soon after the male and female gametes, each with a null allele, come together at fertilization. In light of your answer to part (e), what does this fact tell you about the advantage to an organism of having two copies of their genes?
    g. Do you think that a single pea pod could contain peas with different phenotypes? Explain.
    h. Do you think that a pea pod could be of one color (say, green) while the peas within the pod could be of a different color (say, yellow)? Explain.
32. What would have been the outcome (the genotypic and phenotypic ratios) in the F_2 of Mendel's dihybrid cross shown in Fig. 2.15 if the alleles of the pea color gene (*Y, y*) and the pea shape gene (*R, r*) did not assort independently and instead the alleles inherited from a parent always stayed together as a unit?
33. Recall that Mendel obtained pure-breeding plants with either long or short stems and that hybrids had long stems (Fig. 2.8). Monohybrid crosses produced an F_2 generation with a 3:1 ratio of long stems to short stems, indicating that this difference in stem length is governed by a single gene. The gene that likely controlled this trait in Mendel's plants has been discovered, and it specifies an enzyme called G3βH, which catalyzes the reaction shown in the accompanying figure. The product of the reaction, gibberellin, is a growth hormone that makes plants grow tall. What is the most likely hypothesis to explain the difference between the dominant allele (*L*) and the recessive allele (*l*)?

34. The gene that likely controlled flower color (purple or white) in Mendel's pea plants has also been identified. The flower color gene specifies a protein



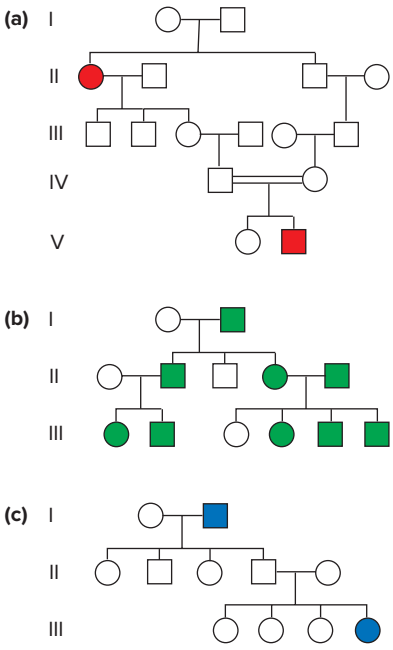
called bHLH required by cells to make three different enzymes (DFR, ANS, and 3GT) that function in the pathway shown in the accompanying figure, leading to synthesis of the purple pigment anthocyanin.

a. What is the most likely explanation for the difference between the dominant allele (*P*) and the recessive allele (*p*) of the gene responsible for these flower colors?

b. Given the biochemical pathway shown, could a different gene have been the one governing Mendel's flower colors?

Section 2.3

35. For each of the following human pedigrees, indicate whether the inheritance pattern is recessive or dominant. What feature(s) of the pedigree did you use to determine the mode of inheritance? Give the genotypes of affected individuals and of individuals who carry the disease allele but are not affected.



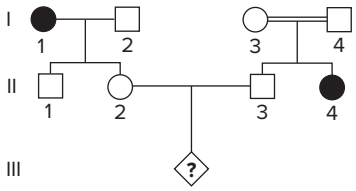
36. Consider the pedigree that follows for cutis laxa, a connective tissue disorder in which the skin hangs in loose folds.

a. Assuming that the trait is rare, what is the apparent mode of inheritance?

b. What is the probability that individual II-2 is a carrier?

c. What is the probability that individual II-3 is a carrier?

d. What is the probability that individual III-1 is affected by the disease?



37. A young couple went to see a genetic counselor because each had a sibling with cystic fibrosis. (Cystic fibrosis is a recessive disease, and neither member of the couple nor any of their four parents is affected.)

a. What is the probability that the female of this couple is a carrier?

b. What are the chances that their child will have cystic fibrosis?

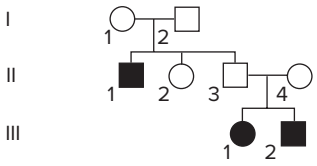
c. What is the probability that their child will be a carrier of the cystic fibrosis disease allele?

38. Huntington disease is a rare fatal, degenerative neurological disease in which individuals start to show symptoms in their 40s. It is caused by a dominant allele. Joe, a man in his 20s, just learned that his father has Huntington disease.

a. What is the probability that Joe will also develop the disease?

b. Joe and his new wife have been eager to start a family. What is the probability that their first child will eventually develop the disease?

39. Is the disease shown in the following pedigree caused by a dominant or a recessive allele? Why? Based on this limited pedigree, do you think the disease allele is rare or common in the population? Why?

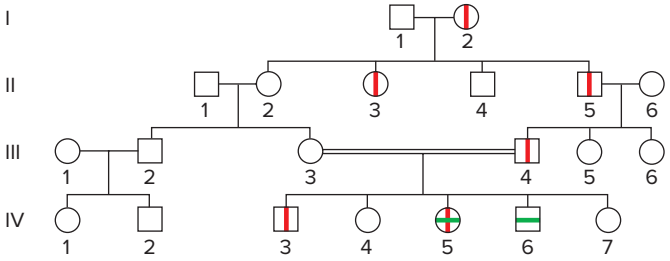


40. Figure 2.22 shows the inheritance of Huntington disease in a family from a small village near Lake Maracaibo in Venezuela. The village was founded by a small number of immigrants, and generations of their descendants have remained concentrated in this isolated location. The allele for Huntington disease has remained unusually prevalent there.

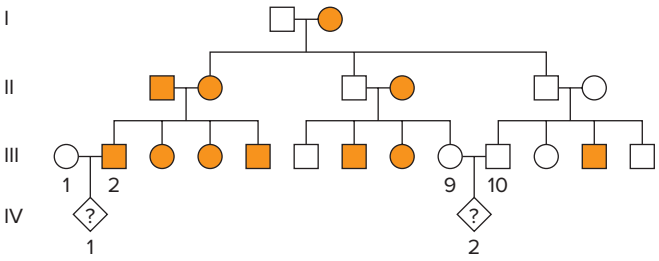
a. Why could you not conclude definitively that the disease is the result of a dominant or a recessive allele solely by looking at this pedigree?

b. Is there any information you could glean from the family's history that might imply the disease is due to a dominant rather than a recessive allele?

41. Consider the cystic fibrosis pedigree in Figure 2.24a.
    a. Assuming that one of the individuals in generation I was a carrier, and that no one from outside the family was a carrier, what was the probability that any single child of the consanguineous couple in generation V would have cystic fibrosis? (Assume that none of their children is born yet, so you don't know that VI-4 has the disease.)
    b. Assuming that one of the individuals in generation I was a carrier and that 1/1000 people in the population is a carrier, and knowing that VI-4 has the disease, how likely was it that VII-1 would be affected?
42. The common grandfather of two first cousins has hereditary hemochromatosis, a recessive condition causing an abnormal buildup of iron in the body. Neither of the cousins has the disease nor do any of their relatives.
    a. If the first cousins had a child, what is the chance that the child would have hemochromatosis? Assume that the unrelated, unaffected parents of the cousins are not carriers.
    b. How would your calculation change if you knew that 1 out of every 10 unaffected people in the population (including the unrelated parents of these cousins) was a carrier for hemochromatosis?
43. People with nail-patella syndrome have poorly developed or absent kneecaps and nails. Individuals with alkaptonuria have arthritis as well as urine that darkens when exposed to air. Both nail-patella syndrome and alkaptonuria are rare phenotypes. In the following pedigree, vertical red lines indicate individuals with nail-patella syndrome, while horizontal green lines denote individuals with alkaptonuria.
    a. What are the most likely modes of inheritance of nail-patella syndrome and alkaptonuria? What genotypes can you ascribe to each of the individuals in the pedigree for both of these phenotypes?
    b. In a mating between IV-2 and IV-5, what is the chance that the child produced would have both nail-patella syndrome and alkaptonuria? Nail-patella syndrome alone? Alkaptonuria alone? Neither defect?



44. Midphalangeal hair (hair on top of the middle segment of the fingers) is a common phenotype caused by a dominant allele *M*. Homozygotes for the recessive allele (*mm*) lack hair on the middle segment of their fingers. Among 1000 families in which both parents had midphalangeal hair, 1853 children showed the trait while 209 children did not. Explain this result.
45. A man with Huntington disease (he is heterozygous *HD* HD^+) and a normal woman have two children.
    a. What is the probability that only the second child has the disease?
    b. What is the probability that only one of the children has the disease?
    c. What is the probability that none of the children has the disease?
    d. Answer (a) through (c) assuming that the couple had 10 children.
    e. What is the probability that 4 of the 10 children in the family in (d) have the disease?
46. Explain why disease alleles for cystic fibrosis (*CF*) are recessive to the normal alleles (CF^+), yet the disease alleles responsible for Huntington disease (*HD*) are dominant to the normal alleles (HD^+).
47. The following pedigree shows the inheritance of red hair in a family in Scotland. Red hair is caused by homozygosity for a recessive allele of a gene called *MC1R*. Although worldwide red hair is the rarest of human hair colors, red hair is not uncommon in Scotland. In fact, 40% of Scots without red hair are nonetheless carriers of the red hair allele.



    a. Why does red hair show a horizontal inheritance pattern in this particular pedigree even though the trait is caused by a recessive allele?
    b. Assuming that individual III-2 has a child with the Scottish woman shown (III-1) who is not a close relative, what is the probability that this child (IV-1) will have red hair?
    c. What is the probability that the child of first cousins III-9 and III-10 (IV-2), will have red hair?

chapter 3

Extensions to Mendel's Laws

In this array of green, brown, and red lentils, some of the seeds have speckled patterns, while others are clear.

chapter outline

UNLIKE THE PEA traits that Mendel examined, most human characteristics do not fall neatly into just two opposing phenotypic categories. These *complex traits,* such as skin and hair color, height, athletic ability, and many others, seem to defy Mendelian analysis. The same can be said of traits expressed by many of the world's food crops: Their size, shape, succulence, and nutrient content vary over a wide range of values.

Lentils (*Lens culinaris*) provide a graphic illustration of this variation. A type of legume, lentils are grown in many parts of the world as a rich source of both protein and carbohydrate. The mature plants set fruit in the form of diminutive pods that contain two small seeds. These seeds can be ground into meal or used in soups, salads, and stews. Lentils come in an intriguing array of colors and patterns (**Fig. 3.1**), and commercial growers always seek to produce combinations to suit the cuisines of different cultures. But crosses between pure-breeding lines of lentils result in some startling surprises. A cross between pure-breeding tan and pure-breeding gray parents, for example, yields an all-brown F_1 generation. When these hybrids self-pollinate, the F_2 plants produce not only tan, gray, and brown lentils, but also green.

Beginning in the first decade of the twentieth century, geneticists subjected many kinds of plants and animals to controlled breeding tests, using Mendel's 3:1 phenotypic ratio as a guideline. If the traits under analysis behaved as predicted by Mendel's laws, then they were assumed to be determined by a single gene with alternative dominant and recessive alleles. Many traits, however, did not behave in this way. For some, no definitive dominance and recessiveness could be observed, or more than two alleles could be found in a particular cross (Fig. 3.1). Other traits turned out to be determined by two genes. Yet other traits were *multifactorial*, that is, determined by several different genes, or by the interaction of genes with the environment. The seed color of lentils is a multifactorial trait because color is controlled by multiple genes.

Because traits can arise from an intricate network of interactions, they do not always generate straightforward Mendelian phenotypic ratios. Nonetheless, simple extensions of Mendel's hypotheses can clarify the relationship between genotype and

Figure 3.1 Some phenotypic variation poses a challenge to Mendelian analysis. Lentils show complex speckling patterns that are controlled by a gene that has more than two alleles.

phenotype, allowing explanation of the observed deviations without challenging Mendel's basic laws.

One general theme stands out from these breeding studies: To make sense of the enormous phenotypic variation of the living world, geneticists usually try to limit the number of variables under investigation at any one time. Mendel did this by using pure-breeding, inbred strains of peas that differed from each other by one or a few traits, so that the action of single genes could be detected. Similarly, twentieth-century geneticists used inbred populations of fruit flies, mice, and other experimental organisms to study specific traits. Of course, geneticists cannot study people in this way. Human populations are typically far from inbred, and researchers cannot ethically perform breeding experiments on people. As a result, the genetic basis of much human variation remained a mystery. The advent of molecular biology in the 1970s provided new tools that geneticists now use to unravel the genetics of complex human traits, as will be described in later chapters.

3.1 Extensions to Mendel for Single-Gene Inheritance

learning objectives

1. Categorize allele interactions as completely dominant, incompletely dominant, or codominant.
2. Recognize progeny ratios that imply the existence of recessive lethal alleles.
3. Predict from the results of crosses whether a gene is polymorphic or monomorphic in a population.

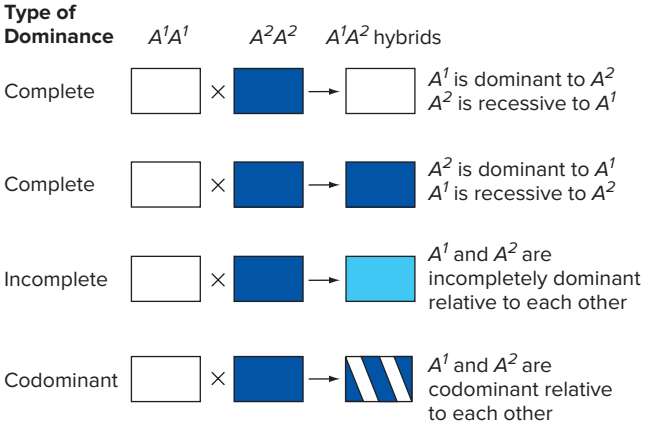
William Bateson was an early interpreter and defender of Mendel. Bateson, who coined the terms *genetics*, *allelomorph* (later shortened to *allele*), *homozygote*, and *heterozygote*, entreated the audience at a 1908 lecture: "Treasure your exceptions! … Keep them always uncovered and in sight. Exceptions are like the rough brickwork of a growing building which tells that there is more to come and shows where the next construction is to be." Consistent exceptions to simple Mendelian ratios revealed unexpected patterns of single-gene inheritance. By distilling the significance of these patterns, Bateson and other early geneticists extended the scope of Mendelian analysis and obtained a deeper understanding of the relationship between genotype and phenotype. We now look at the major extensions to Mendelian analysis elucidated over the last century.

Dominance Is Not Always Complete

A consistent working definition of dominance and recessiveness depends on the F_1 hybrids that arise from a mating between two pure-breeding lines. If a hybrid is identical to one parent for the trait under consideration, the allele carried by that parent is deemed dominant to the allele carried by the parent whose trait is not expressed in the hybrid. If, for example, a mating between a pure-breeding white line and a pure-breeding blue line produces F_1 hybrids that are white, the white allele of the gene for color is dominant to the blue allele. If the F_1 hybrids are blue, the blue allele is dominant to the white one (**Fig. 3.2**).

Mendel described and relied on complete dominance in sorting out his ratios and laws, but it is not the only kind of dominance he observed. Figure 3.2 diagrams two situations in which neither allele of a gene is completely dominant. As the figure shows, crosses between true-breeding strains can produce hybrids with phenotypes that differ from both parents. We now explain how these phenotypes arise.

Figure 3.2 Different dominance relationships. The phenotype of the heterozygote defines the dominance relationship between two alleles of the same gene (here, A^1 and A^2). Dominance is complete when the hybrid resembles one of the two pure-breeding parents. Dominance is incomplete when the hybrid resembles neither parent; its novel phenotype is usually intermediate. Codominance occurs when the hybrid shows traits from both pure-breeding parents.



Incomplete dominance: The F_1 hybrid resembles neither pure-breeding parent

A cross between pure late-blooming and pure early-blooming pea plants results in an F_1 generation that blooms in between the two extremes. This is just one of many examples of **incomplete dominance,** in which the hybrid does not resemble either pure-breeding parent. F_1 hybrids that differ from both parents often express a phenotype that is intermediate between those of the pure-breeding parents. Thus, with incomplete dominance, neither parental allele is dominant or recessive to the other; both contribute to the F_1 phenotype. Mendel observed plants that bloomed midway between two extremes when he cultivated various types of pure-breeding peas for his hybridization studies, but he did not pursue the implications. Blooming time was not one of the seven characteristics he chose to analyze in detail, almost certainly because in peas, the time of bloom was not as clear-cut as seed shape or flower color.

In many plant species, flower color serves as a striking example of incomplete dominance. With the floret clusters of snapdragons, for instance, a cross between pure-breeding red-flowered parents and pure-breeding white yields hybrids with pink blossoms, as if a painter had mixed red and white pigments to get pink (**Fig. 3.3a**). If allowed to self-pollinate, the F_1 pink-blooming plants produce F_2 progeny bearing red, pink, and white flowers in a ratio of 1:2:1 (**Fig. 3.3b**). This is the familiar *genotypic ratio* of an ordinary single-gene F_1 self-cross. What is new is that because the heterozygotes look unlike either homozygote, the *phenotypic ratios* are an exact reflection of the genotypic ratios.

The simplest biochemical explanation for this type of incomplete dominance is that each allele of the gene under analysis specifies an alternative form of a protein molecule with an enzymatic role in red pigment production. The white allele (A^2) does not give rise to a functional enzyme, but the red allele (A^1) does. Thus, in snapdragons, two red alleles per cell (A^1A^1) produce a double dose of a red-producing enzyme, which generates enough pigment to make the flowers look fully red. In the heterozygote (A^1A^2), one copy of the red allele per cell results in only enough pigment to make the flowers look pink. In the homozygote for the white allele (A^2A^2), where there is no functional enzyme and thus no red pigment, the flowers appear white.

Codominance: The F_1 hybrid exhibits traits of both parents

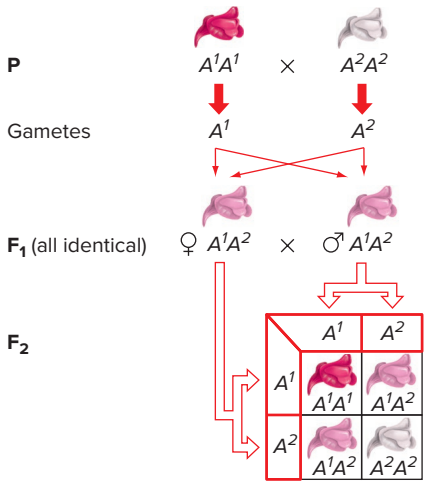
A cross between pure-breeding spotted lentils and pure-breeding dotted lentils produces heterozygotes that are both spotted and dotted (**Fig. 3.4a**). These F_1 hybrids illustrate a second significant departure from complete dominance. The progeny look like both parents, which means that neither the spotted nor the dotted allele is dominant or recessive to the other. Because both traits show up equally in the heterozygote's phenotype, the alleles are termed **codominant.** Self-pollination of the spotted/dotted F_1 generation generates F_2 progeny in the ratio of 1 spotted : 2 spotted/dotted : 1 dotted. The Mendelian 1:2:1 ratio among these F_2 progeny establishes that the spotted and dotted traits are determined by alternative alleles of a single gene. Once again, because the heterozygotes can be distinguished from both homozygotes, the phenotypic and genotypic ratios coincide.

In humans, some of the complex membrane-anchored molecules that distinguish different types of red blood cells exhibit codominance. For example, one gene (I) with alleles I^A and I^B controls the presence of a sugar polymer

Figure 3.3 Pink flowers are the result of incomplete dominance. (a) Color differences in these snapdragons reflect the activity of one pair of alleles. **(b)** The F_1 hybrids from a cross of pure-breeding red and white strains of snapdragons have pink blossoms. Flower colors in the F_2 appear in the ratio of 1 red : 2 pink : 1 white. This ratio signifies that the alleles of a single gene determine these three colors.

(a) *Antirrhinum majus* (snapdragons)

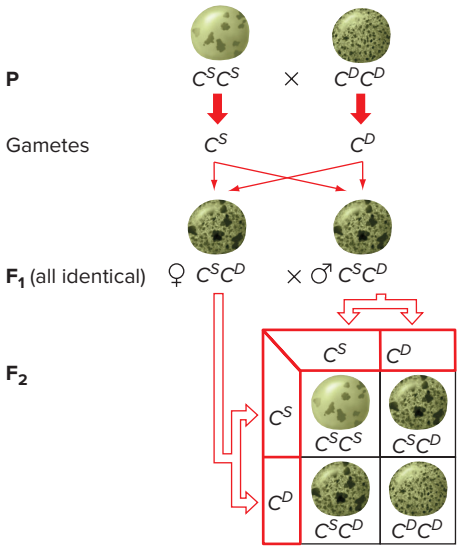
(b) A Punnett square for incomplete dominance



1 A^1A^1 (red) : 2 A^1A^2 (pink) : 1 A^2A^2 (white)

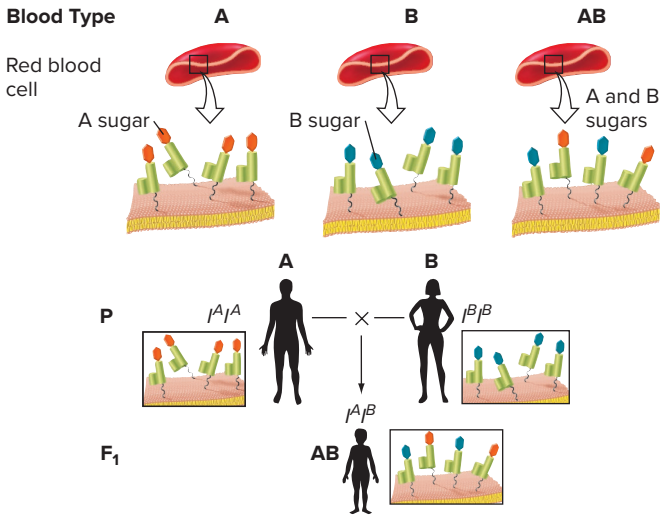
Figure 3.4 In codominance, F_1 hybrids display the traits of both parents. **(a)** A cross between pure-breeding spotted lentils and pure-breeding dotted lentils produces heterozygotes that are both spotted and dotted. Each genotype has its own corresponding phenotype, so the F_2 ratio is 1:2:1. **(b)** The I^A and I^B blood group alleles are codominant because the red blood cells of an I^AI^B heterozygote have both kinds of sugars at their surface.

(a) Codominant lentil coat patterns



1 C^SC^S (spotted) : 2 C^SC^D (spotted/dotted) : 1 C^DC^D (dotted)

(b) Codominant blood group alleles



that protrudes from the red blood cell membrane. Each of the alternative alleles encodes a slightly different form of an enzyme that causes production of a slightly different form of the complex sugar. In heterozygous individuals, the red blood cells carry both the I^A-determined and the I^B-determined sugars on their surface, whereas the cells of homozygous individuals display the products of either I^A or I^B alone (**Fig. 3.4b**). As this example illustrates, when both alleles produce a functional gene product, they are usually codominant for phenotypes analyzed at the molecular level.

Figure 3.2 summarizes the differences between complete dominance, incomplete dominance, and codominance for phenotypes reflected in color variations. Determinations of dominance relationships depend on the phenotype that appears in the F_1 generation. With complete dominance, F_1 progeny look like one of the true-breeding parents. Complete dominance, as we saw in Chapter 2, results in a 3:1 ratio of phenotypes in the F_2. With incomplete dominance, hybrids resemble neither of the parents and thus display neither pure-breeding trait. With codominance, the phenotypes of both pure-breeding lines show up simultaneously in the F_1 hybrid. Both incomplete dominance and codominance yield 1:2:1 F_2 ratios.

Mendel's law of segregation still holds

The dominance relations of a gene's alleles do not affect the alleles' transmission. Whether two alternative alleles of a single gene show complete dominance, incomplete dominance, or codominance depends on the kinds of proteins determined by the alleles and the biochemical function of those proteins in the cell. These phenotypic dominance relations, however, have no bearing on the segregation of the alleles during gamete formation.

As Mendel proposed, cells still carry two copies of each gene, and these copies—a pair of either similar or dissimilar alleles—segregate during gamete formation. Fertilization then restores two alleles to each cell without reference to whether the alleles are the same or different. Variations in dominance relations thus do not detract from Mendel's laws of segregation. Rather, they reflect differences in the way gene products control the production of phenotypes, adding a level of complexity to the tasks of interpreting the visible results of gene transmission and inferring genotype from phenotype.

A Gene May Have More Than Two Alleles

Mendel analyzed *either-or* traits controlled by genes with two alternative alleles, but for many traits, more than two alternatives exist. Here, we look at three such traits: human ABO blood types, lentil seed coat patterns, and human histocompatibility antigens.

ABO blood types

If a person with blood type A mates with a person with blood type B, it is possible in some cases for the couple to have a child that is neither A nor B nor AB, but a fourth blood type called O. The reason? The gene for the ABO blood types has three alleles: I^A, I^B, and i (**Fig. 3.5a**). Allele I^A gives rise to blood type A by specifying an enzyme that adds sugar A, I^B results in blood type B by specifying an enzyme that adds sugar B; i does not produce a functional sugar-adding enzyme. Alleles I^A and I^B are both dominant to i, and blood type O is therefore a result of homozygosity for allele i.

Figure 3.5 ABO blood types are determined by three alleles of one gene. (a) Six genotypes produce the four blood group phenotypes. **(b)** Blood serum contains antibodies against foreign red blood cell molecules. **(c)** If a recipient's serum has antibodies against the sugars on a donor's red blood cells, the blood types of recipient and donor are incompatible, and coagulation of red blood cells will occur during transfusions. In this table, a plus (+) indicates compatibility, and a minus (–) indicates incompatibility. Antibodies in the donor's blood usually do not cause problems because the amount of transfused antibody is small.

(a)

Genotypes	Corresponding Phenotypes: Type(s) of Molecule on Cell
$I^A I^A$ $I^A i$	A
$I^B I^B$ $I^B i$	B
$I^A I^B$	AB
ii	O

(b)

Blood Type	Antibodies in Serum
A	Antibodies against B
B	Antibodies against A
AB	No antibodies against A or B
O	Antibodies against A and B

(c)

Blood Type of Recipient	Donor Blood Type (Red Cells) A	B	AB	O
A	+	–	–	+
B	–	+	–	+
AB	+	+	+	+
O	–	–	–	+

Note in Fig. 3.5a that the A phenotype can arise from two genotypes, $I^A I^A$ or $I^A i$. The same is true for the B blood type, which can be produced by $I^B I^B$ or $I^B i$. But a combination of the two alleles $I^A I^B$ generates blood type AB.

We can draw several conclusions from these observations. First, as already stated, a given gene may have more than two alleles, or *multiple alleles*; in our example, the series of alleles is denoted I^A, I^B, and i.

Second, although the ABO blood group gene has three alleles, each person carries only two of the alternatives—$I^A I^A$, $I^B I^B$, $I^A I^B$, $I^A i$, $I^B i$, or ii. Thus six possible ABO genotypes exist. Because each individual carries no more than two alleles for each gene, no matter how many alleles are in a series, Mendel's law of segregation remains intact, because in a sexually reproducing organism, the two alleles of a gene separate during gamete formation.

Third, an allele is not inherently dominant or recessive; its dominance or recessiveness is always relative to a second allele. In other words, dominance relations are unique to a pair of alleles. In our example, I^A is completely dominant to i, but it is codominant with I^B. Given these dominance relations, the six genotypes possible with I^A, I^B, and i generate four different phenotypes: blood groups A, B, AB, and O. With this background, you can understand how a type A and a type B parent could produce a type O child: The parents must be $I^A i$ and $I^B i$ heterozygotes, and the child receives an i allele from each parent.

An understanding of the genetics of the ABO system has had profound medical and legal repercussions. Matching ABO blood types is a prerequisite of successful blood transfusions, because people make antibodies to foreign blood cell molecules. A person whose cells carry only A molecules, for example, produces anti-B antibodies; B people manufacture anti-A antibodies; AB individuals make neither type of antibody; and O individuals produce both anti-A and anti-B antibodies (**Fig. 3.5b**). These antibodies cause coagulation of cells displaying the foreign molecules (**Fig. 3.5c**). As a result, people with blood type O have historically been known as *universal donors* because their red blood cells carry no surface molecules that will stimulate an antibody attack in a transfusion recipient. In contrast, people with blood type AB are considered *universal recipients,* because they make neither anti-A nor anti-B antibodies, which, if present, would target the surface molecules of incoming blood cells.

Information about ABO blood types can also be used as legal evidence in court, to exclude the possibility of paternity or criminal guilt. In a paternity suit, for example, if the mother is type A and her child is type B, logic dictates that the I^B allele must have come from the father, whose genotype may be $I^A I^B$, $I^B I^B$, or $I^B i$. In 1944, the actress Joan Barry (phenotype A) sued Charlie Chaplin (phenotype O) for support of a child (phenotype B) whom she claimed he fathered. The scientific evidence indicated that Chaplin could not have been the father, since he was apparently ii and did not carry an I^B allele. This evidence was admissible in court, but the jury was not convinced, and Chaplin had to pay. Today, the molecular genotyping of DNA (*DNA fingerprinting,* see Chapter 11) provides a powerful tool to help establish paternity, guilt, or innocence, but juries still often find it difficult to evaluate such evidence.

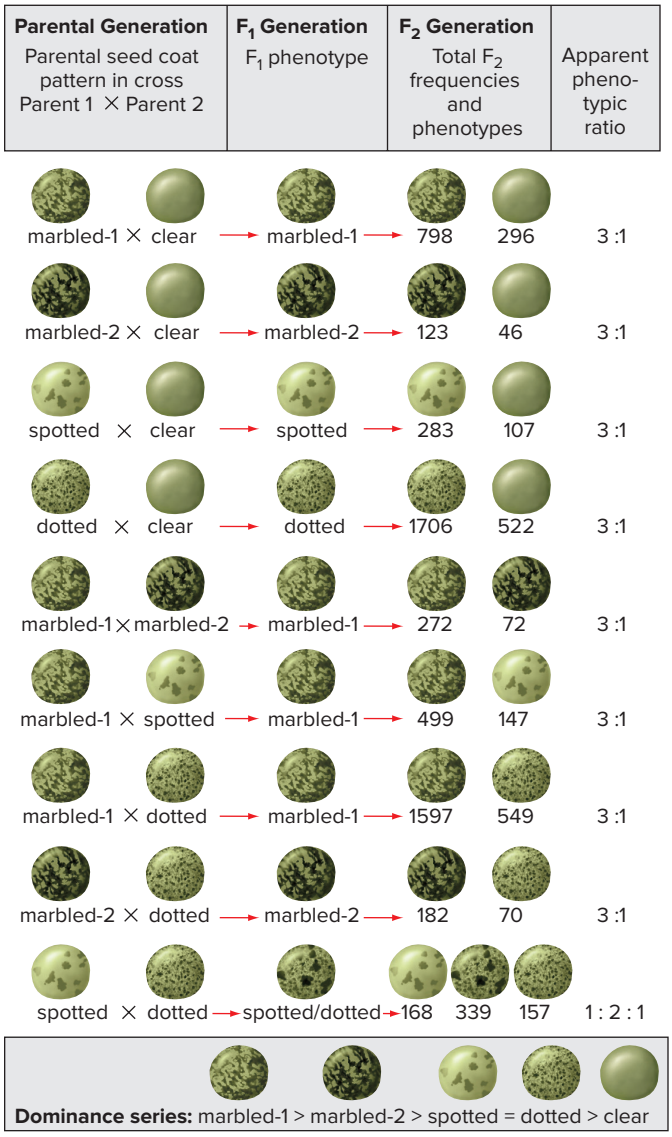
Lentil seed coat patterns

Lentils offer another example of multiple alleles. A gene for seed coat pattern has five alleles: spotted, dotted, clear (pattern absent), and two types of marbled. Reciprocal crosses between pairs of pure-breeding lines of all patterns (marbled-1 × marbled-2, marbled-1 × spotted, marbled-2 × spotted, and so forth) have clarified the dominance relations of all possible pairs of the alleles to reveal a **dominance series** in which alleles are listed in order from most dominant to most recessive. For example, crosses of marbled-1 with

marbled-2, or of marbled-1 with spotted or dotted or clear, produce the marbled-1 phenotype in the F_1 generation and a ratio of three marbled-1 to one of any of the other phenotypes in the F_2. These results indicate that the marbled-1 allele is completely dominant to each of the other four alleles.

Analogous crosses with the remaining four phenotypes reveal the dominance series shown in **Fig. 3.6.** Recall that dominance relations are meaningful only when comparing two alleles: An allele, such as marbled-2, can be recessive to a second allele (marbled-1) but dominant to a third and fourth (dotted and clear). The fact that all tested pairings of lentil seed coat pattern alleles yielded a 3:1 ratio in the F_2 generation (except for spotted × dotted, which yielded the 1:2:1 phenotypic ratio reflective of codominance) indicates that these lentil seed coat patterns are determined by different alleles of the same gene.

Figure 3.6 How to establish the dominance relations between multiple alleles. Pure-breeding lentils with different seed coat patterns are crossed in pairs, and the F_1 progeny are self-fertilized to produce an F_2 generation. The 3:1 or 1:2:1 F_2 monohybrid ratios from all of these crosses indicate that different alleles of a single gene determine all the traits. The phenotypes of the F_1 hybrids establish the dominance relationships (*bottom*). Spotted and dotted alleles are codominant, but each is recessive to the marbled alleles and is dominant to clear.



Histocompatibility in humans

In some multiple allelic series, each allele is codominant with every other allele, and every distinct genotype therefore produces a distinct phenotype. This happens particularly with traits defined at the molecular level. An extreme example is the group of three major genes that encode a family of related cell surface molecules in humans and other mammals known as *histocompatibility antigens*. Carried by all of the body's cells except the red blood cells and sperm, histocompatibility antigens play a crucial role in facilitating a proper immune response that destroys intruders (viral or bacterial, for example) while leaving the body's own tissues intact. Because each of the three major histocompatibility genes (called *HLA-A, HLA-B,* and *HLA-C* in humans) has between 400 and 1200 alleles, the number of possible allelic combinations in an individual creates a powerful potential for the phenotypic variation of cell surface molecules. Other than identical (that is, *monozygotic*) twins, no two people are likely to carry the same array of histocompatibility antigens on the surfaces of their cells.

The extreme variation in these proteins has important medical consequences, because people can make antibodies to *non-self* histocompatibility antigens different from their own. These antibodies can lead to rejection of transplanted organs. Doctors thus attempt to match as closely as possible the histocompatibility antigen types of transplant donors and recipients. Family members usually make the best organ donors, as the closer the genetic relationship between two people, the more likely they are to share *HLA* alleles.

Mutations Are the Source of New Alleles

How do the multiple alleles of an allelic series arise? The answer is that chance alterations of the genetic material, known as **mutations,** arise spontaneously in nature. Once they occur in gamete-producing cells, they are inherited faithfully. Mutations that have phenotypic consequences can be counted, and such counting reveals that they occur at low frequency. The frequency of gametes carrying a new mutation in a particular gene varies anywhere from 1 in 10,000 to 1 in 1,000,000. This range exists because different genes have different mutation rates.

Mutations make it possible to follow gene transmission. If, for example, a mutation specifies an alteration in an enzyme that normally produces yellow so that it now makes green, the new phenotype (green) will make it possible to recognize the new mutant allele. In fact, it takes at least two

different alleles, that is, some form of variation, to "see" the transmission of a gene. Thus, in segregation studies, geneticists can analyze only genes with variants; they have no way of following a gene that comes in only one form. If all peas were yellow, Mendel would not have been able to decipher the transmission patterns of the gene for the seed color trait. We discuss mutations in greater detail in Chapter 7.

Allele frequencies and monomorphic genes

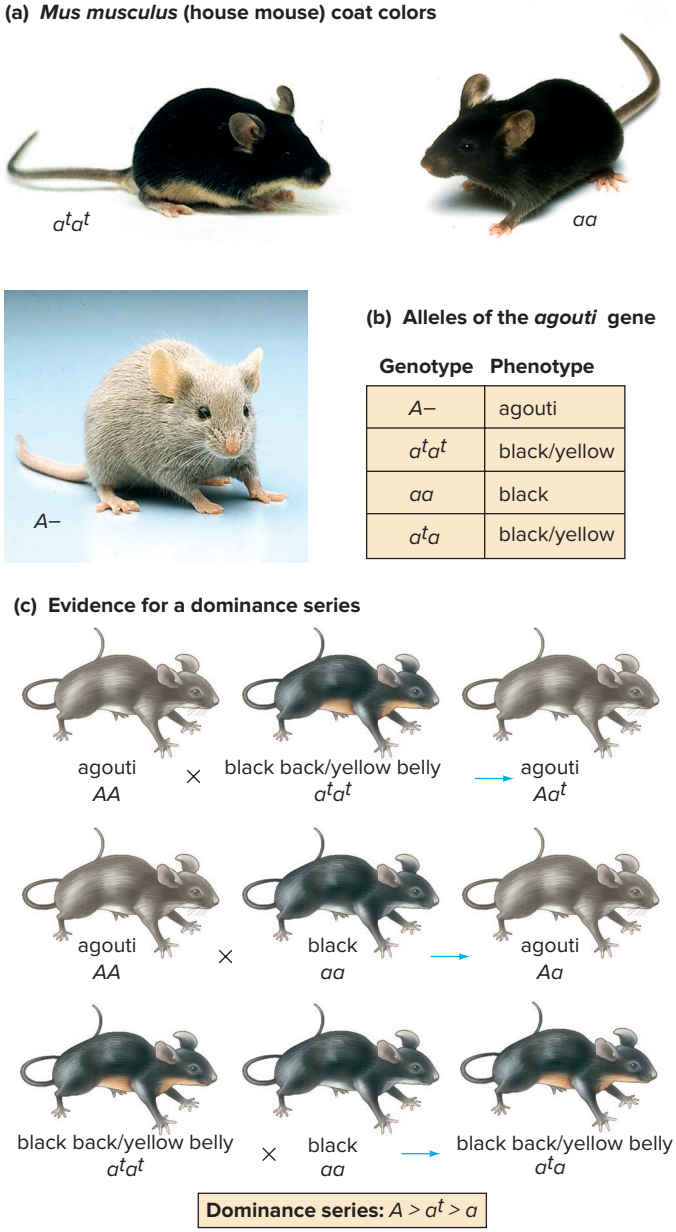
Because each organism carries two copies of every gene, you can calculate the number of copies of a gene in a given population by multiplying the number of individuals by 2. Each allele of the gene accounts for a percentage of the total number of gene copies, and that percentage is known as the **allele frequency.** The most common alleles in a population are usually called the **wild-type alleles,** often designated by a superscript plus sign ($^+$). An allele is considered wild-type if it is present in the population at a frequency greater than 1%. A rare allele in the same population is considered a **mutant allele.** (Note that the definitions of *wild-type* versus *mutant alleles* are not static. A newly induced mutation generates a mutant allele whose frequency can increase and over time, the allele can become wild-type.)

In mice, for example, one of the main genes determining coat color is the *agouti* gene. The wild-type allele (*A*) produces fur with each hair having yellow and black bands that blend together from a distance to give the appearance of dark gray, or agouti. Researchers have identified in the laboratory 14 distinguishable mutant alleles for the *agouti* gene. One of these (a^t) is recessive to the wild type and gives rise to a black coat on the back and a yellow coat on the belly; another (*a*) is also recessive to *A* and produces a pure black coat (**Fig. 3.7**). In nature, wild-type agoutis (*AA*) survive to reproduce, while very few black-backed or pure black mutants (a^ta^t or *aa*) do so because their dark coat makes it hard for them to evade the eyes of predators. As a result, *A* is present at a frequency of much more than 99% and is thus the only wild-type allele in mice for the *agouti* gene. A gene with only one common, wild-type allele is **monomorphic.**

Allele frequencies and polymorphic genes

In contrast, some genes have more than one common allele, which makes them **polymorphic.** For example, in the ABO blood type system, all three alleles—I^A, I^B, and *i*—have appreciable frequencies in most human populations. Although all three of these alleles can be considered to be wild-type, geneticists instead usually refer to the high-frequency alleles of a polymorphic gene as **common variants.** Certain rare genes are so polymorphic that hundreds of allelic variants can be found in populations. We have already discussed the case of the *HLA* histocompatibility genes in humans, which encode cell surface proteins that help the immune system deal with pathogenic invaders such as bacteria and viruses. Some scientists think that evolution favors the emergence of new *HLA* gene alleles to ensure that no single pathogen among the many to which we are exposed in the environment could destroy the entire human population. That is, at least a few individuals with particular *HLA* gene alleles would be protected from any given pathogen.

Figure 3.7 The mouse *agouti* gene: One wild-type allele, many mutant alleles. **(a)** Black-backed, yellow-bellied (*top left*); black (*top right*); and agouti (*bottom*) mice. **(b)** Genotypes and corresponding phenotypes for alleles of the *agouti* gene. **(c)** Crosses between pure-breeding lines reveal a dominance series. Interbreeding of the F_1 hybrids (*not shown*) yields 3:1 phenotypic ratios of F_2 progeny, indicating that *A*, a^t, and *a* are in fact alleles of one gene.



One Gene May Contribute to Several Characteristics

Mendel derived his laws from studies in which one gene determined one trait, but, always the careful observer, he himself noted possible departures. In listing the traits selected for his pea experiments, Mendel remarked that specific seed coat colors are always associated with specific flower colors.

The phenomenon of a single gene determining a number of distinct and seemingly unrelated characteristics is known as **pleiotropy.** Because geneticists now know that each gene determines a specific protein (or RNA) and that each gene product can have a cascade of effects on an organism, we can understand how pleiotropy arises. Among the Maori people of New Zealand, for example, many men develop respiratory issues and are also sterile. These men are said to exhibit a **syndrome**—a group of problems that are usually seen together. Researchers have found that the fault lies with the recessive allele of a single gene. The gene's normal dominant allele specifies a protein necessary for the action of cilia and flagella, both of which are hairlike structures extending from the surfaces of some cells. In men who are homozygous for the recessive allele, cilia that normally clear the airways fail to work effectively, and flagella that normally propel sperm fail to do their job. Thus, one gene determines a protein that affects both respiratory function and reproduction. Because most proteins act in a variety of tissues and influence multiple biochemical processes, mutations in almost any gene may have pleiotropic effects.

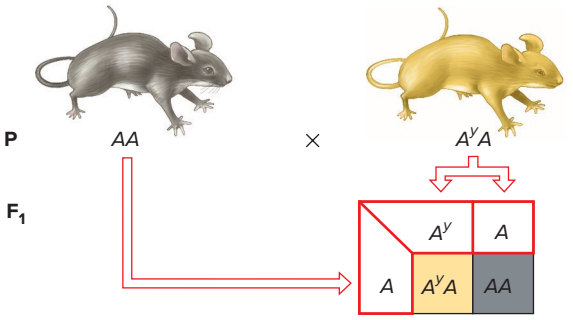
Recessive lethal alleles

A significant variation of pleiotropy occurs in alleles that not only produce a visible phenotype but also affect viability. Mendel assumed that all genotypes are equally viable—that is, they have the same likelihood of survival. If all genotypes were not equally viable, and a large percentage of, say, homozygotes for a particular allele died before germination or birth, you would not be able to count them. This lethality would alter the 1:2:1 genotypic ratios and the 3:1 phenotypic ratios predicted for the F_2 generation.

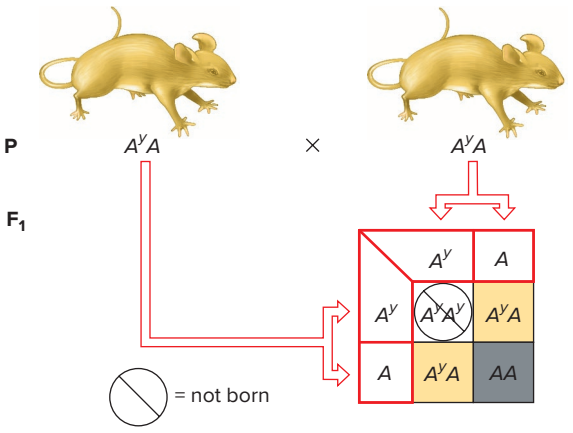
Consider the inheritance of coat color in mice. As mentioned earlier, wild-type agouti (*AA*) animals have black hairs with a yellow stripe that appear dark gray to the eye. One of the 14 known mutant alleles of the *agouti* gene, A^Y, gives rise to mice with a much lighter, almost yellow color. When pure-breeding *AA* mice are mated to yellow mice, the offspring always emerge in a 1:1 ratio of the two coat colors (**Fig. 3.8a**). From this result, we can draw three conclusions: (1) All yellow mice must carry the wild-type *A* allele even though they do not express the agouti phenotype; (2) yellow is therefore dominant to agouti; and (3) all yellow mice are A^YA heterozygotes.

Figure 3.8 A^Y **is a pleiotropic and recessive lethal allele.** **(a)** A cross between inbred agouti mice and yellow mice yields a 1:1 ratio of yellow to agouti progeny. The yellow mice are therefore A^YA heterozygotes, and for the trait of coat color, A^Y (for yellow) is dominant to *A* (for agouti). (Note we assume that if they could survive, A^YA^Y mice would have the same coat color as A^YA mice.) **(b)** Yellow mice do not breed true. In a yellow × yellow cross, the 2:1 ratio of yellow to agouti progeny indicates that the A^Y allele is a recessive lethal.

(a) All yellow mice are heterozygotes.



(b) Two copies of A^Y cause lethality.



Note again that dominance and recessiveness are defined in the context of each pair of alleles. Even though, as previously mentioned, agouti (*A*) is dominant to the a^t and *a* mutations for black coat color, it can still be recessive to the yellow coat color allele. The yellow mice in the preceding cross are A^yA heterozygotes, and the agoutis, *AA* homozygotes. So far, no surprises. But a mating of yellow to yellow produces a skewed phenotypic ratio of two yellow mice to one agouti (**Fig. 3.8b**). Among these progeny, matings between agouti mice show that the agoutis are all pure-breeding and therefore *AA* homozygotes as expected. However, no pure-breeding yellow mice appear among the progeny. When the yellow mice are mated to each other, they unfailingly produce 2/3 yellow and 1/3 agouti offspring, a ratio of 2:1, so the yellow mice must be heterozygotes (A^YA). In short, one can never obtain pure-breeding yellow mice (A^YA^Y).

How can we explain this phenomenon? The Punnett square in Fig. 3.8b suggests an answer. Two copies of the

A^y allele prove fatal to the animal carrying them, whereas one copy of the allele produces a yellow coat. This means that the A^y allele affects two different traits: It is dominant to A in the determination of coat color, but it is recessive to A in the production of lethality. An allele, such as A^y, that negatively affects the survival of a homozygote is known as a **recessive lethal allele.** Note that the same two alleles (A^y and A) can display different dominance relationships when looked at from the point of view of different phenotypes; we return later to this important point.

Because the A^y allele is dominant for yellow coat color, it is easy to detect carriers of this particular recessive lethal allele in mice. Such is not the case, however, for the vast majority of recessive lethal mutations, as these usually do not simultaneously show a visible dominant phenotype for some other trait. Lethal mutations can arise in many different genes, and as a result, most animals, including humans, carry some recessive lethal mutations. Such mutations usually remain hidden, except in rare cases of homozygosity, which in people are often caused by consanguineous matings (that is, matings between close relatives). If a mutation produces an allele that prevents production of a crucial molecule, homozygous individuals will not make any of the vital molecule and will not survive. Heterozygotes, by contrast, with only one copy of the deleterious mutation and one wild-type allele, can produce 50% of the wild-type amount of the normal molecule; this is usually sufficient to sustain normal cellular processes such that life goes on.

Delayed lethality

In the preceding discussion, we have described recessive alleles that result in the death of homozygotes prenatally; that is, *in utero.* With some mutations, however, homozygotes may survive beyond birth and die later from the deleterious consequences of the genetic defect. An example is seen in human infants with Tay-Sachs disease. The seemingly normal newborns remain healthy for five to six months but then develop blindness, paralysis, mental impairment, and other symptoms of a deteriorating nervous system; the disease usually proves fatal by the age of six. Tay-Sachs disease results from the absence of an active lysosomal enzyme called hexosaminidase A, leading to the accumulation of a toxic waste product inside nerve cells. The approximate incidence of Tay-Sachs among live births is 1/35,000 worldwide, but it is 1/3000 among Jewish people of Eastern European descent. Reliable tests that detect carriers, in combination with genetic counseling and educational programs, have all but eliminated the disease in the United States.

Recessive alleles causing prenatal or early childhood lethality can be passed on to subsequent generations only by heterozygous carriers because affected homozygotes die before they can mate. However, for late-onset diseases causing death in adults, homozygous patients can pass on the lethal allele before they become debilitated. An example is provided by the degenerative disease Friedreich ataxia: Some homozygotes first display symptoms of ataxia (loss of muscle coordination) at age 30–35 and die about five years later from heart failure.

Dominant alleles causing late-onset lethality can also be transmitted to subsequent generations; Figure 2.22 illustrates this fact for the inheritance of Huntington disease. By contrast, if the lethality caused by a dominant allele occurs instead during fetal development or early childhood, the allele will not be passed on, so all dominant early lethal mutant alleles must be new mutations.

Table 3.1 summarizes Mendel's basic assumptions about dominance, the number and viability of one gene's alleles, and the effects of each gene on phenotype, and then compares these assumptions with the extensions contributed by his twentieth-century successors. Through carefully controlled monohybrid crosses, these later geneticists analyzed the transmission patterns of the alleles of single genes, challenging and then confirming the law of segregation.

TABLE 3.1 For Traits Determined by One Gene: Alterations of the 3:1 Monohybrid Ratio

What Mendel Described	Extension	Extension's Effect on Heterozygous Phenotype	Extension's Effect on Ratios Resulting from an $F_1 \times F_1$ Cross
Complete dominance	Incomplete dominance Codominance	Unlike either homozygote	Phenotypes coincide with genotypes in a ratio of 1:2:1
Two alleles	Multiple alleles	Multiplicity of phenotypes	A series of 3:1 or 1:2:1 ratios
All alleles are equally viable	Recessive lethal alleles	Heterozygotes survive but may have visible phenotypes	2:1 instead of 3:1
One gene determines one trait	Pleiotropy: One gene influences several traits	Several traits affected in different ways, depending on dominance relations	Different ratios, depending on dominance relations for each affected trait

A Comprehensive Example: Sickle-Cell Disease Illustrates Many Extensions to Mendel's View of Single-Gene Inheritance

Sickle-cell disease is the result of a faulty hemoglobin molecule. Hemoglobin is composed of two types of polypeptide chains, alpha (α)-globin and beta (β)-globin, each specified by a different gene: *Hbα* for *α*-globin and *Hbβ* for β-globin. Normal red blood cells are packed full of millions upon millions of hemoglobin molecules, each of which picks up oxygen in the lungs and transports it to all the body's tissues.

Multiple alleles

The β-globin gene has a normal wild-type allele ($Hb\beta^A$) that gives rise to fully functional β-globin, as well as close to 400 mutant alleles that have been identified so far. Some of these mutant alleles result in the production of hemoglobin that carries oxygen only inefficiently. Other mutant alleles prevent the production of β-globin, causing a hemolytic (blood-destroying) disease called *β-thalassemia*. Here, we discuss the most common mutant allele of the β-globin gene, $Hb\beta^S$, which specifies an abnormal polypeptide that causes sickling of red blood cells (**Fig. 3.9a**).

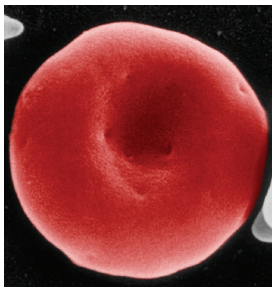
Pleiotropy

The $Hb\beta^S$ allele of the β-globin gene affects more than one trait (**Fig. 3.9b**). Hemoglobin molecules in the red blood cells of homozygous $Hb\beta^S$ $Hb\beta^S$ individuals undergo an aberrant transformation after releasing their oxygen. Instead of remaining soluble in the cytoplasm, they aggregate to form long fibers that deform the red blood cell from a normal biconcave disk to a sickle shape (see Fig. 3.9a). The deformed cells clog small blood vessels, reducing oxygen flow to the tissues and giving rise to muscle cramps, shortness of breath, and fatigue. The sickled cells are also fragile and easily broken. Consumption of fragmented cells by phagocytic white blood cells leads to a low red blood cell count, a condition called *anemia.*

On the positive side, $Hb\beta^S$ $Hb\beta^S$ homozygotes are resistant to malaria because the organism that causes the disease, *Plasmodium falciparum,* can multiply rapidly in normal red blood cells but cannot do so in cells that sickle. Infection by *P. falciparum* causes sickle-shaped cells to break down before the malaria organism has a chance to multiply.

Figure 3.9 Pleiotropy of sickle-cell anemia: Dominance relations vary with the phenotype under consideration. **(a)** A normal red blood cell (*top*) is easy to distinguish from the sickled cell in the scanning electron micrograph at the *bottom*. **(b)** Different levels of analysis identify various phenotypes. Dominance relationships between the $Hb\beta^S$ and $Hb\beta^A$ alleles of the *Hbβ* gene vary with the phenotype and sometimes even change with the environment.

a (top): © BSIP/Newscom; a (bottom): Source: Janice Haney Carr/CDC



(a)

Phenotypes at Different Levels of Analysis	Normal $Hb\beta^A$ $Hb\beta^A$	Carrier $Hb\beta^A$ $Hb\beta^S$	Diseased $Hb\beta^S$ $Hb\beta^S$	Dominance Relations at Each Level of Analysis
Red blood cell shape at sea level	Normal	Normal	Sickled cells present	$Hb\beta^A$ is dominant $Hb\beta^S$ is recessive
Red blood cell concentration at sea level	Normal	Normal	Lower	
β-globin polypeptide production	A protein	A and S proteins	S proteins	$Hb\beta^A$ and $Hb\beta^S$ are codominant
Red blood cell shape at high altitudes	Normal	Sickled cells present	Severe sickling	
Red blood cell concentration at high altitudes	Normal	Lower	Very low, anemia	$Hb\beta^A$ and $Hb\beta^S$ show incomplete dominance
Susceptibility to malaria	Normal susceptibility	Resistant	Resistant	$Hb\beta^S$ is dominant $Hb\beta^A$ is recessive

(b)

Recessive lethality

People who are homozygous for the recessive $Hb\beta^S$ allele often develop heart failure because of stress on the circulatory system. Many sickle-cell sufferers die in childhood, adolescence, or early adulthood.

Different dominance relations

Comparisons of heterozygous carriers of the sickle-cell allele—individuals whose cells contain one $Hb\beta^A$ and one $Hb\beta^S$ allele—with homozygous $Hb\beta^A$ $Hb\beta^A$ (normal) and homozygous $Hb\beta^S$ $Hb\beta^S$ (diseased) individuals make it possible to distinguish different dominance relationships for different phenotypic aspects of sickle-cell anemia (Fig. 3.9b).

At the molecular level—the production of β-globin proteins—both alleles are expressed such that $Hb\beta^A$ and $Hb\beta^S$ are codominant. At the cellular level, in their effect on red blood cell shape, the $Hb\beta^A$ and $Hb\beta^S$ alleles show either complete dominance or codominance depending on altitude. Under normal oxygen conditions, the great majority of a heterozygote's red blood cells have the normal biconcave shape ($Hb\beta^A$ is dominant to $Hb\beta^S$). When oxygen levels drop, however, sickling occurs in some $Hb\beta^A$ $Hb\beta^S$ cells ($Hb\beta^A$ and $Hb\beta^S$ are codominant). During World War II, soldiers who were heterozygous carriers and who were airlifted in transport planes to cross the Pacific experienced sickling crises for this reason.

Considering the trait of resistance to malaria, the $Hb\beta^S$ allele is dominant to the $Hb\beta^A$ allele. The reason is that infected $Hb\beta^A$ $Hb\beta^S$ cells are resistant to malaria because they break down before the malarial organism has a chance to reproduce, just like the $Hb\beta^S$ $Hb\beta^S$ cells described previously. But luckily for the heterozygote, for the phenotypes of anemia or death, $Hb\beta^S$ is recessive to $Hb\beta^A$. A corollary of this observation is that in its effect on general health under normal environmental conditions and its effect on red blood cell count, the $Hb\beta^A$ allele is dominant to $Hb\beta^S$.

Thus, for the β-globin gene, as for other genes, dominance and recessiveness are not an inherent quality of alleles in isolation; rather, they are specific to each pair of alleles and to the level of physiology at which the phenotype is examined. When discussing dominance relationships, it is therefore essential to define the particular phenotype under analysis.

The complicated dominance relationships between the $Hb\beta^A$ and $Hb\beta^S$ alleles help explain the puzzling observation that the normally deleterious allele $Hb\beta^S$ is widespread in certain populations. In areas where malaria is endemic, heterozygotes are better able to survive and pass on their genes than are either type of homozygote. $Hb\beta^S$ $Hb\beta^S$ individuals often die of sickle-cell disease, while those with the genotype $Hb\beta^A$ $Hb\beta^A$ often die of malaria. Heterozygotes, however, are relatively immune to both conditions, so high frequencies of both alleles persist in tropical environments where malaria is found. We explore this phenomenon in more quantitative detail in Chapter 21 on population genetics.

essential concepts

- Two alleles of a single gene may exhibit *complete dominance*, in which heterozygotes resemble the homozygous dominant parent; *incomplete dominance*, in which heterozygotes have an intermediate phenotype; and *codominance,* in which heterozygotes display aspects of each homozygous phenotype.
- New alleles of a gene arise by mutation. Alleles with a frequency greater than 1% in a population are *wild-type;* alleles that are less frequent are *mutant.*
- When two or more wild-type alleles (*common variants*) exist for a gene, the gene is *polymorphic;* a gene with only one wild-type allele is *monomorphic.*
- In *pleiotropy,* one gene contributes to multiple traits. The dominance relationship between any two alleles can vary depending on the trait.
- Homozygotes for a *recessive lethal allele* that fails to provide an essential function will die. If a recessive lethal allele has dominant effects on a visible trait, two-thirds of the surviving progeny of a cross between heterozygotes will display this trait.

3.2 Extensions to Mendel for Two-Gene Inheritance

learning objectives

1. Conclude from the results of crosses whether a single gene or two genes control a trait.
2. Infer from the results of crosses the existence of interactions between alleles of different genes including: additivity, epistasis, redundancy, and complementation.

Two genes can interact in several ways to determine a single trait, such as the color of a flower, a seed coat, a chicken's feathers, a dog's fur, or the shape of a plant's leaves. In a dihybrid cross like Mendel's, each type of interaction produces its own signature of phenotypic ratios. In the following examples, the alternate alleles of each of the two genes are completely dominant (such as *A* and *B*) and recessive (*a* and *b*). For simplicity, we sometimes refer to a gene name using the symbol for the dominant allele, for example, gene *A*. In addition, we refer to the protein

product of allele *A* as protein A (no italics), and when appropriate, that of allele *a* as protein a (no italics).

Additive Interactions Between Two Genes Controlling a Single Trait Can Produce Novel Phenotypes

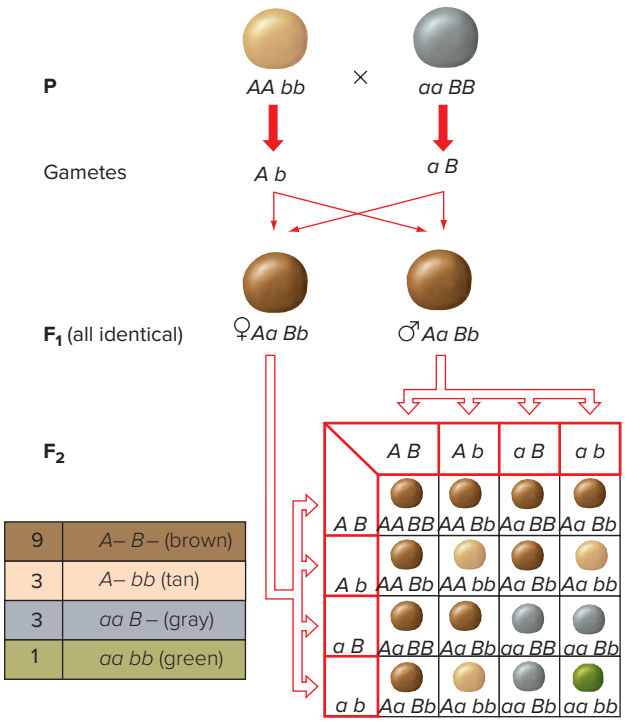
In the chapter opening, we described a mating of tan and gray lentils that produced a uniformly brown F_1 generation and then an F_2 generation containing brown, tan, gray, and green lentil seeds. An understanding of how this can happen emerges from experimental results demonstrating that the ratio of the four F_2 colors is 9 brown : 3 tan : 3 gray : 1 green (**Fig. 3.10a**). Recall from Chapter 2 that this is the same ratio Mendel observed in his analysis of the F_2 generations from dihybrid crosses following two independently assorting genes. In Mendel's studies, each of the four classes consisted of plants that expressed a combination of two unrelated traits. With lentils, however, we are looking at a single trait—seed color. The simplest explanation for the parallel ratios is that a combination of genotypes at two independently assorting genes interacts *additively* to produce the phenotype of seed color in lentils.

Results obtained from self-crosses with the various types of F_2 lentil plants support a two-gene explanation. Self-crosses of F_2 green individuals show that they are pure-breeding, producing an F_3 generation that is entirely green. Tan individuals generate either all tan offspring, or a mixture of tan offspring and green offspring. Grays similarly produce either all gray, or gray and green. Self-crosses of brown F_2 individuals can have four possible outcomes: all brown, brown plus tan, brown plus gray, or all four colors (**Fig. 3.10b**). The two-gene hypothesis explains why:

- only one green genotype exists: pure-breeding *aa bb,* but
- two types of tans exist: pure-breeding *AA bb* as well as tan- and green-producing *Aa bb,* and
- two types of grays exist: pure-breeding *aa BB* and gray- and green-producing *aa Bb,* yet
- four types of browns exist: true-breeding *AA BB,* brown- and tan-producing *AA Bb,* brown- and gray-producing *Aa BB,* and *Aa Bb* dihybrids that give rise to plants producing lentils of all four colors.

Figure 3.10 How two genes interact to produce seed colors in lentils. **(a)** In a cross of pure-breeding tan and gray lentils, all the F_1 hybrids are brown, but four different phenotypes appear among the F_2 progeny. The 9:3:3:1 ratio of F_2 phenotypes suggests that seed coat color is determined by two independently assorting genes. **(b)** Expected results of selfing individual F_2 plants of the indicated phenotypes to produce an F_3 generation, if seed coat color results from the interaction of two genes. The third column shows the proportion of the F_2 population that would be expected to produce the observed F_3 phenotypes. **(c)** Other two-generation crosses involving pure-breeding parental lines also support the two-gene hypothesis. In this table, the F_1 hybrid generation has been omitted.

(a) A dihybrid cross with lentil coat colors



9	*A– B–* (brown)
3	*A– bb* (tan)
3	*aa B–* (gray)
1	*aa bb* (green)

(b) Self-pollination of the F_2 to produce an F_3

Phenotypes of F_2 Individuals	Observed F_3 Phenotypes	Expected Proportion of F_2 Population*
Green	Green	1/16
Tan	Tan	1/16
Tan	Tan, green	2/16
Gray	Gray, green	2/16
Gray	Gray	1/16
Brown	Brown	1/16
Brown	Brown, tan	2/16
Brown	Brown, gray	2/16
Brown	Brown, gray, tan, green	4/16

*This 1:1: 2 : 2 :1: 1: 2 : 2 : 4 F_2 genotypic ratio corresponds to a 9 brown :3 tan : 3 gray :1 green F_2 phenotypic ratio.

(c) Sorting out the dominance relations by select crosses

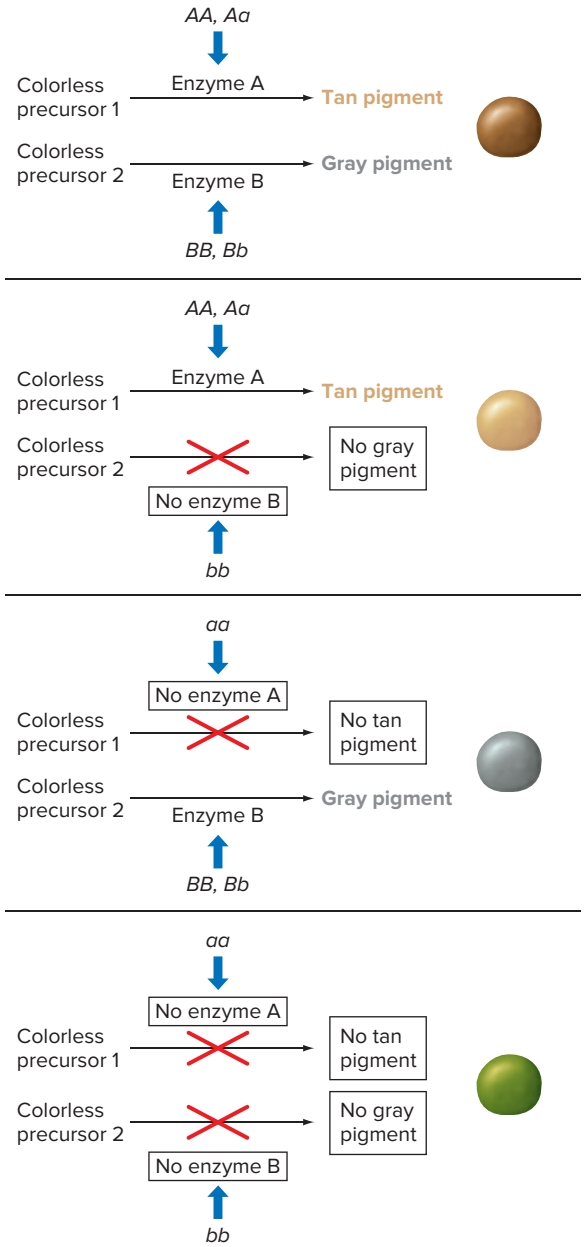
Seed Coat Color of Pure-Breeding Parents	F_2 Phenotypes and Frequencies	Ratio
Tan × green	231 tan, 85 green	3:1
Gray × green	2586 gray, 867 green	3:1
Brown × gray	964 brown, 312 gray	3:1
Brown × tan	255 brown, 76 tan	3:1
Brown × green	57 brown, 18 gray, 13 tan, 4 green	9:3:3:1

In short, for the two genes that determine seed color, both dominant alleles must be present to yield brown (*A– B–*); the dominant allele of one gene produces tan (*A– bb*); the dominant allele of the other specifies gray (*aa B–*); and the complete absence of dominant alleles (that is, the double recessive) yields green (*aa bb*). Thus, the four color phenotypes arise from four **genotypic classes,** with each class defined in terms of the presence or absence of the dominant alleles of two genes: (1) both present (*A– B–*), (2) one present (*A– bb*), (3) the other present (*aa B–*), and (4) neither present (*aa bb*). Note that the *A–* notation means that the second allele of this gene can be either *A* or *a,* while *B–* denotes a second allele of either *B* or *b.* Note also that only with a two-gene system in which the dominance and recessiveness of alleles at both genes is complete can the nine different genotypes of the F_2 generation be categorized into the four phenotypic classes described. With incomplete dominance or co-dominance, the F_2 genotypes could not be grouped together in this simple way, as they would give rise to more than four phenotypes.

Further crosses between plants carrying lentils of different colors confirmed the two-gene hypothesis (**Fig. 3.10c**). Thus, the 9:3:3:1 phenotypic ratio of brown to tan to gray to green in an F_2 descended from pure-breeding tan and pure-breeding gray lentils tells us not only that dominant and recessive alleles of two genes assort independently and interact to produce the seed color, but also that each genotypic class (*A– B–, A– bb, aa B–,* and *aa bb*) determines a particular phenotype.

How can we explain the 9:3:3:1 phenotypic ratio in terms of the action of the protein products of these two genes? We cannot answer this question definitively because the genes controlling lentil seed color have not been identified at the molecular level, and the biochemical pathways in which they function are not known. However, information available about the mechanisms of seed color inheritance in other plant species allows us to formulate a plausible model for this system in lentils (**Fig. 3.11**). The model illustrates an important implication of the 9:3:3:1 ratio: The two independently assorting genes controlling the same trait probably function additively in independent biochemical pathways, so that in this case, tan (the product of one pathway) + gray (the product of the other pathway) = brown.

Figure 3.11 A biochemical model for the inheritance of lentil seed colors. The seed has an opaque outer layer (the seed coat) and an inner layer (the cotyledon). The green chlorophyll in the cotyledon is not visible if the seed coat is colored. Allele *A* encodes enzyme A. Allele *a* does not produce this enzyme. Allele *B* of a second gene encodes a different enzyme; *b* produces none of this enzyme. Seeds appear brown if the tan and gray pigments are both present. In the absence of both enzymes (*aa bb*), the seed coat is nonpigmented, so the green chlorophyll in the cotyledon will show through. The 9:3:3:1 ratio implies that the *A* and *B* genes operate in independent biochemical pathways.



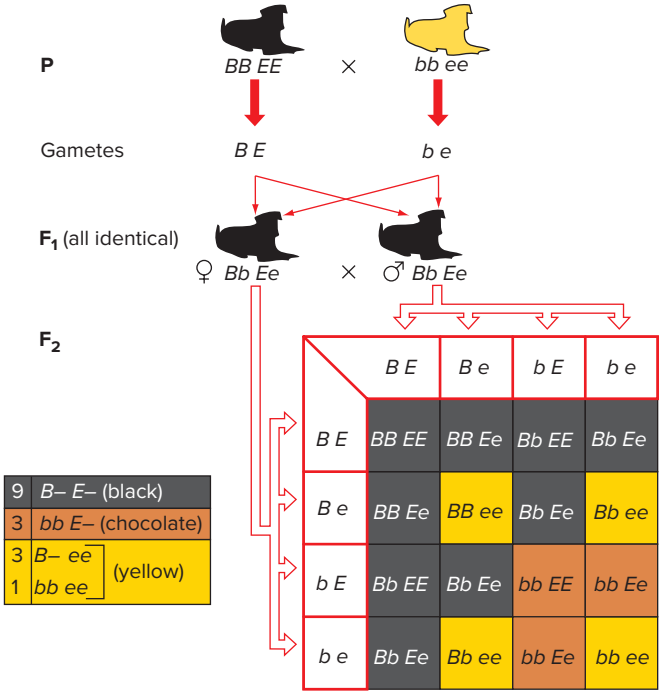
Epistasis: Alleles of One Gene Can Mask the Phenotypic Effects of Alleles of Another Gene

Sometimes, when two genes control a single trait, the four Mendelian genotypic classes produce fewer than four observable phenotypes because one gene masks the phenotypic effects of another. A gene interaction in which an allele at one gene hides the effects of alleles at another gene is known as **epistasis;** the allele that is doing the masking is *epistatic* to the gene that is being masked (the *hypostatic* gene).

Figure 3.12 Recessive epistasis determines coat color in Labrador retrievers. **(a)** Labrador retriever colors. **(b)** Yellow Labrador retrievers are homozygous for the recessive *e* allele, which masks the effects of the *B* or *b* alleles of a second coat color gene. In *E*– dogs, a *B*– genotype produces black and a *bb* genotype produces brown.

(a) Chocolate, yellow, and black Labrador retrievers

(b) A dihybrid cross showing recessive epistasis



Recessive epistasis

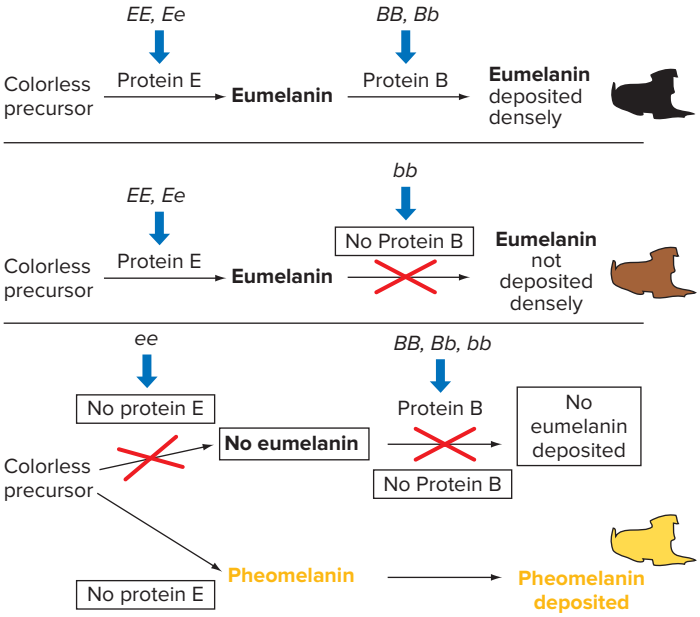
We present here three examples of **recessive epistasis,** where homozygosity for a recessive allele of one gene hides the effect of a second gene. In other words, when an individual is homozygous for the epistatic recessive allele of the first gene, the phenotype is independent of the alleles present at the second (hypostatic) gene. The final example in this section describes a surprising phenomenon in which recessive epistasis is reciprocal between the two genes that determine the trait.

Yellow Labrador retrievers The sleek, short-haired coat of Labrador retrievers can be black, chocolate brown, or yellow (**Fig. 3.12a**). Which color appears depends on the allelic combinations of two independently assorting coat color genes (**Fig. 3.12b**). When the dominant *E* allele of the first gene is present, the *B* allele of the second gene determines black, and the recessive *bb* homozygote is chocolate. However, a double dose of the recessive allele (*ee*) hides the effect of any combination of the black or chocolate alleles to yield yellow. Thus, the recessive *ee* homozygous genotype is epistatic to any allelic combination at the second, hypostatic gene, *B*.

Let's look at the phenomenon in greater detail. Crosses between pure-breeding black retrievers (*BB EE*) and one type of pure-breeding yellow retriever (*bb ee*) create an F_1 generation of dihybrid black retrievers (*Bb Ee*). Crosses between these F_1 dihybrids produce an F_2 generation with nine black dogs (*B*– *E*–) for every three brown (*bb E*–) and four yellow (– – *ee*) (Fig. 3.12b). Note that only three phenotypic classes exist because the two genotypic classes without a dominant *E* allele—the three *B*– *ee* and the one *bb ee*—combine to produce a yellow phenotype. The telltale ratio of recessive epistasis in the F_2 generation is thus 9:3:4, with the 4 representing a combination of 3 (*B*– *ee*) + 1 (*bb ee*). Because the *ee* genotype masks the influence of the other gene for coat color, you cannot tell by looking at a yellow Labrador whether its genotype at the *B* locus is *B*– (black) or *bb* (chocolate).

Scientists understand with some precision the biochemical pathways in which different alleles of the *B* and *E* genes operate (**Fig. 3.13**). All coat color in dogs comes from two pigments synthesized from a common precursor: a dark pigment called eumelanin and a light pigment called pheomelanin. When Labrador retrievers have at least one copy of the *E* allele, the resultant protein E ensures that the animals will make only eumelanin and no pheomelanin. The protein specified by the *B* allele is required for eumelanin synthesis and its deposition in the hair, while the protein made by the *b* allele is less efficient. As a result, chocolate *E*– *bb* dogs have less eumelanin in their hairs than black dogs with at least one *B* allele (*E*– *B*–). But in the absence of the E protein (in *ee* dogs), only pheomelanin is synthesized, and so the dogs appear yellow. It is easy to

Figure 3.13 A biochemical explanation for coat color in Labrador retrievers. Protein E activates an enzyme that generates eumelanin from a colorless precursor. When protein E is present, only eumelanin is produced. Protein B deposits eumelanin densely so that the hair is black. Pigment is deposited less densely without protein B, producing brown hair (chocolate). In the absence of protein E, no eumelanin is produced and instead, pheomelanin (yellow pigment) is synthesized. Homozygous *ee* dogs are always yellow regardless of the gene *B* genotype. The reason is that protein B affects eumelanin only, but these *ee* animals have no eumelanin.



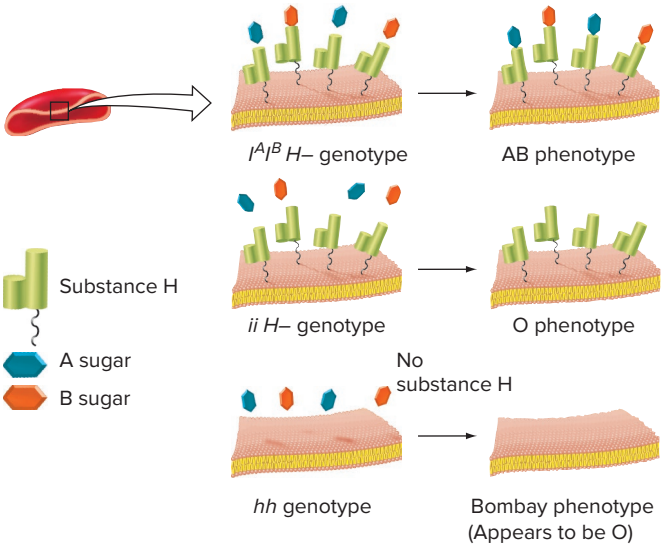
see in Fig. 3.13 why *ee* is epistatic to both alleles of gene *B*: In *ee* dogs, no eumelanin is present, so the dogs are yellow regardless of whether they are *B*– or *bb*.

The Bombay phenotype in humans An understanding of recessive epistasis made it possible to resolve an intriguing puzzle in human genetics. In rare instances, two parents who appear to have blood type O, and thus would be predicted to be genotype *ii,* produce a child who is either blood type A (genotype I^Ai) or blood type B (genotype I^Bi). This phenomenon occurs because an extremely rare trait, called the Bombay phenotype after its discovery in Bombay, India, superficially resembles blood type O. As **Fig. 3.14a** shows, the Bombay phenotype actually arises from homozygosity for a mutant recessive allele (*hh*) of a second gene that masks the effects of any ABO alleles that might be present.

Here's how it works at the molecular level (Fig. 3.14a). In the construction of the red blood cell surface molecules that determine blood type, type A individuals make an enzyme that adds polysaccharide A onto a sugar polymer known as substance H; type B individuals make an altered form of the enzyme that adds polysaccharide B onto the sugar polymer H; and type O individuals make neither A-adding nor B-adding enzyme and thus have an exposed substance H in the membranes of their red blood cells. All people of A, B, or O phenotype carry at least one dominant wild-type *H* allele for the second gene and thus produce some substance H. In contrast, the rare Bombay-phenotype individuals, with

Figure 3.14 Recessive epistasis in humans causes a rare blood type. **(a)** Homozygosity for the *h* Bombay allele is epistatic to the *I* gene determining ABO blood types. *hh* individuals fail to produce substance H, which is needed for the addition of A or B sugars at the surface of red blood cells. **(b)** Because *h* is epistatic to *I*, rare individuals may appear to have blood type O despite having an I^A or I^B allele. When the masked *I* allele is expressed in their *Hh* progeny, these people may be surprised by their child's blood type.

(a) Molecular basis of the Bombay phenotype



(b) How epistasis causes unexpected inheritance pattern of ABO blood type

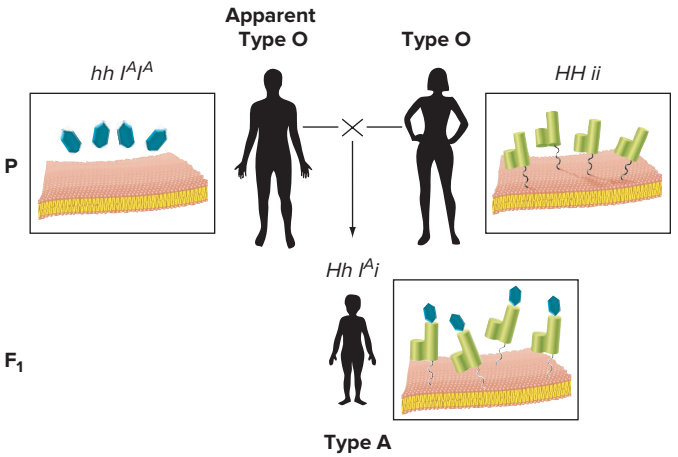
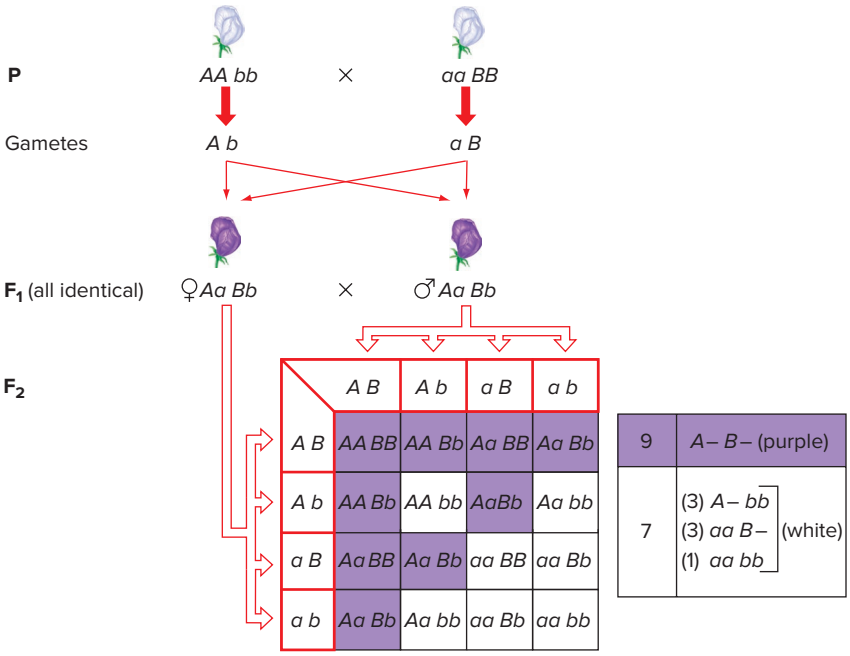


Figure 3.15 Dominant alleles of two genes needed for purple color in sweet peas. (a) White and purple sweet pea flowers. **(b)** The 9:7 ratio of purple to white F_2 plants indicates that at least one dominant allele of each gene is necessary for the development of purple color.

(a) ***Lathyrus odoratus*** **(sweet peas)**

(b) A dihybrid cross showing reciprocal recessive epistasis



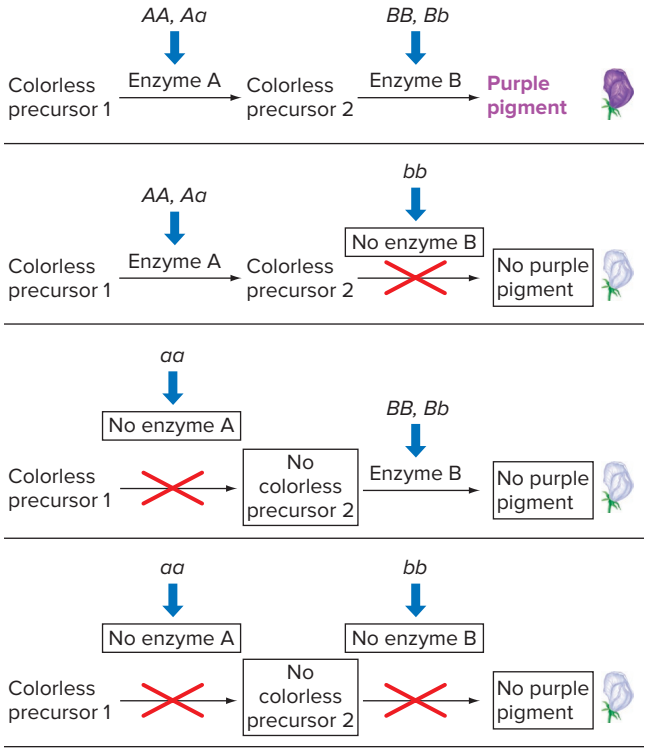
genotype *hh* for the second gene, do not make substance H at all. Thus, even if these people make an enzyme that would add A or B to this polysaccharide base, they have nothing to add it onto; as a result, Bombay-phenotype individuals appear to be type O. For this reason, homozygosity for the recessive *h* allele of the H-substance gene masks the effects of the *ABO* gene, making the *hh* genotype epistatic to any combination of I^A, I^B, and *i* alleles (except for *ii*.).

A person who carries I^A, I^B, or both I^A and I^B but is also an *hh* homozygote for the H-substance gene may appear to be type O, but he or she will be able to pass along an I^A or I^B allele in sperm or egg. An offspring receiving, let's say, an I^A allele for the ABO gene and a recessive *h* allele for the H-substance gene from the father plus an *i* allele and a dominant *H* allele from the mother would have blood type A (genotype $I^A i\ Hh$), even though neither of the parents is phenotype A or AB (**Fig. 3.14b**).

White sweet pea flowers In the first decade of the twentieth century, William Bateson conducted a cross between two lines of pure-breeding white-flowered sweet peas (**Fig. 3.15a**). Quite unexpectedly, all of the F_1 progeny were purple (**Fig. 3.15b**). Self-pollination of these novel hybrids produced a ratio of 9 purple: 7 white in the F_2 generation. The explanation? Two genes work in tandem to produce purple sweet pea flowers, and a dominant allele of each gene must be present to produce that color.

A simple biochemical hypothesis for these results is shown in **Fig. 3.16.** Because it takes two enzymes catalyzing

Figure 3.16 A biochemical explanation for reciprocal recessive epistasis in the generation of sweet pea color. Enzymes specified by the dominant alleles of two genes are both necessary to produce pigment. The recessive alleles of both genes specify no enzymes. In *aa* homozygotes, no intermediate precursor 2 is generated, so even if enzyme B is available, it cannot produce purple pigment.



sequential biochemical reactions to change a colorless precursor into a purple pigment, only the *A– B–* genotypic class, which produces active forms of both required enzymes, can generate colored flowers. The other three genotypic classes (*A– bb, aa B–,* and *aa bb*) become grouped together with respect to phenotype because they do not specify functional forms of one or the other requisite enzyme and thus give rise to no color, which is the same as white. It is easy to see how the 7 part of the 9:7 ratio encompasses the 3:3:1 of the 9:3:3:1 F_2 ratio.

The 9:7 ratio is the phenotypic signature of this type of **reciprocal recessive epistasis** in which the dominant alleles of two genes acting together (*A– B–*) produce color or some other trait, while the other three genotypic classes (*A– bb, aa B–,* and *aa bb*) do not (see **Fig. 3.15b**). Given that the phenotype associated with either allele *A* or allele *B* is purple, then we can say that *aa* is epistatic to *B*, and *bb* is epistatic to *A*. If the sweet peas are either *aa* or *bb*, their flowers will be white regardless of whether or not they have a dominant allele of the other gene.

Dominant epistasis

Epistasis can also be caused by a dominant allele. Depending on the details of the biochemical pathway involved, dominant epistasis can result in either of two different phenotypic ratios.

Squash fruit color In summer squash, two genes influence the color of the fruit (**Fig. 3.17a**). With one gene, the dominant allele (*A–*) determines yellow, while homozygotes for the recessive allele (*aa*) are green. A second gene's dominant allele (*B–*) produces white, while *bb* fruit may be either yellow or green, depending on the genotype of the first gene. In the interaction between these two genes, the presence of *B* hides the effects of either *A–* or *aa,* producing white fruit, and *B–* is thus epistatic to any genotype of the *A* gene. The recessive *b* allele has no effect on fruit color determined by gene *A*. Epistasis in which the dominant allele of one gene hides the effects of another gene is called **dominant epistasis.** In a cross between white F_1 dihybrids (*Aa Bb*), the F_2 phenotypic ratio is 12 white : 3 yellow : 1 green (Fig. 3.17a). The 12 includes two genotypic classes: 9 *A– B–* and 3 *aa B–*.

Figure 3.17 Dominant epistasis may result in a 12:3:1 phenotypic ratio. **(a)** In summer squash, the dominant *B* allele causes white color and is sufficient to mask the effects of any combination of *A* and *a* alleles. As a result, yellow (*A–*) or green (*aa*) color is expressed only in *bb* individuals. **(b)** The *A* allele encodes enzyme A, while the *a* allele specifies no enzyme. Therefore, yellow pigment is present in *A–* squash and green pigment in *aa* squash. Deposition of either pigment depends on protein b encoded by allele *b*, the normal (wild-type) allele of a second gene. However, the mutant dominant allele *B* encodes an abnormal version B of this protein that prevents pigment deposition, even when the normal protein b is present. Therefore, in order to be colored, the squash must have protein b but not protein B (genotype *bb*).

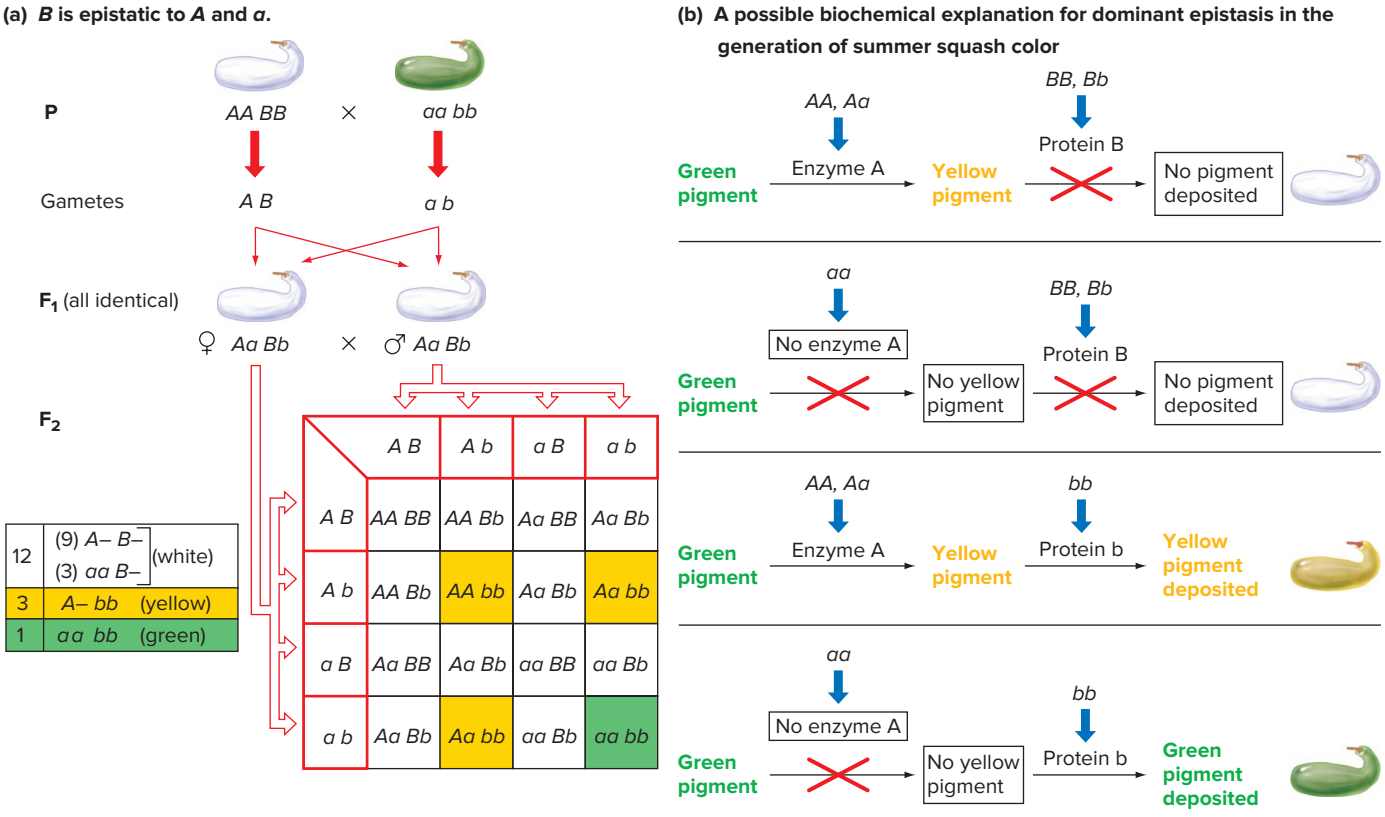
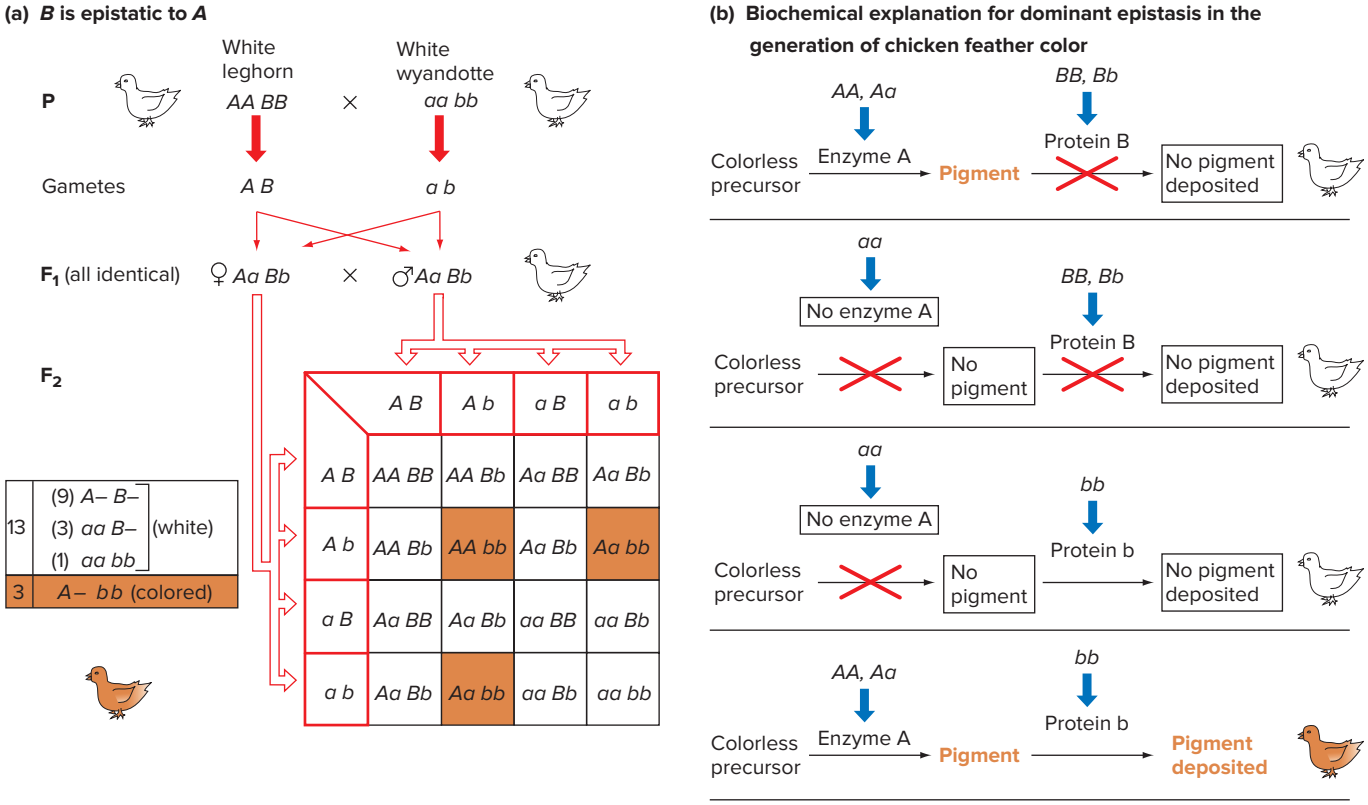


Figure 3.18 Dominant epistasis may also result in a 13:3 phenotypic ratio. **(a)** In the F_2 generation resulting from a dihybrid cross between white leghorn and white wyandotte chickens, the ratio of white birds to birds with color is 13:3. This ratio emerges because at least one copy of *A* and the absence of *B* is needed to produce color. **(b)** Enzyme A, encoded by allele *A*, is needed to synthesize pigment. Allele *a* encodes no enzyme. Pigment deposition in the feathers depends on protein b encoded by allele *b*, the normal (wild-type) allele of a second gene. The mutant dominant allele *B*, however, encodes an abnormal version of the protein that prevents pigment deposition, even when the normal protein b is present.



