

# Contents

Preface xix

Acknowledgments xxvii

---

## PART I

### Molecular Basis of Heredity: An Overview 1

#### Perspective I 1

Chromosomes 3

    The Cell Cycle 3

    Meiosis and the Formation of Gametes 3

    Chromosome Structure 11

The Inheritance of Single Traits 17

    Independent Segregation and Independent Assortment 17

    The Relation Between Genes and Chromosomes 19

    Recombination 19

    The Relation Between Genes and Proteins 20

Genes and DNA 23

Genetics—A Molecular Science 23

    The Intracellular Flow of Genetic Information 25

    The Structure and Maintenance of DNA Genomes 27

    Expression and Regulation of a Genome's Phenotype 29

**Chapter 1**  
**The Genetic Molecules** 35

- 1.1 Structure and Behavior of DNA 36
  - a. Constituents and Chemical Linkages in DNA 36
  - b. The Double Helical Structure of DNA 40
  - c. Alternative Forms of the DNA Double Helix 43
  - d. The Size of DNA Molecules 43
  - e. Variations in DNA Shape and Strandedness 44
  - f. Denaturation and Renaturation of DNA 47
  - g. Packaging of DNA in Chromosomes 49
- 1.2 Structure and Behavior of RNA 54
  - a. Types of RNA and Their Occurrence 54
  - b. Constituents and Chemical Linkages in RNA 54
  - c. RNA Structure 55
  - d. RNA Denaturation and Renaturation 58
  - e. RNA-DNA Hybrid Helices 58
- 1.3 Structure of Proteins 59
  - a. Constituents and Chemical Linkages in Proteins 59
  - b. Protein Size and Shape 62
  - c. Determinants of Protein Conformation 64

**Chapter 2**  
**Replication, Maintenance, and Modification of the Genome** 73

- 2.1 DNA Replication 74
  - a. DNA's Template Function During Replication 74
  - b. Replication Initiates at Discrete Locations 75
  - c. DNA Replication Is Semiconservative 80
  - d. Complementary Base Copying, Deoxynucleotidyl Transfer, and DNA Ligation in DNA Replication 84
  - e. Key Enzymes in DNA Synthesis 85
  - f. Replication Requires Unwinding the Helix 89
  - g. Initiation of New DNA Chains and Their Extension at Replication Forks 94
  - h. Termination of DNA Replication and Resolution of the Daughter Helices 102
- 2.2 Replication of RNA into DNA 103
  - a. Replication of Retroviral Genomes 103
  - b. Some DNA Viruses Use Reverse Transcription for Replication 107
- 2.3 DNA Repair 107
  - a. Repair by Reversing the Modification 108
  - b. Repair by Replacing Modified Residue(s) 111
  - c. The Importance of DNA Repair 112

- 2.4 DNA Recombination 113
  - a. Types of Recombination 114
  - b. General Recombination Between Homologous DNAs 118
  - c. Necessary Enzymes in General Recombination 122
  - d. Site-Specific Recombination 122
- 2.5 RNA Replication 124

## Chapter 3

### The Logic and Machinery of Gene Expression 129

- 3.1 A Synopsis of Gene Expression 131
  - a. Transcription of DNA into RNA 131
  - b. Relation of Nucleotide Triplets to Amino Acids 131
  - c. Recognition of Codons by tRNAs 132
  - d. Proper Initiation of Translation 132
  - e. Codon Translation and Amino Acid Assembly 133
  - f. Regulation of Gene Expression at Various Stages of RNA and Protein Formation 133
- 3.2 Transcription: The Transfer of DNA Sequence Information to RNA 134
  - a. Copying DNA Sequences into RNA 135
  - b. DNA-Dependent RNA Polymerases 138
  - c. Transcription Initiates at Characteristic Nucleotide Sequences 140
  - d. Termination and Release of RNA Chains 142
- 3.3 RNA Processing in Prokaryotes 145
  - a. The Grouping of rRNA and tRNA Genes 145
  - b. Trimming the rRNA-tRNA Cotranscripts 145
  - c. Producing Mature tRNAs from Larger Transcripts 147
- 3.4 The Genetic Code 149
  - a. Amino Acid Sequences in Proteins Correspond to Nucleotide Sequences in Genes 150
  - b. Matching Amino Acids to Their Codons 151
  - c. Solving the Genetic Code 153
  - d. Redundancy in the Genetic Code 156
  - e. Universality of the Genetic Code 157
- 3.5 The Translation Machinery 157
  - a. Attachment of Amino Acids to Cognate tRNAs 158
  - b. Ribosomes Match Aminoacyl-tRNAs to Codons and Assemble Protein Chains 163
- 3.6 mRNA Translation in Prokaryotes 167
  - a. Requirements for Initiation 168
  - b. Polypeptide Chain Elongation 169
  - c. Termination of Polypeptide Chain Elongation 173

