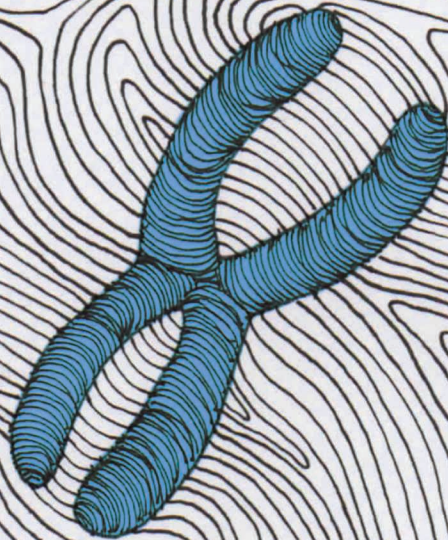
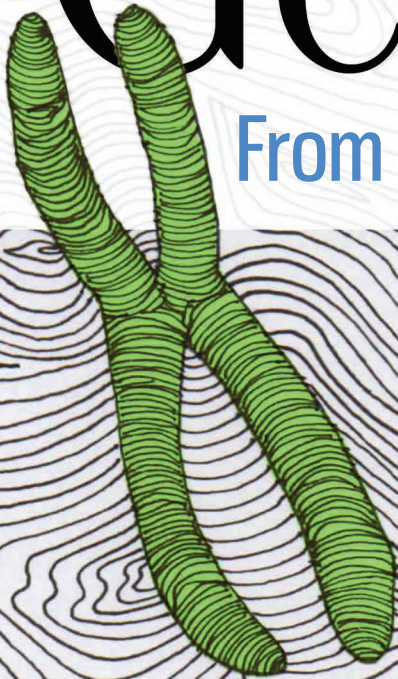


SIXTH EDITION

Genetics

From Genes to Genomes



Mc
Graw
Hill
Education

HARTWELL ~ GOLDBERG ~ FISCHER ~ HOOD

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

Published by McGraw-Hill Education, 2 Penn Plaza, New York, NY 10121. Copyright © 2018 by McGraw-Hill Education. All rights reserved. Printed in the United States of America. Previous editions © 2015, 2011, and 2008. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of McGraw-Hill Education, including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

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*

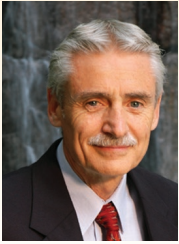
All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

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://lcn.loc.gov/2017017175>

The Internet addresses listed in the text were accurate at the time of publication. The inclusion of a website does not indicate an endorsement by the authors or McGraw-Hill Education, and McGraw-Hill Education does not guarantee the accuracy of the information presented at these sites.

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.

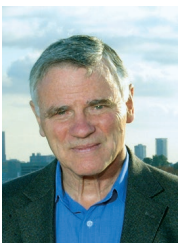
Dr. Hartwell is a member of the National Academy of Sciences and has received the Albert Lasker Basic Medical Research Award, the Gairdner Foundation International Award, the Genetics Society Medal, and the 2001 Nobel Prize in Physiology or Medicine.



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

1 Genetics: The Study of Biological Information 1

PART I

Basic Principles: How Traits Are Transmitted 14

- 2 Mendel's Principles of Heredity 14
- 3 Extensions to Mendel's Laws 45
- 4 The Chromosome Theory of Inheritance 89
- 5 Linkage, Recombination, and the Mapping of Genes on Chromosomes 133

PART II

What Genes Are and What They Do 181

- 6 DNA Structure, Replication, and Recombination 181
- 7 Anatomy and Function of a Gene: Dissection Through Mutation 219
- 8 Gene Expression: The Flow of Information from DNA to RNA to Protein 270

PART III

Analysis of Genetic Information 316

- 9 Digital Analysis of DNA 316
- 10 Genome Annotation 341
- 11 Analyzing Genomic Variation 365

PART IV

How Genes Travel on Chromosomes 406

- 12 The Eukaryotic Chromosome 406
- 13 Chromosomal Rearrangements and Changes in Chromosome Number 436
- 14 Bacterial Genetics 487
- 15 Organellar Inheritance 521

PART V

How Genes Are Regulated 547

- 16 Gene Regulation in Prokaryotes 547
- 17 Gene Regulation in Eukaryotes 583

PART VI

Using Genetics 618

- 18 Manipulating the Genomes of Eukaryotes 618
- 19 The Genetic Analysis of Development 646
- 20 The Genetics of Cancer 681

PART VII

Beyond the Individual Gene and Genome 713

- 21 Variation and Selection in Populations 713
- 22 The Genetics of Complex Traits 746

About the Authors iii
Preface ix
Acknowledgements xx

Introduction: Genetics in the Twenty-First Century 1

chapter 1

Genetics: The Study of Biological Information 1

- 1.1 DNA: Life's Fundamental Information Molecule 2
- 1.2 Proteins: The Functional Molecules of Life Processes 3
- 1.3 Molecular Similarities of All Life-Forms 4
- 1.4 The Modular Construction of Genomes 6
- 1.5 Modern Genetic Techniques 8
- 1.6 Human Genetics and Society 10