The squash genes *A* and *B* have not been identified at the molecular level, and the biochemical pathway in which they interact is unknown. However, based on knowledge of similar phenomena in other plants, a likely biochemical pathway underlying the 12:3:1 phenotypic ratio is shown in **Fig. 3.17b.**

Chicken feather color A variant ratio indicating dominant epistasis is seen in the feather color of certain chickens (**Fig. 3.18a**). White leghorns have a doubly dominant *AA BB* genotype for feather color; white wyandottes are homozygous recessive for both genes (*aa bb*). A cross between these two pure-breeding white strains produces an all-white dihybrid (*Aa Bb*) F_1 generation, but birds with color in their feathers appear in the F_2, and the ratio of white to colorful is 13:3 (Fig. 3.18a). We can explain this ratio by assuming a kind of dominant epistasis in which *B* is epistatic to *A;* the *A* allele produces color only in the absence of *B;* and the *a, B,* and *b* alleles produce no color. The interaction is characterized by a 13:3 ratio because the 9 *A– B–,* 3 *aa B–,* and 1 *aa bb* genotypic classes combine to produce only one phenotype: white. The biochemical pathway known to underly the 13:3 ratio for chicken feather color is shown in **Fig. 3.18b.**

Important points regarding epistasis

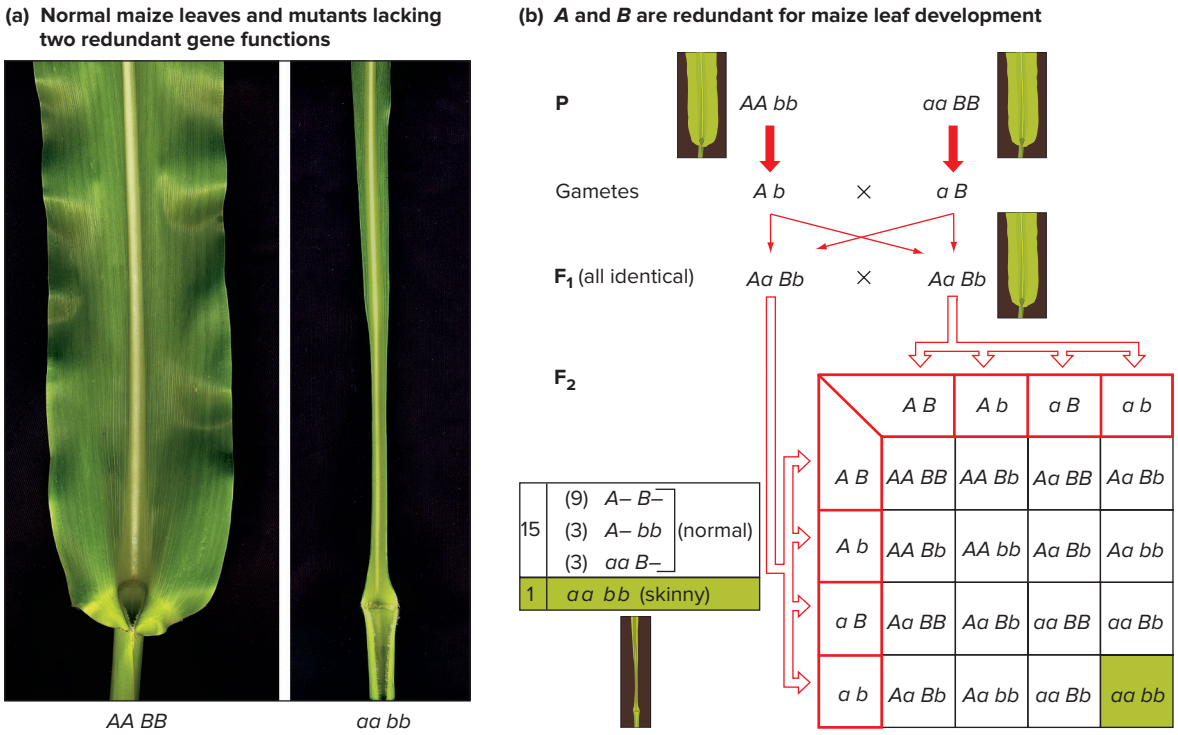
Several important points emerge from the examples of recessive and dominant epistasis we have discussed:

- Epistasis is an interaction between alleles of different genes, not between alleles of the same gene.
- In dihybrid crosses, the F_2 phenotypic ratios resulting from epistasis depend on the functions of the specific alleles and the particular biochemical pathways in which the genes participate.

In the Labrador retriever and sweet pea examples of recessive epistasis, the completely dominant alleles of both genes specify normally functional protein, while the recessive alleles are either nonfunctional or specify weakly functional protein. Nevertheless, the phenotypic ratios among the F_2 of a dihybrid cross differ in the Labradors and peas because the underlying biochemical pathways are not identical. Likewise, the two dominant epistasis

Figure 3.19 Redundant genes result in a 15:1 phenotypic ratio. **(a)** Normal maize leaves (*AA BB*) and a leaf lacking both dominant alleles *A* and *B* (*aa bb*). **(b)** In maize, either dominant allele *A* or *B* is sufficient for normal leaf development. Only the absence of both dominant alleles (*aa bb*) results in malformed, thin leaves. The result is a 15:1 ratio in the dihybrid cross F_2.

a: © Dr. Michael J. Scanlon, Cornell University



examples (squash and chicken colors) yield different F_2 phenotypic ratios because of differences in the biochemical pathways involved.

- Recessive epistasis usually indicates that the dominant alleles of the two genes function in the same pathway to achieve a common outcome. In the Labrador retriever example, *B* and *E* both function to generate black hairs.
- Dominant epistasis usually indicates that the dominant alleles of the two genes have antagonistic functions. Both in the cases of squash and chicken color, the dominant allele of gene *B* prevents deposition of a pigment whose synthesis depends on the dominant allele of gene *A*.

Redundancy: One or More Genes in a Pathway Are Superfluous

In maize, two genes, *A* and *B,* control leaf development. Normal broad leaves develop as long as the plant has either a dominant *A* allele or a dominant *B* allele (*A*– *B*–, *A*– *bb*, or *aa B*–). However, the leaves of plants that have neither dominant allele (*aa bb*) are skinny because they contain too few cells (**Fig. 3.19a**). Given that leaves are malformed only in the absence of both *A* and *B* (*aa bb*), the F_2 phenotypic ratio signifying **redundant gene action** is 15:1 (**Fig. 3.19b**).

The proteins (A and B) encoded by the dominant alleles act in parallel, redundant pathways that recruit precursor cells to become part of the leaf (**Fig. 3.20**). That is, if either pathway functions, the leaves will develop their normal broad shape.

Often, as in this case, redundant genes specify nearly identical proteins that perform the same function. Why does the organism have two genes that do the same thing? One answer is that redundant genes often arise by chance evolutionary processes that duplicate genes, as will be explained in Chapter 10.

Summary: A Variety of Different Biochemical Pathways Can Produce Any Given Altered Mendelian Ratio

So far we have seen that when two independently assorting genes interact to determine a trait, the 9:3:3:1 ratio of the four Mendelian genotypic classes in the F_2 generation can produce a variety of phenotypic ratios, depending on the nature of the gene interactions. The result may be four, three, or two phenotypes, composed of different combinations of

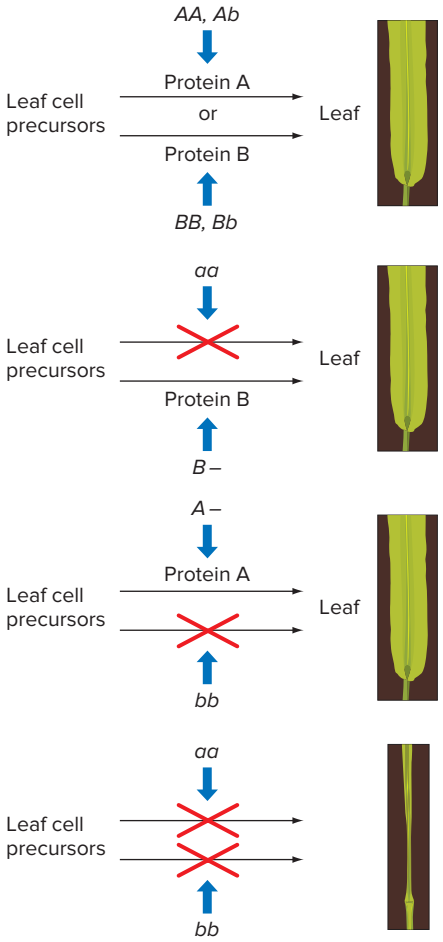


Figure 3.20 A biochemical explanation for redundant gene action. The dominant alleles *A* and *B* specify proteins that function in independent pathways to instruct cells to become part of the leaf. The recessive alleles *a* and *b* specify no proteins. Because either pathway is sufficient, only plants that lack both dominant alleles have thin leaves.

the four genotypic classes. **Table 3.2** summarizes some of the possibilities, correlating the phenotypic ratios with the genetic phenomena they reflect.

It is important to appreciate that wild-type and mutant alleles of genes participating in many different types of biochemical pathways may produce any specific F_2 phenotypic ratio shown in Table 3.2, such as 9:7 or 12:3:1. Thus, if you observe a certain ratio in a cross, you cannot infer the underlying pathway, although you can exclude some possibilities. On the other hand, as you will see in the problems at the end of this chapter, if you know the pathway's biochemistry, you can predict accurately the phenotypic ratios among the progeny of a cross involving the genes that determine the trait.

Incomplete Dominance or Codominance Can Expand Phenotypic Variation

We have identified to this point several variations on the theme of two-gene inheritance:

- alleles of different genes can interact *additively* to generate novel phenotypes;
- one gene's alleles can mask the effects of alleles at another gene (*epistasis*);
- different genes may have *redundant* functions so that a dominant allele of either gene is sufficient for the production of a particular normal phenotype.

All but the first of these interactions between different genes resulted in the merging of two or more of Mendel's four genotypic classes into one phenotypic class. For example, when genes are redundant, *A– B–*, *A– bb*, and *aa B–* have the same phenotype. In examining each of these

TABLE 3.2 Summary of Two-Gene Interactions

Gene Interaction	Example	F_2 Genotypic Ratios from an F_1 Dihybrid Cross				F_2 Phenotypic Ratio
		A– B–	*A– bb*	*aa B–*	*aa bb*	
Additive: Four distinct F_2 phenotypes	Lentil: seed coat color (see Fig. 3.10a)	9	3	3	1	9:3:3:1
Recessive epistasis: When homozygous, recessive allele of one gene masks both alleles of another gene	Labrador retriever: coat color (see Fig. 3.12b)	9	3	3	1	9:3:4
Reciprocal recessive epistasis: When homozygous, recessive allele of each gene masks the dominant allele of the other gene	Sweet pea: flower color (see Fig. 3.15b)	9	3	3	1	9:7
Dominant epistasis I: Dominant allele of one gene hides effects of both alleles of the other gene	Summer squash: color (see Fig. 3.17a)	9	3	3	1	12:3:1
Dominant epistasis II: Dominant allele of one gene hides effects of dominant allele of other gene	Chicken feathers: color (see Fig. 3.18a)	9	3	3	1	13:3
Redundancy: Only one dominant allele of either of two genes is necessary to produce phenotype	Maize: leaf development (see Fig. 3.19b)	9	3	3	1	15:1

categories, for the sake of simplicity, we have looked at examples in which one allele of each gene in a pair showed complete dominance over the other. But for any type of gene interaction, the alleles of one or both genes may exhibit incomplete dominance or codominance, and these possibilities increase the potential for phenotypic diversity. For example, **Fig. 3.21** shows how incomplete dominance at both genes in a dihybrid cross results not in a collapse of several genotypic classes into one but rather an expansion—each of the nine genotypes in the dihybrid cross F_2 corresponds to a different phenotype.

A simple biochemical explanation for the phenotypes in Fig. 3.21 is similar to that for incomplete dominance in Fig. 3.3b, where the amount of red pigment produced was proportional to the amount of an enzyme. The difference here is that purple pigmentation requires the action of two enzymes, A and B, and one is more efficient than the other, resulting in one gene (in this case, the *A* gene) contributing more to the purple phenotype than the other gene.

Figure 3.21 With incomplete dominance, the interaction of two genes can produce nine different phenotypes for a single trait. In this example, two genes produce purple pigments. Alleles A^1 and A^2 of the first gene exhibit incomplete dominance, as do alleles B^1 and B^2 of the second gene. The two alleles of each gene can generate three different phenotypes, so double heterozygotes can produce nine (3 × 3) different colors in a ratio of 1:2:2:1:4:1:2:2:1.

♀ $A^1A^2\ B^1B^2$ × ♂ $A^1A^2\ B^1B^2$

F_1 (all identical)

F_2

	$A^1\ B^1$	$A^1\ B^2$	$A^2\ B^1$	$A^2\ B^2$
$A^1\ B^1$	$A^1A^1\ B^1B^1$	$A^1A^1\ B^1B^2$	$A^1A^2\ B^1B^1$	$A^1A^2\ B^1B^2$
$A^1\ B^2$	$A^1A^1\ B^1B^2$	$A^1A^1\ B^2B^2$	$A^1A^2\ B^1B^2$	$A^1A^2\ B^2B^2$
$A^2\ B^1$	$A^1A^2\ B^1B^1$	$A^1A^2\ B^1B^2$	$A^2A^2\ B^1B^1$	$A^2A^2\ B^1B^2$
$A^2\ B^2$	$A^1A^2\ B^1B^2$	$A^1A^2\ B^2B^2$	$A^2A^2\ B^1B^2$	$A^2A^2\ B^2B^2$

1	$A^1A^1\ B^1B^1$	purple shade 9
2	$A^1A^1\ B^1B^2$	purple shade 8
2	$A^1A^2\ B^1B^1$	purple shade 7
1	$A^1A^1\ B^2B^2$	purple shade 6
4	$A^1A^2\ B^1B^2$	purple shade 5
1	$A^2A^2\ B^1B^1$	purple shade 4
2	$A^1A^2\ B^2B^2$	purple shade 3
2	$A^2A^2\ B^1B^2$	purple shade 2
1	$A^2A^2\ B^2B^2$	purple shade 1 (white)

Although the possibilities for variation are manifold, none of the observed departures from Mendelian phenotypic ratios contradicts Mendel's genetic laws of segregation and independent assortment. The alleles of each gene still segregate as he proposed. Interactions between the alleles of many genes simply make it harder to unravel the complex relation of genotype to phenotype.

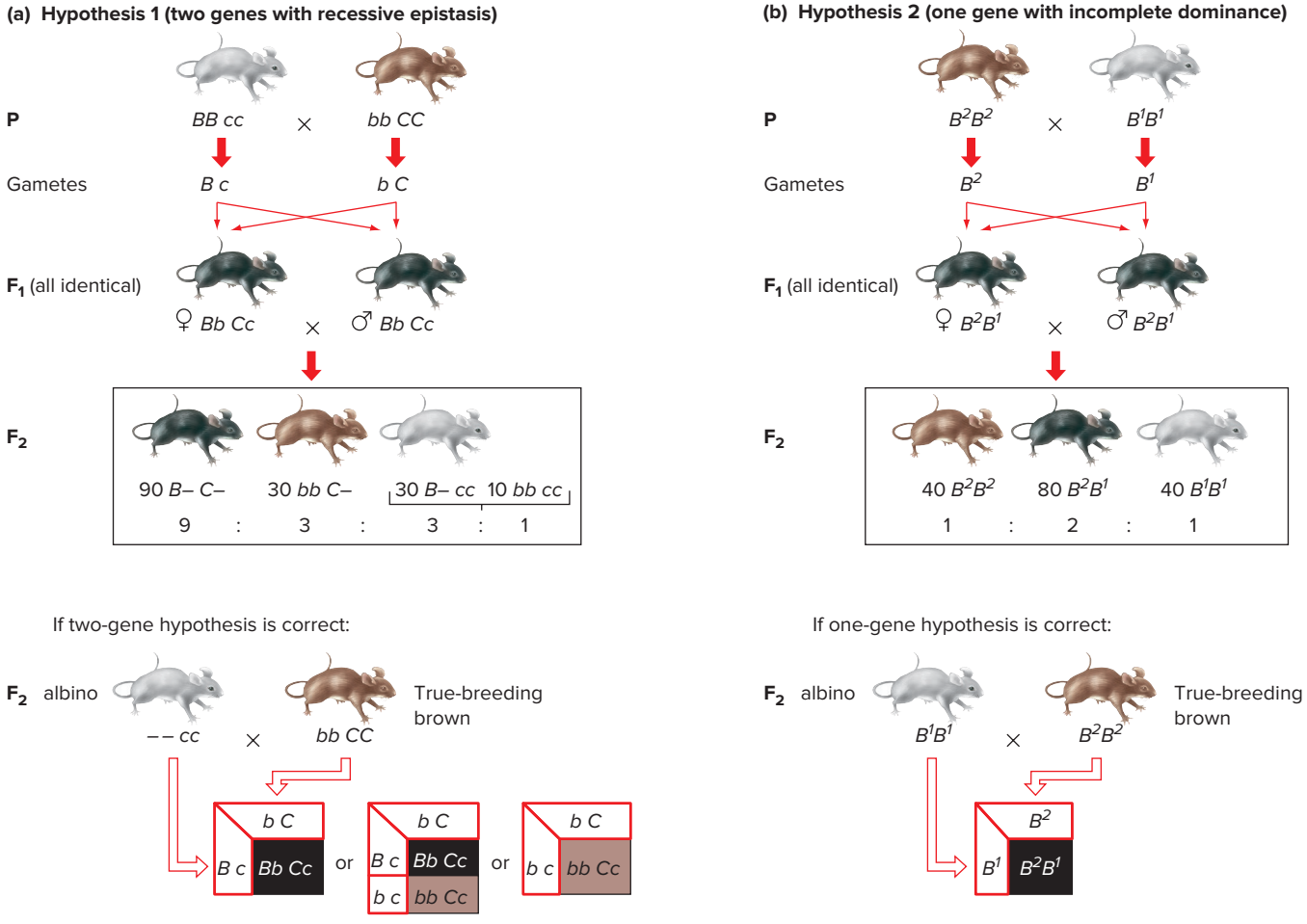
Breeding Studies Help Geneticists Determine Whether One or Two Genes Determine a Trait

How do geneticists know whether a particular trait is caused by the alleles of one gene or by two genes interacting in one of a number of possible ways? Breeding tests can usually resolve the issue. Phenotypic ratios diagnostic of a particular mode of inheritance (for instance, the 9:7 or 13:3 ratios indicating that two genes are interacting) can provide the first clues and suggest hypotheses. Further breeding studies can then show which hypothesis is correct.

As an example, a mating of one strain of pure-breeding white albino mice with pure-breeding brown results in black hybrids; and a cross between the black F_1 hybrids produces 90 black, 30 brown, and 40 albino offspring. What is the genetic constitution of these phenotypes? We could assume that we are seeing the 9:3:4 ratio of recessive epistasis and hypothesize that two interacting genes (call them *B* and *C*) control color. In this model, each gene has completely dominant and recessive alleles, and the homozygous recessive of one gene is epistatic to both alleles of the other gene (**Fig. 3.22a**). This idea makes sense, but it is not the only hypothesis consistent with the data.

We might also explain the data—160 progeny in a ratio of 90:30:40—by the activity of one gene (**Fig. 3.22b**). According to this one-gene hypothesis, albinos would be homozygotes for one allele (B^1B^1), brown mice would be homozygotes for a second allele (B^2B^2), and black mice would be heterozygotes (B^1B^2) that have their own novel phenotype because B^1 and B^2 are incompletely dominant. Under this system, a mating of black (B^1B^2) to black (B^1B^2) would be expected to produce 1 B^2B^2 brown : 2 B^1B^1 black : 1 B^1B^1 albino, or 40 brown : 80 black : 40 albino. Is it possible that the 30 brown, 90 black, and 40 albino mice actually counted were obtained from the inheritance of a single gene? Intuitively, the answer is yes because the ratios 40:80:40 and 30:90:40 do not seem that different. We know that if we flip a coin 100 times, it doesn't always come up 50 heads : 50 tails; sometimes it's 60:40 just by chance. So, how can we decide between the two-gene and the one-gene model?

Figure 3.22 Specific breeding tests can help decide between hypotheses. Either of two models could explain the results of a cross tracking coat color in mice. **(a)** In one hypothesis, two genes interact with recessive epistasis to produce a 9:3:4 ratio. **(b)** In the other hypothesis, a single gene with incomplete dominance between the alleles generates the observed results. One way to decide between these models is to cross each of several albino F_2 mice with true-breeding brown mice. The two-gene model predicts several different outcomes depending on the *– – cc* albino's genotype at the *B* gene. The one-gene model predicts that all progeny of all the crosses will be black.



The answer is that we can use other types of crosses to verify or refute the hypotheses. For instance, if the one-gene hypothesis were correct, a mating of F_2 albinos with pure-breeding brown mice similar to those of the parental generation would produce all black heterozygotes [brown (*BB*) × albino (*bb*) = all black (*Bb*)] (Fig. 3.22b). But if the two-gene hypothesis is correct, with recessive mutations at an albino gene (called *C*) epistatic to all expression from the *B* gene, different matings of pure-breeding brown (*bb CC*) with the F_2 albinos (*– – cc*) will give different results—all progeny are black; half are black and half brown; all are brown—depending on the albino's genotype at the *B* gene (see Fig. 3.22a). In fact, when the experiment is actually performed, the diversity of results confirms the two-gene hypothesis.

Locus Heterogeneity: Mutations in Any One of Several Genes May Cause the Same Phenotype

Close to 50 different genes have mutant alleles that can cause deafness in humans. Many genes generate the developmental pathway that brings about hearing, and a loss of function in any part of the pathway, for instance, in one small bone of the middle ear, can result in deafness. In other words, it takes a dominant wild-type allele at each of these 50 genes to produce normal hearing. Thus, deafness is a **heterogeneous trait:** A mutation at any one of a number of genes can give rise to the same phenotype. We saw earlier (Fig. 3.15b) that whiteness of sweet pea flowers is also a heterogeneous trait; *AA bb* and *aa BB* flowers,

each homozygous for recessive, nonfunctional alleles of different genes, were both white.

Evidence for locus heterogeneity in human pedigrees

Careful examination of many family pedigrees can reveal whether **locus heterogeneity**—a property of a trait where mutations in any one of two or more genes results in the same mutant phenotype—explains the inheritance pattern of a trait. In the case of deafness, for example, whether a particular nonhearing man and a particular nonhearing woman carry mutations in the same gene or different ones can be determined if they have children together. If they have only children who can hear, the parents most likely carry mutations at two different genes, and the children carry one normal, wild-type allele for both of those genes (**Fig. 3.23a**). By contrast, if all of their children are deaf, it is likely that both parents are homozygous for a mutation in the same gene, and all of their children are also homozygous for this same mutation (**Fig. 3.23b**).

Complementation and complementation tests

The method outlined in Fig. 3.23 for discovering whether a particular phenotype arises from mutations in the same or separate genes is a naturally occurring version of an experimental genetic tool called the **complementation test.** Simply put, when what appears to be an identical *recessive* phenotype arises in two separate breeding lines, geneticists want to know whether mutations in the same gene are responsible for the phenotype in both lines. They answer this question by setting up a mating between affected individuals from the two lines. If offspring receiving the two mutations—one from each parent—express the wild-type phenotype, **complementation** has occurred. The observation of complementation means that the original mutations affected two different genes, and for both genes, the normal allele from one parent can provide what the mutant allele of the same gene from the other parent cannot. Note that a finding of complementation implies that the trait in question must be heterogeneous.

You previously saw an example of complementation in Fig. 3.15b. There, the white parental plants were homozygous for nonfunctional alleles of different genes required for purple pigment synthesis. The F_1 were purple because the gamete of each parent provided the wild-type allele that the other lacked. The pedigree for deafness in Fig. 3.23a, in which all the children of two deaf parents had normal hearing, provides another example of this same phenomenon. By contrast, if all the offspring of affected parents express the mutant phenotype, no complementation has occurred. Each offspring received two recessive mutant alleles—one from each parent—of the same gene (Fig. 3.23b). A lack of complementation does not exclude the possibility that a trait could be heterogeneous, but instead it simply indicates that the parents involved in the particular cross had mutant alleles of the same gene.

You can quiz your understanding of the related concepts of locus heterogeneity and complementation by considering a form of albinism known as ocular-cutaneous albinism (OCA). People with this inherited condition have little or no pigment in their skin, hair, and eyes (**Fig. 3.24a**). The horizontal inheritance pattern seen in **Fig. 3.24b** suggests that OCA is determined by the recessive allele of one gene, with albino family members being homozygotes for that allele. But a 1952 paper on albinism reported a family in which two albino parents produced three normally pigmented children (**Fig. 3.24c**). How would you explain this phenomenon?

The answer is that albinism is another example of locus heterogeneity: Mutant alleles at any one of several different genes can cause the condition. The reported mating was, in effect, an inadvertent complementation test. The complementation observed showed that one albino parent was homozygous for an OCA-causing mutation in gene *A*, while the other albino parent was homozygous for an OCA-causing mutation in a different gene, *B*.

Figure 3.23 Locus heterogeneity in humans: Mutations in many genes can cause deafness. **(a)** Two deaf parents can have hearing offspring if the mother and father are homozygous for recessive mutations in different genes. **(b)** Two deaf parents with mutations in the same gene may produce all deaf children.

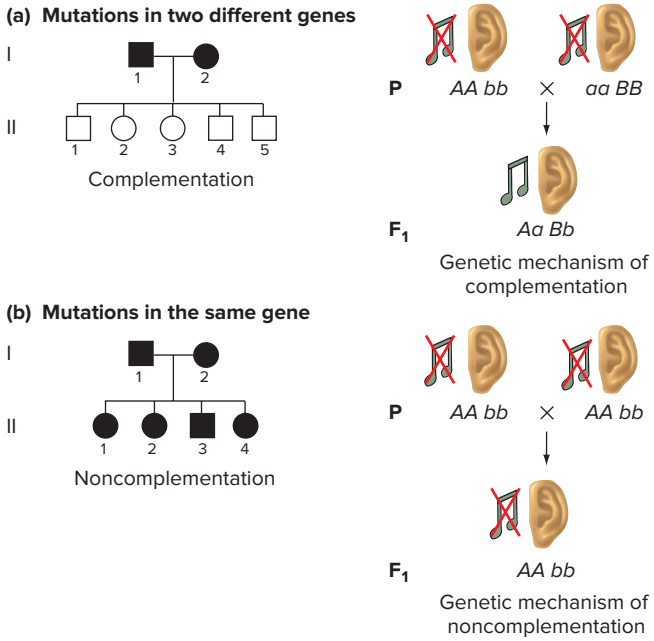
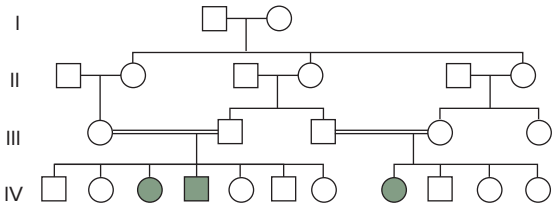


Figure 3.24 Family pedigrees help unravel the genetic basis of ocular-cutaneous albinism (OCA). **(a)** An albino Nigerian girl and her sister celebrating the conclusion of the All Africa games. **(b)** A pedigree following the inheritance of OCA in an inbred family indicates that the trait is recessive. **(c)** A family in which two albino parents have nonalbino children demonstrates that homozygosity for a recessive allele of either of two genes can cause OCA.

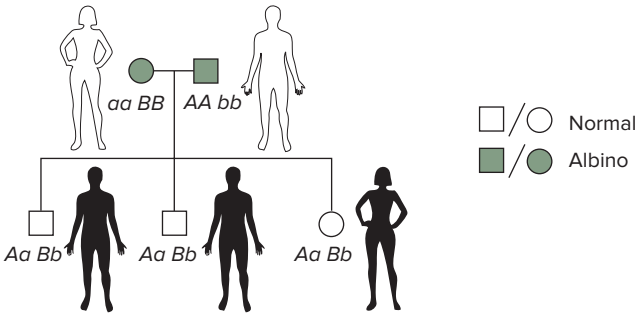
a: © Radu Sigheti/Reuters

(a) Ocular-cutaneous albinism (OCA)

(b) OCA is recessive



(c) Complementation for albinism



essential concepts

- Two genes may interact to affect a single trait; these interactions may be detected by ratios that can be predicted from Mendelian principles.
- Retention of the 9:3:3:1 phenotypic ratio usually indicates that two genes function in independent pathways and their alleles interact *additively.*
- In *epistasis,* an allele at one gene can hide traits otherwise caused by alleles at another gene.
- When genes display *redundancy* for a trait, one dominant and normally functioning allele of either gene is sufficient to generate the normal phenotype.
- Many traits exhibit *locus heterogeneity,* in which homozygosity for mutations at any one of several genes can produce the same mutant phenotype.
- *Complementation* occurs in the progeny of pure-breeding parents with the same mutant phenotype if the parents are homozygous for recessive, nonfunctional alleles of different genes whose products function in a common pathway.

3.3 Extensions to Mendel for Multifactorial Inheritance

learning objectives

1. Discuss the factors that can cause different individuals with the same genotype to be phenotypically dissimilar.
2. Explain how Mendelian genetics is compatible with the fact that many traits, such as human height and skin colors, exhibit continuous variation.

The inheritance of many traits appears to be more complex than can be explained by the participation of only one or two genes in patterns compatible with straightforward Mendelian principles. Of course, one reason for this complexity is that more than two genes can influence certain traits. But a second reason is that genes are not the only players: The environment and chance events can sometimes exert considerable effects on traits that are otherwise genetically determined. In this section we discuss **multifactorial traits**—traits determined by several different genes, or by the interaction of genes with the environment.

The Same Genotype Does Not Always Produce the Same Phenotype

In our discussion of gene interactions so far, we have considered examples in which a genotype reliably fashions a particular phenotype. But this is not always what happens. Sometimes a genotype is not expressed at all; that is, even though the genotype is present, the expected phenotype does not appear. Other times, the trait caused by a genotype is expressed to varying degrees or in a variety of ways in different individuals. Factors that alter the phenotypic expression of genotype include *modifier genes,* the environment, and chance. These factors complicate the interpretation of breeding experiments.

Penetrance and expressivity

Retinoblastoma, the most malignant form of eye cancer, arises from a dominant mutation in one gene, but only about 75% of people who carry the mutant allele develop the disease. Geneticists use the term **penetrance** to describe the proportion of individuals with a particular genotype who show the expected phenotype. Penetrance can be *complete* (100%), as in the traits that Mendel studied, or *incomplete,* as in retinoblastoma. For retinoblastoma, the penetrance is ~75%.

In some people with retinoblastoma, only one eye is affected, while in other individuals with the phenotype, both eyes are diseased. **Expressivity** refers to the degree or intensity with which a particular genotype is expressed in a phenotype. Expressivity can be *variable,* as in retinoblastoma (one or both eyes affected), or *unvarying,* as in pea color (all *yy* peas are green). As we will see, the incomplete penetrance and variable expressivity of retinoblastoma are mainly the result of chance, but in other cases, it is other genes and/or the environment that cause variations in phenotype. **Figure 3.25** summarizes in graphic form the differences between complete penetrance, incomplete penetrance, variable expressivity, and unvarying expressivity.

Modifier genes

Not all genes that influence the appearance of a trait contribute equally to the phenotype. Major genes have a large influence, while **modifier genes** have a more subtle, secondary effect. Modifier genes alter the phenotypes produced by the alleles of other genes. No formal distinction exists between major and modifier genes. Rather, a continuum exists between the two, and the cutoff is arbitrary. Scientists sometimes call the set of unknown modifier genes that influence the action of known genes the **genetic background.**

Modifier genes influence the length of a mouse's tail. The mutant *T* allele of the tail-length gene causes a shortening of the normally long wild-type tail. But not all mice carrying the *T* mutation have the same length tail. A comparison of several inbred lines points to modifier genes as the cause of this variable expressivity. In one inbred line, mice carrying the *T* mutation have tails that are approximately 75% as long as normal tails; in another inbred line, the tails are 50% normal length; and in a third line, the tails are only 10% as long as wild-type tails. Because all members of each inbred line grow the same length tail, no matter what the environment (for example, diet, cage temperature, or bedding), geneticists conclude it is genes and not the environment or chance that determine the length of a mutant mouse's tail. Different inbred lines most likely carry different alleles of the modifier genes that determine exactly how short the tail will be when the *T* mutation is present; that is, these lines have different genetic backgrounds.

Environmental effects on phenotype

Temperature is one element of the environment that can have a visible effect on phenotype. For example, temperature influences the unique coat color pattern of Siamese cats (**Fig. 3.26**). These domestic felines are homozygous for one of the multiple alleles of a gene that encodes an enzyme catalyzing the production of the dark pigment melanin. The form of the enzyme generated by the variant *Siamese* allele does not function at the cat's normal core body temperature. It becomes active only at the lower temperatures found in the cat's extremities, where it promotes the production of melanin, which darkens the animal's ears, nose, paws, and tail. The enzyme is thus *temperature sensitive.* Under the normal environmental conditions in temperate climates, the Siamese phenotype does not vary much in expressivity from one cat to another, but one can imagine the expression of a very different phenotype—no dark extremities—in equatorial deserts, where the ambient temperature is at or above normal body temperature.

Temperature can also affect survivability. In one type of experimentally bred fruit fly (*Drosophila melanogaster*), some individuals develop and multiply normally at temperatures between 18°C and 29°C; but if the thermometer climbs beyond that cutoff for a short time, they become reversibly paralyzed; and if the temperature remains high for more than a few hours, they die. These insects carry a temperature-sensitive allele of the *shibire*

Figure 3.25 Phenotypes may show variations in penetrance and expressivity. A genotype is completely penetrant when all individuals with that genotype have the same phenotype (*green*). Some genotypes are incompletely penetrant—some individuals with the same genotype show the phenotype and others do not. Genotypes may also show variable expressivity, meaning that individuals with the same genotype may show the trait but to different degrees.

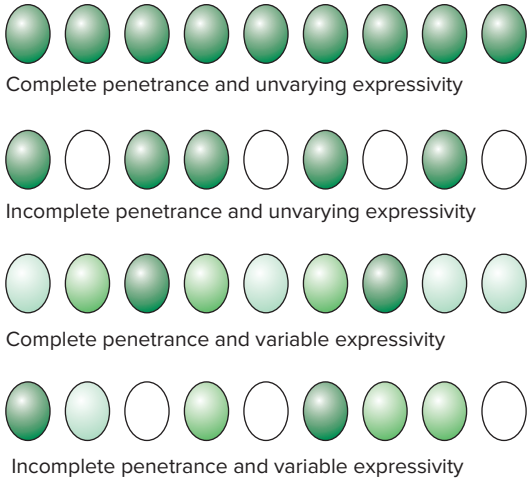
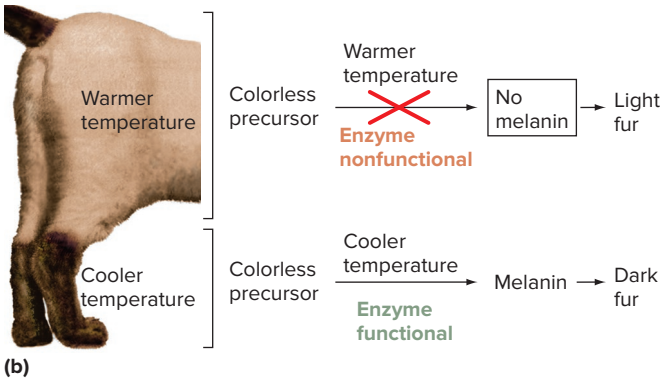


Figure 3.26 In Siamese cats, temperature affects coat color. **(a)** A Siamese cat. **(b)** Melanin is produced only in the cooler extremities. This phenomenon occurs because Siamese cats are homozygous for a mutation that renders an enzyme involved in melanin synthesis temperature sensitive. The mutant enzyme is active at lower temperatures but inactive at higher temperatures.

(a)



(b)

gene, which specifies a protein essential for nerve cell transmission. This type of allele is known as a **conditional lethal** because it is lethal only under certain conditions. The range of temperatures under which the insects remain viable are **permissive conditions;** the lethal temperatures above that are **restrictive conditions.** Thus, at one temperature, the allele gives rise to a phenotype that is indistinguishable from the wild type, while at another temperature, the same allele generates a mutant phenotype (in this case, lethality). Flies with the wild-type *shibire* allele are viable even at the higher temperatures. The fact that some mutations are lethal only under certain conditions clearly illustrates that the environment can affect the penetrance of a phenotype.

Even in genetically normal individuals, exposure to chemicals or other environmental agents can have phenotypic consequences that are similar to those caused by mutant alleles of specific genes. A change in phenotype arising in such a way is known as a **phenocopy.** By definition, phenocopies are not heritable because they do not result from a change in a gene. In humans, ingestion of the sedative thalidomide by pregnant women in the early 1960s produced a phenocopy of a rare dominant trait called *phocomelia.* By disrupting limb development in otherwise normal fetuses, the drug mimicked the effect of the phocomelia-causing mutation. When this problem became evident, thalidomide was withdrawn from the market.

Some types of environmental change may have a positive effect on an organism's survivability. In the following example, a straightforward application of medical science artificially reduces the penetrance of a mutant phenotype. Children born with the recessive trait known as *phenylketonuria,* or *PKU,* will develop a range of neurological problems, including convulsive seizures and mental impairment, unless they are put on a special diet. Homozygosity for the mutant PKU allele eliminates the activity of a gene encoding the enzyme phenylalanine hydroxylase. This enzyme normally converts the amino acid phenylalanine to the amino acid tyrosine. Absence of the enzyme causes a buildup of phenylalanine, and this buildup results in neurological problems. Today, a reliable blood test can detect the condition in newborns. Once a baby with PKU is identified, a protective diet that excludes phenylalanine is prescribed. The diet must also provide enough calories to prevent the infant's body from breaking down its own proteins, thereby releasing the damaging amino acid from within. Such dietary therapy—a simple change in the environment—now enables many PKU infants to develop into healthy adults.

Finally, two of the top killer diseases in the United States—cardiovascular disease and lung cancer—also illustrate how the environment can alter phenotype by influencing both expressivity and penetrance. People may inherit a propensity to heart disease, but the environmental factors of diet and exercise contribute to the occurrence (penetrance) and seriousness (expressivity) of their condition. Similarly, some people are born genetically prone to lung cancer, but whether or not they develop the disease (penetrance) is strongly determined by whether they choose to smoke.

Thus, various aspects of an organism's environment, including temperature, diet, and exercise, interact with its genotype to generate the phenotype, the ultimate combination of traits that determines what a plant or animal looks like and how it behaves.

The effects of random events on penetrance and expressivity

Whether a carrier of the retinoblastoma mutation described earlier develops the phenotype, and whether the cancer affects one or both eyes, depend on additional genetic events that occur at random. To produce retinoblastoma, these events must alter the second allele of the gene in specific body cells. Examples of random events that can trigger the onset of the disease include cosmic rays (to which humans are constantly exposed) that alter the genetic material in retinal cells or mistakes made during cell division in the retina. Chance events provide the second *hit*—a mutation in the second copy of the retinoblastoma gene—necessary to turn a normal retinal cell into a cancerous one. The phenotype of retinoblastoma thus results from a specific heritable mutation in a specific gene, but the incomplete penetrance and variable expressivity of the disease depend on random genetic events that affect the other allele in certain cells. The relationship between genotype and phenotype as it applies to cancer will be discussed fully in Chapter 20.

By contributing to incomplete penetrance and variable expressivity, modifier genes, the environment, and chance give rise to phenotypic variation. The probability of penetrance and the level of expressivity cannot be derived from the original Mendelian principles of segregation and independent assortment; they are established empirically by observation and counting.

Mendelian Principles Can Also Explain Continuous Variation

In Mendel's experiments, height in pea plants was determined by two segregating alleles of one gene (in the wild, plant height is determined by many genes, but in Mendel's inbred populations, the alleles of all but one of these genes were invariant). The phenotypes that resulted from these alternative alleles were clear-cut, either short or tall, and pea plant height was therefore known as a **discontinuous trait** (or **discrete trait**). In contrast, because human populations are not inbred, height in people is determined by alleles of many different genes whose interaction with each other and the environment produces continuous phenotypic variation; height in humans is thus an example of a **continuous trait** (or **quantitative trait**). Within human populations, individual heights vary over a range of values that when charted on a graph produce a bell curve (**Fig. 3.27a**). In fact, many human traits, including height, weight, and skin color, show continuous variation, rather than the clear-cut alternatives analyzed by Mendel.

Continuous traits often appear to blend and unblend. Think for a moment of skin color. Children of marriages between people of African and Northern European descent, for example, often seem to be a blend of their parents' skin colors. Progeny of these F_1 individuals produce offspring displaying a wide range of skin pigmentation; a few may be as light as the original Northern European parent, a few as dark as the original African parent, but most will fall in a range between the two (**Fig. 3.27b**). For such reasons, early human geneticists were slow to accept Mendelian analysis. Because they were working with *outbred populations* (populations in which individuals differ in alleles of many genes), these scientists found very few examples of either-or Mendelian traits in normal, healthy people.

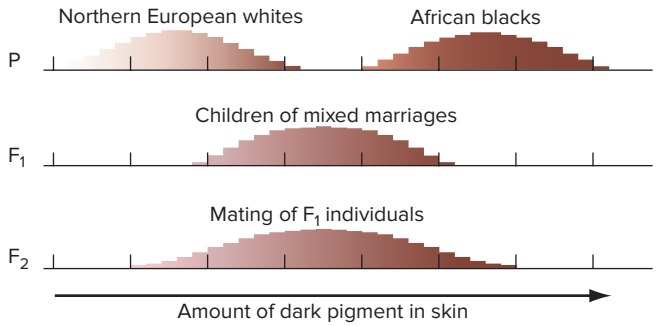
By 1930, however, studies of corn and tobacco demonstrated conclusively that it is possible to provide a Mendelian explanation of continuous variation by simply increasing the number of genes contributing to a phenotype. The more genes, the more phenotypic classes, and the more classes, the more the variation appears continuous.

As a hypothetical example, consider a series of genes (*A, B, C,* . . .) all affecting the height of pole beans. For

Figure 3.27 Continuous traits in humans. **(a)** Women runners at the start of a 5th Avenue mile race in New York City demonstrate that height is a trait showing continuous variation. **(b)** The skin color of most F_1 offspring is usually between the parental extremes, while the F_2 generation exhibits a broader distribution of continuous variation.

a: © Rudi Von Briel/PhotoEdit

(a)



(b)

GENETICS AND SOCIETY

Crowd: © Image Source/Getty Images RF

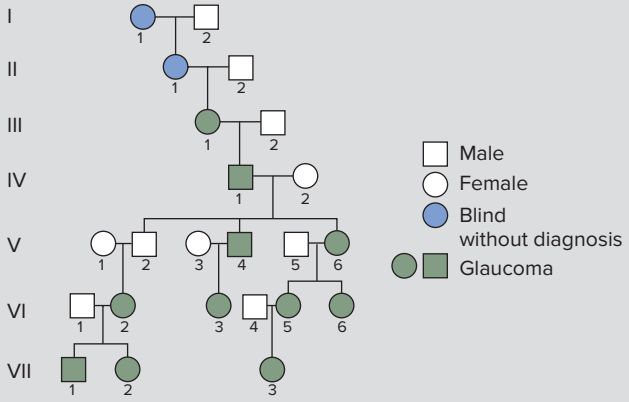
Disease Prevention Versus the Right to Privacy

In one of the most extensive human pedigrees ever assembled, a team of researchers traced a familial pattern of blindness back through five centuries of related individuals to its origin in a couple who died in a small town in northwestern France in 1495. More than 30,000 French men and women alive today descended from that one fifteenth-century couple, and within this direct lineage reside close to half of all reported French cases of hereditary juvenile glaucoma. The massive genealogical tree for the trait (when posted on the office wall, it was over 100 feet long) showed that the genetic defect follows a simple Mendelian pattern of transmission determined by the dominant allele of a single gene (**Fig. A**). The pedigree also showed that the dominant genetic defect displays incomplete penetrance. Not all people receiving the dominant allele become blind; these sighted carriers may unknowingly pass the blindness-causing dominant allele to their children.

Unfortunately, people do not know they have the disease until their vision starts to deteriorate. By that time, their optic fibers have sustained irreversible damage, and blindness is all but inevitable. Surprisingly, the existence of medical therapies that make it possible to arrest the nerve deterioration created a quandary in the late 1980s. Because effective treatment has to begin before symptoms of impending blindness show up, information in the pedigree could have helped doctors pinpoint people who are at risk, even if neither of their parents is blind. The researchers who compiled the massive family history therefore wanted to give physicians the names of at-risk individuals living in their area, so that doctors could monitor them and recommend treatment if needed. However, a long-standing French law protecting personal privacy forbids public circulation of the names in genetic pedigrees. The French government agency interpreting this law maintained that if the names in the glaucoma pedigree were made public, potential carriers of the disease might suffer discrimination in hiring or insurance.

France thus faced a serious ethical dilemma: On the one hand, giving out names could save perhaps thousands of people from blindness; on the other hand, laws designed to protect personal privacy precluded the dissemination of specific names. The solution adopted by the French government at the time was a massive educational program to alert the general public to the problem so that concerned families could seek medical advice. This approach addressed the legal issues but was only partially helpful in dealing with the medical problem, because many affected individuals escaped detection.

Figure A Pedigree showing the transmission of juvenile glaucoma. A small part of the genealogic tree: The vertical transmission pattern over seven generations shows that a dominant allele of a single gene causes juvenile glaucoma. The lack of glaucoma in V-2 followed by its reappearance in VI-2 reveals that the trait is incompletely penetrant.

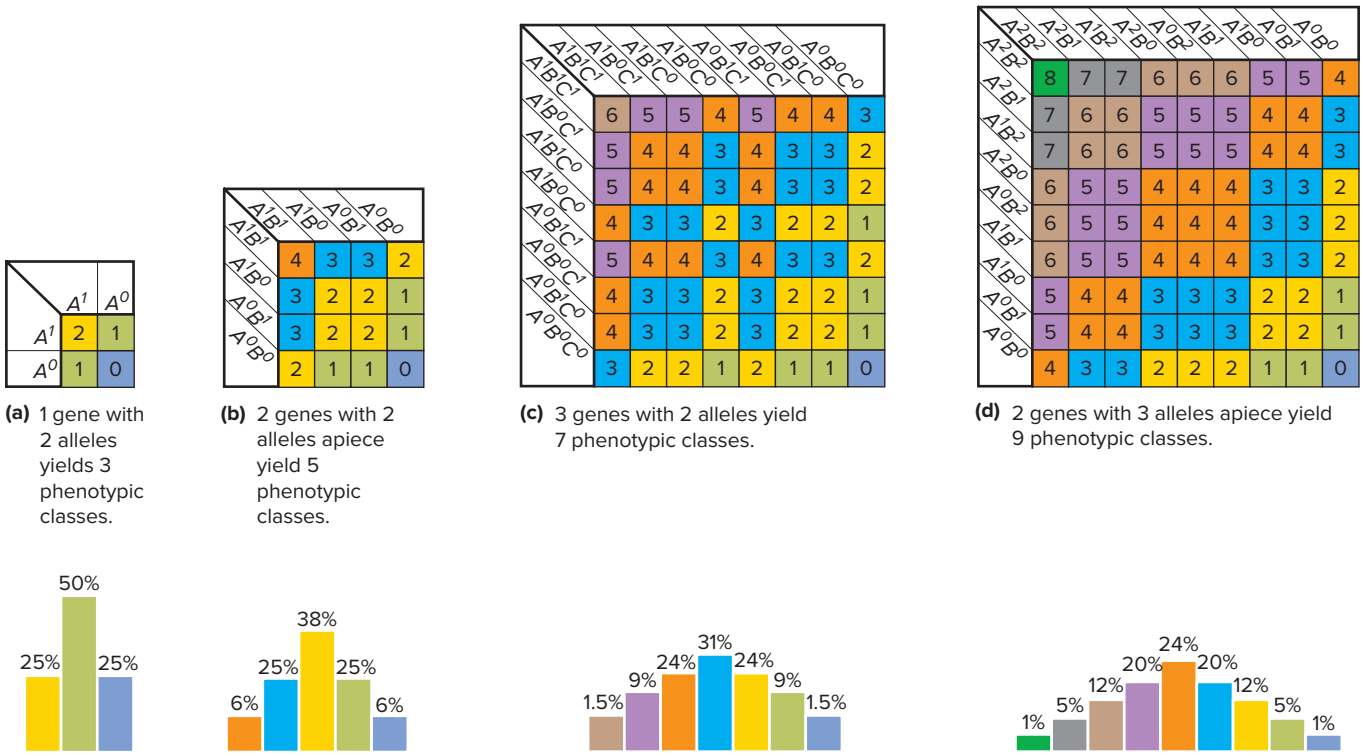


By 1997, molecular geneticists had identified the gene whose dominant mutant allele causes juvenile glaucoma. This gene specifies a protein called myocilin whose normal function in the eye is at present unknown. The mutant allele encodes a form of myocilin that folds incorrectly and then accumulates abnormally in the tiny canals through which eye fluid normally drains into the bloodstream. Misfolded myocilin blocks the outflow of excess vitreous humor, and the resulting increased pressure within the eye (glaucoma) eventually damages the optic nerve, leading to blindness.

Knowledge of the specific disease-causing mutations in the *myocilin* gene has more recently led to the development of diagnostic tests based on the direct analysis of genotype. (We describe methods for direct genotype analysis in Chapter 11.) These DNA-based tests can not only identify individuals at risk, but they can also improve disease management. Detection of the mutant allele before the optic nerve is permanently damaged allows for timely treatment. If these tests become sufficiently inexpensive in the future, they could resolve France's ethical dilemma. Doctors could routinely administer the tests to all newborns and immediately identify nearly all potentially affected children; private information in a pedigree would thus not be needed.

each gene, two alleles exist; a *0* allele that contributes nothing to height and a *1* allele that increases the height of a plant by one unit. All alleles exhibit incomplete dominance relative to alternative alleles of the same gene. The phenotypes determined by all these genes are additive. What would be the result of a two-generation cross between pure-breeding plants carrying only *0* alleles at each height gene and pure-breeding plants carrying only *1* alleles at each height gene? If only one gene were responsible for height, and if environmental effects could be discounted, the F_2 population would be distributed among three classes: homozygous A^0A^0 plants with 0 height (they

Figure 3.28 A Mendelian explanation of continuous variation. The more genes or alleles, the more possible phenotypic classes, and the greater the similarity to continuous variation. In these examples, several pairs of incompletely dominant alleles have additive effects. Percentages shown at the *bottom* denote frequencies of each genotype expressed as fractions of the total population.



(a) 1 gene with 2 alleles yields 3 phenotypic classes.

(b) 2 genes with 2 alleles apiece yield 5 phenotypic classes.

(c) 3 genes with 2 alleles yield 7 phenotypic classes.

(d) 2 genes with 3 alleles apiece yield 9 phenotypic classes.

lie prostrate on the ground); heterozygous A^0A^1 plants with a height of 1; and homozygous A^1A^1 plants with a height of 2 (**Fig. 3.28a**). This distribution of heights over three phenotypic classes does not make a continuous curve. But for two genes, five phenotypic classes will appear in the F_2 generation (**Fig. 3.28b**); for three genes, seven classes (**Fig. 3.28c**); and for four genes, nine classes (*not shown*).

The distributions produced by three and four genes thus begin to approach continuous variation, and if we add a small contribution from environmental variation, a smoother curve will appear. After all, we would expect bean plants to grow better in good soil, with ample sunlight and water. The environmental component effectively converts the stepped bar graph to a continuous curve by producing some variation in expressivity within each genotypic class. Moreover, additional variation might arise from more than two alleles at some genes (**Fig. 3.28d**), unequal contribution to the phenotype by the various genes involved (review Fig. 3.21), interactions with modifier genes, and chance. Thus, from what we now know about the relation between genotype and phenotype, it is possible to see how only a handful of genes that behave according to known Mendelian principles can easily generate continuous variation.

Continuous (or quantitative) traits vary over a range of values and can usually be measured: the length of a tobacco flower in millimeters, the amount of milk produced by a cow per day in liters, or the height of a person in meters. Continuous traits are usually **polygenic**—controlled by multiple genes—and show the additive effects of a large number of alleles, which creates an enormous potential for variation within a population. Differences in the environments encountered by particular individuals contribute even more variation. We discuss quantitative, multifactorial traits in detail in Chapter 22.

A Comprehensive Example: Multiple Alleles of Several Genes Determine Dog Coat Color

Domestic dogs have a wide variety of coat colors and markings. The *E* and *B* genes described previously in Labrador retrievers (recall Fig. 3.13) are only two of at least 12 genes that control dog coat color and pattern. The roles of seven of these genes—*E*, *B*, *A*, *K*, *D*, *S*, and *M*—are the best understood. **Table 3.3** lists the proteins these genes specify and the nature of various alleles found in domesticated dogs.

TABLE 3.3 Some of the Genes Affecting Domestic Dog Coat Color and Pattern

Gene	Protein	Dominance Series of Alleles		Phenotypes
Pigment-type switch genes				
Gene *A Agouti*	Agouti signaling protein (ASIP)	$A^y > a^w > a^t > a$	A^y a^w a^t a	Fawn (lots of light pigment on hair) Agouti (light stripe on dark hair) Tan belly (only hairs on belly have some light pigment) Black or brown (no light stripe on hairs)
Gene *E Extension*	Melanocortin receptor (MC1R)	$E^m > E > e$	E^m E e	Black mask on fawn or brindle Eumelanin (dark) and pheomelanin (yellow) pigments Only pheomelanin (cream, tan, red)
Gene *K Kurokami*	Beta-defensin	$K^b > k^{br} > k^y$	K^b k^{br} k^y	Solid color Brindled Gene *A* markings expressed normally
Dilution genes				
Gene *D Dilute*	Melanophilin (MLPH)	$D > d$	D d	Colors not dilute Colors dilute
Gene *B Brown*	Tyrosine-related protein (TYRP1)	$B > b$	B b	Black: eumelanin deposited densely Brown: eumelanin deposited less densely
Pigment cell development and survival genes				
Gene *S Spotting*	Micropthalmia-associated transcription factor (MITF)	$S > s^p$	S s^p	No white markings Colored patches on white background
Gene *M Merle*	Premelanosome protein (PMEL)	$M^1 = M^2$	M^1 M^2	Coat color diluted (homozygote has various health problems) Normal color

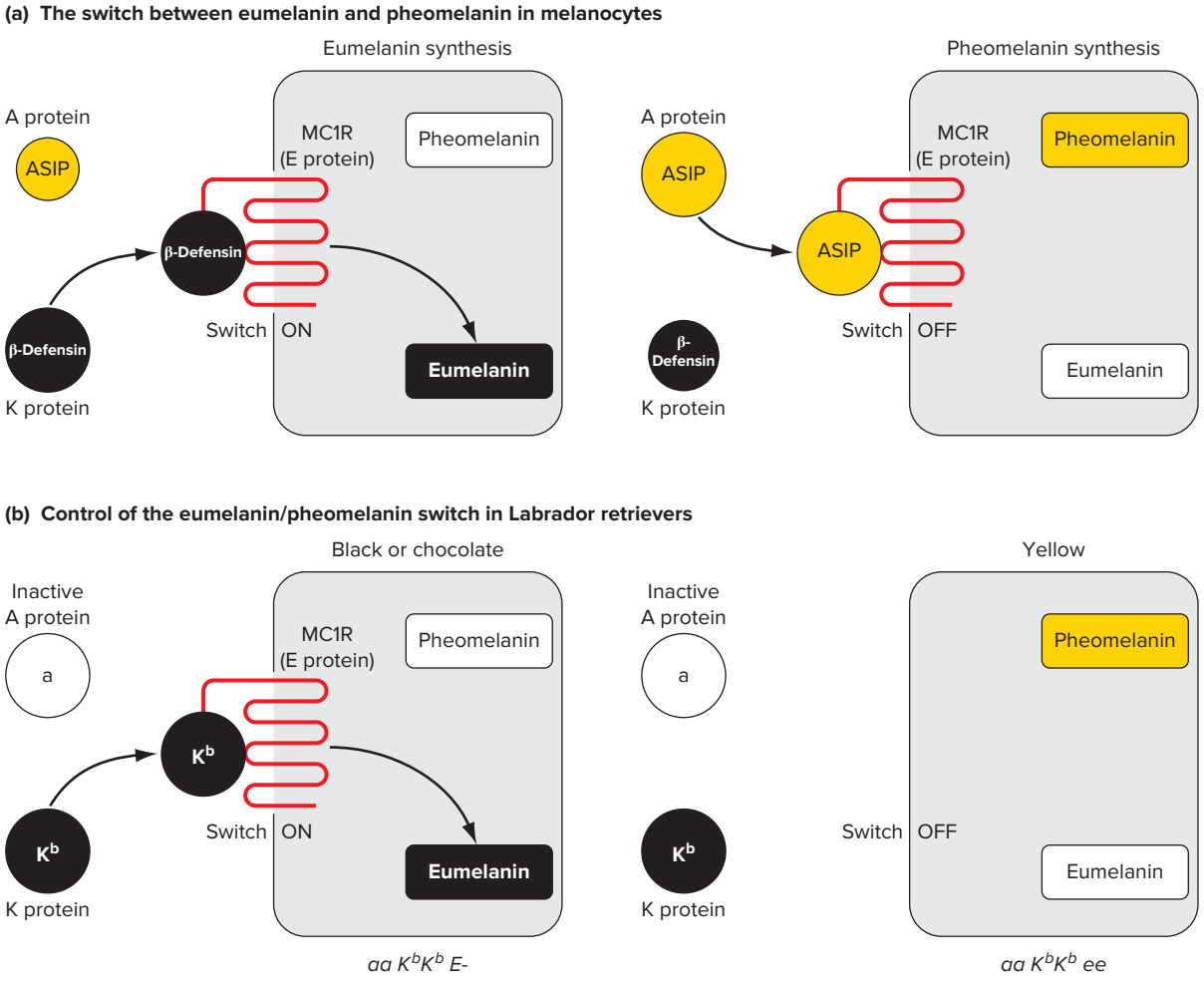
Genes *A, E,* and *K* control the switch from eumelanin to pheomelanin production

Skin cells called *melanocytes* make the pigments deposited in each dog hair. Melanocytes can produce either a dark pigment (eumelanin), or a light pigment (pheomelanin). The MC1R protein, specified by the *E* gene, spans the cell membrane and acts a switch that determines which pigment a melanocyte produces (**Fig. 3.29a**). Melanocytes produce eumelanin only when MC1R is switched on; pheomelanin is produced when the switch is turned off. By binding to MC1R at the cell surface, two proteins made by nearby skin cells control the MC1R switch. Binding of ASIP (specified by gene *A*) turns the switch off, but when Beta-defensin (specified by gene *K*) successfully outcompetes ASIP for MC1R binding, the switch flips on. The different alleles of the *E, A,* and *K* genes found in different dog breeds result in various colors and pigment distribution patterns on each hair and over the dog's body as a whole.

As described earlier, Labrador retrievers have two different gene *E* alleles; *E* specifies MC1R while *e* is nonfunctional. All Labrador retrievers are homozygous for the *a* allele of gene *A,* which specifies an inactive ASIP protein, and for the K^b allele of gene *K*, whose product is a functional Beta-defensin. Labradors that are *E*– produce eumelanin because MC1R is present and is switched on by the K^b Beta-defensin protein (**Fig. 3.29b**). These *E*– dogs are either black or chocolate depending on their gene *B* allele (as will be explained). In contrast, the melanocytes of *ee* Labradors have no MC1R, so the switch cannot be turned on; pheomelanin is made by default, and the dogs are yellow. (The particular shade of yellow in *ee* dogs, which varies from cream to red, is controlled by other genes not yet identified at the molecular level.)

Other dog breeds can have various alleles of genes *A, E* and *K*. Four different alleles of the *A* gene form a dominance series (Table 3.3). As just described, the *a* allele makes a nonfunctional protein. The proteins made by the other three alleles direct the pigment switch with different efficiencies or in different parts of the dog's body. Although the *A* gene in dogs is the same as the *A* gene in mice (review Fig. 3.7), the A^Y allele behaves differently. In contrast with mice, dogs may be homozygous A^YA^Y; they are an overall light brown color called fawn because the hairs contain a

Figure 3.29 Genes *E, A,* and *K* control the switch between light and dark pigment synthesis in melanocytes. MCR1 is specified by gene *E*. **(a)** At *left,* Beta-defensin (specified by gene *K*) out-competes ASIP for MC1R binding and switches MC1R on; melanocytes produce eumelanin (dark pigment). At *right,* melanocytes produce pheomelanin (light pigment) when ASIP (specified by gene *A*) out-competes Beta-defensin for MC1R binding. When bound by ASIP, MCR1 does not signal. **(b)** At *left,* the version of Beta-defensin made by the Labrador's K^b alleles always binds MC1R because it out-competes the version of ASIP made by their *a* alleles. As a result, eumelanin is synthesized and the dog is either black or chocolate (depending on its alleles of gene *B*). At *right,* the melanocytes of yellow Labrador retrievers lack MC1R (their *ee* alleles are nonfunctional), and pheomelanin is made by default.



lot of pheomelanin in addition to eumelanin. The a^w allele (like the *A* allele in mice) gives the dogs an overall gray (agouti) color. In an agouti dog (or mouse), the hairs are mainly black with a single yellow stripe. In both species, the a^t allele results in lighter hair on the belly and solid dark hairs on the back.

The *ee* genotype is epistatic to A^Y, a^w, and a^t because these *A* gene phenotypes require the hair to have some dark pigment. Some dogs breeds have an allele of gene *E* called E^m that is dominant to both *E* and *e* and directs the formation of a dark mask (**Fig. 3.30**). The E^m allele specifies a version of MC1R thought to cause more eumelanin production than normal, and the melanocytes around the muzzle are most sensitive to the effects of this increased eumelanin.

Three alleles of the *K* gene determine whether or not the *A* gene markings are visible, and we consider two of them here. The K^b allele present in Labradors specifies a version of Beta-defensin that is made in all melanocytes and can out-compete any of the different ASIP proteins made by gene *A* alleles. Thus in a K^b– dog, eumelanin is made all of the time, regardless of whether or not ASIP is present. If an *E*– dog is also K^b– (like black or chocolate Labrador retrievers), the dog is solidly dark regardless of its *A* gene alleles: K^b is epistatic to all alleles of gene *A*. In contrast, the version of Beta-defensin made by the k^y allele

Figure 3.30 Dog color pattern is a polygenic trait. The major alleles determining three kinds of dog coat color patterns.


Mask (E^m-)


Piebald ($s^p s^p$)


Merle (M^1M^2)

allows ASIP to inhibit MC1R sometimes, permitting pheomelanin production. As a result, $k^y k^y$ homozygotes allow expression of the fawn, agouti, or tan traits associated with the A^Y, a^w, or a^t alleles of the *A* gene.

Genes *B* and *D* control deposition of all pigments

Gene *B* specifies TYRP1, a multifunctional protein required for eumelanin synthesis and for depositing pigment in melanocytes. The *B* allele makes fully functional TYRP1, and the *b* allele specifies a less active version. As described earlier, *E*– *B*– Labrador retrievers are black because eumelanin is densely deposited, and their *E*– *bb* counterparts are brown because less eumelanin is produced and it is deposited less densely. Because the role of TYRP1 depends on the presence of MC1R (made by gene *E*), *ee* is epistatic to both *B* and *b* (recall Figs. 3.12 and 3.13).

Gene *D* specifies Melanophilin (MLPH), another protein required for pigment deposition. The recessive allele of gene *D* specifies a version of MLPH that functions less efficiently than the MLPH specified by the dominant allele. The lower amount of MLPH activity in *dd* homozygotes results in less pigment deposition and thus dilution of the colors dictated by the other genes. The dominant (normal) allele (*D*) does not dilute the colors. For example, an *E*– *B*– *D*– dog is black, while an *E*– *B*– *dd* dog is light black.

Genes *S* and *M* control spotting

Dogs homozygous for the recessive allele of gene *S* (that is, $s^p s^p$) are white with large spots of color; this pattern is called piebald (Fig. 3.30). As long as a dog has one dominant *S* allele, it will not be piebald. Gene *S* specifies a protein called MITF, a *transcription factor* needed to express (transcribe) many genes specifying enzymes needed for pigment production. The s^p allele makes a version of MITF that is less active than normal. Melanocyte precursor cells with only low levels of MITF die, resulting in white areas of the skin with no melanocytes and thus no pigment. By chance, some melanocyte precursor cells have sufficient MITF to survive, producing colored spots; the color is determined by genes other than *S*.

A second gene called *M* also controls the patterning of pigmentation, and it has codominant alleles M^1 and M^2. The M^2 allele is normal—it specifies a protein called PMEL required for eumelanin deposition. The M^1 allele makes an abnormal PMEL protein that interferes with eumelanin deposition and thus dilutes color. M^1M^2 heterozygotes, called *merle* dogs, have patches of diluted color (the M^1 phenotype) and patches of normal color (the M^2 phenotype) (Fig. 3.30). Breeders would never mate two merle dogs because the *M* gene is pleiotropic. The M^1 allele has recessive deleterious effects: So-called double merle dogs (M^1M^1) usually have serious health problems, including defects in hearing and vision. The eye and ear problems are due to the death of retinal and ear pigment cells caused by the abnormal PMEL protein.

This example of coat color in dogs gives some idea of the potential for variation from just half of the genes known to affect coat color. Amazingly, this is just the tip of the iceberg. When you realize that both dogs and humans carry roughly 27,000 genes, the number of interactions that connect the various alleles of these genes in the expression of phenotype is in the millions, if not the billions. The potential for variation and diversity among individuals is staggering indeed.

essential concepts

- In *incomplete penetrance,* a phenotype is expressed in fewer than 100% of individuals having the same genotype. In *variable expressivity,* a phenotype is expressed at different levels among individuals with the same genotype.
- A *continuous trait* can have any value of expression between two extremes. Most traits of this type are *polygenic,* that is, determined by the interactions of multiple genes.
- The environment and random events can interact with genes to influence the expression of many so-called *multifactorial traits.*

WHAT'S NEXT

Part of Mendel's genius was to look at the genetic basis of variation through a very narrow window, focusing his first glimpse of the mechanisms of inheritance on simple yet fundamental phenomena. Mendel worked on just a handful of traits in inbred populations of one species. For each trait, he manipulated one gene with one completely dominant and one recessive allele that determined two distinguishable, or discontinuous, phenotypes. Both the dominant and recessive alleles showed complete penetrance and negligible differences in expressivity.

In the first few decades of the twentieth century, many biologists questioned the general applicability of Mendelian analysis, for it seemed to shed little light on the complex inheritance patterns of most plant and animal traits or on the mechanisms producing continuous variation. Simple embellishments, however, clarified the genetic basis of continuous variation and provided explanations for other apparent exceptions to Mendelian analysis as described in this chapter. Each embellishment extends the range of Mendelian analysis and deepens our understanding of the genetic basis of variation. And no matter how broad the view, Mendel's basic conclusions, embodied in his first law of segregation, remain valid.

But what about Mendel's second law that genes assort independently? As it turns out, its application is not as universal as that of the law of segregation. Many genes do assort independently, but some do not; rather, the two genes appear to be linked and transmitted together from generation to generation. An understanding of this fact emerged from studies that located Mendel's hereditary units, the genes, in specific cellular organelles, the chromosomes. In describing how researchers deduced that genes travel with chromosomes, Chapter 4 establishes the physical basis of inheritance, including the segregation of alleles, and clarifies why some genes assort independently while others do not.

SOLVED PROBLEMS

I. Imagine you purchased an albino mouse (genotype *cc*) in a pet store. The *c* allele is epistatic to other coat color genes. How would you go about determining the genotype of this mouse at the brown locus? (In pigmented mice, *BB* and *Bb* are black, *bb* is brown.)

Answer

This problem requires an understanding of gene interactions, specifically epistasis. You have been placed in the role of experimenter and need to design crosses that will answer the question. To determine the alleles of the *B* gene present, you need to eliminate the blocking action of the *cc* genotype. Because only the recessive *c* allele is epistatic, when a *C* allele is present, no epistasis will occur. To introduce a *C* allele during the mating, the test mouse you mate to your albino can have the genotype *CC* or *Cc*. (If the mouse is *Cc,* half the progeny will be albino and will not contribute useful information, but the non-albinos from this cross would be informative.) What alleles of the *B* gene should the test mouse carry? To make this decision, work through the expected results using each of the possible genotypes.

Test mouse genotype		Albino mouse	Expected non-albino progeny
BB	×	*BB*	all black
	×	*Bb*	all black
	×	*bb*	all black
Bb	×	*BB*	all black
	×	*Bb*	3/4 black, 1/4 brown
	×	*bb*	1/2 black, 1/2 brown
bb	×	*BB*	all black
	×	*Bb*	1/2 black, 1/2 brown
	×	*bb*	all brown

From these hypothetical crosses, you can see that a test mouse with either the *Bb* or *bb* genotype would yield distinct outcomes for each of the three possible albino mouse genotypes. However, a *bb* test mouse would be more

useful and less ambiguous. First, it is easier to identify a mouse with the *bb* genotype because a brown mouse must have this homozygous recessive genotype. Second, the results are completely different for each of the three possible genotypes when you use the *bb* test mouse. (In contrast, a *Bb* test mouse would yield both black and brown progeny whether the albino mouse was *Bb* or *bb;* the only distinguishing feature is the ratio.) To determine the full genotype of the albino mouse, you should cross it to a brown mouse (which could be *CC bb* or *Cc bb*).

II. In a particular kind of ornamental flower, the wild-type flower color is deep purple, and the plants are true-breeding. In one true-breeding mutant stock, the flowers have a reduced pigmentation, resulting in a lavender color. In a different true-breeding mutant stock, the flowers have no pigmentation and are thus white. When a lavender-flowered plant from the first mutant stock was crossed to a white-flowered plant from the second mutant stock, all the F_1 plants had purple flowers. The F_1 plants were then allowed to self-fertilize to produce an F_2 generation. The 277 F_2 plants were 157 purple : 71 white : 49 lavender.

   a. Explain how flower color is inherited. Is this trait controlled by the alleles of a single gene?

   b. What kinds of progeny would be produced if lavender F_2 plants were allowed to self-fertilize?

Answer

a. Are any modes of single-gene inheritance compatible with the data? The observations that the F_1 plants look different from either of their parents and that the F_2 generation is composed of plants with three different phenotypes exclude complete dominance. The ratio of the three phenotypes in the F_2 plants has some resemblance to the 1:2:1 ratio expected from codominance or incomplete dominance, but the results would then imply that purple plants must be heterozygotes. This conflicts with the information provided that purple plants are true-breeding.

Consider now the possibility that two genes are involved. From a cross between plants heterozygous for two genes (*W* and *P*), the F_2 generation would contain a 9:3:3:1 ratio of the genotypes *W– P–, W– pp, ww P–,* and *ww pp* (where the dash indicates that the allele could be either a dominant or a recessive form). Would any combinations of the 9:3:3:1 ratio be close to that seen in the F_2 generation in this example? The numbers appear to fit best with a 9:4:3 ratio. What hypothesis would support combining two of the classes (3 + 1)? If *w* is epistatic to the *P* gene, then the *ww P–* and *ww pp* genotypic classes would have the same white phenotype. With this explanation, 1/3 of the F_2 lavender plants would be *WW pp,* and the remaining 2/3 would be *Ww pp.*

b. Upon self-fertilization, *WW pp* plants would produce only lavender (*WW pp*) progeny, while *Ww pp* plants would produce a 3:1 ratio of lavender (*W– pp*) and white (*ww pp*) progeny.

III. Huntington disease is a rare dominant condition in humans that results in a slow but inexorable deterioration of the nervous system. The disease shows what might be called age-dependent penetrance, which is to say that the probability that a person with the Huntington genotype will express the phenotype varies with age. Assume that 50% of those inheriting the *HD* allele will express the symptoms by age 40. Susan is a 35-year-old woman whose father has Huntington disease. She currently shows no symptoms. What is the probability that Susan will show symptoms in five years?

Answer

This problem involves probability and penetrance. Two conditions are necessary for Susan to show symptoms of the disease. A 1/2 (50%) chance exists that she inherited the mutant allele from her father, and if she does inherit the disease allele, a 1/2 (50%) chance exists that she will express the phenotype by age 40. Because these are independent events, the probability is the product of the individual probabilities, or 1/4.

PROBLEMS

Vocabulary

1. For each of the terms in the left column, choose the best matching phrase in the right column.

a. epistasis	1. one gene affecting more than one phenotype
b. modifier genes	2. the alleles of one gene mask the effects of alleles of another gene
c. conditional lethal	3. both parental phenotypes are expressed in the F_1 hybrids
d. permissive condition	4. a heritable change in a gene
e. reduced penetrance	5. genes whose alleles alter phenotypes produced by the action of other genes
f. multifactorial trait	6. less than 100% of the individuals possessing a particular genotype express it in their phenotype
g. incomplete dominance	7. environmental condition that allows conditional lethals to live

h. codominance	8. a trait produced by the interaction of alleles of at least two genes or from interactions between gene and environment
i. mutation	9. individuals with the same genotype have related phenotypes that vary in intensity
j. pleiotropy	10. a genotype that is lethal in some situations (for example, high temperature) but viable in others
k. variable expressivity	11. the heterozygote resembles neither homozygote

Section 3.1

2. In four-o'clocks, the allele for red flowers is incompletely dominant to the allele for white flowers, so heterozygotes have pink flowers. What ratios of flower colors would you expect among the offspring of the following crosses: (a) pink × pink, (b) white × pink, (c) red × red, (d) red × pink, (e) white × white, and (f) red × white? If you specifically wanted to produce pink flowers, which of these crosses would be most efficient?
3. The *Aa* heterozygous snapdragons in Fig. 3.3 are pink, while *AA* homozygotes are red. However, Mendel's *Pp* heterozygous pea flowers were every bit as purple as those of *PP* homozygotes (Fig. 2.8). Assuming that the *A* allele and the *P* allele specify functional enzymes, and the *a* and *p* alleles specify no protein at all, explain why the alleles of gene *A* and the alleles of gene *P* interact so differently.
4. Recall from Chapter 2 (Fig. 2.20) that Mendel's *R* gene specifies an enzyme called Sbe1 that forms branched starches. The dominant allele (*R*) makes protein, and the recessive allele (*r*) is nonfunctional. When considering the phenotype of round or wrinkled peas, *R* is completely dominant to *r: RR* and *Rr* peas are both equally round and *rr* peas are wrinkled. Imagine that the phenotype described is instead the average number of Sbe1 protein molecules in a pea. How would you describe the dominance relation between *R* and *r* in this case?
5. In the fruit fly *Drosophila melanogaster,* very dark (ebony) body color is determined by the *e* allele. The e^+ allele produces the normal wild-type, honey-colored body. In heterozygotes for the two alleles (but not in e^+e^+ homozygotes), a dark marking called the trident can be seen on the thorax, but otherwise the body is honey-colored. The e^+ snd *e* alleles are thus considered to be incompletely dominant.
   a. When female e^+e^+ flies are crossed to male e^+e flies, what is the probability that progeny will have the dark trident marking?
   b. Animals with the trident marking mate among themselves. Of 300 progeny, how many would be expected to have a trident, how many ebony bodies, and how many honey-colored bodies?
6. A cross between two plants that both have yellow flowers produces 80 offspring plants, of which 38 have yellow flowers, 22 have red flowers, and 20 have white flowers. If one assumes that this variation in color is due to inheritance at a single locus, what is the genotype associated with each flower color, and how can you describe the inheritance of flower color?
7. In radishes, color and shape are each controlled by a single locus with two incompletely dominant alleles. Color may be red (*RR*), purple (*Rr*), or white (*rr*) and shape can be long (*LL*), oval (*Ll*), or round (*ll*). What phenotypic classes and proportions would you expect among the offspring of a cross between two plants heterozygous at both loci?
8. A wild legume with white flowers and long pods is crossed to one with purple flowers and short pods. The F_1 offspring are allowed to self-fertilize, and the F_2 generation has 301 long purple, 99 short purple, 612 long pink, 195 short pink, 295 long white, and 98 short white. How are these traits being inherited?
9. Assuming no involvement of the Bombay phenotype (in case you've already read ahead to Section 3.2):
   a. If a girl has blood type O, what could be the genotypes and corresponding phenotypes of her parents?
   b. If a girl has blood type B and her mother has blood type A, what genotype(s) and corresponding phenotype(s) could the other parent have?
   c. If a girl has blood type AB and her mother is also AB, what are the genotype(s) and corresponding phenotype(s) of any male who could *not* be the girl's father?
10. Several genes in humans in addition to the ABO gene (*I*) give rise to recognizable antigens on the surface of red blood cells. The *MN* and *Rh* genes are two examples. The *Rh* locus can contain either a positive or a negative allele, with positive being dominant to negative. *M* and *N* are codominant alleles of the *MN* gene. The following chart shows several mothers and their children. For each mother-child pair, choose the father of the child from among the males in the right column, assuming one child per male.

	Mother	Child	Males
a.	O M Rh(pos)	B MN Rh(neg)	O M Rh(neg)
b.	B MN Rh(neg)	O N Rh(neg)	A M Rh(pos)
c.	O M Rh(pos)	A M Rh(neg)	O MN Rh(pos)
d.	AB N Rh(neg)	B MN Rh(neg)	B MN Rh(pos)

11. Alleles of the gene that determines seed coat patterns in lentils can be organized in a dominance series: marbled > spotted = dotted (codominant alleles) > clear. A lentil plant homozygous for the marbled seed coat pattern allele was crossed to one homozygous for the spotted pattern allele. In another cross, a homozygous dotted lentil plant was crossed to one homozygous for clear. An F_1 plant from the first cross was then mated to an F_1 plant from the second cross.
    a. What phenotypes in what proportions are expected from this mating between the two F_1 types?
    b. What are the expected phenotypes of the F_1 plants from the two original parental crosses?
12. One of your fellow students tells you that there is no way to know that the spotted and dotted patterns on the lentils in Fig. 3.4a are due to codominant alleles (C^S and C^D) of a single gene *C*. He claims that spotting could be controlled by gene *S*, with a completely dominant allele *S* that directs spotting and a recessive allele *s* that directs no spots. Likewise, he claims that dotting could be controlled by a separate gene *D*, with a completely dominant allele *D* that directs dotting and a recessive allele *d* that directs no dots. Is he correct, or does the information in Fig. 3.4a argue against this idea? Explain.
13. In a population of rabbits, you find three different coat color phenotypes: chinchilla (C), himalaya (H), and albino (A). To understand the inheritance of coat colors, you cross individual rabbits with each other and note the results in the following table.

Cross number	Parental phenotypes	Phenotypes of progeny
1	H × H	3/4 H : 1/4 A
2	H × A	1/2 H : 1/2 A
3	C × C	3/4 C : 1/4 H
4	C × H	all C
5	C × C	3/4 C : 1/4 A
6	H × A	all H
7	C × A	1/2 C : 1/2 A
8	A × A	all A
9	C × H	1/2 C : 1/2 H
10	C × H	1/2 C : 1/4 H : 1/4 A

    a. What can you conclude about the inheritance of coat color in this population of rabbits?
    b. Ascribe genotypes to the parents in each of the 10 crosses.
    c. What kinds of progeny would you expect, and in what proportions, if you crossed the chinchilla parents in crosses 9 and 10?
14. In clover plants, the pattern on the leaves is determined by a single gene with multiple alleles that are related in a dominance series. The gene is not pleiotropic. Seven different alleles of this gene are known; an allele that determines the absence of a pattern is recessive to the other six alleles, each of which produces a different pattern. All heterozygous combinations of alleles show complete dominance.
    a. How many different kinds of leaf patterns (including the absence of a pattern) are possible in a population of clover plants in which all seven alleles are represented?
    b. What is the largest number of different genotypes that could be associated with any one phenotype? Is there any phenotype that could be represented by only a single genotype?
    c. In a particular field, you find that the large majority of clover plants lack a pattern on their leaves, even though you can identify a few plants representative of all possible pattern types. Explain this finding.
15. Fruit flies with one allele for curly wings (*Cy*) and one allele for normal wings (Cy^+) have curly wings. When two curly-winged flies were crossed, 203 curly-winged and 98 normal-winged flies were obtained. In fact, all crosses between curly-winged flies produce nearly the same curly : normal ratio among the progeny.
    a. What is the approximate phenotypic ratio in these offspring?
    b. Suggest an explanation for these data.
    c. If a curly-winged fly was mated to a normal-winged fly, how many flies of each type would you expect among 180 total offspring?
16. In certain plant species such as tomatoes and petunias, a highly polymorphic *incompatibility* gene *S* with more than 100 known alleles prevents self-fertilization and promotes outbreeding. In this form of incompatibility, a plant cannot accept sperm carrying an allele identical to either of its own incompatibility alleles. If, for example, pollen carrying sperm with allele S^1 of the incompatibility gene lands onto the stigma (a female organ) of a plant that also carries the S^1 allele, the sperm cannot fertilize any eggs in that plant. (This phenomenon occurs because the pollen grain on the stigma cannot grow a pollen tube to allow the sperm to unite with the egg.)

    For the following crosses, indicate whether any progeny would be produced, and if so, list all possible genotypes of these progeny.

    a. ♂ $S^1 S^2$ × ♀ $S^1 S^2$
    b. ♂ $S^1 S^2$ × ♀ $S^2 S^3$
    c. ♂ $S^1 S^2$ × ♀ $S^3 S^4$
    d. Explain how this mechanism of incompatibility would prevent plant self-fertilization.

e. How does this incompatibility system ensure that all plants will be heterozygotes for different alleles of the *S* gene?

f. How do you know that peas are not governed by this incompatibility mechanism?

g. Explain why evolution would favor the emergence of new incompatibility alleles, making the gene increasingly polymorphic in populations of tomatoes or petunias.

17. In a species of tropical fish, a colorful orange and black variety called montezuma occurs. When two montezumas are crossed, 2/3 of the progeny are montezuma and 1/3 are the wild-type, dark grayish green color. Montezuma is a single-gene trait, and montezuma fish are never true-breeding.

    a. Explain the inheritance pattern seen here and show how your explanation accounts for the phenotypic ratios given.

    b. In this same species, the morphology of the dorsal fin is altered from normal to ruffled by homozygosity for a recessive allele designated *f*. What progeny would you expect to obtain, and in what proportions, from the cross of a montezuma fish homozygous for normal fins to a green, ruffled fish?

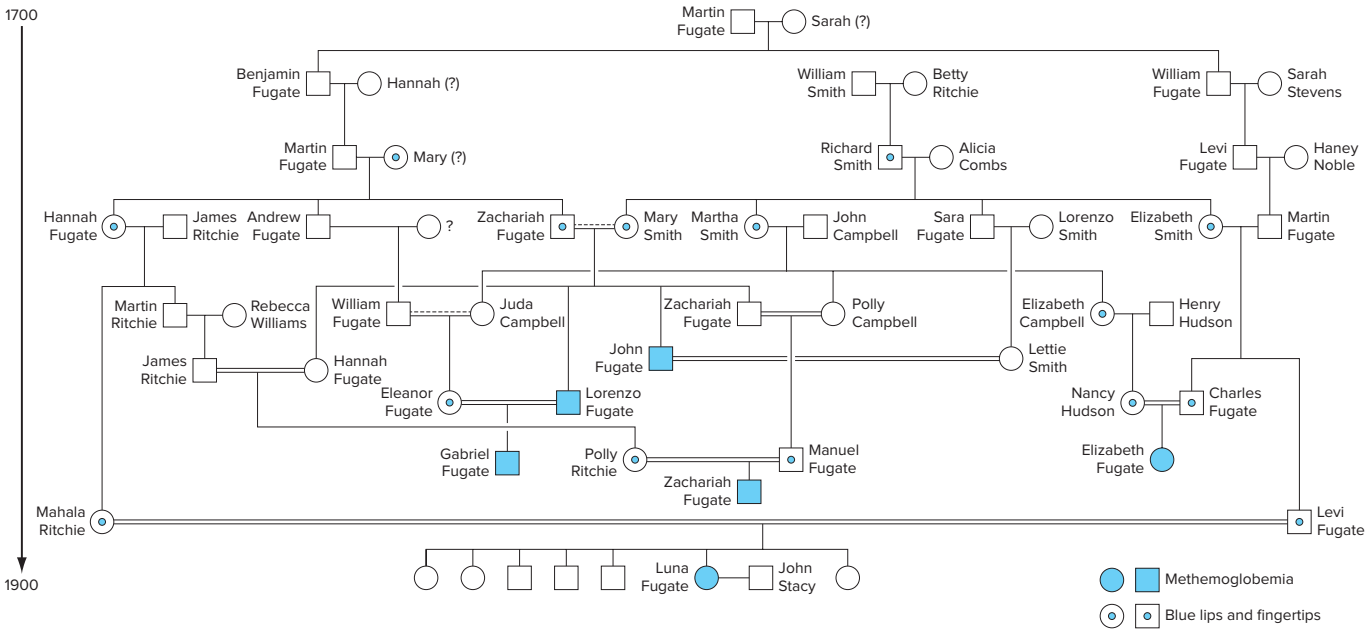
    c. What phenotypic ratios of progeny would be expected from the crossing of two of the montezuma progeny from part (b)?

18. People heterozygous for normal and nonfunctional mutant alleles of the *SMARCAD1* gene have a condition known as *adermatoglyphia:* they have no fingerprints. Sometimes adermatoglyphia is called *immigration delay disease* because people lacking fingerprints have trouble obtaining passports.

    Homozygotes for nonfunctional mutant alleles do not exist—they are never born. Describe the dominance relation between the mutant and wild-type alleles of *SMARCAD1*.

19. Using old Fugate family Bibles and the Perry County, Kentucky historical record, a hematologist in the 1960s constructed the pedigree of the *Blue People of Troublesome Creek*. Many members of the Fugate family had blue skin, a rare but harmless condition known as *methemoglobemia;* other people in the pedigree had blue lips and fingertips but their skin was otherwise normal. The blue color is due to lack of function of the enzyme NADH diaphorase, which repairs hemoglobin damaged by oxidation. Unrepaired hemoglobin accumulates as blue pigment.

© James Devaney/Getty Images



a. Based on the pedigree, describe the dominance relation between the wild-type and mutant alleles of the gene for NADH diaphorase.

b. The pedigree indicates certain people who were known to have only blue lips and fingertips. However, the historical record is incomplete. Which other people in the diagram must have had this phenotype? Explain any ambiguities that exist. (In case you've already read Section 3.3, assume that the blue lip and fingertip phenotype is fully penetrant.)

c. Two of the matings in the pedigree are shown as possibly consanguineous, as indicated by a dotted horizontal line above a solid horizontal line. The reason for the uncertainty is that the historical record does not say whether or not Mary [Mary (?)], the wife of the Martin Fugate at the top left of the diagram, was a Ritchie or a Smith or was instead unrelated to either family. Explain why a geneticist would think that Mary is likely a Ritchie or a Smith.

d. All of the Blue People (people with methemoglobemia) in the pedigree are Fugates, yet the blue mutation did not originate in the Fugate family. Which person or people introduced the mutant NADH diaphorase allele(s) into the Fugate family?

Section 3.2

20. A rooster with a particular comb morphology called *walnut* was crossed to a hen with a type of comb morphology known as *single*. The F_1 progeny all had walnut combs. When F_1 males and females were crossed to each other, 93 walnut and 11 single combs were seen among the F_2 progeny, but there were also 29 birds with a new kind of comb called *rose* and 32 birds with another new comb type called *pea*.

    a. Explain how comb morphology is inherited.

    b. What progeny would result from crossing a homozygous rose-combed hen with a homozygous pea-combed rooster? What phenotypes and ratios would be seen in the F_2 progeny?

    c. A particular walnut rooster was crossed to a pea hen, and the progeny consisted of 12 walnut, 11 pea, 3 rose, and 4 single chickens. What are the likely genotypes of the parents?

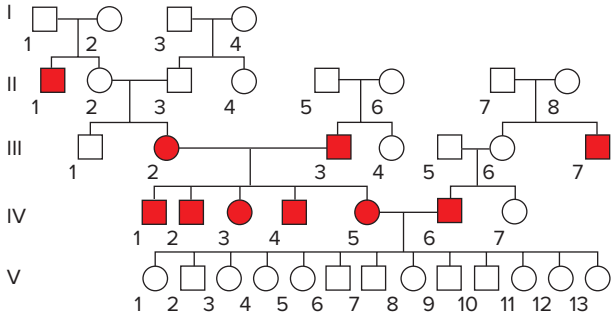
    d. A different walnut rooster was crossed to a rose hen, and all the progeny were walnut. What are the possible genotypes of the parents?

21. A black mare was crossed to a chestnut stallion and produced a bay son and a bay daughter. The two offspring were mated to each other several times, and they produced offspring of four different coat colors: black, bay, chestnut, and liver. Crossing a liver grandson back to the black mare gave a black foal, and crossing a liver granddaughter back to the chestnut stallion gave a chestnut foal. Explain how coat color is being inherited in these horses.

22. Filled-in symbols in the pedigree that follows designate individuals who are deaf.

    a. Study the pedigree and explain how deafness is being inherited.

    b. What is the genotype of the individuals in generation V? Why are they not deaf?



23. You perform a cross between two true-breeding strains of zucchini. One has green fruit and the other has yellow fruit. The F_1 plants are all green, but when these are crossed, the F_2 plants consist of 9 green : 7 yellow.

    a. Explain this result. What were the genotypes of the two parental strains?

    b. Indicate the phenotypes, with frequencies, of the progeny of a testcross of the F_1 plants.

    c. Describe the epistasis interactions observed.

    d. Suppose that the dominant alleles specify functional enzymes, and the recessive alleles are nonfunctional. Propose a biochemical pathway that could explain the gene interactions.

    e. Is it possible to cross two different pure-breeding yellow zucchini strains and obtain all green progeny? What would be the genotypes of the parents and progeny?

    f. Assuming that wild-type zucchini are green, how would you describe the phenomenon that occurred in the F_1 of part (e)?

24. Two true-breeding white strains of the plant *Illegitimati noncarborundum* were mated, and the F_1 progeny were all white. When the F_1 plants were allowed to self-fertilize, 126 white-flowered and 33 purple-flowered F_2 plants grew.

    a. How would you describe inheritance of flower color? Describe how specific alleles influence each other and therefore affect phenotype.

    b. A white F_2 plant is allowed to self-fertilize. Of the progeny, 3/4 are white-flowered, and 1/4 are

purple-flowered. What is the genotype of the white F_2 plant?

c. A purple F_2 plant is allowed to self-fertilize. Of the progeny, 3/4 are purple-flowered, and 1/4 are white-flowered. What is the genotype of the purple F_2 plant?

d. Two white F_2 plants are crossed with each other. Of the progeny, 1/2 are purple-flowered, and 1/2 are white-flowered. What are the genotypes of the two white F_2 plants?

