

CONTENTS

Preface xiii

Introduction xv

1 DNA Chemistry and Biology	1
Basic Properties of DNA	1
Covalent Structure	1
Double Helical Structure	1
Methylated Bases	5
Plasticity in DNA Structure	7
DNA Synthesis	8
DNA as a Flexible Set of Chemical Reagents	15
Basic DNA Biology	19
Genome Sizes	25
Number of Genes	27
Sources and Additional Readings	27
2 A Genome Overview at the Level of Chromosomes	29
Basic Properties of Chromosomes	29
Bacterial Chromosomes	29
Chromosomes of Eukaryotic Organisms	32
Centromeres	32
Telomeres	34
Dynamic Behavior of Telomeres	35
Chromatin and the Higher-Order Structure of Chromosomes	36
Chromosomes in the Cell Cycle	39
Genome Organization	40
Chromosome Purification	43
Chromosome Number	51
Unusual Characteristics of Sex Chromosomes and Mitochondria	54
Synteny	59
Sources and Additional Readings	63
3 Analysis of DNA Sequences by Hybridization	64
Basic Requirements for Selectivity and Sensitivity	64
Detection of Specific DNA Sequences	66
Equilibria between DNA Double and Single Strands	67
Thermodynamics of the Melting of Short Duplexes	71

Thermodynamics of Imperfectly Paired Duplexes	74
Kinetics of the Melting of Short Duplexes	77
Kinetics of Melting of Long DNA	79
Kinetics of Double-Strand Formation	80
Complexity	85
Hybridization on Filters	86
Sensitive Detection	90
Sources and Additional Readings	97

4 Polymerase Chain Reaction and Other Methods for In Vitro DNA Amplification 98

Why Amplify DNA?	98
Basic Principles of the Polymerase Chain Reaction (PCR)	98
Noise in PCR: Contamination	103
PCR Noise: Mispriming	104
Misincorporation	106
Long PCR	106
Incorporating Extra Functionalities	107
Single-Sided PCR	107
Reducing Complexity with PCR	112
Additional Variants of the Basic PCR Reaction	114
Total Genome Amplification Methods	116
Application of PCR to Detect Molecules Other Than DNA	119
DNA Amplification without Thermal Cycling and Other Alternatives to PCR	122
Future of PCR	127
Sources and Additional Readings	128

5 Principles of DNA Electrophoresis 131

Physical Fractionation of DNA	131
Separation of DNA in the Ultracentrifuge	131
Electrophoretic Size Separations of DNA	132
Electrophoresis without Gels	133
Motions of DNA Molecules in Gels	135
Complex Effects of Gel Structure and Behavior	136
Biased Reptation Model of DNA Behavior in Gels	138
Pulsed Field Gel Electrophoresis (PFG)	140
Macroscopic Behavior of DNA in PFG	146
Inadequacy of Reptation Models for PFG	148
DNA Trapping Electrophoresis	155
Secondary Pulsed Field Gel Electrophoresis (SPFG)	157
Entry of DNAs into Gels	158
Sources and Additional Readings	164

6 Genetic Analysis 165

Why We Need Genetics	165
Basic Strategy for Genetic Analysis in the Human:	
Linkage Mapping	165

A Glossary of Genetic Terms	170
Relationship between the Physical and the Genetic Maps	174
Power of Mouse Genetics	178
Weakness of Human Genetics	178
Linkage Analysis Ignoring Recombination	180
Linkage Analysis with Recombination	183
Interval Mapping	185
Finding Genes by Genetic Mapping	188
Moving from Weak Linkage Closer to a Gene	190
Linkage Disequilibrium	191
Complications in Linkage Disequilibrium and Genetic Maps in General	193
Distortions in the Genetic Map	194
Current State of the Human Genetic Map	195
Genetics in the Pseudoautosomal Region	197
Why Genetics Needs DNA Analysis	201
Detection of Homozygous Regions	204
Sources and Additional Readings	206
7 Cytogenetics and Pseudogenetics	208
Why Genetics Is Insufficient	208
Somatic Cell Genetics	208
Subchromosomal Mapping Panels	210
Radiation Hybrids	212
Single-Sperm PCR	215
In Situ Hybridization	218
High-Resolution FISH	224
Chromosome Painting	229
Chromosome Microdissection	230
Sources and Additional Readings	232
8 Physical Mapping	234
Why High Resolution Physical Maps Are Needed	234
Restriction Maps	235
Ordered Libraries	239
Restriction Nuclease Genomic Digests	241
HTF Islands	245
Ordering Restriction Fragments	246
Identifying the DNA Fragments Generated by a Rare-Cutting Restriction Enzyme	248
Mapping in Cases Where Fragment Lengths Can Be Measured Directly	252
Generation of Larger DNA Fragment Sizes	253
Linking Clones	254
Jumping Libraries	257
Partial Digestion	259
Exploiting DNA Polymorphisms to Assist Mapping	262
Placing Small Fragments on Maps	264