PART I



Basic Principles: How Traits Are Transmitted 14

© Lawrence Manning/Corbis RF

chapter 2

Mendel's Principles of Heredity 14

- 2.1 The Puzzle of Inheritance 15
- 2.2 Genetic Analysis According to Mendel 20
- 2.3 Mendelian Inheritance in Humans 30

■ *Genetics and Society: Developing Guidelines for Genetic Screening 34*

chapter 3

Extensions to Mendel's Laws 45

- 3.1 Extensions to Mendel for Single-Gene Inheritance 46
- 3.2 Extensions to Mendel for Two-Gene Inheritance 55
- 3.3 Extensions to Mendel for Multifactorial Inheritance 68

■ *Genetics and Society: Disease Prevention Versus the Right to Privacy 72*

chapter 4

The Chromosome Theory of Inheritance 89

- 4.1 Chromosomes: The Carriers of Genes 90
- 4.2 Sex Chromosomes and Sex Determination 94
- 4.3 Mitosis: Cell Division That Preserves Chromosome Number 98
- 4.4 Meiosis: Cell Divisions That Halve Chromosome Number 103
- 4.5 Gametogenesis 109
- 4.6 Validation of the Chromosome Theory 113
- 4.7 Sex-Linked and Sexually Dimorphic Traits in Humans 118

■ *Genetics and Society: Prenatal Genetic Diagnosis 93*

■ *Fast Forward: Transgenic Mice Prove That SRY Is the Maleness Factor 96*

■ *Fast Forward: Visualizing X Chromosome Inactivation in Transgenic Mice 121*

chapter 5

Linkage, Recombination, and the Mapping of Genes on Chromosomes 133

- 5.1 Gene Linkage and Recombination 134
- 5.2 Recombination: A Result of Crossing-Over During Meiosis 138
- 5.3 Mapping: Locating Genes Along a Chromosome 143
- 5.4 The Chi-Square Test and Linkage Analysis 153
- 5.5 Tetrad Analysis in Fungi 156
- 5.6 Mitotic Recombination and Genetic Mosaics 164

■ *Fast Forward: Mapping the Crossovers that Generate the Chromosomes of Individual Human Sperm 150*

■ *Fast Forward: Gene Mapping May Lead to a Cure for Cystic Fibrosis 152*

■ *Tools of Genetics: The Chi-Square Test for Goodness of Fit 155*

■ *Genetics and Society: Mitotic Recombination and Cancer Formation 167*

PART II



What Genes Are and What They Do 181

© Adrian Neal/Getty Images RF

chapter 6

DNA Structure, Replication, and Recombination 181

- 6.1 Experimental Evidence for DNA as the Genetic Material 182
- 6.2 The Watson and Crick Double Helix Model of DNA 187
- 6.3 Genetic Information in Nucleotide Sequence 193
- 6.4 DNA Replication 194
- 6.5 Homologous Recombination at the DNA Level 201
- 6.6 Site-Specific Recombination 209

chapter 7

Anatomy and Function of a Gene: Dissection Through Mutation 219

- 7.1 Mutations: Primary Tools of Genetic Analysis 220
- 7.2 Molecular Mechanisms That Alter DNA Sequence 225
- 7.3 DNA Repair Mechanisms 234
- 7.4 What Mutations Tell Us About Gene Structure 239
- 7.5 What Mutations Tell Us About Gene Function 247
- 7.6 A Comprehensive Example: Mutations That Affect Vision 255

■ *Fast Forward: Trinucleotide Repeat Diseases: Huntington Disease and Fragile X Syndrome 229*

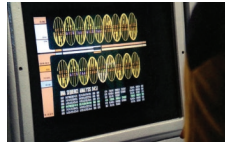
chapter 8

Gene Expression: The Flow of Information from DNA to RNA to Protein 270

- 8.1 The Genetic Code 271
- 8.2 Transcription: From DNA to RNA 280
- 8.3 Translation: From mRNA to Protein 289
- 8.4 Differences in Gene Expression Between Prokaryotes and Eukaryotes 297
- 8.5 The Effects of Mutations on Gene Expression and Function 299

■ *Genetics and Society: HIV and Reverse Transcription 284*

PART III



Analysis of Genetic Information 316

© CBS Photo Archive/Getty Images

chapter 9

Digital Analysis of DNA 316

- 9.1 Fragmenting DNA 317
- 9.2 Cloning DNA Fragments 322
- 9.3 Sequencing DNA 327
- 9.4 Sequencing Genomes 330

■ *Tools of Genetics: Serendipity in Science: The Discovery of Restriction Enzymes 319*

chapter 10

Genome Annotation 341

- 10.1 Finding the Genes in Genomes 342
- 10.2 Genome Architecture and Evolution 347
- 10.3 Bioinformatics: Information Technology and Genomes 354
- 10.4 A Comprehensive Example: The Hemoglobin Genes 355

chapter 11

Analyzing Genomic Variation 365

- 11.1 Variation Among Genomes 366
- 11.2 Genotyping a Known Disease-Causing Mutation 371
- 11.3 Sampling DNA Variation in a Genome 376
- 11.4 Positional Cloning 380
- 11.5 The Era of Whole-Genome Sequencing 386