25. Suppose the intermediate called *Colorless precursor 2* in the pathway shown in Fig. 3.16 was blue instead of colorless.

a. What would be the phenotypic ratio of the F_2? (Blue color is distinct from purple.)

b. Describe the type of genetic interaction that corresponds to this new phenotypic ratio.

26. Explain the difference between epistasis and dominance. How many loci are involved in each case?

27. The dominant allele *H* reduces the number of body bristles in fruit flies, giving rise to a *hairless* phenotype. In the homozygous condition, *H* is lethal. The dominant allele *S* has no effect on bristle number except in the presence of *H*, in which case a single *S* allele suppresses the hairless phenotype, thus restoring the bristles. However, *S* is also lethal in homozygotes.

a. What ratio of flies with normal bristles to hairless individuals would we find in the live progeny of a cross between two normal flies both carrying the *H* allele in the suppressed condition?

b. When the hairless progeny of the previous cross are crossed with one of the parental normal flies from part (a) (meaning a fly that carries *H* in the suppressed condition), what phenotypic ratio would you expect to find among their live progeny?

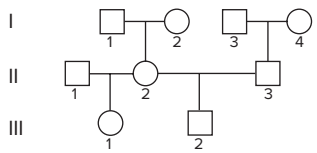
28. *Secretors* (genotypes *SS* and *Ss*) secrete their A and B blood group antigens into their saliva and other body fluids, while *nonsecretors* (*ss*) do not. What would be the apparent phenotypic blood group proportions among the offspring of an I^AI^B *Ss* woman and an I^AI^A *Ss* man if typing was done using saliva?

29. Normally, wild violets have yellow petals with dark brown markings and erect stems. Imagine you discover a plant with white petals, no markings, and prostrate stems. What experiment could you perform to determine whether the non-wild-type phenotypes are due to several different mutant genes or to the pleiotropic effects of alleles at a single locus? Explain how your experiment would settle the question.

30. A woman who is blood type B has a child whose blood type is A; her husband is blood type O. Despite his wife's claims of innocence, the irate father claims that the child is not his. Do you think that the wife is necessarily guilty of adultery? Explain.

31. The following table shows the responses of blood samples from the individuals in the pedigree to anti-A and anti-B sera. A plus (+) in the anti-A row indicates that the red blood cells of that individual were clumped by anti-A serum and therefore the individual made A antigens, and a minus (−) indicates no clumping. The same notation is used to describe the test for the B antigens.

	I-1	I-2	I-3	I-4	II-1	II-2	II-3	III-1	III-2
anti-A	+	+	–	+	–	–	+	+	–
anti-B	+	–	+	+	–	–	+	–	–



a. Deduce the blood type of each individual from the data in the table.

b. Assign genotypes for the blood groups as accurately as you can from these data, explaining the pattern of inheritance shown in the pedigree. Assume that all genetic relationships are as presented in the pedigree (that is, there are no cases of false paternity).

32. Three different pure-breeding strains of corn that produce ears with white kernels were crossed to each other. In each case, the F_1 plants were all red, while both red and white kernels were observed in the F_2 generation in a 9:7 ratio. These results are summarized in the following table.

	F_1	F_2
white-1 × white-2	red	9 red : 7 white
white-1 × white-3	red	9 red : 7 white
white-2 × white-3	red	9 red : 7 white

a. How many genes are involved in determining kernel color in these three strains?

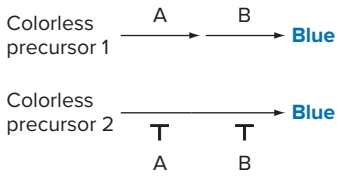
b. Define your symbols and show the genotypes for the pure-breeding strains white-1, white-2, and white-3.

c. Diagram the cross between white-1 and white-2, showing the genotypes and phenotypes of the F_1 and F_2 progeny. Explain the observed 9:7 ratio.

33. In mice, the A^Y allele of the *agouti* gene is a recessive lethal allele, but it is dominant for yellow coat color. What phenotypes and ratios of offspring would you expect from the cross of a mouse heterozygous at the *agouti* gene (genotype A^YA) and also at the *albino* gene (*Cc*) to an albino mouse (*cc*) heterozygous at the *agouti* gene (A^YA)?

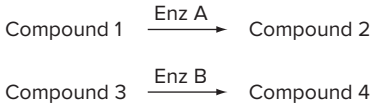
34. A student whose hobby was fishing pulled a very unusual carp out of Cayuga Lake: It had no scales on its body. She decided to investigate whether this strange nude phenotype had a genetic basis. She therefore obtained some inbred carp that were pure-breeding for the wild-type scale phenotype (body covered with scales in a regular pattern) and crossed them with her nude fish. To her surprise, the F_1 progeny consisted of a 1:1 ratio of wild-type fish and fish with a single linear row of scales on each side.
    a. Can a single gene with two alleles account for this result? Why or why not?
    b. To follow up on the first cross, the student allowed the linear fish from the F_1 generation to mate with each other. The progeny of this cross consisted of fish with four phenotypes: linear, wild type, nude, and scattered (the latter had a few scales scattered irregularly on the body). The ratio of these phenotypes was 6:3:2:1, respectively. How many genes appear to be involved in determining these phenotypes?
    c. In parallel, the student allowed the phenotypically wild-type fish from the F_1 generation to mate with each other and observed, among their progeny, wild-type and scattered carp in a ratio of 3:1. How many genes with how many alleles appear to determine the difference between wild-type and scattered carp?
    d. The student confirmed the conclusions of (c) by crossing those scattered carp with her pure-breeding wild-type stock. Diagram the genotypes and phenotypes of the parental, F_1, and F_2 generations for this cross and indicate the ratios observed.
    e. The student attempted to generate a true-breeding nude stock of fish by inbreeding. However, she found that this was impossible. Every time she crossed two nude fish, she found nude and scattered fish in the progeny, in a 2:1 ratio. (The scattered fish from these crosses bred true.) Diagram the phenotypes and genotypes of this gene in a nude × nude cross and explain the altered Mendelian ratio.
    f. The student now felt she could explain all of her results. Diagram the genotypes in the linear × linear cross performed by the student in (b). Show the genotypes of the four phenotypes observed among the progeny and explain the 6:3:2:1 ratio.

35. Suppose that blue flower color in a plant species is controlled by two genes, *A* and *B*. The dominant alleles *A* and *B* specify proteins that function in the pathways shown below. The A and B proteins are both required to make blue pigment from a colorless precursor. A and B proteins also independently inhibit the production of blue pigment from a different colorless precursor; that is, the presence of either protein A or protein B is sufficient to prevent blue pigment production from precursor 2. The recessive mutant alleles *a* and *b* specify no protein. Two different pure-breeding mutant strains with white flowers were crossed and complementation was observed so that all the F_1 were blue.

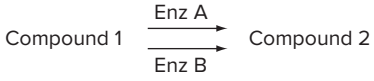


    a. What are the genotypes of each white mutant strain and the F_1?
    b. If the F_1 are selfed, what would be the phenotypic ratio of the F_2?

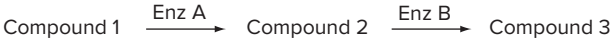
36. This problem examines possible biochemical explanations for variations of Mendel's 9:3:3:1 ratio. Except where indicated, compounds 1, 2, 3, and 4 have different colors, as do mixtures of these compounds. A and B are enzymes that catalyze the indicated steps of the pathway. Alleles *A* and *B* specify functional enzymes A and B, respectively; these are completely dominant to alleles *a* and *b*, which do not specify any of the corresponding enzyme. If functional enzyme is present, assume that the compound to the left of the arrow is converted completely to the compound to the right of the arrow. For each pathway, what phenotypic ratios would you expect among the progeny of a dihybrid cross of the form *Aa Bb* × *Aa Bb*?
    a. Independent pathways



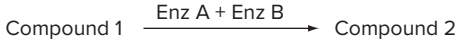
    b. Redundant pathways



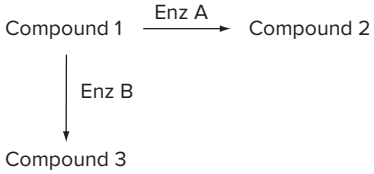
    c. Sequential pathway



    d. Enzymes A and B both needed to catalyze the reaction indicated.



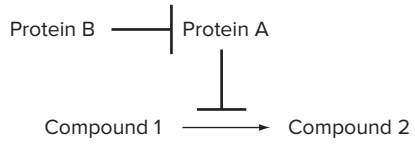
    e. Branched pathways (assume enough of compound 1 for both pathways)



f. Now consider independent pathways as in (a), but the presence of compound 2 masks the colors due to all other compounds.

g. Next consider the sequential pathway shown in (c), but compounds 1 and 2 are the same color.

h. Finally, examine the pathway that follows. Here, compounds 1 and 2 have different colors. The protein encoded by A prevents the conversion of compound 1 to compound 2. The protein encoded by B prevents protein A from functioning.



37. Considering your answers to Problem 36, does the existence of a particular variation of a 9:3:3:1 ratio among the F_2 progeny allow you to infer the operation of a specific biochemical mechanism responsible for these phenotypes? Inversely, if you know a biochemical mechanism of gene interaction, can you predict the ratios of the phenotypes you would see among the F_2 progeny?

Section 3.3

38. You picked up two mice (one female and one male) that had escaped from experimental cages in the animal facility. One mouse is yellow in color, and the other is brown agouti. (Agouti hairs have bands of yellow, while non-agouti hairs are solid-colored.) You know that this mouse colony has animals with different alleles at only three coat color genes: the agouti (*A*) or non-agouti (*a*) or yellow (A^Y) alleles of the *A* gene ($A^Y > A > a$; A^Y is a recessive lethal), the black (*B*) or brown (*b*) alleles of the *B* gene ($B > b$), and the albino (*c*) or non-albino (*C*) alleles of the *C* gene ($C > c$; *cc* is epistatic to all other phenotypes). However, you don't know which alleles of these genes are actually present in each of the animals that you've captured. To determine the genotypes, you breed the two escaped mice together. The first litter has only three pups. One is albino, one is brown (non-agouti), and the third is black agouti.

a. What alleles of the *A, B,* and *C* genes are present in the two mice you caught?

b. After raising several litters from these two parents, you have many offspring. How many different coat color phenotypes (in total) do you expect to see expressed in the population of offspring? What are the phenotypes and corresponding genotypes?

39. Figure 3.21 and Fig. 3.28b both show traits that are determined by two genes, each of which has two incompletely dominant alleles. But in Fig. 3.21 the gene interaction produces nine different phenotypes, while the situation depicted in Fig. 3.28b shows only five possible phenotypic classes. How can you explain this difference in the amount of phenotypic variation?

40. Three genes in fruit flies affect a particular trait, and one dominant allele of *each* gene is necessary to get a wild-type phenotype.

a. What phenotypic ratios would you predict among the progeny if you crossed triply heterozygous flies?

b. You cross a particular wild-type male in succession with three tester strains. In the cross with one tester strain (*AA bb cc*), only 1/4 of the progeny are wild type. In the crosses involving the other two tester strains (*aa BB cc* and *aa bb CC*), half of the progeny are wild type. What is the genotype of the wild-type male?

41. The garden flower *Salpiglossis sinuata* (painted tongue) comes in many different colors. Several crosses are made between true-breeding parental strains to produce F_1 plants, which are in turn self-fertilized to produce F_2 progeny.

Parents	F_1 phenotypes	F_2 phenotypes
red × blue	all red	102 red, 33 blue
lavender × blue	all lavender	149 lavender, 51 blue
lavender × red	all bronze	84 bronze, 43 red, 41 lavender
red × yellow	all red	133 red, 58 yellow, 43 blue
yellow × blue	all lavender	183 lavender, 81 yellow, 59 blue

a. State a hypothesis explaining the inheritance of flower color in painted tongues.

b. Assign genotypes to the parents, F_1 progeny, and F_2 progeny for all five crosses.

c. In a cross between true-breeding yellow and true-breeding lavender plants, all of the F_1 progeny are bronze. If you used these F_1 plants to produce an F_2 generation, what phenotypes in what ratios would you expect? Are there any genotypes that might produce a phenotype that you cannot predict from earlier experiments, and if so, how might this alter the phenotypic ratios among the F_2 progeny?

42. In foxgloves, three different petal phenotypes exist: white with red spots (WR), dark red (DR), and light red (LR). Two different kinds of true-breeding WR strains (WR-1 and WR-2) can be distinguished by two-generation intercrosses with true-breeding DR and LR strains:

			F_2		
	Parental	F_1	WR	LR	DR
1	WR-1 × LR	all WR	480	39	119
2	WR-1 × DR	all WR	99	0	32
3	DR × LR	all DR	0	43	132
4	WR-2 × LR	all WR	193	64	0
5	WR-2 × DR	all WR	286	24	74

a. What can you conclude about the inheritance of the petal phenotypes in foxgloves?

b. Ascribe genotypes to the four true-breeding parental strains (WR-1, WR-2, DR, and LR).

c. A WR plant from the F_2 generation of cross 1 is now crossed with an LR plant. Of 500 total progeny from this cross, there were 253 WR, 124 DR, and 123 LR plants. What are the genotypes of the parents in this WR × LR mating?

43. In a culture of fruit flies, matings between any two flies with hairy wings (wings abnormally containing additional small hairs along their edges) always produce both hairy-winged and normal-winged flies in a 2:1 ratio. You now take hairy-winged flies from this culture and cross them with four types of normal-winged flies; the results for each cross are shown in the following table. Assuming that only two possible alleles of the hairy-winged gene exist (one for hairy wings and one for normal wings), what can you say about the genotypes of the four types of normal-winged flies?

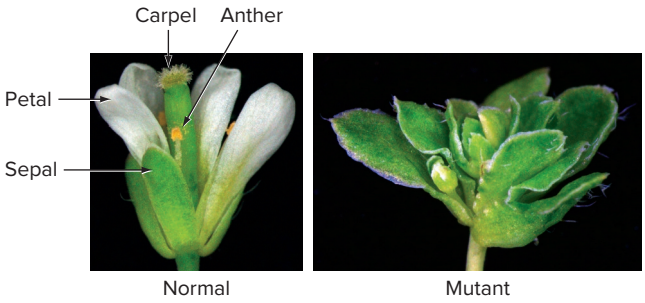
	Progeny obtained from cross with hairy-winged flies	
Type of normal-winged flies	**Fraction with normal wings**	**Fraction with hairy wings**
1	1/2	1/2
2	1	0
3	3/4	1/4
4	2/3	1/3

44. As shown in the picture that follows, flowers of the plant *Arabidopsis thaliana* (mustard weed) normally contain four different types of organs: sepals (leaves), petals, anthers (male sex organs), and carpels (female sex organs). The mutant strain shown in the picture at the right has abnormal flower morphology—the flower is made up entirely of sepals! Three genes (called *SEP1*, *SEP2*, and *SEP3*) function redundantly in the pathway for generating petals, anthers, and carpels. For normal flower morphology, the plant requires only one dominant, normally functioning allele of any one of these genes: *SEP1* (*A*) or *SEP2* (*B*) or *SEP3* (*C*). Recessive mutant alleles of these genes (*a*, *b*, or *c*) specify no protein.

a. What is the genotype of the mutant plant below?

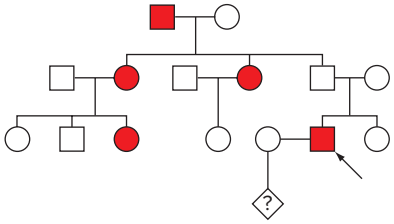
b. In a trihybrid cross of the type *AA bb cc* × *aa BB CC*, where all of the F_1 are *Aa Bb Cc*, what is the expected fraction of normal plants among the F_2 progeny?

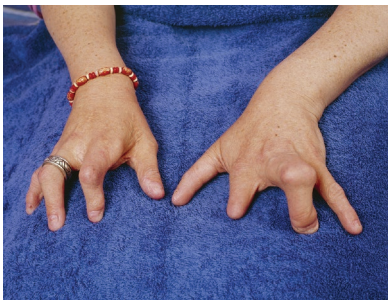
c. Suggest a model to explain how the *Arabidopsis thaliana* genome came to acquire three redundant genes.



© Sandra Biewers, www.sysflo.edu

45. A couple wants to know the probability that their expected child will suffer from split-hand deformity, which affects the prospective father, who is indicated by an arrow in the pedigree shown. (The arrow means that he is the *proband*—the person in the family who first brought the disorder to the attention of medical professionals.) This trait, shown in the following photo, is rare in the population, and the prospective parents are not related to each other.





© Maria Platt-Evans/Science Source

a. What is the mode of inheritance of this trait?

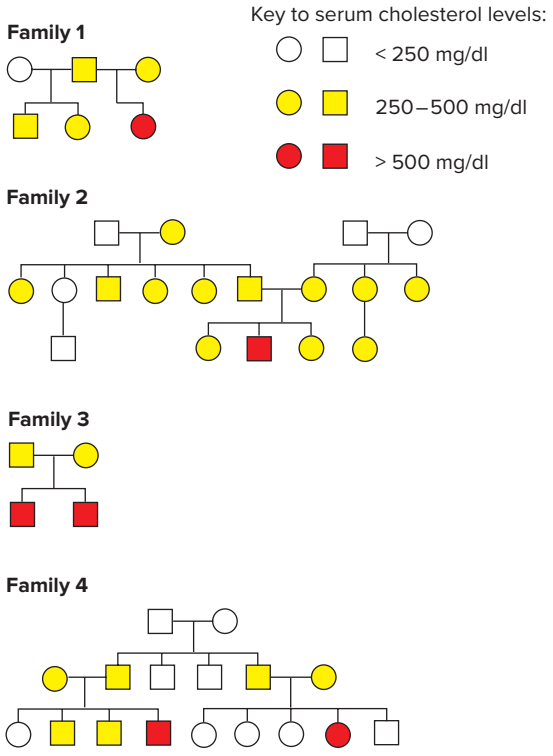
b. What is the penetrance of this trait [that is, the ratio between the number of individuals in the pedigree who display the trait (numerator), and the number of individuals you know from the pedigree *must* have the trait-determining genotype regardless of whether they have the trait or not (denominator)]?

c. Using your answer to part (b) above, what would you tell the parents about the numerical likelihood their expected child will have split-hand deformity?

d. Why is it possible that the likelihood the child will be affected is actually less than the number you

just answered in part (c)? You should specify the lowest numerical likelihood that could possibly be consistent with the data.

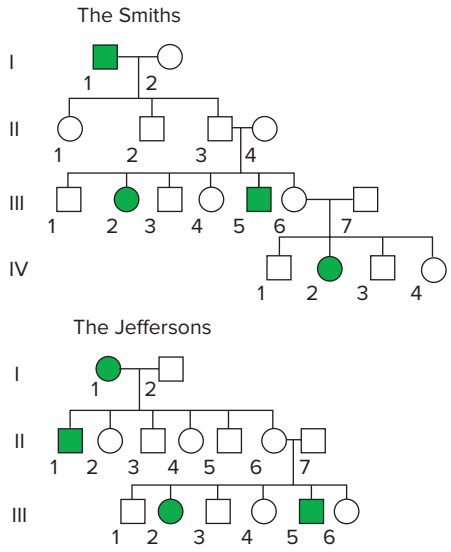
46. This problem illustrates why classical geneticists in the days before DNA analysis usually needed to work with traits showing complete penetrance. Consider the sweet peas shown in Fig. 3.15, where the *A– B–* genotypic class normally produces purple flowers and all other genotypic classes have white flowers.
    a. If the parental generation is *AA bb* × *aa BB,* what phenotypic ratio do you expect in the F_2 generation, assuming complete penetrance?
    b. Suppose now that only 75% of *A– B–* individuals have purple flowers (that is, the penetrance of this trait is 75%). What phenotypic ratio do you now expect among the F_2 plants?
    c. In doing these types of crosses, what kinds of results (other than an unexpected F_2 ratio) might suggest that penetrance of the purple phenotype is incomplete?
47. Spherocytosis is an inherited blood disease in which the erythrocytes (red blood cells) are spherical instead of biconcave. This condition can be inherited in a dominant fashion, with *ANK1* (the nonfunctional mutant allele) dominant to $ANK1^+$. In people with spherocytosis, the spleen recognizes the spherical red blood cells as defective and removes them from the bloodstream, leading to anemia. The spleen in different people removes the spherical erythrocytes with different efficiencies. Some people with spherical erythrocytes suffer severe anemia and some mild anemia, yet others have spleens that function so poorly no symptoms of anemia exist at all. When 2400 people with the genotype *ANK1* $ANK1^+$ were examined, it was found that all of them had spherical erythrocytes, 2250 had anemia of varying severity, and 150 had no symptoms. (Assume that *ANK1 ANK1* homozygotes do not exist.)
    a. Does this description of people with spherocytosis represent incomplete penetrance, variable expressivity, or both? Explain your answer. Can you derive any values from the numerical data to measure penetrance or expressivity?
    b. Suggest a treatment for spherocytosis and describe how the incomplete penetrance and/or variable expressivity of the condition might affect this treatment.
48. Familial hypercholesterolemia (FH) is an inherited trait in humans that results in higher-than-normal serum cholesterol levels [measured in milligrams of cholesterol per deciliter of blood (mg/dl)]. People with serum cholesterol levels that are roughly twice normal have a 25 times higher frequency of heart attacks than unaffected individuals. People with serum cholesterol levels three or more times higher than normal have severely blocked arteries and almost always die before they reach the age of 20. The following pedigrees show the occurrence of FH in four Japanese families:
    a. What is the most likely mode of inheritance of FH based on these data? Do any individuals in any of these pedigrees not fit your hypothesis? What special conditions might account for such individuals?
    b. Why do individuals in the same phenotypic class (unfilled, yellow, or red symbols) show such variation in their levels of serum cholesterol?



49. You have come into contact with two unrelated patients who express what you think is a rare phenotype—a dark spot on the bottom of the foot. According to a medical source, this phenotype is seen in 1 in every 100,000 people in the population. The two patients give their family histories to you, and you generate the pedigrees that follow.
    a. Given that this trait is rare, do you think the inheritance is dominant or recessive? Are there any special conditions that appear to apply to the inheritance?
    b. Which nonexpressing members of these families must carry the mutant allele?

c. If this trait is instead quite common in the population, what alternative explanation would you propose for the inheritance?

d. Based on this new explanation in (c), which non-expressing members of these families must have the genotype normally causing the trait?



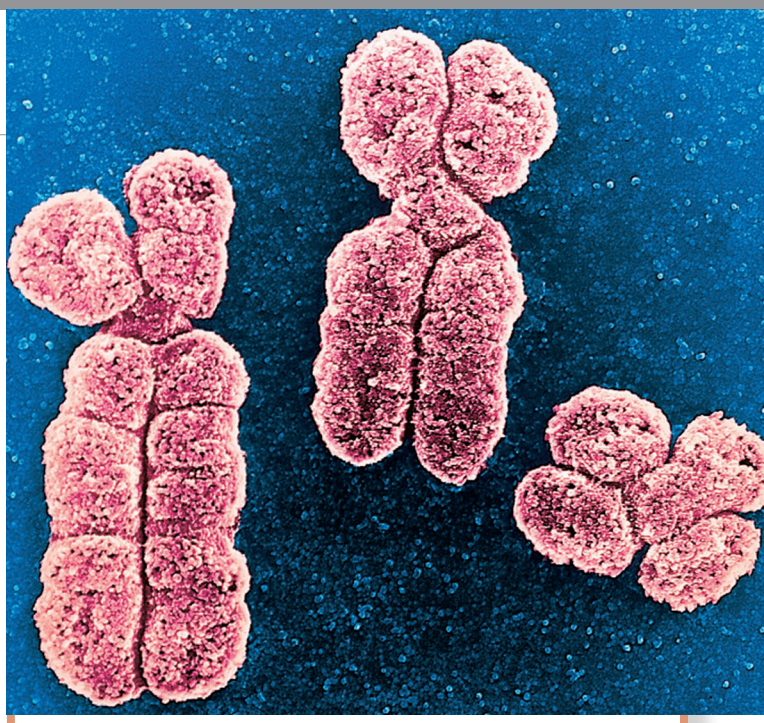
50. Polycystic kidney disease is a dominant trait that causes the growth of numerous cysts in the kidneys. The condition eventually leads to kidney failure. A child with polycystic kidney disease is born to a couple, neither of whom shows the disease. What possibilities might explain this outcome?

51. Identical (monozygotic) twins have similar, but not identical, fingerprints. Given that all the alleles of all the genes of identical twins are the same, explain how this outcome is possible.

52. Using each of the seven coat color genes discussed in the text (listed in Table 3.3), propose a possible genotype for each of the three Labrador retrievers in Fig. 3.12a. Keep in mind that the Labrador retrievers are pure-breeding for uniformly colored coats without spots or eye masks. Explain any ambiguities in your genotype assignments.

chapter 4

The Chromosome Theory of Inheritance



Each of these three human chromosomes carries hundreds of genes.

chapter outline

IN THE SPHERICAL, membrane-bound nuclei of plant and animal cells prepared for viewing under the microscope, chromosomes appear as brightly colored, threadlike bodies. The nuclei of normal human cells carry 23 pairs of chromosomes for a total of 46. Noticeable differences in size and shape exist among the 23 pairs, but within each pair, the two chromosomes appear to match exactly. (The only exceptions are the male's sex chromosomes, designated X and Y, which constitute an unmatched pair.)

Down syndrome was the first human genetic disorder attributable not to a gene mutation but to an abnormal number of chromosomes. Children born with Down syndrome have 47 chromosomes in each somatic cell nucleus because they carry three, instead of the normal pair, of a very small chromosome referred to as number 21. The aberrant genotype, called *trisomy 21*, gives rise to an abnormal phenotype, including a wide skull that is flatter than normal at the back, an unusually large tongue, learning disabilities caused by the abnormal development of the hippocampus and other parts of the brain, and a propensity to respiratory infections as well as heart disorders, rapid aging, and leukemia (**Fig. 4.1**).

How can one extra copy of a chromosome that is itself of normal size and shape cause such wide-ranging phenotypic effects? The answer has two parts. First and foremost, chromosomes are the cellular structures responsible for transmitting genetic information. In this chapter, we describe how geneticists concluded that chromosomes are the carriers of genes, an idea that became known as the **chromosome theory of inheritance.** The second part of the answer is that proper development depends not just on what type of genetic material is present, but also on how much of it there is. Thus the mechanisms governing gene transmission during cell division must vigilantly maintain each cell's chromosome number.

Cell division proceeds through the precise chromosome-parceling mechanisms of *mitosis* (for somatic, or body cells) and *meiosis* (for gametes—eggs and sperm). When

Figure 4.1 Down syndrome: One extra chromosome 21 has widespread phenotypic consequences.
Trisomy 21 usually causes changes in physical appearance as well as in the potential for learning. Many children with Down syndrome, such as the fifth grader at the center of the photograph, can participate fully in regular activities.

the machinery does not function properly, errors in chromosome distribution can have dire repercussions on the individual's health and survival. Down syndrome, for example, is the result of a failure of chromosome segregation during meiosis. The meiotic error gives rise to an egg or sperm carrying an extra chromosome 21 which, if incorporated in the zygote at fertilization, is passed on via mitosis to every cell of the developing embryo. Trisomy—three copies of a chromosome instead of two—can occur with other chromosomes as well, but in nearly all of these cases, the condition is prenatally lethal and results in a miscarriage.

Two themes emerge in our discussion of meiosis and mitosis. First, direct microscopic observations of chromosomes during gamete formation led early twentieth-century investigators to recognize that chromosome movements parallel the behavior of Mendel's genes, so chromosomes are likely to carry the genetic material. This chromosome theory of inheritance was proposed in 1902 and was confirmed in the following 15 years through elegant experiments performed mainly using the fruit fly *Drosophila melanogaster*. Second, the chromosome theory transformed the concept of a gene from an abstract particle to a physical reality—part of a chromosome that could be seen and manipulated.

4.1 Chromosomes: The Carriers of Genes

learning objectives

1. Differentiate among somatic cells, gametes, and zygotes with regard to the number and origin of their chromosomes.
2. Distinguish between homologous and nonhomologous chromosomes.
3. List the differences between sister chromatids and nonsister chromatids.

One of the first questions asked at the birth of an infant—is it a boy or a girl?—acknowledges that male and female normally are mutually exclusive characteristics like the yellow versus green of Mendel's peas. What's more, among humans and most other sexually reproducing species, a roughly 1:1 ratio exists between the two sexes. Both males and females produce cells specialized for reproduction—sperm or eggs—that serve as a physical link to the next generation. In bridging the gap between generations, these **gametes** must each contribute half of the genetic material for making a normal, healthy son or daughter. Whatever part of the gamete carries this material, its structure and function must be able to account for the either-or aspect of sex determination as well as the generally observed 1:1 ratio of males to females. These two features of sex determination were among the earliest clues to the cellular basis of heredity.

Genes Reside in the Nucleus

The nature of the specific link between sex and reproduction remained a mystery until Anton van Leeuwenhoek, one of the earliest and most astute of microscopists, discovered in 1667 that semen contains spermatozoa (literally *sperm animals*). He imagined that these microscopic creatures might enter the egg and somehow achieve fertilization, but it was not possible to confirm this hypothesis for another 200 years.

Then, during a 20-year period starting in 1854 (about the same time Gregor Mendel was beginning his pea experiments), microscopists studying fertilization in frogs and sea urchins observed the union of male and female gametes and recorded the details of the process in a series of drawings. These drawings, as well as later micrographs (photographs taken through a microscope), clearly show that egg and sperm nuclei are the only elements contributed equally by maternal and paternal gametes. This observation implies that something in the nucleus contains the hereditary material. In humans, the nuclei of the gametes are less than 2 millionths of a meter in diameter. It is indeed remarkable that the genetic link between generations is packaged within such an exceedingly small space.

Genes Reside in Chromosomes

Further investigations, some dependent on technical innovations in microscopy, suggested that yet smaller, discrete structures within the nucleus are the repository of genetic information. In the 1880s, for example, a newly discovered combination of organic and inorganic dyes revealed the existence of the long, brightly staining, threadlike bodies within the nucleus that we call **chromosomes** (literally *colored bodies*). It was now possible to follow the movement of chromosomes during different kinds of cell division.

In embryonic cells, the chromosomal threads split lengthwise in two just before cell division, and each of the two newly forming daughter cells receives one-half of every split thread. The kind of nuclear division followed by cell division that results in two daughter cells containing the same number and type of chromosomes as the original parent cell is called **mitosis** (from the Greek *mitos* meaning *thread* and *-osis* meaning *formation* or *increase*).

In the cells that give rise to male and female gametes, the chromosomes composing each pair become segregated, so that the resulting gametes receive only one chromosome from each chromosome pair. The kind of nuclear division that generates egg or sperm cells containing half the number of chromosomes found in other cells within the same organism is called **meiosis** (from the Greek word for *diminution*).

Fertilization: The union of haploid gametes to produce diploid zygotes

In the first decade of the twentieth century, cytologists—scientists who use the microscope to study cell structure—showed that the chromosomes in a fertilized egg actually consist of two matching sets, one contributed by the maternal gamete, the other by the paternal gamete. The corresponding maternal and paternal chromosomes appear alike in size and shape, forming pairs (with one exception—the *sex chromosomes*—which we discuss in a later section).

Gametes and other cells that carry only a single set of chromosomes are called **haploid** (from the Greek word for *single*). Zygotes and other cells carrying two matching sets are **diploid** (from the Greek word for *double*). The number of chromosomes in a normal haploid cell is designated by the shorthand symbol ***n*.** The number of chromosomes in a normal diploid cell is then **2*n*. Figure 4.2** shows diploid cells as well as the haploid gametes that arise from them in *Drosophila,* where $2n = 8$ and $n = 4$. In humans, $2n = 46$; $n = 23$.

You can see how the halving of chromosome number during meiosis and gamete formation, followed by the union of two gametes' chromosomes at fertilization, normally allows a constant $2n$ number of chromosomes to be maintained from generation to generation in all individuals of a species. The chromosomes of every pair must segregate from each other during meiosis so that the haploid gametes will each have one complete set of chromosomes. After fertilization forms the zygote, the process of mitosis then ensures that all the somatic cells of the developing individual have identical diploid chromosome sets.

Species variations in the number and shape of chromosomes

Scientists analyze the chromosomal makeup of a cell when the chromosomes are most visible—at a specific moment in the cell cycle of growth and division, just before the nucleus divides. At this point, known as *metaphase* (described in detail later), individual chromosomes have duplicated and condensed from thin threads into compact rodlike structures. Each chromosome now consists of two identical halves known as **sister chromatids** (**Fig. 4.3**).

The specific location at which sister chromatids are attached to each other is called the **centromere**. Each sister chromatid has its own centromere (Fig. 4.3), but in the duplicated chromosome, the two sister centromeres are pulled together so tightly that they form a *constriction* within which they cannot be resolved from each other, even in images obtained in the scanning electron microscope (see the picture at the beginning of the chapter).

Figure 4.2 Diploid versus haploid: 2*n* versus *n*. Fruit fly somatic cells are diploid: They carry a maternal and paternal copy of each chromosome. Meiosis generates haploid gametes with only one copy of each chromosome. In *Drosophila,* diploid cells have eight chromosomes ($2n = 8$), while gametes have four chromosomes ($n = 4$). Note that the chromosomes in this diagram are pictured before their replication. The X and Y chromosomes determine the sex of the individual.

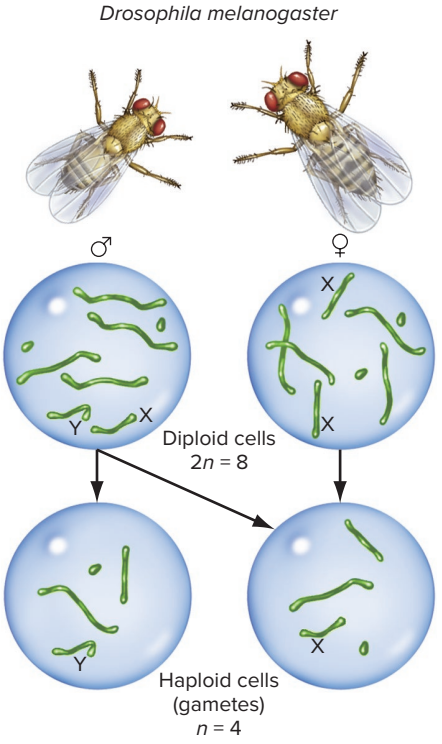


Figure 4.3 Metaphase chromosomes can be classified by centromere position. Before cell division, each chromosome replicates into two sister chromatids connected at their centromeres. In highly condensed metaphase chromosomes, the centromeres can appear near the middle (a *metacentric chromosome*), very near an end (an *acrocentric chromosome*), or anywhere in between. In a diploid cell, one homologous chromosome in each pair is from the mother and the other from the father.

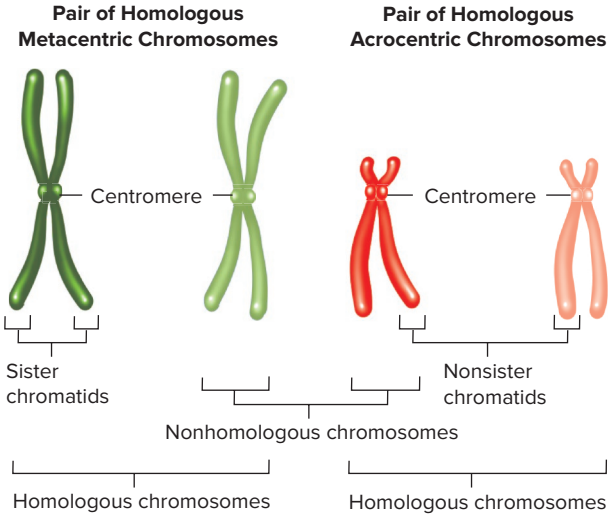
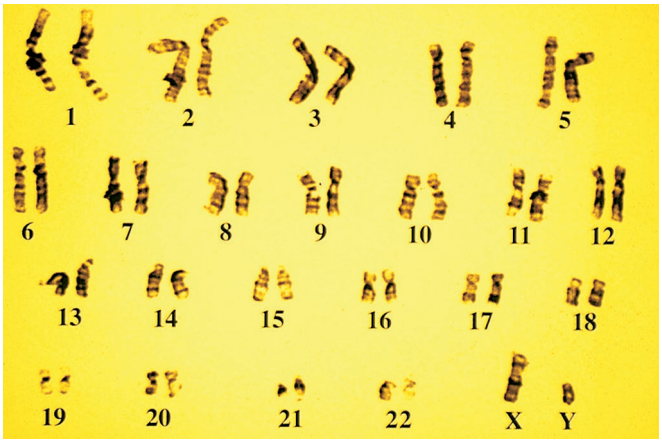


Figure 4.4 Karyotype of a human male. Photos of metaphase human chromosomes are paired and arranged in order of decreasing size. In a normal human male karyotype, 22 pairs of autosomes are present, as well as an X and a Y ($2n = 46$). Homologous chromosomes share the same characteristic pattern of dark and light bands.



Geneticists often describe chromosomes according to the location of the centromere (Fig. 4.3). In **metacentric** chromosomes, the centromere is more or less in the middle; in **acrocentric** chromosomes, the centromere is very close to one end. Chromosomes thus always have two *arms* separated by a centromere, but the relative sizes of the two arms can vary in different chromosomes.

Cells in metaphase can be fixed and stained with one of several dyes that highlight the chromosomes and accentuate the centromeres. The dyes also produce characteristic banding patterns made up of lighter and darker regions. Chromosomes that match in size, shape, and banding are called **homologous chromosomes,** or **homologs.** The two homologs of each pair contain the same set of genes, although for some of those genes, they may carry different alleles. The differences between alleles occur at the molecular level and don't show up in the microscope.

Figure 4.3 introduces a system of notation employed throughout this book, using color to indicate degrees of relatedness between chromosomes. Thus, sister chromatids, which are identical duplicates, appear in the same shade of the same color. Homologous chromosomes, which carry the same genes but may vary in the identity of particular alleles, are pictured in different shades (light or dark) of the same color. **Nonhomologous chromosomes,** which carry completely unrelated sets of genetic information, appear in different colors.

To study the chromosomes of a single organism, geneticists arrange micrographs of the stained chromosomes in homologous pairs of decreasing size to produce a **karyotype.** Karyotype assembly can now be speeded and automated by computerized image analysis. **Figure 4.4** shows the karyotype of a human male, with 46 chromosomes arranged in 22 matching pairs of chromosomes and one nonmatching pair. The 44 chromosomes in matching pairs are known as **autosomes.** The two unmatched chromosomes in this male karyotype are called *sex chromosomes* because they determine the sex of the individual. (We discuss sex chromosomes in more detail in subsequent sections.)

Modern methods of DNA analysis can reveal differences between the maternally and paternally derived chromosomes of a homologous pair and can thus track the origin of the extra chromosome 21 that causes Down syndrome in individual patients. In 80% of cases, the third chromosome 21 comes from the egg; in 20%, from the sperm. The Genetics and Society box entitled *Prenatal Genetic Diagnosis* describes how physicians use karyotype analysis and a technique called *amniocentesis* to diagnose Down syndrome prenatally, roughly three months after a fetus is conceived.

Through thousands of karyotypes on normal individuals, cytologists have verified that the cells of each species carry a distinctive diploid number of chromosomes. For example, Mendel's peas contain 14 chromosomes in 7 pairs in each diploid cell, the fruit fly *Drosophila melanogaster* carries 8 chromosomes (4 pairs), macaroni wheat has 28 (14 pairs), giant sequoia trees 22 (11 pairs), goldfish 94 (47 pairs), dogs 78 (39 pairs), and people 46 (23 pairs). Differences in the size, shape, and number of chromosomes reflect differences in the assembled genetic material that determines what each species looks like and how it functions. As these figures show, the number of chromosomes does not always correlate with the size or complexity of the organism.

In the next section, you will see that the discovery that chromosomes carry information about an individual's sex led to the realization that chromosomes carry the genes that determine all traits.

GENETICS AND SOCIETY

Crowd: © Image Source/Getty Images RF

Prenatal Genetic Diagnosis

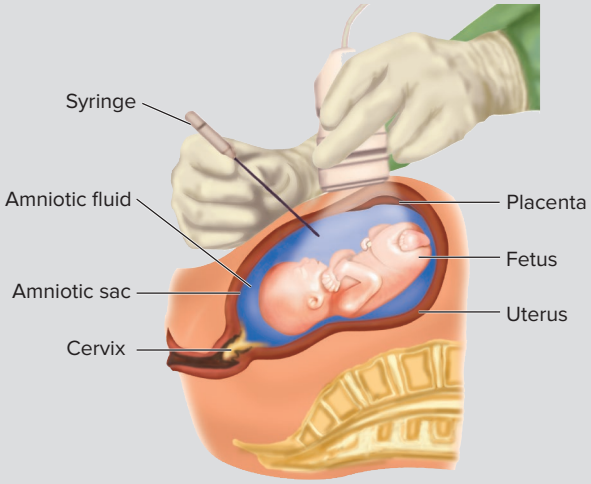
With new technologies for observing chromosomes and the DNA in genes, modern geneticists can define an individual's genotype directly. Doctors can use this basic strategy to diagnose, before birth, whether or not a baby will be born with a genetic condition.

The methods first developed for prenatal diagnosis were to obtain fetal cells whose DNA and chromosomes could be analyzed for genotype. The most frequently used method for acquiring these cells is **amniocentesis** (**Fig. A**). To carry out this procedure, a doctor inserts a needle through a pregnant woman's abdominal wall into the amniotic sac in which the fetus is growing; this procedure is performed about 16 weeks after the woman's last menstrual period. By using ultrasound imaging to guide the location of the needle, the doctor then withdraws some of the amniotic fluid in which the fetus is suspended into a syringe. This fluid contains living cells called *amniocytes* that were shed by the fetus. When placed in a culture medium, these fetal cells undergo several rounds of mitosis and increase in number. Once enough fetal cells are available, clinicians look at the chromosomes and genes in those cells. In later chapters, we describe techniques that allow the direct examination of the DNA constituting particular disease genes.

Amniocentesis also allows the diagnosis of Down syndrome through the analysis of chromosomes by karyotyping. Because the risk of Down syndrome increases rapidly with the age of the mother, more than half the pregnant women in North America who are over the age of 35 currently undergo amniocentesis. Although the goal of this karyotyping is usually to learn whether the fetus is trisomic for chromosome 21, many other abnormalities in chromosome number or shape may show up when the karyotype is examined.

More recently, scientists have been able to analyze the genotype of fetuses from the mother's blood, bypassing the need to obtain fetal cells. This procedure is made possible because the mother's blood contains cell-free fetal DNA. Fetal cells leak into the mother's bloodstream and then break down, releasing their DNA. Modern DNA sequencing techniques allow geneticists not only to genotype this material for particular disease-associated alleles, but even to determine the fetus's entire genome sequence. The analysis of fetal DNA obtained from the mother's blood is still experimental, but it likely will replace amniocentesis in the near future because drawing blood from the mother is inexpensive and noninvasive. The normal risk of miscarriage at 16 weeks' gestation is about 2–3%, and amniocentesis increases that risk by about 0.5% (about 1 in 200 precedures). In contrast, analyzing cell-free DNA from the mother's blood cannot harm the fetus.

Figure A Obtaining fetal cells by amniocentesis. A physician guides the insertion of the needle into the amniotic sac (aided by ultrasound imaging) and extracts amniotic fluid containing fetal cells into the syringe.



The availability of amniocentesis and **cell-free fetal DNA analysis** for prenatal diagnosis is intimately entwined with the personal and societal issue of abortion. The large majority of amniocentesis procedures are performed with the understanding that a fetus whose genotype indicates a genetic disorder, such as Down syndrome, will be aborted. Some prospective parents who are opposed to abortion still elect to undergo amniocentesis so that they can better prepare for an affected child, but this is rare.

The ethical and political aspects of the abortion debate influence many of the practical questions underlying prenatal diagnosis. For example, parents must decide which genetic conditions would be sufficiently severe that they would be willing to abort the fetus. From the economic point of view, society must decide who should pay for prenatal diagnosis procedures. In current practice, the risks and costs of amniocentesis generally restrict its application to women over age 35 or to mothers whose fetuses are at high risk for a testable genetic condition because of family history. The personal and societal equations determining the frequency of prenatal testing may, however, need to be overhauled in the not-too-distant future because technological advances such as the analysis of cell-free fetal DNA will minimize the costs and risks.

essential concepts

- *Chromosomes* are cellular structures specialized for the storage and transmission of genetic material.
- *Genes* are located on chromosomes and travel with them during cell division and gamete formation.
- *Somatic cells* carry a precise number of *homologous pairs* of chromosomes, which is characteristic of the species.
- In *diploid* organisms, one *homolog* of a pair is of maternal origin, and the other paternal.

4.2 Sex Chromosomes and Sex Determination

learning objectives

1. Predict the sex of humans with different complements of X and Y chromosomes.
2. Describe the basis of sex reversal in humans.
3. Compare the means of sex determination in different organisms.

Walter S. Sutton, a young American graduate student at Columbia University in the first decade of the twentieth century, was one of the earliest cytologists to realize that particular chromosomes carry the information for determining sex. In one study, he obtained cells from the testes of the great lubber grasshopper (*Brachystola magna;* **Fig. 4.5**) and followed them through the meiotic divisions that produce sperm. He observed that prior to meiosis, precursor cells within the testes of a great lubber grasshopper contain a total of 24 chromosomes. Of these, 22 are found in 11 matched pairs and are thus autosomes. The remaining two chromosomes are unmatched. He called the larger of these the X chromosome and the smaller the Y chromosome.

After meiosis, the sperm produced within these testes are of two equally prevalent types: one-half have a set of 11 autosomes plus an X chromosome, while the other half have a set of 11 autosomes plus a Y. By comparison, all of the eggs produced by females of the species carry an 11-plus-X set of chromosomes like the set found in the first class of sperm. When a sperm with an X chromosome fertilizes an egg, an XX female grasshopper results; when a Y-containing sperm fuses with an egg, an XY male develops. Sutton concluded that the X and Y chromosomes determine sex.

Figure 4.5 The great lubber grasshopper. In this mating pair, the smaller male is astride the female.

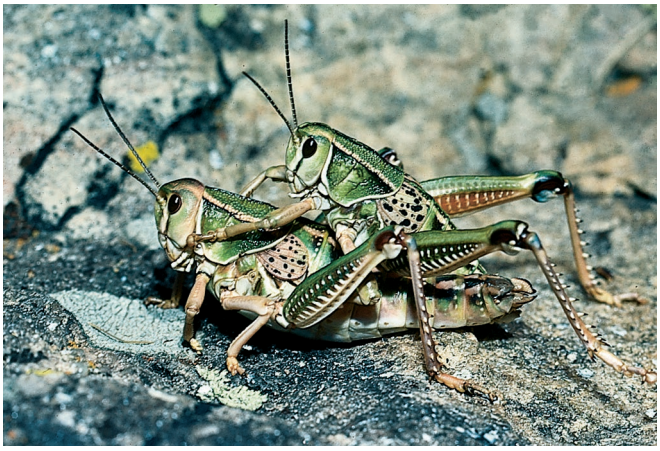
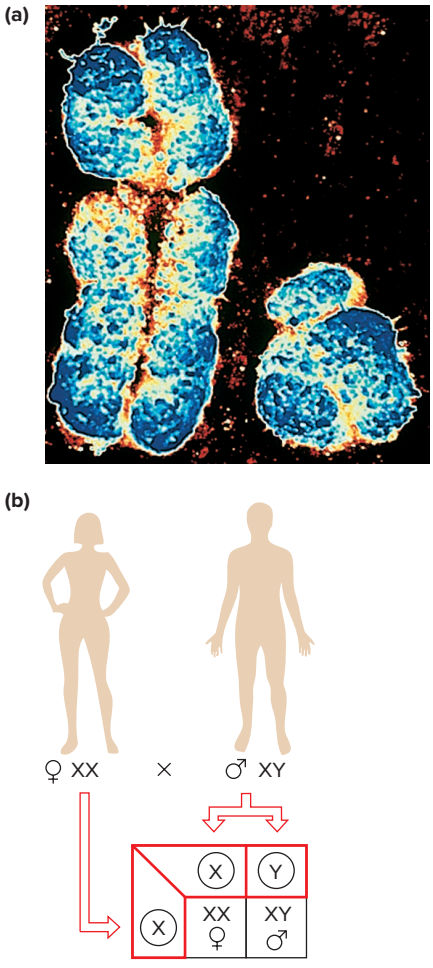


Figure 4.6 The X and Y chromosomes determine sex in humans. (a) This colorized micrograph shows the human X chromosome on the *left* and the human Y on the *right.* **(b)** Children can receive only an X chromosome from their mother, but they can inherit either an X or a Y from their father.



Several researchers studying other organisms soon verified that in many sexually reproducing species, two distinct chromosomes—known as the **sex chromosomes**—provide the basis of sex determination. One sex carries two copies of the same chromosome (a matching pair), while the other sex has one of each type of sex chromosome (an unmatched pair). The cells of normal human females, for example, contain 23 pairs of chromosomes. The two chromosomes of each pair, including the sex-determining X chromosomes, appear to be identical in size and shape. In males, however, one unmatched pair of chromosomes is present: the larger of these is the X; the smaller, the Y (Fig. 4.4 and **Fig. 4.6a**). Apart from this difference in sex chromosomes, the two sexes are not distinguishable at any other pair of chromosomes. Thus, geneticists can designate women as XX and men as XY and represent sexual reproduction as a simple cross between XX and XY.

If sex is an inherited trait determined by a pair of sex chromosomes that separate to different cells during gamete formation, then an XX × XY cross could account for both the mutual exclusion of sexes and the near 1:1 ratio of males to females, which are hallmark features of sex determination (**Fig. 4.6b**). And if chromosomes carry information defining the two contrasting sex phenotypes, we can easily infer that chromosomes also carry genetic information specifying other characteristics as well.

In Humans, the *SRY* Gene Determines Maleness

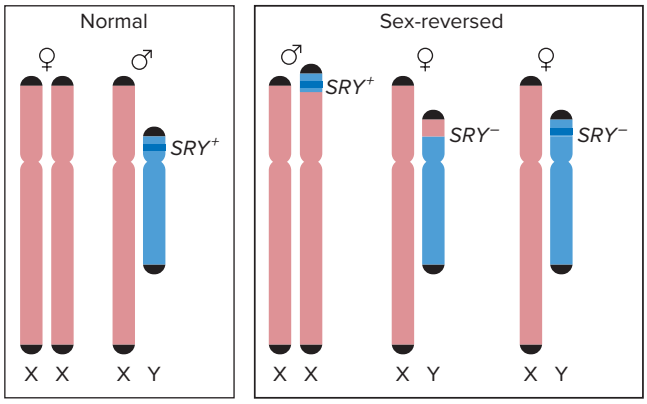
You have just seen that humans and other mammals have a pair of sex chromosomes that are identical in the XX female but different in the XY male. Several studies have shown that in humans, it is the presence or absence of the Y that actually makes the difference; that is, any person carrying a Y chromosome will look like a male. For example, rare humans with two X and one Y chromosome (XXY) are males displaying certain abnormalities collectively called *Klinefelter syndrome.* Klinefelter males are typically tall, thin, and sterile, and they sometimes show mental retardation. That these individuals are males shows that two X chromosomes are insufficient for female development in the presence of a Y.

In contrast, humans carrying an X and no second sex chromosome (XO) are females with *Turner syndrome.* Turner females are usually sterile, lack secondary sexual characteristics such as pubic hair, are of short stature, and have folds of skin between their necks and shoulders (webbed necks). Even though these individuals have only one X chromosome, they develop as females because they have no Y chromosome.

In 1990, researchers discovered that it is not the entire Y chromosome, but rather a single Y-chromosome-specific gene called ***SRY*** (*sex determining region of Y*) that is the primary determinant of maleness. The evidence implicating *SRY* came from so-called **sex reversal:** the existence of XX males and XY females (**Fig. 4.7**). In many sex-reversed XX males, one of the two X chromosomes carries a portion of the Y chromosome. Although in different XX males, different portions of the Y chromosome are found on the X, one particular gene—*SRY*—is always present. Sex-reversed XY females, in contrast, always have a Y chromosome lacking a functional *SRY* gene; the portion of the Y chromosome containing *SRY* is either replaced by a portion of the X chromosome, or the Y contains a nonfunctional mutant copy of *SRY* (Fig. 4.7). Later experiments with mice confirmed that *SRY* indeed determines maleness. These experiments are described in the Fast Forward Box *Transgenic Mice Prove That* SRY *Is the Maleness Factor.*

SRY is one of about 110 protein-coding genes on the Y chromosome. The two ends of the Y chromosome are called the **pseudoautosomal regions (PARs)** because homologous DNA sequences are present at the ends of the X chromosome

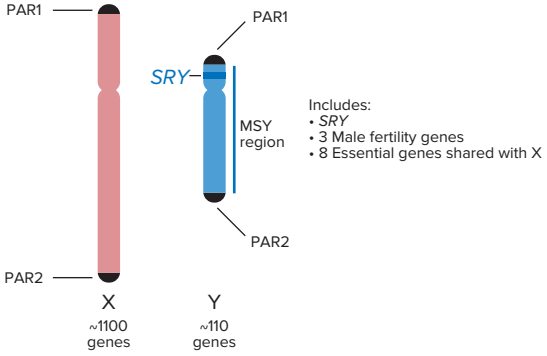
Figure 4.7 Sex reversal. Sex-reversed XX males have a part of the Y including the *SRY* gene on one of their X chromosomes. Sex-reversed XY females lack *SRY* on their Y chromosome either because it has been replaced by part of the X chromosome or because it has been inactivated by mutation.



(**Fig. 4.8**). The two PARs (PAR1 and PAR2) together contain about 30 genes, copies of which are found on both the X and Y chromosomes.

Most of the Y chromosome, however, is called the male-specific region (MSY) (Fig. 4.8); the functions of only some of the genes in the MSY are understood. The MSY includes four Y-specific (and therefore male-specific) genes: *SRY* and three genes required for spermatogenesis. The name MSY is somewhat misleading because eight of the genes in the MSY also exist on the X chromosome, but unlike the PAR genes, they are not grouped together in one region of either the X or Y. These eight MSY genes affect the functions of cells and tissues all over the body. In fact, several of these MSY genes shared with X are essential for male viability because without the Y-linked copies, the single gene copies on the X chromosome do not supply sufficient protein. (Females normally express both alleles of the X-linked copies of these eight genes, as these genes escape a phenomenon described later in this chapter.)

Figure 4.8 Human sex chromosomes have both shared and unique genes. PAR1 and PAR2 (*black*) are homologous regions of the X and Y chromosomes that together contain about 30 genes. The MSY region contains genes needed for maleness itself (*SRY*), genes for male fertility, and essential genes shared with the X required for male viability because their X-linked counterparts alone do not produce enough protein.



FAST FORWARD

Sprinters: © Robert Michael/Corbis RF

Transgenic Mice Prove That *SRY* Is the Maleness Factor

Genes similar to human *SRY* have been identified on the Y chromosomes of nearly all mammalian species. In 1991, researchers used mouse transgenic technology to show definitively that the *SRY* gene is the crucial determinant of maleness. A *transgenic mouse* is one whose genome contains copies of a gene that came from another individual—or even from another species. Such genes are called *transgenes*. One focus of genetic engineering is technology for the manipulation and insertion of transgenes.

To determine if *SRY* is sufficient to determine maleness, researchers wanted to introduce copies of the mouse *SRY* gene into the genome of chromosomally female (XX) mice. If *SRY* is the crucial determinant of maleness, then XX mice containing an *SRY* transgene would nevertheless be male.

First, the scientists isolated the DNA of the mouse *SRY* gene using cloning technology to be discussed in later chapters. Next, using a method called *pronuclear injection,* transgenic mice were generated that contained the *SRY* gene on one of their autosomes. To perform pronuclear injection, researchers collected many fertilized mouse eggs from mated females and injected the sperm or egg nucleus (called a *pronucleus* when in the zygote) with hundreds of copies of the *SRY* gene DNA (**Fig. A**). Enzymes in the pronucleus integrated the DNA into random locations in the genome (Fig. A).

After the injected zygotes matured into early embryos, they were implanted into surrogate mothers. When the mice were born, cells were taken from their tails and tested for the presence of the *SRY* transgene using molecular biology techniques.

Figure B shows at the right a transgenic mouse (*transformed* with *SRY*) obtained in this study. Although it is chromosomally XX, it is phenotypically male. This result demonstrates conclusively that the *SRY* gene alone is sufficient to determine maleness.

Figure A Using pronuclear injection to generate mice transgenic for the SRY gene.

(1): © Brigid Hogan, Howard Hughes Medical Institute, Vanderbilt University; (2): © Charles River Laboratories

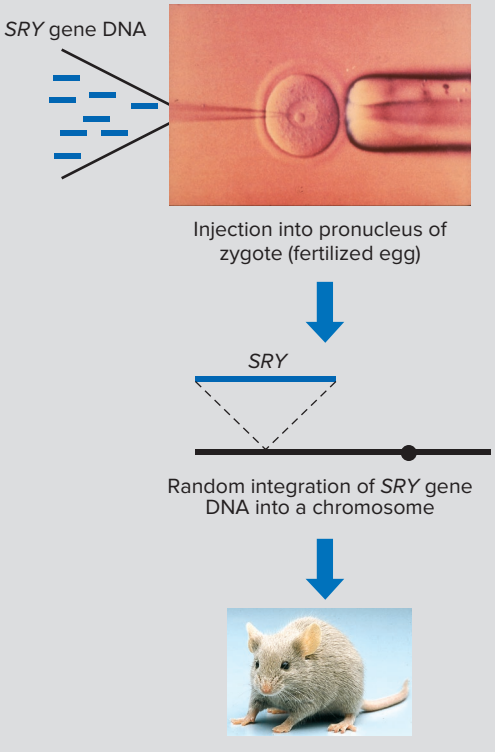


Figure B An XX mouse transformed with *SRY* is phenotypically male. Both the transformed XX mouse at the *right* and its normal XY littermate at the *left* have normal male genitalia. Arrows point to the penis.

© Medical Research Council/Science Source

The X chromosome contains about 1100 genes, most of which have nothing to do with sex; they specify proteins needed by both males and females.

Why does having an *SRY* gene mean that you will be male and not having *SRY* mean that you will be female? Approximately six weeks after fertilization, SRY protein activates testes development in XY (or sex-reversed XX) embryos. The embryonic testes secrete hormones that trigger the development of male sex organs and prevent the formation of female sex organs. In the absence of SRY protein, an ovary develops instead of a testis, and other female sex organs develop by default.

TABLE 4.1 Sex Determination in Fruit Flies and Humans

	Complement of Sex Chromosomes						
	XXX	**XX**	**XXY**	**XO**	**XY**	**XYY**	**OY**
Drosophila	Dies	Normal female	Normal female	Sterile male	Normal male	Normal male	Dies
Humans	Nearly normal female	Normal female	Klinefelter male (sterile); tall, thin	Turner female (sterile); short, webbed neck	Normal male	Normal or nearly normal male	Dies

Humans can tolerate extra X chromosomes (e.g., XXX) better than can *Drosophila* because in humans all but one X chromosome becomes a Barr body, as discussed later in this chapter. Complete absence of an X chromosome is lethal to both fruit flies and humans. Additional Y chromosomes have little effect in either species. Although the Y chromosome in *Drosophila* does not determine whether a fly looks like a male, it is necessary for male fertility; XO flies are thus sterile males.

Species Vary Enormously in Sex Determining Mechanisms

Other species show variations on this XX versus XY chromosomal strategy of sex determination. In fruit flies, for example, although normal females are XX and normal males XY (see Fig. 4.2), it is ultimately the number of X chromosomes (and not the presence or absence of the Y) that determines sex. The different responses of humans and *Drosophila* to the same unusual complements of sex chromosomes (**Table 4.1**) reveal that the mechanisms for sex determination differ in flies and humans. XXY flies are female because they have two X chromosomes, but XXY humans are male because they have a Y. Conversely, because they have one X chromosome, XO flies are male, while XO humans are female because they lack a Y.

The XX = female / XY = male strategy of sex determination is by no means universal. In some species of moths, for example, the females are XX, but the males are XO. In *C. elegans* (one species of nematode), males are similarly XO, but XX individuals are not females; they are instead self-fertilizing hermaphrodites that produce both eggs and sperm. In birds and butterflies, males have the matching sex chromosomes, while females have an unmatched set; in such species, geneticists represent the sex chromosomes as ZZ in the male and ZW in the female. The sex having two different sex chromosomes is termed the **heterogametic sex** because it gives rise to two different types of gametes; conversely, the sex with two similar sex chromosomes is the **homogametic sex.** The gametes of the heterogametic sex would contain either X or Y in the case of male humans, and either Z or W in the case of female birds; the gametes of the homogametic sex would contain only an X (humans) or only a Z (birds). Yet other variations include the complicated sex-determination mechanisms of bees and wasps, in which females are diploid and males haploid, and the systems of certain fish, in which sex is determined by changes in the environment, such as fluctuations in temperature. **Table 4.2** summarizes some of the astonishing variety in the ways that different species have solved the problem of assigning sex to individuals.

In spite of these many differences between species, early researchers concluded that chromosomes can carry the genetic information specifying sexual characteristics—and probably many other traits as well. Sutton and other early adherents of the chromosome theory realized that the perpetuation of life itself therefore depends on the proper distribution of chromosomes during cell division. In the next sections, you will see that the behavior of chromosomes during mitosis and meiosis is exactly that expected of cellular structures carrying genes.

TABLE 4.2 Mechanisms of Sex Determination

	♀	♂
Humans and *Drosophila*	XX	XY
Moths and *C. elegans*	XX (hermaphrodites in *C. elegans*)	X0
Birds and Butterflies	ZW	ZZ
Bees and Wasps	Diploid	Haploid
Lizards and Alligators	Cool temperature	Warm temperature
Tortoises and Turtles	Warm temperature	Cool temperature
Anemone Fish	Older adults	Young adults

In the species in the *top three rows,* sex is determined by sex chromosomes. The species in the *bottom four rows* have identical chromosomes in the two sexes, and sex is determined instead by environmental or other factors. Anemone fish (*bottom row*) undergo a sex change from male to female as they age.

essential concepts

- Many sexually reproducing organisms have *sex chromosomes* that are sex-specific and that determine sex.
- In humans, male sex determination is triggered by a Y-linked gene called *SRY*; female sex determination occurs in XX embryos by default.
- Mechanisms of sex determination vary remarkably; in some species sex is determined by environmental factors rather than by specific chromosomes.

4.3 Mitosis: Cell Division That Preserves Chromosome Number

learning objectives

1. Describe the key chromosome behaviors during mitosis.
2. Diagram the forces and structures that dictate chromosomal movement during mitosis.

The fertilized human egg is a single diploid cell that preserves its genetic identity unchanged through more than 100 generations of cells as it divides again and again to produce a full-term infant ready to be born. As the newborn infant develops into a toddler, a teenager, and an adult, yet more cell divisions fuel continued growth and maturation. Mitosis, the nuclear division that apportions chromosomes in equal fashion to two daughter cells, is the cellular mechanism that preserves genetic information through all these generations of cells. In this section, we take a close look at how the nuclear division of mitosis fits into the overall scheme of cell growth and division.

If you were to peer through a microscope and follow the history of one cell through time, you would see that for much of your observation, the chromosomes resemble a mass of extremely fine tangled string—called **chromatin**—surrounded by the **nuclear envelope.** Each convoluted thread of chromatin is composed mainly of DNA (which carries the genetic information) and protein (which serves as a scaffold for packaging and managing that information, as described in Chapter 12). You would also be able to distinguish one or two darker areas of chromatin called *nucleoli* (singular, **nucleolus,** literally *small nucleus*); nucleoli play a key role in the manufacture of ribosomes, organelles that function in protein synthesis. During the period between cell divisions, the chromatin-laden nucleus houses a great deal of invisible activity necessary for the growth and survival of the cell. One particularly important part of this activity is the accurate duplication of all the chromosomal material.

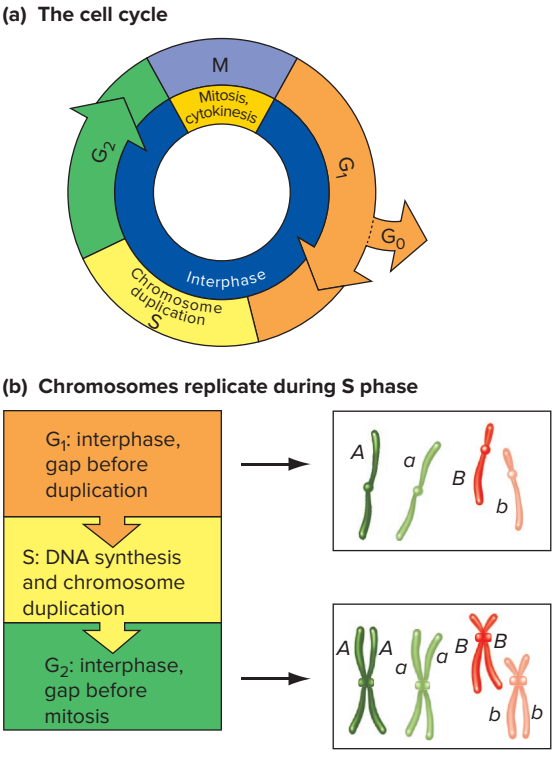
With continued vigilance, you would observe a dramatic change in the nuclear landscape during one very short period in the cell's life history: The chromatin condenses into discrete threads, and then each chromosome compacts even further into the twin rods clamped together at their centromeres that can be identified in karyotype analysis (review Fig. 4.3). Each rod in a duo is called a **chromatid;** as described earlier, it is an exact duplicate of the other sister chromatid to which it is connected. Continued observation would reveal the doubled chromosomes beginning to jostle around inside the cell, eventually lining up at the cell's midplane. At this point, the sister chromatids of each chromosome separate to opposite poles of the now elongating cell, where they become identical sets of chromosomes. Each of the two identical sets eventually ends up enclosed in a separate nucleus in a separate cell. The two cells, known as *daughter cells,* are thus genetically identical.

The repeating pattern of cell growth (an increase in size) followed by division (the splitting of one cell into two) is called the **cell cycle** (**Fig. 4.9**). Only a small part of the cell cycle is spent in division (or **M phase**); the period between divisions is called **interphase.**

During Interphase, Cells Grow and Replicate Their Chromosomes

Interphase consists of three parts: gap 1 **(G_1)**, synthesis **(S)**, and gap 2 **(G_2)** (Fig. 4.9). G_1 lasts from the birth of a new cell to the onset of chromosome replication; for the genetic material, it is a period when the chromosomes are neither duplicating nor dividing. During this time, the cell achieves most of its growth by using the information from its genes to make and assemble the materials it needs to function normally. G_1 varies in length more than any other phase of the cell cycle. In rapidly dividing cells of the human embryo, for example, G_1 is as short as a few hours. In contrast, mature brain cells

Figure 4.9 The cell cycle: An alternation between interphase and mitosis. (a) Chromosomes replicate to form sister chromatids during synthesis (S phase); the sister chromatids segregate to daughter cells during mitosis (M phase). The gaps between the S and M phases, during which most cell growth takes place, are called the G_1 and G_2 phases. In multicellular organisms, some terminally differentiated cells stop dividing and arrest in a G_0 stage. **(b)** Interphase consists of the G_1, S, and G_2 phases together.



become arrested in a resting form of G_1 known as G_0 and do not normally divide again during a person's lifetime.

Synthesis (S) is the time when the cell duplicates its genetic material by synthesizing DNA. During duplication, each chromosome doubles to produce identical sister chromatids that will become visible when the chromosomes condense at the beginning of mitosis. The two sister chromatids remain joined to each other at their centromeres. (Note that this joined structure is considered a single chromosome as long as the connection between sister chromatids is maintained.) The replication of chromosomes during S phase is crucial; the genetic material must be copied exactly so that both daughter cells receive identical sets of chromosomes.

Gap 2 (G_2) is the interval between chromosome duplication and the beginning of mitosis. During this time, the cell may grow (usually less than during G_1); it also synthesizes proteins that are essential to the subsequent steps of mitosis itself.

In addition, during interphase an array of fine microtubules crucial for many biochemical processes becomes visible outside the nucleus. The microtubules radiate out into the cytoplasm from a single organizing center known as the **centrosome,** usually located near the nuclear envelope. In animal cells, the discernible core of each centrosome is a pair of small, darkly staining bodies called **centrioles** (**Fig. 4.10a**); the microtubule-organizing center of plants does not contain centrioles. During the S and G_2 stages of interphase, the centrosomes replicate, producing two centrosomes that remain in extremely close proximity.

During Mitosis, Sister Chromatids Separate and Two Daughter Nuclei Form

Although the rigorously choreographed events of nuclear and cellular division occur as a dynamic and continuous process, scientists traditionally analyze the process in separate stages marked by visible cytological events. The artist's sketches in Fig. 4.10 illustrate these stages in the nematode *Ascaris,* whose diploid cells contain only four chromosomes (two pairs of homologous chromosomes).

Prophase: Chromosomes condense (Fig. 4.10a)

During all of interphase, the cell nucleus remains intact, and the chromosomes are indistinguishable aggregates of chromatin. At **prophase** (from the Greek *pro-* meaning *before*), the gradual emergence, or **condensation,** of individual chromosomes from the undifferentiated mass of chromatin marks the beginning of mitosis. Each condensing chromosome has already been duplicated during interphase and thus consists of sister chromatids attached at their centromeres. At this stage in *Ascaris* cells, therefore, four chromosomes exist with a total of eight chromatids.

The progressive appearance of an array of individual chromosomes is a truly impressive event. Interphase DNA molecules as long as 3–4 cm condense into discrete chromosomes whose length is measured in microns (millionths of a meter). This process is equivalent to compacting a 200 m length of thin string (as long as two football fields) into a cylinder 8 mm long and 1 mm wide.

Another visible change in chromatin also takes place during prophase: The darkly staining nucleoli begin to break down and disappear. As a result, the manufacture of ribosomes ceases, providing one indication that general cellular metabolism shuts down so that the cell can focus its energy on chromosome movements and cellular division.

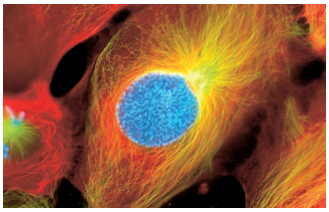
Several important events that characterize prophase occur outside the nucleus in the cytoplasm. The centrosomes, which replicated during interphase, now move apart and become clearly distinguishable as two separate entities in the light microscope. At the same time, the interphase scaffolding of long, stable microtubules disappears and is replaced by a set of dynamic microtubules that rapidly grow from and shrink back toward their centrosomal organizing centers. The centrosomes continue to move apart, migrating around the nuclear envelope toward opposite ends of the nucleus, apparently propelled by forces exerted between interdigitated microtubules extending from the two centrosomes.

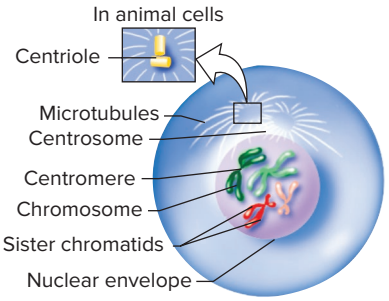
Prometaphase: The spindle forms (Fig. 4.10b)

Prometaphase (*before middle stage*) begins with the breakdown of the nuclear envelope, which allows microtubules extending from the two centrosomes to invade the nucleus. Chromosomes attach to these microtubules through the **kinetochore,** a structure in the centromere region of each chromatid that is specialized for conveyance. Each kinetochore contains proteins that act as molecular motors, enabling the chromosome to slide along the microtubule. When the kinetochore of a chromatid originally contacts a microtubule at prometaphase, the kinetochore-based motor moves the entire chromosome toward the centrosome from which that microtubule radiates. Microtubules growing from the two centrosomes capture chromosomes by connecting first to the kinetochore of one of the two sister chromatids, chosen at random. As a result, it is sometimes possible to observe groups of chromosomes congregating in the vicinity of each centrosome. In this early part of prometaphase, for each chromosome, one chromatid's kinetochore is attached to a microtubule, but the sister chromatid's kinetochore remains unattached.

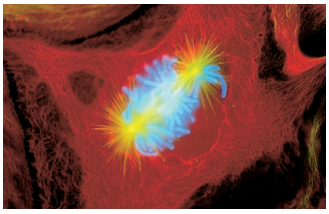
During prometaphase, three different types of microtubule fibers together form the **mitotic spindle.** All of these microtubule classes originate from the centrosomes, which function as the two *poles* of the spindle apparatus. Microtubules that extend between a centrosome and the kinetochore of a chromatid are called **kinetochore microtubules,** or *centromeric fibers*. Microtubules from each centrosome that are directed toward the middle of the cell are **polar microtubules;** polar microtubules originating in opposite centrosomes interdigitate near the cell's equator. Finally, short

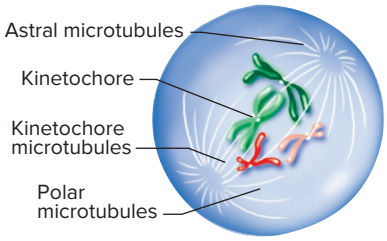
Figure 4.10 Mitosis maintains the chromosome number of the parent cell in the two daughter nuclei. In the photomicrographs of newt lung cells at the left, chromosomes are stained *blue* and microtubules appear either *green* or *yellow*. Note that the drawings are of *Ascaris* cells ($2n = 4$).



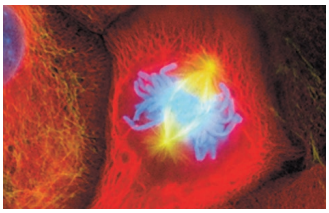


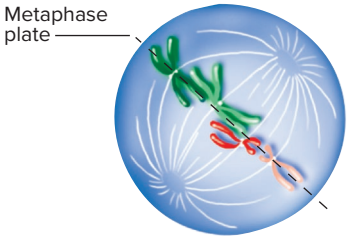
(a) Prophase: (1) Chromosomes condense and become visible; (2) centrosomes move apart toward opposite poles and generate new microtubules; (3) nucleoli begin to disappear.



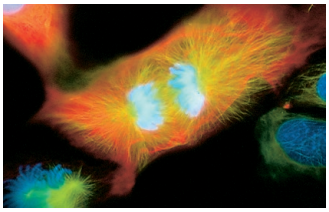


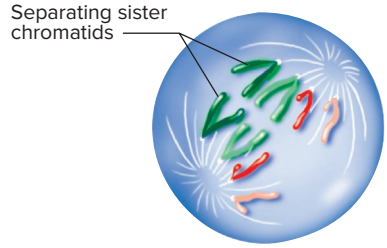
(b) Prometaphase: (1) Nuclear envelope breaks down; (2) microtubules from the centrosomes invade the nucleus; (3) sister chromatids attach to microtubules from opposite centrosomes.



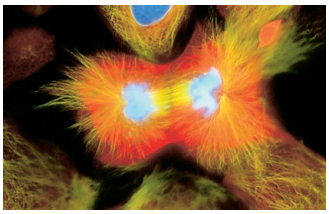


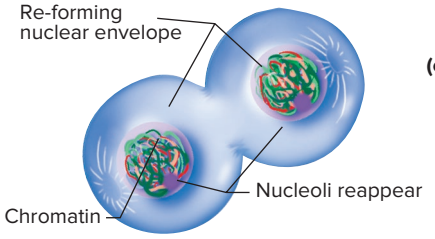
(c) Metaphase: Chromosomes align on the metaphase plate with sister chromatids facing opposite poles.



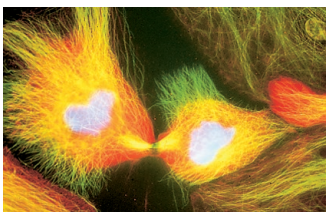


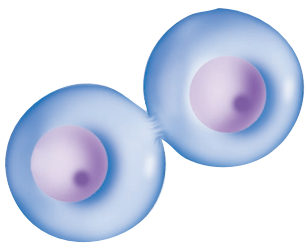
(d) Anaphase: (1) The connection between the centromeres of the sister chromatids is severed; (2) the now separated sister chromatids move to opposite poles.





(e) Telophase: (1) Nuclear membranes and nucleoli re-form; (2) spindle fibers disappear; (3) chromosomes uncoil and become a tangle of chromatin.





(f) Cytokinesis: The cytoplasm divides, splitting the elongated parent cell into two daughter cells with identical nuclei.

astral microtubules extend out from the centrosome toward the cell's periphery.

Soon before the end of prometaphase, the kinetochore of each chromosome's previously unattached sister chromatid now associates with microtubules extending from the opposite centrosome. This event orients each chromosome such that one sister chromatid faces one pole of the cell and the other faces the opposite pole. Experimental manipulation has shown that if both kinetochores become attached to microtubules from the same pole, the configuration is unstable; one of the kinetochores will detach repeatedly from the spindle until it associates with microtubules from the other pole. The attachment of sister chromatids to opposite spindle poles is the only stable arrangement.

Metaphase: Chromosomes align at the cell's equator (Fig. 4.10c)

During **metaphase** (*middle stage*), the connection of sister chromatids to opposite spindle poles sets in motion a series of jostling movements that cause the chromosomes to move toward an imaginary equator halfway between the two poles. The imaginary midline is called the **metaphase plate.** When the chromosomes are aligned along it, the forces pulling sister chromatids toward opposite poles are in a balanced equilibrium maintained by tension across the chromosomes. Tension results from the fact that the sister chromatids are pulled in opposite directions while they are still connected to each other by the tight cohesion of their centromeres. Tension compensates for any chance movement away from the metaphase plate by restoring the chromosome to its position equidistant between the poles.

Anaphase: Sister chromatids move to opposite spindle poles (Fig. 4.10d)

The nearly simultaneous severing of the centromeric connections between the sister chromatids of all chromosomes indicates that **anaphase** (from the Greek *ana-* meaning *up* as in *up toward the poles*) is underway. The separation of sister chromatids allows each chromatid to be pulled toward the spindle pole to which it is linked by kinetochore microtubules; as the chromatid moves toward the pole, its kinetochore microtubules shorten. Because the arms of the chromatids lag behind the kinetochores, metacentric chromatids have a characteristic V shape during anaphase. The attachment of sister chromatids to microtubules emanating from opposite spindle poles means that the genetic information migrating toward one pole is exactly the same as its counterpart moving toward the opposite pole.

Telophase: Identical sets of chromosomes are enclosed in two nuclei (Fig. 4.10e)

The final transformation of chromosomes and the nucleus during mitosis happens at **telophase** (from the Greek *telo-* meaning *end*). Telophase is like a rewind of prophase. The spindle fibers begin to disperse; a nuclear envelope forms around the group of chromatids at each pole; and one or more nucleoli reappear. The former chromatids now function as independent chromosomes, which decondense (uncoil) and dissolve into a tangled mass of chromatin. Mitosis, the division of one nucleus into two identical nuclei, is over.

Cytokinesis: The cytoplasm divides (Fig. 4.10f)

In the final stage of cell division, the daughter nuclei emerging at the end of telophase are packaged into two separate daughter cells. This final stage of division is called **cytokinesis** (literally *cell movement*). During cytokinesis, the elongated parent cell separates into two smaller independent daughter cells with identical nuclei. Cytokinesis usually begins during anaphase, but it is not completed until after telophase.

The mechanism by which cells accomplish cytokinesis differs in animals and plants. In animal cells, cytoplasmic division depends on a **contractile ring** that pinches the cell into two approximately equal halves, similar to the way the pulling of a string closes the opening of a bag of marbles (**Fig. 4.11a**). Intriguingly, some types of molecules that form the contractile ring also participate in the mechanism responsible for muscle contraction. In plants, whose cells are surrounded by a rigid cell wall, a membrane-enclosed disk, known as the *cell plate,* forms inside the cell near the equator and then grows rapidly outward, thereby dividing the cell in two (**Fig. 4.11b**).

During cytokinesis, a large number of important organelles and other cellular components, including ribosomes, mitochondria, membranous structures such as Golgi

Figure 4.11 Cytokinesis: The cytoplasm divides, producing two daughter cells. (a) In this dividing frog zygote, the contractile ring at the cell's periphery has contracted to form a *cleavage furrow* that will eventually pinch the cell in two. **(b)** In this dividing onion root cell, a cell plate that began forming near the equator of the cell expands to the periphery, separating the two daughter cells.
a: © Don W. Fawcett/Science Source; b: © McGraw-Hill Education/Al Telser

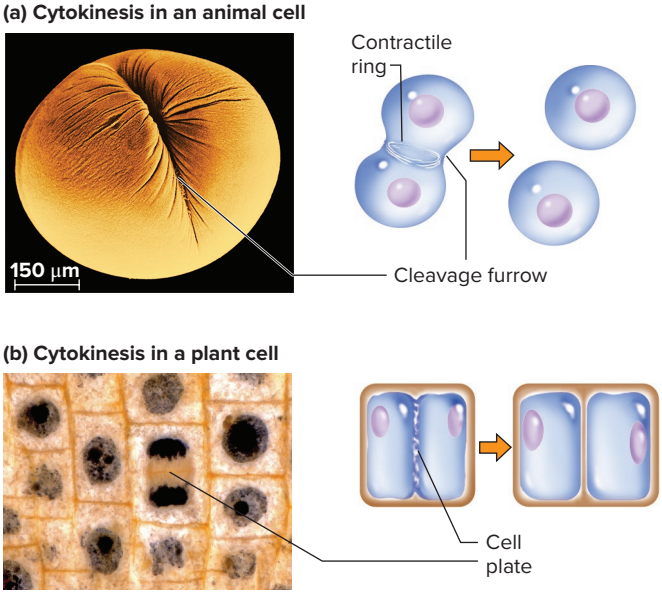
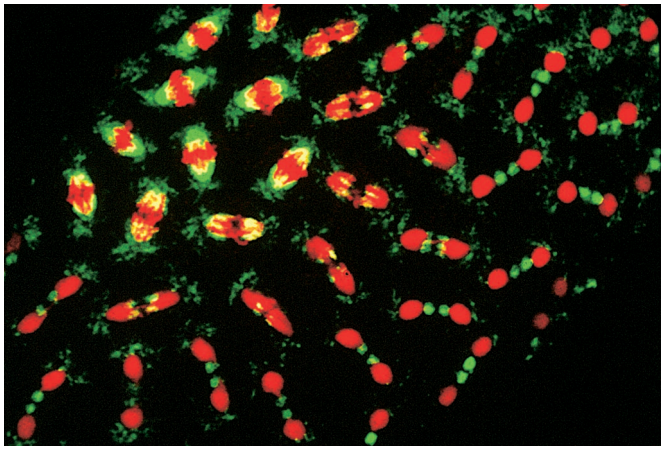


Figure 4.12 If cytokinesis does not follow mitosis, one cell may contain many nuclei. In fertilized *Drosophila* eggs, 13 rounds of mitosis take place without cytokinesis. The result is a single-celled syncytial embryo that contains several thousand nuclei. The photograph shows part of an embryo with dividing nuclei; chromosomes are in *red,* and spindle fibers are in *green.* Nuclei at the *upper left* are in metaphase, while nuclei toward the *bottom right* are progressively later in anaphase. Membranes eventually grow around these nuclei, dividing the embryo into cells.



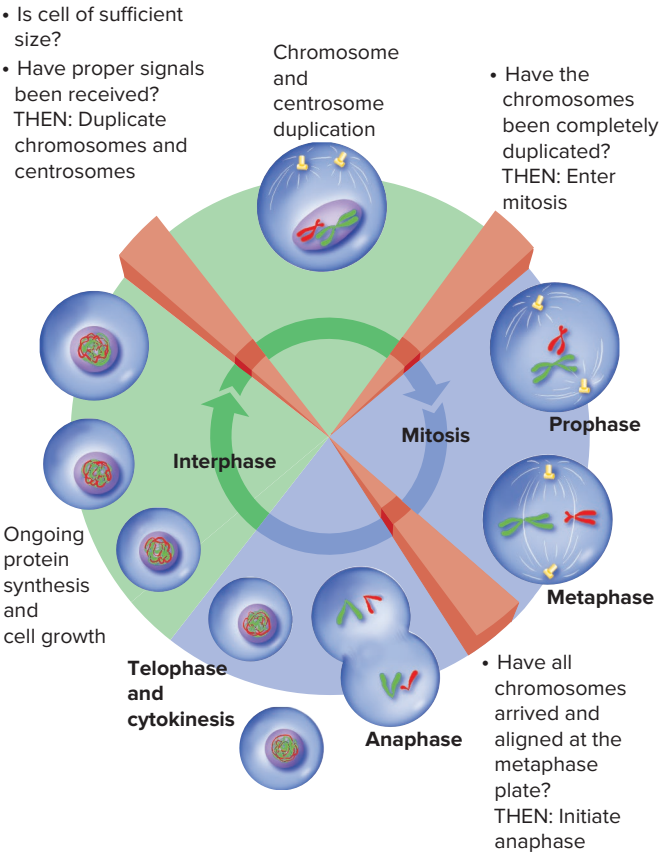
bodies, and (in plants) chloroplasts, must be parceled out to the emerging daughter cells. The mechanism accomplishing this task does not appear to predetermine which organelle is destined for which daughter cell. Instead, because most cells contain many copies of these cytoplasmic structures, each new cell is bound to receive at least a few representatives of each component. This original complement of structures is enough to sustain the cell until synthetic activity can repopulate the cytoplasm with organelles.

Sometimes cytoplasmic division does not immediately follow nuclear division, and the result is a cell containing more than one nucleus. An animal cell with two or more nuclei is known as a **syncytium.** The early embryos of fruit flies are multinucleated syncytia (**Fig. 4.12**), as are the precursors of spermatozoa in humans and many other animals. A multinucleate plant tissue is called a **coenocyte;** coconut milk is a nutrient-rich food composed of coenocytes.

Regulatory Checkpoints Ensure Correct Chromosome Separation

The cell cycle is a complex sequence of precisely coordinated events. In higher organisms, a cell's "decision" to divide depends on both intrinsic factors, such as conditions within the cell that register a sufficient size for division, and signals from the environment, such as hormonal cues or contacts with neighboring cells that encourage or restrain division. Once a cell has initiated events leading to division, usually during the G_1 period of interphase, everything else follows like clockwork. A number of **checkpoints**—moments at

Figure 4.13 Checkpoints help regulate the cell cycle. Cellular checkpoints (*red wedges*) ensure that important events in the cell cycle occur in the proper sequence. At each checkpoint, the cell determines whether prior events have been completed before it can proceed to the next step of the cell cycle. (For simplicity, we show only two chromosomes per cell.)



which the cell evaluates the results of previous steps—allow the sequential coordination of cell-cycle events (**Fig. 4.13**). For example, the enzymes operating in one type of checkpoint monitor DNA replication to ensure that cells do not begin mitosis until all the chromosomes have been completely copied. If this checkpoint did not exist, at least one of the daughter cells would lose DNA every cell cycle.

In a second illustration of the molecular basis of checkpoints, even a single kinetochore that has not attached to spindle fibers generates a molecular signal that prevents the sister chromatids of all chromosomes from separating at their centromeres. This signal makes the beginning of anaphase dependent on the prior proper alignment of all the chromosomes at metaphase. As a result of this cell-cycle checkpoint, each daughter cell reliably receives the right number of chromosomes.

Breakdown of the mitotic machinery can produce division mistakes that have crucial consequences for the cell. Improper chromosome segregation, for example, can cause serious malfunction or even the death of daughter cells. Gene mutations that disrupt mitotic structures, such as the

spindle, kinetochores, or centrosomes, are one source of improper segregation. Other problems occur in cells where the normal restraints on cell division, such as checkpoints, have broken down. Such cells may divide uncontrollably, leading to a tumor. We present the details of cell-cycle regulation, checkpoint controls, and cancer formation in Chapter 20.

essential concepts

- Through *mitosis,* diploid cells produce identical diploid progeny cells.
- At *metaphase,* the *sister chromatids* are being pulled at their kinetochores toward opposite spindle poles; these poleward forces are balanced because the chromatids are connected at their centromeres.
- At the beginning of *anaphase,* the connections between sister centromeres are severed so sister chromatids separate and move to opposite spindle poles.
- Cell cycle *checkpoints* help ensure correct duplication and separation of chromosomes.

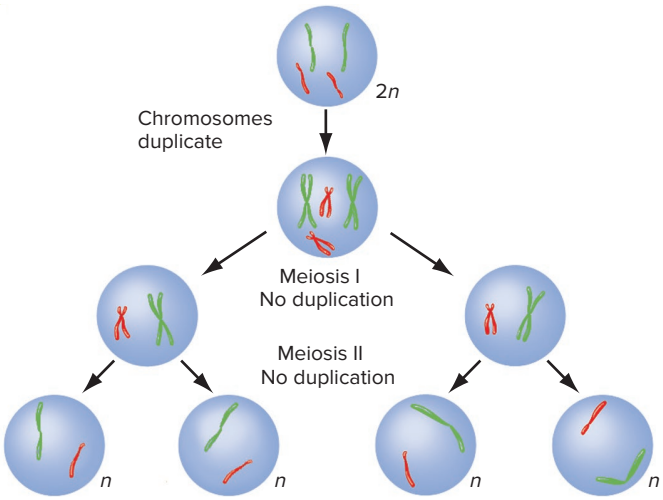
4.4 Meiosis: Cell Divisions That Halve Chromosome Number

learning objectives

1. Describe the key chromosome behaviors during meiosis that lead to haploid gametes.
2. Compare chromosome behaviors during mitosis and meiosis.
3. Explain how the independent alignment of homologs, and also crossing-over during the first meiotic division, each contribute to the genetic diversity of gametes.

During the many rounds of cell division within an embryo, most cells either grow and divide via the mitotic cell cycle just described, or they stop growing and become arrested in G_0. These mitotically dividing and G_0-arrested cells are the so-called **somatic cells** whose descendants continue to make up the vast majority of each organism's tissues throughout the lifetime of the individual. Early in the embryonic development of animals, however, a group of cells is set aside for a different fate. These are the **germ cells:** cells destined for a specialized role in the production of gametes. Germ cells arise later in plants, during floral development instead of during embryogenesis. The germ cells become incorporated in the reproductive organs—ovaries and testes in animals, ovaries and anthers in flowering plants—where they ultimately undergo meiosis, the special two-part cell division that produces gametes (eggs and sperm) containing half the number of chromosomes other body cells have.

Figure 4.14 An overview of meiosis: The chromosomes replicate once, while the nuclei divide twice. In this figure, all four chromatids of each chromosome pair are shown in the same shade of the same color. Note that the chromosomes duplicate before meiosis I, but they do not duplicate between meiosis I and meiosis II.



The union of haploid gametes at fertilization yields diploid offspring that carry the combined genetic heritage of two parents. Sexual reproduction therefore requires the alternation of haploid and diploid generations of cells. If gametes were diploid rather than haploid, the number of chromosomes would double in each successive generation. In humans, for example, the children would have 92 chromosomes per cell, the grandchildren 184, and so on. Meiosis prevents this lethal, exponential accumulation of chromosomes.

In Meiosis, the Chromosomes Replicate Once but the Nucleus Divides Twice

Unlike mitosis, meiosis consists of two successive nuclear divisions, logically named *division I of meiosis* and *division II of meiosis*, or simply **meiosis I** and **meiosis II.** With each round, the cell passes through a prophase, metaphase, anaphase, and telophase. In meiosis I, the parent nucleus divides to form two daughter nuclei; in meiosis II, each of the two daughter nuclei divides, resulting in four nuclei (**Fig. 4.14**). These four nuclei—the final products of meiosis—become partitioned in four separate daughter cells because cytokinesis occurs after both rounds of division. The chromosomes duplicate at the start of meiosis I, but they do not duplicate in meiosis II, which explains why the gametes contain half the number of chromosomes found in somatic cells. A close look at each round of meiotic division reveals the mechanisms by which each gamete comes to receive one full haploid set of chromosomes.

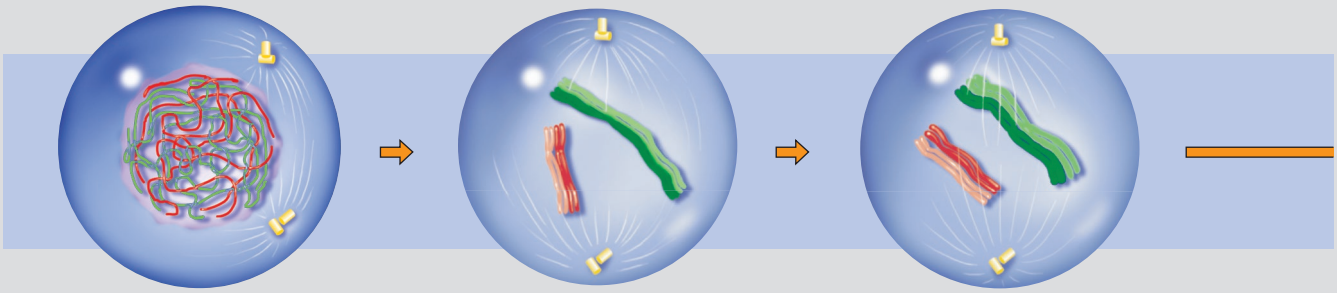
During Meiosis I, Homologs Pair, Exchange Parts, and Then Segregate

The events of meiosis I are unique among nuclear divisions (**Fig. 4.15,** meiosis I). The process begins with the replication

FEATURE FIGURE 4.15

Meiosis: One Diploid Cell Produces Four Haploid Cells

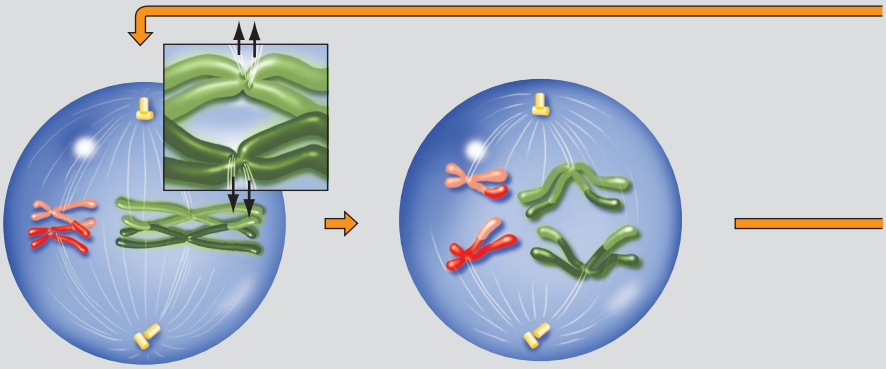
Meiosis I: A reductional division



Prophase I: Leptotene
1. Chromosomes thicken and become visible, but the chromatids remain invisible.
2. Centrosomes begin to move toward opposite poles.

