discovering

GENOMICS, PROTEOMICS, & BIOINFORMATICS

A. MALCOLM CAMPBELL & LAURIE J. HEYER

CONTENTS

FOREWORD ix

PREFACE xi

ACKNOWLEDGMENTS xiv

UNIT ONE

Genome Sequences 1

CHAPTER 1

Genome Sequence Acquisition and Analysis 2

1.1 Defining Genomes 3

What Is Genomics? 3

How Are Whole Genomes Sequenced? 3

Math Minute 1.1 *What Is an E-Value? 5*

Why Do the Databases Contain So Many Partial Sequences? 6

How Do We Make Sense of All These Bases? 9

Box 1.1 *Which Draft Sequence Is Better? 9*

Can We Predict Protein Functions? 10

How Well Are Genes Conserved in Diverse Species? 12

How Do You Know Which Bases Form a Gene? 14

How Many Proteins Can One Gene Make? 15

Summary 1.1 16

1.2 What Have We Learned from the Human Genome Draft Sequences? 17

Overview of Human Genome First Draft 17

Summary Statements 17

Box 1.2 *Whose DNA Did We Sequence? 17*

Math Minute 1.2 *How Do You Fit a Line to Data? 18*

Can We Describe a Typical Human Gene? 20

When Are the Data Sufficient? 24

Can the Genome Alter Gene Expression Without Changing the DNA Sequence? 25

Summary 1.2 28

Chapter 1 Conclusions 28

References 28

CHAPTER 2

Genome Sequences Answer Interesting Questions 30

2.1 Evolution of Genomes 31

How Did Eukaryotes Evolve? 31

Math Minute 2.1 *Are the Hit Numbers Significantly Different? 33*

What Is the Origin of Our Species? 41

Math Minute 2.2 *How Do You Know if the Tree Is Right? 45*

Summary 2.1 46

2.2 Genomic Identifications 46
How Can We Identify Biological Weapons? 47
How Long Can DNA Survive? 49
How Did Tuberculosis Reach North America? 50
How Are Newly Emerging Diseases Identified? 53
Summary 2.2 57

2.3 Biomedical Genome Research 57
Can We Use Genomic Sequences to Make New Vaccines? 57
Can We Make New Types of Antibiotics? 59
Can We Invent New Types of Medication? 62
How Can *E. coli* Be Lethal and in Our Intestines at the Same Time? 64
Math Minute 2.3 *How Can You Tell if Base Compositions Are Different? 65*
Summary 2.3 66
Chapter 2 Conclusions 66
References 66

CHAPTER 3
Genomic Variations 69

3.1 Environmental Case Study 70
Can Genomic Diversity Affect Global Warming? 70
Math Minute 3.1 *How Do You Measure Genetic Variation? 72*
Math Minute 3.2 *How Are Populations Modeled? 74*
Summary 3.1 76

3.2 Human Genomic Variation 76
How Much Variation Is in the Human Genome? 76
Math Minute 3.3 *Are All SNPs Really SNPs? 78*
Why Should We Care About SNPs? 79
Box 3.1 *What's the Difference Between a Mutation and an Allele? 80*
Are There Any Known Examples of SNPs That Cause Diseases? 82
Are There Any Known Changes in Nondisease QTL Due to SNPs? 84
Box 3.2 *Patent Law and Genomics 85*
Why the SNP Frenzy? Pharmacogenomics! 86
Summary 3.2 88

3.3 The Ultimate Genomic Phenotype—Death? 89
Why Do We Age? 89
Are There Hidden Costs for a Prolonged Life? 90
Do Bacteria Experience Genomic Trade-offs Too? 91
Summary 3.3 93

3.4 Ethical Consequences of Genomic Variations 93
Are Genetically Modified Organisms Bad? 93
Is Genetic Testing Good? 95
Are There Simple Applications for Complex Genomes? 99
Box 3.3 *Should I Get a Genetic Test? 99*
Summary 3.4 101
Chapter 3 Conclusions 102
References 102

UNIT TWO
Genome Expression 105

CHAPTER 4
Basic Research with DNA Microarrays 106

4.1 Introduction to Microarrays 107
What Happened to My Home Brew? 107
Math Minute 4.1 *How Do You Transform the Data to Avoid Fractions? 112*
Math Minute 4.2 *How Do You Measure Similarity Between Expression Patterns? 113*
Math Minute 4.3 *How Do You Cluster Genes? 114*
What Can We Learn from Stressed-out Yeast? 124
Why Are There So Many Copies of Some Genes but Not Others? 125
How Well Do Promoters Control Gene Expression? 126
Are Promoters Able to Work in Reverse? 127
Summary 4.1 128

4.2 Alternative Uses of DNA Microarrays 128
Why Do So Many Unrelated Genes Share the Same Expression Profile? 128
Math Minute 4.4 *Is It Useful to Compare the Columns of a Gene Expression Matrix? 129*
Can Cells Verify Their Own Genes? 131
Summary 4.2 133
Chapter 4 Conclusions 135
References 135

CHAPTER 5
Applied Research with DNA Microarrays 137

5.1 Cancer and Genomic Microarrays 138
Are There Better Ways to Diagnose Cancer? 138
Math Minute 5.1 *What Are Signature Genes, and How Do You Use Them? 139*
Can Breast Cancer Be Categorized with Microarrays, too? 141
What Genomic Changes Occur in Cancer Cells? 143
Summary 5.1 146

5.2 Improving Health Care with DNA Microarrays 146
Why Is the Tuberculosis Vaccine Less Effective Now? 146
How Does This Drug Work? 149
Can We Predict Which Drugs Will Be Effective in Different Cancers? 152
What Happens When You Accumulate Fat? 154
Summary 5.2 158
Chapter 5 Conclusions 158
References 158