■ *Tools of Genetics: The Lod Score Statistic 384*

PART IV



(left) © Texas A&M University/AP Photo;
(right) © Alpha/ZUMAPRESS/Newscom

How Genes Travel on Chromosomes 406

chapter 12

The Eukaryotic Chromosome 406

- 12.1 Chromosomal DNA and Proteins 407
- 12.2 Chromosome Structure and Compaction 408

- 12.3 Chromosomal Packaging and Gene Expression 414
- 12.4 Replication of Eukaryotic Chromosomes 419
- 12.5 Chromosome Segregation 423
- 12.6 Artificial Chromosomes 426

chapter 13

Chromosomal Rearrangements and Changes in Chromosome Number 436

- 13.1 Rearrangements of Chromosomal DNA 437
- 13.2 The Effects of Rearrangements 443
- 13.3 Transposable Genetic Elements 453
- 13.4 Aberrations in Chromosome Number: Aneuploidy 460
- 13.5 Variation in Number of Chromosome Sets: Euploidy 465
- 13.6 Genome Restructuring and Evolution 470

■ **Fast Forward: Programmed DNA Rearrangements and the Immune System 440**

chapter 14

Bacterial Genetics 487

- 14.1 The Enormous Diversity of Bacteria 488
- 14.2 Bacterial Genomes 490
- 14.3 Bacteria as Experimental Organisms 495
- 14.4 Gene Transfer in Bacteria 497
- 14.5 Using Genetics to Study Bacterial Life 509
- 14.6 A Comprehensive Example: How *N. gonorrhoeae* Became Resistant to Penicillin 511

■ **Genetics and Society: The Human Microbiome Project 494**

chapter 15

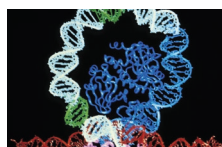
Organellar Inheritance 521

- 15.1 Mitochondria and Their Genomes 522
- 15.2 Chloroplasts and Their Genomes 525
- 15.3 The Relationship Between Organellar and Nuclear Genomes 528
- 15.4 Non-Mendelian Inheritance of Mitochondria and Chloroplasts 530
- 15.5 Mutant Mitochondria and Human Disease 535

■ **Genetics and Society: Mitochondrial DNA Tests as Evidence of Kinship in Argentine Courts 532**

■ **Fast Forward: Mitochondrial Eve 536**

PART V



© SPL/Science Source

How Genes Are Regulated 547

chapter 16

Gene Regulation in Prokaryotes 547

- 16.1 The Elements of Prokaryotic Gene Expression 548
- 16.2 Regulation of Transcription Initiation via DNA-Binding Proteins 549
- 16.3 RNA-Mediated Mechanisms of Gene Regulation 562
- 16.4 Discovering and Manipulating Bacterial Gene Regulatory Mechanisms 565
- 16.5 A Comprehensive Example: Control of Bioluminescence by Quorum Sensing 570

chapter 17

Gene Regulation in Eukaryotes 583

- 17.1 Overview of Eukaryotic Gene Regulation 584
- 17.2 Control of Transcription Initiation Through Enhancers 585
- 17.3 Epigenetics 594
- 17.4 Regulation After Transcription 600
- 17.5 A Comprehensive Example: Sex Determination in *Drosophila* 605

■ **Tools of Genetics: The Gal4/UAS_G Binary Gene Expression System 592**

PART VI



© Michael Goldberg, Cornell University, Ithaca, NY

Using Genetics 618

chapter 18

Manipulating the Genomes of Eukaryotes 618

- 18.1 Creating Transgenic Organisms 619
- 18.2 Uses of Transgenic Organisms 622

- 18.3 Targeted Mutagenesis 627
- 18.4 Human Gene Therapy 634
- *Tools of Genetics: Cloning by Somatic Cell Nuclear Transfer* 625
- *Tools of Genetics: How Bacteria Vaccinate Themselves Against Viral Infections with CRISPR/Cas9* 633
- *Genetics and Society: Should We Alter Human Germ-Line Genomes?* 637

chapter 19

The Genetic Analysis of Development 646

- 19.1 Model Organisms: Prototypes for Developmental Genetics 647
- 19.2 Mutagenesis Screens 649
- 19.3 Determining Where and When Genes Act 654
- 19.4 Ordering Genes in a Pathway 657
- 19.5 A Comprehensive Example: Body Plan Development in *Drosophila* 659

chapter 20

The Genetics of Cancer 681

- 20.1 Characteristics of Cancer Cells 682
- 20.2 The Genetic Basis of Cancers 684
- 20.3 How Cell Division Is Normally Controlled 687
- 20.4 How Mutations Cause Cancer Phenotypes 694
- 20.5 Personalized Cancer Treatment 699
- *Tools of Genetics: Analysis of Cell-Cycle Mutants in Yeast* 691