Prophase I: Zygotene
1. Homologous chromosomes enter *synapsis*.
2. The *synaptonemal complex* forms.

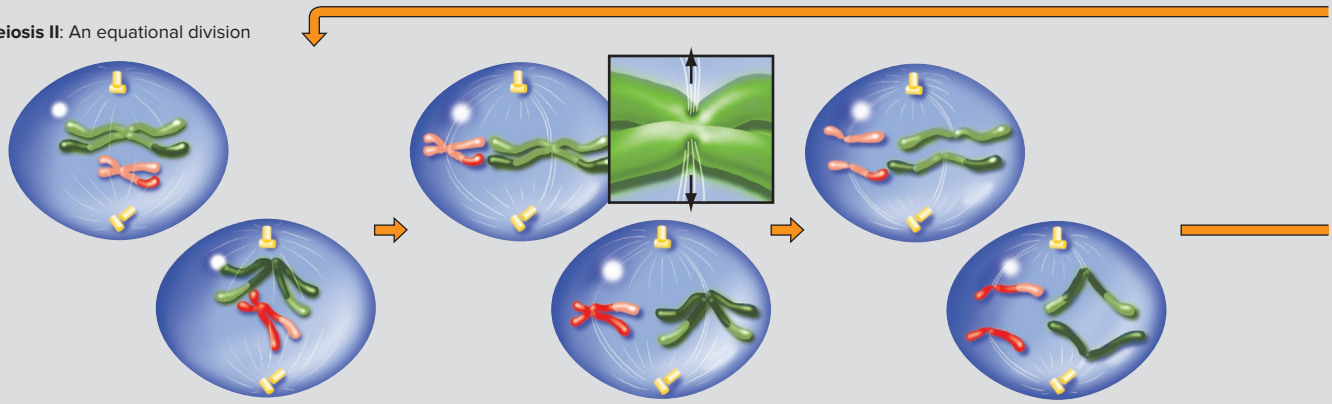
Prophase I: Pachytene
1. Synapsis is complete.
2. *Crossing-over*, genetic exchange between nonsister chromatids of a homologous pair, occurs.



Metaphase I
1. Tetrads line up along the *metaphase plate*.
2. Each chromosome of a homologous pair attaches to fibers from opposite poles.
3. Sister chromatids attach to fibers from the same pole.

Anaphase I
1. Sister centromeres remain connected to each other.
2. The chiasmata dissolve.
3. Homologous chromosomes move to opposite poles.

Meiosis II: An equational division



Prophase II
1. Chromosomes condense.
2. Centrioles move toward the poles.
3. The nuclear envelope breaks down at the end of prophase II (*not shown*).

Metaphase II
1. Chromosomes align at the metaphase plate.
2. Sister chromatids attach to spindle fibers from opposite poles.

Anaphase II
1. Sister centromeres detach from each other, allowing sister chromatids to move to opposite poles.

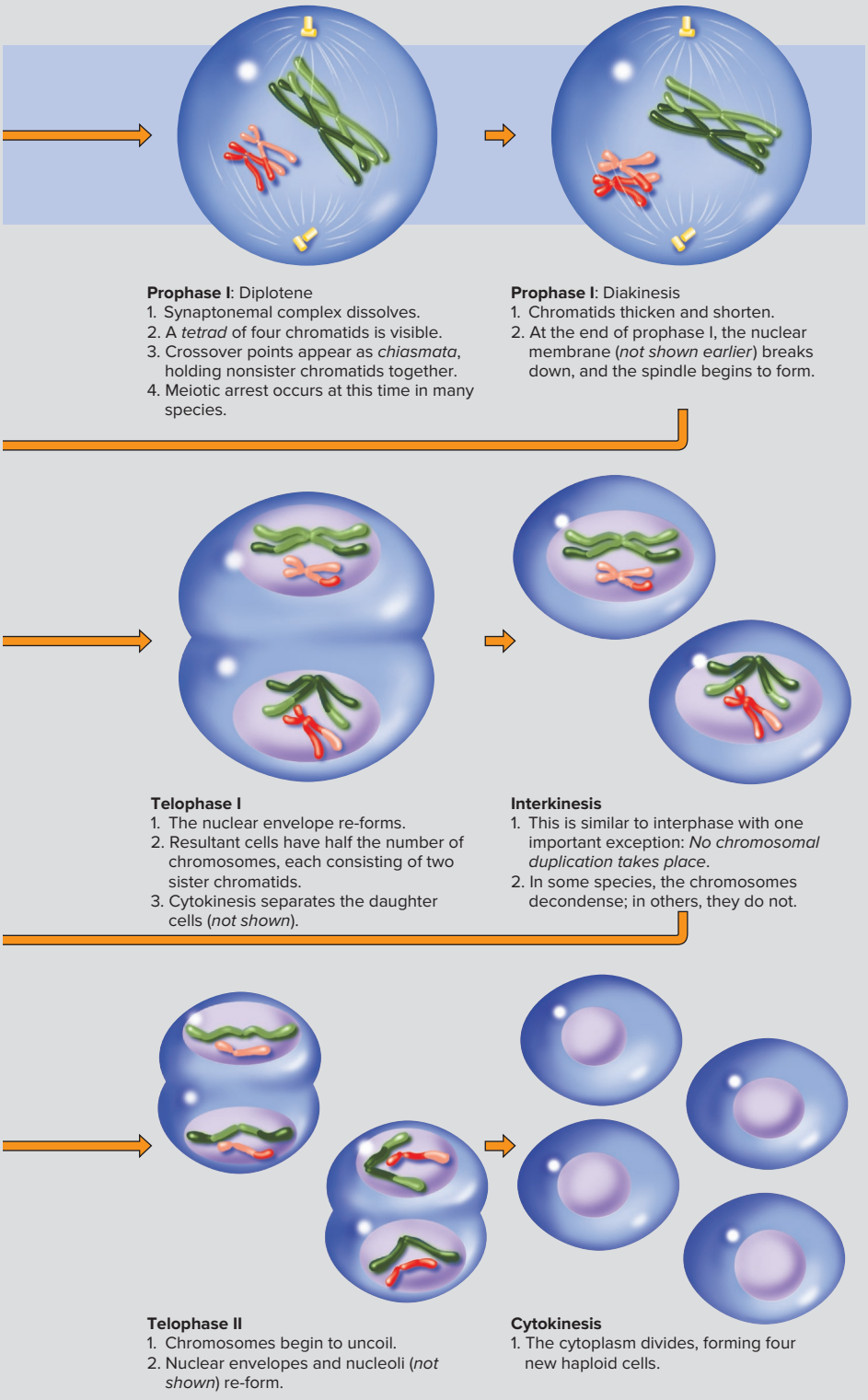


Figure 4.15 To aid visualization of the chromosomes, the figure is simplified in two ways: (1) The nuclear envelope is not shown during prophase of either meiotic division. (2) The chromosomes are shown as fully condensed at zygotene; in reality, full condensation is not achieved until diakinesis.

of chromosomes, after which each one consists of two sister chromatids. A key to understanding meiosis I is the observation that the centromeres of these sister chromatids remain connected throughout the entire division, rather than separating from each other as in mitosis.

As meiosis I proceeds, homologous chromosomes align across the cellular equator to form a coupling that ensures proper chromosome segregation later in the division. Moreover, during the time homologous chromosomes face each other across the equator, the maternal and paternal chromosomes of each homologous pair may exchange parts, creating new combinations of alleles at different genes along the chromosomes. Afterward, the two homologous chromosomes, each still consisting of two sister chromatids connected at their centromeres, are pulled to opposite poles of the spindle. As a result, it is homologous chromosomes (rather than sister chromatids as in mitosis) that segregate into different daughter cells at the conclusion of the first meiotic division. With this overview in mind, let us take a closer look at the specific events of meiosis I, remembering that we analyze a dynamic, flowing sequence of cellular events by breaking it down somewhat arbitrarily into the easily pictured, traditional phases.

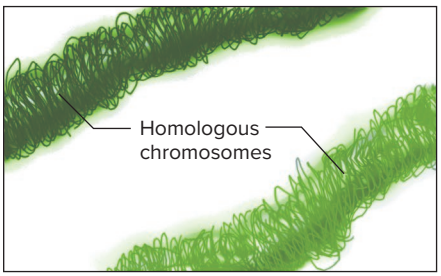
Prophase I: Homologs condense and pair, and crossing-over occurs

Among the crucial events of **prophase I** are the condensation of chromatin, the pairing of homologous chromosomes, and the reciprocal exchange of genetic information between these paired homologs. Figure 4.15 shows a generalized view of prophase I; however, research suggests that the exact sequence of events may vary in different species. These complicated processes can take many days, months, or even years to complete. For example, in the female germ cells of several species, including humans, meiosis is suspended at prophase I for many years until ovulation (as will be discussed further in Section 4.5).

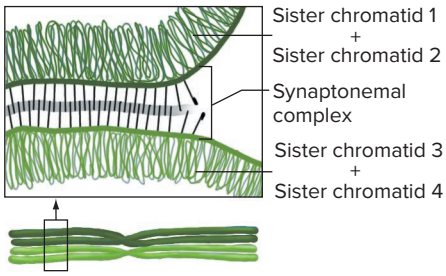
Leptotene (from the Greek for *thin* and *delicate*) is the first definable substage of prophase I, the time when the long, thin chromosomes begin to thicken (see **Fig. 4.16a** for a more detailed view). Each chromosome has already duplicated prior to prophase I (as in mitosis) and thus consists of two sister chromatids affixed at their centromeres. At this point, however, these sister chromatids are so tightly bound together that they are not yet visible as separate entities.

Zygotene (from the Greek for *conjugation*) begins as each chromosome seeks out its homologous partner and the matching chromosomes become zipped together in a

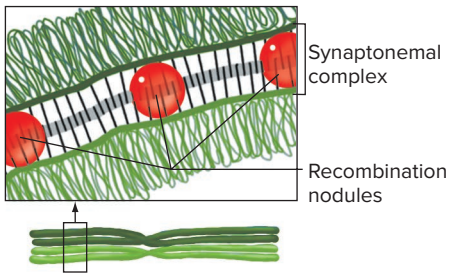
Figure 4.16 Prophase I of meiosis at very high magnification.



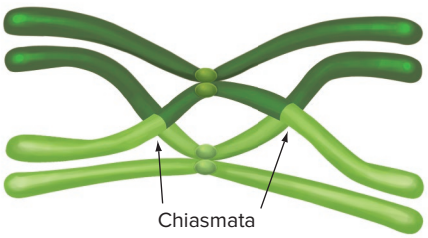
(a) Leptotene: Threadlike chromosomes begin to condense and thicken, becoming visible as discrete structures. Although the chromosomes have duplicated, the sister chromatids of each chromosome are not yet visible in the microscope.



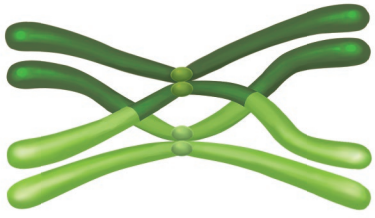
(b) Zygotene: Chromosomes are clearly visible and begin pairing with homologous chromosomes along the synaptonemal complex to form a bivalent, or tetrad.



(c) Pachytene: Full synapsis of homologs. Recombination nodules appear along the synaptonemal complex.



(d) Diplotene: Bivalent pulls apart slightly, but homologous chromosomes remain connected due to recombination at crossover sites (chiasmata).



(e) Diakinesis: Further condensation of the bivalent.

process known as **synapsis.** The "zipper" itself is an elaborate protein structure called the **synaptonemal complex** that aligns the homologs with remarkable precision, juxtaposing the corresponding genetic regions of the chromosome pair (**Fig. 4.16b**).

Pachytene (from the Greek for *thick* or *fat*) begins at the completion of synapsis when homologous chromosomes are united along their length. Each synapsed chromosome pair is known as a **bivalent** (because it encompasses two chromosomes), or a **tetrad** (because it contains four chromatids). On one side of the bivalent is a maternally derived chromosome, on the other side a paternally derived one. Because X and Y chromosomes are not identical, they do not synapse completely. However, the pseudoautosomal regions previously shown in Fig. 4.8 provide small stretches of similarity between the X and the Y chromosomes that allow them to pair with each other during meiosis I in males.

During pachytene, structures called **recombination nodules** begin to appear along the synaptonemal complex, and an exchange of parts between nonsister (that is, between maternal and paternal) chromatids occurs at these nodules (see **Fig. 4.16c** for details). Such an exchange is known as **crossing-over;** it results in the **recombination** of genetic material. As a result of crossing-over, chromatids may no longer be of purely maternal or paternal origin; however, no genetic information is gained or lost, so all chromatids retain their original size.

Diplotene (from the Greek for *twofold* or *double*) is signaled by the gradual dissolution of the synaptonemal zipper complex and a slight separation of regions of the homologous chromosomes (see **Fig. 4.16d**). The aligned homologous chromosomes of each bivalent nonetheless remain very tightly merged at intervals along their length called **chiasmata** (singular, *chiasma*), which represent the sites where crossing-over occurred.

Diakinesis (from the Greek for *double movement*) is accompanied by further condensation of the chromatids. Because of this chromatid thickening and shortening, it can now clearly be seen that each tetrad consists of four separate chromatids, or viewed in another way, that the two homologous chromosomes of a bivalent are each composed of two sister chromatids held together at their centromeres (see **Fig. 4.16e**). Nonsister chromatids that have undergone crossing-over remain closely associated at chiasmata. The end of diakinesis is analogous to the prometaphase of mitosis: The nuclear envelope breaks down, and the microtubules of the spindle apparatus begin to form.

Metaphase I: Paired homologs attach to spindle fibers from opposite poles

During mitosis, each sister chromatid has a kinetochore that becomes attached to microtubules emanating from opposite spindle poles. During meiosis I, the situation is different. The kinetochores of sister chromatids fuse, so that each chromosome contains only a single functional kinetochore. During **metaphase I** (see Fig. 4.15, meiosis I), it is the kinetochores of homologous chromosomes that attach to microtubules from opposite spindle poles. As a result, in chromosomes aligned at the metaphase plate, the kinetochores of maternally and paternally derived chromosomes are subject to pulling forces from opposite spindle poles, balanced by the physical connections between homologs at chiasmata. Each bivalent's alignment and hookup is independent of that of every other bivalent, so the chromosomes facing each pole are a random mix of maternal and paternal origin.

Anaphase I: Homologs move to opposite spindle poles

At the onset of **anaphase I,** the chiasmata joining homologous chromosomes dissolve, which allows the maternal and paternal homologs to begin to move toward opposite spindle poles (see Fig. 4.15, meiosis I). Note that in anaphase of the first meiotic division, the sister centromeres do not separate as they do in mitosis. Thus, from each homologous pair, one chromosome consisting of two sister chromatids joined at their centromeres segregates to each spindle pole.

Recombination through crossing-over plays an important role in the proper segregation of homologous chromosomes during the first meiotic division. The chiasmata hold the homologs together and thus ensure that their kinetochores remain attached to opposite spindle poles throughout metaphase. When recombination does not occur within a bivalent, mistakes in hookup and conveyance may cause homologous chromosomes to move to the same pole, instead of segregating to opposite poles. In some organisms, however, proper segregation of nonrecombinant chromosomes nonetheless occurs through other pairing mechanisms. Investigators do not yet completely understand the nature of these processes, and they are currently evaluating several models to explain them.

Telophase I: Nuclear envelopes re-form

The telophase of the first meiotic division, or **telophase I,** takes place when nuclear membranes begin to form around the chromosomes that have moved to the poles. Each of the incipient daughter nuclei contains one-half the number of chromosomes in the original parent nucleus, but each chromosome consists of two sister chromatids joined at their centromeres (see Fig. 4.15, meiosis I). Because the number of chromosomes is reduced to one-half the normal diploid number, meiosis I is often called a **reductional division.**

In most species, cytokinesis follows telophase I, with daughter nuclei becoming enclosed in separate daughter cells. A short interphase then ensues. During this time, the chromosomes usually decondense, in which case they must recondense during the prophase of the subsequent second

meiotic division. In some species, however, the chromosomes simply stay condensed. Most importantly, no S phase exists during the interphase between meiosis I and meiosis II; that is, the chromosomes do not replicate during meiotic interphase. The relatively brief interphase between meiosis I and meiosis II is known as **interkinesis.**

During Meiosis II, Sister Chromatids Separate to Produce Haploid Gametes

The second meiotic division (meiosis II) proceeds in a fashion very similar to that of mitosis, but because the number of chromosomes in each dividing nucleus has already been reduced by half, the resulting daughter cells are haploid. The same process occurs in each of the two daughter cells generated by meiosis I, producing four haploid cells at the end of this second meiotic round (see Fig. 4.15, meiosis II).

Prophase II: The chromosomes condense

If the chromosomes decondensed during the preceding interphase, they recondense during **prophase II.** At the end of prophase II, the nuclear envelope breaks down, and the spindle apparatus re-forms.

Metaphase II: Chromosomes align at the metaphase plate

The kinetochores of sister chromatids attach to microtubule fibers emanating from opposite poles of the spindle apparatus, just as in mitotic metaphase. Nonetheless, two significant features of **metaphase II** distinguish it from mitosis. First, the number of chromosomes is one-half that in mitotic metaphase of the same species. Second, in most chromosomes, the two sister chromatids are no longer strictly identical because of the recombination through crossing-over that occurred during meiosis I. The sister chromatids still contain the same genes, but they may carry different combinations of alleles.

Anaphase II: Sister chromatids move to opposite spindle poles

Just as in mitosis, severing of the connection between sister centromeres allows the sister chromatids to move toward opposite spindle poles during **anaphase II.**

Telophase II: Nuclear membranes re-form, and cytokinesis follows

Membranes form around each of four daughter nuclei in **telophase II,** and cytokinesis places each nucleus in a separate cell. The result is four haploid gametes. Note that at the end of meiosis II, each daughter cell (that is, each gamete) has the same number of chromosomes as the parental cell present at the beginning of this division. For this reason, meiosis II is termed an **equational division.**

Mistakes in Meiosis Produce Defective Gametes

Segregational errors during either meiotic division can lead to aberrations, such as trisomies, in the next generation. If, for example, the homologs of a chromosome pair do not segregate during meiosis I (a mistake known as **nondisjunction**), they may travel together to the same pole and eventually become part of the same gamete. Such an error may at fertilization result in any one of a large variety of possible trisomies. Most autosomal trisomies in humans, as we already mentioned, are lethal *in utero;* one exception is trisomy 21, the genetic basis of Down syndrome. Like trisomy 21, extra sex chromosomes may also be viable but cause a variety of mental and physical abnormalities, such as those seen in Klinefelter syndrome (see Table 4.1).

Meiosis Contributes to Genetic Diversity

The wider the assortment of different gene combinations among members of a species, the greater the chance that at least some individuals will carry combinations of alleles that allow survival in a changing environment. Two aspects of meiosis contribute to genetic diversity in a population. First, because only chance governs which paternal or maternal homologs migrate to the two poles during the first meiotic division, different gametes carry a different mix of maternal and paternal chromosomes. **Figure 4.17a** shows how different patterns of homolog migration produce different mixes of parental chromosomes in the gametes. The amount of potential variation generated by this random independent assortment increases with the number of chromosomes. In *Ascaris,* for example, where $n = 2$ (the chromosome complement shown in Fig. 4.17a), the random assortment of homologs could produce only 2^2, or four types of gametes. In a human being, however, where $n = 23$, this mechanism alone could generate 2^{23}, or more than 8 million genetically different kinds of gametes.

A second feature of meiosis, the reshuffling of genetic information through crossing-over during prophase I, ensures an even greater amount of genetic diversity in gametes. Because crossing-over recombines maternally and paternally derived genes, each chromosome in each different gamete could consist of different combinations of maternal and paternal alleles (**Fig. 4.17b**).

Of course, sexual reproduction adds yet another means of producing genetic diversity. At fertilization, any one of a vast number of genetically diverse sperm can fertilize an egg with its own distinctive genetic constitution. It is thus not very surprising that, with the exception of identical twins, the 6 billion people in the world are each genetically unique.

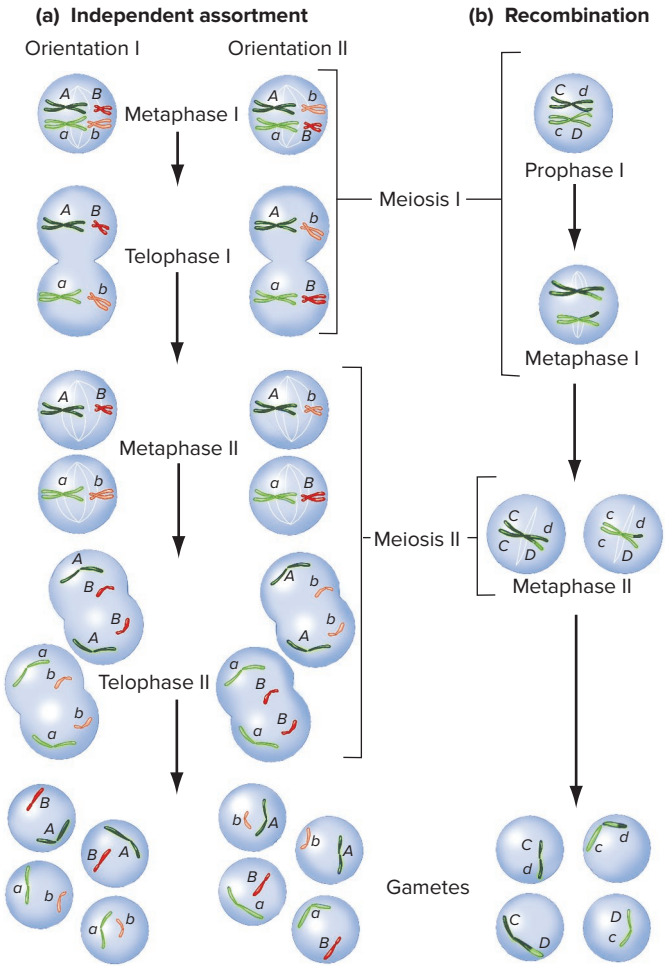


Figure 4.17 How meiosis contributes to genetic diversity. (a) The variation resulting from the independent assortment of nonhomologous chromosomes increases with the number of chromosomes in the genome. **(b)** Crossing-over between homologous chromosomes ensures that each gamete is unique.

Mitosis and Meiosis: A Comparison

Mitosis occurs in all types of eukaryotic cells (that is, cells with a membrane-bounded nucleus) and is a conservative mechanism that preserves the genetic *status quo.* Mitosis followed by cytokinesis produces growth by increasing the number of cells. It also promotes the continual replacement of roots, stems, and leaves in plants and the regeneration of blood cells, intestinal tissues, and skin in animals.

Meiosis, on the other hand, occurs only in sexually reproducing organisms, in just a few specialized germ cells within the reproductive organs that produce haploid gametes. It is not a conservative mechanism; rather, the extensive combinatorial changes arising from meiosis are one source of the genetic variation that fuels evolution. **Table 4.3** illustrates the significant contrasts between the two mechanisms of cell division.

essential concepts

- In *meiosis,* chromosomes replicate once (before meiosis I), but the nucleus divides twice (meiosis I and II).
- During *metaphase I,* homologous chromosomes connect to opposite spindle poles. The independent alignment of each pair of homologs ensures the independent assortment of genes carried on different chromosomes.
- *Crossing-over* during the first meiotic division maintains the connection between homologous chromosomes until *anaphase I* and contributes to the genetic diversity of gametes.
- Sister chromatids separate from each other during *meiosis II* so that gametes have only one copy of each chromosome.
- *Fertilization*—the union of egg and sperm—restores the diploid number of chromosomes (2*n*) to the zygote.
- Errors during meiosis may produce gametes with missing or extra chromosomes, which often is lethal to offspring.

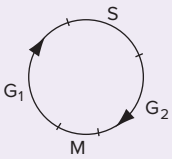
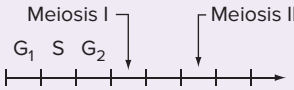
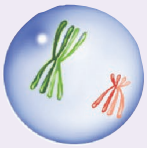
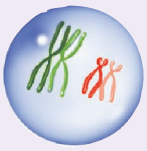
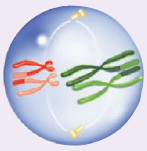
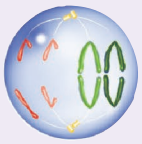
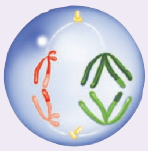
4.5 Gametogenesis

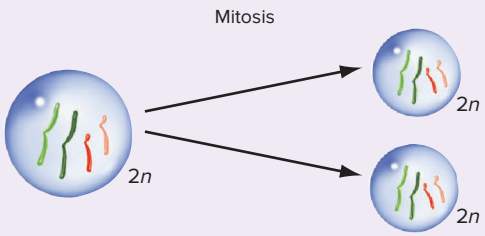
learning objectives

1. Compare the processes of oogenesis and spermatogenesis in humans.
2. Distinguish between the sex chromosome complements of human female and male germ-line cells at different stages of gametogenesis.

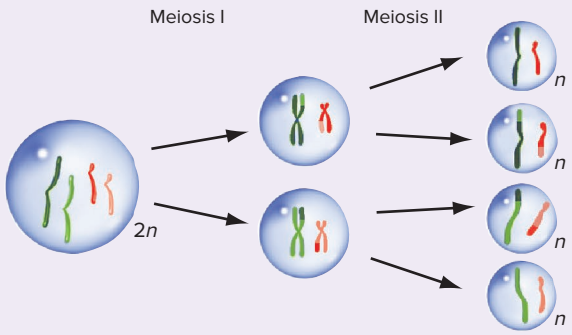
In all sexually reproducing animals, the embryonic germ cells (collectively known as the **germ line**) undergo a series of mitotic divisions that yield a collection of specialized diploid cells, which subsequently divide by meiosis to produce haploid cells. As with other biological processes, many variations on this general pattern have been observed. In some species, the haploid cells resulting from meiosis are the gametes themselves, while in other species, those cells must undergo a specific plan of differentiation to fulfill that function. Moreover, in certain organisms, the four haploid products of a single meiosis do not all become gametes. Gamete formation, or **gametogenesis,** thus gives rise to haploid gametes marked not only by the events of meiosis *per se* but also by cellular events that precede and follow meiosis. Here we illustrate gametogenesis with a description of egg and sperm formation in humans. The details of gamete formation in several other organisms appear throughout the book in discussions of specific experimental studies.

TABLE 4.3 Comparing Mitosis and Meiosis

Mitosis	Meiosis
Occurs in somatic cells and germ-line precursor cells Haploid and diploid cells can undergo mitosis One round of division	Occurs in germ cells as part of the sexual cycle Two rounds of division, meiosis I and meiosis II Only diploid cells undergo meiosis
 Mitosis is preceded by S phase (chromosome duplication).	 Chromosomes duplicate prior to meiosis I but not before meiosis II.
 Homologous chromosomes do not pair.	 During prophase of meiosis I, homologous chromosomes pair (synapse) along their length.
Genetic exchange between homologous chromosomes is very rare.	 Crossing-over occurs between homologous chromosomes during prophase of meiosis I.
 Sister chromatids attach to spindle fibers from opposite poles during metaphase.	 Homologous chromosomes (not sister chromatids) attach to spindle fibers from opposite poles during metaphase I.
 The centromeres of the sister chromatids separate at the beginning of anaphase.	 The centromeres of the sister chromatids remain tightly attached during meiosis I.
	 Sister chromatids attach to spindle fibers from opposite poles during metaphase II.
	 The centromeres of the sister chromatids separate at the beginning of anaphase II.

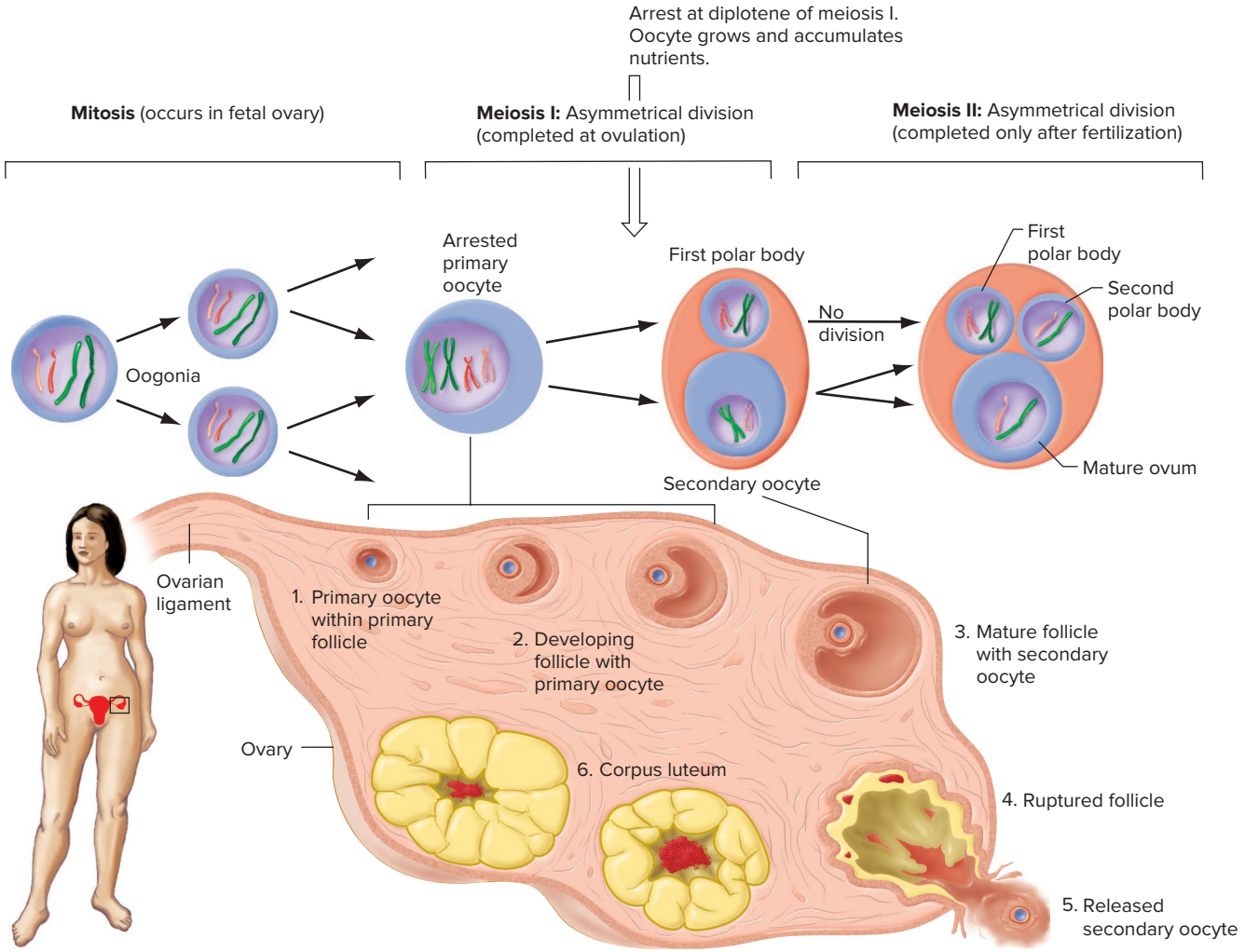


Mitosis produces two new daughter cells, identical to each other and the original cell. Mitosis is thus genetically conservative.



Meiosis produces four haploid cells, one (egg) or all (sperm) of which can become gametes. None of these is identical to each other or to the original cell, because meiosis results in combinatorial change.

Figure 4.18 In humans, egg formation begins in the fetal ovaries and arrests during the prophase of meiosis I. Fetal ovaries contain about 500,000 primary oocytes arrested in the diplotene substage of meiosis I. If the egg released during a menstrual cycle is fertilized, meiosis is completed. Only one of the three cells produced by meiosis serves as the functional gamete, or ovum.



Oogenesis in Humans Produces One Ovum from Each Primary Oocyte

The end product of egg formation in humans is a large, nutrient-rich **ovum** whose stored resources can sustain the early embryo. The process, known as **oogenesis** (**Fig. 4.18**), begins when diploid germ cells in the ovary, called **oogonia** (singular, *oogonium*), multiply rapidly by mitosis and produce a large number of **primary oocytes,** which then undergo meiosis.

For each primary oocyte, meiosis I results in the formation of two daughter cells that differ in size, so this division is asymmetric. The larger of these cells, the **secondary oocyte,** receives over 95% of the cytoplasm. The other small sister cell is known as the first **polar body.** During meiosis II, the secondary oocyte undergoes another asymmetrical division to produce a large haploid **ovum** and a small, haploid second polar body. The first polar body usually arrests its development. The two small polar bodies apparently serve no function and disintegrate, leaving one large haploid ovum as the functional gamete. Thus, only one of the three (or rarely, four) products of a single meiosis serves as a female gamete. A normal human ovum carries 22 autosomes and an X sex chromosome.

Oogenesis begins in the fetus. By six months after conception, the fetal ovaries are fully formed and contain about half a million primary oocytes arrested in the diplotene substage of prophase I. These cells, with their homologous chromosomes locked in synapsis, were thought for decades to be the only oocytes the female will produce. If so, a girl is born with all the oocytes she will ever possess. Remarkably, recent research has brought this long-held theory into question. Scientists have shown that germ-line precursor cells removed from adult ovaries can produce new eggs in a petri dish. However, it is not yet known whether these eggs are viable nor if these germ-line cells normally produce eggs in adults.

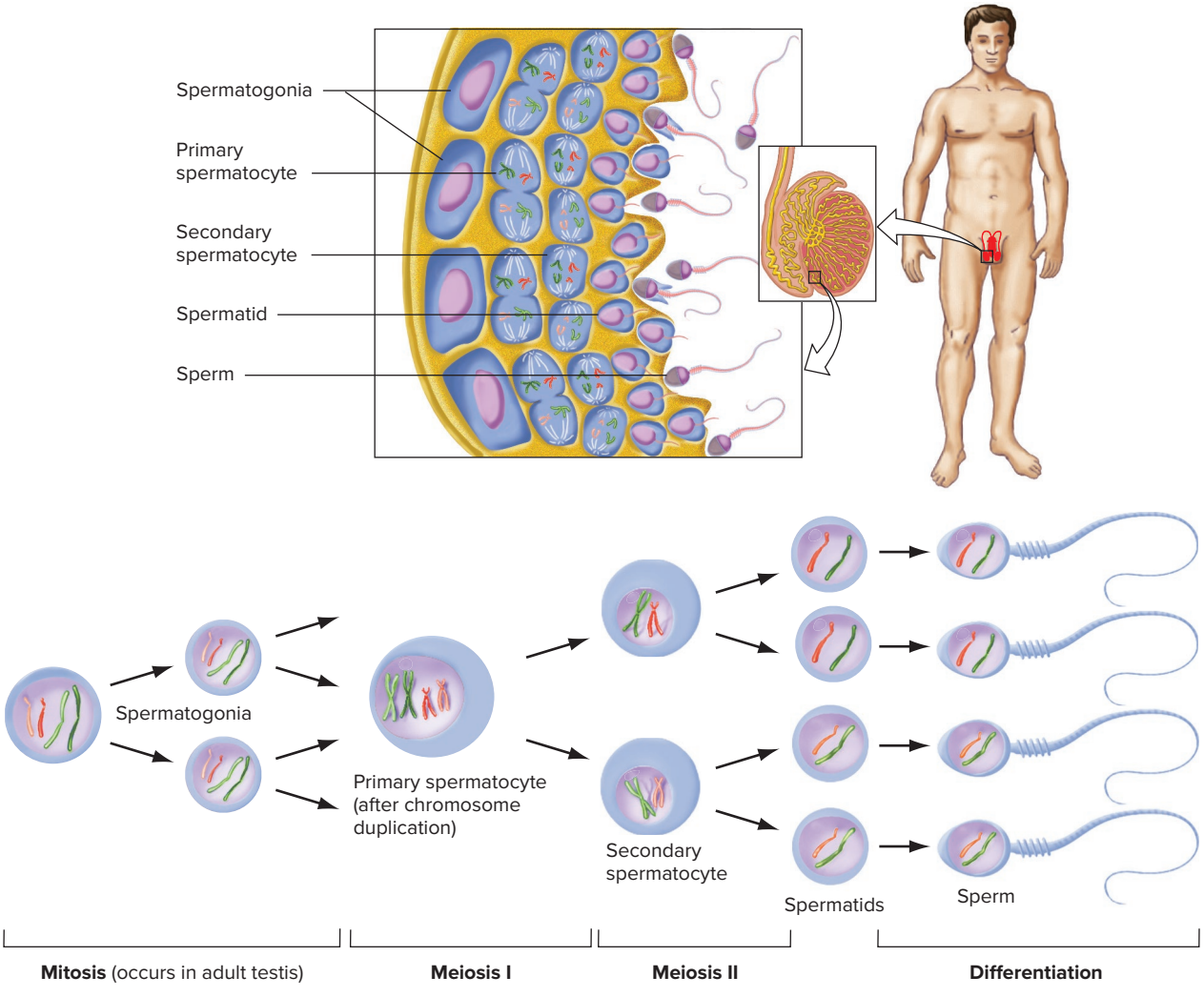
From the onset of puberty at about age 12, until menopause some 35–40 years later, most women release one primary oocyte each month (from alternate ovaries), amounting to roughly 480 oocytes released during the reproductive years. The remaining primary oocytes disintegrate during menopause. At ovulation, a released oocyte completes meiosis I and proceeds as far as the metaphase of meiosis II. If the oocyte is then fertilized, that is, penetrated by a sperm nucleus, it quickly completes meiosis II. The nuclear membranes of the sperm and ovum dissolve, allowing their chromosomes to form the single diploid nucleus of the zygote, and the zygote divides by mitosis to produce a functional embryo. In contrast, unfertilized oocytes exit the body during the menses stage of the menstrual cycle.

The long interval before completion of meiosis in oocytes released by women in their 30s, 40s, and 50s may contribute to the observed correlation between maternal age and meiotic segregational errors, including those that produce trisomies. Women in their mid-20s, for example, run a very small risk of trisomy 21; only 0.05% of children born to women of this age have Down syndrome. During the later childbearing years, however, the risk rises rapidly; at age 35, it is 0.9% of live births, and at age 45, it is 3%. You would not expect this age-related increase in risk if meiosis were completed before the mother's birth.

Spermatogenesis in Humans Produces Four Sperm from Each Primary Spermatocyte

The production of sperm, or **spermatogenesis** (**Fig. 4.19**), begins in the male testes in germ cells known as **spermatogonia.** Mitotic divisions of the spermatogonia

Figure 4.19 Human sperm form continuously in the testes after puberty. Spermatogonia are located near the exterior of seminiferous tubules in a human testis. Once they divide to produce the primary spermatocytes, the subsequent stages of spermatogenesis—meiotic divisions in the spermatocytes and maturation of spermatids into sperm—occur successively closer to the middle of the tubule. Mature sperm are released into the central lumen of the tubule for ejaculation.



produce many diploid cells, the **primary spermatocytes.** Unlike primary oocytes, primary spermatocytes undergo a symmetrical meiosis I, producing two **secondary spermatocytes,** each of which undergoes a symmetrical meiosis II. At the conclusion of meiosis, each original primary spermatocyte thus yields four equivalent haploid **spermatids.** These spermatids then mature by developing a characteristic whiplike tail and by concentrating all their chromosomal material in a head, thereby becoming functional **sperm.** A human sperm, much smaller than the ovum it will fertilize, contains 22 autosomes and *either* an X *or* a Y sex chromosome.

The timing of sperm production differs radically from that of egg formation. The meiotic divisions allowing conversion of primary spermatocytes to spermatids begin only at puberty, but meiosis then continues throughout a man's life. The entire process of spermatogenesis takes about 48–60 days: 16–20 for meiosis I, 16–20 for meiosis II, and 16–20 for the maturation of spermatids into fully functional sperm. Within each testis after puberty, millions of sperm are always in production, and a single ejaculate can contain up to 300 million. Over a lifetime, a man can produce billions of sperm, almost equally divided between those bearing an X and those bearing a Y chromosome.

essential concepts

- Diploid *germ cell* precursors proliferate by mitosis and then undergo meiosis to produce haploid *gametes*.
- Human females are born with *oocytes* arrested in prophase of meiosis I. Meiosis resumes at *ovulation* but is not completed until fertilization. *Spermatogenesis* begins at puberty and continues through the lifetimes of human males.
- The two meiotic divisions of *oogenesis* are asymmetrical, so a *primary oocyte* results in a single egg. The two meiotic divisions of spermatogenesis are symmetrical, so a *primary spermatocyte* results in four *sperm*.
- All human oocytes contain a single X chromosome; human sperm contain either an X or a Y.

4.6 Validation of the Chromosome Theory

learning objectives

1. Describe the key events of meiosis that explain Mendel's first and second laws.
2. Infer from the results of crosses whether or not a trait is sex-linked.
3. Predict phenotypes associated with nondisjunction of sex chromosomes.

We have presented thus far two circumstantial lines of evidence in support of the chromosome theory of inheritance. First, the phenotype of sexual morphology is associated with the inheritance of particular chromosomes. Second, the events of mitosis, meiosis, and gametogenesis ensure a constant number of chromosomes in the somatic cells of all members of a species over time; one would expect the genetic material to exhibit this kind of stability even in organisms with very different modes of reproduction. Final acceptance of the chromosome theory depended on researchers going beyond the circumstantial evidence to a rigorous demonstration of two key points: (1) that the inheritance of genes corresponds with the inheritance of chromosomes in every detail, and (2) that the transmission of particular chromosomes coincides with the transmission of specific traits other than sex determination.

Mendel's Laws Correlate with Chromosome Behavior During Meiosis

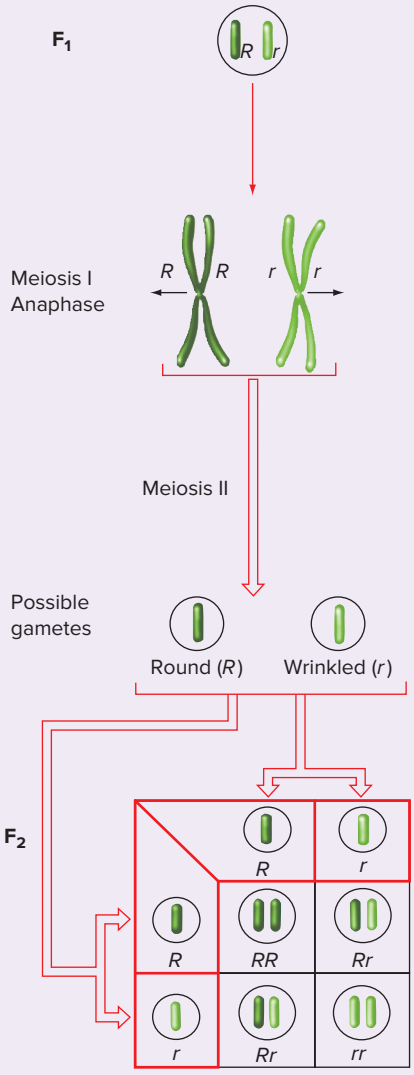
Walter Sutton first outlined the chromosome theory of inheritance in 1902–1903, building on the theoretical ideas and experimental results of Theodor Boveri in Germany, E. B. Wilson in New York, and others. In a 1902 paper, Sutton speculated that "the association of paternal and maternal chromosomes in pairs and their subsequent separation during the reducing division (that is, meiosis I) . . . may constitute the physical basis of the Mendelian law of heredity." In 1903, he suggested that chromosomes carry Mendel's hereditary units for the following reasons:

1. Every cell contains two copies of each kind of chromosome, and two copies of each kind of gene.
2. The chromosome complement, like Mendel's genes, appears unchanged as it is transmitted from parents to offspring through generations.
3. During meiosis, homologous chromosomes pair and then separate to different gametes, just as the alternative alleles of each gene segregate to different gametes.
4. Maternal and paternal copies of each chromosome pair move to opposite spindle poles without regard to the assortment of any other homologous chromosome pair, just as the alternative alleles of unrelated genes assort independently.
5. At fertilization, an egg's set of chromosomes unites with a randomly encountered sperm's set of chromosomes, just as alleles obtained from one parent unite at random with those from the other parent.
6. In all cells derived from the fertilized egg, one-half of the chromosomes and one-half of the genes are of maternal origin, the other half of paternal origin.

The two parts of **Table 4.4** show the intimate relationship between the chromosome theory of inheritance and

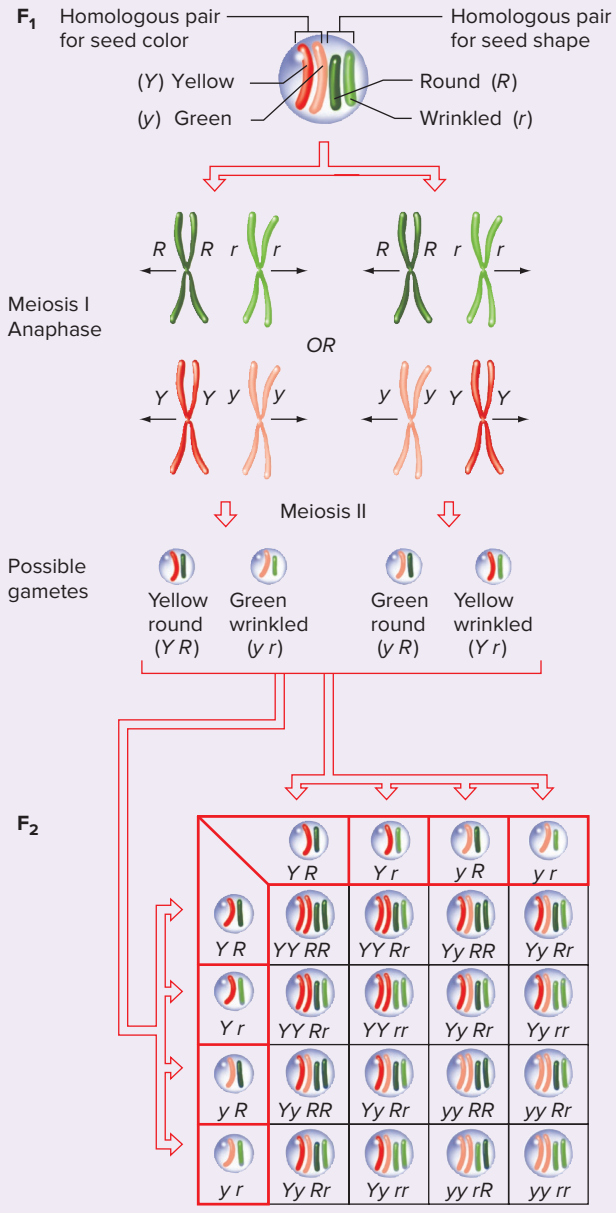
TABLE 4.4 How the Chromosome Theory of Inheritance Explains Mendel's Laws

(a) The Law of Segregation



In an F_1 hybrid plant, the allele for round peas (*R*) is found on one chromosome, and the allele for wrinkled peas (*r*) is on the homologous chromosome. The pairing between the two homologous chromosomes during prophase through metaphase of meiosis I makes sure that the homologs will separate to opposite spindle poles during anaphase I. At the end of meiosis II, two types of gametes have been produced: half have *R*, and half have *r*, but no gametes have both alleles. Thus, the separation of homologous chromosomes at meiosis I corresponds to the segregation of alleles. As the Punnett square shows, fertilization of 50% *R* and 50% *r* eggs with the same proportion of *R* and *r* sperm leads to Mendel's 3:1 ratio in the F_2 generation.

(b) The Law of Independent Assortment



One pair of homologous chromosomes carries the gene for seed shape (alleles *R* and *r*). A second pair of homologous chromosomes carries the gene for seed color (alleles *Y* and *y*). Each homologous pair aligns at random at the metaphase plate during meiosis I, independently of the other homologous pair. Thus, two equally likely configurations are possible for the migration of any two chromosome pairs toward the poles during anaphase I. As a result, a dihybrid individual will generate four equally likely types of gametes with regard to the two traits in question. The Punnett square affirms that independent assortment of traits carried by nonhomologous chromosomes produces Mendel's 9:3:3:1 ratio.

Mendel's laws of segregation and independent assortment. If Mendel's genes for pea shape and pea color are assigned to different (that is, nonhomologous) chromosomes, the behavior of chromosomes can be seen to parallel the behavior of genes. Walter Sutton's observation of these parallels led him to propose that chromosomes and genes are physically connected in some manner. Meiosis ensures that each gamete will contain only a single chromatid of a bivalent and thus only a single allele of any gene on that chromatid (Table 4.4a). The independent behavior of two bivalents during meiosis means that the genes carried on different chromosomes will assort into gametes independently (Table 4.4b).

From a review of Fig. 4.17a, which follows two different chromosome pairs through the process of meiosis, you might wonder whether crossing-over abolishes the clear correspondence between Mendel's laws and the movement of chromosomes. The answer is no. Each chromatid of a homologous chromosome pair contains only one copy of a given gene, and only one chromatid from each pair of homologs is incorporated into each gamete. Because alternative alleles remain on different chromatids even after crossing-over has occurred, alternative alleles still segregate to different gametes as demanded by Mendel's first law.

Furthermore, because the orientations of nonhomologous chromosomes are completely random with respect to each other during both meiotic divisions, the genes on different chromosomes assort independently even if crossing-over occurs, as demanded by Mendel's second law. In Fig. 4.17a, you can see that without recombination, each of the two random alignments of the nonhomologous chromosomes results in the production of only two of the four gamete types: *AB* and *ab* for one orientation, and *Ab* and *aB* for the other orientation. With recombination, each of the alignments of alleles in Fig. 4.17a may in fact generate all four gamete types. (Imagine a crossover switching the positions of *A* and *a* nonsister chromatids in Fig. 4.17a). Thus, both the random alignment of nonhomologous chromosomes and crossing-over contribute to the phenomenon of independent assortment.

Specific Traits Are Transmitted with Specific Chromosomes

The fate of a theory depends on whether its predictions can be validated. Because genes determine traits, the prediction that chromosomes carry genes could be tested by breeding experiments that would show whether transmission of a specific chromosome coincides with transmission of a specific trait. Cytologists knew that one pair of chromosomes, the sex chromosomes, determines whether an individual is male or female. Would similar correlations exist for other traits?

A gene determining eye color on the *Drosophila* X chromosome

Thomas Hunt Morgan, an American experimental biologist with training in embryology, headed the research group whose findings eventually established a firm experimental base for the chromosome theory. Morgan chose to work with the fruit fly *Drosophila melanogaster* because it is extremely prolific and has a very short generation time, taking only 12 days to develop from a fertilized egg into a mature adult capable of producing hundreds of offspring. Morgan fed his flies mashed bananas and housed them in empty milk bottles capped with wads of cotton.

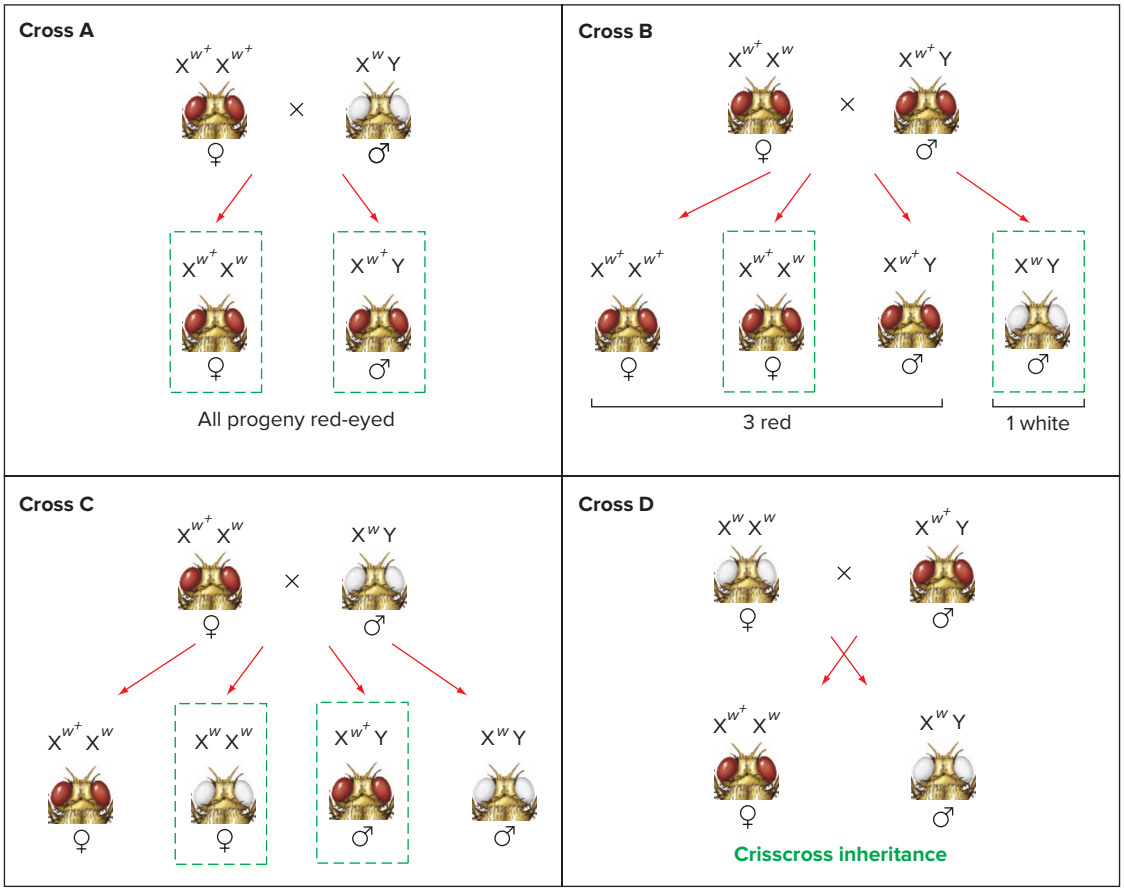
In 1910, a white-eyed male appeared among a large group of flies with brick-red eyes. A mutation had apparently altered a gene determining eye color, changing it from the normal wild-type allele specifying red to a new allele that produced white. When Morgan allowed the white-eyed male to mate with its red-eyed sisters, all the flies of the F_1 generation had red eyes; the red allele was clearly dominant to the white (**Fig. 4.20,** cross A).

Establishing a pattern of nomenclature for *Drosophila* geneticists, Morgan named the gene identified by the abnormal white eye color the *white* gene, for the mutation that revealed its existence. The normal wild-type allele of the *white* gene, abbreviated w^+, is for brick-red eyes, while the counterpart mutant *w* allele results in white eye color. The superscript + signifies the wild type. By writing the gene name and abbreviation in lowercase, Morgan symbolized that the mutant *w* allele is recessive to the wild-type w^+. (If a *Drosophila* mutation results in a dominant non-wild-type phenotype, the first letter of the gene name or of its abbreviation is capitalized; thus the mutation known as *Bar* eyes is dominant to the wild-type Bar^+ allele. (See the Appendix *Guidelines for Gene Nomenclature.*)

Morgan then crossed the red-eyed males of the F_1 generation with their red-eyed sisters (Fig. 4.20, cross B) and obtained an F_2 generation with the predicted 3:1 ratio of red to white eyes. But there was something askew in the pattern: Among the red-eyed offspring, there were two females for every one male, and all the white-eyed offspring were males. This result was surprisingly different from the equal transmission to both sexes of the Mendelian traits discussed in Chapters 2 and 3. In these fruit flies, the ratio of eye colors was not the same in male and female progeny.

By mating F_2 red-eyed females with their white-eyed brothers (Fig. 4.20, cross C), Morgan obtained some females with white eyes, which then allowed him to mate a white-eyed female with a red-eyed wild-type male (Fig. 4.20, cross D). The result was exclusively red-eyed daughters and white-eyed sons. The pattern seen in cross D is known as **crisscross inheritance** because the males inherit their eye color from their mothers, while the daughters inherit their eye color from their fathers. Note in Fig. 4.20 that the

Figure 4.20 A *Drosophila* eye color gene is located on the X chromosome. X-linkage explains the inheritance of alleles of the *white* gene in this series of crosses performed by Thomas Hunt Morgan. The progeny of crosses A, B, and C outlined with green dotted boxes are those used as the parents in the next cross of the series.



results of the reciprocal crosses red female × white male (cross A) and white female × red male (cross D) are not identical, again in contrast with Mendel's findings.

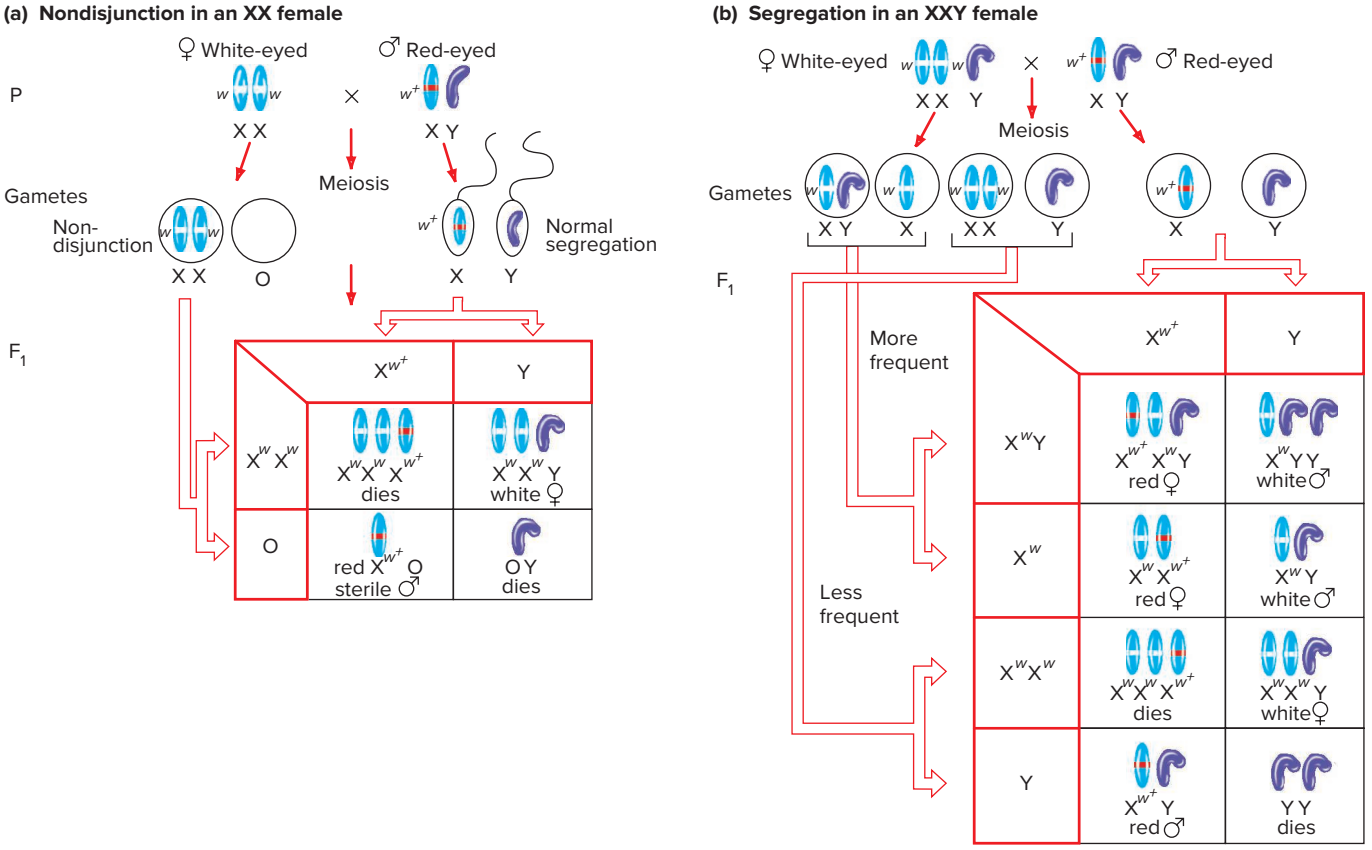
From the data, Morgan reasoned that the *white* gene for eye color is **X-linked,** that is, carried by the X chromosome. (Note that while symbols for genes and alleles are italicized, symbols for chromosomes are not.) The Y chromosome carries no allele of this gene for eye color. Males, therefore, have only one copy of the gene, which they inherit from their mother along with their only X chromosome; their Y chromosome must come from their father. Thus, males are **hemizygous** for this eye color gene, because their diploid cells have half the number of alleles carried by the female on her two X chromosomes.

If the single *white* gene on the X chromosome of a male is the wild-type w^+ allele, he will have red eyes and a genotype that can be written $X^{w^+}Y$. [Here we designate the chromosome (X or Y) together with the allele it carries, to emphasize that certain genes are X-linked.] In contrast to an $X^{w^+}Y$ male, a hemizygous X^wY male would have white eyes. Females with two X chromosomes can be one of three genotypes: $X^w X^w$ (white-eyed), $X^wX^{w^+}$ (red-eyed because w^+ is dominant to w), or $X^{w^+}X^{w^+}$ (red-eyed). As shown in Fig. 4.20, Morgan's assumption that the gene for eye color is X-linked explains the results of his breeding experiments. Crisscross inheritance, for example, occurs because the only X chromosome in sons of a white-eyed mother (X^wX^w) must carry the w allele, so the sons will be white-eyed. In contrast, because daughters of a red-eyed ($X^{w^+}Y$) father must receive a w^+-bearing X chromosome from their father, they should all have red eyes.

Validation of the chromosome theory from the analysis of nondisjunction

Although Morgan's work strongly supported the hypothesis that the gene for eye color lies on the X chromosome, he himself continued to question the validity of the chromosome theory until Calvin Bridges, one of his top students, found another key piece of evidence. Bridges repeated the cross Morgan had performed between white-eyed females and red-eyed males, but this time he did the experiment on a larger scale. As expected, the progeny of this cross consisted mostly of red-eyed females and white-eyed males. However, about 1 in every 2000 males had red eyes, and about the same small fraction of females had white eyes.

Figure 4.21 Nondisjunction: Rare mistakes in meiosis help confirm the chromosome theory. **(a)** Rare events of nondisjunction in an XX female produce XX and O eggs. The results of normal disjunction in the female are not shown. XO males are sterile because the missing Y chromosome is needed for male fertility in *Drosophila*. **(b)** In an XXY female, the three sex chromosomes can pair and segregate in two ways, producing progeny with unusual sex chromosome complements.



Bridges hypothesized that these exceptions arose through rare events in which the X chromosomes fail to separate during meiosis in females. He called such failures in chromosome segregation *nondisjunction*. Mistakes leading to nondisjunction can occur during either meiosis I or meiosis II, but in either case nondisjunction would result in some eggs with two X chromosomes and others with none. As **Fig. 4.21a** shows, fertilization of these chromosomally abnormal eggs could produce four types of zygotes: XXY (with two X chromosomes from the egg and a Y from the sperm), XXX (with two Xs from the egg and one X from the sperm), XO (with the lone sex chromosome from the sperm and no sex chromosome from the egg), and OY (with the only sex chromosome again coming from the sperm).

When Bridges examined the sex chromosomes of the rare white-eyed females produced in his large-scale cross, he found that they were indeed XXY individuals who must have received two X chromosomes and with them two *w* alleles from their white-eyed X^wX^w mothers. The exceptional red-eyed males emerging from the cross were XO; their eye color showed that they must have obtained their sole sex chromosome from their $X^{w+}Y$ fathers. In this study then, transmission of the *white* gene alleles followed the predicted behavior of X chromosomes during rare meiotic mistakes, indicating that the X chromosome carries the gene for eye color. These results also suggested that zygotes with the two other abnormal sex chromosome karyotypes expected from nondisjunction in females (XXX and OY) die during embryonic development and thus produce no progeny.

Because XXY white-eyed females have three sex chromosomes rather than the normal two, Bridges reasoned they would produce four kinds of eggs: XY and X, or XX and Y (**Fig. 4.21b**). You can visualize the formation of these four kinds of eggs by imagining that when the three chromosomes pair and disjoin during meiosis, two chromosomes must go to one pole and one chromosome to the other. With this kind of segregation, only two results are possible: Either one X and the Y go to one pole and the second X to the other (yielding XY and X gametes), or the two Xs go to one pole and the Y to the other (yielding XX and Y gametes). The first of these two scenarios occurs more often because it comes about when the two similar X chromosomes pair with each other, ensuring that they will go to opposite poles during the first meiotic division. The second, less likely possibility happens only if the two X chromosomes fail to pair with each other.

Bridges next predicted that fertilization of these four kinds of eggs from an XXY female by normal sperm would generate an array of sex chromosome karyotypes associated with specific eye colors in the progeny. Bridges verified all his predictions when he analyzed the eye colors and sex chromosomes of a large number of offspring. For instance, he showed cytologically that all of the white-eyed females emerging from the cross in Fig. 4.21b had two X chromosomes and one Y chromosome, while one-half of the white-eyed males had a single X chromosome and two Y chromosomes. Bridges' painstaking observations provided compelling evidence that specific genes do in fact reside on specific chromosomes.

The Chromosome Theory Integrates Many Aspects of Gene Behavior

Mendel had assumed that genes are located in cells. The chromosome theory assigned the genes to a specific kind of structure within cells and explained alternative alleles as physically matching parts of homologous chromosomes. In so doing, the theory provided an explanation of Mendel's laws. The mechanism of meiosis ensures that the matching parts of homologous chromosomes will segregate to different gametes (except in rare instances of nondisjunction), accounting for the segregation of alleles predicted by Mendel's first law. Because each homologous chromosome pair aligns independently of all others at meiosis I, genes carried on different chromosomes will assort independently, as predicted by Mendel's second law.

The chromosome theory is also able to explain the creation of new alleles through *mutation,* a spontaneous change in a particular gene (that is, in a particular part of a chromosome). If a mutation occurs in the germ line, it can be transmitted to subsequent generations.

Finally, through mitotic cell divisions in the embryo and after birth, each cell in a multicellular organism receives the same chromosomes—and thus the same maternal and paternal alleles of each gene—as the zygote received from the egg and sperm at fertilization. In this way, an individual's genome—the chromosomes and genes he or she carries—remains constant throughout life.

essential concepts

- Segregation of homologous chromosomes into daughter cells at meiosis I explains Mendel's first law.
- Independent alignment of homologs with respect to each other and crossing-over of nonsister chromatids during meiosis I explain Mendel's second law.
- In organisms with XX/XY sex determination, males are *hemizygous* for X-linked genes, while females have two copies.

4.7 Sex-Linked and Sexually Dimorphic Traits in Humans

learning objectives

1. Determine from pedigree analysis whether human traits are X-linked or autosomal.
2. Explain how human cells compensate for the X-linked gene dosage difference in XX and XY nuclei.

A person unable to tell red from green would find it nearly impossible to distinguish the rose, scarlet, and magenta in the flowers of a garden bouquet from the delicately variegated greens in their foliage, or to complete a complex electrical circuit by fastening red-clad metallic wires to red ones and green to green. Such a person has most likely inherited some form of red-green color blindness, a recessive condition that runs in families and affects mostly males. Among Caucasians in North America and Europe, 8% of men but only 0.44% of women have this vision defect. **Figure 4.22** suggests to readers with normal color vision what people with red-green color blindness actually see.

In 1911, E. B. Wilson, a contributor to the chromosome theory of inheritance, combined family studies of the inheritance of color blindness with recent knowledge of the roles of the X and Y chromosomes in sex determination to make the first assignment of a human gene to a particular chromosome. The gene for red-green color blindness, he said, lies on the X because the condition usually passes from a maternal grandfather through an unaffected carrier mother to roughly 50% of the grandsons.

Several years after Wilson made this gene assignment, pedigree analysis established that various forms of hemophilia, or *bleeders disease* (in which the blood fails to clot properly), also result from mutations on the X chromosome that give rise to a relatively rare, recessive trait. In this context, rare means *infrequent in the population.* The family histories under review, including one following the descendants of Queen Victoria of England (**Fig. 4.23a**), showed that relatively rare X-linked traits appear more often in males than in females and often skip generations. The clues that suggest X-linked recessive inheritance in a pedigree are summarized in **Table 4.5.**

Unlike color blindness and hemophilia, some—although very few—of the known rare mutations on the X chromosome are dominant to the wild-type allele. With such dominant X-linked mutations, more females than males show the aberrant phenotype. This phenomenon occurs because all the daughters of an affected male but none of the sons will have the condition, while one-half the sons and one-half the daughters of an affected female will receive the dominant allele and therefore show the phenotype (see Table 4.5).

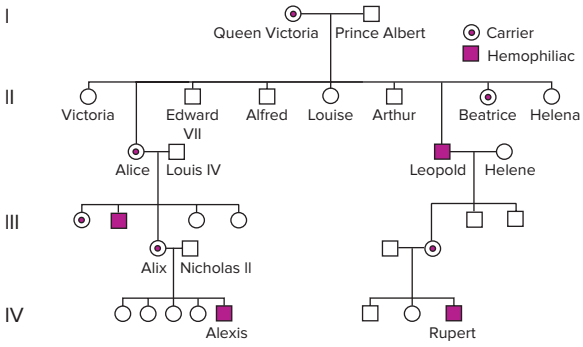
Figure 4.22 Red-green color blindness is an X-linked recessive trait in humans. How the world looks to a person with either normal color vision **(a)** or a kind of red-green color blindness known as deuteranopia **(b)**.
(both): Color deficit simulation courtesy of Vischeck (www.vischeck.com). Source image courtesy of NASA

(a)

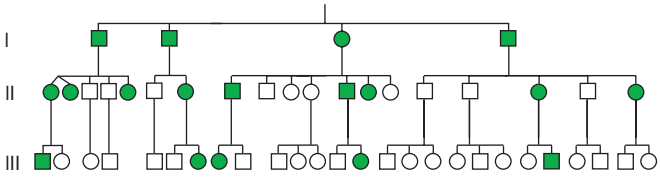
(b)

Figure 4.23 X-linked traits may be recessive or dominant. **(a)** Pedigree showing inheritance of the recessive X-linked trait hemophilia in Queen Victoria's family. **(b)** Pedigree showing the inheritance of the dominant X-linked trait hypophosphatemia, commonly referred to as vitamin D–resistant rickets.

(a) X-linked recessive: Hemophilia



(b) X-linked dominant: Hypophosphatemia



Vitamin D–resistant rickets, or hypophosphatemia, is an example of an X-linked dominant trait. **Figure 4.23b** presents the pedigree of a family affected by this disease.

In XX Human Females, One X Chromosome Is Inactivated

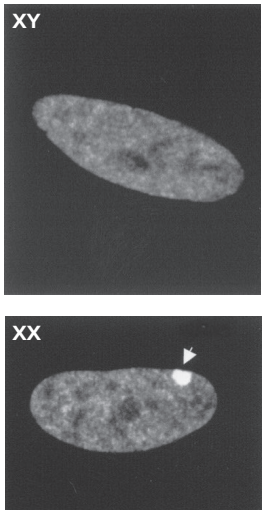
The XX and XY system of sex determination presents human cells with a curious problem that requires a solution

TABLE 4.5 Pedigree Patterns Suggesting Sex-Linked Inheritance

X-Linked Recessive Trait
1. The trait appears in more males than females because a female must receive two copies of the rare defective allele to display the phenotype, whereas a hemizygous male with only one copy will show it.
2. The mutation will never pass from father to son because sons receive only a Y chromosome from their father.
3. An affected male passes the X-linked mutation to all his daughters, who are thus carriers. Each son of these carrier females has a one-half chance to inherit the defective allele and thus the trait.
4. The trait often skips a generation as the mutation passes from grandfather through a carrier daughter to grandson.
5. The trait can appear in successive generations when a sister of an affected male is a carrier. If she is, each of her sons has a one-half chance of being affected.
6. With the rare affected (homozygous) female, all her sons will be affected and all her daughters will be carriers.
X-Linked Dominant Trait
1. More females than males show the aberrant trait.
2. The trait is seen in every generation as long as affected males have female children.
3. All the daughters but none of the sons of an affected male will be affected. This criterion is the most useful for distinguishing an X-linked dominant trait from an autosomal dominant trait.
4. The sons and daughters of an affected female each have a one-half chance of being affected.
5. For incompletely dominant X-linked traits, carrier females may show the trait in less extreme form than males with the defective allele.
Y-Linked Trait
1. The trait is seen only in males.
2. All male descendants of an affected man will exhibit the trait.
3. Not only do females not exhibit the trait, they also cannot transmit it.

Figure 4.24 Barr bodies are densely staining particles in XX cell nuclei. The arrow points to a Barr body in the nucleus of an XX cell treated with a DNA stain. The Barr body appears bright *white* in this negative image. Unlike the other chromosomes, the Barr body is highly condensed and attached to the nuclear envelope. XY cells have no Barr bodies.

a-b: From: Hong et al. (17 July 2001), "Identification of an autoimmune serum containing antibodies against the Barr body," *PNAS,* 98(15): 8703-8708, Fig 1A-B. © 2001 National Academy of Sciences, U.S.A.

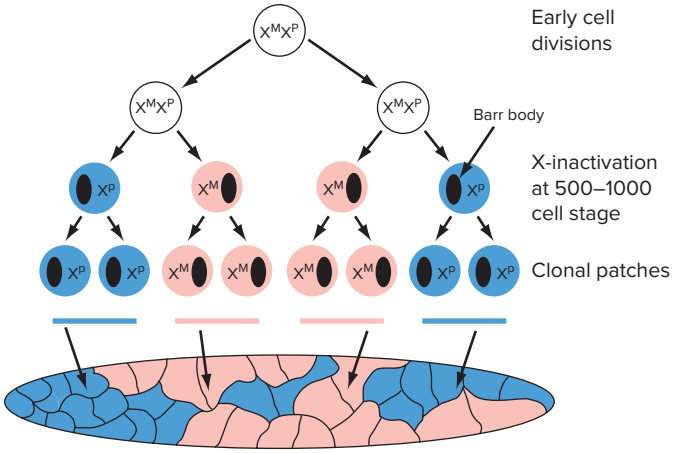


called **dosage compensation.** As mentioned earlier, the X chromosome contains about 1100 genes, and the proteins that they specify need to be present in the same amounts in male and female cells. To compensate for female cells having two copies of each X-linked gene and male cells having only one, XX cells inactivate one of their two X chromosomes. Almost all of the genes on the inactivated X chromosome are turned off, so no gene products can be made. X inactivation occurs at about two weeks after fertilization, when an XX human embryo is composed of only 500–1000 cells. At that time, each cell chooses one X chromosome at random to condense into a so-called **Barr body** and thereby inactivate it. Barr bodies, named after the cytologist Murray Barr who discovered them, appear as small, dark chromosomes in interphase cells treated with a DNA stain that allows chromosomes to be visible under a light microscope (**Fig. 4.24**).

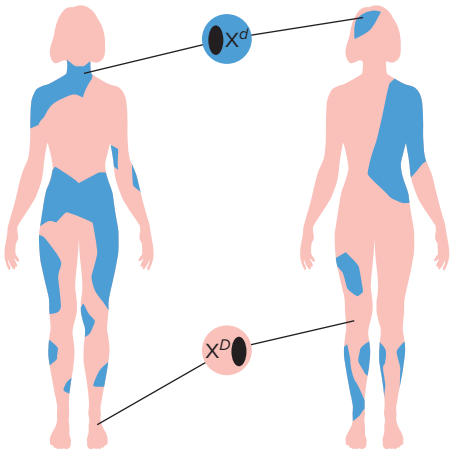
Each embryonic cell "decides" independently which X chromosome will be inactivated—either the X inherited from the mother or the paternal X. Once the determination is made, it is clonally perpetuated so that all of the millions of cells descended by mitosis from a particular embryonic cell condense the same X chromosome to a Barr body (**Fig. 4.25a**). Human females are thus a patchwork of cells, some containing a maternally derived active X chromosome, and the others an active paternal X (**Fig. 4.25b**). The Fast Forward Box *Visualizing X Chromosome Inactivation in Transgenic Mice* explains how scientists have recently developed technology in mice to visualize, both

Figure 4.25 X chromosome dosage compensation makes human females a patchwork for X-linked gene expression. (a) Early in embryogenesis, each XX cell inactivates one randomly chosen X chromosome by condensing it into a Barr body (*black* oval). The same X chromosome remains a Barr body in all descendants of each cell. X^M = maternal X chromosome; X^P = paternal X chromosome. **(b)** The twins shown here are heterozygotes (*Dd*) for the X-linked recessive condition anhidrotic ectodermal dysplasia, which prevents sweat gland development. Patches of skin in *blue* lack sweat glands because the chromosome with the wild-type allele (*D*) is inactivated and the recessive *d* allele is nonfunctional.

(a) Perpetuation of X chromosome inactivation after cell divisions



(b) X chromosome inactivation results in patchwork females



outside and inside the body, the clonal patches of cells that express the genes on one X chromosome or the other.

The phenomenon of **X chromosome inactivation** may have interesting effects on the traits controlled by X-linked genes. When females are heterozygous at an X-linked gene, parts of their bodies are in effect hemizygous for one allele, and parts are hemizygous for the other allele in terms of gene function. Moreover, which body parts are functionally hemizygous for one allele or the other is random; even identical twins, who have identical alleles of all of their genes, will

FAST FORWARD

Visualizing X Chromosome Inactivation in Transgenic Mice

Scientists have recently used molecular techniques and transgenic technology (similar to that described in the earlier Fast Forward Box *Transgenic Mice Prove That* SRY *Is the Maleness Factor*) to visualize the pattern of X chromosome inactivation in mice. The researchers generated XX mice containing two different *transgenes* (in this case, genes from a different species). One of these transgenes was a jellyfish gene that specifies green fluorescent protein (GFP); the other was a gene from red coral that makes red fluorescent protein (RFP) **(Fig. A).**

In the XX mice, the GFP gene is located on the X chromosome from the mother, and the RFP gene resides on the X chromosome from the father. Clonal patches of cells are either green or red depending on which X chromosome was turned into a Barr body in the original cell that established the patch **(Fig. B).**

Different XX mice display different green and red patchwork patterns, providing a clear demonstration of the random nature of X chromosome inactivation. The patchwork patterns reflect the cellular memory of which X chromosome was inactivated in the founder cell for each clonal patch. Geneticists currently use these transgenic mice to decipher the genetic details of how cells "remember" which X to inactivate after each cell division.

Figure A Cells of transgenic mice glow either green or red in response to X chromosome inactivation. The mouse carries a green (GFP) transgene inserted in the maternal X chromosome (X^M), and a red (RFP) transgene in the paternal X chromosome (X^P). Cells in which X^P is inactivated (*top*) glow green; cells glow red (*bottom*) when X^M is inactivated.

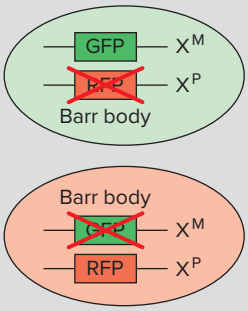
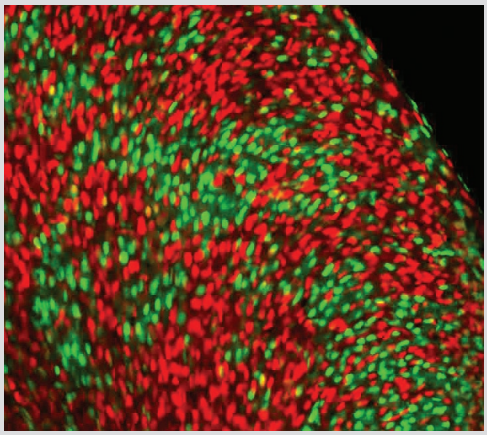


Figure B Heart cells of a transgenic mouse reveal a clonal patchwork of X inactivation. Patches of red or green cells represent cellular descendants of the founders that randomly inactivated one of their X chromosomes.



have a different pattern of X chromosome inactivation. In Fig. 4.25b, females heterozygous for the X-linked recessive trait anhidrotic epidermal dysplasia have patches of skin that lack sweat glands interspersed with patches of normal skin; the phenotype of a patch depends upon which X chromosome is inactivated. Each patch is a clone of skin cells derived from a single embryonic cell that made the decision to inactivate one of the X chromosomes. In a second example, women heterozygous for an X-linked recessive hemophilia allele are called *carriers* of the disease allele, even though they may have some symptoms of hemophilia. The severity of the condition depends on the particular random pattern of cells that inactivated the disease allele and cells that inactivated the normal allele. In Chapter 3, we discussed how chance events work through genes to affect phenotype; X inactivation is a perfect example of such an event.

Recall that the two tips of the X chromosome, the pseudoautosomal regions (PARs), contain genes also present at the tips of the Y chromosome (Fig. 4.8). In order to equalize the dosage of these genes in XX and XY cells, the PAR genes on the Barr body X chromosome escape inactivation. This feature of dosage compensation may explain at least in part why XXY males (Klinefelter syndrome) and XO females (Turner syndrome) have abnormal morphological features. Although one of the two X chromosomes in XXY males becomes a Barr body, Klinefelter males have three doses (rather than the normal two) of the genes in the PAR regions. The single X chromosome in XO cells does not become a Barr body, yet these cells have only one dose of the PAR genes (rather than two in XX females).

X chromosome inactivation is common to mammals, and we will present the molecular details of this process in

later chapters. It is nonetheless important to realize that other organisms compensate for sex chromosome differences in alternative ways. Fruit flies, for example, hyperactivate the single X chromosome in XY (male) cells, so that most X chromosome genes produce twice as much protein product as each X chromosome in a female. The nematode *C. elegans,* in contrast, ratchets down the level of gene activity on each of the X chromosomes in XX hermaphrodites relative to the single X in XO males.

Maleness and Male Fertility Are the Only Known Y-Linked Traits in Humans

Theoretically, phenotypes caused by mutations on the Y chromosome should also be identifiable by pedigree analysis. Such traits would pass from an affected father to all of his sons, and from them to all future male descendants. Females would neither exhibit nor transmit a Y-linked phenotype (see Table 4.5). However, besides the determination of maleness itself, as well as contributions to sperm formation and thus male fertility, no clear-cut Y-linked visible traits have turned up in humans. The paucity of known Y-linked traits reflects the fact that, as mentioned earlier, the small Y chromosome contains very few genes. Indeed, one would expect the Y chromosome to have only a limited effect on phenotype because normal XX females do perfectly well without it.

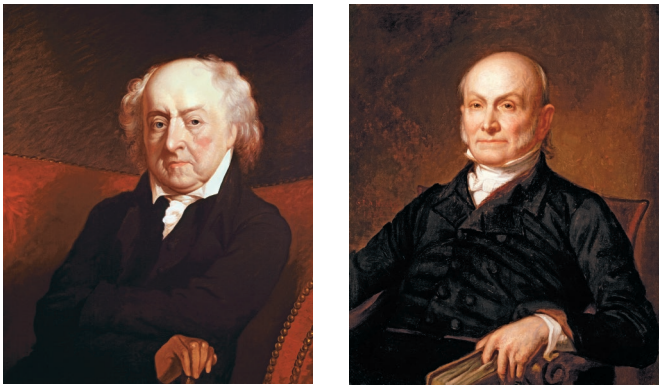
Autosomal Genes Contribute to Sexual Dimorphism

Not all genes that produce sexual dimorphism (differences in the two sexes) reside on the X or Y chromosomes. Some autosomal genes govern traits that appear in one sex but not the other, or traits that are expressed differently in the two sexes.

Sex-limited traits affect a structure or process that is found in one sex but not the other. Mutations in genes for sex-limited traits can influence only the phenotype of the sex that expresses those structures or processes. A vivid example of a sex-limited trait occurs in *Drosophila* males homozygous for an autosomal recessive mutation known as *stuck,* which affects the ability of mutant males to retract their penis and release the claspers by which they hold on to female genitalia during copulation. The mutant males have difficulty separating from females after mating. In extreme cases, both individuals die, forever caught in their embrace. Because females lack penises and claspers, homozygous *stuck* mutant females can mate normally.

Sex-influenced traits show up in both sexes, but the expression of such traits may differ between the two sexes because of hormonal differences. Pattern baldness, a condition in which hair is lost prematurely from the top of the head but not from the sides (**Fig. 4.26**), is a sex-influenced trait in humans. Although pattern baldness is a complex trait that can be affected by many genes, an autosomal gene appears to play an important role in certain families. Men in these families who are heterozygous for the balding allele lose their hair while still in their 20s, whereas heterozygous women do not show any significant hair loss. In contrast, homozygotes in both sexes become bald (though the onset of baldness in homozygous women is usually much later in life than in homozygous men). This sex-influenced trait is thus dominant in men, recessive in women.

Figure 4.26 Male pattern baldness, a sex-influenced trait. **(a)** John Adams (1735–1826), second president of the United States, at about age 60. **(b)** John Quincy Adams (1767–1848), son of John Adams and the sixth president of the United States, at about the same age. The father-to-son transmission suggests that male pattern baldness in the Adams family is likely determined by an allele of an autosomal gene.



(a) **(b)**

Mutations in Sex Determination Pathway Genes Can Result in Intersexuality Disorders

We previously saw that the *SRY* gene on the Y chromosome is essential to maleness because it initiates testis development early in embryogenesis. But the functions of many genes are required for testis development, or for subsequent events that rely on hormones made in the testes for the development of sexual organs. Some of these genes are autosomal and some are X-linked; in either case, an XY individual with mutant alleles for any of these genes may have unusual intersexual phenotypes.

In one important example, XY people with nonfunctional mutant alleles of the X-linked *AR* gene specifying the androgen receptor have a disorder known as complete androgen insensitivity syndrome (CAIS). These

XY individuals have testes that make the hormone testosterone, but in the absence of the androgen receptor to which it binds, the testosterone has no effect. Without the androgen receptor, these people cannot develop male genitalia (penis and scrotum) nor male internal duct systems (the vas deferens, seminal vesicles, and ejaculatory ducts); instead, their external genitalia assume the default female state (labia and clitoris). However, the testes make another hormone that prevents the formation of female internal duct systems (including the fallopian tubes, uterus, and vagina). The result is that persons with CAIS are externally female but sterile because they lack the internal duct systems of either sex.

essential concepts

- *Sex-linked* (*X-linked*) traits show sex-specific inheritance patterns because sons always inherit their father's Y chromosome, while daughters always inherit their father's X chromosome.
- Random inactivation of either the maternal or paternal X chromosome in XX cells ensures that male and female mammalian cells express equivalent amounts of the proteins encoded by most X-linked genes.
- Mutations of genes—whether autosomal or X-linked—can have different effects in males and females.

WHAT'S NEXT

T. H. Morgan and his students, collectively known as the *Drosophila* group, acknowledged that Mendelian genetics could exist independently of chromosomes. "Why then, we are often asked, do you drag in the chromosomes? Our answer is that because the chromosomes furnish exactly the kind of mechanism that Mendelian laws call for, and since there is an ever-increasing body of information that points clearly to the chromosomes as the bearers of the Mendelian factors, it would be folly to close one's eyes to so patent a relation. Moreover, as biologists, we are interested in heredity not primarily as a mathematical formulation, but rather as a problem concerning the cell, the egg, and the sperm."

The *Drosophila* group went on to find several X-linked mutations in addition to white eyes. One made the body yellow instead of brown, another shortened the wings, yet another made bent instead of straight body bristles. These findings raised several compelling questions. First, if the genes for all of these traits are physically linked together on the X chromosome, does this linkage affect their ability to assort independently, and if so, how? Second, does each gene have an exact chromosomal address, and if so, does this specific location in any way affect its transmission? In Chapter 5 we describe how the *Drosophila* group and others analyzed the transmission patterns of genes on the same chromosome in terms of known chromosome movements during meiosis, and how they then used the information obtained to localize genes to specific chromosomal positions.

SOLVED PROBLEMS

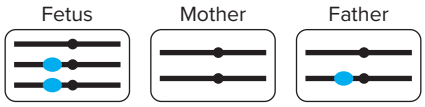
I. In humans, chromosome 16 sometimes has a heavily stained area in the long arm near the centromere. This feature can be seen through the microscope but has no effect on the phenotype of the person carrying it. When such a "blob" exists on a particular copy of chromosome 16, it is a constant feature of that chromosome and is inherited.

A couple conceived a child, but the fetus had multiple abnormalities and was miscarried. When the chromosomes of the fetus were studied, it was discovered that it had three copies of chromosome 16 (it was *trisomic* for chromosome 16), and that two of the three chromosome 16s had large blobs. Both chromosome 16 homologs in the mother lacked blobs, but the father was heterozygous for blobs. Which parent experienced nondisjunction, and in which meiotic division did it occur?

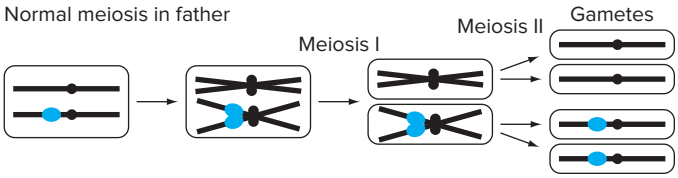
Answer

This problem requires an understanding of nondisjunction during meiosis. When individual chromosomes contain some distinguishing feature that allows one homolog to be distinguished from another, it is possible to follow the path of the two homologs through meiosis. Because the fetus had two chromosome 16s with the blob, we can conclude

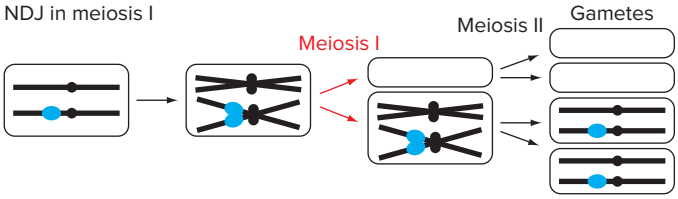
that the extra chromosome came from the father (the only parent with a blobbed chromosome).



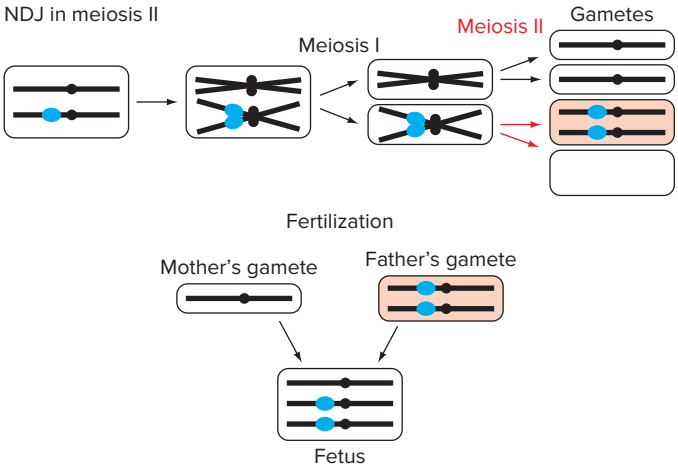
We know that nondisjunction must be involved because normal meiosis in the father would generate gametes with only a single chromosome 16 with the blob; as a result, the fetus could have only a single chromosome 16 with the blob.



In which meiotic division did the nondisjunction occur? When nondisjunction occurs during meiosis I, homologs fail to segregate to opposite poles. If this occurred in the father, the chromosome with the blob and the normal chromosome 16 would segregate into the same cell (a secondary spermatocyte). After meiosis II, the gametes resulting from this cell would carry both types of chromosomes. If such sperm fertilized a normal egg, the zygote would have two copies of the normal chromosome 16 and one copy of the chromosome with a blob.



On the other hand, if nondisjunction occurred during meiosis II in the father in a secondary spermatocyte containing the blobbed chromosome 16, sperm with two copies of the blob-marked chromosome would be produced. After fertilization with a normal egg, the result would be a zygote of the type seen in this spontaneous abortion.



Therefore, the nondisjunction occurred in meiosis II in the father.

II. (a) What sex ratio would you expect among the offspring of a cross between a normal male mouse and a female mouse heterozygous for a recessive X-linked lethal gene? (b) What would be the expected sex ratio among the offspring of a cross between a normal hen and a rooster heterozygous for a recessive Z-linked lethal allele?