Reaching the Ends of the Physical Map: Cloning Telomeres	265
Optical Mapping	269
Bottom-Up Library Ordering	269
Measurements of Progress in Building Ordered Libraries	275
Survey of Restriction Map and Ordered Library Construction	277
Sources and Additional Readings	284
9 Enhanced Methods for Physical Mapping	285
Why Better Mapping Methods Are Needed	285
Larger Yeast Artificial Chromosomes (YACs)	285
How Far Can YACs Go?	288
Vector Obsolescence	290
Hybrid Mapping Strategies: Cross-connections between Libraries	291
Screening by PCR versus Hybridization	296
Tiered Sets of Samples	298
Simple Pooling Strategies for Finding a Clone of Interest	300
Sequence-Specific Tags	301
Pooling in Mapping Strategies	303
Probe Pooling in <i>S. pombe</i> Mapping	305
False Positives with Simple Pooling Schemes	311
More General Pooling Schemes	312
Alternate Array Configurations	316
Inner Product Mapping	318
Sliced PFG Fractionations as Natural Pools of Samples	320
Restriction Landmark Genome Scanning	320
Prognosis for the Future of Genome Mapping	322
Sources and Additional Readings	323
10 DNA Sequencing: Current Tactics	325
Why Determine DNA Sequence	325
Design of DNA Sequencing Projects	326
Ladder Sequencing Tactics	327
Issues in Ladder Sequencing	330
Current Fluorescent DNA Sequencing	334
Variations in Contemporary DNA Sequencing Tactics	336
Errors in DNA Sequencing	341
Automated DNA Sequencing Chemistry	345
Future Improvements in Ladder Sequencing	348
Approaches to DNA Sequencing by Mass Spectrometry	349
Rate-Limiting Steps in Current DNA Sequencing	358
Sources and Additional Readings	359
11 Strategies for Large-Scale DNA Sequencing	361
Why Strategies Are Needed	361
Shotgun DNA Sequencing	361

Directed Sequencing with Walking Primers	363
Priming with Mixtures of Short Oligonucleotides	365
Ordered Subdivision of DNA Targets	368
Transposon-Mediated DNA Sequencing	368
Delta Restriction Cloning	370
Nested Deletions	371
Primer Jumping	373
Primer Multiplexing	375
Multiplex Genomic Walking	376
Global Strategies	377
Sequence-Ready Libraries	379
Sequencing cDNA Libraries	380
Dealing with Uneven cDNA Distribution	381
Large-Scale cDNA Sequencing	384
What Is Meant by a Complete Genome Sequence?	389
Sequencing the Fifth Base	390
Sources and Additional Readings	392
12 Future DNA Sequencing without Length Fractionation	394
Why Try to Avoid Length Fractionations?	394
Single-Molecule Sequencing	394
Sequencing by High-Resolution Microscopy	397
Stepwise Enzymatic Sequencing	400
DNA Sequencing by Hybridization (SBH)	403
Branch Point Ambiguities	404
SBH Using Oligonucleotide Chips	406
Sequencing by Hybridization to Sample Chips	410
Early Experiences with SBH	412
Data Acquisition and Analysis	415
Obstacles to Successful SBH	417
SBH in Comparative DNA Sequencing	420
Oligonucleotide Stacking Hybridization	421
Other Approaches for Enhancing SBH	424
Positional Sequencing by Hybridization (PSBH)	425
Combination of SBH with Other Sequencing Methods	430
Sources and Additional Readings	431
13 Finding Genes and Mutations	433
Detection of Altered DNA Sequences	433
Finding Genes	434
Diagnostics at the DNA Level	448
Analysis of DNA Sequence Differences	455
Heteroduplex Detection	456
Diagnosis of Infectious Disease	462
Detection of New Mutations	463
Sources and Additional Readings	467

14	Sequence-Specific Manipulation of DNA	470
	Exploiting the Specificity of Base-Base Recognition	470
	Structure of Triple-Stranded DNA	470
	Triplex-Mediated DNA Cleavage	476
	Sequence-Specific DNA Capture	480
	Triplex-Mediated DNA Capture	480
	Affinity Capture Electrophoresis	486
	Use of Backbone Analogues in Sequence-Specific DNA Manipulation	489
	Sequence-Specific Cloning Procedures	492
	Identification or Cloning of Sequences Based on Differences in Expression Level	499
	Coincidence Cloning	500
	Human Interspersed Repeated DNA Sequences	506
	Distribution of Repeats Along Chromosomes	509
	PCR Based on Repeating Sequences	510
	Repeat Expansion Detection	516
	Aptamer Selection Strategies	517
	Oligonucleotides as Drugs	520
	Sources and Additional Readings	523
15	Results and Implications of Large-Scale DNA Sequencing	526
	Costing the Genome Project	526
	Finding Genes	530
	More Robust Methods for Finding Genes by DNA Sequence Analysis	532
	Neural Net Analysis of DNA Sequences	535
	Survey of Past Large-Scale DNA Sequencing Projects	540
	Finding Errors in DNA Sequences	545
	Searching for the Biological Function of DNA Sequences	547
	Searching for the Biological Function of Genes	548
	Methods for Comparing Sequences	551
	Dynamic Programming	557
	Gaining Additional Power in Sequence Comparisons	560
	Domains and Motifs	561
	Interpreting Noncoding Sequence	563
	Diversity of DNA Sequences	564
	Sources and Additional Readings	565
	Appendix: Databases	569
	Index	575