PART VII



© Sue Flood/Oxford Scientific/Getty Images

Beyond the Individual Gene and Genome 713

chapter 21

Variation and Selection in Populations 713

- 21.1 The Hardy-Weinberg Law: Predicting Genetic Variation in “Ideal” Populations 714
- 21.2 What Causes Allele Frequencies to Change in Real Populations? 721
- 21.3 Ancestry and the Evolution of Modern Humans 731

chapter 22

The Genetics of Complex Traits 746

- 22.1 Heritability: Genetic Versus Environmental Influences on Complex Traits 747
- 22.2 Mapping Quantitative Trait Loci (QTLs) 757
- *Tools of Genetics: The Chi-Square Test for Independence* 763

Guidelines for Gene Nomenclature A-1

Glossary G-1

Index I-1

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.

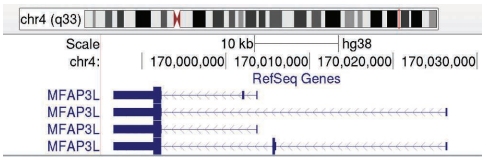
- **One Voice Genetics:** *Genes to Genomes* has a friendly, engaging reading style that helps students master the concepts throughout this book. The writing style provides the student with the focus and continuity required to make the book successful in the classroom.
- **Visualizing Genetics** The 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>

- Describe in approximate terms the genomic location of *MFAP3L*.
- Is the gene transcribed in the direction from the centromere-to-telomere or from the telomere-to-centromere?
- How many alternative splice forms of *MFAP3L* mRNA are indicated by the data?
- 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.
- 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.
- 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.
- 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 Mitochondrial 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.

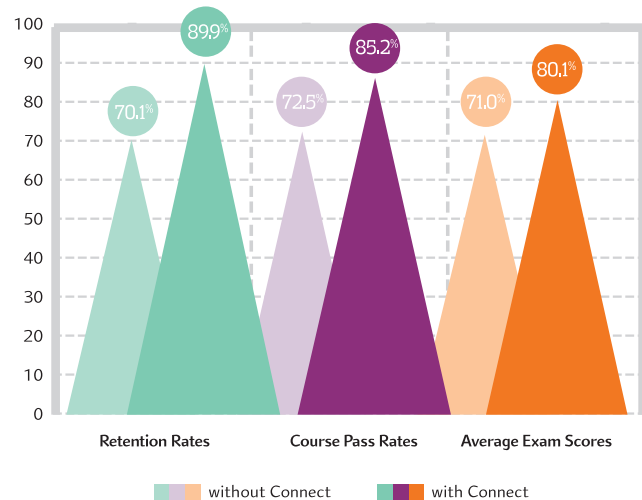
McGraw-Hill Connect® is a highly reliable, easy-to-use homework and learning management solution that utilizes learning science and award-winning adaptive tools to improve student results.

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.

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

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%**.

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.

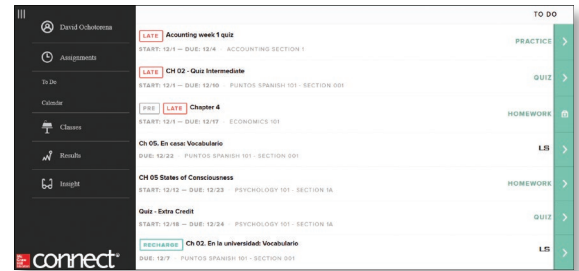


Robust Analytics and Reporting

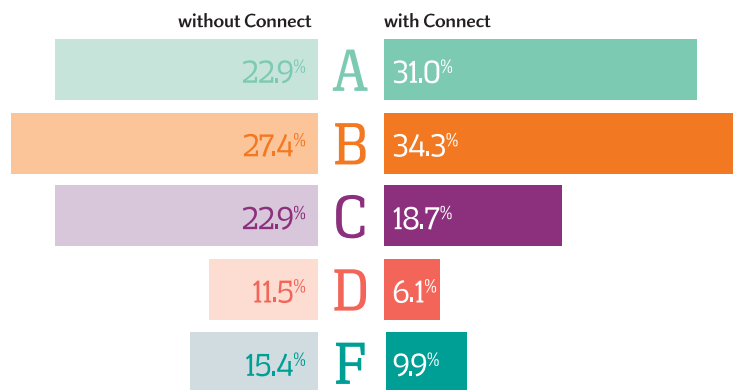
- Connect Insight® generates easy-to-read reports on individual students, the class as a whole, and on specific assignments.
- The Connect Insight dashboard delivers data on performance, study behavior, and effort. Instructors can quickly identify students who struggle and focus on material that the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