- 3.7 Some Notable Features of the Translation Process 173
  - a. Simultaneous Translation of mRNAs by More Than One Ribosome 174
  - b. Translation of Bacterial mRNAs Can Occur During Transcription 174
  - c. Ribosomes Recycle After Translating a Coding Sequence 174
  - d. Codon-Anticodon Interactions 175
- 3.8 mRNA Translation in Eukaryotes 180
  - a. Special Modifications in Eukaryotic mRNAs 181
  - b. Initiation of Translation by Small Ribosomal Subunits at the 5' Capped Ends of mRNAs 181
  - c. Polypeptide Chain Elongation and Termination 182
- 3.9 Inhibitors of Transcription and Translation 183
  - a. Inhibition of RNA Polymerase 183
  - b. Inhibition of Translation 183
- 3.10 The Fate of Newly Synthesized Proteins 186
  - a. Posttranslational Alterations in Polypeptide Chains 187
  - b. Targeting of Eukaryotic Proteins into and Through Cell Membranes 188
  - c. Transport of Proteins to Eukaryotic Cellular Organelles 192
  - d. Transport of Proteins in Prokaryotes 194
- 3.11 Regulation of Gene Expression 196
  - a. Regulation of RNA Levels During Biosynthesis 196
  - b. Coordinate Regulation of Prokaryotic Gene Expression 197
  - c. Regulation of Lactose Operon Expression 199
  - d. Regulation of Tryptophan Operon Expression 202
  - e. Temporal Control of Gene Expression in the Life Cycle of Bacteriophage  $\lambda$  207
  - f. Translational Regulation of the Expression of Some Gene Products 210

References for Part I 215

---

## PART II The Recombinant DNA Breakthrough 223

Perspective II 223

- Introducing New Genetic Information into Bacteria 225
  - Bacterial Transformation 225
  - Conjugation 227
  - Transduction 227

- The Principle of Cloning 231
- The Recombinant DNA Concept 233
- Important Discoveries 234
  - Bacterial Plasmids 234
  - Restriction Endonucleases 238

## Chapter 4

### The Tools: Enzymes 243

- 4.1 Nucleases 244
  - a. General Properties 244
  - b. Single Strand Specific Nucleases 245
  - c. *Bal 31* Nuclease 246
  - d. RNase H 246
- 4.2 The Restriction Endonucleases 246
  - a. Three Types of Restriction Endonucleases 247
  - b. A Typical Type II Restriction Endonuclease 248
  - c. Several Groups of Type II Restriction Endonucleases 250
  - d. Mapping DNA Segments with Type II Restriction Endonucleases 251
  - e. Protection by Methylation 253
- 4.3 Phosphomonoesterases 254
- 4.4 Polynucleotide Kinase 255
- 4.5 DNA Ligase 255
- 4.6 DNA Polymerase I 256
  - a. A Versatile Enzyme 256
  - b. Nick Translation 256
  - c. Filling In 257
- 4.7 RNA-Dependent DNA Polymerases (Reverse Transcriptases) 257
- 4.8 Terminal Deoxynucleotidyl Transferase 258
  - a. Polymerization Without a Template 258
  - b. Synthesizing Cohesive Ends 259
- 4.9 Poly A Polymerase 260

## Chapter 5

### The Tools: Host-Vector Systems 261

- 5.1 *E. Coli* Systems—the Host Cells 263
  - a. A Versatile Host 263
  - b. A Hospitable Host 263
  - c. An Accessible Host 264
  - d. Some Examples 264