CHAPTER 6
Proteomics 161

6.1 Introduction 162
What Do All These Proteins Do? 162
Which Proteins Are Needed in Different Conditions? 166
Math Minute 6.1 *How Do You Know if You Have Sampled Enough Cells? 168*
Can You Live Without Some Proteins? 170
Summary 6.1 171

6.2 Protein 3D Structures 171
Does a Protein's Shape Reveal Its Function? 172
Can We Use Structures to Develop Better Drugs? 173
Can One Protein Kill You? 174
6.2 Summary 176

6.3 Protein Interaction Networks 176
Which Proteins Interact with Each Other? 176
How Can We Measure Protein Interactions? 177
Math Minute 6.2 *Is Sup35 a Central Protein in the Network? 179*
Is It Possible to Understand Proteome-wide Interactions? 181
Summary 6.3 183

6.4 Measuring Proteins 183
How Do We Know Which Proteins Are Present? 184
What Proteins Do Our White Blood Cells Need to Kill a Pathogen? 187
How Much of Each Protein Is Present? 189
Can We Make *Protein* Chips? 195
Are All Cells Equal? 198
What Does a Proteome Produce? 200
Summary 6.4 202
Chapter 6 Conclusions 202
References 202

UNIT THREE
Whole Genome Perspective 205

CHAPTER 7
Genomic Circuits in Single Genes 206

7.1 Dissecting a Gene's Circuitry 207
How Do Genomes Control Individual Genes? 207

How does a Gene Control Location, Timing, and Quantity of Transcription? 210
What Does Module G Do? 216
Can We Apply Engineering and Computer Science Concepts to Genes? 226
Summary 7.1 229

7.2 Integrating Single-Gene Circuits 229
How Can We Describe to Others What We Know About a Genome Circuit? 229
Technical Hints 230
Can We Visualize Circuits for Protein Interaction and DNA Binding? 230
Summary 7.2 230
Chapter 7 Conclusions 230
References 231

CHAPTER 8
Integrated Genomic Circuits 232

8.1 Simple Integrated Circuits 233
Can Genes Form Toggle Switches and Make Choices? 233
Math Minute 8.1 *How Are Stochastic Models Applied to Cellular Processes? 234*
Can Humans Engineer a Genetic Toggle Switch? 238
Can Humans Build a Synthetic Circadian Clock from a Toggle Switch Design? 240
If Toggle Switches Are So Noisy, How Can Multicellular Organisms Develop? 241
Redundancy: Is It Really Beneficial to Have More Than One Copy of a Gene? 242
Summary 8.1 244

8.2 Complex Integrated Circuits 244
Are Circuits the Key to Learning? 244
Math Minute 8.2 *Is It Possible to Predict Steady-state Behavior? 250*
Can We Understand Cancer Better by Understanding Its Circuitry? 257
If Circuits Are Interconnected, Does Gene Order Matter? 259
Summary 8.2 263
Chapter 8 Conclusions 263
References 263

CHAPTER 9
Modeling Whole-Genome Circuits 265

9.1 Is Genomics a New Perspective? 266
The People Involved: Who Is Doing Systems Biology? 266
The Quality of the Message: What Questions Do Systems Biologists Ask? 267

9.2 Can We Model Entire Eukaryotes with a Systems Approach? 267
Genomics versus Proteomics 271
Building a Systems Model 272
Context of the Message 273

9.3 Will Systems Biology Go Systemic? 274
Chapter 9 Conclusions 274
References 275

UNIT FOUR
Transition from Genetics to Genomics: Medical Case Studies 277

CHAPTER 10
What's Wrong with My Child? 278

10.1 First Patients 279
Phase I: Clinical Presentation 279
Phase II: Family Pedigree 280
Phase III: Karyotyping and Linkage Analysis 280
Phase IV: DNA Sequence Analysis 281
Summary 10.1 283

10.2 The Next Steps in Understanding the Disease 284
We Need an Animal Model System 284
What Was That Other Protein I Got Lots of Hits For? 284
Does Utrophin Play a Role in Muscular Dystrophy, Too? 284

What Does Dystrophin Do Anyway? 285
Math Minute 10.1 *What's Special about This Graph? 286*
Why Do DMD Patients' Muscles Deteriorate After the First Three Years? 287
Is It Possible to Have DMD and Be Wild-Type for Dystrophin? 288
How Can They Have Muscular Dystrophy if Their Dystrophin Genes Are Normal? 288
Math Minute 10.2 *What Do You Mean by Highly Unlikely? 289*
Where Is the Muscular Dystrophy Field Now? 293
Math Minute 10.3 *Is cGMP Production Elevated? 301*
Summary 10.2: Your Final Thoughts 303
Chapter 10 Conclusions 303
References 304

CHAPTER 11
Why Can't I Just Take a Pill to Lose Weight? 306

Hungry for Knowledge 307
Saturday, 21 October. 7:30 A.M. 307
Building a Model for Weight Homeostasis 308
Cloning the Leptin Gene 308
Functional Tests for Leptin 310
Time to Visit Grandma 311
Grandma Gives You Homework! 311
Chapter 11 Conclusions 319
References 319

CHAPTER 12
Why Can't We Cure More Diseases? 320

How to Develop a New Medication 321
Define the Problem and Devise a Solution 321
Focus 1: Location, Location, Location 321
Focus 2: Delivery Vehicles 321
Focus 3: Specificity—"If It Ain't Broke, Don't Fix It" 323
Math Minute 12.1 *What's the Right Dose? 324*
Eye of Newt . . . ? 326
Don't Treat the Symptom, Treat the Cause 327
Chapter 12 Conclusions 329
References 329

GLOSSARY 331
CREDITS 341
INDEX 345