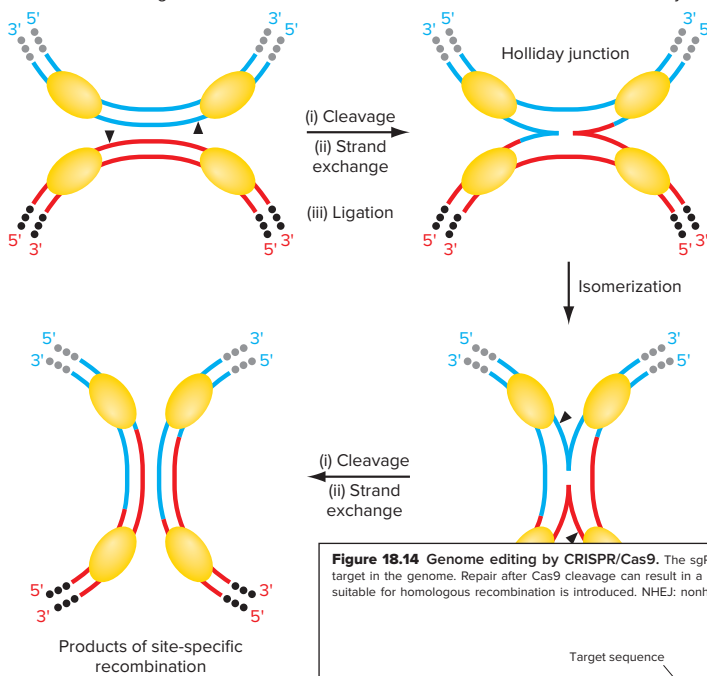
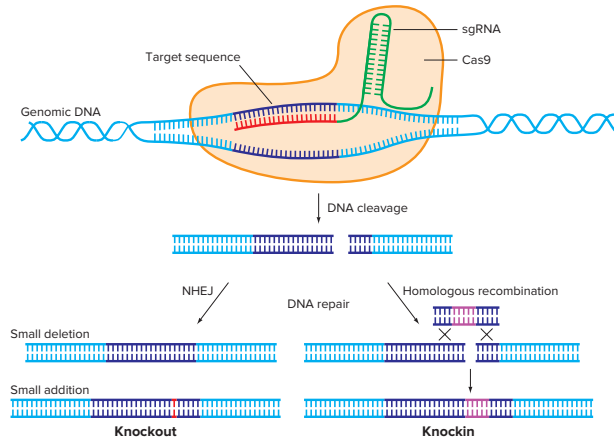


Figure 18.14 Genome editing by CRISPR/Cas9. The sgRNA sequence is designed to bring the Cas9 endonuclease to a specific target in the genome. Repair after Cas9 cleavage can result in a knockout or a knockin, depending on whether or not a DNA fragment suitable for homologous recombination is introduced. 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^M is inactivated (top) glow green; cells glow red (bottom) when X^M is inactivated.

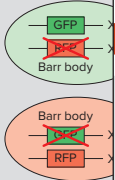
Figure B Heart cells of a transgenic mouse reveal a clonal patchwork of X inactivation. Patches of red or green cells represent cellular descendants of the founders that randomly inactivated one of their X chromosomes.

© Hao Wu and Jeremy Nathans, Molecular Biology and Genetics, Neuroscience, and HHMI, Johns Hopkins Medical School.

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



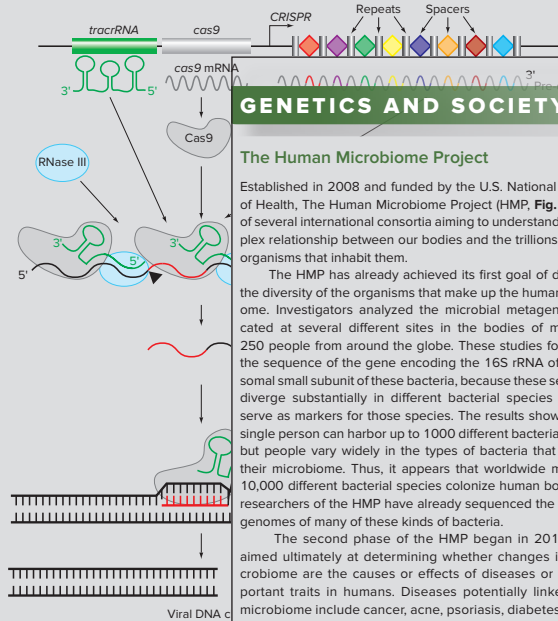
How Bacteria Vaccinate Themselves Against Viral Infections with CRISPR/Cas9

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.



GENETICS AND SOCIETY

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 ex-

Tools of Genetics Essays

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

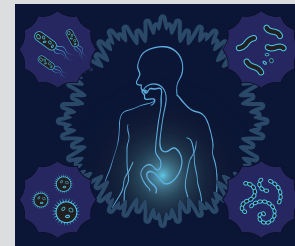


Figure A
© Arina Smitnova/Alamy RF

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.

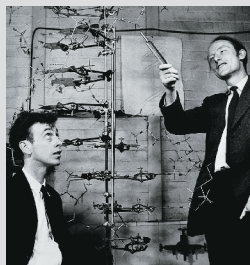
Visualizing Genetics

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

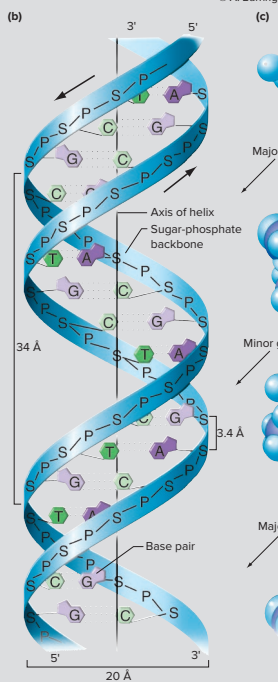
FEATURE FIGURE 6.11

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**).