Answer

This problem deals with sex-linked inheritance and sex determination.

a. Mice have a sex determination system of XX = female and XY = male. A normal male mouse (X^RY) × a heterozygous female mouse (X^RX^r) would result in X^RX^R, X^RX^r, X^RY, and X^rY mice. The X^rY mice would die, so there would be a 2:1 ratio of females to males.

b. The sex determination system in birds is ZZ = male and ZW = female. A normal hen (Z^RW) × a heterozygous rooster (Z^RZ^r) would result in Z^RZ^R, Z^RZ^r, Z^RW, and Z^rW chickens. Because the Z^rW offspring do not live, the ratio of females to males would be 1:2.

III. A woman with normal color vision whose father was color-blind mates with a man with normal color vision.

a. What do you expect to see among their offspring?

b. What would you expect if it was the normal man's father who was color-blind?

Answer

This problem involves sex-linked inheritance.

a. The woman's father has a genotype of $X^{cb}Y$. Because the woman had to inherit an X from her father, she must have an X^{cb} chromosome, but because she has normal color vision, her other X chromosome must be X^{CB}. The man she mates with has normal color vision and therefore has an $X^{CB}Y$ genotype. Their children could with equal probability be $X^{CB}X^{CB}$ (normal female), $X^{CB}X^{cb}$ (carrier female), $X^{CB}Y$ (normal male), or $X^{cb}Y$ (color-blind male).

b. If the man with normal color vision had a color-blind father, the X^{cb} chromosome would not have been passed on to him, because a male does not inherit an X chromosome from his father. The man has the genotype $X^{CB}Y$ and cannot pass on the color-blind allele.

PROBLEMS

Vocabulary

1. Choose the best matching phrase in the right column for each of the terms in the left column.

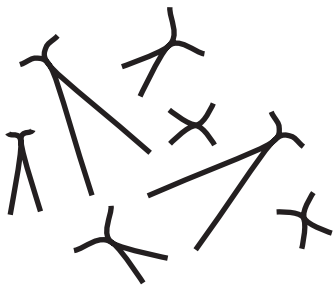
a. meiosis	1. X and Y
b. gametes	2. chromosomes that do not differ between the sexes
c. karyotype	3. one of the two identical halves of a replicated chromosome
d. mitosis	4. microtubule organizing centers at the spindle poles
e. interphase	5. cells in the testes that undergo meiosis
f. syncytium	6. division of the cytoplasm
g. synapsis	7. haploid germ cells that unite at fertilization
h. sex chromosomes	8. an animal cell containing more than one nucleus
i. cytokinesis	9. pairing of homologous chromosomes
j. anaphase	10. one diploid cell gives rise to two diploid cells
k. chromatid	11. the array of chromosomes in a given cell
l. autosomes	12. the part of the cell cycle during which the chromosomes are not visible
m. centromere	13. one diploid cell gives rise to four haploid cells
n. centrosomes	14. cell produced by meiosis that does not become a gamete
o. polar body	15. the time during mitosis when sister chromatids separate
p. spermatocytes	16. site of the closest connection between sister chromatids

Section 4.1

2. Humans have 46 chromosomes in each somatic cell.
   a. How many chromosomes does a child receive from its father?
   b. How many autosomes and how many sex chromosomes are present in each somatic cell?
   c. How many chromosomes are present in a human ovum?
   d. How many sex chromosomes are present in a human ovum?

Section 4.2

3. The figure that follows shows the metaphase chromosomes of a male of a particular species. These chromosomes are prepared as they would be for a karyotype, but they have not yet been ordered in pairs of decreasing size.
   a. How many chromosomes are shown?
   b. How many chromatids are shown?
   c. How many centromeres are shown? (Count each sister centromere separately.)
   d. How many pairs of homologous chromosomes are shown?
   e. How many chromosomes on the figure are metacentric? Acrocentric?
   f. What is the likely mode of sex determination in this species? What would you predict to be different about the karyotype of a female in this species?

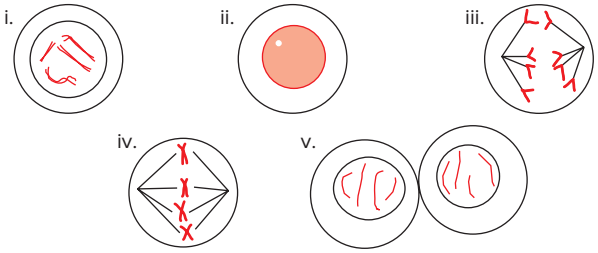


4. Human XX males who are sex-reversed because they have a mutant X chromosome like that shown in Fig. 4.7 often learn of their condition when they want to have children and discover that they are sterile. Can you explain why they are sterile?
5. Researchers discovered recently that the sole function of the SRY protein is to activate an autosomal gene called *Sox9* in the presumptive gonad (before it has "decided" to become a testis or an ovary).
   a. What would be the sex of an XY individual homozygous for nonfunctional mutant alleles of *Sox9?* Explain.
   b. Given your answer to part (a), why is *SRY*, rather than *Sox9*, considered the male determining factor? (*Hint:* What do you think would happen if you did an experiment like the one in the Fast Forward Box *Transgenic Mice Prove That* SRY *Is the Maleness Factor,* except that you used a *Sox9* transgene instead of *SRY?*)

Section 4.3

6. One oak tree cell with 14 chromosomes undergoes mitosis. How many daughter cells are formed, and what is the chromosome number in each cell?

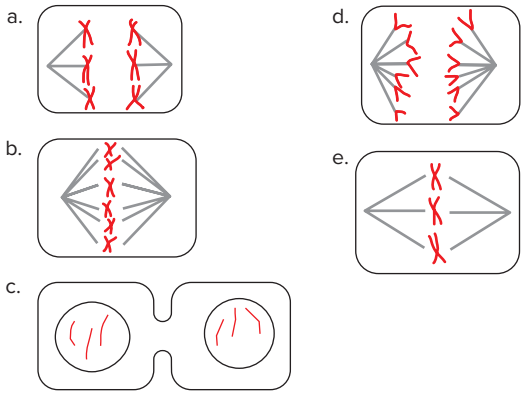
7. Indicate which of the cells numbered i–v matches each of the following stages of mitosis:
   a. anaphase
   b. prophase
   c. metaphase
   d. G_2
   e. telophase/cytokinesis



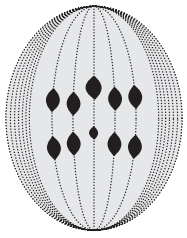
8. a. What are the four major stages of the cell cycle?
   b. Which stages are included in interphase?
   c. What events distinguish G_1, S, and G_2?
9. Answer the questions that follow for each stage of the cell cycle (G_1, S, G_2, prophase, metaphase, anaphase, telophase). If necessary, use an arrow to indicate a change that occurs during a particular cell cycle stage (for example, 1 → 2 or yes → no).
   a. How many chromatids make up each chromosome during this stage?
   b. Is the nucleolus present?
   c. Is the mitotic spindle organized?
   d. Is the nuclear membrane present?
10. Can you think of anything that would prevent mitosis from occurring in a cell whose genome is haploid?

Section 4.4

11. One oak tree cell with 14 chromosomes undergoes meiosis. How many cells will result from this process, and what is the chromosome number in each cell?
12. Which type(s) of cell division (mitosis, meiosis I, meiosis II) reduce(s) the chromosome number by half? Which type(s) of cell division can be classified as reductional? Which type(s) of cell division can be classified as equational?
13. Complete the following statements using as many of the following terms as are appropriate: mitosis, meiosis I (first meiotic division), meiosis II (second meiotic division), and none (not mitosis nor meiosis I nor meiosis II).
   a. The spindle apparatus is present in cells undergoing _____.
   b. Chromosome replication occurs just prior to _____.
   c. The cells resulting from _____ in a haploid cell have a ploidy of *n*.
   d. The cells resulting from _____ in a diploid cell have a ploidy of *n*.
   e. Homologous chromosome pairing regularly occurs during _____.
   f. Nonhomologous chromosome pairing regularly occurs during _____.
   g. Physical recombination leading to the production of recombinant progeny classes occurs during _____.
   h. The separation of sister centromeres occurs during _____.
   i. Nonsister chromatids are found in the same cell during _____.
14. The five cells shown in figures a–e are all from the same individual. For each cell, indicate whether it is in mitosis, meiosis I, or meiosis II. What stage of cell division is represented in each case? What is *n* in this organism?



15. One of the first microscopic observations of chromosomes in cell division was published in 1905 by Nettie Stevens. Because it was hard to reproduce photographs at the time, she recorded these observations as *camera lucida* sketches. One such drawing, of a completely normal cell division in the mealworm *Tenebrio molitor,* is shown here. The techniques of the time were relatively unsophisticated by today's standards, and they did not allow her to resolve chromosomal structures that must have been present.



   a. Describe in as much detail as possible the kind of cell division and the stage of division depicted in the drawing.

b. What chromosomal structure(s) cannot be resolved in the drawing?

c. How many chromosomes are present in normal *Tenebrio molitor* gametes?

16. A person is simultaneously heterozygous for two autosomal genetic traits. One is a recessive condition for albinism (alleles *A* and *a*); this albinism gene is found near the centromere on the long arm of an acrocentric autosome. The other trait is the dominantly inherited Huntington disease (alleles *HD* and HD^{+}). The Huntington gene is located near the telomere of one of the arms of a metacentric autosome. Draw all copies of the two relevant chromosomes in this person as they would appear during metaphase of (a) mitosis, (b) meiosis I, and (c) meiosis II. In each figure, label the location on every chromatid of the alleles for these two genes, assuming that no recombination takes place.

17. Assuming (i) that the two chromosomes in every homologous pair carry different alleles of some genes, and (ii) that no crossing-over takes place, how many genetically different offspring could any one human couple potentially produce? Which of these two assumptions (i or ii) is more realistic?

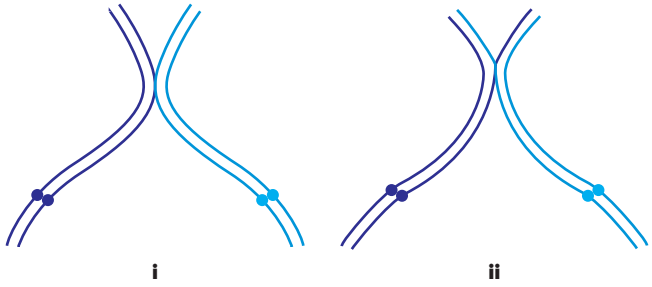
18. In the moss *Polytrichum commune,* the haploid chromosome number is 7. A haploid male gamete fuses with a haploid female gamete to form a diploid cell that divides and develops into the multicellular *sporophyte.* Cells of the sporophyte then undergo meiosis to produce haploid cells called *spores.* What is the probability that an individual spore will contain a set of chromosomes all of which came from the male gamete? (Assume that no recombination occurs.)

19. Can you think of anything that would prevent meiosis from occurring in an organism whose genome is always haploid?

20. Sister chromatids are held together through metaphase of mitosis by complexes of *cohesin* proteins that form rubber band–like rings bundling the two sister chromatids. Cohesin rings are found both at centromeres and at many locations scattered along the length of the chromosomes. The rings are destroyed by protease enzymes at the beginning of mitotic anaphase, allowing the sister chromatids to separate.

a. Cohesin complexes between sister chromatids are also responsible for keeping homologous chromosomes together until anaphase of meiosis I. With this point in mind, which of the two diagrams that follow (i or ii) properly represents the arrangement of chromatids during prophase through metaphase of meiosis I? Explain.

b. What does your answer to part (a) allow you to infer about the nature of cohesin complexes at the centromere versus those along the chromosome arms? Suggest a molecular hypothesis to explain your inference.



21. The pseudoautosomal regions (PARs) of the X and Y chromosomes enable the sex chromosomes to pair and synapse during meiosis in males. Given the location of the *SRY* gene near PAR1, can you propose a mechanism for how the mutant X and Y chromosomes in Fig. 4.7 (in which part of the X is on the Y, and part of the Y is on the X) may have arisen during meiosis?

22. Remarkably, the platypus has 10 sex chromosomes, the largest number found in any mammal. The female platypus has five pairs of different X chromosomes (X1–X5), while the male has X1–X5, and also five different Ys (Y1–Y5). During meiosis in the male, the five Xs always end up together in one gamete, and the five Ys always end up together in another gamete. To achieve this segregation, during prophase of meiosis I the sex chromosomes form a long chain, always in the order X1 Y1 X2 Y2 X3 Y3 X4 Y4 X5 Y5, in which the chromosomes are held together through pseudoautosomal regions (PARs).

a. How many different PARs must exist to allow the formation of these chains? (*Hint:* To answer this question, try drawing the chain of chromosomes.)

b. In terms of pairing ability of the PARs, explain the structural differences between the human and platypus sex chromosomes.

Section 4.5

23. Somatic cells of chimpanzees contain 48 chromosomes.

How many chromatids and chromosomes are present at: (a) anaphase of mitosis, (b) anaphase I of meiosis, (c) anaphase II of meiosis, (d) G_1 prior to mitosis, (e) G_2 prior to mitosis, (f) G_1 prior to meiosis I, and (g) prophase of meiosis I?

How many chromatids or chromosomes are present in: (h) an oogonial cell prior to S phase, (i) a spermatid, (j) a primary oocyte arrested prior to ovulation, (k) a secondary oocyte arrested prior to fertilization, (l) a second polar body, and (m) a chimpanzee sperm?

24. In humans:
    a. How many sperm develop from 100 primary spermatocytes?
    b. How many sperm develop from 100 secondary spermatocytes?
    c. How many sperm develop from 100 spermatids?
    d. How many ova develop from 100 primary oocytes?
    e. How many ova develop from 100 secondary oocytes?
    f. How many ova develop from 100 polar bodies?
25. Women sometimes develop benign tumors called *ovarian teratomas* or *dermoid cysts* in their ovaries. Such a tumor begins when a primary oocyte escapes from its prophase I arrest and finishes meiosis I within the ovary. (Normally meiosis I does not finish until the primary oocyte is expelled from the ovary upon ovulation.) The secondary oocyte then develops as if it were an embryo. Development is disorganized, however, and results in a tumor containing differentiated diploid tissues, including teeth, hair, bone, muscle, and nerve. If a dermoid cyst forms in a woman whose genotype is *Aa,* what are the possible genotypes of the cyst, assuming no recombination?
26. In a certain strain of turkeys, unfertilized eggs sometimes develop parthenogenetically to produce diploid offspring. (Females have ZW and males have ZZ sex chromosomes. Assume that WW cells are inviable.) What distribution of sexes would you expect to see among the parthenogenetic offspring according to each of the following models for how parthenogenesis occurs?
    a. The eggs develop from oogonial cells that never undergo meiosis.
    b. The eggs go all the way through meiosis and then duplicate their chromosomes to become diploid.
    c. The eggs go through meiosis I, and the chromatids separate to create diploidy.
    d. The egg goes all the way through meiosis and then fuses at random with one of its three polar bodies (this scenario assumes the first polar body goes through meiosis II).

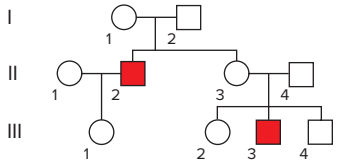
Section 4.6

27. Imagine you have two pure-breeding lines of canaries, one with yellow feathers and the other with brown feathers. In crosses between these two strains, yellow female × brown male gives only brown sons and daughters, while brown female × yellow male gives only brown sons and yellow daughters. Propose a hypothesis to explain these results.
28. A system of sex determination known as *haplodiploidy* is found in honeybees. Females are diploid, and males (drones) are haploid. Male offspring result from the development of unfertilized eggs. Sperm are produced by mitosis in males and fertilize eggs in the females. Ivory eye is a recessive characteristic in honeybees; wild-type eyes are brown.
    a. What progeny would result from an ivory-eyed queen and a brown-eyed drone? Give both genotype and phenotype for progeny produced from fertilized and nonfertilized eggs.
    b. What would result from crossing a daughter from the mating in part (a) with a brown-eyed drone?
29. In *Drosophila,* the autosomal recessive *brown* eye color mutation displays interactions with both the X-linked recessive *vermilion* mutation and the autosomal recessive *scarlet* mutation. Flies homozygous for *brown* and simultaneously hemizygous or homozygous for *vermilion* have white eyes. Flies simultaneously homozygous for both the *brown* and *scarlet* mutations also have white eyes. Predict the F_1 and F_2 progeny of crossing the following true-breeding parents:
    a. vermilion females × brown males
    b. brown females × vermilion males
    c. scarlet females × brown males
    d. brown females × scarlet males
30. Barred feather pattern is a Z-linked dominant trait in chickens. What offspring would you expect from (a) the cross of a barred hen to a nonbarred rooster? (b) the cross of an F_1 rooster from part (a) to one of his sisters?
31. When Calvin Bridges observed a large number of offspring from a cross of white-eyed female *Drosophila* to red-eyed males, he found very rare white-eyed females and red-eyed males among the offspring. He was able to show that these exceptions resulted from nondisjunction, such that the white-eyed females had received two Xs from the egg and a Y from the sperm, while the red-eyed males had received no sex chromosome from the egg and an X from the sperm. What progeny would have arisen from these same kinds of nondisjunctional events if they had occurred in the male parent? What would their eye colors have been?
32. In a vial of *Drosophila,* a research student noticed several female flies (but no male flies) with *bag* wings each consisting of a large, liquid-filled blister instead of the usual smooth wing blade. When bag-winged females were crossed with wild-type males, 1/3 of the progeny were bag-winged females, 1/3 were normal-winged females, and 1/3 were normal-winged males. Explain these results.

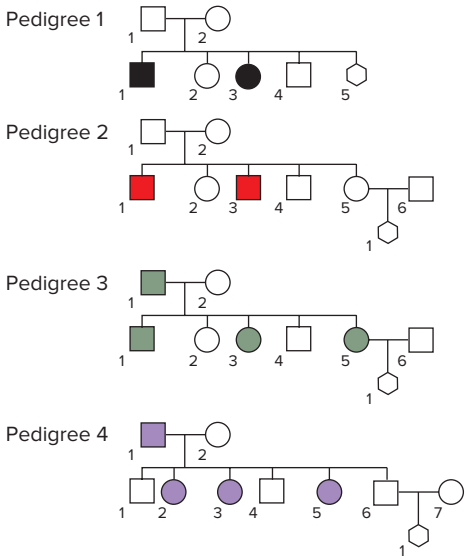
33. In 1919, Calvin Bridges began studying an X-linked recessive mutation causing eosin-colored eyes in *Drosophila.* Within an otherwise true-breeding culture of eosin-eyed flies, he noticed rare variants that had much lighter cream-colored eyes. By intercrossing these variants, he was able to make a true-breeding cream-eyed stock. Bridges now crossed males from this cream-eyed stock with true-breeding wild-type females. All the F_1 progeny had red (wild-type) eyes. When F_1 flies were intercrossed, the F_2 progeny were 104 females with red eyes, 52 males with red eyes, 44 males with eosin eyes, and 14 males with cream eyes. Assume that these numbers represent an 8:4:3:1 ratio.
    a. Formulate a hypothesis to explain the F_1 and F_2 results, assigning phenotypes to all possible genotypes.
    b. What do you predict in the F_1 and F_2 generations if the parental cross is between true-breeding eosin-eyed males and true-breeding cream-eyed females?
    c. What do you predict in the F_1 and F_2 generations if the parental cross is between true-breeding eosin-eyed females and true-breeding cream-eyed males?
34. In *Drosophila,* a cross was made between a yellow-bodied male with vestigial (not fully developed) wings and a wild-type female (brown body). The F_1 generation consisted of wild-type males and wild-type females. F_1 males and females were crossed, and the F_2 progeny consisted of 16 yellow-bodied males with vestigial wings, 48 yellow-bodied males with normal wings, 15 males with brown bodies and vestigial wings, 49 wild-type males, 31 brown-bodied females with vestigial wings, and 97 wild-type females. Explain the inheritance of the two genes in question based on these results.
35. As we learned in this chapter, the *white* mutation of *Drosophila* studied by Thomas Hunt Morgan is X-linked and recessive to wild type. When true-breeding white-eyed males carrying this mutation were crossed with true-breeding purple-eyed females, all the F_1 progeny had wild-type (red) eyes. When the F_1 progeny were intercrossed, the F_2 progeny emerged in the ratio 3/8 wild-type females: 1/4 white-eyed males: 3/16 wild-type males: 1/8 purple-eyed females: 1/16 purple-eyed males.
    a. Formulate a hypothesis to explain the inheritance of these eye colors.
    b. Predict the F_1 and F_2 progeny if the parental cross was reversed (that is, if the parental cross was between true-breeding white-eyed females and true-breeding purple-eyed males).

Section 4.7

36. The following is a pedigree of a family in which a rare form of color blindness is found (filled-in symbols). Indicate as much as you can about the genotypes of all the individuals in the pedigree.



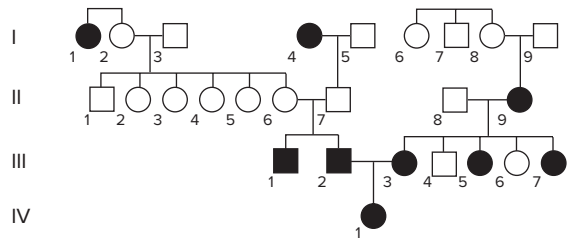
37. Each of the four pedigrees that follow represents a human family within which a genetic disease is segregating. Affected individuals are indicated by filled-in symbols. One of the diseases is transmitted as an autosomal recessive condition, one as an X-linked recessive, one as an autosomal dominant, and one as an X-linked dominant. Assume all four traits are rare in the population and completely penetrant.
    a. Indicate which pedigree represents which mode of inheritance, and explain how you know.
    b. For each pedigree, what would you tell the parents about the chance that their child (indicated by the hexagon shape) will have the condition?



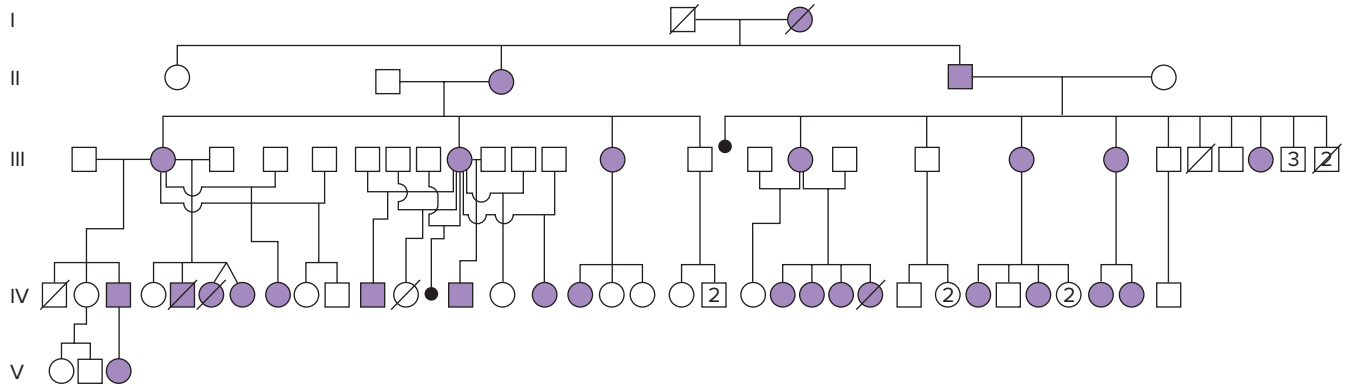
38. The pedigree that follows indicates the occurrence of albinism in a group of Hopi Indians, among whom the trait is unusually frequent. Assume that the trait is fully penetrant (all individuals with a genotype that could give rise to albinism will display this condition).
    a. Is albinism in this population caused by a recessive or a dominant allele?
    b. Is the gene sex-linked or autosomal?

What are the genotypes of the following individuals?

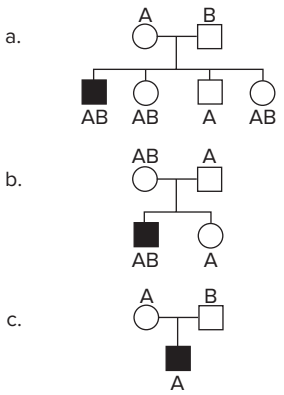
c. individual I-1
d. individual I-8
e. individual I-9
f. individual II-6
g. individual II-8
h. individual III-4



39. Duchenne muscular dystrophy (DMD) is caused by a relatively rare X-linked recessive allele. It results in progressive muscular wasting and usually leads to death before age 20. In this problem, an affected person is one with the severe form of DMD caused by hemizygosity or homozygosity for the disease allele.
    a. What is the probability that the first son of a woman whose brother is affected will be affected?
    b. What is the probability that the second son of a woman whose brother is affected will be affected, if her first son was affected?
    c. What is the probability that a child of an unaffected man whose brother is affected will be affected?
    d. An affected man mates with his unaffected first cousin; there is otherwise no history of DMD in this family. If the mothers of this man and his mate were sisters, what is the probability that the couple's first child will be an affected boy? An affected girl? An unaffected child?
    e. If the two related parents of the couple in part (d) were brother and sister (instead of sisters), what is the probability that the couple's first child will be an affected boy? An affected girl? An unaffected child?
40. The X-linked gene responsible for DMD encodes a protein called dystrophin that is required for muscle function. Dystrophin protein is not secreted—it remains in the cells that produce it. Given what you know about Barr body formation, do you think that females heterozygous for the recessive DMD disease allele could have the disease in some parts of their bodies and not others?
41. Males have hemophilia when they are hemizygous for a nonfunctional recessive mutant allele of the X-linked gene for clotting factor VIII. Factor VIII is normally secreted into the blood serum by cells in the bone marrow (and other specialized cells) that produce it.
    a. Do you think that females heterozygous for the hemophilia disease allele could have hemophilia in some parts of their bodies and not others?
    b. If such a female carrier of hemophilia suffered a cut, would her blood coagulate (form clots) faster, slower, or in about the same time as that of an individual homozygous for a normal allele of the factor VIII gene? Would the rate of clotting vary significantly among heterozygous females?
42. In the Fast Forward Box *Visualizing X Chromosome Inactivation in Transgenic Mice,* suppose the investigators had looked at the expression of green and red fluorescent protein in early mouse embryos, when the embryos have fewer than 500 cells. What patterns would they likely have observed? (Assume that the transgenes make gene product this early in development.)
43. The following pedigree shows five generations of a family that exhibits congenital hypertrichosis, a rare condition in which affected individuals are born with unusually abundant amounts of hair on their faces and upper bodies. The two small black dots in the pedigree indicate miscarriages.
    a. What can you conclude about the inheritance of hypertrichosis in this family, assuming complete penetrance of the trait?
    b. On what basis can you exclude other modes of inheritance?
    c. With how many fathers did III-2 and III-9 have children?



44. Consider the following pedigrees from human families containing a male with Klinefelter syndrome (a set of abnormalities seen in XXY individuals; indicated with shaded boxes). In each, *A* and *B* refer to codominant alleles of the X-linked *G6PD* gene. The *phenotypes* of each individual (A, B, or AB) are shown on the pedigree. Indicate if nondisjunction occurred in the mother or father of the son with Klinefelter syndrome for each of the three examples. Can you tell if the nondisjunction was in the first or second meiotic division?



45. Several different antigens can be detected in blood tests. The following four traits were tested for each individual shown:

ABO type	(I^A and I^B codominant, i recessive)
Rh type	(Rh^+ dominant to Rh^-)
MN type	(M and N codominant)
$Xg^{(a)}$ type	[$Xg^{(a+)}$ dominant to $Xg^{(a-)}$]

All of these blood type genes are autosomal, except for $Xg^{(a)}$, which is X-linked.

Mother	AB	Rh^-	MN	$Xg^{(a+)}$
Daughter	A	Rh^+	MN	$Xg^{(a-)}$
Alleged father 1	AB	Rh^+	M	$Xg^{(a+)}$
Alleged father 2	A	Rh^-	N	$Xg^{(a-)}$
Alleged father 3	B	Rh^+	N	$Xg^{(a-)}$
Alleged father 4	O	Rh^-	MN	$Xg^{(a-)}$

a. Which, if any, of the alleged fathers could be the real father?

b. Would your answer to part (a) change if the daughter had Turner syndrome (the abnormal phenotype seen in XO individuals)? If so, how?

46. The ancestry of a white female tiger bred in a city zoo is depicted in the pedigree following part (e) of this problem. White tigers are indicated with unshaded symbols. (As you can see, there was considerable inbreeding in this lineage. For example, the white tiger Mohan was mated with his daughter.) In answering the following questions, assume that *white* is determined by allelic differences at a single gene and that the trait is fully penetrant. Explain your answers by citing the relevant information in the pedigree.

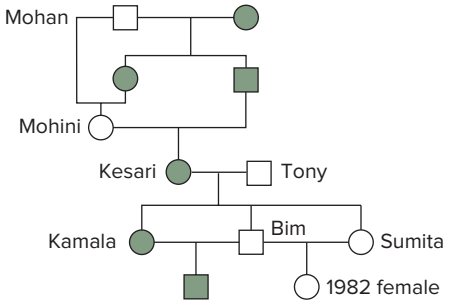
a. Could white coat color be caused by a Y-linked allele?

b. Could white coat color be caused by a dominant X-linked allele?

c. Could white coat color be caused by a dominant autosomal allele?

d. Could white coat color be caused by a recessive X-linked allele?

e. Could white coat color be caused by a recessive autosomal allele?

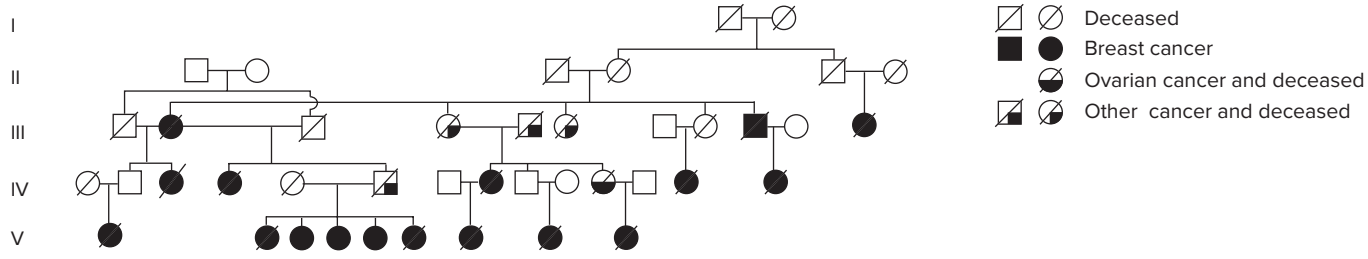


47. The pedigree that follows shows the inheritance of various types of cancer in a particular family. Molecular analyses (described in subsequent chapters) indicate that with one exception, the cancers occurring in the patients in this pedigree are associated with a rare mutation in a gene called *BRCA2*.

a. Which individual is the exceptional cancer patient whose disease is not associated with a *BRCA2* mutation?

b. Is the *BRCA2* mutation dominant or recessive to the normal *BRCA2* allele in terms of its cancer-causing effects?

c. Is the *BRCA2* gene likely to reside on the X chromosome, the Y chromosome, or an autosome? How definitive is your assignment of the chromosome carrying *BRCA2*?



d. Is the penetrance of the cancer phenotype complete or incomplete?

e. Is the expressivity of the cancer phenotype unvarying or variable?

f. Are any of the cancer phenotypes associated with the *BRCA2* mutation sex-limited or sex-influenced?

g. How can you explain the absence of individuals diagnosed with cancer in generations I and II?

48. In 1995, doctors reported a Chinese family in which retinitis pigmentosa (progressive degeneration of the retina leading to blindness) affected only males. All six sons of affected males were affected, but all of the five daughters of affected males (and all of the children of these daughters) were unaffected.

    a. What is the likelihood that this form of retinitis pigmentosa is due to an autosomal mutation showing complete dominance?

    b. What other possibilities could explain the inheritance of retinitis pigmentosa in this family? Which of these possibilities do you think is most likely?

49. In cats, the dominant *O* allele of the X-linked orange gene is required to produce orange fur; the recessive *o* allele of this gene yields black fur. Tortoiseshell cats have coats with patches of orange fur alternating with patterns of black fur. Approximately 90% of all tortoiseshell cats are females.

    a. Explain why tortoiseshell cats are nearly always female.

    b. What types of crosses would be expected to produce female tortoiseshell cats?

    c. Suggest a hypothesis to explain the origin of male tortoiseshell cats.

    d. Calico cats (most of which are females) have patches of white, orange, and black fur. Suggest a hypothesis for the origin of calico cats.

Tortoiseshell Calico

50. In marsupials like the opposum or kangaroo, X inactivation selectively inactivates the paternal X chromosome.

    a. Predict the possible coat colors of the progeny of both sexes if a female marsupial homozygous for a mutant allele of an X-linked coat color gene was mated with a male hemizygous for the alternative wild-type alleles of this gene.

    b. Predict the possible coat colors of progeny of both sexes if a male marsupial hemizygous for an allele of an X-linked coat color gene was mated with a female homozygous for the alternative wild-type allele of this gene.

    c. Why are the terms *recessive* and *dominant* not useful in describing the alleles of X-linked coat color genes in marsupials?

    d. Why would marsupials heterozygous for two alleles of an X-linked coat color gene not have patches of fur of two different colors as did the tortoiseshell cats described in the previous problem?

51. The pedigree diagram below shows a family in which many individuals are affected by a disease called Leri-Weill dyschondrosteosis (LWD). People with LWD are short in stature due to leg bone deformities; arm bones are also malformed in some individuals. The mutant gene responsible for LWD was identified in 1998 as *SHOX*, a gene located in a pseudoautosomal region (PAR1) of the X and Y chromosomes.

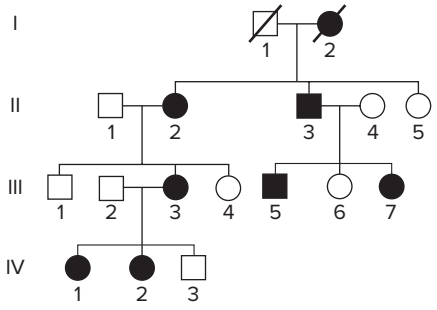
    a. Is the *SHOX* allele that causes LWD dominant or recessive? Explain. (*Note:* Sex reversal is not involved.)

    b. Even though *SHOX* is located on the X chromosome, the pedigree is atypical for an X-linked allele. What features of the pedigree are incompatible with X-linkage?

    c. For each affected individual in the pedigree, determine whether the *SHOX* disease allele is on the X or the Y.

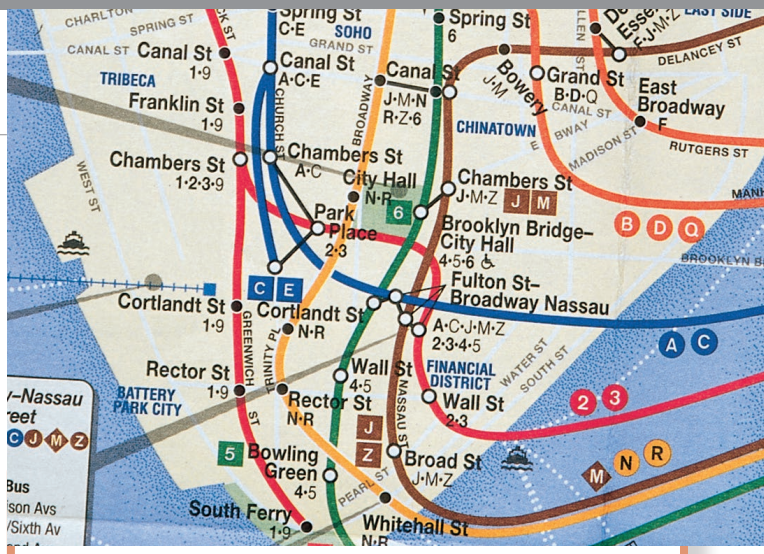
    d. Explain the inheritance pattern of the *SHOX* disease allele and the $SHOX^+$ normal allele in pedigree.

    e. Diagram the crossover event that generated the Y chromosome in individual III-5. Your diagram should indicate the positions of the *SHOX* (disease) and $SHOX^+$ (normal) alleles on the X and Y chromosomes in the germ-line cells of individual II-3, and SRY^+ on Y.



chapter 5

Linkage, Recombination, and the Mapping of Genes on Chromosomes



Maps illustrate the spatial relationships of objects, such as the locations of subway stations along subway lines. Genetic maps portray the positions of genes along chromosomes.

chapter outline

IN 1928, DOCTORS completed a four-generation pedigree tracing two known X-linked traits: red-green color blindness and hemophilia A (the more serious X-linked form of *bleeders disease*). The maternal grandfather of the family exhibited both traits, which means that his single X chromosome carried mutant alleles of the two corresponding genes. As expected, neither color blindness nor hemophilia showed up in his sons and daughters, but two grandsons and one great-grandson inherited both of the X-linked conditions (**Fig. 5.1a**). The fact that none of the descendants manifested one of the traits without the other suggests that the mutant alleles did not assort independently during meiosis. Instead they traveled together in the gametes forming one generation and then into the gametes forming the next generation, producing grandsons and great-grandsons with an X chromosome specifying both color blindness and hemophilia. Genes that travel together more often than not exhibit **genetic linkage.**

In contrast, another pedigree following color blindness and the slightly different B form of hemophilia, which also arises from a mutation on the X chromosome, revealed a different inheritance pattern. A grandfather with hemophilia B and color blindness had four grandsons, but only one of them exhibited both conditions. In this family, the genes for color blindness and hemophilia appeared to assort independently, producing in the male progeny all four possible combinations of the two traits—normal vision and normal blood clotting, color blindness and hemophilia, color blindness and normal clotting, and normal vision and hemophilia—in approximately equal frequencies (**Fig. 5.1b**). Thus, even though the mutant alleles of the two genes were on the same X chromosome in the grandfather, they had to separate to give rise to grandsons III-2 and III-3. This separation of genes on the same chromosome is the result of **recombination,** the occurrence in progeny of new gene combinations not seen in previous generations. (Note that *recombinant progeny* can result in either of two ways: from the recombination of genes on the same chromosome during gamete formation, discussed

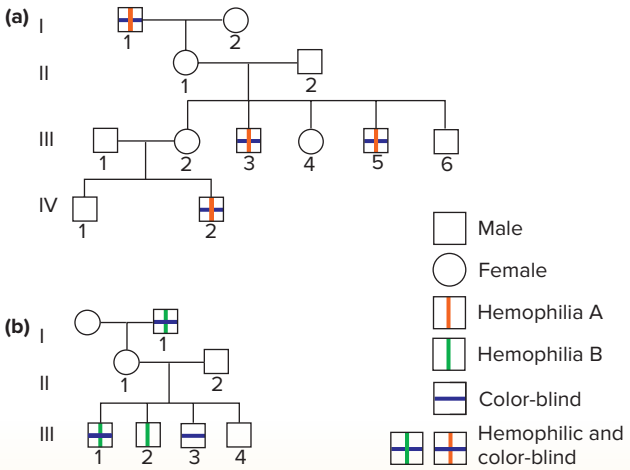


Figure 5.1 Pedigrees indicate that color blindness and two forms of hemophilia are X-linked traits. (a) Transmission of red-green color blindness and hemophilia A. The traits travel together through the pedigree, indicating their genetic linkage. **(b)** Transmission of red-green color blindness and hemophilia B. Even though both genes are X-linked, the mutant alleles are inherited together in only one of four grandsons in generation III. These two pedigrees suggest that the gene for color blindness is close to the hemophilia A gene but far away from the hemophilia B gene.

in this chapter, or from the independent assortment of genes on nonhomologous chromosomes, described in Chapter 4.)

Two important themes should be kept in mind as we follow the transmission of genes linked on the same chromosome. The first is that the farther apart two genes are, the greater is the probability of separation through recombination. Extrapolating from this general rule, you can see that the gene for hemophilia A is likely very close to the gene for red-green color blindness because, as Fig. 5.1a shows, the two rarely separate. By comparison, the gene for hemophilia B must lie far away from the color blindness gene because, as Fig. 5.1b indicates, new combinations of alleles of the two genes occur quite often. A second theme is that geneticists can use data about how often genes separate during transmission to map the genes' relative locations on a chromosome. Such mapping is a key to sorting out and tracking down the components of complex genetic networks; it is also crucial to geneticists' ability to isolate and characterize genes at the molecular level.

5.1 Gene Linkage and Recombination

learning objectives

1. Define linkage with respect to gene loci and chromosomes.
2. Differentiate between parental and recombinant gametes.
3. Conclude from ratios of progeny in a dihybrid cross whether two genes are linked.
4. Explain how a testcross can provide evidence for or against linkage.

If people have roughly 27,000 genes but only 23 pairs of chromosomes, most human chromosomes must carry hundreds, if not thousands, of genes. This is certainly true of the human X chromosome, which contains about 1100 protein-coding genes, as just described in Chapter 4. Recognition that many genes reside on each chromosome raises an important question. If genes on *different* chromosomes assort independently because nonhomologous chromosomes align independently on the spindle during meiosis I, how do genes on the *same* chromosome assort?

Some Genes on the Same Chromosome Do Not Assort Independently—Instead, They Are Linked

We begin our analysis with X-linked *Drosophila* genes because they were the first to be assigned to a specific chromosome. As we outline various crosses, remember that females carry two X chromosomes, and thus two alleles for each X-linked gene. Males, in contrast, have only a single X chromosome (from the female parent), and thus only a single allele for each of these genes.

We look first at two X-linked genes that determine a fruit fly's eye color and body color. These two genes are said to be **syntenic** because they are located on the same chromosome. The *white* gene was introduced in Chapter 4; you will recall that the dominant wild-type allele w^+ specifies red eyes, while the recessive mutant allele w confers white eyes. The alleles of the *yellow* body color gene are y^+ (the dominant wild-type allele for brown bodies) and y (the recessive mutant allele for yellow bodies). To avoid confusion, note that lowercase y and y^+ refer to alleles of the *yellow* gene, while capital Y refers to the Y chromosome (which does not carry genes for either eye or body color). You should also pay attention to the slash symbol (/), which is used to separate genes found on the two chromosomes of a pair (either the X and Y chromosomes as in this case, or a pair of X chromosomes or homologous autosomes). Thus

w y / Y represents the genotype of a male with an X chromosome bearing *w* and *y,* as well as a Y chromosome; phenotypically this male has white eyes and a yellow body.

Detecting linkage by analyzing the gametes produced by a dihybrid

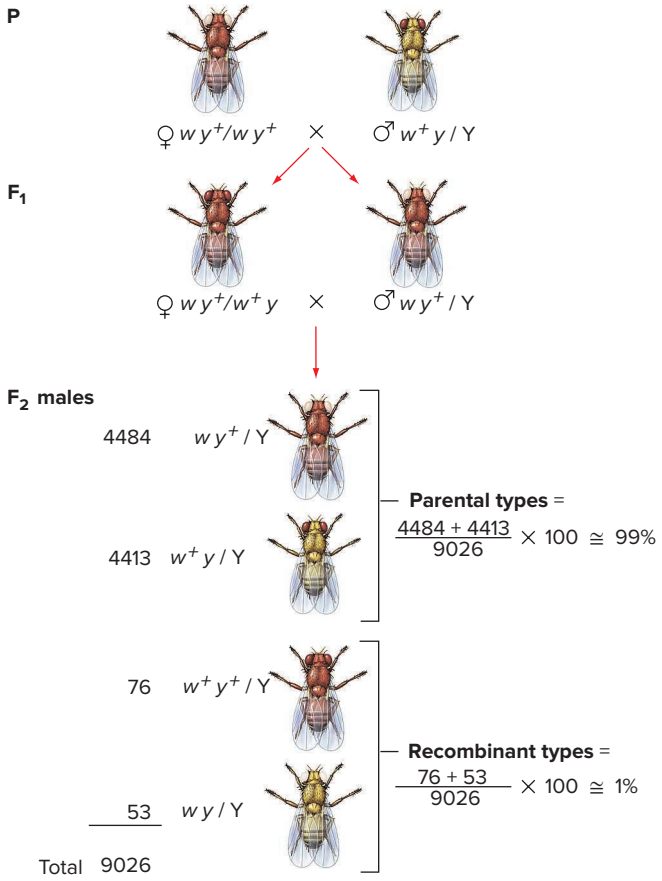
In a cross between a female with mutant white eyes and a wild-type brown body ($w\ y^+/w\ y^+$) and a male with wild-type red eyes and a mutant yellow body ($w^+\ y$ / Y), the F_1 offspring are divided evenly between brown-bodied females with normal red eyes ($w\ y^+/w^+\ y$) and brown-bodied males with mutant white eyes ($w\ y^+$/Y) (**Fig. 5.2**). Note that the male progeny look like their mother because their phenotype directly reflects the genotype of the single X chromosome they received from her. The same is not true for the F_1 females, who received *w* and y^+ on the X from their mother and w^+ *y* on the X from their father. These F_1 females are thus dihybrids. With two alleles for each X-linked gene, one derived from each parent, the dominance relations of each pair of alleles determine the female phenotype.

Now comes the significant cross for answering our question about the assortment of genes on the same chromosome. If these two *Drosophila* genes for eye and body color assort independently, as predicted by Mendel's second law, the dihybrid F_1 females should make four kinds of gametes, with four different combinations of genes on the X chromosome—$w\ y^+$, $w^+\ y$, $w^+\ y^+$, and $w\ y$. These four types of gametes should occur with equal frequency, that is, in a ratio of 1:1:1:1. If it happens this way, approximately half of the gametes will be of the two **parental types,** carrying either the $w\ y^+$ allele combination seen in the original female of the P generation or the w^+y allele combination seen in the original male of the P generation. The remaining half of the gametes will be of two **recombinant types,** in which reshuffling has produced either w^+y^+ or $w\ y$ allele combinations not seen in the P generation parents of the F_1 females.

We can see whether the 1:1:1:1 ratio of the four kinds of gametes actually materializes by counting the different types of male progeny in the F_2 generation, as these sons receive their only X-linked genes from their maternal gamete. The bottom part of Fig. 5.2 depicts the results of a breeding study that produced 9026 F_2 males. The relative numbers of the four X-linked gene combinations passed on by the dihybrid F_1 females' gametes reflect a significant departure from the 1:1:1:1 ratio expected of independent assortment. By far, the largest numbers of gametes carry the parental combinations $w\ y^+$ and w^+y. Of the total 9026 male flies counted, 8897, or almost 99%, had these genotypes. In contrast, the new combinations w^+y^+ and $w\ y$ made up little more than 1% of the total.

We can explain why the two genes fail to assort independently in one of two ways. The $w\ y^+$ and $w^+\ y$ combinations could be preferred because some intrinsic chemical affinity exists between these particular alleles. Alternatively, these combinations of alleles might show up most often because they are parental types. That is, the F_1 female inherited *w* and y^+ together from her P generation mother, and w^+ and *y* together from her P generation father; the F_1 female is then more likely to pass on these parental combinations of alleles, rather than the recombinant combinations, to her own progeny.

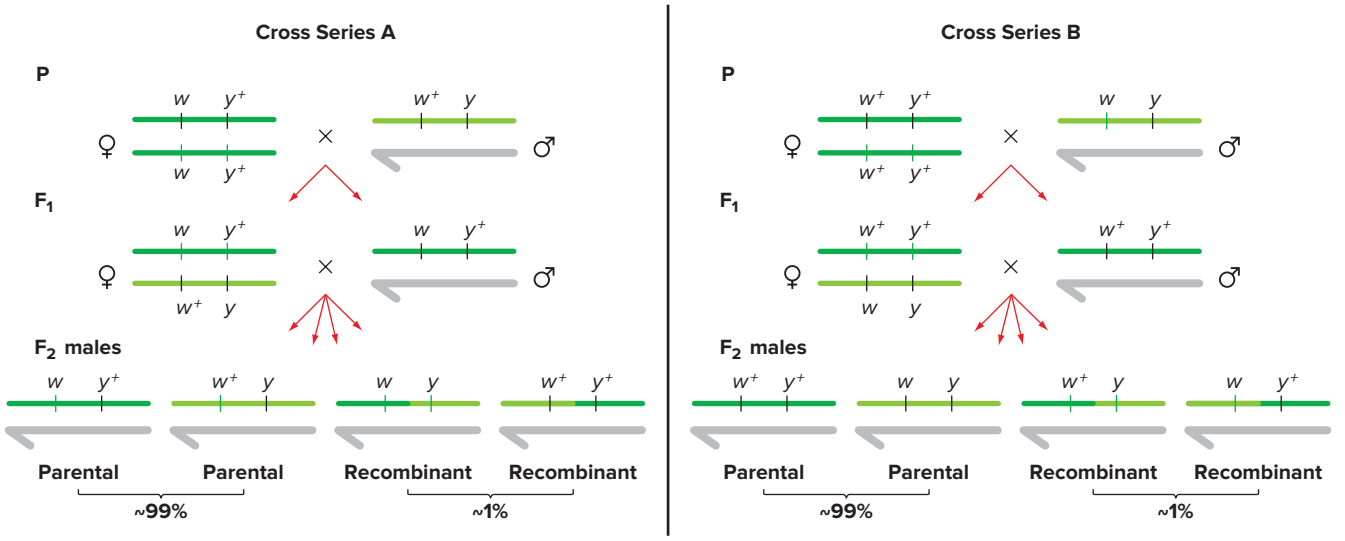
Figure 5.2 When genes are linked, parental combinations outnumber recombinant types. Doubly heterozygous $w\ y^+/\ w^+\ y$ F_1 females produce four types of male offspring. Sons that look like the father ($w^+\ y$ / Y) or mother ($w\ y^+$/ Y) of the F_1 females are parental types. Other sons (w^+y^+/ Y or $w\ y$ / Y) are recombinant types. For these closely-linked genes, many more parental types are produced than recombinant types.



Linkage: A preponderance of parental classes of gametes

A second set of crosses involving the same genes but with a different arrangement of alleles explains why the dihybrid F_1 females do not produce a 1:1:1:1 ratio of the four possible types of gametes (see Cross Series B in **Fig. 5.3**). In this second set of crosses, the original parental generation consists of red-eyed, brown-bodied females ($w^+\ y^+$ / $w^+\ y^+$) and white-eyed, yellow-bodied males ($w\ y$ / Y), and the resultant F_1 females are all $w^+\ y^+$ / $w\ y$ dihybrids. To

Figure 5.3 Designations of *parental* and *recombinant* relate to past history. Figure 5.2 has been redrawn here as **Cross Series A** for easier comparison with **Cross Series B,** in which the dihybrid F_1 females received different allelic combinations of the *white* and *yellow* genes. Note that the parental and recombinant classes in the two cross series are the opposite of each other. The percentages of recombinant and parental types are nonetheless similar in both experiments, showing that the frequency of recombination is independent of the arrangement of alleles.



find out what kinds and ratios of gametes these F_1 females produce, we need to look at the telltale F_2 males.

This time, as Cross Series B in Fig. 5.3 shows, $w^+ y$ / Y and $w\ y^+$ / Y are the recombinants that account for little more than 1% of the total, while $w\ y$ / Y and $w^+ y^+$ / Y are the parental combinations, which again add up to almost 99%. You can see that no preferred association of w^+ and y or of y^+ and w exists in this cross. Instead, a comparison of the two experiments with these particular X chromosome genes demonstrates that the observed frequencies of the various types of progeny depend on how the arrangement of alleles in the F_1 females originated. We have redrawn Fig. 5.2 as Cross Series A in Fig. 5.3 so that you can make this comparison more directly. Note that in both experiments, it is the **parental classes**—the combinations originally present in the P generation—that show up most frequently in the F_2 generation. The reshuffled **recombinant classes** occur less often. It is important to appreciate that the designation of *parental* and *recombinant* gametes or progeny of a doubly heterozygous F_1 female is operational, that is, determined by the particular set of alleles she receives from each of her parents.

When genes assort independently, the numbers of parental and recombinant F_2 progeny are equal because a doubly heterozygous F_1 individual produces an equal number of all four types of gametes. By comparison, two genes are considered **linked** when the number of F_2 progeny with parental genotypes exceeds the number of F_2 progeny with recombinant genotypes. Instead of assorting independently, the genes behave as if they are connected to each other much of the time. The genes for eye and body color that reside on the X chromosome in *Drosophila* are an extreme illustration of the linkage concept. The two genes are so tightly coupled that the parental combinations of alleles—$w\ y^+$ and $w^+\ y$ (in Cross Series A of Fig. 5.3) or $w^+\ y^+$ and $w\ y$ (in Cross Series B)—are reshuffled to form recombinants in only 1 out of every 100 gametes formed. In other words, the two parental allele combinations of these tightly linked genes are inherited together 99 times out of 100.

Gene-pair-specific variation in the degree of linkage

Linkage is not always this tight. In *Drosophila,* a mutation for miniature wings (m) is also found on the X chromosome. A cross of red-eyed females with normal wings ($w^+ m^+ / w^+ m^+$) and white-eyed males with miniature wings ($w\ m$ / Y) yields an F_1 generation containing all red-eyed, normal-winged flies. The genotype of the dihybrid F_1 females is $w^+ m^+ / w\ m$. Of the F_2 males, 67.2% are parental types ($w^+ m^+$ and $w\ m$), while the remaining 32.8% are recombinants ($w\ m^+$ and $w^+\ m$).

This preponderance of parental combinations among the F_2 genotypes reveals that the two genes are linked: The parental combinations of alleles travel together more often than not. But compared to the 99% linkage between the w and y genes for eye color and body color, the linkage of w to m is not that tight. The parental combinations for color and wing size are reshuffled in roughly 33 (instead of 1) out of every 100 gametes.

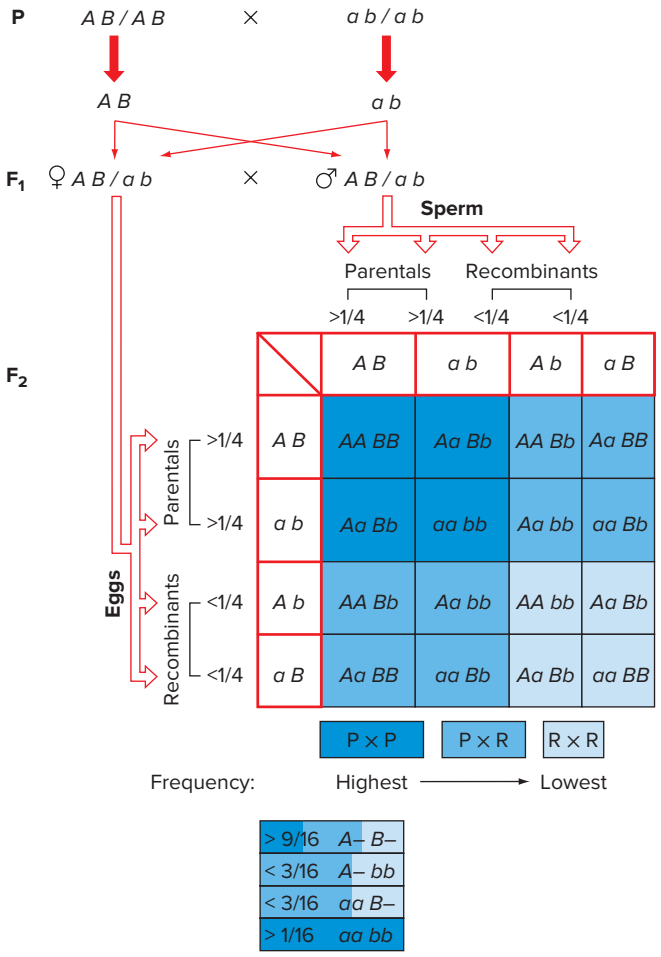
Linkage of autosomal traits

Linked autosomal genes are not inherited according to the 9:3:3:1 Mendelian ratio expected for two independently assorting, noninteracting genes, each with one completely

dominant and one recessive allele. Mendel observed the 9:3:3:1 phenotypic ratio in the F_2 of his dihybrid crosses because the four possible gamete types (*A B*, *A b*, *A B*, and *a b*) were produced at equal frequency by both parents. Equal numbers of each of the four gamete types—independent assortment—means that each one of the 16 boxes in the Punnett square for the F_2 is an equally likely fertilization with a frequency of 1/16 (recall Fig. 2.15).

Had Mendel's two genes been linked, the phenotypic ratio in the F_2 would no longer have been 9:3:3:1 because the parental gametes would have been present at greater frequency than the recombinant gametes. **Figure 5.4** shows the consequences of linkage if the F_1 dihybrid individuals were both of genotype *A B / a b* : The 9/16 and 1/16 classes of F_2 would have increased at the expense of the two 3/16 classes. Conversely, if the alleles of the parents are configured differently (*A b / A b* × *a B / a B*) and the F_1 are therefore *A b / a B*, then the two 3/16 genotypic classes would increase at the expense of the 9/16 and 1/16 classes (*not shown*). Linkage thus undoes the basis of the 9:3:3:1 ratio. Unequal numbers of the four gamete types are produced, so each box of the Punnett square in Fig. 5.4 no longer represents an equally likely fertilization.

Figure 5.4 The 9:3:3:1 ratio is altered when genes *A* and *B* are linked. For linked genes, the F_2 genotypic classes produced most often by parental gametes increase in frequency at the expense of the other classes. In the *A B/a b* dihybrid cross shown here, the *A– B–* and *aa bb* classes in the F_2 will occur at higher frequencies, and the two other classes (*A– bb* and *aa B–*) at lower frequencies than predicted by the 9:3:3:1 ratios. Note that the blue colors and the relative sizes of the boxes in the Punnett square denote the frequencies at which particular genotypic classes will appear in the F_2 generation.

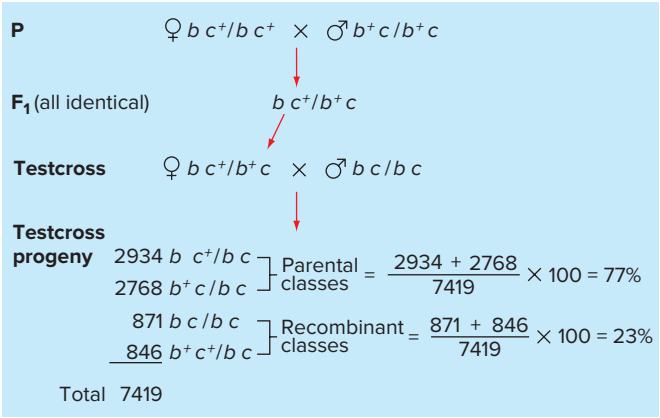


Testcrosses Simplify the Detection of Linkage

Early twentieth-century geneticists found it difficult to interpret crosses involving autosomal genes such as that shown in Fig. 5.4 because it was hard to trace which alleles came from which parent. For example, all the F_2 in Fig. 5.4 with genotype *A– B–* would have the same phenotype, but they could have arisen from fertilizations involving two parental gametes (*dark blue squares*), two recombinant gametes (*light blue squares*), or one of each kind (*medium blue squares*). However, by setting up testcrosses in which one parent was homozygous for the recessive alleles of both genes, as detailed in the next section, geneticists can easily analyze the gene combinations received in the gametes from the other, doubly heterozygous parent.

Fruit flies, for example, carry an autosomal gene for body color (in addition to the X-linked *y* gene); the wild type is once again brown, but a recessive mutation in this gene gives rise to black (*b*). A second gene on the same autosome helps determine the shape of a fruit fly's wing, with the wild type having straight edges and a recessive mutation (*c*) producing curves. **Figure 5.5** depicts a cross between

Figure 5.5 Autosomal genes can also exhibit linkage. A testcross shows that the recombination frequency for the body color (*b*) and wing shape (c) pair of *Drosophila* genes is 23%. Because parentals outnumber recombinants, the *b* and *c* genes are genetically linked and must be on the same autosome.



two pure-breeding strains: black-bodied females with straight wings ($b\ c^+ / b\ c^+$) and brown-bodied males with curved wings ($b^+\ c / b^+\ c$). All the F_1 progeny are double heterozygotes ($b\ c^+ / b^+\ c$) and are phenotypically wild type.

In a testcross of the F_1 females with $b\ c / b\ c$ males, all the offspring receive the recessive *b* and *c* alleles from their father. The phenotypes of the offspring thus indicate the kinds of gametes received from the mother. For example, a black fly with normal wings would be genotype $b\ c^+ / b\ c$; because we know it received the $b\ c$ combination from its father, it must have received $b\ c^+$ from its mother. As Fig. 5.5 shows, roughly 77% of the testcross progeny in one experiment received parental gene combinations (that is, allelic combinations transmitted to the F_1 females by the gametes of each of her parents), while the remaining 23% were recombinants. Because the parental classes outnumbered the recombinant classes, we can conclude that the autosomal genes for black body and curved wings are linked.

essential concepts

- Genes on the same chromosome that do not assort independently are said to be *linked.*
- *Parental gametes* contain alleles inherited together from a single grandparent; *recombinant gametes* contain alleles inherited from different grandparents.
- The hallmark of linkage is that a dihybrid produces more parental gametes than recombinant gametes; as a result, the progeny ratio in a dihybrid cross involving linked genes is not 9:3:3:1.
- *Testcrosses* clarify linkage because each phenotypic class of progeny corresponds to each gamete type produced by the dihybrid parent.

5.2 Recombination: A Result of Crossing-Over During Meiosis

learning objectives

1. Explain the physical process by which recombination takes place.
2. Describe the role of chiasmata in chromosome segregation during meiosis.
3. Discuss the relationship between the recombination frequency and the map distance separating two loci on a chromosome.
4. Explain why the value of the recombination frequency between any two genes is limited to 50%.

It is easy to understand how genes that are physically connected on the same chromosome can be transmitted together and thus show genetic linkage. It is not as obvious why all linked genes always show some recombination in a sample population of sufficient size. Do the chromosomes participate in a physical process that gives rise to the reshuffling of linked genes that we call recombination? The answer to this question is of more than passing interest as it provides a basis for gauging relative distances between pairs of genes on a chromosome.

In 1909, the Belgian cytologist Frans Janssens described structures he had observed in the light microscope during prophase of the first meiotic division. He called these structures *chiasmata;* as described in Chapter 4, they seemed to represent regions in which nonsister chromatids of homologous chromosomes cross over each other (review Fig. 4.16). Making inferences from a combination of genetic and cytological data, Thomas Hunt Morgan suggested that the chiasmata observed through the light microscope were sites of chromosome breakage and exchange resulting in genetic recombination.

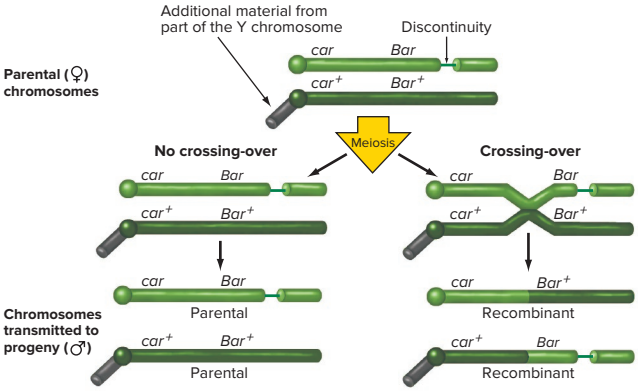
Reciprocal Exchanges Between Homologs Are the Physical Basis of Recombination

Morgan's idea that the physical breaking and rejoining of chromosomes during meiosis was the basis of genetic recombination seemed reasonable. But although Janssens's chiasmata could be interpreted as signs of the process, before 1930 no one had produced visible evidence that crossing-over between homologous chromosomes actually occurs. The identification of **physical markers,** or cytologically visible abnormalities that make it possible to keep track of specific chromosome parts from one generation to the next, enabled researchers to turn the logical deductions about recombination into facts derived from experimental evidence.

In 1931, Harriet Creighton and Barbara McClintock, who studied corn, and Curt Stern, who worked with *Drosophila,* published the results of experiments showing that genetic recombination indeed depends on the reciprocal exchange of parts between maternal and paternal chromosomes. Stern, for example, bred female flies with two different X chromosomes, each containing a distinct physical marker near one of the ends. These same females were also doubly heterozygous for two X-linked **genetic markers**—alleles of genes that could serve as points of reference in determining whether particular progeny were the result of recombination.

Figure 5.6 diagrams the chromosomes of these heterozygous females. One X chromosome carried mutations producing carnation eyes (a dark ruby color, abbreviated *car*) that were kidney-shaped (*Bar*); in addition, this chromosome was marked physically by a visible discontinuity, which resulted when the end of the X chromosome was broken off and attached to an autosome. The other X chromosome had wild-type alleles (+) for both the *car* and the

Figure 5.6 Evidence that recombination results from reciprocal exchanges between homologous chromosomes. Genetic recombination between the *car* and *Bar* genes on the *Drosophila* X chromosome is accompanied by the exchange of physical markers observable in the microscope. Note that this depiction of crossing-over is a simplification, as genetic recombination actually occurs after each chromosome has replicated into sister chromatids.



Bar genes, and its physical marker consisted of part of the Y chromosome that had become connected to the X-chromosome centromere.

Figure 5.6 illustrates how the chromosomes in these *car Bar* / car^+ Bar^+ females were transmitted to male progeny. According to the experimental results, all sons showing a phenotype determined by one or the other parental combination of genes (either *car Bar* or car^+ Bar^+) had an X chromosome that was structurally indistinguishable from one of the original X chromosomes in the mother. In recombinant sons, however, such as those that manifested carnation eye color and normal eye shape (*car* Bar^+ / Y), an identifiable exchange of the abnormal features marking the ends of the homologous X chromosomes accompanied the recombination of genes. The evidence thus tied an instance of genetic recombination to the crossing-over of specifically marked parts of particular chromosomes. This experiment demonstrated elegantly that genetic recombination is associated with the actual reciprocal exchange of segments between homologous chromosomes during meiosis.

Why Recombination?

In Chapter 4, we discussed one advantage that recombination provides for organisms on the earth measured over evolutionary time: Recombination contributes to genetic diversity by reshuffling the alleles of genes between homologous chromosomes (review Fig. 4.17). However, crossing-over also plays another, even more crucial role to ensure that chromosomes segregate properly when they are transmitted between parents and their progeny. As you will see, if recombination did not occur, nondisjunction during meiosis I would be a common, rather than a rare, occurrence, and species could never retain the same number of chromosomes in successive generations.

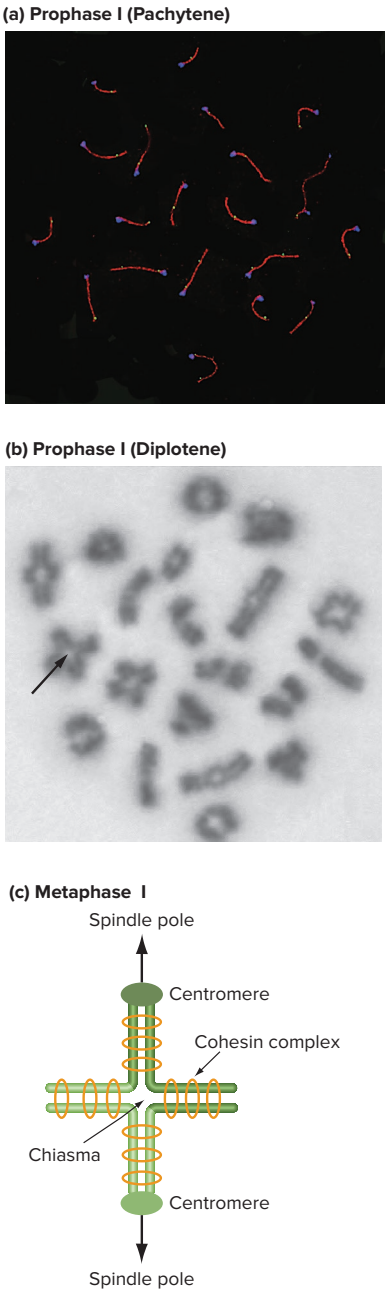
The issue is that proper chromosome segregation requires homologous chromosomes to be pulled to opposite spindle poles, which in turn requires the homologous chromosomes not only to pair with each other during prophase, but also to be linked to each other physically through metaphase until they separate at anaphase. The meiosis I spindle can exert tension on the chromosomes only if the homologs are pulled in opposite directions but remain joined by a physical link. If the tension did not exist, homologous chromosomes would not "know" that they were connected to opposite spindle poles. Without tension, both chromosomes therefore could often connect to fibers from the same spindle pole, and nondisjunction would occur. What then provides the physical link between homologous chromosomes until anaphase of meiosis I?

You might think from Fig. 4.16 that the synaptonemal complex or recombination nodules form the necessary link between homologous chromosomes. **Figure 5.7a** shows an actual fluorescence micrograph of these structures during the middle of prophase I (the pachytene substage). The synaptonemal complexes that help homologous chromosomes to pair with each other are shown in *red*. Although the DNA of the chromosomes is not illustrated in this figure, each red line represents a bivalent (tetrad) made of two homologous chromosomes, each of which has previously been duplicated into sister chromatids. Studded at intervals along the synaptonemal complex are recombination nodules that contain the enzymes responsible for the actual crossing-over; one of these proteins is stained in *green*. However, contrary to expectations, neither synaptonemal complexes nor recombination nodules can link homologous chromosomes until anaphase I begins, for the simple reason that these structures both disappear by the end of prophase I.

Figure 5.7b and **c** illustrate that the homologous chromosomes are still connected to each other even after the synaptonemal complexes and recombination nodules have dissolved. As discussed in Chapter 4, chiasmata mark the sites where recombination actually occurred earlier in prophase I (that is, where nonsister chromatids from homologous chromosomes exchanged places). However, crossing-over by itself is insufficient to keep the homologous chromosomes together. As seen in the artist's diagram in Fig. 5.7c, the physical linkage between homologous chromosomes involves molecular complexes called *cohesin* that make connections between sister chromatids soon after the chromosomes have replicated. Once a crossover takes place, it is cohesin complexes distal to the crossover point (that is, farther away from the centromere than the chiasmata) that keep the homologous chromosomes together at the metaphase plate and thus ensure proper chromosome segregation.

Cohesin plays many roles in the biology of chromosomes; for example, not only is it found along chromosomal

Figure 5.7 Recombination helps ensure proper chromosome segregation during meiosis I. (a) Mouse chromosomes during mid-prophase of meiosis I in a primary spermatocyte. A protein component of the synaptonemal complex is *red,* a component of recombination nodules is *green,* and a component of centromeres is *blue.* Each bivalent has at least one recombination nodule, although some are hard to see. **(b)** Mouse chromosomes in late prophase of meiosis I (diakinesis). Note that each bivalent has at least one chiasma (the *arrow* points to one example), indicating that crossing-over occurred earlier. **(c)** Artist's representation of cohesin complexes (*orange rings*) along the chromosome arms of a bivalent during metaphase of meiosis I; cohesin at the centromeres is *not shown*. Cohesin complexes distal to the crossover point keep sister chromatids together. *Arrows* pointing toward the poles indicate the forces that try to pull the homologous chromosomes apart.



arms, but it also is a key component of the centromeres themselves. We will discuss in detail how cohesin connects sister chromatids during mitosis and meiosis in Chapter 12.

The importance of crossing-over to proper chromosome segregation is underlined by the fact that each bivalent in Figs. 5.7a and b has at least one recombination nodule or chiasma. In fact, a mechanism called *interference* that occurs in almost all sexually reproducing organisms helps ensure that each chromosome pair undergoes at least one crossover, thus preventing nondisjunction of any chromosome except when rare mistakes occur. We discuss interference in more detail later in this chapter.

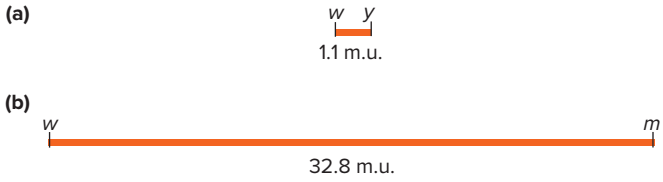
Recombination Frequency Reflects the Distance Between Two Genes

Thomas Hunt Morgan's intuitions that chiasmata represent sites of physical crossing-over between chromosomes and that such crossing-over may result in recombination, led him to the following logical deduction: Different gene pairs exhibit different linkage frequencies because genes are arranged in a line along a chromosome. The closer together two genes are on the chromosome, the smaller their chance of being separated by an event that cuts and recombines the line of genes. To look at it another way, if we assume for the moment that chiasmata can form anywhere along a chromosome with equal likelihood, then the probability of a crossover occurring between two genes increases with the distance separating them. If this is so, the frequency of genetic recombination also must increase with the distance between genes.

To illustrate the point, imagine pinning to a wall a 10-inch piece of ribbon containing tiny black dots along its length and then repeatedly throwing a dart to see where you will cut the ribbon. You would find that practically every throw of the dart separates a dot at one end of the ribbon from a dot at the other end, while few if any throws separate any two particular dots positioned right next to each other.

Alfred H. Sturtevant, one of Morgan's students, took this idea one step further. He proposed that the percentage of total progeny that were recombinant types, the **recombination frequency (RF),** could be used as a gauge of the physical distance separating any two genes on the same chromosome. Sturtevant arbitrarily defined one RF percentage point as the unit of measure along a chromosome; later, another geneticist named the unit a **centimorgan (cM)** after T. H. Morgan. Mappers often refer to a centimorgan as a **map unit (m.u.)**. Although the two terms are interchangeable, researchers prefer one or the other, depending on their experimental organism. *Drosophila* geneticists, for example, use map units while human geneticists use centimorgans. In Sturtevant's system, 1% RF = 1 cM = 1 m.u.

Figure 5.8 Recombination frequencies are the basis of genetic maps. **(a)** 1.1% of the gametes produced by a female doubly heterozygous for the genes *w* and *y* are recombinant. The recombination frequency (RF) is thus 1.1%, and the genes are approximately 1.1 map units (m.u.) or 1.1 centimorgans (cM) apart. **(b)** The distance between the *w* and *m* genes is longer: 32.8 m.u. (or 32.8 cM).



A review of the two pairs of X-linked *Drosophila* genes we analyzed earlier shows how Sturtevant's proposal works. Because the X-linked genes for eye color (*w*) and body color (*y*) recombine in 1.1% of F_2 progeny, they are 1.1 m.u. apart (**Fig. 5.8a**). In contrast, the X-linked genes for eye color (*w*) and wing size (*m*) have a recombination frequency of 32.8 and are therefore 32.8 m.u. apart (**Fig. 5.8b**).

It is easy to see why the fraction of recombinant gametes is a measure of the distance between two genes when you consider individual meioses that can occur in the germline cells of a dihybrid. If genes *A* and *B* are close together on a chromosome, meiosis will usually take place as shown in **Fig. 5.9a,** with no crossovers between the genes (an **NCO** meiosis). In a dihybrid (such as *A B / a b*), the outcome of an NCO meiosis will be four parental gametes (Fig. 5.9a). Occasionally, during a meiosis a single crossover will occur between the two genes (an **SCO** meiosis), resulting in two recombinant and two parental gametes (**Fig. 5.9b**). These crossovers will be rare because most of the length of the chromosome lies outside of the region between genes *A* and *B*. As the distance between the two genes increases, the frequency of SCO meioses increases, and the fraction of recombinant gametes increases.

You can also understand from Fig. 5.9 why the two types of parental gametes (*A B* and *a b* in this example) are produced at roughly equivalent frequencies, and why the two types of recombinant gametes (*A b* and *a B*) appear in approximately equal numbers with respect to each other also. Whenever two nonsister chromatids *do not* cross over between genes *A* and *B*, one of each type of parental chromosome is generated. Likewise, every time two nonsister chromatids *do* cross over between the two genes, both reciprocal recombinants are produced.

As a unit of measure, the map unit is simply an index of recombination probabilities assumed to reflect distances between genes. According to this index, the *y* and *w* genes are much closer together than the *m* and *w* genes. Geneticists have used this logic to map thousands of genetic markers to the chromosomes of *Drosophila,* building recombination maps step-by-step with closely linked markers.

Figure 5.9 Recombinant gametes are produced less frequently than parental gametes when two genes are linked. **(a)** Most meioses occur with no crossovers between tightly linked genes *A* and *B,* resulting in four parental gametes (shaded *orange*). **(b)** Occasionally, a single crossover occurs between *A* and *B* during meiosis, resulting in two recombinant gametes (shaded *blue*) and two parental gametes (*orange*).

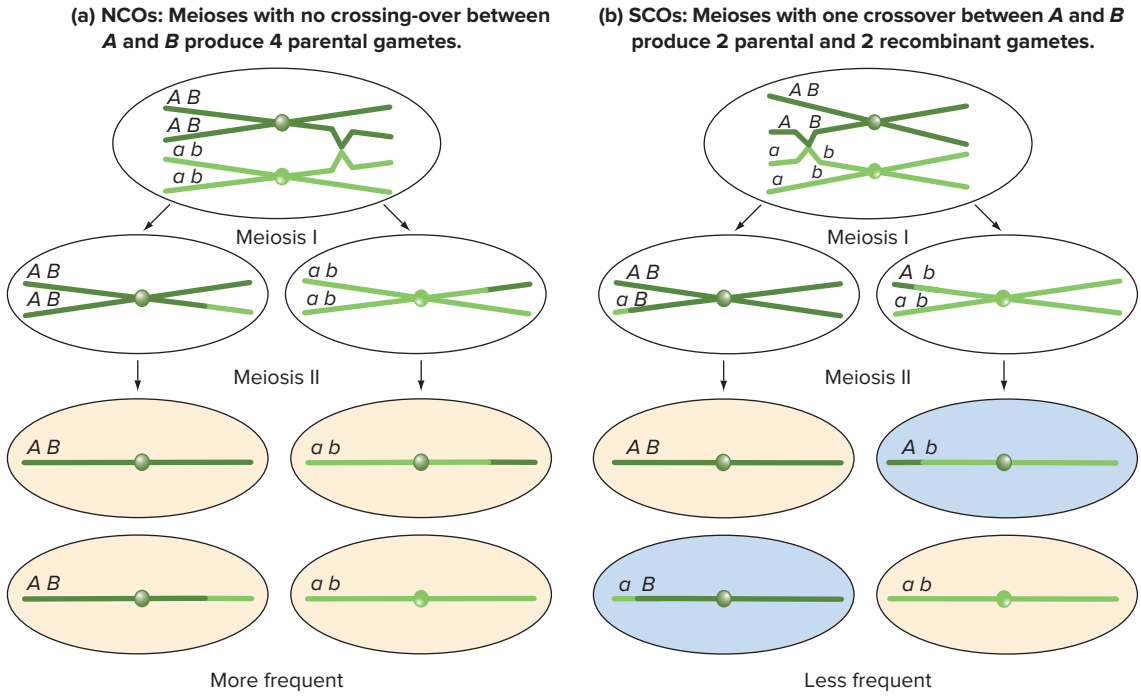
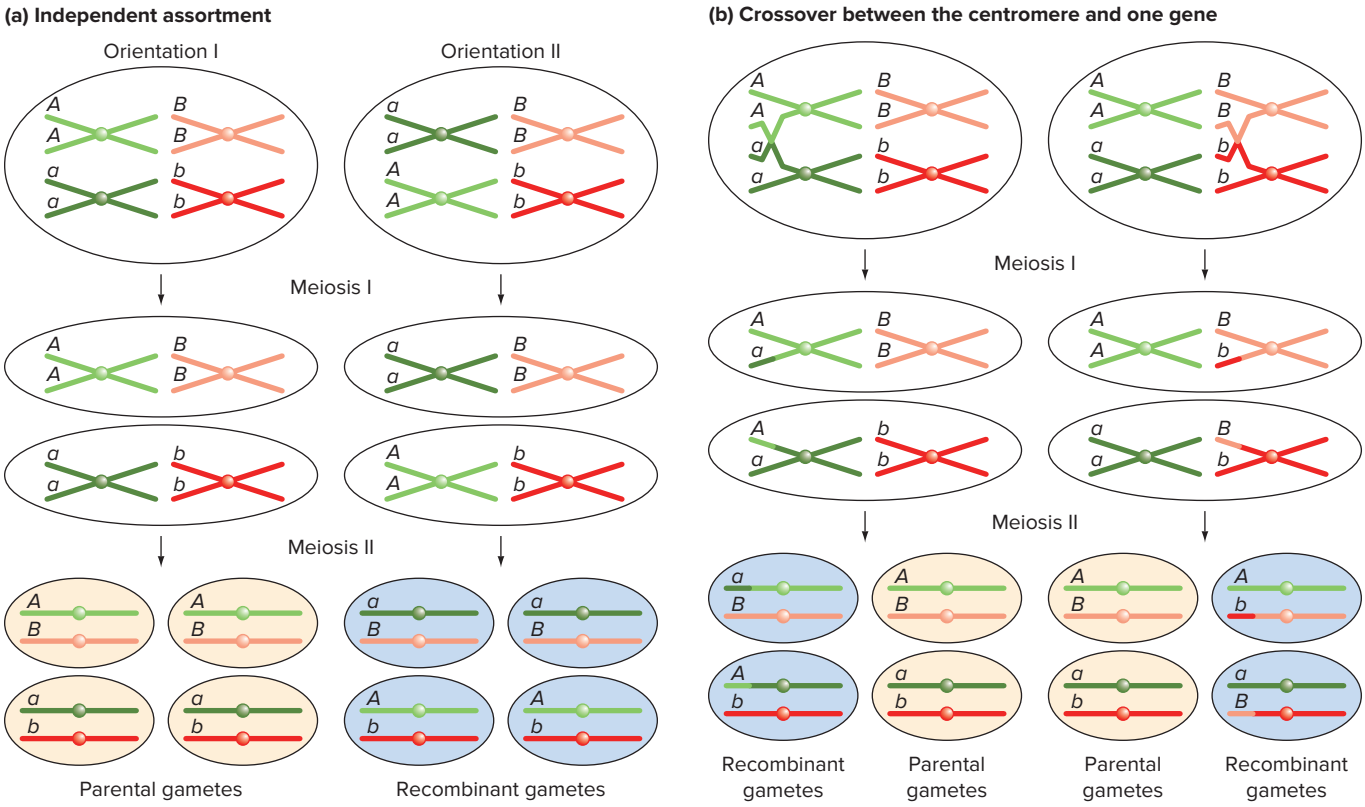


Figure 5.10 Why RF = 50% in a dihybrid for genes on nonhomologous chromosomes. **(a)** Nonhomologous chromosomes line up randomly with respect to one another so that meiosis results in all parental gametes (*left*) or all recombinant gametes (*right*) with equal frequency. **(b)** Meioses that occur with one crossover between one gene and the centromere generate all four gamete types with equal frequency.



Recombination Frequencies Between Two Genes Never Exceed 50%

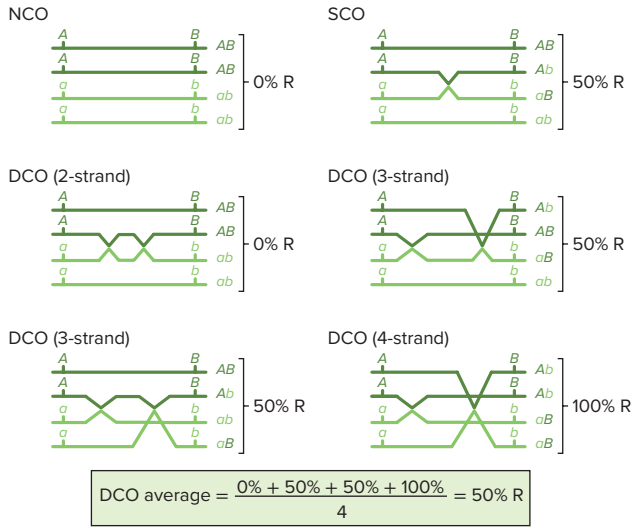
If the definition of linkage is that the proportion of recombinant classes is smaller than that of parental classes, a recombination frequency of less than 50% indicates linkage. But what can we conclude about the relative location of genes if roughly equal numbers of parental and recombinant progeny exist? And does it ever happen that recombinants are in the majority?

We already know one situation that can give rise to a recombination frequency of 50%. Genes located on different (that is, nonhomologous) chromosomes will obey Mendel's law of independent assortment for two reasons. First, the two chromosomes can line up on the spindle during meiosis I in either of two equally likely configurations. As a result, when summed over the products of many meioses, the number of parental and recombinant gametes will be equal (**Fig. 5.10a**). Second, if a crossover occurs between any one gene and the centromere, that meiosis will produce two parental and two recombinant gametes (**Fig. 5.10b**). Thus, if genes *A* and *B* are on nonhomologous chromosomes, a dihybrid will produce all four possible types of gametes (*A B, A b, a B,* and *a b*) with approximately equal frequency.

Importantly, experiments have established that genes located very far apart on the same chromosome also show recombination frequencies of approximately 50%. To understand why this is true, consider the different meioses that contribute to the pool of gametes counted in a testcross experiment. When two genes are very close together on the same chromosome, only two kinds of meioses are likely to occur: those with no crossovers (an NCO meiosis in Fig. 5.9a and **Fig. 5.11**) and rarer meioses with a single crossover (SCO in Fig. 5.9b and Fig. 5.11). An NCO meiosis yields only parental gametes, while an SCO event produces 50% recombinant gametes. These closely spaced genes are linked because some meioses (NCOs) do not make any recombinant-type gametes.

When the two genes are farther apart, SCOs become more frequent, and in some meioses, two crossovers between *A* and *B* (**DCO** meioses) can also occur (Fig. 5.11). DCO meioses can be one of four different types; two, three, or all four nonsister chromatids may cross over. Because the four DCO events are equally frequent, the average fraction of recombinant gametes produced by DCOs is 50% (see equation at the bottom of Fig. 5.11). The same is true for triple crossover events, quadruple crossover events, etc. (*not shown*). Thus, even if the two genes are far enough apart on the same chromosome that at least one crossover

Figure 5.11 Meioses possible in a dihybrid for genes on the same chromosome. Meioses with no crossovers (NCOs) between *A* and *B* result in all parental chromosomes. Meioses with an SCO between *A* and *B* result in half parental and half recombinant (R) chromosomes. Four kinds of DCOs are equally likely, and collectively they result in half parental and half recombinant gametes (*bottom*). In a 2-strand DCO event, one pair of nonsister chromatids undergoes two crossovers. In either of two types of 3-strand DCOs, a single chromatid of one homolog crosses over with each of the two chromatids of the other homolog. In a 4-strand DCO, each pair of nonsister chromatids recombines.



occurs between them in every meiosis, the pool of gametes produced by a dihybrid would be only 50% recombinant types. You can see now that if two genes on the same chromosome are so far apart that no meioses occur as NCOs, then the alleles of the genes assort independently, just as if they were on different chromosomes.

Even though crosses between two genes lying very far apart on a chromosome may show no linkage at all, you can demonstrate that they are on the same chromosome if you can tie each of the widely separated genes to one or more common intermediaries. **Table 5.1** summarizes the relationship between the relative locations of two genes and the presence or absence of linkage as measured by recombination frequencies.

TABLE 5.1 Properties of Linked Versus Unlinked Genes

Linked Genes
Parentals > Recombinants (RF < 50%)
Linked genes must be syntenic and sufficiently close together on the same chromosome so that they do not assort independently.
Unlinked Genes
Parentals = Recombinants (RF = 50%)
Occurs either when two genes are on different chromosomes or when they are sufficiently far apart on the same chromosome that at least one crossover occurs between them in every meiosis.

essential concepts

- *Recombination* occurs when chromatids of homologous chromosomes exchange parts during prophase of meiosis I.
- *Crossing-over* helps establish physical linkages between homologous chromosomes needed to prevent nondisjunction.
- The *recombination frequency* (*RF*) indicates how often two genes are transmitted together. For *linked* genes, RF < 50%.
- Two genes are *linked* if they are physically so close that some meioses occur with no crossovers between them.
- The maximum RF of 50% (independent assortment) occurs when two genes are on different chromosomes, or when two syntenic genes are so far apart that there is at least one crossover between them in every meiosis.

5.3 Mapping: Locating Genes Along a Chromosome

learning objectives

1. Establish relative gene positions using two-point cross data.
2. Refine genetic maps based on data from three-point testcrosses.
3. Explain how a genetic map (in map units) is related to actual physical distance (in base pairs of DNA).
4. Describe the relationship between linkage groups and chromosomes.

Maps are images of the relative positions of objects in space. Whether depicting the floor plan of New York's Metropolitan Museum of Art, the layout of the Roman Forum, or the location of cities served by the railways of Europe, maps turn measurements into patterns of spatial relationships that add a new level of meaning to the original data of distances. Maps that assign genes to specific locations on particular chromosomes called **loci** (singular **locus**) are no exception. By transforming genetic data into spatial arrangements, maps sharpen our ability to predict the inheritance patterns of specific traits.

Geneticists have been obsessed with mapping genes because knowing a gene's location gives scientists the ability to identify the segment of chromosomal DNA corresponding to a gene. In later chapters of this book, you will see how knowledge of a gene's location can be used to isolate its DNA, and furthermore, how molecular geneticists can use a gene's DNA to understand the gene's function.

We have seen that recombination frequency (RF) is a measure of the distance separating two genes along a

chromosome. We now examine how data from many crosses following two and three genes at a time can be compiled and compared to generate accurate, comprehensive gene/chromosome maps.

Comparisons of Two-Point Crosses Establish Relative Gene Positions

In his senior undergraduate thesis, Morgan's student A. H. Sturtevant asked whether data obtained from a large number of two-point crosses (crosses tracing two genes at a time) would support the idea that genes form a definite linear series along a chromosome. Sturtevant began by looking at X-linked genes in *Drosophila.* **Figure 5.12a** lists his recombination data for several two-point crosses. Recall that the distance between two genes that yields 1% recombinant progeny—an RF of 1%—is 1 m.u.

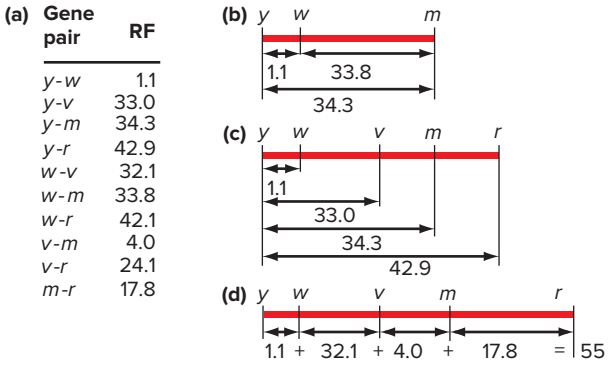
As an example of Sturtevant's reasoning, consider the three genes *w, y,* and *m.* If these genes are arranged in a line (instead of a more complicated branched structure, for example), then one of them must be in the middle, flanked on either side by the other two. The greatest genetic distance should separate the two genes on the outside, and this value should roughly equal the sum of the distances separating the middle gene from each outside gene. The data Sturtevant obtained are consistent with this idea, implying that *w* lies between *y* and *m* (**Fig. 5.12b**). Note that the left-to-right orientation of this map was selected at random; the map in Fig. 5.12b would be equally correct if it portrayed *y* on the right and *m* on the left.

By following exactly the same procedure for each set of three genes considered in pairs, Sturtevant established a self-consistent order for all the genes he investigated on *Drosophila*'s X chromosome (**Fig. 5.12c;** once again, the left-to-right arrangement is an arbitrary choice). By checking the data for every combination of three genes, you can assure yourself that this ordering makes sense. The fact that the recombination data yield a simple linear map of gene position supports the idea that genes reside in a unique linear order along a chromosome.

Limitations of two-point crosses

Though of great importance, the pairwise mapping of genes has several shortcomings that limit its usefulness. First, in crosses involving only two genes at a time, it may be difficult to determine gene order if some gene pairs lie very close together. For example, in mapping *y, w,* and *m,* 34.3 m.u. separate the outside genes *y* and *m,* while nearly as great a distance (33.8 m.u.) separates the middle *w* from the outside *m* (Fig. 5.12b). Before being able to conclude with any confidence that *y* and *m* are truly farther apart, that is, that the small difference between the values of 34.3 and 33.8 is not the result of sampling error, you would have to examine a very large number of flies and subject the data to a statistical test, such as the *chi-square test* that will be explained in the next section.

Figure 5.12 Mapping genes by comparisons of two-point crosses. (a) Sturtevant's data for the distances between pairs of X-linked genes in *Drosophila.* **(b)** Because the distance between *y* and *m* is greater than the distance between *w* and *m,* the order of genes must be *y-w-m.* **(c)** and **(d)** Maps for five genes on the *Drosophila* X chromosome. The left-to-right orientation is arbitrary. Note that the numerical position of the *r* gene depends on how it is calculated. The best genetic maps are obtained by summing many small intervening distances as in (d).