- 5.2 *E. Coli* Systems—Plasmid Vectors 265
  - a. The Modular Structures of Plasmids 265
  - b. Designing Vectors for Selection 268
  - c. A Plasmid Vector—pBR322 272
  - d. Different Vectors for Different Purposes 275
- 5.3 *E. Coli* Systems—Bacteriophage Vectors 276
  - a. Some Differences Between Plasmid and Phage Vectors 276
  - b.  $\lambda$  Phage 276
  - c.  $\lambda$  Phage Vectors 278
  - d. Packaging  $\lambda$  Vectors into Phage Particles 279
  - e. Phage M13 280
  - f. M13 Vectors 283
- 5.4 *E. Coli* Systems—Plasmid-Phage Combination Vectors 286
  - a. Cosmids 286
  - b. Phasmids 286
- 5.5 Other Prokaryotic Host-Vector Systems 288
  - a. Gram-Negative Organisms 289
  - b. Gram-Positive Organisms 289
  - c. Shuttle Vectors 291
- 5.6 Eukaryotic Host-Vector Systems: Yeast 291
  - a. Versatility and Convenience 291
  - b. Vectors That Replicate in Yeast 293
  - c. Permanent Transformation by Recombination with the Yeast Genome 295
- 5.7 Eukaryotic Host-Vector Systems: Animals 296
  - a. The Transformation of Animal Cells 296
  - b. SV40 Vectors 302
  - c. Bovine Papillomavirus Vectors 307
  - d. Retrovirus Vectors 310
- 5.8 Eukaryotic Host-Vector Systems: Plants 314
  - a. General Considerations 314
  - b. pTi—A Tumor-Inducing Plasmid 315
  - c. Designing Recombinant DNA Vectors with pTi 318

## Chapter 6

### The Means: Constructing, Cloning, and Selecting Recombinant DNA 321

- 6.1 Inserts 321
  - a. General Considerations 321
  - b. Inserts from Genomic DNA 322
  - c. Synthetic Inserts 325
  - d. Copying RNA into DNA 332

|     |   |     |
|-----|---|-----|
| 6.2 | Ligating Vector to Insert                               | 334 |
| a.  | Joining Ends  | 335 |
| b.  | Attaching Cohesive Ends                                 | 338 |
| 6.3 | Infection, Transfection, and Cloning                    | 339 |
| a.  | Moving Recombinant Molecules from Test Tube to Cell     | 339 |
| b.  | Cloning   | 340 |
| 6.4 | Screening Cloned Populations of Recombinants            | 341 |
| a.  | Finding the Right Clone                                 | 341 |
| b.  | Annealing with a Complementary Polynucleotide           | 341 |
| c.  | Testing for Gene Expression in Cells                    | 347 |
| 6.5 | Libraries   | 352 |
| a.  | Genomic Libraries                                       | 354 |
| b.  | cDNA Libraries  | 356 |
| 6.6 | Examples of Strategies Used for Cloning Genes and cDNAs | 358 |

## Chapter 7

### The Products: Characterizing and Manipulating Recombinants 369

|     |  |     |
|-----|--|-----|
| 7.1 | The Gross Anatomy of a Cloned Insert                           | 370 |
| a.  | Insert Size  | 370 |
| b.  | Mapping Restriction Endonuclease Sites                         | 371 |
| c.  | Subcloning   | 371 |
| d.  | Locating the Segment of Interest in the Insert                 | 372 |
| 7.2 | The Fine Anatomy of a DNA Segment—Primary Nucleotide Sequence  | 373 |
| a.  | General Principles   | 374 |
| b.  | Chemical Sequencing  | 375 |
| c.  | Sequencing by Enzymatic Copying                                | 379 |
| 7.3 | Computer Analysis of DNA Sequences                             | 383 |
| a.  | Storing Primary Sequence Data                                  | 383 |
| b.  | Structural Analysis  | 384 |
| c.  | Biological Significance  | 386 |
| 7.4 | Locating Cloned Segments in Genomes                            | 388 |
| a.  | Molecular Location   | 388 |
| b.  | Chromosomal Location   | 391 |
| c.  | The Fidelity of Cloning  | 395 |
| 7.5 | Determining the Number of Copies of a DNA Sequence in a Genome | 396 |
| a.  | Estimating Copy Number Using DNA Blots                         | 396 |
| b.  | Estimating Copy Number from the Kinetics of DNA Annealing      | 396 |
| c.  | Estimating Copy Number by Annealing to Saturation              | 400 |

|     |   |     |
|-----|---|-----|
| 7.6 | Altering Cloned Segments: Constructing Mutants                                | 400 |
| a.  | General Considerations  | 400 |
| b.  | Deletion Mutants  | 401 |
| c.  | Insertion Mutants   | 403 |
| d.  | Point Mutations   | 404 |
| 7.7 | Analyzing the Function of Cloned DNA Segments                                 | 407 |
| a.  | Characterizing Intracellular Transcripts Corresponding to Cloned DNA Segments | 408 |
| b.  | Testing Cloned DNAs for Function  | 411 |
| 7.8 | Synthesizing Polypeptides Encoded by Cloned Eukaryotic DNA Segments           | 413 |
| a.  | The Choice of an Expression System  | 413 |
| b.  | Expression Vectors Used in <i>E. Coli</i>                                     | 415 |
| c.  | Expression Vectors Used in Yeast  | 418 |
| d.  | Expression Vectors for Use in Animal Cells                                    | 419 |
| 7.9 | Enzymatic Amplification of DNA and RNA Segments                               | 420 |

|                        |     |
|------------------------|-----|
| References for Part II | 426 |
|------------------------|-----|