(a) © A. Barrington Brown/Science Source



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.

Feature Figures

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

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.

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

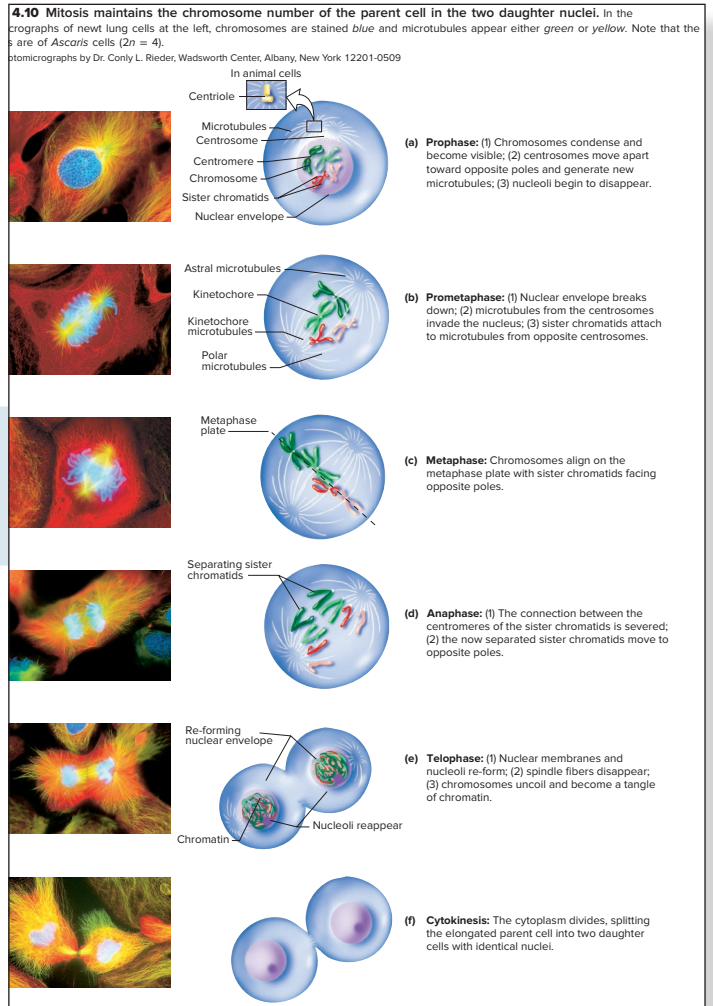
(5): © Lee Silver, Princeton University

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 *EcoRI* in lane A and with *RsaI* in lane B. Smears containing hundreds of thousands of fragments are produced with an average size of about 4.1 kb for *EcoRI* and 256 bp for *RsaI*. In C, D, and E, the chromosome of bacteriophage λ was cut with *HindIII*, *EcoRI*, and *RsaI*, 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.

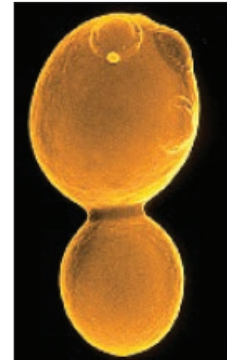
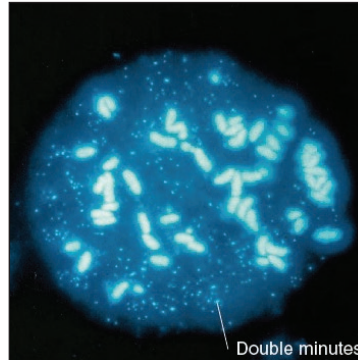
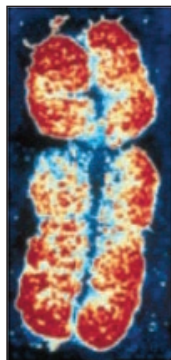
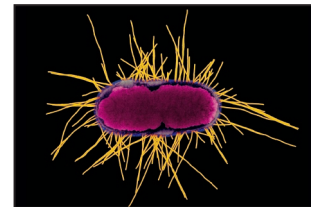
Process Figures