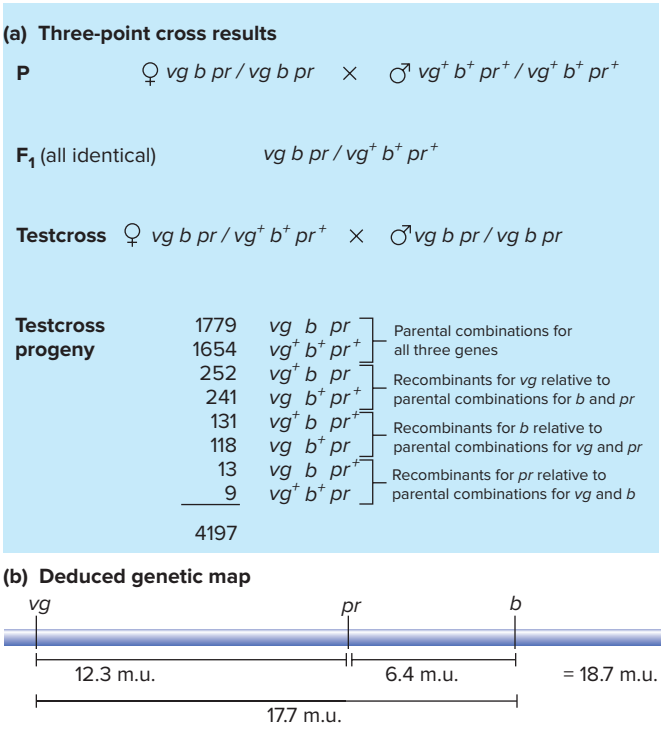
A second problem with Sturtevant's mapping procedure is that the actual distances in his map do not always add up, even approximately. As an example, suppose that the locus of the *y* gene at the far left of the map is regarded as position 0 (Fig. 5.12c). The *w* gene would then lie near position 1, and *m* would be located in the vicinity of 34 m.u. But what about the *r* gene, named for a mutation that produces rudimentary (very small) wings? Based solely on its distance from *y,* as inferred from the $y \leftrightarrow r$ data in Fig. 5.12a, we would place it at position 42.9 (Fig. 5.12c). However, if we calculate its position as the sum of all intervening distances inferred from the data in Fig. 5.12a, that is, as the sum of $y \leftrightarrow w$ plus $w \leftrightarrow v$ plus $v \leftrightarrow m$ plus $m \leftrightarrow r$, the locus of *r* becomes $1.1 + 32.1 + 4.0 + 17.8 = 55.0$ (**Fig. 5.12d**). What can explain this difference, and which of these two values is closer to the truth? Three-point crosses help provide some of the answers.

Three-Point Crosses Provide Faster and More Accurate Mapping

The simultaneous analysis of three markers makes it possible to obtain enough information to position the three genes in relation to each other from just one set of crosses. To describe this procedure, we look at three genes linked on one of *Drosophila*'s autosomes.

A homozygous female with mutations for vestigial wings (*vg*), black body (*b*), and purple eye color (*pr*) was mated to a wild-type male (**Fig. 5.13a**). All the triply heterozygous F_1 progeny, both male and female, had normal

Figure 5.13 Analyzing the results of a three-point cross. **(a)** Results from a three-point testcross of F_1 females simultaneously heterozygous for *vg, b,* and *pr.* **(b)** The gene in the middle must be *pr* because the longest distance is between the other two genes: *vg* and *b.* The most accurate map distances are calculated by summing shorter intervening distances, so 18.7 m.u. is a more accurate estimate of the genetic distance between *vg* and *b* than is 17.7 m.u.



phenotypes, indicating that the mutations are autosomal recessive. In a testcross of the F_1 females with males having vestigial wings, black body, and purple eyes, the progeny were of eight different phenotypes reflecting eight different genotypes. The order in which the genes in each phenotypic class are listed in Fig. 5.13a is completely arbitrary. Thus, instead of *vg b pr,* one could write *b vg pr* or *vg pr b* to indicate the same genotype. Remember that at the outset we do not know the gene order; deducing it is the goal of the mapping study.

In analyzing the data, we look at two genes at a time (recall that the recombination frequency is always a function of a pair of genes). For the pair *vg* and *b,* the parental combinations are *vg b* and $vg^+\ b^+$; the recombinants are *vg* b^+ and $vg^+\ b$. To determine whether a particular class of progeny is parental or recombinant for *vg* and *b,* we do not care whether the flies are *pr* or pr^+. Thus, to the nearest tenth of a map unit, the *vg* ↔ *b* distance, calculated as the percentage of recombinants in the total number of progeny, is

$$\frac{252 + 241 + 131 + 118}{4197} \times 100$$
$$= 17.7 \text{ m. u. } (vg \leftrightarrow b \text{ distance}).$$

Similarly, because recombinants for the *vg–pr* gene pair are *vg* pr^+ and vg^+ *pr,* the interval between these two genes is

$$\frac{252 + 241 + 13 + 9}{4197} \times 100$$
$$= 12.3 \text{ m.u. } (vg \leftrightarrow pr \text{ distance})$$

while the distance separating the *b–pr* pair is

$$\frac{131 + 118 + 13 + 9}{4197} \times 100$$
$$= 6.4 \text{ m.u. } (b \leftrightarrow pr \text{ distance}).$$

These recombination frequencies show that *vg* and *b* are separated by the largest distance (17.7 m.u., as compared with 12.3 and 6.4) and must therefore be the outside genes, flanking *pr* in the middle (**Fig. 5.13b**). But as with the X-linked *y* and *r* genes analyzed by Sturtevant, the distance separating the outside *vg* and *b* genes (17.7) does not equal the sum of the two intervening distances (12.3 + 6.4 = 18.7). In the next section, we learn that the reason for this discrepancy is the rare occurrence of double crossovers.

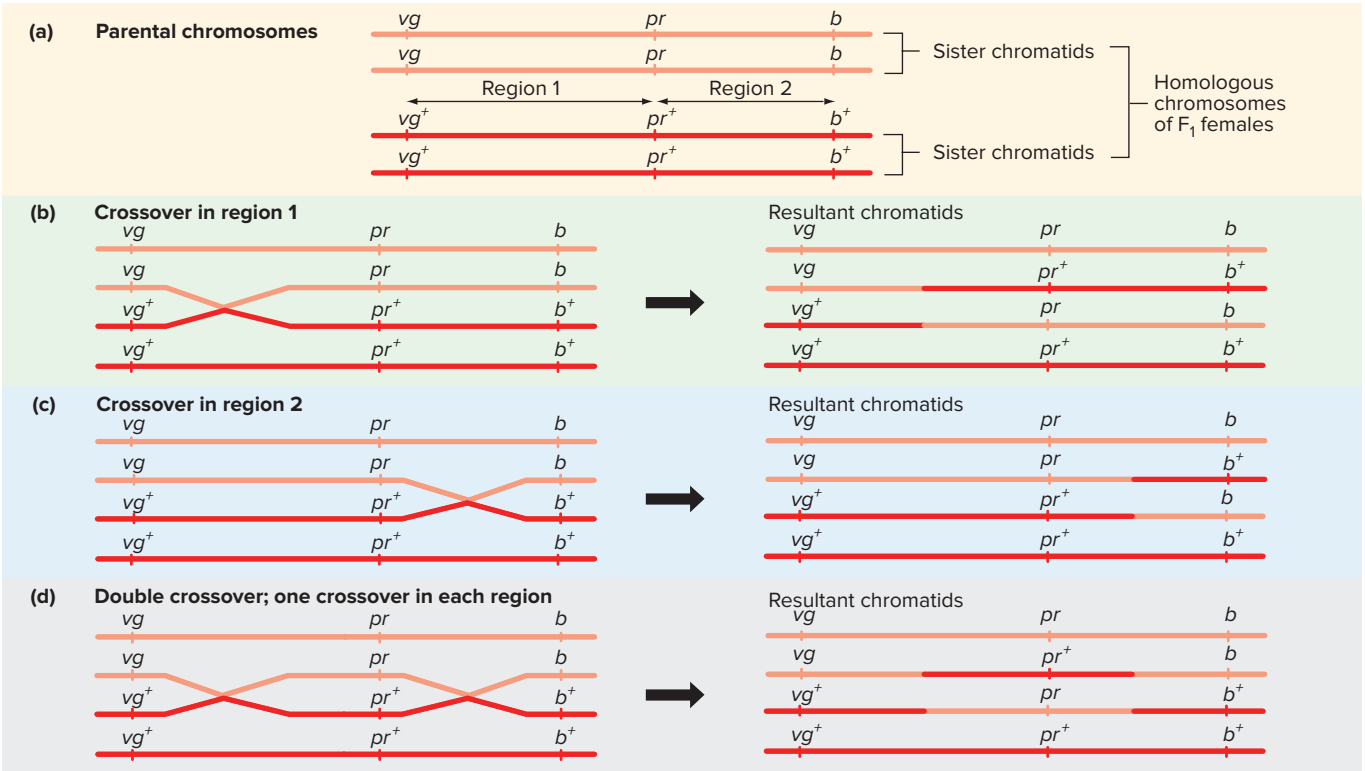
Correction for double crossovers

Figure 5.14 depicts the homologous autosomes of the F_1 females that are heterozygous for the three genes *vg, pr,* and *b.* A close examination of the chromosomes reveals the kinds of crossovers that must have occurred to generate the classes and numbers of progeny observed. In this and subsequent figures, the chromosomes depicted are in late prophase/early metaphase of meiosis I, when there are four chromatids for each pair of homologous chromosomes. As we have suggested previously and demonstrate more rigorously in Chapter 6, prophase I is the stage at which recombination takes place. Note that *region 1* is the space between *vg* and *pr,* and the space between *pr* and *b* is *region 2.*

Recall that the progeny from the testcross performed earlier fall into eight groups (review Fig. 5.13). Flies in the two largest groups carry the same configurations of genes as did their grandparents in the P generation: *vg b pr* and $vg^+\ b^+\ pr^+$; they thus represent the parental classes (Fig. 5.14a). The next two groups—vg^+ *b pr* and *vg* b^+ pr^+—are composed of recombinants that must be the reciprocal products of a crossover in region 1 between *vg* and *pr* (Fig. 5.14b). Similarly the two groups containing vg^+ *b* pr^+ and *vg* b^+ *pr* flies must have resulted from recombination in region 2 between *pr* and *b* (Fig. 5.14c).

But what about the two smallest groups made up of rare *vg b* pr^+ and vg^+ b^+ *pr* recombinants? What kinds of chromosome exchange could account for them? Most likely, they result from two different crossover events occurring simultaneously, one in region 1, the other in region 2 (Fig. 5.14d). The gametes produced by such double crossovers still have the parental configuration for the

Figure 5.14 Inferring the location of a crossover event. Once you establish the order of genes involved in a three-point cross, it is easy to determine which crossover events gave rise to particular recombinant gametes. Note that double crossovers are needed to generate gametes in which the gene in the middle has recombined relative to the parental combinations for the genes at the ends. Such events include the 2-strand DCO in part (d) as well as 3-strand DCOs (*not shown*).



outside genes *vg* and *b,* even though not one but two exchanges must have occurred between them.

Because of the existence of double crossovers, the *vg* ↔ *b* distance of 17.7 m.u. calculated in the previous section does not reflect all of the recombination events producing the gametes that gave rise to the observed progeny. To correct for this oversight, it is necessary to adjust the recombination frequency by adding the double crossovers twice, because each individual in the double crossover groups is the result of two exchanges between *vg* and *b*. The corrected distance is

$$\frac{252 + 241 + 131 + 118 + 13 + 13 + 9 + 9}{4197} \times 100$$

$$= 18.7 \text{ m.u.}$$

This value makes sense because you have accounted for all of the crossovers that occur in region 1 as well as all of the crossovers in region 2. As a result, the corrected value of 18.7 m.u. for the distance between *vg* and *b* is now exactly the same as the sum of the distances between *vg* and *pr* (region 1) and between *pr* and *b* (region 2).

As previously discussed, when Sturtevant originally mapped several X-linked genes in *Drosophila* by two-point crosses, the locus of the rudimentary wings (*r*) gene was ambiguous. A two-point cross involving *y* and *r* gave a recombination frequency of 42.9, but the sum of all the intervening distances was 55.0 (review Fig. 5.12). This discrepancy occurred because the two-point cross ignored double crossovers that might have occurred in the large interval between the *y* and *r* genes. The data summing the smaller intervening distances accounted for at least some of these double crossovers by catching recombinations of gene pairs between *y* and *r*. Moreover, each smaller distance is less likely to encompass a double crossover than a larger distance, so each number for a smaller distance is inherently more accurate.

Note that even a three-point cross like the one for *vg, pr,* and *b* ignores the possibility of two recombination events taking place in, say, region 1. For greatest accuracy, it is always best to construct a map using many genes separated by relatively short distances.

Interference: Fewer double crossovers than expected

In a three-point cross following three linked genes, of the eight possible genotypic classes, the two parental classes contain the largest number of progeny, while the two double recombinant classes, resulting from double crossovers, are always the smallest (see Fig. 5.11). We can understand why

double-crossover progeny are the rarest by looking at the probability of their occurrence. If an exchange in region 1 of a chromosome does not affect the probability of an exchange in region 2, the probability that both will occur simultaneously is the product of their separate probabilities (recall the product rule in Chapter 2). For example, if progeny resulting from recombination in region 1 alone account for 10% of the total progeny (that is, if region 1 is 10 m.u.) and progeny resulting from recombination in region 2 alone account for 20%, then the probability of a double crossover (one event in region 1, the second in region 2) is $0.10 \times 0.20 = 0.02$, or 2%. This makes sense because the likelihood of two rare events occurring simultaneously is even smaller than that of either rare event occurring alone.

If eight classes of progeny are obtained in a three-point cross, the two classes containing the fewest progeny must have arisen from double crossovers. The numerical frequencies of observed double crossovers, however, almost never coincide with expectations derived from the product rule. Let's look at the actual numbers from the cross we have been discussing. The probability of a single crossover between *vg* and *pr* is 0.123 (corresponding to 12.3 m.u.), and the probability of a single crossover between *pr* and *b* is 0.064 (6.4 m.u.). The product of these probabilities is

$$0.123 \times 0.064 = 0.0079 = 0.79\%.$$

But the observed proportion of double crossovers (see Fig. 5.11) was

$$\frac{13 + 9}{4197} \times 100 = 0.52\%.$$

The fact that the number of observed double crossovers is less than the number expected if the two exchanges are independent events suggests that the occurrence of one crossover reduces the likelihood that another crossover will occur in an adjacent part of the chromosome. This phenomenon—of crossovers not occurring independently—is called **chromosomal interference.**

As was shown in Fig. 5.7, interference likely exists to ensure that every pair of homologous chromosomes undergoes at least one crossover. It is crucial that every pair of homologous chromosomes sustain one or more crossovers because such events help the chromosomes orient properly at the metaphase plate during the first meiotic division. Indeed, homologous chromosome pairs without crossovers often segregate improperly. If only a limited number of crossovers can occur during each meiosis, and interference lowers the number of crossovers on large chromosomes, then the remaining possible crossovers are more likely to occur on small chromosomes. This increases the probability that at least one crossover will take place on every homologous pair. Though the molecular mechanism underlying interference is not yet clear, recent experiments suggest that interference is mediated by the synaptonemal complex.

Interference is not uniform and may vary even for different regions of the same chromosome. Investigators can obtain a quantitative measure of the amount of interference in different chromosomal intervals by first calculating a **coefficient of coincidence,** defined as the ratio between the actual frequency of double crossovers observed in an experiment and the number of double crossovers expected on the basis of independent probabilities.

$$\text{coefficient of coincidence} = \frac{\text{frequency observed}}{\text{frequency expected}}$$

For the three-point cross involving *vg, pr,* and *b,* the coefficient of coincidence is

$$\frac{0.52}{0.79} = 0.66.$$

The definition of interference itself is

$$\text{Interference} = 1 - \text{coefficient of coincidence}.$$

In this case, the interference is

$$1 - 0.66 = 0.34.$$

To understand the meaning of interference, it is helpful to contrast what happens when there is no interference with what happens when interference is complete. If interference is 0, the frequency of observed double crossovers equals expectations, and crossovers in adjacent regions of a chromosome occur independently of each other. If interference is complete (that is, if interference = 1), no double crossovers occur in the experimental progeny because one exchange effectively prevents another. As an example, in a particular three-point cross in mice, the recombination frequency for the pair of genes on the left (region 1) is 20, and for the pair of genes on the right (region 2), it is also 20. Without interference, the expected rate of double crossovers in this chromosomal interval is

$$0.20 \times 0.20 = 0.04 \text{ , or } 4\%,$$

but when investigators observed 1000 progeny of this cross, they found 0 double recombinants instead of the expected 40.

A method to determine the gene in the middle

The smallest of the eight possible classes of progeny in a three-point cross are the two that contain double recombinants generated by double crossovers. It is possible to use the composition of alleles in these double crossover classes to determine which of the three genes lies in the middle, even without calculating any recombination frequencies.

Consider again the progeny of a three-point testcross looking at the *vg, pr,* and *b* genes. The F_1 females are *vg pr b / vg^+ pr^+ b^+*. As Fig. 5.14d demonstrated, testcross progeny resulting from double crossovers in the trihybrid females of the F_1 generation received gametes from their mothers carrying the allelic combinations *vg pr^+ b* and *vg^+*

pr b^+. In these individuals, the alleles of the *vg* and *b* genes retain their parental associations (*vg b* and vg^+ b^+), while the *pr* gene has recombined with respect to both the other genes (*pr* b^+ and pr^+ *b*, *vg* pr^+ and vg^+ *pr*). The same is true in all three-point crosses: In those gametes formed by double crossovers, the gene whose alleles have recombined relative to the parental configurations of the other two genes must be the one in the middle.

Three-Point Crosses: A Comprehensive Example

The technique of looking at double recombinants to discover which gene has recombined with respect to both other genes allows immediate clarification of gene order even in otherwise difficult cases. Consider the three X-linked genes *y, w,* and *m* that Sturtevant located in his original mapping experiment (see Fig. 5.12). Because the distance between *y* and *m* (34.3 m.u.) appeared slightly larger than the distance separating *w* and *m* (33.8 m.u.), he concluded that *w* was the gene in the middle. But because of the small difference between the two numbers, his conclusion was subject to questions of statistical significance. If, however, we look at a three-point cross following *y, w,* and *m,* these questions disappear.

Figure 5.15 tabulates the classes and numbers of male progeny arising from females heterozygous for the *y, w,* and *m* genes. Because these male progeny receive their only X chromosome from their mothers, their phenotypes indicate directly the gametes produced by the heterozygous females. In each row of the figure's table, the genes appear in an arbitrary order that does not presuppose knowledge of the actual map. As you can see, the two classes of progeny listed at the top of the table outnumber the remaining six classes, which indicates that all three genes are linked to each other. Moreover, these largest groups, which are the parental classes, show that the two X chromosomes of the heterozygous females were w^+ y^+ *m* and *w y* m^+.

Among the male progeny in Fig. 5.15, the two smallest classes, representing the double crossovers, have X chromosomes carrying w^+ *y* m^+ and *w* y^+ *m* combinations, in which the *w* alleles are recombined relative to those of *y* and *m.* The *w* gene must therefore lie between *y* and *m,* verifying Sturtevant's original assessment.

To complete a map based on the *w y m* three-point cross, you can calculate the interval between *y* and *w* (region 1)

$$\frac{49 + 41 + 1 + 2}{6823} \times 100 = 1.3 \text{ m.u.}$$

as well as the interval between *w* and *m* (region 2)

$$\frac{1203 + 1092 + 2 + 1}{6823} \times 100 = 33.7 \text{ m.u.}$$

The genetic distance separating *y* and *m* is the sum of

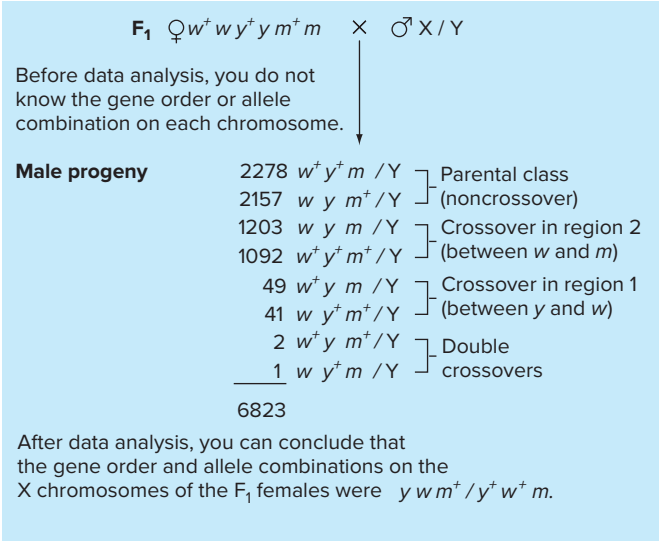
$$1.3 + 33.7 = 35.0 \text{ m.u.}$$

Note that you could also calculate the distance between *y* and *m* directly by including double crossovers twice, to account for the total number of recombination events detected between these two genes.

$$\text{RF} = (1203 + 1092 + 49 + 41 + 2 + 2 + 1 + 1)/6823 \times 100 = 35.0 \text{ m.u.}$$

This method yields the same value as the sum of the two intervening distances (region 1 + region 2).

Figure 5.15 How three-point crosses verify Sturtevant's map of the *Drosophila* X chromosome. The parental classes correspond to the two X chromosomes in the F_1 female. The genotype of the double recombinant classes shows that *w* must be the gene in the middle.



How Do Genetic Maps Correlate with Physical Reality?

Many types of experiments presented throughout this book show resoundingly that the *order of genes* revealed by recombination mapping always reflects the order of those same genes along the DNA molecule of a chromosome. In contrast, the actual *physical distances between genes*—that is, the amount of DNA separating them—does not always correspond linearly to genetic map distances.

Underestimation of physical distances between genes by recombination frequency

You have already seen that DCOs between two genes may go undetected in a testcross experiment, resulting in undercounting of the number of crossovers between the two

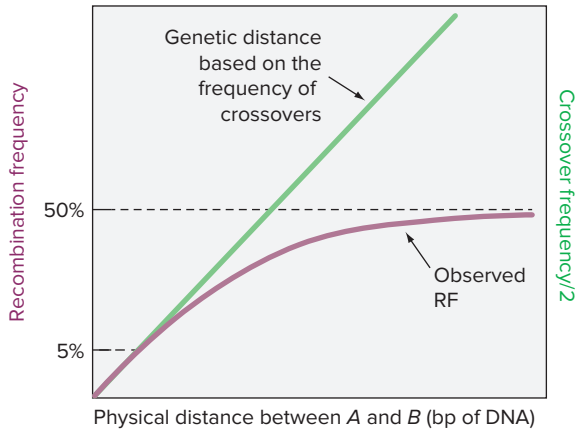
genes, and therefore an underestimation of the distance between them. This is not much of a problem when two genes are close enough together that DCOs take place infrequently. However, as the distance between two genes increases, double and multiple crossovers occur often enough to affect the relationship between recombination frequency and map distance. This relationship cannot be linear because, as we have already seen, the RF for a two-point cross cannot exceed 50% regardless of how far apart two genes are on the same chromosome.

A second look at Fig. 5.11 makes it easy to see how DCOs result in underestimation of gene distances by RF. When genes *A* and *B* are close together, most meioses are NCOs, and the occasional meiosis is an SCO. Each SCO produces exactly two recombinant gametes, and so a perfect linear correspondence exists between the number of crossovers and the number of recombinant gametes (1 crossover: 2 recombinants). However, when genes *A* and *B* are farther apart, DCOs occur. Only one of the four equally frequent DCOs (4-strand) preserves the linear relationship between crossovers and recombinant gametes: Two crossovers occur in a 4-strand DCO and four recombinant gametes (2 crossovers: 4 recombinants). In contrast, the other three types of DCOs result in fewer than four recombinant gametes.

Figure 5.16 illustrates this discrepancy between the actual number of crossovers (*green* line) and observed RF (*purple* line) as a function of the amount of DNA separating the two genes. As you can see, the two graphs are nearly identical for distances of 5 m.u. or less. At genetic distances this small, the RF seen in a two-point cross is an accurate measure of the physical distance. The two graphs then diverge increasingly from each other, so the RF becomes a less precise estimate at genetic distances greater than 5 m.u.

Geneticists have developed mathematical equations called *mapping functions* to compensate for the inaccuracies inherent in relating recombination frequencies to physical distances. However, the corrections for large distances are at best imprecise because mapping functions are based on simplifying assumptions that are not completely true. Thus, the best way to create an accurate map is by summing many smaller intervals, locating widely separated genes through linkage to common intermediaries. Maps are subject to continual refinement as more and more newly discovered genes are included.

Figure 5.16 Recombination frequency underestimates crossover frequency. The lines representing the RF observed in a testcross (*purple*) and the crossover frequency (*green*) are nearly coincident when genes *A* and *B* are close together. As two genes become farther apart, the RF value increasingly underestimates the actual crossover frequency; eventually no NCOs will take place and the maximum RF of 50% will be observed. The *green* line represents crossover frequency/2 because every SCO produces 1/2 parental and 1/2 recombinant chromosomes.

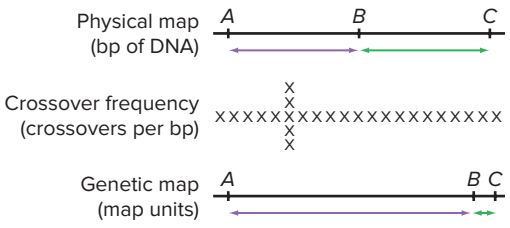


Nonuniform crossover frequencies

Although we have been assuming thus far that crossovers are just as likely to occur between any two base pairs along a chromosome, recombination is not in fact random. In human DNA, for example, most crossovers take place in so-called **recombination hotspots**—small regions of DNA where the frequency of recombination is much higher than average. As shown in **Fig. 5.17,** genes with hotspots between them (*A* and *B*) will be more distant from each other on a genetic map (measured in m.u.) than another gene pair (*B* and *C*) without a hotspot between them, even though the physical distance (measured in bp of DNA) separating each gene pair is the same. As hotspots are relatively frequent (a hotspot appears about every 50,000 bp in human chromosomes), recombination frequency is nevertheless a reasonable estimation of physical distance between most genes.

Frequencies of recombination may also differ from species to species. We know this because recent elucidation of the complete DNA sequences of several organisms' genomes has allowed investigators to compare the physical distances between genes with genetic map distances. They found that in humans, a map unit corresponds on average to

Figure 5.17 Recombination hotspots. Genes *A* and *B* are separated by the same number of base pairs as genes *B* and *C*. Because *A* and *B* flank a recombination hotspot, they appear much more distant from each other than do genes *B* and *C* on a genetic map.



FAST FORWARD

Sprinters: © Robert Michael/Corbis RF

Mapping the Crossovers that Generate the Chromosomes of Individual Human Sperm

Using DNA analysis technologies that will be described in Chapters 9 through 11, scientists can now examine the base pair sequences of the whole genomes of single sperm. In one such study, by comparing the DNA sequences in each of the pairs of homologous chromosomes in a man's somatic cells with those of individual sperm the same man produces, researchers could locate specific recombination events that occurred in the man's primary spermatocytes.

The homologous chromosomes any person inherits from his or her father or mother differ in about 1 out of every 1000 base pairs. The base pair differences in different genomes are called *SNPs* (pronounced *snips;* for *s*ingle *n*ucleotide *p*olymorphisms). From comparisons of the genome sequences of many individuals, approximately 50,000,000 locations in the genome have been identified where SNPs commonly can occur. The different base pair sequences of SNPs are considered different alleles of the SNP locus (Fig. A). Researchers can zero in on SNP loci and determine which alleles of millions of SNPs are present in a genome.

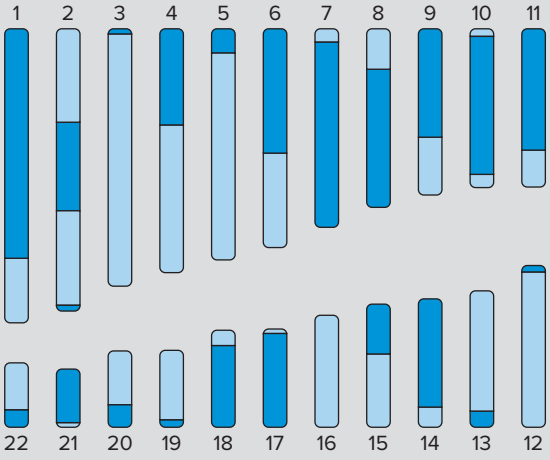
In order to map recombination sites, first the scientists developed new technology to isolate individual chromosomes from diploid somatic cells. Once isolated, the SNP alleles of individual homologs could be determined (**Fig. A**).

Figure A DNA sequences of homologs reveal SNP loci. At a particular SNP locus, maternal (*M*) and paternal (*P*) homologs can have different alleles (for example, an A–T base pair or a G–C base pair).

Next, the researchers determined which SNP alleles were present in individual sperm genomes. Then, by comparing the SNP alleles present on each chromosome in an individual sperm with the corresponding SNP alleles on each homolog of the man's somatic cells (prior to crossing over during meiosis), the locations of crossovers were revealed (**Fig. B**). By analyzing the crossovers in 91 different sperm, the researchers found that about 1 crossover per chromosome took place in each gamete, and crossover hotspots were detected.

The information obtained from this study and others like it is useful to scientists studying the biochemistry of recombination. In addition, you will see in later chapters of the book that the ability to determine the base pair sequence of individual chromosomes and individual gamete genomes has widespread applications in the study of mutation and human evolution.

Figure B Crossover map of a single sperm's autosomes. The autosomes (chromosomes *1–22*) of a sperm are depicted, where the *dark blue* and *light blue* regions correspond to the different homologs in the man's somatic cells (see Fig. A). Most chromosomes are the products of a single crossover; in this example, chromosomes 2 and 10 are exceptions.



about 1 million base pairs. In yeast, however, where the frequency of recombination per length of DNA is much higher than in humans, one map unit is approximately 2500 base pairs. Thus, although map units are useful for estimating relative distances between the genes of an organism, 1% RF can reflect very different expanses of DNA in different organisms.

Recombination frequencies sometimes vary even between the two sexes of a single species. In humans, the frequency of crossovers is about twofold higher in the female germ line than in males. This fact means that the same two genes will appear roughly twice as far apart on a genetic map generated by measuring RF in female meiosis than they would if crossing-over during male meiosis were measured instead. The Fast Forward Box *Mapping the Crossovers that Generate the Chromosomes of Individual Human Sperm* explains how new technology allows analysis of the DNA sequences of individual human sperm genomes. Researchers now can detect crossovers directly in each chromosome of single sperm. The results of these

analyses reveal that most chromosomes in human sperm have undergone only a single crossover.

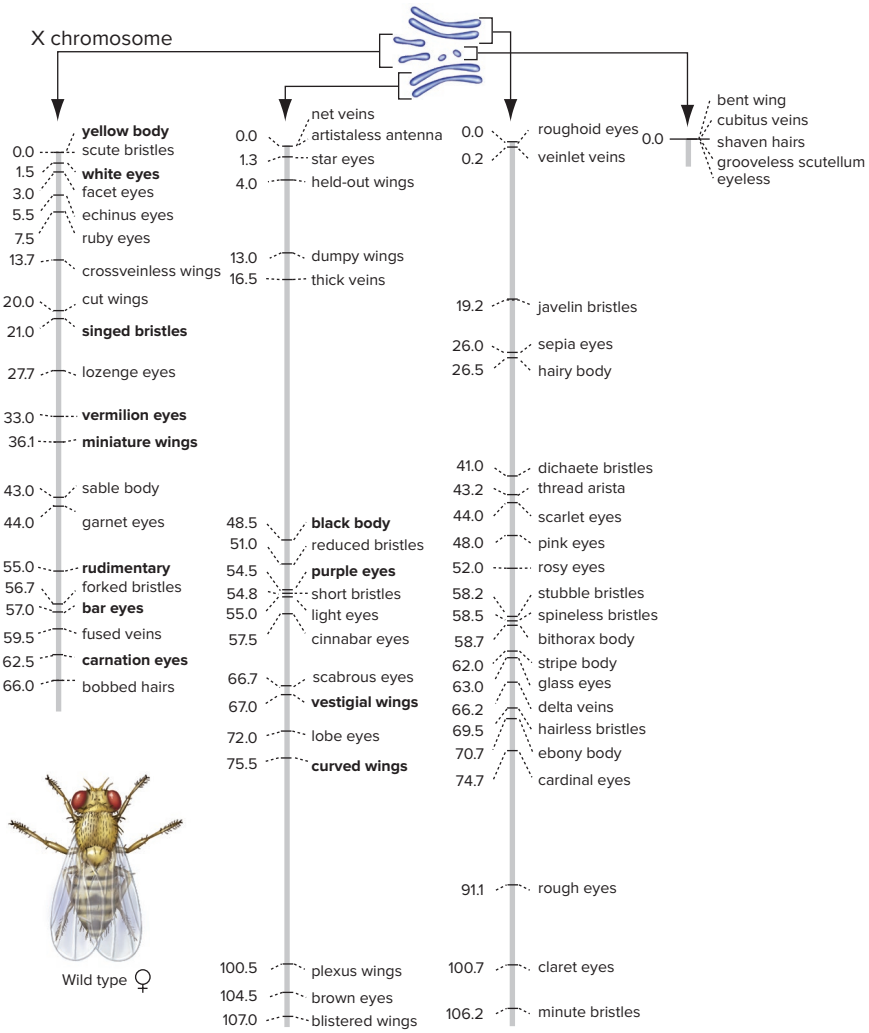
Drosophila provides an extreme example: No recombination occurs during meiosis in males. If you review the examples already discussed in this chapter, you will discover that they all measure recombination among the progeny of doubly heterozygous *Drosophila* females. Problem 19 at the end of this chapter shows how geneticists can exploit the absence of recombination in *Drosophila* males to establish rapidly that genes far apart on the same chromosome are indeed syntenic.

Multiple-Factor Crosses Help Establish Linkage Groups

Genes chained together by linkage relationships are known collectively as a **linkage group.** When enough genes have been assigned to a particular chromosome, the terms *chromosome* and *linkage group* become synonymous. If you can demonstrate that gene *A* is linked to gene *B, B* to *C, C* to *D,* and *D* to *E,* you can conclude that all these genes are syntenic. When the genetic map of a genome becomes so dense that it is possible to show that any gene on a chromosome is linked to another gene on the same chromosome, the number of linkage groups equals the number of pairs of homologous chromosomes in the species. Humans have 23 linkage groups, mice have 20, and fruit flies have 4 (**Fig. 5.18**).

The total genetic distance along a chromosome, which is obtained by adding many short distances between genes, may be much more than 50 m.u. For example, the two long *Drosophila* autosomes are both slightly more than 100 m.u. in length (Fig. 5.18), while the longest human chromosome is approximately 270 m.u. Recall, however, that even with the longest chromosomes, *pairwise* crosses between genes

Figure 5.18 ***Drosophila melanogaster*** **has four linkage groups.** A genetic map of the fruit fly, showing the position of many genes affecting body morphology, including those used as examples in this chapter (*highlighted in bold*). Because so many *Drosophila* genes have been mapped, each of the four chromosomes can be represented as a single linkage group.



FAST FORWARD

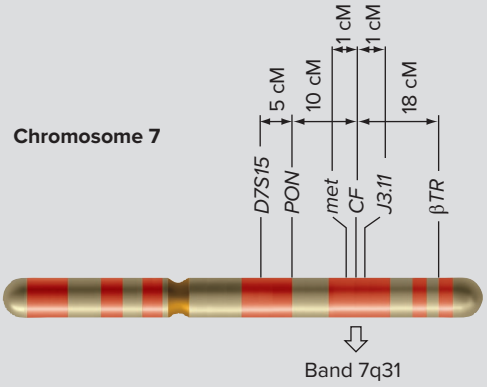
GENE MAPPING MAY LEAD TO A CURE FOR CYSTIC FIBROSIS

For 40 years after the symptoms of cystic fibrosis were first described in 1938, no molecular clue—no visible chromosomal abnormality transmitted with the disease, no identifiable protein defect carried by affected individuals—suggested the genetic cause of the disorder. As a result, no effective treatment existed for the 1 in 2000 Caucasian Americans born with the disease, most of whom died before they were 30. In the 1980s, however, geneticists were able to combine recently invented techniques for looking directly at DNA with maps constructed by linkage analysis to pinpoint a precise chromosomal position, or *locus,* for the cystic fibrosis gene.

The mappers of the cystic fibrosis gene faced an overwhelming task. They were searching for a gene that encoded an unknown protein, a gene that had not yet even been assigned to a chromosome. It could lie anywhere among the 23 pairs of chromosomes in a human cell.

- A review of many family pedigrees confirmed that cystic fibrosis is most likely determined by a single gene (*CF*). Investigators collected white blood cells from 47 families with two or more affected children, obtaining genetic data from 106 patients, 94 parents, and 44 unaffected siblings.
- The researchers next tried to discover if any other trait is reliably transmitted with cystic fibrosis. Analyses of the easily obtainable serum enzyme paroxonase showed that its gene (*PON*) is indeed linked to *CF.* At first, this knowledge was not that helpful, because *PON* had not yet been assigned to a chromosome.
- Then, in the early 1980s, geneticists developed a large series of DNA markers, based on new techniques that enabled them to recognize variations in the genetic material. A **DNA marker** is a segment of DNA representing a specific locus that comes in identifiable variations. These allelic variations segregate according to Mendel's laws, which means it is possible to follow their transmission as you would any gene's. Chapter 11 explains the discovery and use of DNA markers in greater detail; for now, it is only important to know that they exist and can be identified.
- By 1986, linkage analyses of hundreds of DNA markers had shown that one marker called *D7S15,* and known to reside on the long arm of chromosome 7, is linked with both *PON* and *CF*. Researchers computed recombination frequencies and found that the distance from the DNA marker to *CF* was 15 cM; from the DNA marker to *PON,* 5 cM; and from *PON* to *CF,* 10 cM. They concluded that the order of the three loci was *D7S15-PON-CF* (**Fig. A**). Because *CF* could lie 15 cM in either of two directions from the DNA marker, the area under investigation was approximately 30 cM. And because the human genome consists of roughly 3000 cM, this step of linkage analysis narrowed the search to 1% of the human genome, in a small region of chromosome 7.
- Finally, investigators discovered linkage with several other markers on the long arm of chromosome 7, called *J3.11, βTR,* and *met*. Two of the markers turned out to be separated from *CF* by a distance of only 1 cM. It now became possible to place *CF* in band 31 of chromosome 7's long arm (band 7q31, Fig. A). By 1989, researchers had used this mapping information to identify and clone the *CF* gene on the basis of its location.
- In 1992, investigators showed that the *CF* gene specifies a cell membrane protein that regulates the flow of chloride ions into and out of cells (review Fig. 2.25). This knowledge has become the basis of new drug therapies to open up ion flow, as well as gene therapies to introduce normal copies of the *CF* gene into the cells of CF patients. Although only in the early stages of development, such gene therapy holds out hope of an eventual cure for cystic fibrosis.

Figure A How molecular markers helped locate the gene for cystic fibrosis (*CF*).



located at the two ends will not produce more than 50% recombinant progeny.

Linkage mapping has practical applications of great importance. For example, the Fast Forward Box *Gene Mapping May Lead to a Cure for Cystic Fibrosis* describes how researchers used linkage information to locate the gene for this important human hereditary disease.

essential concepts

- A series of *two-point crosses* can establish the order of linked genes and the distances between them through pairwise analysis of recombination frequencies.
- *Three-point testcrosses* can refine map distances and reveal the existence of crossover *interference,* a phenomenon that distributes among all chromosomes the limited number of crossovers that occur in each meiosis.
- Although *genetic maps* provide an accurate picture of gene order on a chromosome, the distances measured between genes can be misleading.
- Genes in a linkage group are by definition *syntenic.* With enough mapped genes, the entire chromosome becomes a single *linkage group.*

5.4 The Chi-Square Test and Linkage Analysis

learning objectives

1. Explain the purpose of the chi-square test.
2. Discuss the concept of the null hypothesis and its use in data analysis.
3. Evaluate the significance of experimental data based on the chi-square test.

How do you know from a particular experiment whether two genes assort independently or are genetically linked? At first glance, this question should pose no problem. Discriminating between the two possibilities involves straightforward calculations based on assumptions well supported by observations. For independently assorting genes, a dihybrid F_1 female produces four types of gametes in equal numbers, so one-half of the F_2 progeny are of the parental classes and the other half are of the recombinant classes. In contrast, for linked genes, the two types of parental classes by definition always outnumber the two types of recombinant classes in the F_2 generation.

The problem is that because real-world genetic transmission is based on chance events, in a particular study even unlinked, independently assorting genes can produce deviations from the 1:1:1:1 ratio, just as in 10 tosses of a coin, you may easily get 6 heads and 4 tails (rather than the predicted 5 and 5). Thus, if a breeding experiment analyzing the transmission of two genes shows a deviation from the equal ratios of parentals and recombinants expected of independent assortment, can we necessarily conclude the two genes are linked? Is it instead possible that the results represent a statistically acceptable chance fluctuation from the mean values expected of unlinked genes that assort independently? Such questions become more important in cases where linkage is not all that tight, so that even though the genes are linked, the percentage of recombinant classes approaches 50%.

The Chi-Square Test Evaluates the Significance of Differences Between Predicted and Observed Values

To answer these kinds of questions, statisticians have devised several different ways to quantify the likelihood that an experimentally observed deviation from the predictions of a particular hypothesis could have occurred solely by chance. One of these probabilistic methods is known as the **chi-square test** for *goodness of fit*. This test measures how well observed results conform to predicted ones, and it is designed to account for the fact that the size of an experimental population (the *sample size*) is an important component of statistical significance. To appreciate the role of sample size, let's return to the proverbial coin toss before examining the details of the chi-square test.

In 10 tosses of a coin, an outcome of 6 heads (60%) and 4 tails (40%) is not unexpected because of the effects of chance. However, with 1000 tosses of the coin, a result of 600 heads (60%) and 400 tails (40%) would intuitively be highly unlikely. In the first case, a change in the results of one coin toss would alter the expected 5:5 ratio to the observed 6:4 ratio. In the second case, 100 tosses would have to change from tails to heads to generate the stated deviation from the predicted 500:500 ratio. Chance events could reasonably, and even likely, cause one deviation from the predicted number, but not 100.

Two important concepts emerge from this simple example. First, a comparison of percentages or ratios alone will never allow you to determine whether or not *observed* data are significantly different from *predicted* values. Second, the absolute numbers obtained are important because they reflect the size of the experiment. The larger the sample size, the closer the observed percentages can be expected to match the values predicted by the experimental hypothesis, *if the hypothesis is correct.* The chi-square test is therefore always calculated with numbers—actual data—and not percentages or proportions.

The chi-square test cannot prove a hypothesis, but it can allow researchers to reject a hypothesis. For this reason, a crucial prerequisite of the chi-square test is the framing of a **null hypothesis:** a model that might possibly be refuted by the test and that leads to clear-cut numerical predictions. Although contemporary geneticists use the chi-square test to interpret many kinds of genetic experiments, they use it most often to discover whether data obtained from breeding experiments provide evidence for or against the hypothesis

that two genes are linked. But one problem with the general hypothesis that *genes A and B are linked* is that no precise prediction exists for what to expect in terms of breeding data. The reason is that the frequency of recombinants, as we have seen, varies with each linked gene pair.

In contrast, the alternative hypothesis that *genes A and B are not linked* gives rise to a precise prediction: that alleles of different genes will assort independently and produce a 1:1:1:1 ratio of progeny types in a testcross. So, whenever a geneticist wants to determine whether two genes are linked, he or she actually tests whether the observed data are consistent with the null hypothesis of no linkage. If the chi-square test shows that the observed data differ significantly from those expected with independent assortment—that is, they differ enough not to be reasonably attributable to chance alone—then the researcher can reject the null hypothesis of no linkage and accept the alternative of linkage between the two genes.

The Tools of Genetics Box entitled *The Chi-Square Test for Goodness of Fit* presents the general protocol for this analysis. The final result of the calculations is the determination of the numerical probability—the ***p* value**—that a particular set of observed experimental results represents a chance deviation from the values predicted by a particular hypothesis. If the probability is high, it is likely that the hypothesis being tested explains the data, and the observed deviation from expected results is considered *insignificant.* If the probability is low, the observed deviation from expected results becomes *significant.* When this happens, it is unlikely that the hypothesis under consideration explains the data, and the hypothesis can be rejected.

It's important that you understand why the null hypothesis of no linkage (RF = 50%) is used, as opposed to a null hypothesis that assumes a particular degree of linkage (a particular RF < 50%). As stated earlier, a chi-square test can allow you to reject a null hypothesis, but not to prove it. This fact explains why geneticists test the null hypothesis that RF = 50 rather than a null hypothesis that the RF equals some specific number below 50, say 38, even though both models provide specific numerical predictions. If the deviations of the experimental values are *insignificant* (a high *p* value) relative to the hypothesis being tested, then the results could be consistent with that model, but they also could potentially be consistent with the other, untested hypothesis (RF = 50%) as well. Insignificant results are therefore not helpful. Suppose now that the deviations of the experimental values from the predictions of RF = 38 are *significant* (low *p* value), so that you could reject that hypothesis. This information would be similarly useless because you would not learn anything about the relative positions of the two genes other than that the map distance is not 38 m.u. Only one outcome is of real value: If you can reject the null hypothesis that the two genes are not linked (RF = 50%), then you have learned that they must be syntenic and are close enough together to be genetically linked.

Figure 5.19 Applying the chi-square test to determine if two genes are linked. The null hypothesis is that the two genes are unlinked. For Experiment 1, $p > 0.05$, so it is not possible to reject the null hypothesis. For Experiment 2, with a data set twice the size, $p < 0.05$, so most geneticists would reject the null hypothesis and conclude with greater than 95% confidence that the genes are linked.

Progeny Classes	Experiment 1			Experiment 2		
	O	E	$(O-E)^2/E$	O	E	$(O-E)^2/E$
Parentals *A B*	**18**	12	36/12	**36**	24	144/24
Parentals *a b*	**14**	12	4/12	**28**	24	16/24
Recombinants *A b*	**7**	12	25/12	**14**	24	100/24
Recombinants *a B*	**9**	12	9/12	**18**	24	36/24
Total	**48**	48	74/12	**96**	96	269/24
			⬇			⬇
			$\chi^2 = 6.17$			$\chi^2 = 12.3$
df = 3			$p > 0.10$			$p < 0.01$

Applying the Chi-Square Test to Linkage Analysis: An Example

Figure 5.19 depicts how the chi-square test could be applied to two sets of data obtained from testcross experiments asking whether genes *A* and *B* are linked. The columns labeled *O* for *observed* contain the actual data—the number of each of the four progeny types—from each experiment. In the first experiment, the total number of offspring is 48, so the *expected value* (*E*) for each progeny class, given the null hypothesis of no linkage, is simply 48/4 = 12. Now, for each progeny class, you square the deviation of the observed from the expected value, and divide the result by the expected value. Those calculations are presented in the column $(O-E)^2/E$. All four quotients are summed to obtain the value of chi square (χ^2). In experiment 1, $\chi^2 = 6.17$.

You next determine the **degrees of freedom (df)** for this experiment. Degrees of freedom is a mathematical concept that takes into consideration the number of independently varying parameters. In this example, the offspring fall into four classes. For three of the classes, the number of offspring can have any value, as long as their sum is no more than 48. However, once three of these values are fixed, the fourth value is also fixed, as the total in all four classes must equal 48. Therefore, the df with four classes is one less than the number of classes, or three. Next, you scan the chi-square table (see **Table 5.2**) for $\chi^2 = 6.17$ and df = 3. You find that the corresponding *p* value is greater than 0.10. From any *p* value greater than 0.05, you can conclude that it is not possible to reject the null hypothesis on the basis of this experiment, which means that this data set is not sufficient to demonstrate linkage between *A* and *B*.

If you use the same strategy to calculate a *p* value for the data observed in the second experiment, where there are a

TOOLS OF GENETICS

Blue DNA: © MedicalRF.com

The Chi-Square Test for Goodness of Fit

The general protocol for using the chi-square test for goodness of fit and evaluating its results can be stated in a series of steps. Two preparatory steps precede the actual chi-square calculation.

1. Use the data obtained from a breeding experiment to answer the following questions:
   a. What is the total number of offspring (events) analyzed?
   b. How many different classes of offspring (events) are present?
   c. In each class, what is the number of offspring (events) observed?
2. Calculate how many offspring (events) would be expected for each class if the null hypothesis (here, no linkage) were correct. To do so, multiply the fraction predicted by the null hypothesis (here, 1/4 of each possible progeny type) by the total number of offspring. You are now ready for the chi-square calculation.
3. To calculate chi square, begin with one class of offspring. Subtract the expected number from the observed number to obtain the deviation from the predicted value for the class. Square the result and divide this value by the expected number.
   Do this procedure for all classes and then sum the individual results. The final result is the chi-square (χ^2) value. This step is summarized by the equation:

$$\chi^2 = \Sigma \frac{(\textit{Number observed} - \textit{Number expected})^2}{\textit{Number expected}}$$

   where Σ means *sum of all classes.*
4. Next, you consider the degrees of freedom (df). The df is a measure of the number of independently varying parameters in the experiment (see text). The value of degrees of freedom is one less than the number of classes. Thus, if *N* is the number of classes, then the degrees of freedom (df) = *N* − *1*. If there are four classes, then there are 3 df.
5. Use the chi-square value together with the df to determine a *p* value: the probability that a deviation from the predicted numbers at least as large as that observed in the experiment would occur by chance. Although the *p* value is arrived at through a numerical analysis, geneticists routinely determine the value by a quick search through a table of critical χ^2 values for different degrees of freedom, such as Table 5.2.
6. Evaluate the significance of the *p* value. You can think of the *p* value as the probability that the null hypothesis is true. A value greater than 0.05 indicates that in more than 1 in 20 (or more than 5%) repetitions of an experiment of the same size, the observed deviation from predicted values could have been obtained by chance, even if the null hypothesis is actually true; the data are therefore *not significant* for rejecting the null hypothesis. Statisticians have arbitrarily selected the 0.05 *p* value as the boundary between rejecting or not rejecting the null hypothesis. A *p* value of less than 0.05 means that you can consider the deviation to be *significant,* and you can reject the null hypothesis.

TABLE 5.2 Critical Chi-Square Values

	p Values						
	Cannot Reject the Null Hypothesis				Null Hypothesis Rejected		
Degrees of Freedom	0.99	0.90	0.50	0.10	0.05	0.01	0.001
	χ^2 Values						
1	—	0.02	0.45	2.71	3.84	6.64	10.83
2	0.02	0.21	1.39	4.61	5.99	9.21	13.82
3	0.11	0.58	2.37	6.25	7.81	11.35	16.27
4	0.30	1.06	3.36	7.78	9.49	13.28	18.47
5	0.55	1.61	4.35	9.24	11.07	15.09	20.52

Note: χ^2 values that lie in the *yellow* region of this table allow you to reject the null hypothesis with > 95% confidence, and for recombination experiments, to postulate linkage.

total of 96 offspring, you find a *p* value less than 0.01 (Fig. 5.19). In this case, you can consider the difference between the observed and expected values to be significant. As a result, you can reject the null hypothesis of independent assortment and conclude it is likely that genes *A* and *B* are linked.

Statisticians have arbitrarily selected a *p* value of 0.05 as the boundary between significance and nonsignificance. Values lower than this indicate there would be fewer than 5 chances in 100 of obtaining the same results by random sampling if the null hypothesis were true. A *p* value of less

than 0.05 thus suggests that the data show major deviations from predicted values significant enough to reject the null hypothesis with greater than 95% confidence. More conservative scientists often set the boundary of significance at $p = 0.01$, and they would therefore reject the null hypothesis only if their confidence was greater than 99%.

In contrast, p values greater than 0.01 or 0.05 do not necessarily mean that two genes are unlinked; it may mean only that the sample size is not large enough to provide an answer. With more data, the p value normally rises if the null hypothesis of no linkage is correct and falls if there is, in fact, linkage.

Note that in Fig. 5.19 all of the numbers in the second set of data are simply double the numbers in the first set, with the proportions remaining the same. Thus, by doubling the sample size from 48 to 96 individuals, it was possible to go from no significant difference to a significant difference between the observed and the expected values. In other words, the larger the sample size, the less the likelihood that a certain percentage deviation from expected results happened simply by chance. Bearing this in mind, you can see that it is not appropriate to use the chi-square test when analyzing small samples of less than 10. This issue creates a problem for human geneticists because human families produce only a small number of children. To achieve a reasonable sample size for linkage studies in humans, scientists often pool data from a large number of family pedigrees and use a different statistical analysis called the *Lod score* (see Chapter 11).

We emphasize again that the chi-square test *does not* prove linkage or its absence. What it *does* is provide a quantitative measure of the likelihood that the data from an experiment can be explained by a particular hypothesis. The chi-square analysis is thus a general statistical test for significance; it can be used with many different experimental designs and with hypotheses other than the absence of linkage. As long as it is possible to propose a null hypothesis that leads to a predicted set of values for a defined set of data classes, you can readily determine whether or not the observed data are consistent with the hypothesis.

When experiments lead to rejection of a null hypothesis, you may need to confirm an alternative. For instance, if you are testing whether two opposing traits result from the segregation of two alleles of a single gene, you would expect a testcross between an F_1 heterozygote and a recessive homozygote to produce a 1:1 ratio of the two traits in the offspring. If instead, you observe a ratio of 6:4 and the chi-square test produces a p value of 0.009, you can reject the null hypothesis (a 1:1 ratio). But you are still left with the question of what the absence of a 1:1 ratio means. Two alternatives exist: (1) Individuals with the two possible genotypes are not equally viable, *or* (2) more than one gene affects the trait. The chi-square test cannot tell you which possibility is correct, and you would have to study the matter further. The problems at the end of this chapter illustrate several applications of the chi-square test pertinent to genetics.

essential concepts

- A *null hypothesis* is a model that leads to a discrete numerical prediction.
- The *chi-square test for goodness of fit* helps determine whether two genes are linked by comparing differences between the numbers of progeny of different classes observed in an experiment, and the numbers of progeny of these classes expected from the null hypothesis that genes are unlinked and thus assort independently.
- The *probability value (p)* measures the likelihood that deviations from the predicted values have occurred by chance alone; the null hypothesis is rejected when $p < 0.05$.

5.5 Tetrad Analysis in Fungi

learning objectives

1. Explain the meaning of the term *tetrad* as applied to the asci produced by certain fungi.
2. Differentiate between parental ditype (PD), nonparental ditype (NPD), and tetratype (T).
3. Describe how the relative numbers of PDs and NPDs can be used to establish linkage.
4. Explain how ordered and unordered tetrad analysis can map the positions of genes and (for ordered tetrads) centromeres.

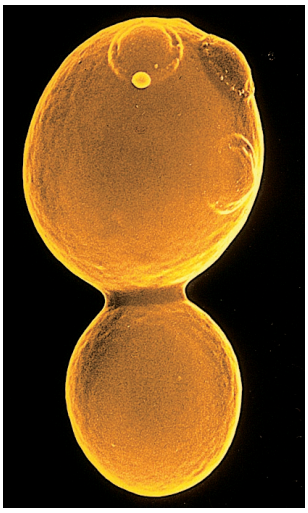
With *Drosophila,* mice, peas, people, and other diploid organisms, each individual represents only one of the four potential gametes generated by each parent in a single meiotic event. Thus, until now, our presentation of linkage, recombination, and mapping has depended on inferences derived from examining the phenotypes of diploid progeny resulting from random unions of meiotic products *en masse* (that is, among a large group). For such diploid organisms, we do not know which, if any, of the parents' other progeny arose from gametes created in the same meiosis. Because of this limitation, the analysis of random products of meiosis in diploid organisms must be based on statistical samplings of large populations.

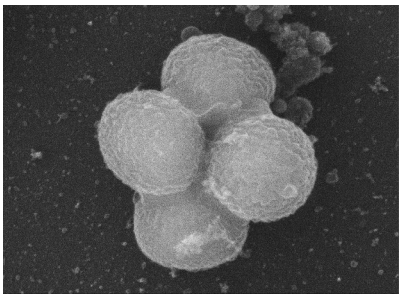
In contrast, various species of fungi provide a unique opportunity for genetic analysis because they house all four haploid products of each meiosis in a sac called an **ascus** (plural, *asci*). These haploid cells, or **ascospores** (also known as *haplospores* or simply *spores*), can germinate and survive as viable haploid individuals that perpetuate themselves by mitosis. The phenotype of such haploid fungi is a direct representation of their genotype, without complications of dominance.

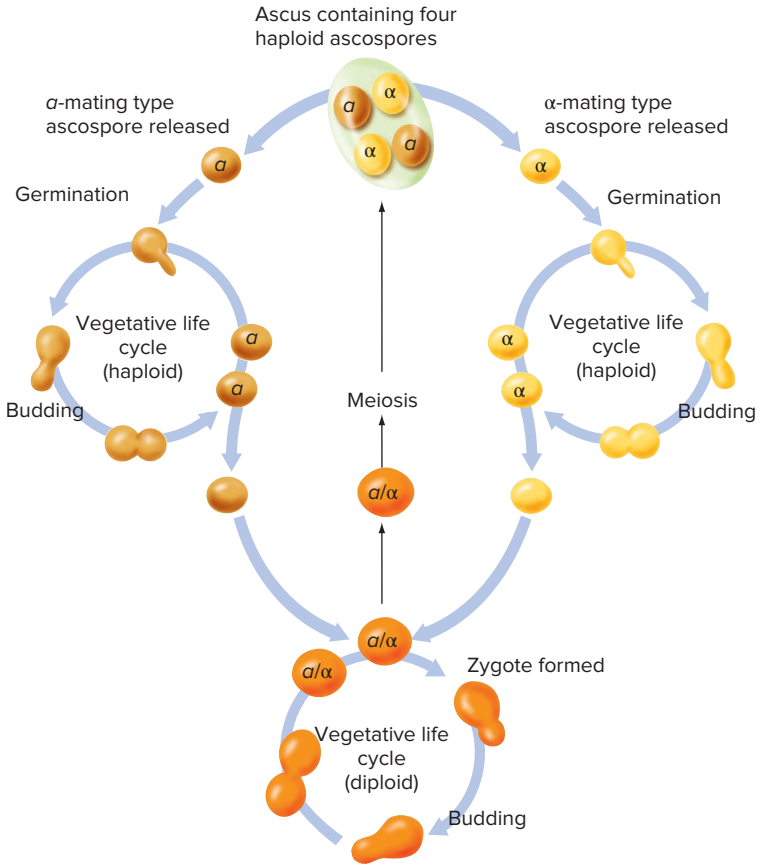
Figure 5.20 illustrates the life cycles of two fungal species that preserve their meiotic products in a sac. One,

Figure 5.20 The life cycles of the yeast *Saccharomyces cerevisiae* and the bread mold *Neurospora crassa*. Both *S. cerevisiae* and *N. crassa* have two mating types that can fuse to form diploid cells that undergo meiosis. **(a)** Yeast cells can grow vegetatively either as haploids or diploids. The products of meiosis in a diploid cell are four haploid ascospores that are arranged randomly in unordered yeast asci. **(b)** The diploid state in *Neurospora* exists only for a short period. Meiosis in *Neurospora* is followed by mitosis, to give eight haploid ascospores in the ascus. The ordered arrangement of spores in *Neurospora* asci reflects the geometry of the meiotic and mitotic spindles. The photographs showing a budding (mitotically dividing) yeast cell (*top*) and a yeast tetrad (*bottom*) in part (a) are at much higher magnification than the photograph displaying *Neurospora* asci in part (b).

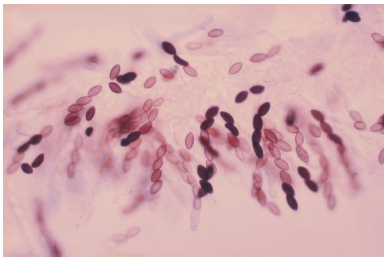
(a) *Saccharomyces cerevisiae*

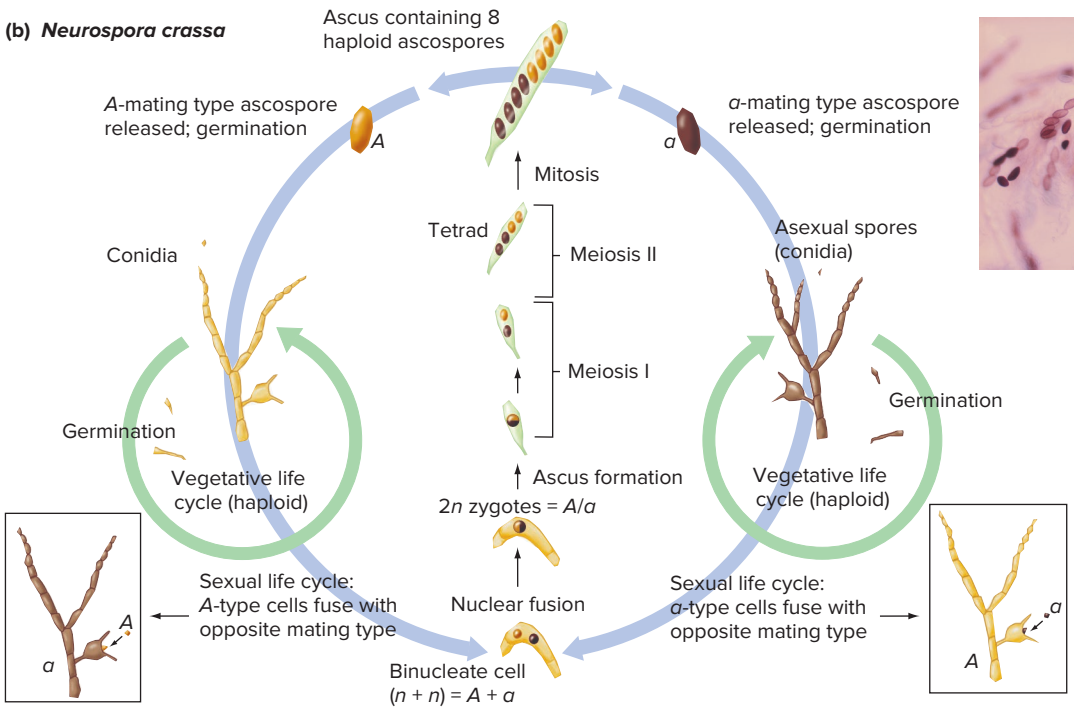






(b) *Neurospora crassa*





the normally unicellular baker's yeast (*Saccharomyces cerevisiae*), is sold in supermarkets and contributes to the texture, shape, and flavor of bread; it generates four ascospores with each meiosis. The other, *Neurospora crassa,* is a mold that renders the bread on which it grows inedible; it too generates four ascospores with each meiosis, but at the completion of meiosis, each of the four haploid ascospores immediately divides once by mitosis to yield four pairs, for a total of eight haploid cells. The two cells in each pair of *Neurospora* ascospores have the same genotype because they arose from mitosis.

Haploid cells of both yeast and *Neurospora* normally reproduce vegetatively (that is, asexually) by mitosis. However, sexual reproduction is possible because the haploid cells come in two mating types, and cells of opposite mating types can fuse to form a diploid zygote (Fig. 5.20). In yeast, these diploid cells are stable and can reproduce through successive mitotic cycles. Stress, such as that caused by a scarcity of essential nutrients, induces the diploid cells of yeast to enter meiosis. In bread mold, the diploid zygote instead immediately undergoes meiosis, so the diploid state is only transient.

Mutations in haploid yeast and mold affect many different traits, including the appearance of the cells and their ability to grow under particular conditions. For instance, yeast cells with the *his4* mutation are unable to grow in the absence of the amino acid histidine, while yeast with the *trp1* mutation cannot grow without an external source of the amino acid tryptophan. Geneticists who specialize in the study of yeast have devised a system of representing genes that is slightly different from the ones for *Drosophila* and mice. They use capital letters (*HIS4*) to designate dominant alleles and lowercase letters (*his4*) to represent recessives. For most of the yeast genes we will discuss, the wild-type alleles are dominant and may also be represented by the alternative shorthand +, while the symbol for the recessive alleles remains the lowercase abbreviation (*his4*). (See *Guidelines for Gene Nomenclature*.) Remember, however, that dominance or recessiveness is relevant only for diploid yeast cells, not for haploid cells that carry only one allele.

An Ascus Contains All Four Products of a Single Meiosis

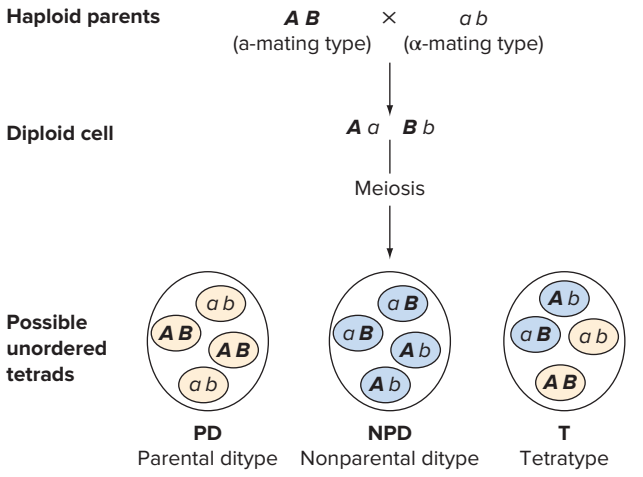
The assemblage of four ascospores (or four pairs of ascospores) in a single ascus is called a **tetrad.** Note that this is a second meaning for the term *tetrad.* In Chapter 4, a tetrad was the four homologous chromatids—two in each chromosome of a bivalent—synapsed during the prophase and metaphase of meiosis I. Here, it is the four products of a single meiosis held together in a sac. Because the four chromatids of a bivalent give rise to the four products of meiosis, the two meanings of tetrad refer to almost the same things. Yeast make **unordered tetrads;** that is, the four meiotic products (the spores) are arranged at random within the ascus. *Neurospora crassa* produce **ordered tetrads,** with the four pairs, or eight haplospores, arranged in a line.

To analyze both unordered and ordered tetrads, researchers can release the spores of each ascus, induce the haploid cells to germinate under appropriate conditions, and then analyze the genetic makeup of the resulting haploid cultures. The data they collect in this way enable them to identify the four products of a single meiosis and compare them with the four products of many other distinct meioses. Ordered tetrads offer another possibility. With the aid of a dissecting microscope, investigators can recover the ascospores in the order in which they occur within the ascus and thereby obtain additional information that is useful for mapping. We look first at the analysis of randomly arranged spores in yeast tetrads as an example. We then describe the additional information that can be gleaned from the microanalysis of ordered tetrads, using *Neurospora* as our model organism.

Tetrads Can Be Characterized as Parental Ditypes (PDs), Nonparental Ditypes (NPDs), or Tetratypes (Ts)

When diploid yeast cells heterozygous for each of two genes are induced to undergo meiosis, three tetrad types can be produced whether the two genes are on the same chromosome or different chromosomes. Consider the cross in **Fig. 5.21,** in which haploid cells of opposite mating types (*a* versus *α*) and with alternate alleles of two genes

Figure 5.21 Three tetrad types produced by meiosis of dihybrid yeast. All three types may be produced whether or not genes *A* and *B* are on the same chromosome and whether or not they are linked. Parental spores are *orange* and recombinant spores are *blue*.



mate to form an *Aa Bb* diploid. One possibility is that all four spores in the resulting tetrad will have the parental configuration of alleles; such a tetrad is called a **parental ditype (PD).** The second kind of tetrad, called a **nonparental ditype (NPD),** contains four recombinant spores, two of each type. The final possibility is a **tetratype (T)** tetrad, which contains four different kinds of spores—two recombinants (one of each type) and two parentals (one of each type). Note that the spores in each yeast ascus are not arranged in any particular order (Fig. 5.21). The classification of a tetrad as PD, NPD, or T is based solely on the number of parental and recombinant spores in the ascus.

Recombination Frequencies May Be Determined by Counting the Number of Each Tetrad Type

In order to determine the RF between the two genes in Fig. 5.21, you could simply break open all the spore cases, pool the spores, and analyze them to determine which ones are parentals and which ones are recombinants. In this case, RF equals the number of recombinant spores divided by the total number of spores (parentals plus recombinants) counted.

Alternatively, as you determine the genotype of each spore, you could keep track of which spores came from the same ascus and count instead the number of each type of tetrad—PD, NPD, or T. When the latter method is used, the recombination frequency is simply the number of NPD-type tetrads (all of the spores in them are recombinant; Fig. 5.21) plus half the number of T-type tetrads (half of the spores in them are recombinant; Fig. 5.21) divided by the total number of tetrads counted:

$$RF = [NPD + 1/2\ (T)]/\text{total tetrads}.$$

Using either method, you will calculate the same value for RF. However, analyzing the products of a fungal cross as tetrads has several advantages. Some of these are technical; for example, in some fungi tetrad analysis enables you to determine the distance between genes and centromeres. But more importantly, analysis of tetrads enables you to develop a deeper appreciation for the events of meiosis. The best way to understand tetrad analysis is to examine how the different tetrad types are generated when the two genes in a dihybrid are on different chromosomes and when they are on the same chromosome.

Unlinked genes on different chromosomes: PDs = NPDs

What kinds of tetrads arise when diploid yeast cells heterozygous for two genes on different chromosomes are induced to undergo meiosis? Consider a mating between a haploid strain of yeast of mating type *a,* carrying the *his4* mutation and the wild-type allele of the *TRP1* gene, and a strain of the opposite mating type α that has the genotype *HIS4 trp1*. The resulting *a/*α diploid cells are *his4/HIS4; trp1/ TRP1,* as shown in **Fig. 5.22a.** [The semicolon (;) in the genotype separates genes on nonhomologous chromosomes.]

Three different types of meiosis can take place, each of which produces a different tetrad type. A PD tetrad will result from one of the two random alignments of homologous chromosomes during meiosis I (**Fig. 5.22b**). The equally likely alternative chromosome alignment yields an NPD tetrad (**Fig. 5.22c**). T tetrads are produced from a crossover between only one of the genes and the centromere (**Fig. 5.22d**).

These results reveal two important facts about the tetrads. First, because the meiosis events shown in panels (b) and (c) of Fig. 5.22 are equally likely, the number of PDs will equal the number of NPDs when the two genes are on different chromosomes. The production of Ts (Fig. 5.22d) does not affect the equality of PDs and NPDs because Ts are produced equally often with the two alternative chromosome alignments. Therefore, T meioses deplete PD and NPD production to the same extent.

Second, as expected for alleles of genes on different chromosomes, RF = 50%. The number of PDs, in which all of the spores are parentals, and the number of NPDs, in which all of the spores are recombinants, are equal. The only other tetrad type, T, contains half recombinants and half parentals.

Figure 5.22e displays the data from one experiment with *his4/HIS4; trp1/TRP1* diploids. (Bear in mind that the column headings of PD, NPD, and T refer to tetrads and not to individual haploid cells.) From these data, you can see that the number of PDs and NPDs are nearly the same. Chi-square analysis would indicate that the results do not differ significantly from the hypothesis that PD = NPD in this experiment.

Linked genes: PDs >> NPDs

You have just seen that if two genes in doubly heterozygous diploid yeast cells are on different chromosomes (unlinked), the number of PD tetrads will approximately equal the number of NPD tetrads. This will not be the outcome for linked genes. When yeast that are dihybrid for linked genes sporulate (that is, undergo meiosis), the number of PDs produced far exceeds the number of NPDs. By analyzing an actual cross involving linked genes, we can see how this statement follows from the events occurring during meiosis.

A haploid yeast strain containing the *arg3* and *ura2* mutations was mated to a wild-type *ARG3 URA2* haploid strain. When the resultant diploid was induced to sporulate,

Figure 5.22 How meiosis can generate three kinds of tetrads when two genes are on different chromosomes.
(a) Parental cross. **(b)** and **(c)** In the absence of recombination, the two equally likely alternative arrangements of two pairs of chromosomes yield either PD or NPD tetrads. T tetrads are made only if either gene recombines with respect to its corresponding centromere, as in **(d).** Numerical data in **(e)** show that the number of PD tetrads ≈ the number of NPD tetrads when the two genes are unlinked. In (b)–(d), parental spores are *orange* and recombinant spores are *blue*.

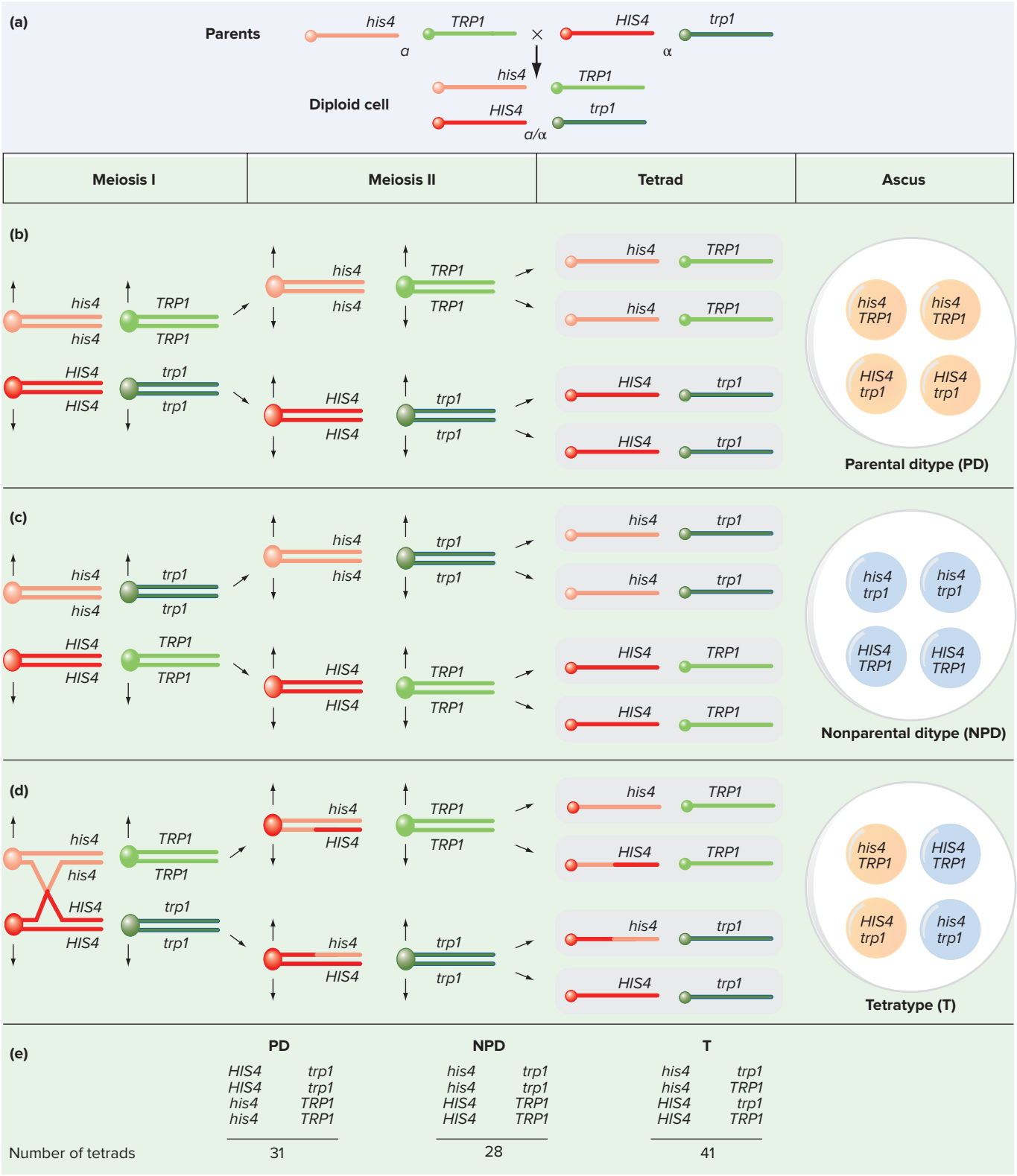


Figure 5.23 When genes are linked, PDs exceed NPDs.

P	*arg3 ura2* (*a*-mating type) ×	*ARG3 URA2* (α-mating type)	
Diploid cell	*arg3 ura2* / *ARG3 URA2*		
	Meiosis		
Products of meiosis	**PD**	**NPD**	**T**
	arg3 ura2 *arg3 ura2* *ARG3 URA2* *ARG3 URA2*	*arg3 URA2* *arg3 URA2* *ARG3 ura2* *ARG3 ura2*	*arg3 ura2* *arg3 URA2* *ARG3 ura2* *ARG3 URA2*
Number of tetrads	127	3	70

the 200 tetrads produced had the distribution shown in **Fig. 5.23.** As you can see, the 127 PD tetrads far outnumber the 3 NPD tetrads, suggesting that the two genes are linked.

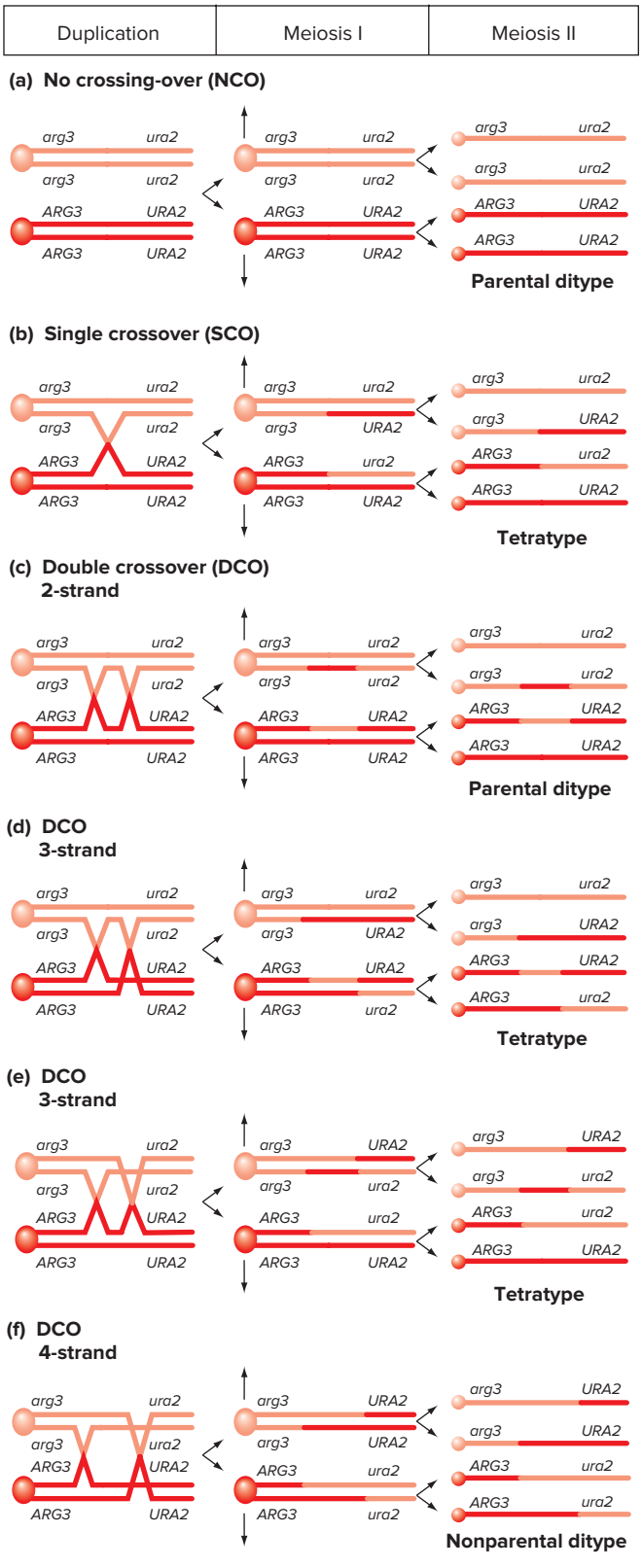
Figure 5.24 shows how we can explain the particular kinds of tetrads observed in terms of the various types of meioses possible. If no crossing-over occurs between the two genes, the resulting tetrad will be PD (Fig. 5.24a). A single crossover between *ARG3* and *URA2* will generate a T tetrad (Fig. 5.24b). But what about double crossovers? As you saw earlier (Fig. 5.11), there are actually four different possibilities, depending on which chromatids participate, and each of the four should occur with equal frequency. A double crossover involving only two chromatids generates a PD tetrad (Fig. 5.24c). Three-strand double crossovers can occur in the two ways depicted in Fig. 5.24d and e; either way, a T tetrad results. Finally, if all four chromatids take part in the two crossovers (one crossover involves two strands and the other crossover, the other two strands), the resulting tetrad is NPD (Fig. 5.24f). Therefore, if two genes are linked, the only way to generate an NPD tetrad is through a four-strand double exchange. When two genes are close together on a chromosome, meioses with one of the four kinds of double crossovers are much rarer than those with no crossing-over or single crossovers, which produce PD and T tetrads, respectively. This explains why, if two genes are linked, the number of PDs must greatly exceed the number of NPDs.

If we calculate the RF from the data in Fig. 5.23 using the equation RF = [NPD + (1/2)T]/total tetrads, we find that

$$RF = [3 + (1/2)70]/200 \times 100 = 19 \text{ m.u.}$$

However, observation of Fig. 5.24 reveals that this equation for RF is not an accurate reflection of the actual number of crossover events when two genes are far enough apart that DCOs occur and NPDs appear. For example, the equation does not count any PDs in the numerator, even though DCO meioses generate some PDs (Fig. 5.24c).

Figure 5.24 How crossovers between linked genes generate different tetrads. **(a)** PDs arise when there is no crossing-over. **(b)** Single crossovers between the two genes yield tetratypes (Ts). **(c)** Double crossovers between linked genes can generate PD, T, or NPD tetrads, depending on which chromatids participate in the crossovers.



Problem 44 at the end of the chapter helps you derive a corrected equation for RF. The corrected RF equation takes into account all of the DCO meioses that contribute to the tetrads resulting from a cross where PD >> NPD, but NPD is greater than zero.

Two genes far apart on a single chromosome: PDs = NPDs

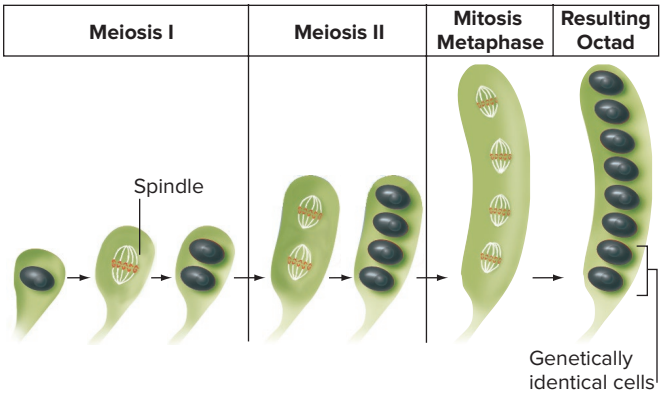
In tetrad analysis, just as in *en masse* linkage analysis, two genes may be so far apart on the same chromosome that they will be indistinguishable from two genes on different chromosomes: In both cases, PD = NPD. If two genes are sufficiently far apart on the chromosome, at least one crossover occurs between them during every meiosis. Under such circumstances, no meioses are NCOs, and therefore all PD tetrads as well as all NPD tetrads come from equally frequent kinds of DCOs (events c and f in Fig. 5.24). Thus, whether two genes are assorting independently because they are on different chromosomes or because they are far apart on the same chromosome, the end result is the same: PD = NPD and RF = 50%.

Ordered Tetrads Help Locate Genes in Relation to the Centromere

Analyses of ordered tetrads, such as those produced by the bread mold *Neurospora crassa,* allow you to map the centromere of a chromosome relative to other genetic markers, information that you cannot normally obtain from unordered yeast tetrads. As described earlier, immediately after specialized haploid *Neurospora* cells of different mating types (*A* and *a*) fuse at fertilization, the diploid zygote undergoes meiosis within the confines of a narrow ascus (review Fig. 5.20b). At the completion of meiosis, each of the four haploid meiotic products divides once by mitosis, yielding an **octad** of eight haploid ascospores. Dissection of the ascus at this point allows one to determine the phenotype of each of the eight haploid cells.

The cross-sectional diameter of the ascus is so small that cells cannot slip past each other. Moreover, during each division after fertilization, the microtubule fibers of the spindle extend outward from the centrosomes parallel to the long axis of the ascus (**Fig. 5.25**). These facts have two important repercussions. First, when each of the four products of meiosis divides once by mitosis, the two genetically identical cells that result lie adjacent to each other. Because of this feature, starting from either end of the ascus, you can count the octad of ascospores as four cell pairs and analyze it as a tetrad. Second, from the precise positioning of the four ascospore pairs within the ascus, you can infer the arrangement of the four chromatids of each homologous chromosome pair during the two meiotic divisions.

Figure 5.25 How ordered tetrads form. Spindles form parallel to the long axis of the growing *Neurospora* ascus, and the cells cannot slide around each other. The order of ascospores thus reflects meiotic spindle geometry. After meiosis, each haploid cell undergoes mitosis, producing an eight-cell ascus (an octad). The octad consists of four pairs of cells; the two cells of each pair are genetically identical.

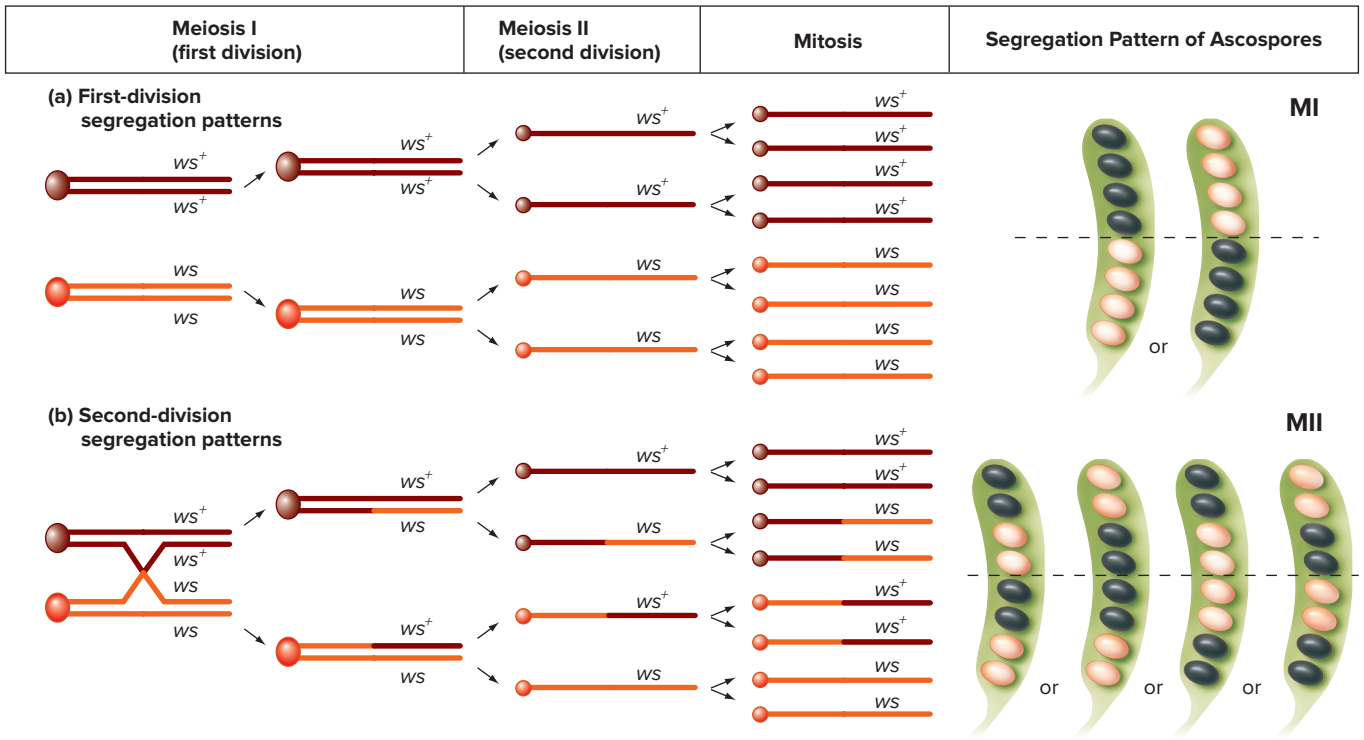


To understand the genetic consequences of the geometry of the ascospores, it is helpful to consider what kinds of tetrads you would expect from the segregation of two alleles of a single gene. (In the following discussion, you will see that *Neurospora* geneticists denote alleles with symbols similar to those used for *Drosophila,* as detailed in the *Guidelines for Gene Nomenclature*.) The mutant *white-spore* allele (*ws*) alters ascospore color from wild-type black to white. In the absence of recombination, the two alleles (ws^+ and *ws*) separate from each other at the first meiotic division because the centromeres separate at that stage. The second meiotic division and subsequent mitosis create asci in which the top four ascospores are of one genotype (for instance ws^+) and the bottom four of the other (*ws*). Whether the top four are ws^+ and the bottom four *ws,* or *vice versa,* depends on the random metaphase I orientation of the homologs that carry the gene relative to the long axis of the developing ascus.

The segregation of two alleles of a single gene at the first meiotic division is thus indicated by an ascus in which an imaginary line drawn between the fourth and the fifth ascospores of the octad cleanly separates haploid products bearing the two alleles. Such an ascus displays a **first-division (MI) segregation pattern** (**Fig. 5.26a**).

Suppose now that during meiosis I, a crossover occurs in a heterozygote between the *white-spore* gene and the centromere. As **Fig. 5.26b** illustrates, this can lead to four equally possible ascospore arrangements, each one depending on a particular orientation of the four chromatids during the two meiotic divisions. In all four cases, both ws^+ and *ws* spores are found on both sides of the imaginary line drawn between ascospores 4 and 5, because cells with only one kind of allele do not arise until the end of the second meiotic division. Octads carrying such configurations of

Figure 5.26 Two segregation patterns in ordered asci. **(a)** In the absence of a crossover between a gene and the centromere, the two alleles of a gene will separate at the first meiotic division. The result is an MI segregation pattern in which each allele appears in spores located on only one side of an imaginary line through the middle of the ascus. **(b)** A crossover between a gene and the centromere produces an MII segregation pattern in which both alleles appear on the same side of the middle line.



spores display a **second-division (MII) segregation pattern.**

Because MII patterns result from meioses in which a crossover has occurred between a gene and the centromere, the relative number of asci with this pattern can be used to determine the gene ↔ centromere distance. In an ascus showing MII segregation, one-half of the ascospores are derived from chromatids that have exchanged parts, while the remaining half arise from chromatids that have not participated in crossovers (Fig. 5.26b). To calculate the distance between a gene and the centromere, you therefore simply divide the percentage of MII octads by 2:

gene ↔ centromere distance = (1/2) MII/total tetrads × 100.

Tetrad Analysis: A Numerical Example

In one experiment, a *thr*$^+$*arg*$^+$ wild-type strain of *Neurospora* was crossed with a *thr arg* double mutant. The *thr* mutants cannot grow in the absence of the amino acid threonine, while *arg* mutants cannot grow without a source of the amino acid arginine; cells carrying the wild-type alleles of both genes can grow in medium that contains neither amino acid. From this cross, 105 octads, considered here as tetrads, were obtained. These tetrads were classified in seven different groups—A, B, C, D, E, F, and G—as shown in **Fig. 5.27a.** For each of the two genes, we can now find the distance between the gene and the centromere.

Figure 5.27 Genetic mapping by ordered-tetrad analysis: An example. **(a)** In ordered-tetrad analysis, tetrad classes are defined not only as PD, NPD, or T but also according to whether they show an MI or MII segregation pattern. Each entry in this table represents a pair of adjacent, identical spores in the actual *Neurospora* octad. Red dots indicate the middle of the asci. **(b)** Genetic map derived from the data in part (a). Ordered-tetrad analysis allows determination of the centromere's position as well as distances between genes.

(a) A *Neurospora* cross

Tetrad group	A	B	C	D	E	F	G
Segregation pattern	*thr arg* *thr arg* • • • • *thr*$^+$*arg*$^+$ *thr*$^+$*arg*$^+$	*thr arg* *thr*$^+$*arg* • • • • *thr*$^+$*arg*$^+$ *thr arg*$^+$	*thr arg* *thr arg*$^+$ • • • • *thr*$^+$*arg* *thr*$^+$*arg*$^+$	*thr arg*$^+$ *thr*$^+$*arg* • • • • *thr*$^+$*arg*$^+$ *thr arg*	*thr arg*$^+$ *thr*$^+$*arg* • • • • *thr*$^+$*arg* *thr arg*$^+$	*thr arg*$^+$ *thr arg*$^+$ • • • • *thr*$^+$*arg* *thr*$^+$*arg*	*thr arg* *thr*$^+$*arg*$^+$ • • • • *thr*$^+$*arg*$^+$ *thr arg*
Total in group	72	16	11	2	2	1	1

(b) Corresponding genetic map

To do this for the *thr* gene, we count the number of tetrads with an MII pattern for that gene. Drawing an imaginary line through the middle of the tetrads, we see that those in groups B, D, E, and G are the result of MII segregations for *thr,* while the remainder show MI patterns. The centromere ↔ *thr* distance is thus:

Half the percentage of MII patterns =

$$\frac{(1/2)(16 + 2 + 2 + 1)}{105} \times 100 = 10 \text{ m.u.}$$

Similarly, the MII tetrads for the *arg* gene are in groups C, D, E, and G, so the distance between *arg* and its centromere is:

$$\frac{(1/2)(11 + 2 + 2 + 1)}{105} \times 100 = 7.6 \text{ m.u.}$$

To ascertain whether the *thr* and *arg* genes are linked, we need to evaluate the seven tetrad groups in a different way, looking at the combinations of alleles for the two genes to see if the tetrads in that group are PD, NPD, or T. We can then ask whether PD >> NPD. Referring again to Fig. 5.27a, we find that groups A and G are PD, because all the ascospores show parental combinations, while groups E and F, with four recombinant spores, are NPD. PD is thus 72 + 1 = 73, while NPD is 1 + 2 = 3. From these data, we can conclude that the two genes are linked.