---

## PART III The Molecular Anatomy, Expression, and Regulation of Eukaryotic Genes 433

### Perspective III 433

#### The Structure and Expression of Eukaryotic Genes 435

Interrupted Genes 436

Regulation of Gene Expression in Eukaryotes 437

#### The Structure of Eukaryotic Genomes 441

Chromosome Maps 441

The Size of Eukaryotic Genomes 444

The Relation Between Particular DNA Sequences and Chromosome Morphology 445

#### The Rearrangement of DNA Sequences 445

Random Genomic Rearrangements 446

Programmed Rearrangements 447

#### Evolutionary Implications 447

Comparative Molecular Genetics 449

The Evolutionary History of Genes 450

The Origin of Genetic Systems 452

#### Conclusion 454

**Chapter 8****The Structure and Regulated Expression of Eukaryotic Genes 455**

- 8.1 Comparative Structural Features of Prokaryotic and Eukaryotic Genes 457
  - a. Prokaryotic Genes 457
  - b. Eukaryotic Genes 459
- 8.2 Structure and Expression of Class I Genes 463
  - a. Transcription Units 463
  - b. Transcription Machinery 465
  - c. rDNA Transcription Control Regions 467
  - d. rDNA Transcription Termination 474
  - e. Termination Promotes Transcription Initiation 476
- 8.3 Structure and Function of Class II Genes 478
  - a. General Considerations 478
  - b. Transcription Machinery 479
  - c. Maturation of mRNAs and U RNAs 481
  - d. Regulated Expression of Viral Genes 489
  - e. Tissue- and Stage-Specific Gene Regulation 504
  - f. Inducible and Repressible Transcription 514
  - g. Transcriptional Regulation of Morphogenesis 532
  - h. Transcriptional Regulation of U RNA Genes 537
- 8.4 Structure and Expression of Class III Genes 539
  - a. General Features 539
  - b. Sequences Regulating Transcription Initiation 541
  - c. Transcription Factor Modulation of Transcription Initiation by RNA Polymerase III 546
  - d. Transcription Termination by RNA Polymerase III 551
  - e. Control of 5S rRNA Gene Expression During Development 552
- 8.5 Dealing with Introns 556
  - a. The Frequency of Introns 557
  - b. Different Kinds of Introns 557
  - c. Autocatalytically Spliced Introns 562
  - d. Splicing Nuclear Pre-mRNA Introns 568
  - e. Splicing tRNAs 576
  - f. Alternative Splicing: Multiple Proteins from One Gene 578
- 8.6 Novel Structural Motifs in Transcription Factors 583
  - a. DNA Binding Domains 584
  - b. Transcription Regulatory Domains 592
- 8.7 Global Influences on Gene Expression 595
  - a. DNA Packing 596
  - b. DNA Topology and Conformation 600
  - c. DNA Methylation 601
  - d. Regulated Utilization of mRNA 612

|   |   |
|---|---|
| Chapter 9                                       |   |
| The Molecular Anatomy of Eukaryotic Genomes 621 |   |
| 9.1   | Architectural Elements 622  |
| a.  | Several Classifications of DNA Segments 622   |
| b.  | The Repetition of DNA Sequences 624   |
| c.  | Evolutionary Implications of Repetition 627   |
| 9.2   | Genes Encoding RNA 631  |
| a.  | Genes for 18S, 5.8S, and 28S rRNA 631   |
| b.  | Genes for 5S rRNA 636   |
| c.  | The Linkage of all Four Yeast rRNA Genes 637  |
| d.  | Genes Encoding tRNAs 638  |
| e.  | Genes for Small Nuclear and Cytoplasmic RNAs 639                                      |
| 9.3   | Genes Encoding Polypeptides 641   |
| a.  | Some General Considerations 641   |
| b.  | Examples of Multigene Families 643  |
| c.  | Histone Genes: Conservation of Coding Sequences and Divergence of Organization 656    |
| 9.4   | Tandem Repetition of DNA Sequences: A Common Characteristic of Eukaryotic Genomes 659 |
| a.  | Genes Constructed of Tandem Repeats of DNA Segments 659                               |
| b.  | Tandem Repetitions Outside of Coding Regions 664                                      |
| c.  | Tandem Repetitions at Centromeres and Telomeres 666                                   |
| d.  | Speculation on the Function of Tandem Repeats 672                                     |
| e.  | Mechanisms for the Formation and Evolution of Tandem Repeats 674                      |
| 9.5   | Repeated Sequences Dispersed in Genomes 677   |
| a.  | Patterns of Interspersion 677   |
| b.  | Interspersed Repeats in Invertebrates 679   |
| c.  | The Very Large Families of Interspersed Repeats in Mammalian Genomes 680              |
| d.  | The Function of Interspersed Repeats 685  |
| e.  | The Remarkable Uniformity of Dispersed Repeated Sequences 687                         |
| 9.6   | Sequences at Centromeres and Telomeres 691  |
| a.  | Sequences at Centromeres 691  |
| b.  | Sequences at Telomeres 694  |
| c.  | Yeast Artificial Chromosomes 698  |
| 9.7   | Genomes of Eukaryotic Organelles: The DNA of Mitochondria and Chloroplasts 699        |
| a.  | Mitochondrial Genomes 700   |
| b.  | The Unusual Mitochondrial DNA of Trypanosomes 707                                     |
| c.  | The DNA of Chloroplasts 709   |
| d.  | The Origin of Organelle DNA 711   |