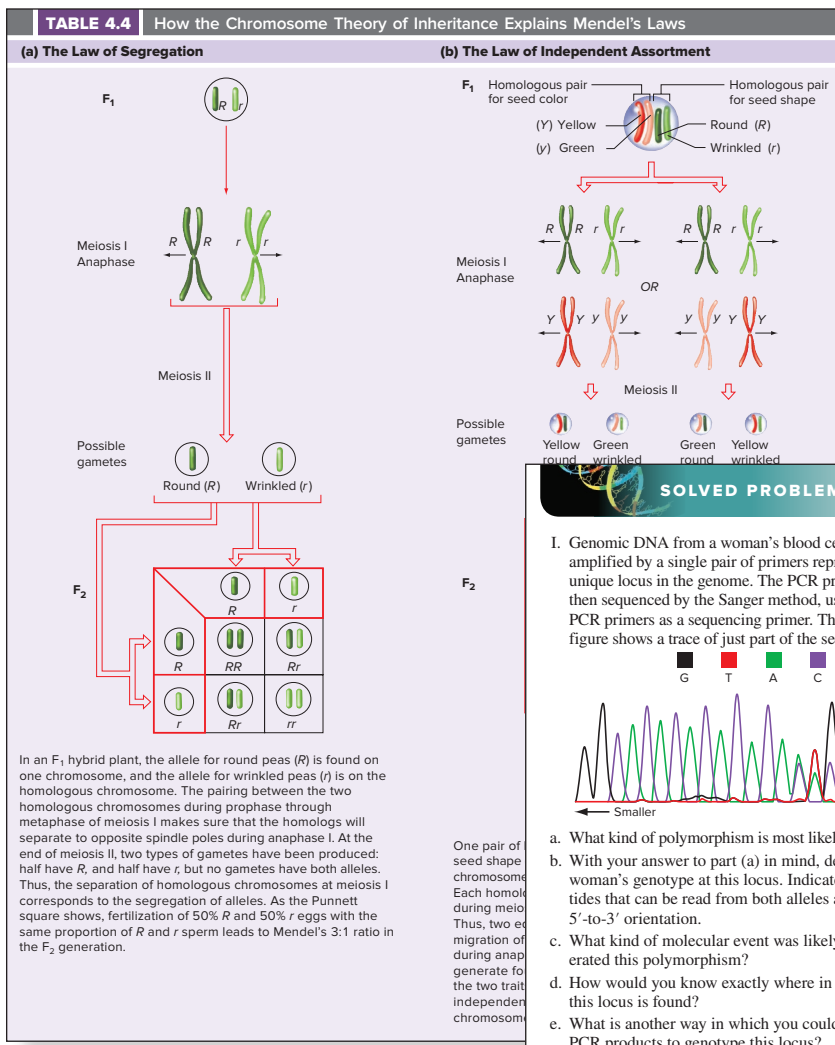
Step-by-step descriptions allow the student to walk through a compact summary of important details.



Micrographs

Stunning micrographs bring the genetics world to life.





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

SOLVED PROBLEMS

I. Genomic DNA from a woman's blood cells is PCR amplified by a single pair of primers representing a unique locus in the genome. The PCR products are then sequenced by the Sanger method, using one of the PCR primers as a sequencing primer. The following figure shows a trace of just part of the sequence read.

- What kind of polymorphism is most likely represented?
- With your answer to part (a) in mind, determine the woman's genotype at this locus. Indicate all nucleotides that can be read from both alleles and their 5'-to-3' orientation.
- What kind of molecular event was likely to have generated this polymorphism?
- How would you know exactly where in the genome this locus is found?
- What is another way in which you could analyze the PCR products to genotype this locus?
- Suppose you wanted to genotype this locus based on single-molecule DNA sequencing of whole genomes as shown in Fig. 9.24. Would a single read suffice for genotyping the locus by this alternative method?

Answer
To solve this problem, you need to understand that PCR will simultaneously amplify both copies of a locus (one on the maternally derived chromosome and one on the paternally derived chromosome), as long as the primer can hybridize to both homologs as is usually the case. The DNA sequence trace has two nucleotides at several positions. This fact indicates that the woman must be a heterozygote and that the PCR is amplifying both alleles of the locus.

- Notice that both alleles contain multiple repeats of the dinucleotide CA. The most likely explanation for the polymorphism is therefore that the locus contains an SSR polymorphism whose alleles have different numbers of CA repeats. One allele has six repeats; the second allele must have more CA units.
- Writing out the first 14 nucleotides of both alleles is straightforward. If the assumption in part (a) is correct, then one allele should have more than six CA repeats. The trace shows evidence for two additional CA repeats in one allele at positions 15–18, for a total of eight CA repeats.
You can then determine the nucleotides beyond the repeats in the shorter allele by subtracting CACA from positions 15–18. The remaining peaks at these positions correspond to ATGT. Note that ATGT can also be found in the longer allele, but now at nucleotides 19–22, just past the two additional CACA repeats. You can determine the last four nucleotides in the shorter allele by subtracting ATGT from positions 19–22, revealing TAGG. The sequences of the two alleles of this SSR locus (indicating only one strand of DNA each) are thus:
Allele 1: 5'...GGCACACACACACAATGTTAGG...3'
Allele 2: 5'...GGCACACACACACACAATGT...3'
- The mechanism thought to be responsible for most SSR polymorphisms is stuttering of DNA polymerase during DNA replication.

DNA: © Design Pics/Bilderbuch RF

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.