What is the map distance between *thr* and *arg?* For this calculation, we need to find the numbers of T and NPD tetrads. Tetratypes are found in groups B, C, and D, and we already know that groups E and F carry NPDs. Using the same formula for map distances as the one previously used for yeast,

$$\text{RF} = \frac{\text{NPD} + 1/2\text{T}}{\text{Total tetrads}} \times 100$$

we get:

$$\text{RF} = \frac{3 + (1/2)(16 + 11 + 2)}{105} \times 100 = 16.7 \text{ m.u.}$$

Because the distance between *thr* and *arg* is larger than that separating either gene from the centromere, the centromere must lie between *thr* and *arg,* yielding the map in **Fig. 5.27b.** The distance between the two genes calculated by the preceding formula (16.7 m.u.) is smaller than the sum of the two gene ↔ centromere distances (10.0 + 7.6 = 17.6 m.u.) because the formula does not account for all of the double crossovers. As always, calculating map positions by adding shorter intervals produces the most accurate genetic maps. The gene ↔ centromere distances are shorter and are therefore more accurate than the *thr/arg* distance calculation in this example.

Table 5.3 summarizes the procedures for mapping genes in fungi producing ordered and unordered tetrads.

TABLE 5.3 Rules for Tetrad Analysis

For Ordered and Unordered Tetrads
Considering genes two at a time, assign tetrads as PD, NPD, or T.
If PD >> NPD, the two genes are linked.
If PD ≈ NPD, the two genes assort independently (they are unlinked).
The map distance between two genes if they are linked $= \frac{\text{NPD} + (1/2)\text{T}}{\text{Total tetrads}} \times 100$
For Ordered Tetrads Only
The map distance between a gene and the centromere $= \frac{(1/2)\text{MII}}{\text{Total tetrads}} \times 100$

essential concepts

- A *tetrad* is the group of four haploid spores within an ascus that results from a single meiosis in fungi.
- In a *parental ditype (PD),* a tetrad has four parental spores; in a *nonparental ditype (NPD),* a tetrad contains four recombinant spores; in a *tetratype (T),* an ascus contains two different parental spores and two different recombinant spores.
- When a dihybrid sporulates, if PD tetrads are equal to NPD tetrads, the genes in question are unlinked; when PDs greatly outnumber NPDs, the genes are linked.
- Analysis of *unordered tetrads* reveals linked genes and the map distance between them; analysis of *ordered tetrads* further allows determination of the distance between a gene and the centromere.

5.6 Mitotic Recombination and Genetic Mosaics

learning objectives

1. Explain how mitotic recombination leads to the mosaic condition termed twin spots.
2. Describe sectored colonies in yeast and their significance in evaluating mitotic recombination.

The recombination of genetic material is a critical feature of meiosis. It is thus not surprising that eukaryotic organisms express a variety of enzymes (to be described in Chapter 6) that specifically initiate meiotic recombination.

Recombination can also occur during mitosis. Unlike what happens in meiosis, however, mitotic crossovers are initiated by mistakes in chromosome replication or by chance exposures to radiation that break DNA molecules, rather than by a well-defined cellular program. As a result, mitotic recombination is a rare event, occurring no more often than once in a million somatic cell divisions. Nonetheless, the growth of a colony of yeast cells or the development of a complex multicellular organism involves so many cell divisions that geneticists can detect these rare mitotic events routinely.

Twin Spots Indicate Mosaicism Caused by Mitotic Recombination

In 1936, the *Drosophila* geneticist Curt Stern inferred the existence of mitotic recombination from observations of *twin spots* in fruit flies. **Twin spots** are adjacent islands of tissue that differ both from each other and from the tissue surrounding them. The distinctive patches arise from homozygous cells with a recessive phenotype growing amid a generally heterozygous cell population displaying the dominant phenotype. In *Drosophila,* the *yellow* (*y*) mutation changes body color from normal brown to yellow, while the *singed bristles* (*sn*) mutation causes body bristles to be short and curled rather than long and straight. Both of these genes are X-linked.

In his experiments, Stern examined *Drosophila* XX females of genotype $y\ sn^+ / y^+\ sn$. These double heterozygotes were generally wild type in appearance, but Stern noticed that some flies carried patches of yellow body color, others had small areas of singed bristles, and still others displayed twin spots: adjacent patches of yellow cells and cells with singed bristles (**Fig. 5.28**). He assumed that mistakes in the mitotic divisions accompanying fly development could have led to these **mosaic** animals containing tissues of different genotypes. Individual yellow or singed patches could arise from chromosome loss or by mitotic nondisjunction. These errors in mitosis would yield XO cells containing only *y* (but not y^+) or *sn* (but not sn^+) alleles; such cells would show one of the recessive phenotypes.

The twin spots must have a different origin. Stern reasoned that they represented the reciprocal products of mitotic crossing-over between the *sn* gene and the centromere. The mechanism is as follows: During mitosis in a diploid cell, after chromosome duplication, homologous chromosomes occasionally—but rarely—pair up with each other. While the chromosomes are paired, nonsister chromatids can exchange parts by crossing-over. The pairing is transient, and the homologous chromosomes soon resume their independent positions on the mitotic metaphase plate. There, the two chromosomes can line up relative to each other in either of two ways (**Fig. 5.29a**). One of these orientations would yield two daughter cells that remain heterozygous for both genes and are thus indistinguishable from the surrounding wild-type cells. The other orientation, however, will generate two homozygous daughter cells: one $y\ sn^+ / y\ sn^+$, the other $y^+sn\ /\ y^+sn$. Because the two daughter cells would lie next to each other, subsequent mitotic divisions would produce adjacent patches of *y* and *sn* tissue (that is, twin spots). Note that if crossing-over occurs between *sn* and *y,* single spots of yellow tissue can form, but a reciprocal singed spot cannot be generated in this fashion (**Fig. 5.29b**).

Figure 5.28 Twin spots: A form of genetic mosaicism. In a $y\ sn^+ / y^+sn$ *Drosophila* female, most of the body is wild type, but aberrant patches showing either yellow color or singed bristles sometimes occur. In some cases, yellow and singed patches are adjacent to each other, a configuration known as *twin spots.*

Single yellow spot | Twin spot | Single singed spot

Sectored Yeast Colonies Can Arise from Mitotic Recombination

Diploid yeast cells that are heterozygous for one or more genes exhibit mitotic recombination in the form of **sectors:** portions of a growing colony that have a different genotype than the remainder of the colony. If a diploid yeast cell of genotype *ADE2 / ade2* is placed on a petri plate, its mitotic descendents will grow into a colony. Usually, such colonies will appear white because the dominant wild-type *ADE2* allele specifies that color. However, many colonies will contain red sectors of diploid *ade2 / ade2* cells (**Fig. 5.30**). These cells are red because a block in the adenine biosynthesis pathway causes them to accumulate red pigment. The red sectors arose as a result of mitotic recombination between the *ADE2* gene and its centromere. (Homozygous *ADE2 / ADE2* cells will also be produced by the same event, but they cannot be distinguished from heterozygotes because both types of cells are white.)

The size of a red sector relative to the size of the colony as a whole indicates when mitotic recombination took place. If a red sector is relatively large, mitotic recombination

Figure 5.29 Mitotic crossing-over. (a) In a *y* sn^+ / y^+ *sn Drosophila* female, a mitotic crossover between the centromere and *sn* can produce two daughter cells, one homozygous for *y* and the other homozygous for *sn,* that can develop into adjacent aberrant patches (twin spots). This outcome depends on a particular distribution of chromatids at anaphase (*top*). If the chromatids are arranged in the equally likely opposite orientation, only phenotypically normal cells will result (*bottom*). **(b)** Crossovers between *sn* and *y* can generate single yellow patches. However, a single mitotic crossover in these females cannot produce a single singed spot if the *sn* gene is closer to the centromere than the *y* gene.

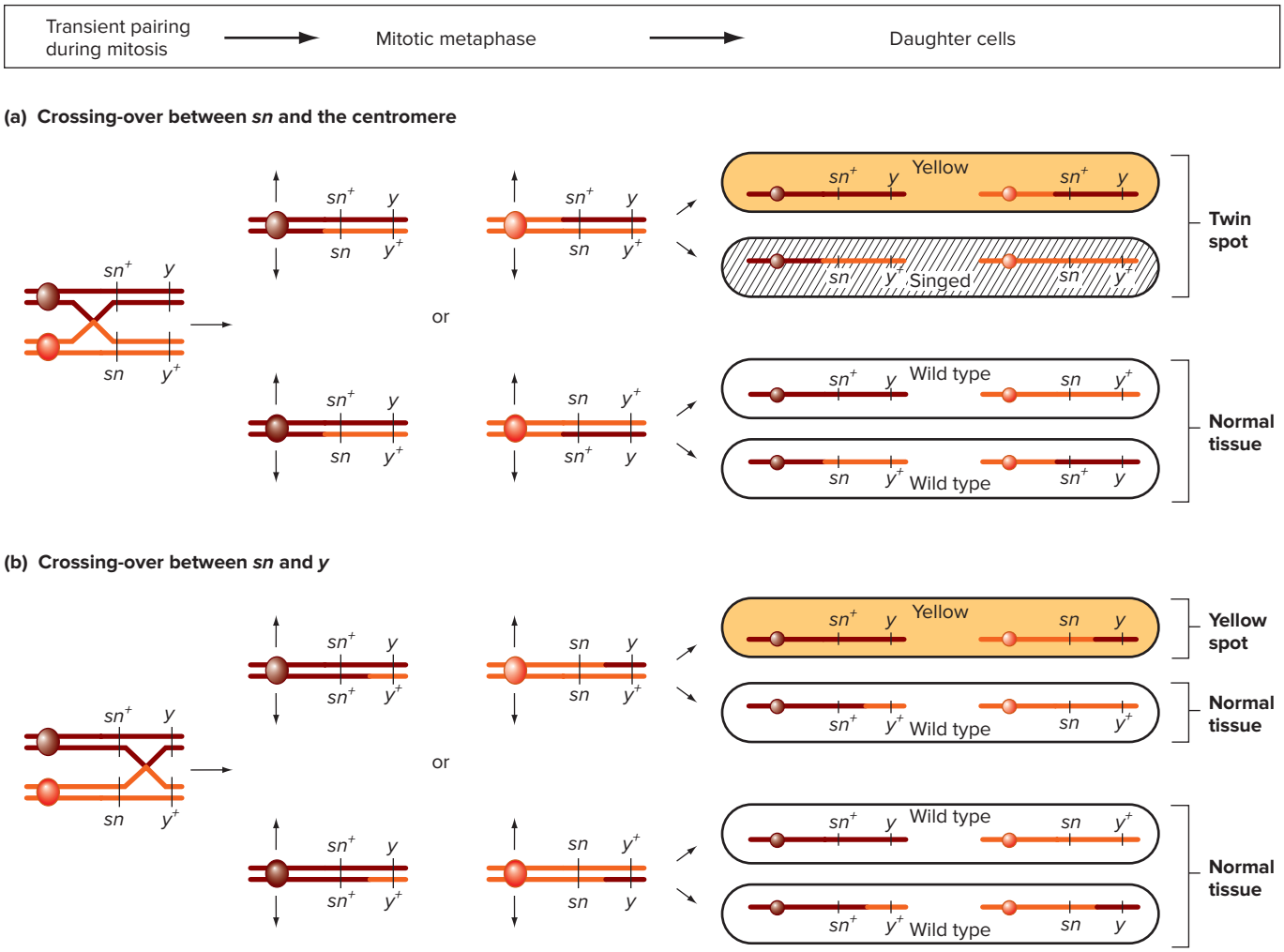
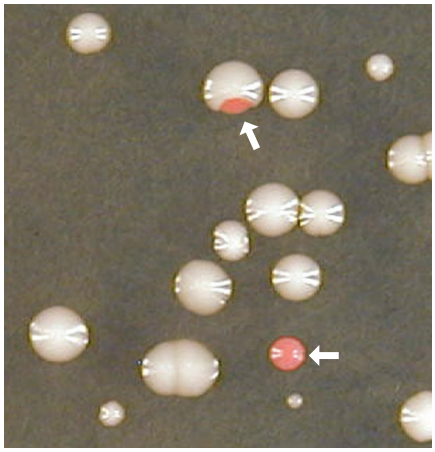


Figure 5.30 Mitotic recombination during the growth of diploid yeast colonies can create sectors. Arrows point to large, red *ade2 / ade2* sectors formed from *ADE2 / ade2* heterozygotes.
Image courtesy of B.A. Montelone, Ph.D. and T.R. Manney, Ph.D



happened during a cell division early in the growth of the colony, giving the resulting daughter cells a long time to proliferate. If a red sector is small, the recombination happened later.

Mitotic Recombination Has Significant Consequences

Problem 51 at the end of this chapter illustrates how geneticists use mitotic recombination to obtain information about the locations of genes relative to each other and to the centromere. Mitotic crossing-over has also been of great value in the study of development because it can generate animals in which different cells have different genotypes (see Problem 52 and also Chapter 19). Finally, as the Genetics and Society Box *Mitotic Recombination and Cancer Formation* explains, mitotic recombination can have major repercussions for human health.

GENETICS AND SOCIETY

Crowd: © Image Source/Getty Images RF

Mitotic Recombination and Cancer Formation

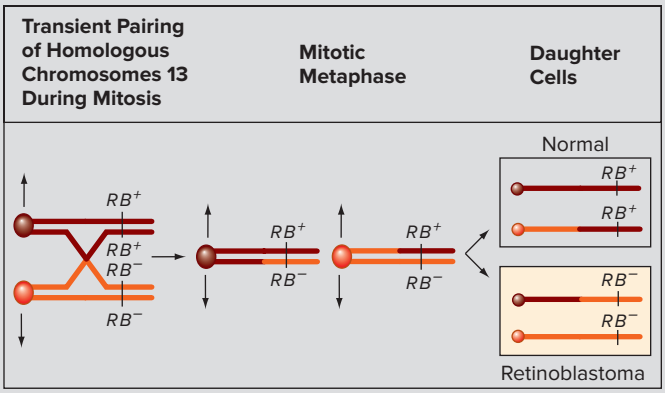
In humans, some tumors, such as those found in retinoblastoma, may arise as a result of mitotic recombination. Recall from the discussion of penetrance and expressivity in Chapter 3 that retinoblastoma is a form of eye cancer. The retinoblastoma gene (*RB*) resides on chromosome 13, where the normal wild-type allele (RB^+) encodes a protein that regulates retinal growth and differentiation. Cells in the eye need at least one copy of the normal wild-type allele to maintain control over cell division. The normal, wild-type RB^+ allele is thus known as a *tumor-suppressor gene*.

People with a genetic predisposition to retinoblastoma are born with only one functional copy of the normal RB^+ allele; their second chromosome 13 carries either a nonfunctional RB^- allele or no *RB* gene at all. If a mutagen (such as radiation) or a mistake in gene replication or segregation destroys or removes the single remaining normal copy of the gene in a retinal cell in either eye, a retinoblastoma tumor will develop at that site. In one study of people with a genetic predisposition to retinoblastoma, cells taken from eye tumors were RB^- homozygotes, while white blood cells from the same people were RB^+/RB^- heterozygotes. As **Fig. A** shows, mitotic recombination between the *RB* gene and the centromere of the chromosome carrying the gene provides one mechanism by which a cell in an RB^+/RB^- individual could become RB^-/RB^-. Once a homozygous RB^- cell is generated, it can divide uncontrollably, leading to tumor formation.

Only 40% of retinoblastoma cases follow the preceding scenario. The other 60% occur in people who are born with two normal copies of the *RB* gene. In such people, it takes two mutational events to cause the cancer. The first of these must convert an RB^+ allele to RB^-, while the second could be a mitotic recombination producing daughter cells that become cancerous because they are homozygous for the newly mutant, nonfunctional allele.

The role of mitotic recombination in the formation of retinoblastoma helps explain the incomplete penetrance and variable expressivity of the disease. People born as RB^+/RB^- heterozygotes may or may not develop the condition (incomplete penetrance). If, as usually happens, they do, they may have tumors in one or both eyes (variable expressivity). It all depends on whether and in what cells of the body mitotic recombination (or some other "homozygosing" event that affects chromosome 13) occurs.

Figure A How mitotic crossing-over can contribute to cancer. Mitotic recombination during retinal growth in an RB^-/RB^+ heterozygote may produce an RB^-/RB^- daughter cell that lacks a functional retinoblastoma gene and thus divides out of control. The crossover must occur between the *RB* gene and its centromere. Only the arrangement of chromatids yielding this result is shown.



essential concepts

- *Twin spots* are a form of *genetic mosaicism;* these spots occur when *mitotic recombination* gives rise to two clones of cells having reciprocal mutant genotypes and phenotypes.
- Mitotic recombination can also produce *sectored colonies* in diploid yeast, in which part of a colony has a recognizable mutant phenotype.

WHAT'S NEXT

Medical geneticists have used their understanding of linkage, recombination, and mapping to make sense of the pedigrees shown at the beginning of this chapter (see Fig. 5.1). The X-linked gene for red-green color blindness must lie very close to the gene for hemophilia A because the two are tightly coupled. In fact, the genetic distance between the two genes is only 3 m.u. The sample size in Fig. 5.1a was so small that none of the individuals in the pedigree were recombinant types. In contrast, even though the hemophilia B locus is also on the X chromosome, it lies far enough away

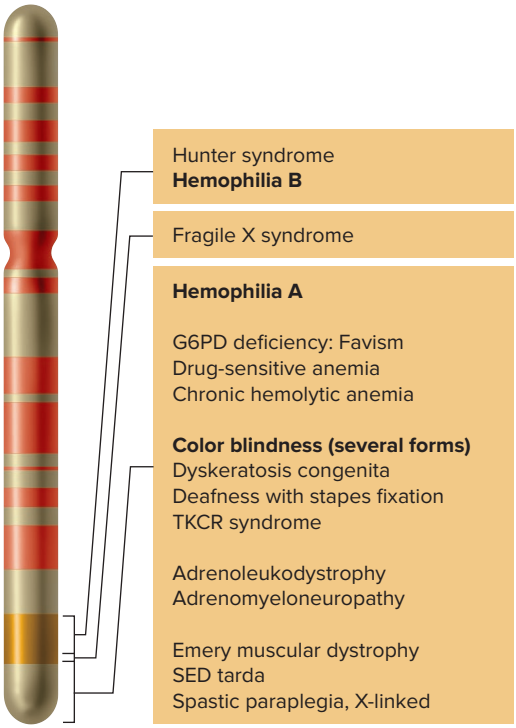
DNA: © Design Pics/Bilderbuch RF

from the red-green color blindness locus that the two genes recombine often. The color blindness and hemophilia B genes may appear to be genetically unlinked in a small sample (as in Fig. 5.1b), but the actual recombination distance separating the two genes is about 36 m.u. Pedigrees pointing to two different forms of hemophilia, one very closely linked to color blindness, the other almost not linked at all, provided one of several indications that hemophilia is determined by more than one gene (**Fig. 5.31**).

Linkage and recombination are universal among lifeforms and must therefore confer important advantages to living organisms. Geneticists reason that linkage provides the potential for transmitting favorable combinations of genes intact to successive generations, while recombination produces great flexibility in generating new combinations of alleles. Some new combinations may help a species adapt to changing environmental conditions, whereas the inheritance of successfully tested combinations can preserve what has worked in the past.

Thus far, this book has examined how genes and chromosomes are transmitted. As important and useful as this knowledge is, it tells us very little about the structure and mode of action of the genetic material. In the next section (Chapters 6–8), we carry our analysis to the level of DNA, the molecule of heredity.

Figure 5.31 A genetic map of part of the human X chromosome.



SOLVED PROBLEMS

I. The *XG* locus on the human X chromosome has two alleles, XG^+ and *XG*. The XG^+ allele causes the presence of the Xg surface antigen on red blood cells, while the recessive *XG* allele does not allow antigen to appear. The *XG* locus is 10 m.u. from the *STS* locus. The STS^+ allele produces normal activity of the enzyme steroid sulfatase, while the recessive *STS* allele results in the lack of steroid sulfatase activity and the disease ichthyosis (scaly skin). A man with ichthyosis and no Xg antigen has a normal daughter with Xg antigen. This daughter is expecting a child.

a. If the child is a son, what is the probability he will lack Xg antigen and have ichthyosis?

b. What is the probability that a son would have both the antigen and ichthyosis?

c. If the child is a son with ichthyosis, what is the probability he will have Xg antigen?

Answer

a. This problem requires an understanding of how linkage affects the proportions of gametes. First designate the genotype of the individual in which recombination during meiosis affects the transmission of alleles: in this problem, the daughter. The X chromosome she inherited from her father (who had icthyosis and no Xg antigen) must be *STS XG*. (No recombination could have separated the genes during meiosis in her father since he has only one X chromosome.) Because the daughter is normal and has the Xg antigen, her other X chromosome (inherited from her mother) must contain the STS^+ and XG^+ alleles. Her X chromosomes can be diagrammed as:

Because the *STS* and *XG* loci are 10 m.u. apart on the chromosome, the recombination frequency is 10%. Ninety percent of the gametes will be parental: *STS XG* or STS^+ XG^+ (45% of each type) and 10% will be recombinant: *STS* XG^+ or STS^+ *XG* (5% of each type). The phenotype of a son directly reflects the genotype of the X chromosome from his mother. Therefore, the probability that he will lack the Xg antigen and have icthyosis (genotype: *STS XG* / Y) is 45/100.

b. The probability that he will have the antigen and ichthyosis (genotype: *STS* XG^+/ Y) is 5/100.

c. Two classes of gametes exist that contain the ichthyosis allele: *STS XG* (45%) and *STS* XG^+ (5%). If the total number of gametes is 100, then 50 will have the *STS* allele. Of those gametes, 5 (or 10%) will have the XG^+ allele. Therefore a 1/10 probability exists that a son with the *STS* allele will have the Xg antigen.

II. *Drosophila* females of wild-type appearance but heterozygous for three autosomal genes are mated with males showing the three corresponding autosomal recessive traits: glassy eyes, coal-colored bodies, and striped thoraxes. One thousand (1000) progeny of this cross are distributed in the following phenotypic classes:

Wild type	27
Striped thorax	11
Coal body	484
Glassy eyes, coal body	8
Glassy eyes, striped thorax	441
Glassy eyes, coal body, striped thorax	29

a. Draw a genetic map based on these data.

b. Show the arrangement of alleles on the two homologous chromosomes in the parent females.

c. Normal-appearing males containing the same chromosomes as the parent females in the preceding cross are mated with females showing glassy eyes, coal-colored bodies, and striped thoraxes. Of 1000 progeny produced, indicate the numbers of the various phenotypic classes you would expect.

Answer

A logical, methodical way to approach a three-point cross is described here.

a. Designate the alleles:

t^+ = wild-type thorax	t = striped thorax
g^+ = wild-type eyes	g = glassy eyes
c^+ = wild-type body	c = coal-colored body

In solving a three-point cross, designate the types of events that gave rise to each group of individuals and the genotypes of the gametes obtained from their mother. [The paternal gametes contain only the recessive alleles of these genes (*t g c*). These alleles from the father allow the traits associated with the recessive maternal alleles to appear in the progeny.]

Progeny	Number	Type of event	Genotype
1. wild type	27	single crossover	t^+ g^+ c^+
2. striped thorax	11	single crossover	t g^+ c^+
3. coal body	484	parental	t^+ g^+ c
4. glassy eyes, coal body	8	single crossover	t^+ g c
5. glassy eyes, striped thorax	441	parental	t g c^+
6. glassy eyes, coal body, striped thorax	29	single crossover	t g c

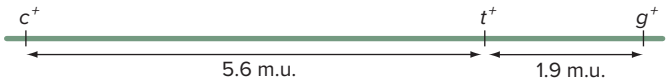
Picking out the parental classes is easy. If all the other classes are rare, the two most abundant categories are those gene combinations that have not undergone recombination. Next, two sets of two phenotypes should exist, one set corresponding to a single crossover between the first and second genes and the other set to a single crossover between the second and third genes. Finally, there should be a pair of classes containing small numbers that result from double crossovers. In this example, no flies are found in the double crossover classes, which would have been the two missing phenotypic combinations: one is glassy eyes, and the other is coal body and striped thorax.

Look at the most abundant classes to determine which alleles were on each chromosome in the female heterozygous parent. One parental class had the phenotype of coal body (484 flies), so one chromosome in the female must have contained the t^+, g^+, and *c* alleles. (Notice that we cannot yet say in what order these alleles are located on the chromosome.) The other parental class was glassy eyes and striped thorax, corresponding to a chromosome with the *t*, *g*, and c^+ alleles.

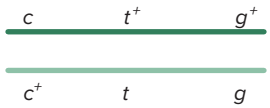
To determine the order of the genes, compare the t^+ *g* c^+ double crossover class (not seen in the data) with the most similar parental class (*t g* c^+). The alleles of *g* and *c* retain their parental associations (*g* c^+), while the *t* gene has recombined with respect to both other genes in the double recombinant class. Thus, the *t* gene is between *g* and *c*.

In order to complete the map, calculate the recombination frequencies between the center gene and each of the genes on the ends. For *g* and *t*, the nonparental combinations of alleles are in classes 2 and 4, so RF = (11 + 8)/1000 = 19/1000, or 1.9%. For *t* and *c*, classes 1 and 6 are nonparental, so RF = (27 + 29)/1000 = 56/1000, or 5.6%.

The genetic map is



b. The alleles on each chromosome were already determined (*c*, g^+, t^+ and c^+, *g*, *t*). Now that the order of loci has also been determined, the arrangement of the alleles can be indicated as follows:



c. Males of the same genotype as the starting female (*c* t^+ g^+/ c^+ *t g*) could produce only two types of gametes: parental types *c* t^+ g^+ and c^+ *t g* because no recombination occurs in male fruit flies. The progeny expected from the mating with a homozygous recessive female are thus 500 coal body and 500 glassy-eyed, striped thorax flies.

III. The following *Neurospora* asci were obtained when a wild-type strain (ad^+ leu^+) was crossed to a double mutant strain that cannot grow in the absence of adenine or leucine (ad^- leu^-). Only one member of each spore pair produced by the final mitosis is shown because the two cells in a pair have the same genotype. Total asci = 120.

Spore pair	Ascus type				
1–2	ad^+ leu^+	ad^+ leu^-	ad^+ leu^+	ad^+ leu^-	ad^- leu^+
3–4	ad^+ leu^+	ad^+ leu^-	ad^+ leu^-	ad^- leu^+	ad^+ leu^+
5–6	ad^- leu^-	ad^- leu^+	ad^- leu^+	ad^- leu^-	ad^- leu^-
7–8	ad^- leu^-	ad^- leu^+	ad^- leu^-	ad^+ leu^+	ad^+ leu^-
# of asci	30	30	40	2	18

a. What genetic event causes the alleles of two genes to segregate to different cells at the second meiotic division, and when does this event occur?

b. Provide the best possible map for the two genes and the relevant centromere(s).

Answer

This problem requires an understanding of tetrad analysis and the process (meiosis) that produces the patterns seen in ordered asci.

a. A crossover between a gene and the centromere causes the segregation of alleles at the second meiotic division. The crossover event itself occurs during prophase of meiosis I.

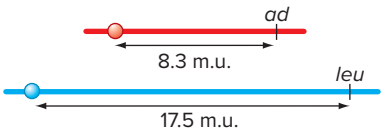
b. Using ordered tetrads (or ordered octads) you can determine whether two genes are linked, the distance between two genes, and the distance between each gene and the centromere of the chromosome on which the gene is located. First designate the five classes of asci shown. The first class is a PD; the second is an NPD; the last three are Ts. Next determine if these genes are linked. The number of PD = number of NPD, so the genes are not linked. When genes are unlinked, the T asci are generated by a crossover between a gene and the centromere. Looking at the *leu* gene, there is an MII pattern of that gene in the third and fourth asci types. Therefore, the fraction of MII asci is:

$$\frac{40 + 2}{120} \times 100 = 35\%.$$

Because only half of the chromatids in the meioses that generated these T asci were involved in the crossover, the map distance between *leu* and the centromere is 35/2, or 17.5 m.u. Asci of the fourth and fifth types show an MII pattern for the *ad* gene:

$$\frac{2 + 18}{120} \times 100 = 16.6\%.$$

Dividing 16.6% by 2 gives the *ad* gene ↔ centromere map distance of 8.3 m.u. The map of these two genes is the following:



PROBLEMS

Vocabulary

1. Choose the phrase from the right column that best fits the term in the left column.

a. recombination	1. a statistical method for testing the fit between observed and expected results
b. linkage	2. an ascus containing spores of four different genotypes
c. chi-square test	3. one crossover along a chromosome makes a second nearby crossover less likely
d. chiasma	4. when two loci recombine in less than 50% of gametes
e. tetratype	5. the relative chromosomal location of a gene
f. locus	6. the ratio of observed double crossovers to expected double crossovers
g. coefficient of coincidence	7. individual composed of cells with different genotypes
h. interference	8. formation of new genetic combinations by exchange of parts between homologs
i. parental ditype	9. when the two alleles of a gene are segregated into different cells at the first meiotic division
j. ascospores	10. an ascus containing only two nonrecombinant kinds of spores
k. first-division segregation	11. structure formed at the spot where crossing-over occurs between homologs
l. mosaic	12. fungal spores contained in a sac

Section 5.1

2. a. A *Drosophila* male from a true-breeding stock with scabrous eyes was mated with a female from a true-breeding stock with javelin bristles. Both scabrous eyes and javelin bristles are autosomal recessive mutant traits. The F_1 progeny all had normal eyes and bristles. F_1 females from this cross were mated with males with both scabrous eyes and javelin bristles. Write all the possible phenotypic classes of the progeny that could be produced from

the cross of the F_1 females with the scabrous, javelin males, and indicate for each class whether it is a recombinant or parental type.

b. The cross in part (a) yielded the following progeny: 77 scabrous eyes and normal bristles; 76 wild type (normal eyes and bristles); 74 normal eyes and javelin bristles; and 73 scabrous eyes and javelin bristles. Are the genes governing these traits likely to be linked, or do they instead assort independently? Why?

c. Suppose you mated the F_1 females from the cross in part (a) to wild-type males. Why would this cross fail to inform you whether the two genes are linked?

d. Suppose you mated females from the true-breeding stock with javelin bristles to males with scabrous eyes and javelin bristles. Why would this cross fail to inform you whether the two genes are linked?

3. With modern molecular methods it is now possible to examine variants in DNA sequence from a very small amount of tissue like a hair follicle or even a single sperm. (See the Fast Forward Box *Mapping the Crossovers that Generate the Chromosomes of Individual Human Sperm*.) You can consider these variants to be alleles of a particular site on a chromosome (a *locus*; *loci* in plural). For example, AAAAAAA, AAACAAA, AAAGAAA, and AAATAAA at the same location (call it *B*) on homologous autosomes in different sperm might be called alleles *1, 2, 3,* and *4* of locus *B* (B^1, B^2, etc.). John's genotype for two loci *B* and *D* is B^1B^3 and D^1D^3. John's father was B^1B^2 and D^1D^4, while his mother was B^3B^3 and D^2D^3.

a. What is (are) the genotype(s) of the parental type sperm John could produce?

b. What is (are) the genotype(s) of the recombinant type sperm John could produce?

c. In a sample of 100 sperm, 51 of John's sperm were found to be B^1 and D^1, while the remaining 49 sperm were B^3 D^3. Can you conclude whether the *B* and *D* loci are linked, or whether they instead assort independently?

4. The Punnett square in Fig. 5.4 shows how Mendel's dihybrid cross results would have been altered had the two genes (*A* and *B*) been linked and had the P generation cross been *A B / A B* × *a b / a b.*

a. What would be the frequency of each F_2 phenotypic class if 80% of the gametes produced by the F_1s were parentals?

b. Answer part (a) assuming that the original, P generation cross was *A b / A b* × *a B / a B*.

Section 5.2

5. In mice, the dominant allele *Gs* of the X-linked gene *Greasy* produces shiny fur, while the recessive wild-type Gs^+ allele determines normal fur. The dominant allele *Bhd* of the X-linked *Broadhead* gene causes skeletal abnormalities including broad heads and snouts, while the recessive wild-type Bhd^+ allele yields normal skeletons. Female mice heterozygous for the two alleles of both genes were mated with wild-type males. Among 100 male progeny of this cross, 49 had shiny fur, 48 had skeletal abnormalities, 2 had shiny fur and skeletal abnormalities, and 1 was wild type.

a. Diagram the cross described and calculate the distance between the two genes.

b. What would have been the results if you had counted 100 female progeny of the cross?

6. In *Drosophila,* males from a true-breeding stock with raspberry-colored eyes were mated to females from a true-breeding stock with sable-colored bodies. In the F_1 generation, all the females had wild-type eye and body color, while all the males had wild-type eye color but sable-colored bodies. When F_1 males and females were mated, the F_2 generation was composed of 216 females with wild-type eyes and bodies, 223 females with wild-type eyes and sable bodies, 191 males with wild-type eyes and sable bodies, 188 males with raspberry eyes and wild-type bodies, 23 males with wild-type eyes and bodies, and 27 males with raspberry eyes and sable bodies. Explain these results by diagramming the crosses and calculating any relevant map distances.

7. If the *a* and *b* loci are 20 m.u. apart in humans and an *A B / a b* woman mates with an *a b / a b* man, what is the probability that their first child will be *A b / a b?*

8. *CC DD* and *cc dd* individuals were crossed to each other, and the F_1 generation was backcrossed to the *cc dd* parent. 997 *Cc Dd,* 999 *cc dd,* 1 *Cc dd,* and 3 *cc Dd* offspring resulted.

a. How far apart are the *c* and *d* loci?

b. What progeny and in what frequencies would you expect to result from testcrossing the F_1 generation from a *CC dd* × *cc DD* cross to *cc dd?*

c. In a typical meiosis, how many crossovers occur between genes *C* and *D?*

d. Assume that the *C* and *D* loci are on the same chromosome, but the offspring from the testcross described in part (b) were 498 *Cc Dd,* 502 *cc dd,* 504 *Cc dd,* and 496 *cc Dd*. How would your answer to part (c) change?

9. In mice, the autosomal locus coding for the β-globin chain of hemoglobin is 1 m.u. from the albino locus.

Assume for the moment that the same is true in humans. The disease sickle-cell anemia is the result of homozygosity for a particular mutation in the β-globin gene.

a. A son is born to an albino man and a woman with sickle-cell anemia. What kinds of gametes will the son form, and in what proportions?

b. A daughter is born to a normal man and a woman who has both albinism and sickle-cell anemia. What kinds of gametes will the daughter form, and in what proportions?

c. If the son in part (a) grows up and marries the daughter in part (b), what is the probability that a child of theirs will be an albino with sickle-cell anemia?

10. In a particular human family, John and his mother both have brachydactyly (a rare autosomal dominant allele causing short fingers). John's father has Huntington disease (another rare autosomal dominant allele). John's wife is phenotypically normal and is pregnant. Two-thirds of people who inherit the Huntington (*HD*) allele show symptoms by age 50, and John is 50 and has no symptoms. Brachydactyly is 90% penetrant.

a. What are the genotypes of John's parents?

b. What are the possible genotypes for John? How likely is John to have each of these genotypes?

c. What is the probability the child will express both brachydactyly and Huntington disease by age 50 if the two genes are unlinked?

d. How will your answer to part (c) change if instead these two loci are 20 m.u. apart?

11. Albino rabbits (lacking pigment) are homozygous for the recessive *c* allele (*C* allows pigment formation). Rabbits homozygous for the recessive *b* allele make brown pigment, while those with at least one copy of *B* make black pigment. True-breeding brown rabbits were crossed to albinos, which were also *BB*. F_1 rabbits, which were all black, were crossed to the double recessive (*bb cc*). The progeny obtained were 34 black, 66 brown, and 100 albino.

a. What phenotypic proportions would have been expected if the *b* and *c* loci were unlinked?

b. How far apart are the two loci?

12. In corn, the allele *A* allows the deposition of anthocyanin (blue) pigment in the kernels (seeds), while *aa* plants have yellow kernels. At a second gene, *W*– produces smooth kernels, while *ww* kernels are wrinkled. A plant with blue smooth kernels was crossed to a plant with yellow wrinkled kernels. The progeny consisted of 1447 blue smooth, 169 blue wrinkled, 186 yellow smooth, and 1510 yellow wrinkled.

a. Are the *a* and *w* loci linked? If so, how far apart are they?

b. What was the genotype of the blue smooth parent? Include the chromosome arrangement of alleles.

c. If a plant grown from a blue wrinkled progeny seed is crossed to a plant grown from a yellow smooth F_1 seed, what kinds of kernels would be expected, and in what proportions?

13. If the *a* and *b* loci are 40 cM apart and an *AA BB* individual and an *aa bb* individual mate:

a. What gametes will the F_1 individuals produce, and in what proportions? What phenotypic classes in what proportions are expected in the F_2 generation (assuming complete dominance for both genes)?

b. If the original cross was *AA bb* × *aa BB,* what gametic proportions would emerge from the F_1? What would be the result in the F_2 generation?

14. Write the number of *different kinds* of phenotypes, excluding sex, you would see among a large number of progeny from an F_1 mating between individuals of identical genotype that are heterozygous for one or two genes (that is, *Aa* or *Aa Bb*) as indicated. No gene interactions means that the phenotype determined by one gene is not influenced by the genotype of the other gene.

a. One gene; *A* completely dominant to *a.*

b. One gene; *A* and *a* codominant.

c. One gene; *A* incompletely dominant to *a.*

d. Two unlinked genes; no gene interactions; *A* completely dominant to *a,* and *B* completely dominant to *b.*

e. Two genes, 10 m.u. apart; no gene interactions; *A* completely dominant to *a,* and *B* completely dominant to *b.*

f. Two unlinked genes; no gene interactions; *A* and *a* codominant, and *B* incompletely dominant to *b.*

g. Two genes, 10 m.u. apart; *A* completely dominant to *a,* and *B* completely dominant to *b;* and with recessive epistasis between *aa* and the alleles of gene *B.*

h. Two unlinked duplicated genes (that is, *A* and *B* perform the same function); *A* and *B* completely dominant to *a* and *b,* respectively.

i. Two genes, 0 m.u. apart; no gene interactions; *A* completely dominant to *a,* and *B* completely dominant to *b.* (Two possible answers exist.)

15. A DNA variant has been found linked to a rare autosomal dominant disease in humans and can thus be used as a marker to follow inheritance of the disease allele. In an *informative family* (in which one parent is heterozygous for both the disease allele and the

DNA marker in a known chromosomal arrangement of alleles, and his or her mate does not have the same alleles of the DNA variant), the reliability of such a marker as a predictor of the disease in a fetus is related to the map distance between the DNA marker and the gene causing the disease.

Imagine that a man affected with the disease (genotype *Dd*) is heterozygous for the V^1 and V^2 forms of the DNA variant, with form V^1 on the same chromosome as the *D* allele and form V^2 on the same chromosome as *d.* His wife is V^3V^3 *dd,* where V^3 is another allele of the DNA marker. Typing of the fetus by amniocentesis reveals that the fetus has the V^2 and V^3 variants of the DNA marker. How likely is it that the fetus has inherited the disease allele *D* if the distance measured in a two-point cross between the *D* locus and the marker locus is (a) 0 m.u., (b) 1 m.u., (c) 5 m.u., (d) 10 m.u., (e) 50 m.u.?

16. Figure 5.7a shows chromosomes during prophase of meiosis I in a mouse primary spermatocyte.
    a. How would you know immediately that this figure shows male meiosis in a mouse and not in a human? (*Hint:* In mice, $n = 20$.)
    b. Are most mouse chromosomes metacentric or acrocentric? Explain.
    c. How many chromatids in total are represented in Fig. 5.7a?
    d. Where is the X-Y bivalent in Fig. 5.7a? (*Note:* Mouse sex chromosomes have only a single pseudoautosomal region, instead of two as in humans.) Diagram this bivalent, showing the X and Y chromosomes, the locations of the centromeres of these chromosomes, and the pseudoautosomal region.
    e. Explain the importance of the pseudoautosomal region(s) of the sex chromosomes of mammals for ensuring proper sex chromosome segregation during meiosis.
17. Figure 5.7b shows bivalents in mouse primary spermatocytes that have previously undergone recombination events, as indicated by the presence of a single chiasma. The artist's image in Fig. 5.7c depicts how cohesin complexes are involved in keeping the homologous chromosomes together within the bivalent. Explain using diagrams why the key cohesin complexes that connect the homologous chromosomes are those located distal to a chiasma (that is, farther away from the centromere) rather than those located proximal to a chiasma (that is, between the centromere and the chiasma).

Section 5.3

18. Cinnabar eyes (*cn*) and reduced bristles (*rd*) are autosomal recessive characters in *Drosophila.* A homozygous wild-type female was crossed to a reduced, cinnabar male, and the F_1 males were then crossed to the F_1 females to obtain the F_2. Of the 400 F_2 offspring obtained, 292 were wild type, 9 were cinnabar, 7 were reduced, and 92 were reduced and cinnabar. Explain these results and estimate the distance between the *cn* and *rd* loci.
19. In *Drosophila,* the autosomal recessive *dp* allele of the *dumpy* gene produces short, curved wings, while the autosomal recessive allele *bw* of the *brown* gene causes brown eyes. In a testcross using females heterozygous for both of these genes, the following results were obtained:

wild-type wings, wild-type eyes	178
wild-type wings, brown eyes	185
dumpy wings, wild-type eyes	172
dumpy wings, brown eyes	181

In a testcross using males heterozygous for both of these genes, a different set of results was obtained:

wild-type wings, wild-type eyes	247
dumpy wings, brown eyes	242

    a. What can you conclude from the first testcross?
    b. What can you conclude from the second testcross?
    c. How can you reconcile the data shown in parts (a) and (b)*?* Can you exploit the difference between these two sets of data to devise a general test for synteny in *Drosophila?*
    d. The genetic distance between *dumpy* and *brown* is 91.5 m.u. How could this value be measured?
20. From a series of two-point crosses, the following map distances were obtained for the syntenic genes *A, B, C,* and *D* in peas:

$B \leftrightarrow C$	23 m.u.
$A \leftrightarrow C$	15 m.u.
$C \leftrightarrow D$	14 m.u.
$A \leftrightarrow B$	12 m.u.
$B \leftrightarrow D$	11 m.u.
$A \leftrightarrow D$	1 m.u.

Chi-square analysis cannot reject the null hypothesis of no linkage for gene *E* with any of the other four genes.

    a. Draw a cross scheme that would allow you to determine the $B \leftrightarrow C$ map distance.
    b. Diagram the best genetic map that can be assembled from this data set.
    c. Explain any inconsistencies or unknown features in your map.
    d. What additional experiments would allow you to resolve these inconsistencies or ambiguities?

21. Map distances were determined for four different genes (*MAT, HIS4, THR4,* and *LEU2*) on chromosome III of the yeast *Saccharomyces cerevisiae:*

HIS4 ↔ *MAT*	37 cM
THR4 ↔ *LEU2*	35 cM
LEU2 ↔ *HIS4*	23 cM
MAT ↔ *LEU2*	16 cM
MAT ↔ *THR4*	20 cM

What is the order of genes on the chromosome?

22. In the tubular flowers of foxgloves, wild-type coloration is red while a mutation called *white* produces white flowers. Another mutation, called *peloria,* causes the flowers at the apex of the stem to be huge. Yet another mutation, called *dwarf,* affects stem length. You cross a white-flowered plant (otherwise phenotypically wild type) to a plant that is dwarf and peloria but has wild-type red flower color. All of the F_1 plants are tall with white, normal-sized flowers. You cross an F_1 plant back to the dwarf and peloria parent, and you see the 543 progeny shown in the chart. (Only mutant traits are noted.)

dwarf, peloria	172
white	162
dwarf, peloria, white	56
wild type	48
dwarf, white	51
peloria	43
dwarf	6
peloria, white	5

a. Which alleles are dominant?

b. What were the genotypes of the parents in the original cross?

c. Draw a map showing the linkage relationships of these three loci.

d. Do the data provide evidence for interference? If so, calculate the coefficient of coincidence and the interference value.

23. In *Drosophila,* the recessive allele *mb* of one gene causes missing bristles, the recessive allele *e* of a second gene causes ebony body color, and the recessive allele *k* of a third gene causes kidney-shaped eyes. (Dominant wild-type alleles of all three genes are indicated with a + superscript.) The three different P generation crosses in the table that follows were conducted, and then the resultant F_1 females from each cross were testcrossed to males that were homozygous for the recessive alleles of both genes in question. The phenotypes of the testcross offspring are tabulated. Determine the best genetic map explaining all the data.

Parental cross	**Testcross offspring of F_1 females**	
$mb^+\ mb^+,\ e^+\ e^+ \times$	normal bristles, normal body	117
mb mb, e e	normal bristles, ebony body	11
	missing bristles, normal body	15
	missing bristles, ebony body	107
$k^+\ k^+,\ e\ e \times k\ k,\ e^+\ e^+$	normal eyes, normal body	11
	normal eyes, ebony body	150
	kidney eyes, normal body	144
	kidney eyes, ebony body	7
$mb^+\ mb^+,\ k^+\ k^+ \times$	normal bristles, normal eyes	203
mb mb, k k	normal bristles, kidney eyes	11
	missing bristles, normal eyes	15
	missing bristles, kidney eyes	193

24. A snapdragon with pink petals, black anthers, and long stems was allowed to self-fertilize. From the resulting seeds, 650 adult plants were obtained. The phenotypes of these offspring are listed here.

78	red	long	tan
26	red	short	tan
44	red	long	black
15	red	short	black
39	pink	long	tan
13	pink	short	tan
204	pink	long	black
68	pink	short	black
5	white	long	tan
2	white	short	tan
117	white	long	black
39	white	short	black

a. Using *P* for one allele and *p* for the other, indicate how flower color is inherited.

b. What numbers of red : pink : white would have been expected among these 650 plants?

c. How are anther color and stem length inherited?

d. What was the genotype of the original plant?

e. Do any of the three genes show independent assortment?

f. For any genes that are linked, indicate the arrangements of the alleles on the homologous chromosomes in the original snapdragon, and estimate the distance between the genes.

25. In *Drosophila,* three autosomal genes have the following map:

a. Provide the data, in terms of the expected number of flies in the following phenotypic classes, when $a^+\ b^+\ c^+$/ *a b c* females are crossed to *a b c* / *a b c*

males. Assume 1000 flies are counted and that no interference exists in this region.

a^+	b^+	c^+
a	b	c
a^+	b	c
a	b^+	c^+
a^+	b^+	c
a	b	c^+
a^+	b	c^+
a	b^+	c

b. If the cross was reversed, such that $a^+ b^+ c^+ / a\ b\ c$ males are crossed to $a\ b\ c / a\ b\ c$ females, how many flies would you expect in the same phenotypic classes?

26. *Drosophila* females heterozygous for each of three recessive autosomal mutations with independent phenotypic effects [thread antennae (*th*), hairy body (*h*), and scarlet eyes (*st*)] were testcrossed to males showing all three mutant phenotypes. The 1000 progeny of this testcross were

thread, hairy, scarlet	432
wild type	429
thread, hairy	37
thread, scarlet	35
hairy	34
scarlet	33

a. Show the arrangement of alleles on the relevant chromosomes in the triply heterozygous females.

b. Draw the best genetic map that explains these data.

c. Calculate any relevant interference values.

27. Male *Drosophila* expressing the autosomal recessive mutations *sc (scute), ec (echinus), cv (crossveinless),* and *b (black)* were crossed to phenotypically wild-type females, and the 3288 progeny listed were obtained. (Only mutant traits are noted.)

653	black, scute, echinus, crossveinless
670	scute, echinus, crossveinless
675	wild type
655	black
71	black, scute
73	scute
73	black, echinus, crossveinless
74	echinus, crossveinless
87	black, scute, echinus
84	scute, echinus
86	black, crossveinless
83	crossveinless
1	black, scute, crossveinless
1	scute, crossveinless
1	black, echinus
1	echinus

a. Diagram the genotype of the female parent.

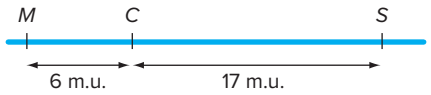
b. Map these loci.

c. Do the data provide evidence of interference? Justify your answer with numbers.

28. a. In *Drosophila,* crosses between F_1 heterozygotes of the form $A\ b / a\ B$ always yield the same ratio of phenotypes in the F_2 progeny regardless of the distance between the two genes (assuming complete dominance for both autosomal genes). What is this ratio? Would this also be the case if the F_1 heterozygotes were $A\ B / a\ b$? (*Hint:* Remember that in *Drosophila,* recombination does not take place during spermatogenesis.)

b. If you intercrossed F_1 heterozygotes of the form $A\ b / a\ B$ in mice, the phenotypic ratio among the F_2 progeny would vary with the map distance between the two genes. Is there a simple way to estimate the map distance based on the frequencies of the F_2 phenotypes, assuming rates of recombination are equal in males and females? Could you estimate map distances in the same way if the mouse F_1 heterozygotes were $A\ B / a\ b$?

29. A true-breeding strain of Virginia tobacco has dominant alleles determining leaf morphology (*M*), leaf color (*C*), and leaf size (*S*). A Carolina strain is homozygous for the recessive alleles of these three genes. These genes are found on the same chromosome as follows:



An F_1 hybrid between the two strains is now backcrossed to the Carolina strain. Assuming no interference:

a. What proportion of the backcross progeny will resemble the Virginia strain for all three traits?

b. What proportion of the backcross progeny will resemble the Carolina strain for all three traits?

c. What proportion of the backcross progeny will have the leaf morphology and leaf size of the Virginia strain but the leaf color of the Carolina strain?

d. What proportion of the backcross progeny will have the leaf morphology and leaf color of the Virginia strain but the leaf size of the Carolina strain?

30. In humans, the correlation between recombination frequency and length of DNA sequence is, on average, 1 million bp per 1% RF. During the process of mapping the Huntington disease gene (*HD*), it was found that *HD* was linked to a DNA marker called *G8* with an RF of 5%. Surprisingly, when the *HD* gene was finally identified, its physical distance from *G8* was found to be about 500,000 bp, instead of the expected 5 million bp. How can this observation be explained?

31. The following list of four *Drosophila* mutations indicates the symbol for the mutation, the name of the gene, and the mutant phenotype:

Allele symbol	Gene name	Mutant phenotype
dwp	*dwarp*	small body, warped wings
rmp	*rumpled*	deranged bristles
pld	*pallid*	pale wings
rv	*raven*	dark eyes and bodies

You perform the following crosses with the indicated results:

Cross #1: dwarp, rumpled females × pallid, raven males
→ dwarp, rumpled males and wild-type females

Cross #2: pallid, raven females × dwarp, rumpled males
→ pallid, raven males and wild-type females

F_1 females from cross #1 were crossed to males from a true-breeding *dwarp rumpled pallid raven* stock. The 1000 progeny obtained were as follows:

pallid	3
pallid, raven	428
pallid, raven, rumpled	48
pallid, rumpled	23
dwarp, raven	22
dwarp, raven, rumpled	2
dwarp, rumpled	427
dwarp	47

Indicate the best map for these four genes, including all relevant data. Calculate interference values where appropriate.

32. a. Explain in a qualitative sense how Fig. 5.7a suggests the existence of interference along mouse chromosomes.
    b. If you could examine many photographs similar to Fig. 5.7a, how would you apply statistics to provide evidence for the existence of interference during meiosis in mice? (No equations required; just outline the logic involved.)
33. The total length of the mouse genetic map is 1386 cM measured in males but 1817 cM measured in females. If Fig. 5.7a on shows a prophase I representative of all primary spermatocytes, how many recombination nodules would you expect to find in a representative prophase I primary oocyte?
34. The Fast Forward Box *Mapping the Crossovers that Generate the Chromosomes of Individual Human Sperm* describes how scientists used the whole genome DNA sequencing of ~100 single sperm from the same man to locate the crossovers involved in the production of each of these sperm. To determine the sequences of the parental chromosomes between which recombination occurred in the man's primary spermatocytes, these researchers isolated and then sequenced the DNA of individual chromosomes from the man's somatic cells. However, although this last step provided useful confirmation, it was not actually essential for determining the two DNA sequences for each chromosome the man inherited, one from his mother and one from his father. Instead, as you will demonstrate, this information can be inferred from the DNA sequences of the ~100 single sperm.
    a. Diagram three syntenic autosomal SNP loci for which the man is heterozygous, and suppose that the middle SNP locus is separated from each of those flanking it by 25 m.u. Use alleles of your choice for each of the three loci in your diagram.
    b. For simplicity, assume that interference = 1. Now, accounting only for these three loci, how many different types of sperm would you see, and in what frequencies?
    c. Given the results shown in the Fast Forward Box, explain why the assumption that interference = 1 is reasonable in this case.
    d. Explain how you would design a computer program that could reconstruct the DNA sequences of each chromosome in the somatic cells of a man, given the DNA sequences of 100 single sperm he produces.

Section 5.4

35. Do the data that Mendel obtained fit his hypotheses? For example, Mendel obtained 315 yellow round, 101 yellow wrinkled, 108 green round, and 32 green wrinkled seeds from the selfing of *Yy Rr* individuals (a total of 556). His hypotheses of segregation and independent assortment predict a 9:3:3:1 ratio in this case. Use the chi-square test to determine whether Mendel's data are significantly different from what he predicted. (The chi-square test did not exist in Mendel's day, so he was not able to test his own data for goodness of fit to his hypotheses.)
36. Two genes control color in corn snakes as follows: *O– B–* snakes are brown, *O– bb* are orange, *oo B–* are black, and *oo bb* are albino. An orange snake was mated to a black snake, and a large number of F_1 progeny were obtained, all of which were brown. When the F_1 snakes were mated to one another, they produced 100 brown offspring, 25 orange, 22 black, and 13 albino.
    a. What are the genotypes of the F_1 snakes?
    b. What proportions of the different colors would have been expected among the F_2 snakes if the two loci assort independently?

c. Do the observed results differ significantly from what was expected, assuming independent assortment is occurring?

d. What is the probability that differences this great between observed and expected values would happen by chance?

37. A mouse from a true-breeding population with normal gait was crossed to a mouse displaying an odd gait called *dancing*. The F_1 animals all showed normal gait.

a. If dancing is caused by homozygosity for the recessive allele of a single gene, what proportion of the F_2 mice should be dancers?

b. If mice must be homozygous for recessive alleles of each of two different genes to have the dancing phenotype, what proportion of the F_2 should be dancers if the two genes are unlinked? (Assume that all the mice in the population with normal gait were homozygous for dominant alleles.)

c. When the F_2 mice were obtained, 42 normal and 8 dancers were seen. Use the chi-square test to determine if these results better fit the one-gene model from part (a) or the two-gene model from part (b).

Section 5.5

38. *Neurospora* of genotype *a + c* are crossed with *Neurospora* of genotype *+ b +*. (Here, *+* is shorthand for the wild-type allele.) The following tetrads are obtained (note that the genotype of the four spore pairs in an ascus are listed, rather than listing all eight spores):

a + c	*a b c*	*+ + c*	*+ b c*	*a b +*	*a + c*
a + c	*a b c*	*a + c*	*a b c*	*a b +*	*a b c*
+ b +	*+ + +*	*+ b +*	*+ + +*	*+ + c*	*+ + +*
+ b +	*+ + +*	*a b +*	*a + +*	*+ + c*	*+ b +*
137	141	26	25	2	3

a. In how many cells has meiosis occurred to yield these data?

b. Give the best genetic map to explain these results. Indicate all relevant genetic distances, both between genes and between each gene and the centromere.

c. Diagram a meiosis that could give rise to one of the three tetrads in the class at the far right in the list.

39. A cross was performed between one haploid strain of yeast with the genotype *a f g* and another haploid strain with the genotype α f^+ g^+ (*a* and α are mating types). The resulting diploid was sporulated, and a random sample of 101 of the resulting haploid spores was analyzed. The following genotypic frequencies were seen:

α	f^+	g^+	31
a	*f*	*g*	29
a	*f*	g^+	14
α	f^+	*g*	13
a	f^+	*g*	6
α	*f*	g^+	6
a	f^+	g^+	1
α	*f*	*g*	1

a. Map the loci involved in the cross.

b. Assuming all three genes are on the same chromosome arm, is it possible that a particular ascus could contain an α *f g* spore but not an *a* f^+ g^+ spore? If so, draw a meiosis that could generate such an ascus.

40. A cross was performed between a yeast strain that requires methionine and lysine for growth (met^- lys^-) and another yeast strain, which is met^+ lys^+. One hundred asci were dissected, and colonies were grown from the four spores in each ascus. Cells from these colonies were tested for their ability to grow on petri plates containing either minimal medium (min), min + lysine (lys), min + methionine (met), or min + lys + met. The asci could be divided into two groups based on this analysis:

Group 1: In 89 asci, cells from two of the four spore colonies could grow on all four kinds of media, while the other two spore colonies could grow only on min + lys + met.

Group 2: In 11 asci, cells from one of the four spore colonies could grow on all four kinds of petri plates. Cells from a second one of the four spore colonies could grow only on min + lys plates and on min + lys + met plates. Cells from a third of the four spore colonies could only grow on min + met plates and on min + lys + met. Cells from the remaining colony could only grow on min + lys + met.

a. What are the genotypes of each of the spores within the two groups of asci?

b. Are the *lys* and *met* genes linked? If so, what is the map distance between them?

c. If you could extend this analysis to many more asci, you would eventually find some asci with a different pattern. For these asci, describe the phenotypes of the four spores. List these phenotypes as the ability of dissected spores to form colonies on the four kinds of petri plates.

41. Two crosses were made in *Neurospora* involving the mating type locus and either the *ad* or *p* genes. In both cases, the mating type locus (*A* or *a*) was one of the loci whose segregation was scored. One cross was *ad A* × ad^+ *a* (cross i), and the other was *p A* × p^+ *a* (cross ii). From cross i, 10 parental ditype,

9 nonparental ditype, and 1 tetratype asci were seen. From cross ii, the results were 24 parental ditype, 3 nonparental ditype, and 27 tetratype asci.

a. What are the linkage relationships between the mating type locus and the other two loci?

b. Although these two crosses were performed in *Neurospora,* you cannot use the data given to calculate centromere-to-gene distances for any of these genes. Why not?

42. Indicate the percentage of tetrads that would have 0, 1, 2, 3, or 4 viable spores after *Saccharomyces cerevisiae a / α* diploids of the following genotypes are sporulated:

a. A true-breeding wild-type strain (with no mutations in any gene essential for viability).

b. A strain heterozygous for a null (completely inactivating) mutation in a single essential gene.

For the remaining parts of this problem, consider crosses between yeast strains of the form $a \times b$, where *a* and *b* are both temperature-sensitive mutations in different essential genes. The cross is conducted under permissive (low-temperature) conditions. Indicate the percentage of tetrads that would have 0, 1, 2, 3, or 4 viable spores subsequently measured under restrictive (high-temperature) conditions.

c. *a* and *b* are unlinked, and both are 0 m.u. from the centromere.

d. *a* and *b* are unlinked; *a* is 0 m.u. from the centromere, while *b* is 10 m.u. from the centromere.

e. *a* and *b* are 0 m.u. apart.

f. *a* and *b* are 10 m.u. apart. Assume all crossovers between *a* and *b* are SCOs (single crossovers).

g. In part (f), if a four-strand DCO (double crossover) occurred between *a* and *b,* how many of the spores in the resulting tetrad would be viable at high temperature?

43. The *a, b,* and *c* loci are all on different chromosomes in yeast. When $a\ b^+$ yeast were crossed to $a^+\ b$ yeast and the resultant tetrads analyzed, it was found that the number of nonparental ditype tetrads was equal to the number of parental ditypes, but there were no tetratype asci at all. On the other hand, many tetratype asci were seen in the tetrads formed after $a\ c^+$ was crossed with $a^+\ c$, and after $b\ c^+$ was crossed with $b^+\ c$. Explain these results.

44. This problem leads you through the derivation of a corrected equation for RF in yeast tetrad analysis that takes into account double crossover (DCO) meioses.

A yeast strain that cannot grow in the absence of the amino acid histidine (his^-) is mated with a yeast strain that cannot grow in the absence of the amino acid lysine (lys^-). Among the 400 unordered tetrads resulting from this mating, 233 were PD, 11 were NPD, and 156 were T.

a. What types of spores are in the PD, NPD, and T tetrads?

b. Are the *his* and *lys* genes linked? How do you know?

c. Using the simple equation RF = 100 × [NPD + (1/2)T]/total tetrads, calculate the distance in map units between the *his* and *lys* genes.

d. If you think about all the kinds of meiotic events that could occur (refer to Fig. 5.24), you can see that the calculation you did in part (c) may substantially underestimate RF. What kinds of meioses (NCO, SCO, or DCO) generated each of the tetrad types in this cross?

e. What incorrect assumptions does the simple RF equation you used in part (c) make about the meiotic events producing each type of tetrad? When could these assumptions nevertheless be correct?

f. Use your answers to part (d) to determine the number of NCO, SCO, and DCO meioses that generated the 400 tetrads.

g. Use your answers to part (f) to write a general equation that relates the number of DCO meioses to the number of the various tetrad types. Then write another general equation that computes the number of SCO meioses as a function of the number of the various tetrad types.

h. Based on your answer to part (f), calculate the average number of crossovers per meiosis (*m*) between *his* and *lys*.

i. Use your answer to (h) to write an equation for *m* in terms of NCO, SCO, and DCO meioses.

j. What is the relationship between RF and *m*?

k. Use your answer to part (j) to write a corrected equation for RF in terms of SCO, DCO, and NCO meioses.

l. Using your answer to part (g), rewrite the corrected RF equation from part (k) in terms of the numbers of the various tetrad types.

m. The equation you just wrote in part (l) is a corrected equation for RF that takes into account double crossovers that would otherwise have been missed. Use this improved formula to calculate a more accurate distance between the *his* and *lys* genes than the one you calculated in part (c).

45. a. In ordered tetrad analysis, what is the maximum RF that you could observe between a gene and the centromere? (*Hint:* What RF would you observe between a gene and the centromere if the gene were distant enough from the centromere so that at least one crossover, and often more than one

crossover, occurs between them in every meiosis?)

b. Can a gene and the centromere be unlinked? Explain.

c. Suppose in an ordered tetrad analysis you observe an RF of 30% between a gene and the centromere. Given your answer to part (a), do think that 30 m.u. is an accurate estimate of the distance between the gene and the centromere?

46. A research group has selected three independent Trp^- haploid strains of *Neurospora,* each of which cannot grow in the absence of the amino acid tryptophan. They first mated these three strains with a wild-type strain of opposite mating type, and then they analyzed the resultant octads. For all three matings, two of the four spore pairs in every octad could grow on minimal medium (that is, in the absence of tryptophan), while the other two spore pairs were unable to grow on this minimal medium.

a. What can you conclude from this result?

In the matings of mutant strains 1 and 2 with wild type, one of the two topmost pairs in some octads had spores that could grow on minimal medium while the other of the two topmost pairs in the same octads had spores that could not grow on minimal medium. In the mating of mutant strain 3 with wild type, either all the spores in the two topmost pairs could grow on minimal medium or all could not grow on minimal medium.

b. What can you conclude from this result?

The researchers next prepared two separate cultures of each mutant strain; one of these cultures was of mating type *A* and the other of mating type *a.* They mated these strains in pairwise fashion, dissected the resultant octads, and determined how many of the individual spores could grow on minimal medium. The results are shown here.

Mating	% of octads with *x* number of spores viable on minimal medium				
	$x = 0$	2	4	6	8
1 × 2	78	22	0	0	0
1 × 3	46	6	48	0	0
2 × 3	42	16	42	0	0

c. For each of the three matings in the table, how many of the 100 octads are PD? NPD? T?

d. Draw a genetic map explaining all of the preceding data. Assume that the sample sizes are sufficiently small that none of the octads are the result of double crossovers.

e. Although this problem describes crosses in *Neurospora,* it does not help in this case to present the matings in the table as ordered octads. Why not?

f. Why in this problem can you obtain gene ↔ centromere distances from the crosses in the table, even though the data are not presented as ordered octads?

Section 5.6

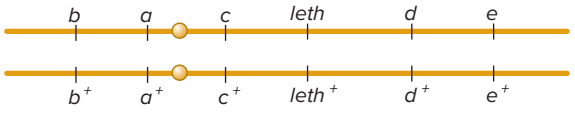
47. A single yeast cell placed on a solid agar will divide mitotically to produce a colony of about 10^7 cells. A haploid yeast cell that has a mutation in the *ade2* gene will produce a red colony; an $ade2^+$ colony will be white. Some of the colonies formed from diploid yeast cells with a genotype of $ade2^+/ ade2^-$ will contain sectors of red within a white colony.

a. How would you explain these sectors?

b. Although the white colonies are roughly the same size, the red sectors within some of the white colonies vary markedly in size. Why? Do you expect the majority of the red sectors to be relatively large or relatively small?

48. Figure 5.29 shows mitotic recombination leading to single spots or twin spots occurring in the G_2 stage of the cell cycle (after the chromosomes have replicated). However, because it usually is initiated by rare, random events of chromosome breakage, mitotic recombination can also take place in G_1, prior to S phase. Redraw Fig. 5.29 with mitotic recombination taking place in G_1 rather than G_2, and demonstrate why any such event could not yield single spots or twin spots.

49. A diploid strain of yeast has a wild-type phenotype but the following genotype:



a, b, c, d, and *e* all represent recessive alleles that yield a visible phenotype, and *leth* represents a recessive lethal mutation. All genes are on the same chromosome, and *a* is very tightly linked to the centromere (indicated by a small circle). Which of the following phenotypes could be found in sectors resulting from mitotic recombination in this cell? (1) *a;* (2) *b;* (3) *c;* (4) *d;* (5) *e;* (6) *b e;* (7) *c d;* (8) *c d e;* (9) *d e;* (10) *a b.* Assume that double mitotic crossovers are too rare to be observed.

50. In *Drosophila,* the *yellow* (*y*) gene is near the telomere of the long arm of the acrocentric X chromosome, while the *singed* (*sn*) gene is located near the middle of the same X chromosome arm. On the wings of female flies of genotype $y\ sn / y^+\ sn^+$, you can very rarely find patches of yellow tissue within which a small subset of cells also have singed bristles.

a. How can you explain this phenomenon?

b. Would you find similar patches on the wings of females having the genotype $y^+\ sn / y\ sn^+$?

51. Neurofibromas are tumors of the skin that can arise when a skin cell that is originally $NF1^+/NF1^-$ loses the $NF1^+$ allele. This wild-type allele encodes a functional protein (called a *tumor suppressor*), while the $NF1^-$ allele encodes a nonfunctional protein.

    A patient of genotype $NF1^+/NF1^-$ has 20 independent tumors in different areas of the skin. Samples are taken of normal, noncancerous cells from this patient, as well as of cells from each of the 20 tumors. Extracts of these samples are analyzed by a technique called gel electrophoresis that can detect variant forms of four different proteins (A, B, C, and D) all encoded by genes that lie on the same autosome as *NF1*. Each protein has a slow (S) and a fast (F) form that are encoded by different alleles (for example, A^S and A^F). In the extract of normal tissue, slow and fast variants of all four proteins are found. In the extracts of the tumors, 12 had only the fast variants of proteins A and D but both the fast and slow variants of proteins B and C; 6 had only the fast variant of protein A but both the fast and slow variants of proteins B, C, and D; and the remaining 2 tumor extracts had only the fast variant of protein A, only the slow variant of protein B, the fast and slow variants of protein C, and only the fast variant of protein D.

    a. What kind of genetic event described in this chapter could cause all 20 tumors, assuming that all the tumors are produced by the same mechanism?

    b. Draw a genetic map describing these data, assuming that this small sample represents all the types of tumors that could be formed by the same mechanism in this patient. Show which alleles of which genes lie on the two homologous chromosomes. Indicate all relative distances that can be estimated. Note that *NF1* is one of the genes you can map in this way.

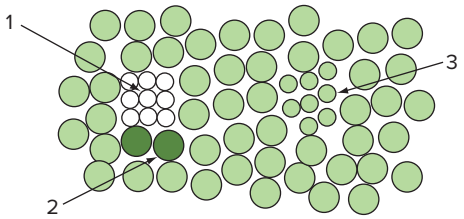
    c. Another mechanism that can lead to neurofibromas in this patient is a mitotic error producing cells with 45 rather than the normal 46 chromosomes. How can this mechanism cause tumors? How do you know, just from the results described, that none of these 20 tumors is formed by such mitotic errors?

    d. Can you think of any other type of error that could produce the results described?

52. Two important methods for understanding the genetic basis for development are mitotic crossing-over and the use of the gene from jellyfish called *GFP* (for green fluorescent protein) that makes these animals glow in the dark. By recombinant DNA techniques described later in the book, you can insert the jellyfish *GFP* gene anywhere into the genome of organisms like *Drosophila* or mice. Cells expressing this *GFP* gene will glow green in the microscope, while those without the *GFP* gene will not glow green.

    Mice homozygous for the recessive mutation *small cells* (*smc*) die as early embryos because their cells divide prematurely before they reach normal size.

    You want to design a mouse carrying one copy of the *GFP* gene and heterozygous for *smc* in which you could generate clones in adult mice by mitotic recombination. In this designer mouse, every cell in every clone that is not green would be homozygous for the *smc* mutation. The figure below shows a field of epithelial cells in the mouse you design. You will see some cells that are normal size and other cells that are small. You will also see cells of three different colors: blank, weakly glowing cells (*light green*), and brightly glowing cells (*dark green*). Most of the cells in the epithelium of this mouse are of normal size and weakly glowing. The epithelium also contains three clones of cells (1, 2, and 3) that have unusual appearances due to the occurrence of mitotic recombination.



    a. Show the chromosomes and centromeres, the alleles smc^+ and *smc,* and GFP^+ (*GFP* gene present) and GFP^- (*GFP* gene absent) in your designer mouse. (As a reminder, this mouse will carry one copy of the *GFP* gene and will be heterozygous for *smc*. Every cell in every clone generated by mitotic recombination that is not green should be homozygous for the *smc* mutation.)

    b. Why do you need to use mitotic recombination to study the function of smc^+ in adult mice?

    c. Why do you see cells of three different colors?

    d. Why are clones 1 and 2 next to each other?

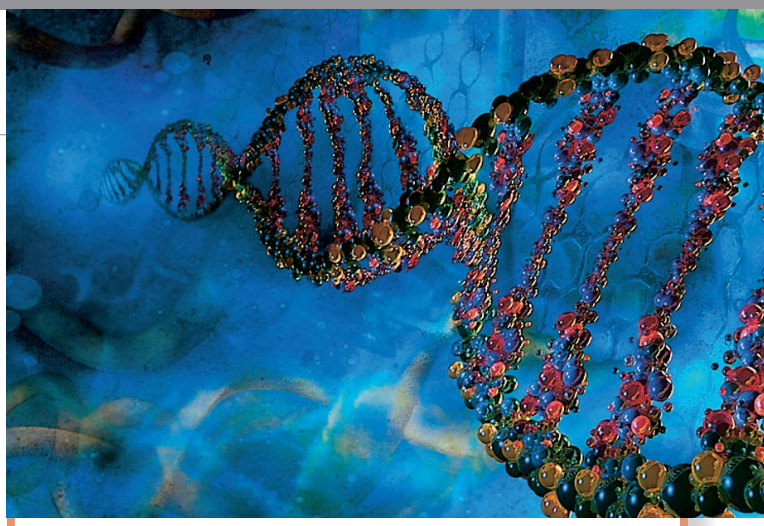
    e. On your map in part (a), place an arrow to show the position of a mitotic recombination event that could give rise to clones 1 and 2.

    f. Why do more cells exist in clone 1 than in clone 2?

    g. On your map in part (a), place an arrow to show the position of a mitotic crossover that could give rise to clone 3.

chapter 6

DNA Structure, Replication, and Recombination



The double-helical structure of DNA provides an explanation for the transmission of genetic information from generation to generation over billions of years.

chapter outline

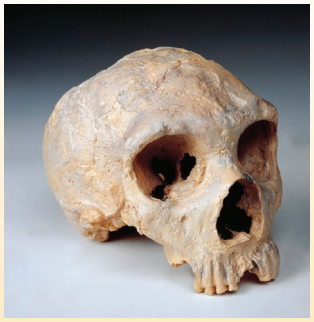
FOR NEARLY 4 BILLION YEARS, the double-stranded **DNA** molecule has served as the bearer of genetic information. It was present in the earliest single-celled organisms and in every other organism that has existed since. Over that long period of time, the *hardware*—the structure of the molecule itself—has not changed. In contrast, evolution has honed and vastly expanded the *software*—the programs of genetic information that the molecule stores, expresses, and transmits from one generation to the next.

Under special conditions of little or no oxygen, DNA can withstand a wide range of temperature, pressure, and humidity and remain relatively intact for hundreds, thousands, even tens of thousands of years. Molecular sleuths have retrieved the evidence: 38,000-year-old DNA from a Neanderthal skeleton (**Fig. 6.1**). Amazingly, this ancient DNA still carries readable sequences—shards of decipherable information that act as time machines for the viewing of genes in this long-vanished species. Comparisons with homologous DNA segments from living people make it possible to identify the precise mutations that have fueled evolution.

For example, comparisons of Neanderthal and human DNA have helped anthropologists settle a long-running debate about the genetic relationship of the two. The evidence shows that Neanderthals and our own species, *Homo sapiens,* last shared a common ancestor between 600,000 and 800,000 years ago. Neanderthal ancestors migrated to Europe about 400,000 years ago while our own ancestors remained in Africa. The two groups remained out of contact until 40,000 years ago, when *Homo sapiens* first arrived in Europe. Within a few millennia, the Neanderthals were extinct. However, their recently recovered DNA suggests that during the 10,000 years that Neanderthals shared Europe with *Homo sapiens*, some interbreeding took place; 1–4% of the genomes of modern non-Africans can be traced to Neanderthals.

Francis Crick, co-discoverer of DNA's double-helical structure and a leading twentieth-century theoretician of molecular biology, wrote that "almost all aspects of life are engineered at the molecular level, and without understanding molecules,

Figure 6.1 Ancient DNA still carries information. Molecular biologists have successfully extracted and determined the sequence of DNA from a 38,000-year-old Neanderthal skull. These findings attest to the chemical stability of DNA, the molecule of inheritance.



we can only have a very sketchy understanding of life itself." For this reason, we shift our perspective in this chapter to an examination of DNA, the molecule of which genes are made.

As we extend our analysis to the molecular level, bear in mind two important themes. First, DNA's genetic functions flow directly from its molecular structure—the way its atoms are arranged in space. Second, all of DNA's genetic functions depend on specialized proteins that interact with it and *read* the information it carries, because DNA itself is chemically inert. In fact, DNA's lack of chemical reactivity makes it an ideal physical container for long-term maintenance of genetic information in living organisms, as well as their nonliving remains.

6.1 Experimental Evidence for DNA as the Genetic Material

learning objectives

1. Describe the chemical components of DNA.
2. Summarize the methods that located DNA in chromosomes.
3. Explain how Avery and his colleagues demonstrated bacterial transformation, and explain the significance of this finding.
4. Describe the blender experiments of Hershey and Chase and what the results revealed about DNA's function.

At the beginning of the twentieth century, geneticists did not know that DNA was the genetic material. It took a cohesive pattern of results from experiments performed over more than 50 years to convince the scientific community that DNA is the molecule of heredity. We now present key pieces of the evidence.

Chemical Studies Locate DNA in Chromosomes

In 1869, Friedrich Miescher extracted a weakly acidic, phosphorus-rich material from the nuclei of human white blood cells and named it *nuclein.* It was unlike any chemical compound reported previously. Nuclein's major component turned out to be DNA, although it also contained some contaminants. The full chemical name of DNA is **deoxyribonucleic acid,** reflecting three characteristics of the substance: One of its constituents is a sugar known as deoxyribose; it is found mainly in cell nuclei; and it is acidic.

After purifying DNA from the nuclein by chemical means, researchers established that contains only four distinct chemical building blocks linked in a long chain (**Fig. 6.2**). The four individual components belong to a class of compounds known as **nucleotides;** the bonds joining one nucleotide to another are covalent **phosphodiester bonds;** and the linked chain of building block subunits is a type of **polymer.**

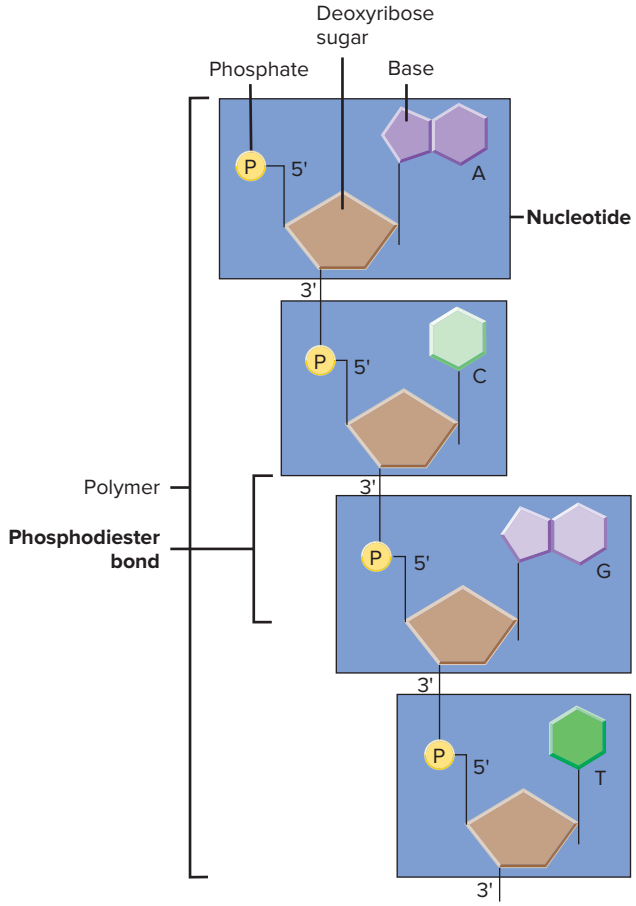
A procedure first reported in 1923 made it possible to discover where in the cell DNA resides. Named the *Feulgen reaction* after its designer, the procedure relies on a chemical which stains DNA red. In a preparation of stained cells, the chromosomes redden, while other areas of the cell remain relatively colorless. The reaction shows that DNA is localized almost exclusively within chromosomes.

The finding that DNA is a component of chromosomes does not itself prove that the molecule has anything to do with genes. Typical eukaryotic chromosomes also contain an even greater amount of protein by weight. Because proteins are built of 20 different amino acids, whereas DNA is made of only four different nucleotides, many researchers thought proteins had greater potential for diversity and were better suited to serve as the genetic material. These same scientists assumed that even though DNA was an important part of chromosome structure, it was too simple to contain the complexity of genes.

Bacterial Transformation Implicates DNA as the Genetic Material

Several studies eventually promoted the idea that DNA would be the chemical substance that carries genetic information. The most important of these used single-celled bacteria as experimental organisms. Bacteria carry their genetic material in a single circular chromosome that lies within the cell without being enclosed in a nuclear membrane. With

Figure 6.2 The chemical composition of DNA. A single strand of a DNA molecule consists of a chain of *nucleotide subunits* (*blue boxes*). Each nucleotide is made of the sugar **deoxyribose** (*tan pentagons*) connected to an inorganic **phosphate group** (*yellow circles*) and to one of four **nitrogenous bases** (*purple* or *green polygons*). The *phosphodiester bonds* that link the nucleotide subunits to each other attach the phosphate group of one nucleotide to the deoxyribose sugar of the preceding nucleotide.

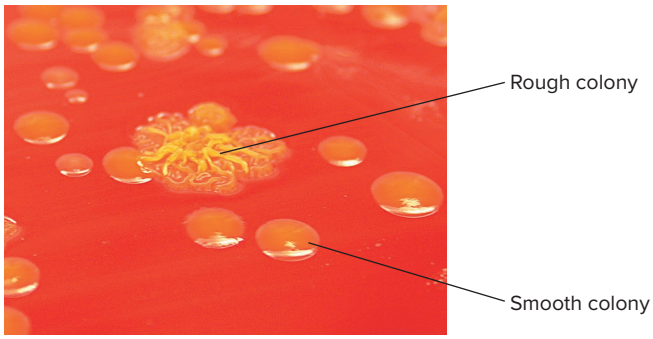


only one chromosome, bacteria do not undergo meiosis to produce germ cells, and they do not apportion their replicated chromosomes to daughter cells by mitosis; rather, they divide by a process known as *binary fission*. In spite of these obvious differences, at least some investigators in the first half of the twentieth century thought that the genetic material of bacteria might be the same as that found in eukaryotic organisms.

One prerequisite of genetic studies in bacteria, as with any species, is the detection of alternative forms of a trait among individuals in a population. In a 1923 study of *Streptococcus pneumoniae* bacteria grown in laboratory media, Frederick Griffith distinguished two bacterial forms: smooth (S) and rough (R). S is the wild type; a mutation in S gives rise to R. From observation and biochemical analysis, Griffith determined that S forms appear

Figure 6.3 Smooth (S) and rough (R) colonies of *S. pneumoniae*.

From: Arnold et al., "New associations with *Pseudomonas luteola* bacteremia: A veteran with a history of tick bites and a trauma patient with pneumonia," *The Internet Journal of Infectious Diseases,* 2005, 4(2): 1-5, Fig. 1 © & Courtesy of Dr. Forest Arnold, University of Louisville. Used with permission.



smooth because they synthesize a polysaccharide capsule that surrounds pairs of cells. R forms, which arise spontaneously as mutants of S, cannot make the capsular polysaccharide, and as a result, their colonies appear to have a rough surface (**Fig. 6.3**). We now know that the R form lacks an enzyme necessary for synthesis of the capsular polysaccharide. Because the polysaccharide capsule helps protect the bacteria from an animal's immune response, the S bacteria are virulent and kill most laboratory animals exposed to them (**Fig. 6.4a**); by contrast, the R forms fail to cause infection (**Fig. 6.4b**). In humans, the virulent S forms of *S. pneumoniae* can cause pneumonia.

The phenomenon of transformation

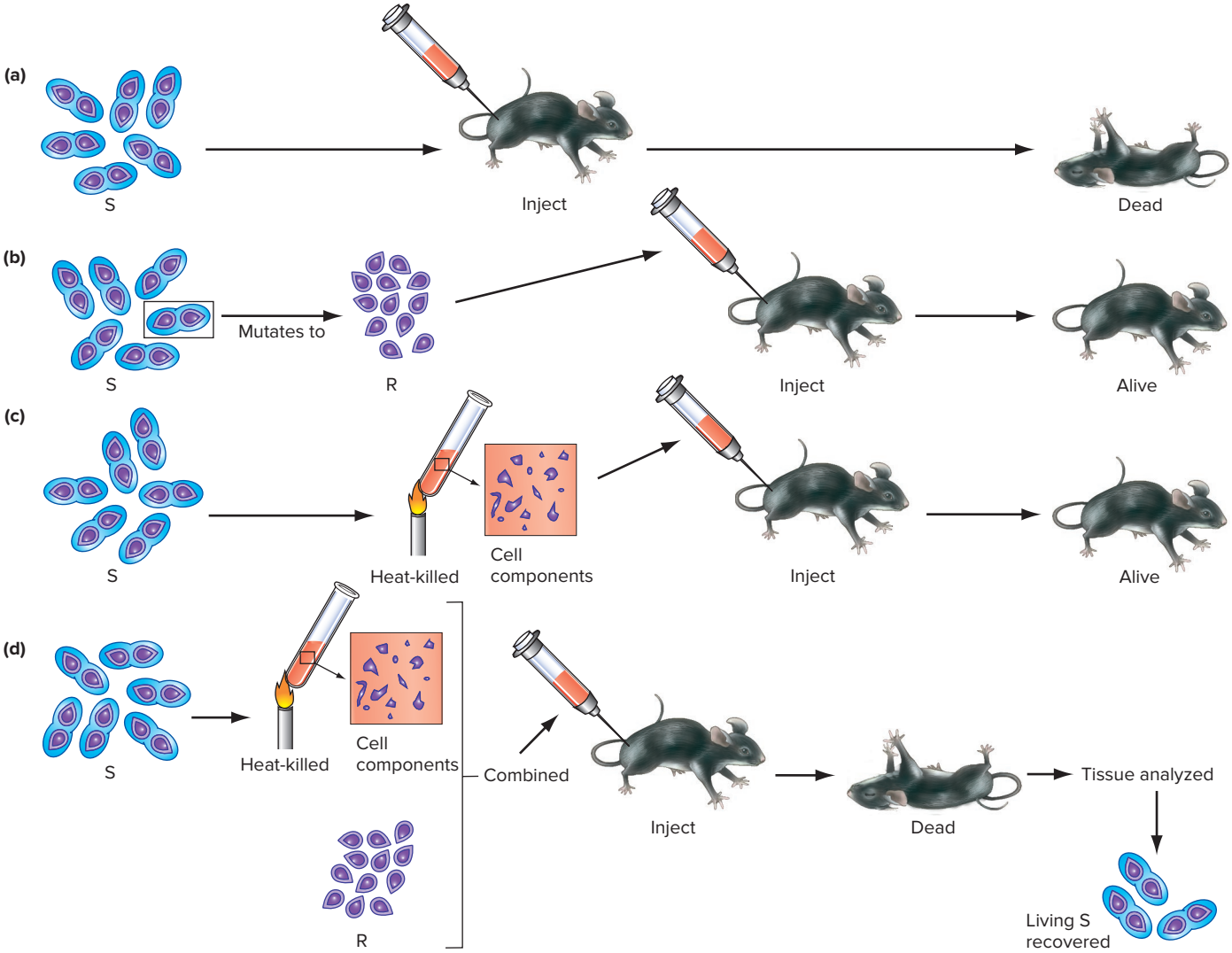
In 1928, Griffith published the astonishing finding that genetic information from dead bacterial cells could somehow be transmitted to live cells. He was working with two types of the *S. pneumoniae* bacteria—live R forms and heat-killed S forms. Neither the heat-killed S forms nor the live R forms produced infection when injected into laboratory mice (Fig. 6.4b and **Fig. 6.4c**), but a mixture of the two killed the animals (**Fig. 6.4d**). Furthermore, bacteria recovered from the blood of the dead animals were living S forms (Fig. 6.4d).

The ability of a substance to change the genetic characteristics of an organism is known as **transformation.** Something from the heat-killed S bacteria must have transformed the living R bacteria into S. This transformation was permanent and most likely genetic, because all future generations of the bacteria grown in culture were the S form.

DNA as the active agent of transformation

By 1929, two other laboratories had repeated these results, and in 1931, investigators in Oswald T. Avery's laboratory found they could achieve transformation without using any animals at all, simply by growing R-form bacteria in

Figure 6.4 Transformation. **(a)** S bacteria are virulent and can cause lethal infections when injected into mice. **(b)** Injections of R mutants by themselves do not cause infections that kill mice. **(c)** Similarly, injections of heat-killed S bacteria do not cause lethal infections. **(d)** Lethal infection does result, however, from injections of live R bacteria mixed with heat-killed S strains; the blood of the dead host mouse contains living S-type bacteria.

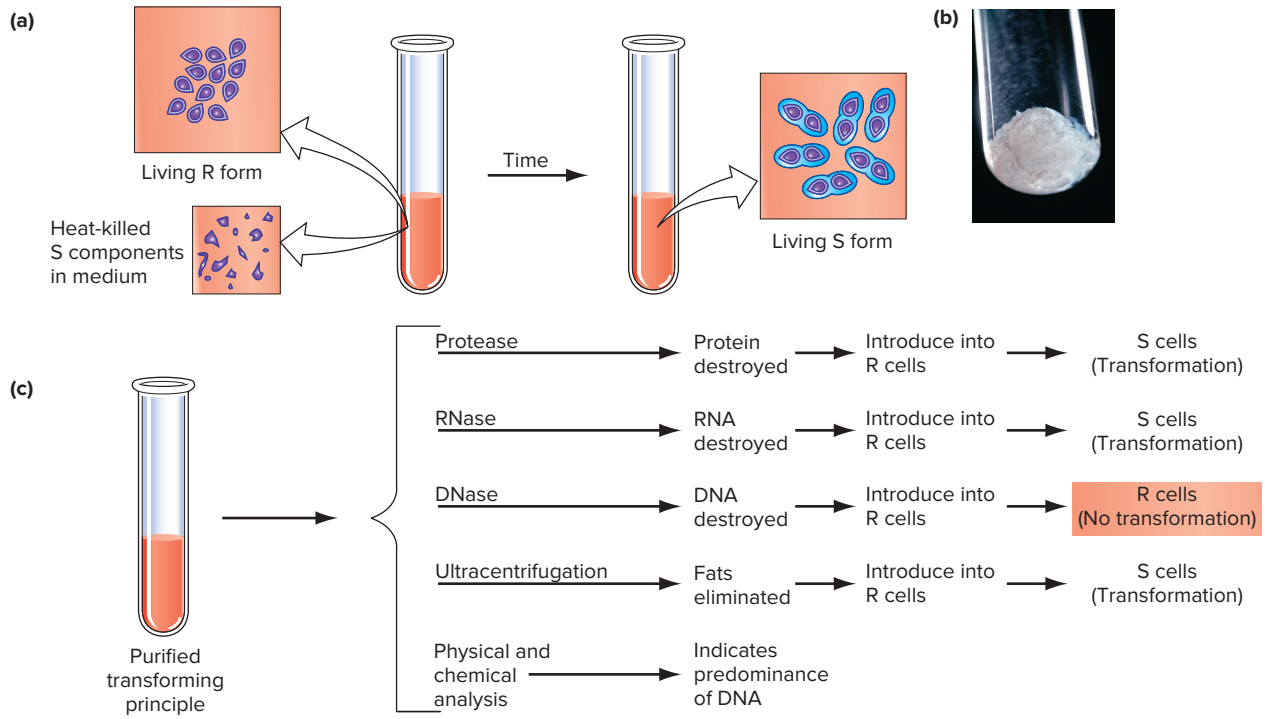


medium in the presence of components from dead S forms (**Fig. 6.5a**). Avery then embarked on a quest that would remain the focus of his work for almost 15 years: "Try to find in that complex mixture, the active principle!" In other words, try to identify the heritable substance in the bacterial extract that induces the transformation of harmless R bacteria into pathogenic S bacteria. Avery dubbed the substance he was searching for the *transforming principle* and spent many years trying to purify it sufficiently to be able to identify it unambiguously. He and his coworkers eventually prepared a tangible, active transforming principle. In the final part of their procedure, a long, whitish wisp materialized from ice-cold alcohol solution and wound around the glass stirring rod to form a fibrous wad of nearly pure transforming principle (**Fig. 6.5b**).

Once purified, the transforming principle had to be characterized. In 1944, Avery and two coworkers, Colin MacLeod and Maclyn McCarty, published the cumulative findings of experiments designed to determine the transforming principle's chemical composition (**Fig. 6.5c**). In these experiments, the purified transforming principle was active at the extraordinarily high dilution of 1 part in 600 million. Although the preparation was almost pure DNA, the investigators nevertheless exposed it to various enzymes to see if some molecule other than DNA might have caused the transformation. Enzymes that degraded RNA, protein, or polysaccharide had no effect on the transforming principle, but an enzyme that degrades DNA completely destroyed its activity. The tentative published conclusion was that the transforming principle appeared to be DNA. In a personal

Figure 6.5 The transforming principle is DNA. **(a)** Bacterial transformation occurs in culture medium containing the remnants of heat-killed S bacteria. Some *transforming principle* from the heat-killed S bacteria is taken up by the live R bacteria, converting (transforming) them into virulent S strains. **(b)** Purified DNA extracted from human white blood cells. **(c)** Chemical fractionation of the transforming principle. Treatment of purified DNA with a DNA-degrading enzyme destroys its ability to cause bacterial transformation, while treatment with enzymes that destroy other kinds of macromolecules has no effect on the transforming principle.

b: © Phanie/Science Source



letter to his brother, Avery went one step further and confided that the transforming principle "may be a gene."

Despite the paper's abundance of concrete evidence, many within the scientific community still resisted the idea that DNA is the molecule of heredity. They argued that perhaps Avery's results reflected the activity of contaminants; or perhaps genetic transformation was not happening at all, and instead, the purified material somehow triggered a physiological switch that transformed bacterial phenotypes. Unconvinced for the moment, these scientists remained attached to the idea that proteins were the prime candidates for the genetic material.

DNA, Not Protein, Contains the Instructions for Virus Propagation

Not everyone shared this skepticism. Alfred Hershey and Martha Chase anticipated that they could assess the relative importance of DNA and protein in gene replication by infecting bacterial cells with viruses called **phages,** short for **bacteriophages** (literally *bacteria eaters*).

Viruses are the simplest of organisms. By structure and function, they fall somewhere between living cells capable of reproducing themselves and macromolecules such as proteins. Because viruses hijack the molecular machinery of their host cell to carry out growth and replication, they can be very small indeed and contain very few genes. For many kinds of phages, each particle consists of roughly equal weights of protein and DNA (**Fig. 6.6a**). These phage particles can reproduce themselves only after infecting a bacterial cell. Thirty minutes after infection, the cell bursts and hundreds of newly made phages spill out (**Fig. 6.6b**). The question is: What substance contains the information used to produce the new phage particles—DNA or protein?

With the invention of the electron microscope in 1939, it became possible to see individual phages, and surprisingly, electron micrographs revealed that the entire phage does not enter the bacterium it infects. Instead, a viral shell—called a *ghost*—remains attached to the outer surface of the bacterial cell wall. Because the empty phage coat remains outside the bacterial cell, one investigator likened phage particles to tiny syringes that bind to the cell surface and inject the material containing the information needed for viral replication into the host cell.