**Chapter 10****Genomic Rearrangements 713**

- 10.1 General Features of Unprogrammed Transpositions 714
  - a. Different Types of Mobile Elements 715
  - b. The Formation of Target Site Duplications 716
  - c. Mobile Elements as Dispersed Repeated Sequences 716
- 10.2 Transposable Elements 718
  - a. Prokaryotic Transposable Elements 718
  - b. The P Elements of *Drosophila* 727
  - c. The Controlling Elements of Maize 733
- 10.3 Retrotransposons 740
  - a. General Characteristics of Class I Retrotransposons 740
  - b. The Ty Elements of Yeast 741
  - c. The Copia-Like Elements of *Drosophila* 750
  - d. The IAP Sequences of Mice 752
  - e. Comparison with Retroviruses 753
  - f. Class II Retrotransposons 757
- 10.4 Retrogenes 762
  - a. Processed Polypeptide Pseudogenes 763
  - b. Processed RNA Pseudogenes 766
- 10.5 Other Unusual Moveable Elements 769
  - a. The Fold-Back Elements of *Drosophila* 769
  - b. Insertions in *Drosophila* rDNA 770
- 10.6 Programmed Rearrangements and the Modulation of Gene Expression 770
  - a. Prokaryotic Models—Translocation through Flip-Flop Inversions 770
  - b. Yeast Mating Types—A Cassette Mechanism 774
  - c. Genes Encoding Vertebrate Immune Proteins 783
  - d. The Variable Surface Antigens of Trypanosomes 799
- 10.7 Programmed Amplification and the Modulation of Gene Expression 808
  - a. Disproportionate Replication: The *Drosophila* Chorion Genes 808
  - b. *Xenopus* rDNA: Rolling Circle Amplification 811
  - c. Amplification of *Tetrahymena* rDNA in Macronuclei 811
- 10.8 Unprogrammed Tandem Amplifications 812
  - a. Amplification Corrects Enzyme Deficiencies 812
  - b. The Structure of Amplified Genes 815
  - c. Amplification Mechanism 818

**References for Part III 823**

---

## **PART IV Understanding and Manipulating Biological Systems 859**

### **Perspective IV 859**

|   |     |
|---|-----|
| Understanding Complex Biological Systems                                      | 860 |
| The Multigene Family for Visual Pigments                                      | 860 |
| Coordinating Interactions Between Many Gene Products                          | 862 |
| The Life Cycles of Viruses  | 863 |
| Mapping Genomes   | 873 |
| Physical and Genetic Maps   | 873 |
| Mapping Large Genomes   | 875 |
| Polymorphic Restriction Endonuclease Fragments                                | 876 |
| Differential Gene Expression in Specific Cells and Tissues                    | 877 |
| Differential Gene Expression in the Pancreas                                  | 879 |
| Posttranslational Modifications as Regulators of the Flow of<br>Gene Products | 881 |
| Control of Gene Expression by DNA Sequence<br>Rearrangement                   | 882 |
| Modifying Biological Systems  | 883 |
| Synthesizing Normal and Modified Eukaryotic Proteins                          | 884 |
| Genetic Engineering   | 887 |
| Modifying Somatic Cells   | 889 |
| Modifying Germ Cells  | 889 |
| New Directions  | 893 |
| Embryological Development and Differentiation of Complex<br>Organisms         | 893 |
| Insertional Mutagenesis   | 896 |
| Future Research   | 897 |
| Future Concerns   | 899 |
| References for Part IV  | 900 |
| Index   | 905 |