

# Contents in Brief

- 1** Sequence Databases, 3
- 2** Mapping Databases, 25
- 3** Information Retrieval from Biological Databases, 55
- 4** Genomic Databases, 81
- 5** Predictive Methods Using DNA Sequences, 115
- 6** Predictive Methods Using RNA Sequences, 143
- 7** Sequence Polymorphisms, 171
- 8** Predictive Methods Using Protein Sequences, 197
- 9** Protein Structure Prediction and Analysis, 223
- 10** Intermolecular Interactions and Biological Pathways, 253

- 11** Assessing Pairwise Sequence Similarity: BLAST and FASTA, 295
- 12** Creation and Analysis of Protein Multiple Sequence Alignments, 325
- 13** Sequence Assembly and Finishing Methods, 341
- 14** Phylogenetic Analysis, 365
- 15** Computational Approaches in Comparative Genomics, 393
- 16** Using DNA Microarrays to Assay Gene Expression, 409
- 17** Proteomics and Protein Identification, 445
- 18** Using Perl to Facilitate Biological Analysis, 475

- 12** Creation and Analysis of Protein Multiple Sequence Alignments, 325  
*Geoffrey J. Barton*
- 13** Sequence Assembly and Finishing Methods, 341  
*Nancy F. Hansen, Pamela Jacques Thomas and Gerard G. Bouffard*
- 14** Phylogenetic Analysis, 365  
*Fiona S. L. Brinkman*
- 15** Computational Approaches in Comparative Genomics, 393  
*Andreas D. Baxevanis*

# Contents

## Foreword, xi

Lee Hood

## Preface, xiii

## Contributors, xiii

## PART ONE BIOLOGICAL DATABASES

### 1 Sequence Databases, 3

Rolf Apweiler

### 2 Mapping Databases, 25

Peter S. White and Tara C. Matisse

### 3 Information Retrieval from Biological Databases, 55

Andreas D. Baxevanis

### 4 Genomic Databases, 81

Tyra G. Wolfsberg

## PART TWO ANALYSIS AT THE NUCLEOTIDE LEVEL

### 5 Predictive Methods Using DNA Sequences, 115

Enrique Blanco and Roderic Guigó

### 6 Predictive Methods Using RNA Sequences, 143

David Mathews and Michael Zuker

### 7 Sequence Polymorphisms, 171

James C. Mullikin and Stephen T. Sherry

## PART THREE ANALYSIS AT THE PROTEIN LEVEL

### 8 Predictive Methods Using Protein Sequences, 197

Yanay Ofran and Burkhard Rost

### 9 Protein Structure Prediction and Analysis, 223

David Wishart

### 10 Intermolecular Interactions and Biological Pathways, 253

Gary D. Bader and Anton J. Enright

## PART FOUR INFERRING RELATIONSHIPS

### 11 Assessing Pairwise Sequence Similarity: BLAST and FASTA, 295

Andreas D. Baxevanis

### 12 Creation and Analysis of Protein Multiple Sequence Alignments, 325

Geoffrey J. Barton

### 13 Sequence Assembly and Finishing Methods, 341

Nancy F. Hansen, Pamela Jacques Thomas and Gerard G. Bouffard

### 14 Phylogenetic Analysis, 365

Fiona S. L. Brinkman

### 15 Computational Approaches in Comparative Genomics, 393

Andreas D. Baxevanis

- 16 Using DNA Microarrays to Assay Gene Expression, 409**  
*John Quackenbush*
- 17 Proteomics and Protein Identification, 445**  
*Mark R. Holmes, Kevin R. Ramkissoon and Morgan C. Giddings*

Appendices, 497

Glossary, 517

Index, 525

**PART FIVE  
DEVELOPING TOOLS**

- 18 Using Perl to Facilitate Biological Analysis, 475**  
*Lincoln D. Stein*

Contents

Foreword, xi  
Les Hood

Preface, xiii

Contributors, xvii

**PART ONE**

**BIOLOGICAL DATABASES**

1 Sequence Databases, 3  
Roy Anderson

2 Mapping Databases, 25  
Peter S. White and Tom C. Moore

3 Information Retrieval from Biological Databases, 55  
Andreas D. Baxevans

4 Genomic Databases, 81  
The G. Wolfberg

**PART TWO**

**ANALYSIS AT THE NUCLEOTIDE LEVEL**

5 Predictive Methods Using DNA Sequences, 115  
Enrique Blanco and Robert Guigo

6 Predictive Methods Using RNA Sequences, 143  
David Matthews and Michael Zuker

7 Sequence Polymorphisms, 171  
James C. Mullikin and Stephen T. Sherry

**PART THREE**

**ANALYSIS AT THE PROTEIN LEVEL**

8 Predictive Methods Using Protein Sequences, 197  
Yanny Ofran and Burkhard Rost

9 Protein Structure Prediction and Analysis, 223  
David Wishart

10 Intermolecular Interactions and Biological Pathways, 253  
Guy D. Bader and Anton J. Enright

**PART FOUR**

**INFERRING RELATIONSHIPS**

11 Assessing Pairwise Sequence Similarity: BLAST and FASTA, 295  
Andreas D. Baxevans

12 Creation and Analysis of Protein Multiple Sequence Alignments, 325  
Geoffrey J. Barton

13 Sequence Assembly and Finishing Methods, 341  
Nancy E. Hansen, Pamela Jacques Thomas and Gerard G. Bouffard

14 Phylogenetic Analysis, 365  
Feroz S. L. Birkman

15 Computational Approaches in Comparative Genomics, 393  
Andreas D. Baxevans