In their famous *Waring blender* experiment of 1952, Alfred Hershey and Martha Chase tested the idea that the ghost left on the cell wall is composed of protein, while the injected material consists of DNA (**Fig. 6.7**). A type of phage known as T2 served as their experimental system. Hershey and Chase grew two separate sets of T2 in bacteria

Figure 6.6 Experiments with viruses provide convincing evidence that genes are made of DNA. (a) and **(b)** Bacteriophage T2 structure and life cycle. The phage particle consists of DNA contained within a protein coat. The virus attaches to the bacterial host cell and injects its genes (the DNA) through the bacterial cell wall into the host cell cytoplasm. Inside the host cell, these genes direct the formation of new phage DNA and proteins, which assemble into progeny phages that are released into the environment when the cell bursts.

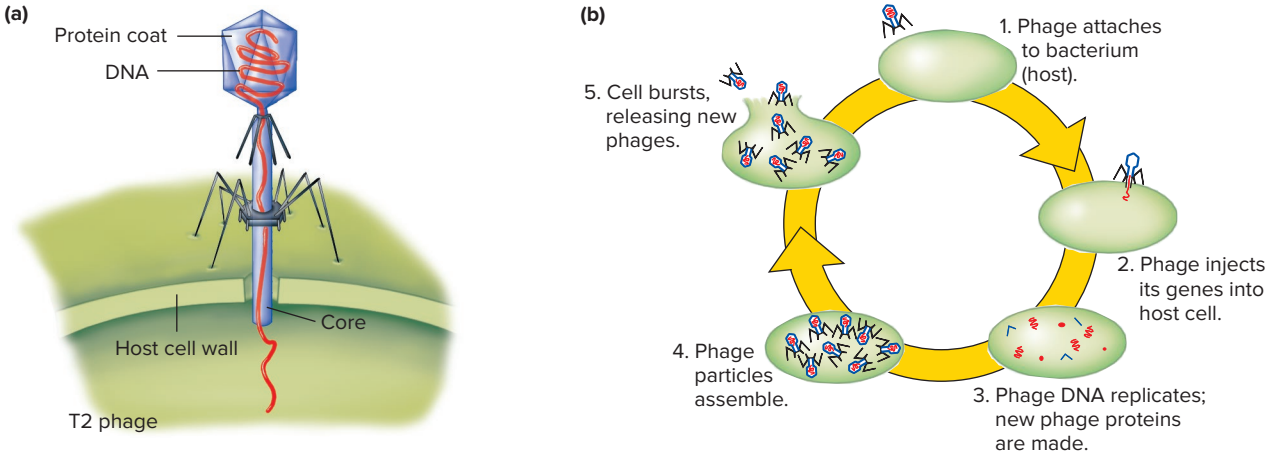
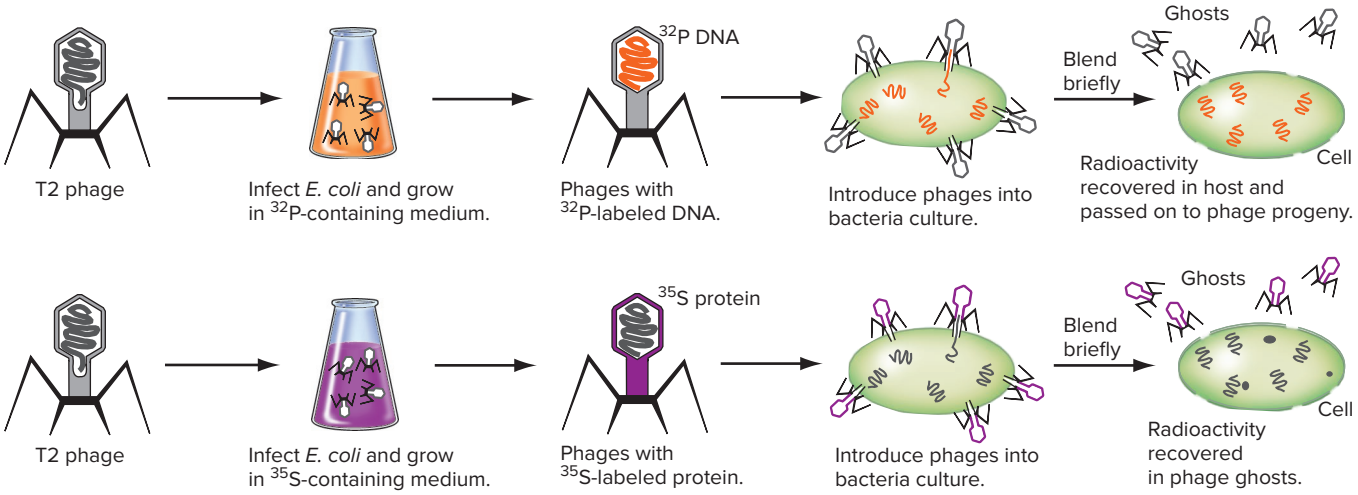


Figure 6.7 The Hershey-Chase Waring blender experiment. T2 bacteriophage particles either with ^{32}P-labeled DNA (*orange*) or with ^{35}S-labeled proteins (*purple*) were used to infect bacterial cells. After a short incubation, Hershey and Chase shook the cultures in a Waring blender and spun the samples in a centrifuge to separate the empty viral ghosts from the heavier infected cells. Most of the ^{35}S-labeled proteins remained with the ghosts, while most of the ^{32}P-labeled T2 DNA was found in the sedimented infected cells.



maintained in two different culture media, one infused with radioactively labeled phosphorus (^{32}P), the other with radioactively labeled sulfur (^{35}S). Because proteins incorporate sulfur but no phosphorus and DNA contains phosphorus but no sulfur, phages grown on ^{35}S would have radioactively labeled protein while particles grown on ^{32}P would have radioactive DNA. The radioactive tags would serve as markers for the location of each material when the phages infected fresh cultures of bacterial cells.

After exposing one fresh culture of bacteria to ^{32}P-labeled phages and another culture to ^{35}S-labeled phages, Hershey and Chase used a Waring blender to disrupt each one, effectively separating the viral ghosts from the bacteria harboring the viral genes. Centrifugation of the cultures then separated the heavier infected cells, which ended up in a pellet at the bottom of the tube, away from the lighter phage ghosts, which remained suspended in the supernatant solution. Most of the radioactive ^{32}P (in DNA) went to the pellet, while most of the radioactive ^{35}S (in protein) remained in the supernatant. This result confirmed that the extracellular ghosts were indeed mostly protein, while the injected viral material specifying production of more phages was mostly DNA. Bacteria containing the radio-labeled phage DNA behaved just

as in a normal phage infection, producing and disgorging hundreds of progeny particles. From these observations, Hershey and Chase concluded that phage genes are made of DNA.

The Hershey-Chase experiment, although less rigorous than the Avery project, had an enormous impact. In the minds of many investigators, it confirmed Avery's results and extended them to viral particles. The spotlight was now on DNA.

essential concepts

- *DNA* is a polymer of *nucleotides* joined by *phosphodiester bonds*. Nucleotides are made of deoxyribose, phosphate, and one of four nitrogenous bases.
- DNA is localized almost exclusively in the chromosomes within the nucleus of a cell.
- Avery and his colleagues showed that a purified DNA preparation from S (virulent) bacteria could *transform* R (nonvirulent) bacteria into the S form; this result was strong evidence for DNA as the genetic material.
- Hershey and Chase grew T2 bacteriophages in the presence of either ^{35}S (labels proteins) or ^{32}P (labels DNA). They showed that it is the ^{32}P-tagged viral DNA that contains the genetic instructions to produce more virus particles.

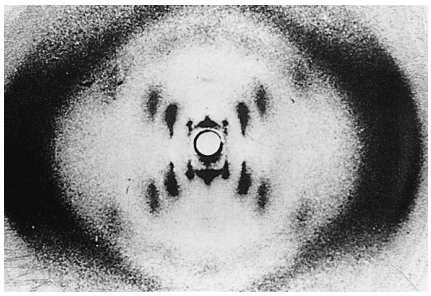
6.2 The Watson and Crick Double Helix Model of DNA

learning objectives

1. Describe the key features of the Watson-Crick model for DNA structure.
2. Explain what is meant by the *antiparallel polarity* of the two strands of DNA within the double helix.
3. Distinguish the different structural forms of DNA from one another.

Under appropriate experimental conditions, purified molecules of DNA can align alongside each other in fibers to produce an ordered structure. And just as a crystal chandelier scatters light to produce a distinctive pattern on the wall, DNA fibers scatter X-rays to produce a characteristic *diffraction pattern* (**Fig. 6.8**). A knowledgeable X-ray crystallographer can interpret DNA's diffraction pattern to deduce certain aspects of the molecule's three-dimensional structure. When in the spring of 1951 the 23-year-old James Watson learned that DNA could project a diffraction pattern, he realized that it "must have a regular structure that could be solved in a straightforward fashion."

Figure 6.8 X-ray diffraction patterns reflect the helical structure of DNA. Photograph of an X-ray diffraction pattern produced by oriented DNA fibers, taken by Rosalind Franklin and Maurice Wilkins in late 1952. The crosswise pattern of X-ray reflections indicates that DNA is helical.



In this section, we analyze DNA's three-dimensional structure, looking first at significant details of the nucleotide building blocks, then at how those subunits are linked together in a polynucleotide chain, and finally, at how two chains associate to form a double helix.

Nucleotides Are the Building Blocks of DNA

DNA is a long polymer composed of subunits known as **nucleotides.** Each nucleotide consists of a *deoxyribose* sugar, a *phosphate*, and one of four *nitrogenous bases*. Detailed knowledge of these chemical constituents and the way they combine played an important role in Watson and Crick's model building.

The components of a nucleotide

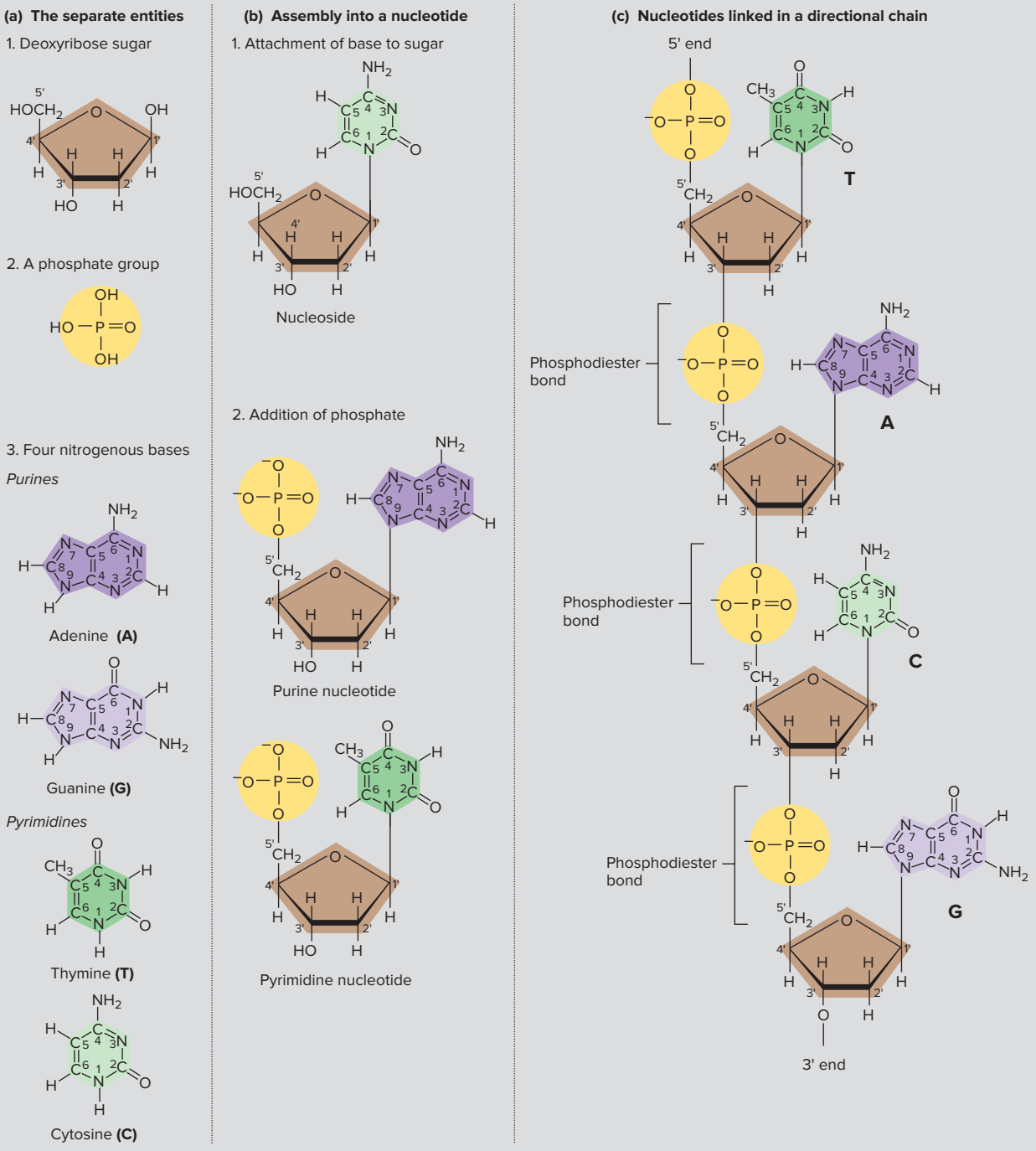
Figure 6.9 depicts the chemical composition and structure of deoxyribose, phosphate, and the four nitrogenous bases; how these components come together to form a nucleotide; and how *phosphodiester bonds* link the nucleotides in a DNA chain. Each individual carbon or nitrogen atom in the central ring structure of a nitrogenous base is assigned a number: 1–9 for **purines**, and 1–6 for **pyrimidines.** The carbon atoms of the deoxyribose sugar are distinguished from atoms within the nucleotide base by the use of primed numbers from 1′ to 5′. Covalent attachment of a base to the 1′ carbon of deoxyribose forms a *nucleoside*. The addition of a phosphate group to the 5′ carbon forms a complete *nucleotide*.

Connecting nucleotides to form a DNA chain

As Fig. 6.9 shows, a DNA chain composed of many nucleotides has **polarity:** an overall direction. Phosphodiester

FEATURE FIGURE 6.9

A Detailed Look at DNA's Chemical Constituents



bonds always form a covalent link between the 3′ carbon of one nucleotide and the 5′ carbon of the following nucleotide. The consistent orientation of the nucleotide building blocks gives a chain overall direction, such that the two ends of a single chain are chemically distinct.

At the **5′ end,** the sugar of the terminal nucleotide has a free 5′ carbon atom, free in the sense that it is not linked to another nucleotide. Depending on how the DNA is synthesized or isolated, the 5′ carbon of the nucleotide at the 5′ end may carry either a hydroxyl or a phosphate group. At the

other end of the chain—the **3′ end**—it is the 3′ carbon of the final nucleotide that is free. Along the chain between the two ends, this 5′-to-3′ polarity is conserved from nucleotide to nucleotide. By convention, a DNA chain is described in terms of its bases, written with the 5′-to-3′ direction going from left to right (unless otherwise noted). The chain depicted in Fig. 6.9c, for instance, would be 5′ TACG 3′.

DNA's information content

Information can be encoded only in a sequence of symbols whose order varies according to the message to be conveyed. Without this sequence variation, there is no potential for carrying information. Because DNA's *backbone* of alternating sugar and phosphate is chemically identical for every nucleotide in a DNA chain, the only difference between nucleotides is in the identity of the nitrogenous base. Thus, the genetic information in DNA must consist of variations in the sequence of the A, G, T, and C bases. The information constructed from the four-letter language of DNA bases is analogous to the information built from the 26-letter alphabet of English or French or Italian. Just as you can combine the 26 letters of the alphabet in different ways to generate the words of a book, so, too, different combinations of the four bases in very long sequences of nucleotides can encode the information for constructing an organism.

The DNA Helix Consists of Two Antiparallel Chains

Watson and Crick's discovery of the structure of the DNA molecule ranks with Darwin's theory of evolution by natural selection and Mendel's laws of inheritance in its contribution to our understanding of biological phenomena. The Watson-Crick structure, first embodied in a model that superficially resembled the Tinker Toys of preschool children, was based on an interpretation of all the chemical and physical data available at the time. Watson and Crick published their findings in the scientific journal *Nature* in April 1953.

Evidence from X-ray diffraction

The diffraction patterns of oriented DNA fibers do not, on their own, contain sufficient information to reveal structure. For instance, the number of diffraction spots, whose intensities and positions constitute the X-ray data (review Fig. 6.8), is considerably lower than the number of unknown coordinates of all the atoms in an oriented DNA molecule. Nevertheless, the photographs do reveal a wealth of structural information to the trained eye. Excellent X-ray images produced by Rosalind Franklin and Maurice Wilkins showed that the molecule is spiral-shaped, or helical; the spacing between repeating units along the axis of the helix is 3.4 Å (3.4×10^{-10} meters); the helix undergoes one complete turn every 34 Å; and the diameter of the molecule is 20 Å. This diameter is roughly twice the width of a single nucleotide, suggesting that a DNA molecule might be composed of two side-by-side DNA chains.

Complementary base pairing

If a DNA molecule contains two side-by-side chains of nucleotides, what forces hold these chains together? Erwin Chargaff provided an important clue with his data on the nucleotide composition of DNA from various species. Despite large variations in the relative amounts of the bases, the ratio of A to T is not significantly different from 1:1, and the ratio of G to C is also the same in every organism (**Table 6.1**). Watson grasped that the roughly 1:1 ratios of A to T and of G to C reflect a significant aspect of the molecule's inherent structure.

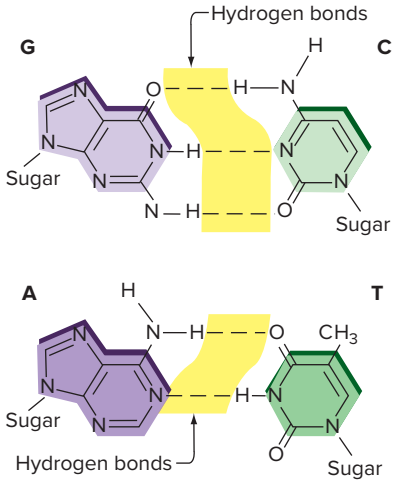
To explain Chargaff's ratios in terms of chemical affinities between A and T and between G and C, Watson made cardboard cutouts of the bases in the chemical forms they assume in a normal cellular environment. He then tried to match these up in various combinations, like pieces in a jigsaw puzzle. He knew that the particular arrangements of

TABLE 6.1 Chargaff's Data on Nucleotide Base Composition in the DNA of Various Organisms

	Percentage of Base in DNA				Ratios	
Organism	**A**	**T**	**G**	**C**	**A:T**	**G:C**
Escherichia coli	24.7	23.6	26.0	25.7	1.05	1.01
Saccharomyces cerevisiae	31.3	32.9	18.7	17.1	0.95	1.09
Caenorhabditis elegans	31.2	29.1	19.3	20.5	1.07	0.96
Drosophila melanogaster	27.3	27.6	22.5	22.5	0.99	1.00
Mus musculus	29.2	29.4	21.7	19.7	0.99	1.10
Homo sapiens	29.3	30.0	20.7	20.0	0.98	1.04

Note that even though the level of any one nucleotide is different in different organisms, the amount of A always approximately equals the amount of T, and the level of G is always similar to that of C. Moreover, the total amount of purines (A plus G) nearly always equals the total amount of pyrimidines (C plus T).

Figure 6.10 Complementary base pairing. An A on one strand can form two (noncovalent) hydrogen bonds with a T on the other strand. G on one strand can form three hydrogen bonds with a C on the other strand. The size and shape of A–T and of G–C base pairs are similar, allowing both to fill the same amount of space between the two backbones of the double helix.



atoms on purines and pyrimidines play a crucial role in molecular interactions as they can participate in the formation of **hydrogen bonds:** weak electrostatic bonds that result in a partial sharing of hydrogen atoms between reacting groups (**Fig. 6.10**). Watson saw that A and T could be paired together such that two hydrogen bonds formed between them. If G and C were similarly paired, hydrogen bonds could also easily connect the nucleotides carrying these two bases. (Watson originally posited two hydrogen bonds between G and C, but there are actually three.) Remarkably, the two pairs—A–T and G–C—had essentially the same shape. This meant that the two pairs could fit in any order between two sugar-phosphate backbones without distorting the structure. This **complementary base pairing** also explained the Chargaff ratios—always equal amounts of A and T, and of G and C. Note that both of these base pairs consist of one purine and one pyrimidine.

Crick connected the chemical facts with the X-ray data, recognizing that because of the geometry of the base-sugar bonds in nucleotides, the orientation of the bases in Watson's pairing scheme could arise only if the bases were attached to backbones running in opposite directions. **Figure 6.11** illustrates and explains the model Watson and Crick proposed in April 1953: DNA as a double helix.

The Double Helix May Assume Alternative Forms

Watson and Crick arrived at the double helix model of DNA structure by building models, not by a direct structural determination from the data alone. And even though Watson has written that "a structure this pretty just had to exist," the beauty of the structure is not necessarily evidence of its correctness. At the time of its presentation, the strongest evidence for its correctness was its physical plausibility, its chemical and spatial compatibility with all available data, and its capacity for explaining many biological phenomena.

B DNA and Z DNA

The majority of naturally occurring DNA molecules have the configuration suggested by Watson and Crick. Such molecules are known as **B-form DNA;** they spiral to the right (**Fig. 6.12a**). DNA structure is, however, more polymorphic than originally assumed. One type, for example, contains nucleotide sequences that cause the DNA to assume what is known as a **Z form** in which the helix spirals to the left and the backbone takes on a zigzag shape (**Fig. 6.12b**). Researchers have observed many kinds of unusual non-B structures *in vitro* (in the test tube, literally *in glass*), and they speculate that some of these might occur at least transiently in living cells. There is some evidence, for instance, that Z DNA might exist in certain chromosomal regions *in vivo* (in the living organism). Whether the Z form and other unusual conformations have any biological role remains to be determined.

Linear and circular DNA

The nuclear chromosomes of all eukaryotic organisms are long, linear double helixes, but some smaller chromosomes are circular (**Fig. 6.13a** and **b**). These include the chromosomes of prokaryotic bacteria, the chromosomes of organelles such as the mitochondria and chloroplasts that are found inside eukaryotic cells, and the chromosomes of some viruses, including the papovaviruses that can cause cancers in animals and humans. Such circular chromosomes consist of covalently closed, double-stranded circular DNA molecules. Although neither strand of these circular double helixes has an end, the two strands are still antiparallel in polarity.

Single-stranded and double-stranded DNA

In some viruses, the genetic material consists of relatively small, single-stranded DNA molecules. Once inside a cell, the single strand serves as a *template* (pattern) for making a second strand, and the resulting double-stranded DNA then governs the production of more virus particles. Examples of viruses carrying single-stranded DNA are bacteriophages ϕX174 and M13, and mammalian parvoviruses, which are associated with fetal death and spontaneous abortion in humans. In both ϕX174 and M13, the single DNA strand is in the form of a covalently closed circle; in the parvoviruses, it is linear (**Fig. 6.13c** and **d**).

Alternative B and Z configurations; circularization of the molecule; and single strands that are converted to double

FEATURE FIGURE 6.11

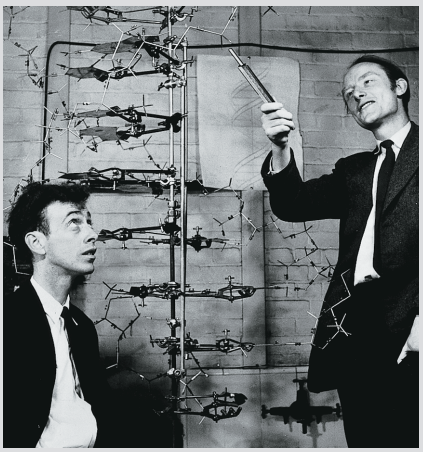
The Double Helix Structure of DNA

(a) Watson and Crick took the known facts about DNA's chemical composition and its physical arrangement in space and constructed a wire-frame model that could explain the molecule's function.

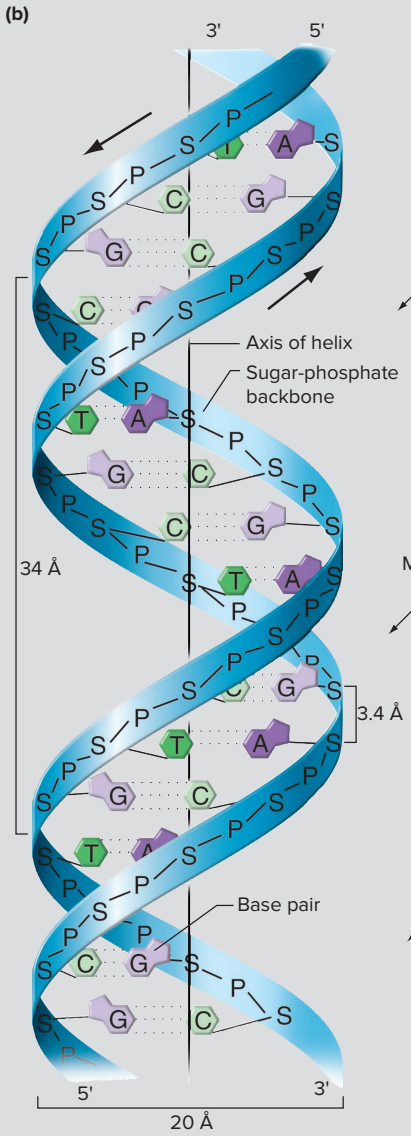
(b) In the model, two DNA chains spiral around an axis with the sugar-phosphate backbones on the outside and flat pairs of bases meeting in the middle. One chain runs 5′ to 3′ upward, while the other runs in the opposite direction of 5′ to 3′ downward. In short, *the two chains are antiparallel.* The two chains wrap around each other once every 10 base pairs, or once every 34 Å. The result is a double helix that looks like a twisted ladder with the two spiraling structural members composed of sugar-phosphate backbones and the perpendicular rungs consisting of base pairs.

(c) In a space-filling representation of the model, the overall shape is that of a grooved cylinder with a diameter of 20 Å. The backbones spiral around the axis of the double helix like threads on a screw. Because two backbones exist, there are two threads, and these two threads are vertically displaced from each other. This displacement of the backbones generates two grooves, one (the **major groove)** much wider than the other (the **minor groove)**.

The two chains of the double helix are held together by hydrogen bonds between complementary base pairs, A–T and G–C. The spatial requirements of the double helix require that each base pair must consist of one small pyrimidine and one large purine, and even then, only for the particular pairings of A–T and G–C. In contrast, A–C and G–T pairs do not fit well and cannot easily form hydrogen bonds. Although any one nucleotide pair forms only two or three hydrogen bonds, the sum of these connections between successive base pairs in a long DNA molecule composed of thousands of nucleotides is a key to the molecule's great chemical stability.



(a)



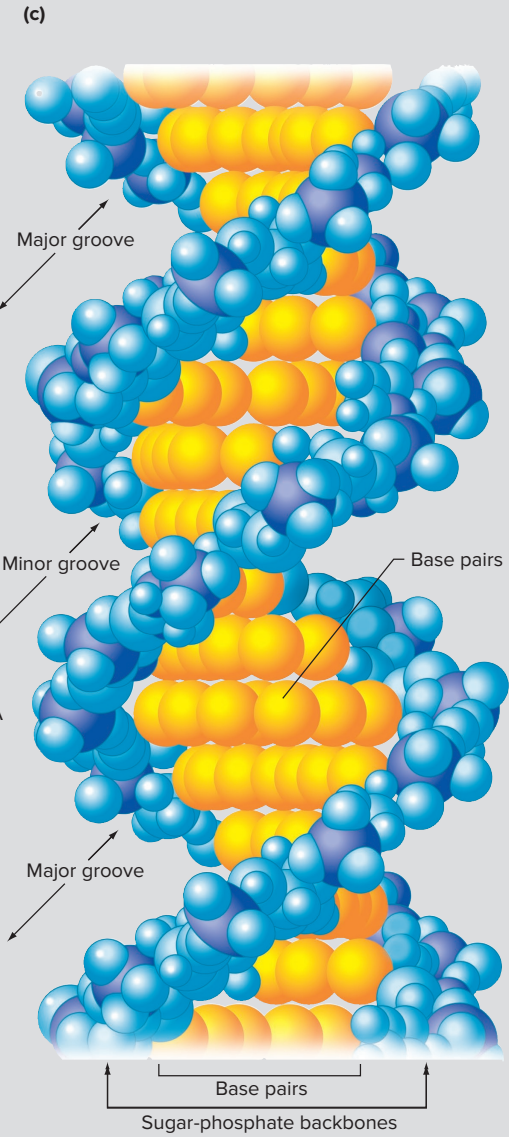
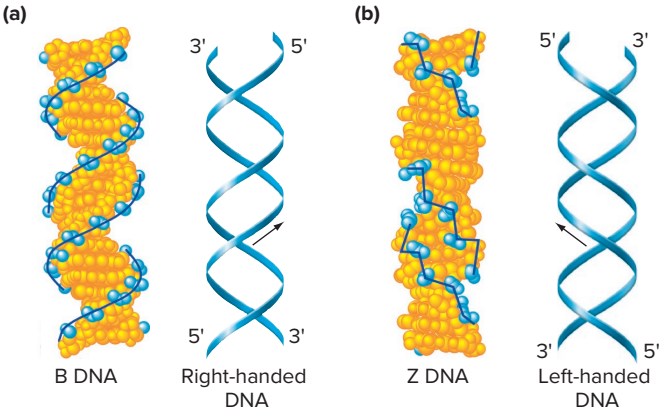


Figure 6.12 Z DNA is one variant of the double helix. **(a)** Typical Watson-Crick B-form DNA forms a right-handed helix with a smooth backbone. **(b)** Z-form DNA is left-handed and has an irregular backbone.



helixes before replication and expression—these are minor variations on the double-helical theme. Despite such experimentally determined departures of detail, the Watson-Crick double helix remains *the* model for thinking about DNA structure. This model describes those features of the molecule that have been preserved through billions of years of evolution.

DNA Structure Is the Foundation of Genetic Function

Without sophisticated computational tools for analyzing base sequence, one cannot distinguish bacterial DNA from human DNA. The reason is that all DNA molecules have the same general chemical properties and physical structure. Proteins, by comparison, are a much more diverse group of molecules with a much greater complexity of structure and function. In his account of the discovery of the double helix, Crick referred to this difference when he said that "DNA is, at bottom, a much less sophisticated molecule than a highly evolved protein and for this reason reveals its secrets more easily."

Four basic DNA secrets are embodied in the following four questions:

1. How does the molecule carry information?
2. How is that information copied for transmission to future generations?
3. What mechanisms allow the information to change?
4. How does DNA-encoded information govern the expression of phenotype?

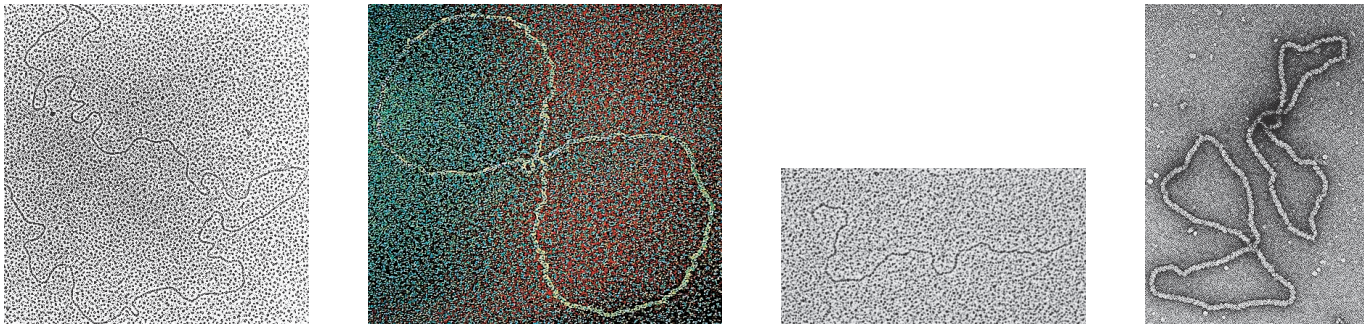
The double-helical structure of DNA provides a potential answer to each of these questions, endowing the molecule with the capacity to carry out all the crucial functions required of the genetic material.

In the remainder of this chapter, we describe how DNA's structure enables it to carry genetic information, replicate that information with great fidelity, and reorganize the information through recombination. How the information changes through mutation and how the information determines phenotype are the subjects of Chapters 7 and 8.

essential concepts

- The DNA molecule is a *double helix* composed of two *antiparallel* strands, in each of which nucleotides are joined by phosphodiester bonds. Hydrogen bonding between the *complementary bases*—A with T, and G with C—holds the two strands together.
- *Antiparallel* means that one strand is oriented in the 5′-to-3′ direction, while the other, *complementary strand* is oriented in the 3′-to-5′ direction.
- Most eukaryotes have double-stranded, linear DNA, but prokaryotes, chloroplasts and mitochondria, and some viruses have double-stranded circular DNA. Certain other viruses contain a single-stranded DNA that can be linear or circular.

Figure 6.13 DNA molecules may be linear or circular, double-stranded or single-stranded. These electron micrographs of naturally occurring DNA molecules show: **(a)** a fragment of a long, linear, double-stranded human chromosome, **(b)** a circular double-stranded papovavirus chromosome, **(c)** a linear single-stranded parvovirus chromosome, and **(d)** circular single-stranded bacteriophage M13 chromosomes.



(a) **(b)** **(c)** **(d)**

6.3 Genetic Information in Nucleotide Sequence

learning objectives

1. Explain how DNA stores complex information.
2. Compare the two ways in which the information in DNA may be accessed by proteins.
3. Describe the structural differences between DNA and RNA.

The information content of DNA resides in the sequence of its bases. The four bases in each chain are like the letters of an alphabet; they may follow each other in any order, and different sequences spell out different "words." Each word has its own meaning, that is, its own effect on phenotype. AGTCAT, for example, means one thing, while CTAGGT means another. Although DNA has only four different letters, or building blocks, the potential for different combinations and thus different sets of information in a long chain of nucleotides is staggering. Some human chromosomes, for example, are composed of chains that are 250 million nucleotides long; because the different bases may follow each other in any order, such chains could contain any one of $4^{250,000,000}$ (which translates to 1 followed by 150,515,000 zeros) potential nucleotide sequences.

Most Genetic Information Is Read from Unwound DNA Chains

The unwinding of a DNA molecule exposes a single sequence of bases on each of two strands (**Fig. 6.14**). Proteins read the information in a single DNA strand by synthesizing a stretch of RNA (a process called *transcription*) or DNA (a process called *replication*) complementary to a specific sequence.

Some Genetic Information Is Accessible Without Unwinding DNA

Some proteins can recognize and bind to specific base pair sequences in double-stranded DNA (**Fig. 6.15**). This information emerges mainly from differences between the four bases that appear in the major and minor grooves. Within the grooves, certain atoms at the periphery of the bases are exposed, and particularly in the major groove, these atoms may assume spatial patterns that provide chemical information. Proteins can access this information to sense the base sequence in a stretch of DNA without disassembling the double helix. Sequence-specific DNA-binding proteins include *transcription factors* that turn genes on and off (Chapters 16 and 17) as well as bacterial *restriction enzymes* that cut DNA at particular sites (Chapter 9).

In Some Viruses, RNA Is the Repository of Genetic Information

DNA carries the genetic information in all cellular forms of life and in many viruses. Prokaryotes such as *Escherichia coli* bacteria carry their DNA in a double-stranded, covalently closed circular chromosome. Eukaryotic cells package their DNA in double-stranded linear chromosomes. DNA viruses carry it in small molecules that are single- or double-stranded, circular, or linear.

By contrast, retroviruses, which include those that cause polio and AIDS, use **ribonucleic acid**, or **RNA** as their genetic material.

Figure 6.14 DNA stores information in the sequence of its bases. A partially unwound DNA double helix. Note that different structural information is available in the double-stranded and unwound regions of the molecule.

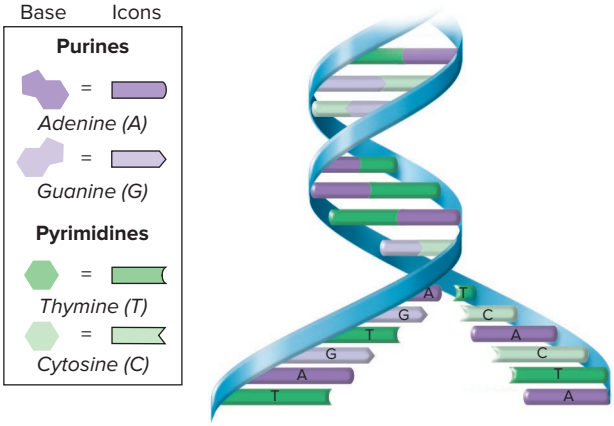


Figure 6.15 Proteins bind to specific sequences in DNA. Computer artwork of the *E. coli* catabolite gene activator protein (CAP) bound to DNA (*green* and *orange*). The structure of CAP is shown as a series of cylinders and ribbons. CAP can recognize specific sites in the major groove of double-helical DNA.

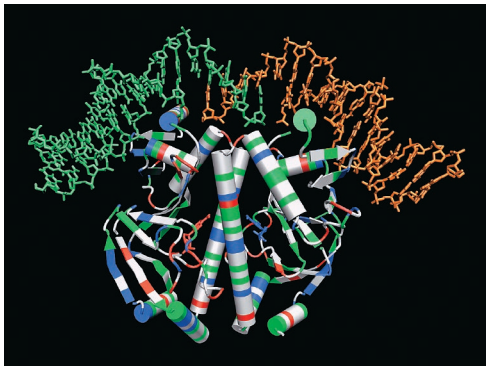


Figure 6.16 Differences in the chemical structure of DNA and RNA. Phosphodiester bonds join ribonucleotides into an RNA chain that differs from DNA in three ways (*bullets* in each column).

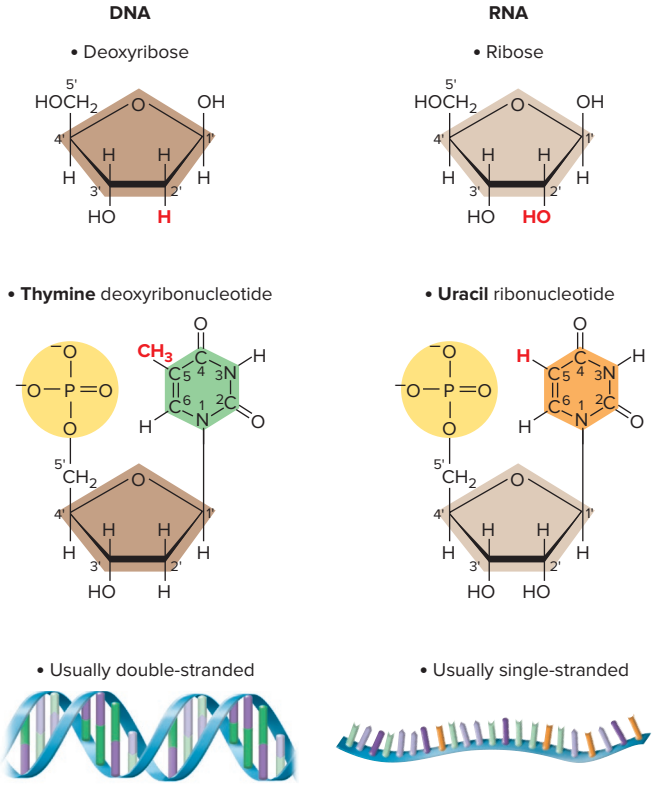
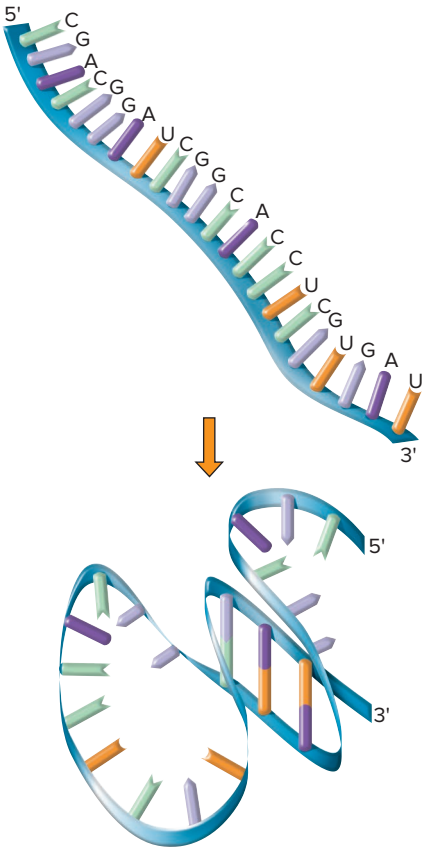


Figure 6.17 Complex folding patterns of RNA molecules. Most RNA molecules are single-stranded but are sufficiently flexible so that some regions can fold back and form base pairs with other parts of the same molecule.



Differences between RNA and DNA

Three major chemical differences distinguish RNA from DNA (**Fig. 6.16**). First, RNA takes its name from the sugar **ribose**, which it incorporates instead of the deoxyribose found in DNA. Second, RNA contains the base uracil (U) instead of the base thymine (T); U, like T, base pairs with A. Finally, most RNA molecules are single-stranded and contain far fewer nucleotides than the very long DNA molecules found in nuclear chromosomes.

Within a single-stranded RNA molecule, folding can bring together two oppositely oriented regions that carry complementary nucleotide sequences to form a short, base-paired stretch within the molecule. This means that, compared to the relatively simple, double-helical shape of a DNA molecule, many RNAs have a complicated structure of short double-stranded segments interspersed with single-stranded loops (**Fig. 6.17**).

RNA has the same ability as DNA to carry information in the sequence of its bases, but it is much less stable than DNA. In addition to serving as the genetic material for an array of viruses, RNA fulfills several vital functions in all cells. For example, it participates in gene expression and protein synthesis, as presented in Chapter 8. RNA also plays a surprisingly significant role in DNA replication, which we now describe.

essential concepts

- DNA carries digital information in the sequence of its four *bases*.
- The *base sequence* of DNA can be read from a single, unwound strand during replication or transcription. In addition, specialized proteins can recognize and bind to short base sequences accessible in the *grooves* of double-stranded DNA.
- *RNA* contains ribose rather than deoxyribose and uracil (U) instead of thymine (T); it also is generally single-stranded instead of double-stranded.

6.4 DNA Replication

learning objectives

1. Describe the key steps in the semiconservative replication of DNA.
2. Explain how the Meselson-Stahl experiment with heavy nitrogen showed that DNA replication is semiconservative.

3. Summarize the key factors DNA polymerase requires to replicate DNA.
4. Outline the steps in the process of DNA replication and how they relate to the requirements of DNA polymerase.
5. Discuss three ways cells preserve the accuracy and integrity of the genetic information in DNA.

In one of the most famous understatements in the scientific literature, Watson and Crick wrote at the end of their 1953 paper proposing the double helix model: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." This copying, as we saw in Chapter 4, must precede the transmission of chromosomes from one generation to the next via meiosis, and it is also the basis of the chromosome duplication prior to each mitosis that allows two daughter cells to receive a complete copy of the genetic information in a progenitor cell.

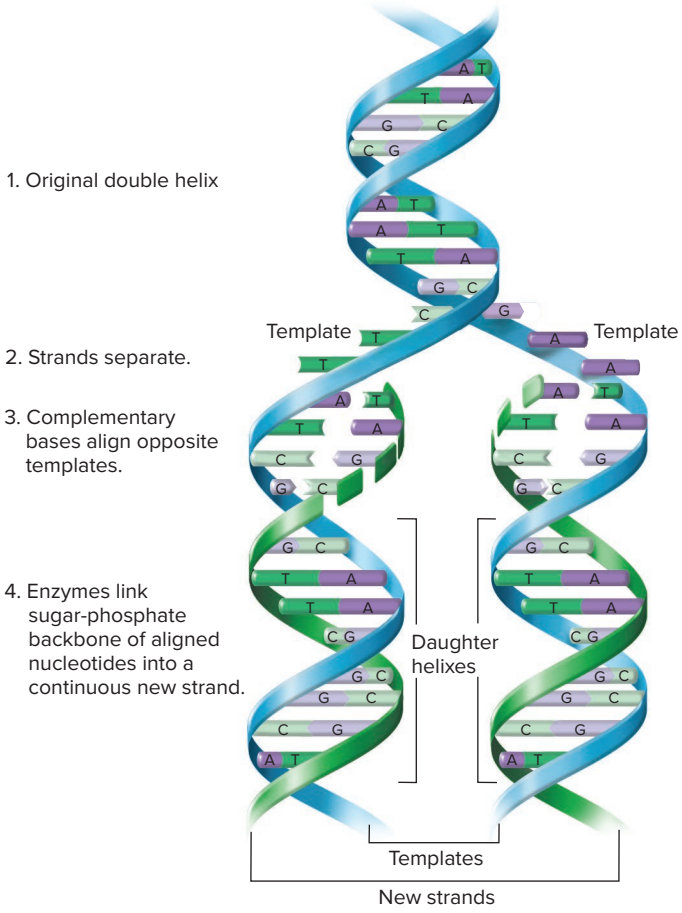
Overview: Complementary Base Pairing Ensures Semiconservative Replication

In the process of replication postulated by Watson and Crick, the double helix unwinds to expose the bases in each strand of DNA. Each of the two separated strands then acts as a **template,** or molecular mold, for the synthesis of a new second strand (**Fig. 6.18**). The newly replicated strands form as complementary bases align opposite the exposed bases on the two parental strands. That is, an A at one position on the original strand signals the addition of a T at the corresponding position on the newly forming strand; a T on the original signifies the addition of an A; similarly, G calls for C, and C calls for G, as demanded by complementary base pairing.

Once the appropriate base has aligned opposite to and formed hydrogen bonds with its complement, enzymes join the base's nucleotide to the preceding nucleotide by a phosphodiester bond, eventually linking a whole new line of nucleotides into a continuous strand. This mechanism of DNA strand separation and complementary base pairing followed by the coupling of successive nucleotides yields two daughter double helixes that each contain one of the original DNA strands intact (this strand is *conserved*) and one completely new strand (**Fig. 6.19a**). For this reason, such a pattern of double helix duplication is called **semiconservative replication:** a copying in which one strand of each new double helix is conserved from the parent molecule and the other is newly synthesized.

Watson and Crick's proposal is not the only replication mechanism imaginable. **Figures 6.19b** and **c** illustrate

Figure 6.18 The model of DNA replication postulated by Watson and Crick. Unwinding of the double helix allows each of the two parental strands to serve as a template for the synthesis of a new strand by complementary base pairing. The end result: A single double helix is transformed into two identical daughter double helixes.

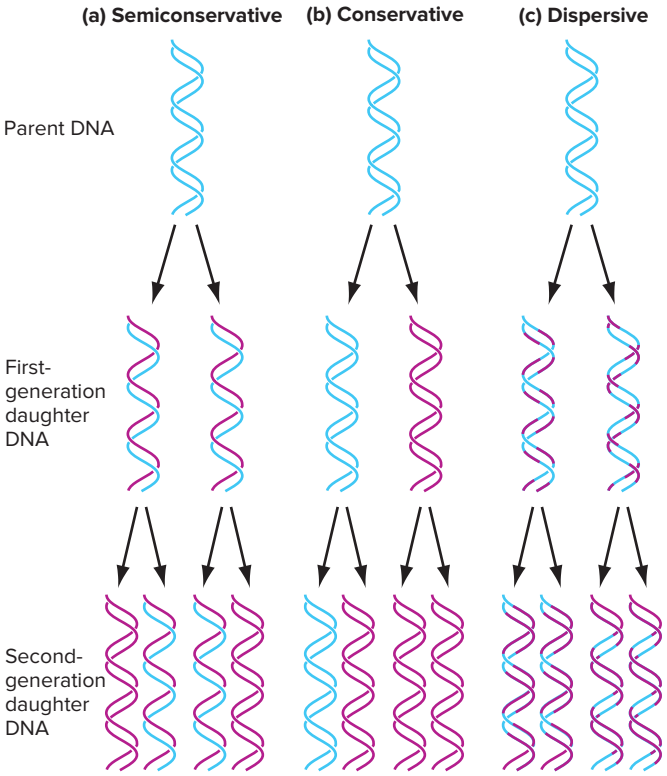


two possible alternatives. With *conservative replication,* one of the two daughter double helixes would consist entirely of original DNA strands, while the other helix would consist of two newly synthesized strands. With *dispersive replication,* both daughter double helixes would carry blocks of original DNA interspersed with blocks of newly synthesized material. These alternatives are less satisfactory because they do not immediately suggest a mechanism for copying the information in the sequence of bases.

Experiments with Heavy Nitrogen Verify Semiconservative Replication

In 1958, Matthew Meselson and Franklin Stahl performed an experiment that confirmed the semiconservative nature

Figure 6.19 Three possible models of DNA replication. DNA from the original double helix is *blue;* newly made DNA is *magenta.* **(a)** Semiconservative replication (the Watson-Crick model). **(b)** Conservative replication: The parental double helix remains intact; both strands of one daughter double helix are newly synthesized. **(c)** Dispersive replication: Both strands of both daughter double helixes contain both original and newly synthesized material.



of DNA replication (**Fig. 6.20**). The experiment depended on being able to distinguish preexisting *parental* DNA from newly synthesized DNA. To accomplish this, Meselson and Stahl controlled the isotopic composition of the nucleotides incorporated in the newly forming strands by taking advantage of the fact that the purine and pyrimidine bases of DNA contain nitrogen atoms. They grew *E. coli* bacteria for many generations on media in which all the nitrogen was the normal isotope ^{14}N; these cultures served as a control. They grew other cultures of *E. coli* for many generations on media in which the only source of nitrogen was the heavy isotope ^{15}N. After several generations of growth on heavy-isotope medium, essentially all the nitrogen atoms in the DNA of these bacterial cells were labeled with (that is, contained) ^{15}N. The cells in some of these cultures were then transferred to new medium in which all the nitrogen was ^{14}N. Any DNA synthesized after the transfer would contain the lighter isotope.

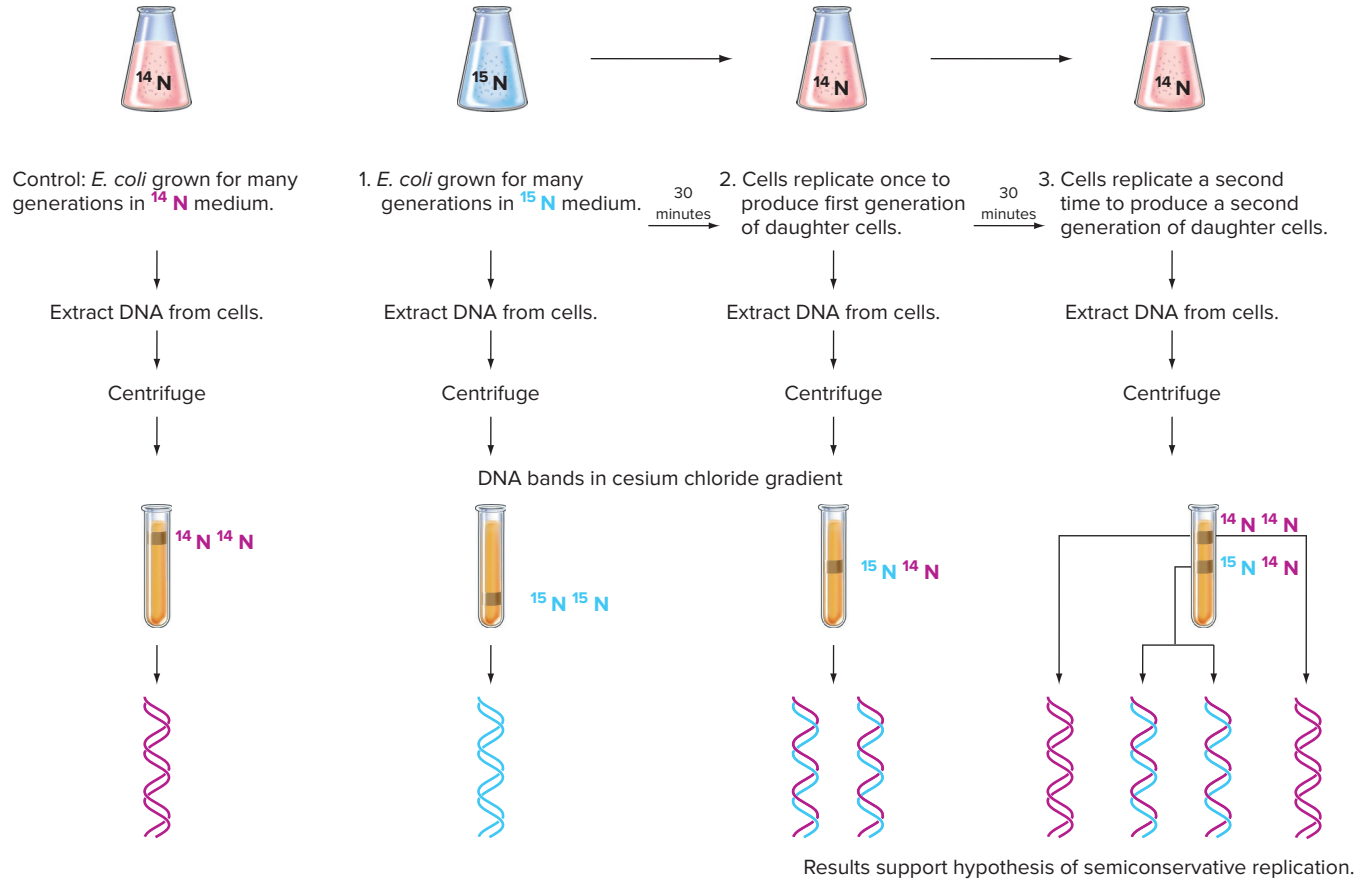
Meselson and Stahl isolated DNA from cells grown in the different nitrogen-isotope cultures and then subjected these DNA samples to *equilibrium density gradient centrifugation,* an analytic technique they had just developed. In a test tube, they dissolved the DNA in a solution of the dense salt cesium chloride (CsCl) and spun these solutions at very high speed (about 50,000 revolutions per minute) in an ultracentrifuge. Over a period of two to three days, the centrifugal force (roughly 250,000 times the force of gravity) causes the formation of a stable gradient of CsCl concentrations, with the highest concentration, and thus the highest CsCl density, at the bottom of the tube. The DNA in the tube forms a sharply delineated band at the position where its own density equals that of the CsCl. Because DNA containing ^{15}N is denser than DNA containing ^{14}N, pure ^{15}N DNA will form a band lower, that is, closer to the bottom of the tube, than pure ^{14}N DNA (Fig. 6.20).

As Fig 6.20 shows, when cells with pure ^{15}N DNA were transferred into ^{14}N medium and allowed to divide once, DNA from the resultant first-generation cells formed a band at a density intermediate between that of pure ^{15}N DNA and that of pure ^{14}N DNA. A logical inference is that the DNA in these cells contains equal amounts of the two isotopes. This finding invalidates the conservative model, which predicts the appearance of bands reflecting only pure ^{14}N and pure ^{15}N with no intermediary band. In contrast, DNA extracted from second-generation cells that had undergone a second round of division in the ^{14}N medium produced two observable bands, one at the density corresponding to equal amounts of ^{15}N and ^{14}N, the other at the density of pure ^{14}N. This result invalidates the dispersive model, which predicts a single band between the two bands of the original generation.

Meselson and Stahl's observations are consistent only with semiconservative replication. In the first generation after transfer from the ^{15}N to the ^{14}N medium, one of the two strands in every daughter DNA molecule carries the heavy isotope label; the other, newly synthesized strand carries the lighter ^{14}N isotope. The band at a density intermediate between that of ^{15}N DNA and ^{14}N DNA represents this isotopic hybrid. In the second generation after transfer, half of the DNA molecules have one ^{15}N strand and one ^{14}N strand, while the remaining half carry two ^{14}N strands. The two observable bands—one at the hybrid position, the other at the pure ^{14}N position—reflect this mix. By confirming the predictions of semiconservative replication, the Meselson-Stahl experiment disproved the conservative and dispersive alternatives. We now know that the semiconservative replication of DNA is nearly universal.

Let's consider precisely how semiconservative replication relates to the structure of chromosomes in eukaryotic cells during the mitotic cell cycle (review Fig. 4.9). Early in interphase, each eukaryotic chromosome contains a single continuous linear double helix of DNA. Later, during the S-phase portion of interphase, the cell replicates the double helix semiconservatively; after this semiconservative replication, each chromosome is composed of two sister chromatids joined at their centromeres. Each sister chromatid is a double helix of DNA, with one strand of parental DNA and one strand of newly synthesized DNA. At the conclusion of

Figure 6.20 How the Meselson-Stahl experiment confirmed semiconservative replication. **(1)** *E. coli* cells were grown in heavy ^{15}N medium. **(2)** and **(3)** The cells were transferred to ^{14}N medium and allowed to divide either once or twice. When DNA from each of these cell preparations was centrifuged in a cesium chloride gradient, the density of the extracted DNA conformed to the predictions of the semiconservative mode of replication, as shown at the *bottom* of the figure, where *blue* indicates heavy DNA and *magenta* depicts light DNA. The results are inconsistent with the conservative and dispersive models for DNA replication (compare with Fig. 6.19).



Results support hypothesis of semiconservative replication.

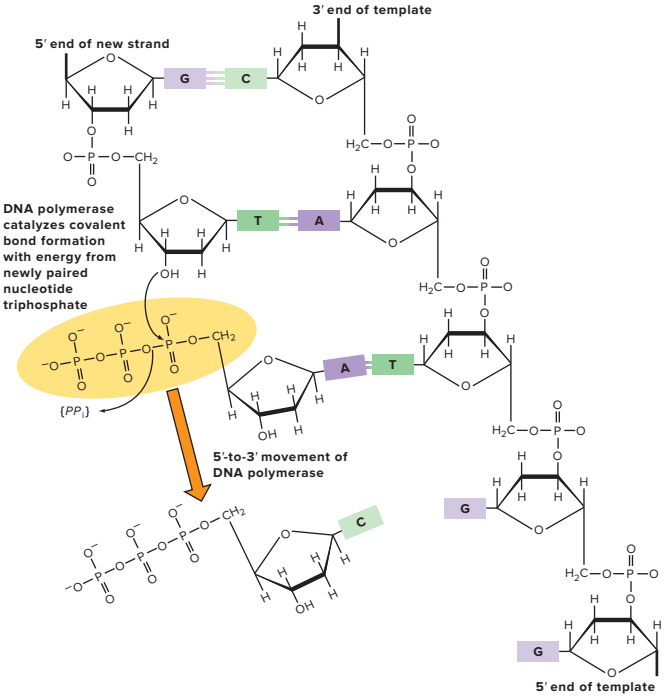
mitosis, each of the two daughter cells receives one sister chromatid from every chromosome in the cell. This process preserves chromosome number and identity during mitotic cell division because the two sister chromatids are identical in base sequence to each other and to the original parental chromosome.

DNA Polymerase Has Strict Operating Requirements

Watson and Crick's model for semiconservative replication is a simple concept to grasp, but the biochemical process through which it occurs is quite complex. Replication does not happen spontaneously any time a mixture of DNA and nucleotides is present. Rather, it occurs at a precise moment in the cell cycle, depends on a network of interacting regulatory elements, requires considerable input of energy, and involves a complex array of the cell's molecular machinery, including the key enzyme **DNA polymerase.** The salient details were deduced primarily in the laboratory of Arthur Kornberg, who won a Nobel prize for this work. The Kornberg group purified individual components of the replication machinery from *E. coli* bacteria. Remarkably, they were eventually able to elicit the reproduction of specific genetic information outside a living cell, in a test tube containing purified enzymes together with a DNA template, *primers* (defined below), and nucleotide triphosphates.

Although the biochemistry of DNA replication was elucidated for a single bacterial species, its essential features are conserved—just like the structure of DNA—within all organisms. The energy required to synthesize every DNA molecule found in nature comes from the high-energy phosphate bonds associated with the four *deoxyribonucleotide triphosphate*s (dATP, dCTP, dGTP, and dTTP; or dNTP as a general term) that provide bases for incorporation into the growing DNA strand. As shown in **Fig. 6.21,** this conserved biochemical feature means that DNA synthesis can proceed only by adding nucleotides to the 3′ end of an existing polynucleotide. With energy released from severing the triphosphate arm of a dNTP substrate molecule, the DNA polymerase enzyme catalyzes the formation of a new phosophodiester bond. Once this bond is formed, the enzyme proceeds to join up the next nucleotide brought into position by complementary base pairing.

Figure 6.21 DNA synthesis proceeds in a 5′-to-3′ direction. The template strand is shown on the *right* in an antiparallel orientation to the new DNA strand under synthesis on the *left*. Here, a free molecule of dATP has formed hydrogen bonds with a complementary thymine base on the template strand. DNA polymerase (*yellow*) cleaves dATP between the first and second phosphate groups. This cleavage releases the energy needed to form a covalent phosphodiester bond between the terminal 3′-OH group on the preceding nucleotide to the first phosphate of the dATP substrate. Pyrophosphate (PP_i) is released as a by-product.



Many proteins in addition to DNA polymerase are required to replicate DNA. However, you will see below that the most important features of DNA replication reflect three strict requirements for DNA polymerase action (**Fig. 6.22**):

(1) The four dNTPs.

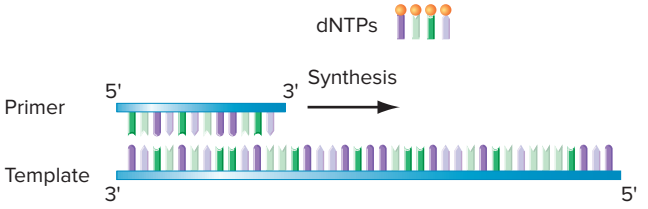
(2) A single-stranded template. Double-stranded DNA must be unwound, and DNA polymerase moves along the template strand in the 3′-to-5′ direction.

(3) A **primer** with a free 3′ hydroxyl group. DNA polymerase adds nucleotides successively to the 3′ end of the growing DNA chain. (That is, DNA polymerase synthesizes DNA only in the 5′-to-3′ direction.) However, DNA polymerase cannot establish the first link in a new chain. Polymerization therefore must start with a *primer,* a short, single-stranded molecule of DNA or RNA a few nucleotides long that base pairs with part of the template strand.

DNA Replication Is a Tightly Regulated, Complex Process

The formation of phosphodiester bonds by DNA polymerase is just one component of the highly coordinated process by

Figure 6.22 Requirements of DNA polymerase. To synthesize DNA, DNA polymerase requires a single-stranded DNA template, a primer that can be RNA or DNA, and free deoxyribonucleotide triphosphates (dNTPs). DNA polymerase adds nucleotides successively onto the 3′ end of the primer as instructed by the complementary nucleotides in the template.



which DNA replication occurs inside a living cell. The entire molecular mechanism, illustrated in **Fig. 6.23,** has two stages: **initiation,** during which proteins open up the double helix and prepare it for complementary base pairing, and **elongation,** during which proteins connect the correct sequence of nucleotides on both newly formed DNA double helixes.

DNA replication is complicated by the strict biochemical mechanism of polymerase function. DNA polymerase can lengthen existing DNA chains only by adding nucleotides to the 3′ hydroxyl group of the DNA strand, as was shown in Figs. 6.21 and 6.22. However, the antiparallel strands of DNA unwind progressively at the two Y-shaped areas called the **replication forks** in **Fig. 6.23a**. As a result, one newly synthesized strand (the **leading strand**) can grow continuously into each of the opening forks. But the other new strand (the **lagging strand**), made at the same fork but synthesized from the other template strand, can only be generated in pieces called **Okazaki fragments** as more and more template is unwound at the fork (**Fig. 6.23b**). These fragments must be joined together later in the process.

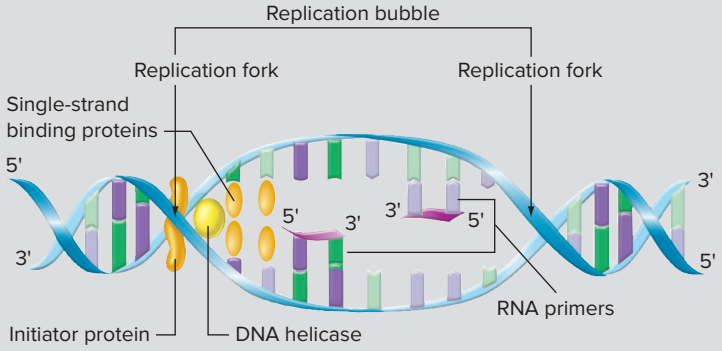
As Fig. 6.23 shows, DNA replication depends on the coordinated activity of many different proteins, including two different DNA polymerases called pol I and pol III (*pol* is short for polymerase). Pol III plays the major role in producing the new strands of complementary DNA, while pol I fills in the gaps between newly synthesized Okazaki segments. Other enzymes contribute to the initiation process: **DNA helicase** unwinds the double helix. A special group of *single-stranded DNA binding proteins* keep the DNA helix open. An enzyme called **primase** generates RNA primers to initiate DNA synthesis. During elongation, the *DNA ligase* enzyme welds together Okazaki fragments.

It took many years for biochemists and geneticists to discover how the tight collaboration of many proteins drives DNA replication. Today scientists think that programmed molecular interactions of this kind underlie many of the biochemical processes that occur in cells. In these processes, a group of proteins, each performing a specialized function, like the workers on an assembly line, cooperate in the manufacture of complex macromolecules.

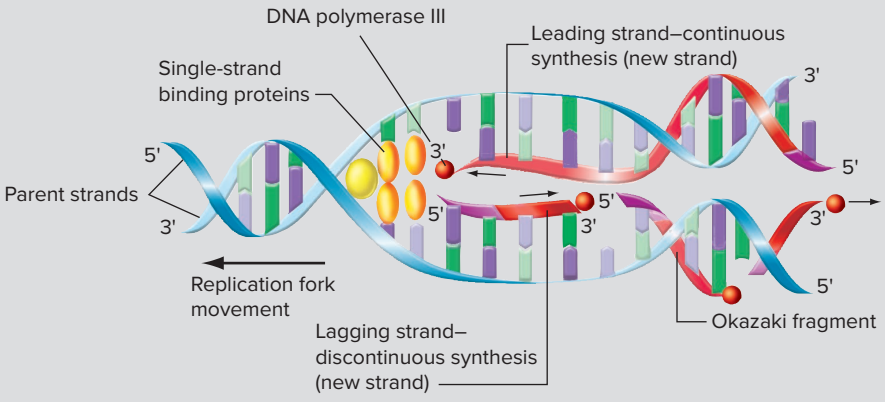
The Mechanism of DNA Replication

(a) Initiation: Preparing the double helix for use as a template. Initiation begins with the unwinding of the double helix at a particular short sequence of nucleotides known as the **origin of replication.** Several proteins bind to the origin, starting with the *initiator protein.* Initiator attracts **DNA helicase**, which unwinds the double helix. The opening up of a region of DNA generates two Y-shaped areas called **replication forks**, one at either end of the unwound area, the **replication bubble.** The single strands will serve as **templates** for synthesizing new strands of DNA.

An enzyme complex known as **DNA polymerase III** adds nucleotides to the 3′ end of a preexisting strand of nucleic acid. The requirement for an already existing chain means that something else must prime the about-to-be-constructed chain. In living cells, that something else is a short stretch of RNA called an **RNA primer,** synthesized by an enzyme called **primase.**



(b) Elongation: Connecting the correct sequence of nucleotides into a continuous new strand of DNA. Through complementary base pairing, the order of bases in the template strand specifies the order of bases in the newly forming strand. DNA polymerase III catalyzes the joining of a new nucleotide to the preceding nucleotide through the formation of a phosphodiester bond, a process known as **polymerization.** DNA polymerase III first joins the correctly paired nucleotide to the 3′ hydroxyl end of the RNA primer, and then it continues to add the appropriate nucleotides to the 3′ end of the growing chain. As a result, the DNA strand under construction grows in the 5′-to-3′ direction, while the DNA polymerase molecule actually moves along the antiparallel template strand in the 3′-to-5′ direction. (The three following figures diagram only the events occurring at the left replication fork.)

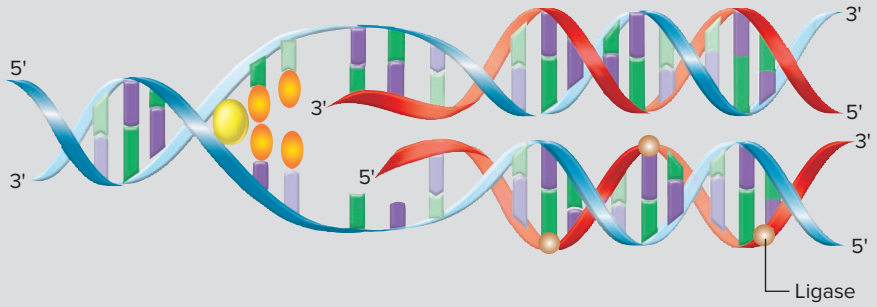


As DNA replication proceeds, helicase progressively unwinds the double helix. DNA polymerase III moves in the same direction as the fork to synthesize the **leading strand.** However, the synthesis of the second new DNA chain — the **lagging strand**— is problematic. The polarity of the lagging strand is opposite that of the leading strand, yet DNA polymerase functions only in the 5′-to-3′ direction. How can this work?

The answer is that the lagging strand is synthesized *discontinuously* in the normal 5′-to-3′ direction as small fragments of about 1000 bases called **Okazaki fragments** (after their discoverers, Reiji and Tuneko Okazaki).

Because DNA polymerase III can add nucleotides only to the 3′ end of an existing strand, each Okazaki fragment is initiated by a short RNA primer. Primase catalyzes the formation of the RNA primer for each upcoming Okazaki fragment as soon as the replication fork has progressed a sufficient distance along the DNA. DNA polymerase III then adds nucleotides to the primer, generating an Okazaki fragment that extends up to the 5′ end of the primer of the previously synthesized fragment.

Finally, **DNA polymerase I** and other enzymes replace the RNA primer of the previously made Okazaki fragment with DNA, and the enzyme **DNA ligase** covalently joins successive Okazaki fragments into a continuous strand of DNA.



The unwinding of DNA beginning at the origin of replication produces two forks (Fig. 6.23a). As a result, replication is generally *bidirectional,* with the replication forks moving in opposite directions as unwinding proceeds. At each fork, polymerase copies both template strands, one in a continuous fashion, the other discontinuously as Okazaki fragments (Fig. 6.23b).

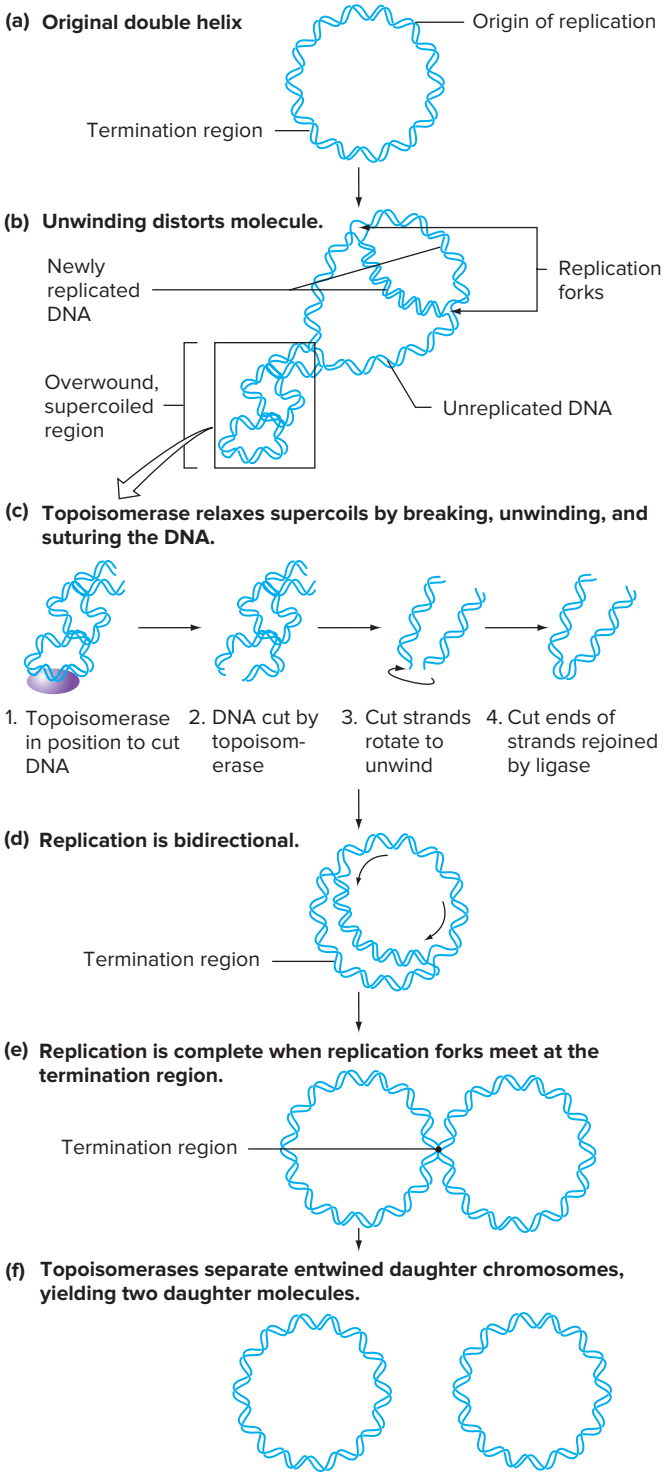
In the circular *E. coli* chromosome, there is only one origin of replication (**Fig. 6.24a**). When its two forks, moving in opposite directions, meet at a designated *termination region* about halfway around the circle from the origin of replication, replication is complete (**Figs. 6.24d–f**).

Not surprisingly, local unwinding of the double helix at a replication fork affects the chromosome as a whole. In *E. coli,* the unwinding of a section of a covalently closed circular chromosome overwinds and distorts the rest of the molecule (**Fig. 6.24b**). Overwinding reduces the number of helical turns to less than the 1-every-10.5-nucleotides characteristic of B-form DNA. The chromosome accommodates the strain of distortion by twisting back upon itself. You can envision the effect by imagining a coiled telephone cord that overwinds and bunches up with use. The additional twisting of the DNA molecule is called **supercoiling.** Movement of the replication fork causes more and more supercoiling.

This cumulative supercoiling, if left unchecked, would wind the chromosome up so tightly that it would impede the progress of the replication fork. A group of enzymes known as **DNA topoisomerases** helps relax the supercoils by nicking one or cutting both strands of the DNA—that is, cleaving the sugar-phosphate backbone between two adjoining nucleotides (**Fig. 6.24c**). Just as a telephone cord freed at the handset end can unwind and restore its normal coiling pattern, the DNA strands, after cleavage, can rotate relative to each other and thereby restore the normal coiling density of one helical turn per 10.5 nucleotide pairs. The activity of topoisomerases thus allows replication to proceed through the entire chromosome by preventing supercoils from accumulating in front of the replication fork. Replication of a circular double helix sometimes produces intertwined daughter molecules whose clean separation also depends on topoisomerase activity (Fig. 6.24e and f).

In the much larger, linear chromosomes of eukaryotic cells, bidirectional replication proceeds roughly as just described but from many origins of replication. The multiple origins ensure that copying is completed within the time allotted (that is, within the S period of the cell cycle). In addition, because the lagging strand is synthesized as Okazaki fragments, replication of the very ends of linear chromosomes is also problematic. But eukaryotic chromosomes have evolved specialized termination structures known as **telomeres,** which ensure the maintenance and accurate replication of the two ends of each linear chromosome. (Chapter 12 presents the details of eukaryotic chromosome replication.)

Figure 6.24 The bidirectional replication of a circular bacterial chromosome: An overview. (a) and **(b)** Replication proceeds in two directions from a single origin, creating two replication forks that move in opposite directions around the circle. Local unwinding of DNA at the replication forks creates supercoiled twists in the DNA in front of the replication fork. **(c)** The action of topoisomerase enzymes helps reduce this supercoiling. **(d)** and **(e)** When the two replication forks meet at the termination region, the entire chromosome has been copied. **(f)** Topoisomerase enzymes separate the two daughter chromosomes.



The Integrity of Genetic Information Must Be Preserved

DNA is the sole repository of the vast amount of information required to specify the structure and function of most organisms. In some species, this information may lie in storage for many years, or it may undergo replication many times before it is called on to generate progeny. During the time of storage and before gamete production, the organism must protect the integrity of the information, for even the most minor change can have disastrous consequences, such as causing severe genetic disease or even death. Each organism ensures the informational fidelity of its DNA in three important ways:

- *Redundancy.* Either strand of the double helix can specify the sequence of the other. This redundancy provides a basis for checking and repairing errors arising either from chemical alterations sustained during storage or from rare malfunctions of the replication machinery.
- *The remarkable precision of the cellular replication machinery.* Evolution has perfected the cellular machinery for DNA replication to the point where errors during copying are exceedingly rare. For example, DNA polymerase has acquired a proofreading ability to prevent unmatched nucleotides from joining a new strand of DNA; as a result, a free nucleotide is attached to a growing strand only if its base is correctly paired with its complement on the parent strand. We examine this proofreading mechanism in Chapter 7.
- *Enzymes that repair chemical damage to DNA.* The cell has an array of enzymes devoted to the repair of nearly every imaginable type of chemical damage. We describe how these enzymes carry out their corrections in Chapter 7.

All of these safeguards help ensure that the information content of DNA will be transmitted intact from generation to generation of cells and organisms. However, as we see next, new combinations of existing information arise naturally as a result of recombination.

essential concepts

- The DNA molecule is reproduced by *semiconservative replication;* the two DNA strands separate, and each acts a *template* for the synthesis of a new complementary strand.
- *DNA polymerase* synthesizes DNA in the 5′-to-3′ direction by adding nucleotides successively onto the 3′ end of a growing DNA chain.
- DNA polymerase requires: (i) a supply of the four *deoxyribonucleotide triphosphates,* (ii) a single-stranded DNA template, and (iii) a *primer* of either DNA or (in cells) RNA with a free 3′ hydroxyl group.
- At the DNA *replication fork*, DNA polymerase synthesizes one new strand (the *leading strand*) continuously, while the other (*lagging strand*) is synthesized as multiple *Okazaki fragments* that are then joined by DNA ligase.
- The integrity and accuracy of information in DNA is preserved by redundancy in the two strands, the precision of the enzymes synthesizing DNA, and the action of enzymes that repair damage to DNA.

6.5 Homologous Recombination at the DNA Level

learning objectives

1. Summarize the evidence from tetrad analysis confirming that recombination occurs at the four-strand stage and involves reciprocal exchange.
2. Explain how we know that DNA breaks and rejoins during recombination.
3. List the key steps of recombination at the molecular level.
4. Explain why recombination events do not always result in crossing-over.
5. Describe how mismatch repair of heteroduplex regions can lead to gene conversion in fungal tetrads.

Mutation, the ultimate source of all new alleles, is a rare phenomenon at any particular nucleotide pair on a chromosome. The most important mechanism for generating genomic diversity in sexually reproducing species is thus the production of new combinations of already existing alleles. This type of diversity increases the chances that at least some offspring of a mating pair will inherit a combination of alleles best suited for survival and reproduction in a changing environment.

New combinations of already existing alleles arise from two different types of meiotic events: (i) independent assortment, in which each pair of homologous chromosomes segregates free from the influence of other pairs, via random spindle attachment; and (ii) crossing-over, in which two homologous chromosomes exchange parts. Independent assortment can produce gametes carrying new allelic combinations of genes on different chromosomes, but for genes on the same chromosome, independent assortment alone will only conserve the existing combinations of alleles. Crossing-over, however, can generate new allelic combinations of linked genes. The evolution of crossing-over thus compensated for what would otherwise be a significant disadvantage of the linkage of the genes within chromosomes.

Historically, geneticists have used the term *recombination* to indicate the production of new combinations of alleles by any means, including independent assortment. But

in the remainder of this chapter, we use **recombination** more narrowly to mean the generation of new allelic combinations through genetic exchange between homologous chromosomes. In this discussion, we refer to the products of crossing-over as **recombinants:** chromosomes that carry a mix of alleles derived from different homologs.

In eukaryotic organisms, recombination has an additional essential function beyond generating new combinations of alleles: It helps ensure proper chromosome segregation during meiosis. Chapter 4 has already described how crossovers, in combination with sister chromatid cohesion, keep homologous chromosomes together as bivalents during the period between prophase I and metaphase I. If homologs fail to recombine, they often cannot orient themselves toward opposite poles of the meiosis I spindle, resulting in nondisjunction.

As we examine recombination at the molecular level, we look first at experiments establishing the basic parameters of crossing-over. We will then describe the molecular details of a crossover event.

Tetrad Analysis Illustrates Key Aspects of Recombination

You saw in Chapter 5 that some fungi like yeast and *Neurospora* generate asci that contain in one sac all the products of individual meioses—that is, tetrads. Analysis of these tetrads allowed geneticists to infer basic information about recombination.

Evidence that recombination takes place at the four-strand stage

Recall that in tetrad analysis, the hallmark of linkage between two genes is that very few NPD tetrads are produced: The number of PDs, and also Ts, is always greater than the number of NPDs. This outcome makes sense because all SCO and some DCO meioses yield Ts, while only one-quarter of the rare DCOs yield NPDs (review Figs. 5.23 and 5.24).

The very low number of NPDs actually observed in crosses establishes that recombination occurs after the chromosomes have replicated, when four chromatids exist for each pair of homologs (**Fig. 6.25a**). If recombination instead took place before chromosome duplication, every single crossover event (that is, every SCO) would yield four recombinant chromatids and generate an NPD tetrad; such a model for recombination would be unable to account for any T tetrads (**Fig. 6.25b**).

Evidence that recombination is usually reciprocal

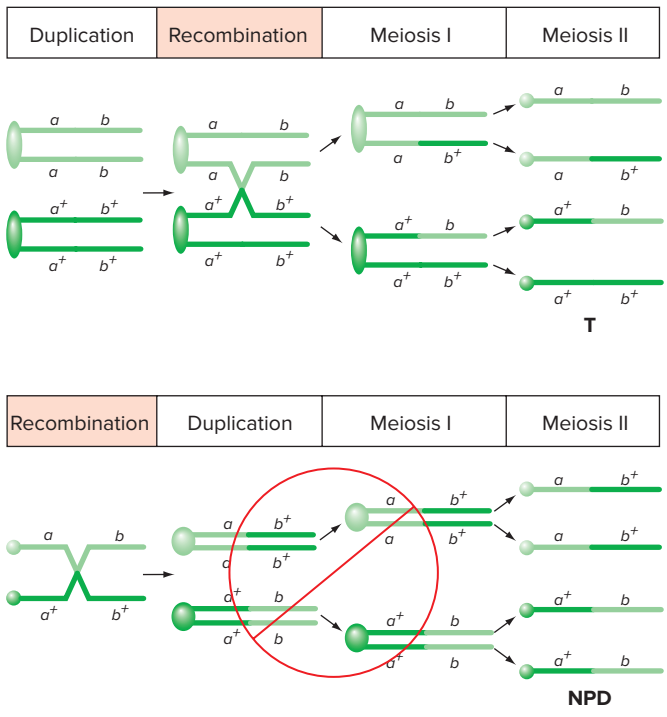
The discussion in Chapter 5 assumed that recombination is reciprocal, with nonsister chromatids from homologous chromosomes exchanging parts equally. That is, whatever is lost from one chromatid is gained from the other, and *vice versa.* We know that this assumption is legitimate because of the results of tetrad analysis.

Suppose you are following linked genes *A* and *B* in a cross between *A B* and *a b* strains of yeast. If the recombination that occurs during meiosis is reciprocal, every tetrad with recombinant progeny should contain equal numbers of both classes of recombinants. Observations have in general confirmed this prediction: Every T tetrad carries one *A b* and one *a B* spore, while every NPD tetrad contains two of each type of recombinant (recall Fig. 5.24). We can thus conclude that meiotic recombination is *almost* always reciprocal. We say *almost* always because as you will see later in this chapter, rare exceptions exist in which a tetrad cannot be classified PD, NPD, or T. These exceptional tetrads helped scientists to understand key features of the recombination mechanism.

DNA Molecules Break and Rejoin During Recombination

When viewed through the light microscope, recombinant chromosomes bearing physical markers appear to result from two homologous chromosomes breaking and exchanging parts as they rejoin (see Fig. 5.7). Because the recombined chromosomes, like all other chromosomes, are composed of

Figure 6.25 Recombination after chromosome replication. **(a)** Because recombination occurs after chromosomes have replicated, most tetrads containing recombinant spores are Ts. **(b)** A disproven model. If recombination occurred before chromosomes replicated and if two genes were linked, most tetrads containing recombinant spores would be NPDs instead of Ts.



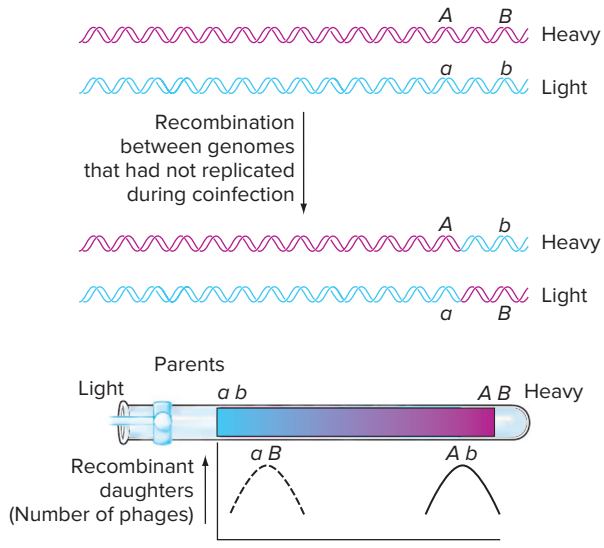
one long DNA molecule, a logical expectation is that they should show some physical signs of this breakage and rejoining at the molecular level. To evaluate this hypothesis, researchers chose a bacterial virus, *lambda,* as their model organism. Lambda had a distinct experimental advantage for this particular study: It is about half DNA by weight, so the density of the whole virus reflects the density of its DNA.

The experimental technique was similar in principle to the one Meselson and Stahl used to monitor a change in DNA density during DNA replication. In this case, however, the researchers (again Matthew Meselson but with a different collaborator, Jean Weigle) monitored DNA density to look at recombination. They used two strains of bacterial viruses that were genetically marked to keep track of recombination. They grew the wild-type strain (*A B*) in medium with heavy isotopes of carbon and nitrogen, and a strain with mutations in two genes (*a b*) in medium with the normal light isotopes of these atoms (**Fig. 6.26**). Meselson and Weigle then infected bacterial cells growing in normal (light isotope) medium with so many phages of each type that every cell was infected with both viral strains. After allowing time for the phages to replicate, recombine, and repackage their DNA into virus particles, the experimenters isolated the viruses released from the lysed cells and analyzed them on a density gradient.

It was important to the design of the experiment that both genes *A* and *B* were close to one end of the viral chromosome (Fig. 6.26). The idea was that some of the original phage chromosomes would undergo recombination before replicating in the light isotope medium. For example, some of the heavy *A B* chromosomes would recombine with light *a b* chromosomes. If crossing-over occurred through breakage and reunion of double-stranded DNA molecules, then some *A b* recombinant phages from the lysed cells should have a density almost as heavy as that of the parent phages that were *A B* (Fig. 6.26). In contrast, few if any recombinants of genotype *a B* should be composed of mostly heavy DNA.

Because the phages had replicated in light medium, recombinant phages could be found throughout the gradient. However, the key result was that a substantial proportion of the *A b* recombinants were indeed found near the heaviest density along with *A B* parent molecules. This result makes sense only if the *A b* chromosomes consisted mostly of double-helical heavy DNA, as expected for the kind of chromosomal breakage and reunion shown in Fig. 6.26.

Figure 6.26 DNA molecules break and rejoin during recombination. Meselson and Weigle infected *E. coli* cells with two different genetically marked strains of bacteriophage lambda previously grown in the presence of heavy (^{13}C and ^{15}N) or light (^{12}C and ^{14}N) isotopes. After growth on light medium, they spun the progeny bacteriophages on a CsCl density gradient. Some *A b* genetic recombinants (but almost no *a B* recombinants) had a heavy density almost the same as that of the *A B* parent.



Crossing-Over at the Molecular Level: A Model

Biochemical experiments performed mostly in yeast have informed our present understanding of the molecular mechanism for meiotic recombination. Researchers have found that the protein Spo11, which plays crucial roles in initiating meiotic recombination in yeast, is homologous to a protein essential for meiotic recombination in nematodes, plants, fruit flies, and mammals. This finding suggests that the mechanism of recombination presented in detail in **Fig. 6.27**—and known as the *double-strand-break repair model*—has been conserved throughout the evolution of eukaryotes.

In the figure, we focus on two nonsister chromatids, even though recombination takes place at the four-strand stage. Furthermore, we use the term *recombination event* to describe the molecular process initiated by Spo11, whether or not it results in crossing-over. As you are about to see, the molecular details of recombination events are such that crossing-over (reciprocal exchange of double-stranded DNA of nonsister chromosomes) results from the Spo11-mediated process only some of the time.

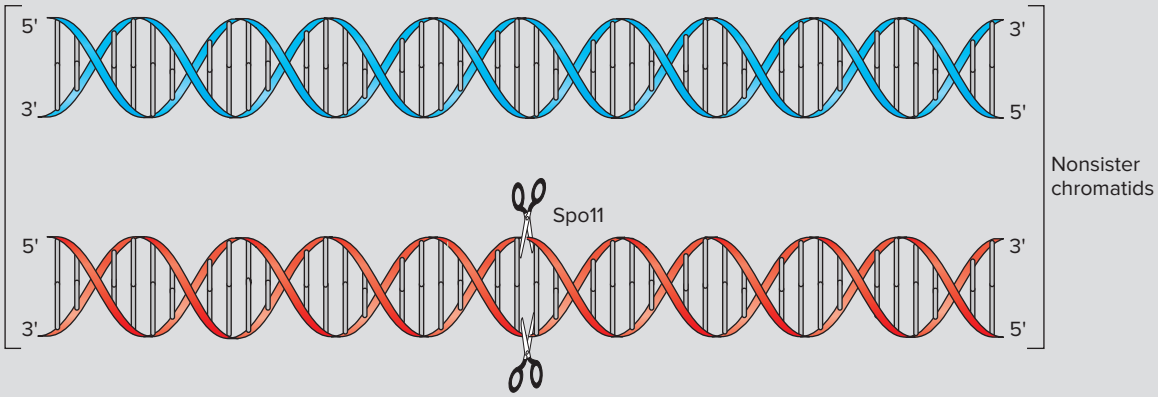
Initiating recombination

A meiotic recombination event begins when Spo11 makes a double-strand break in one of the four chromatids (Fig. 6.27, Step 1). Next, in a process called **resection,** an **exonuclease** (an enzyme that removes nucleotides from an end of a DNA molecule) degrades one strand of DNA from both sides of the cleavage, leaving 3′ single-stranded tails (Fig. 6.27, Step 2). In the next set of reactions, called **strand invasion,** one single-stranded tail displaces the corresponding strand on the nonsister chromatid (Fig. 6.27, Step 3). Strand invasion results in the formation of a **heteroduplex** region (from the Greek *hetero* meaning *other* or *different*) in which the DNA molecule is composed of one strand from each nonsister chromatid (Fig. 6.27, Step 3).

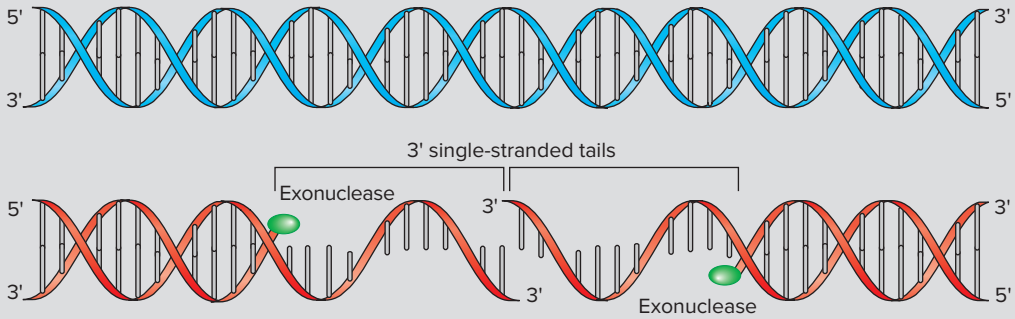
FEATURE FIGURE 6.27

A Model of Recombination at the Molecular Level

Step 1 Double-strand break formation. During meiotic prophase, **Spo11** protein makes a double-strand break on one of the chromatids by cleaving the phosphodiester bonds between adjacent nucleotides on both strands of the DNA. (Note that only the two nonsister chromatids undergoing recombination are shown.)



Step 2 Resection. An **exonuclease** degrades the 5′ ends on each side of the break to produce two 3′ single-stranded tails.



Step 3 First strand invasion. The protein Dmc1 (*orange* ovals) collaborates with other proteins (*not shown*) to help one of the tails invade and open up the other chromatid's double helix. Dmc1 then moves along the double helix, prying it open. The invading strand scans the base sequence in the momentarily unwound stretches of DNA duplex. As soon as it finds a complementary sequence of sufficient length, the two strands form a **heteroduplex** maintained by dozens of hydrogen bonds. The strand displaced by the invading tail forms a **D-loop** (for displacement loop), which is stabilized by binding of replication protein A (RPA) (*yellow* ovals).