© James Strachan/Getty Images

chapter outline

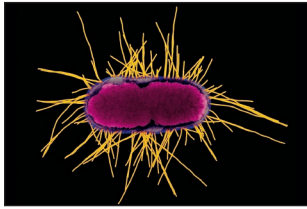
- 1.1 DNA: Life's Fundamental Information Molecule
- 1.2 Proteins: The Functional Molecules of Life Processes
- 1.3 Molecular Similarities of All Life-Forms
- 1.4 The Modular Construction of Genomes
- 1.5 Modern Genetic Techniques
- 1.6 Human Genetics and Society

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.

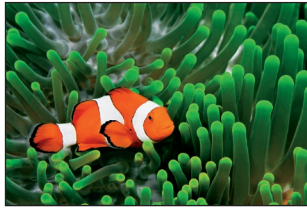
This book introduces you to the field of genetics as currently practiced in the early twenty-first century. Several broad themes recur throughout this presentation. First, we know that biological information is encoded in *DNA*, and that the *proteins* responsible for an organism's many functions are 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): © Kwangshin Kim/Science Source; (b): © Frank & Joyce Burek/Getty Images RF; (c): © Carl D. Walsh/Getty Images RF; (d): © Brand X Pictures/PunchStock RF; (e): © H. Wiesenhofer/PhotoLink/Getty Images RF; (f): © Ingram Publishing RF; (g): Source: Carey James Balboa. https://en.wikipedia.org/wiki/File:Red_eyed_tree_frog_edit2.jpg; (h): © Digital Vision RF



(a) Bacteria



(b) Clown fish



(c) Lion



(d) Oak tree



(e) Poppies



(f) Hummingbird



(g) Red-eyed tree frog



(h) Humans

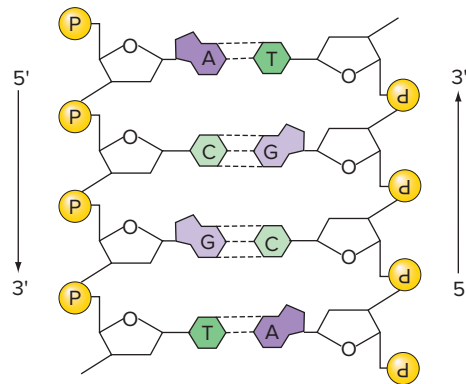
1.1 DNA: Life's Fundamental Information Molecule

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**—**g**uanine, **c**ytosine, **a**denine, and **t**hymine—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.

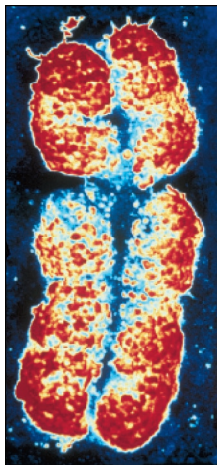
Figure 1.2 Complementary base pairs are a key feature of the DNA molecule. A single strand of DNA is composed of nucleotide subunits each consisting of a deoxyribose sugar (*white pentagons*), a phosphate (*yellow circles*), and one of four nitrogenous bases—adenine, thymine, cytosine, or guanine (designated as *lavender* or *green As, Ts, Cs, or Gs*). 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.

© Biophoto Associates/Science Source



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

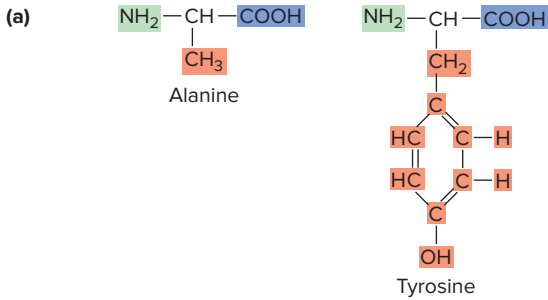
1. Compare the chemical structures of DNA and proteins.
2. Differentiate between the functions of DNA and the functions of proteins.

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 ($-\text{NH}_2$; green) at one end and an acidic carboxyl group ($-\text{COOH}$; 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 β 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.



(b)

Hemoglobin β chain

MVHLTPEEKSAVTALWGKVVNDEVGEALGRLLVVPWTQRLFESFGDLFTPDVAVMGNPKVKAHG
KKVLGAFSDGPAHLDNLKGFTATLSELHCDKLVDPENFRLLGNLVLCVLAHFFGKEFTPPVQAA
YQKVVAGVANALAHKYH

Lactate dehydrogenase

MATIKSELIKNFEEEEIHHNKISIVGTGSVGVACASILLKGLSDELVLVDVDEGKLGKETMDL
QHGSFPMKMPNIVSSKDYLTANSNLVITAGARQKGETRLDLVQRNVSIFKLMIPNITQYSPH
CKLLVITNPVDILTYAWKLSGFKNRVIGSGCNLDSARFRYFIGQRLGIHSESCHLILGEHGD
SSVPVWGSVNIAGVPLKDLNPDIGTDKDEQWENVHKKVISSGYEMVKMGYTSWGISLVSADLT
ESILKNLRRVHPYSTLSKGLYGINEDIFLSVPCILGENGITDLIKVKLTLEEEACLQKSAETLWEIQKELK

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).