

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**

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. Mulkin and Stephen T. Sherry
- 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
In the 21st century, we stand at a grand inflection point in biology—how we view and practice biology has forever changed. This inflection point has

Preface, xiii

In a number of events, perhaps the most important of which is the human genome project. It

Contributors, xiii

Its list and catalyzed the development of high-throughput measurement tools (e.g., the

PART ONE

BIOLOGICAL DATABASES

1 Sequence Databases, 3

Rolf Apweiler
Sequence databases delineating protein/DNA

2 Mapping Databases, 25

Peter S. White and Tara C. Matisse
Information about biological systems has changed how we

3 Information Retrieval from Biological Databases, 55

Andreas D. Baxevanis
Information retrieval from biological databases is a fundamental type of biological information

4 Genomic Databases, 81

Tyra G. Wolfsberg
The genomic database links the genome to facilitate

PART TWO

ANALYSIS AT THE NUCLEOTIDE LEVEL

5 Predictive Methods Using DNA Sequences, 115

Enrique Blanco and Roderic Guigó
Predictive methods for proteins, which are encoded by

6 Predictive Methods Using RNA Sequences, 143

David Mathews and Michael Zuker
Predictive methods for RNA structures and networks

7 Sequence Polymorphisms, 171

James C. Mullikin and Stephen T. Sherry
Sequence polymorphisms are the input into

Glossary, 211

Index, 225

17 Proteomics and Protein Identification, 442
Mark R. Hohne, Kevin R. Ramkissoon and Michael J. O'Neil
John D. Giddings

18 Using Bioinformatics to Facilitate Biologics, 443
Andrea L. Sano
Lionel D. Stuif

PART THREE ANALYSIS AT THE PROTEIN LEVEL

8 Predictive Methods Using Protein Sequences, 197

Yanay Ofran and Burkhard Rost
Predictive methods for protein sequences at each level, in order

9 Protein Structure Prediction and Analysis, 223

David Wishart
Protein structure prediction and analysis have become increasingly cross-disciplinary, as biologists, chemists, computer scientists, and physicians work to

10 Intermolecular Interactions and Biological Pathways, 253

Gary D. Bader and Anton J. Enright
Intermolecular interactions and biological pathways are the computational tools required for this new

PART FOUR INFERRING RELATIONSHIPS

11 Assessing Pairwise Sequence Similarity: BLAST and FASTA, 295

Andreas D. Baxevanis
Assessing pairwise sequence similarity is a fundamental tool for

12 Creation and Analysis of Protein Multiple Sequence Alignments, 325

Geoffrey J. Barton
Multiple sequence alignments are used to compare very different

13 Sequence Assembly and Finishing Methods, 341

Nancy F. Hansen, Pamela Jacques Thomas and Gerard G. Bouffard
Sequence assembly and finishing are the processes of collecting

14 Phylogenetic Analysis, 365

Fiona S. L. Brinkman
Phylogenetic analysis is the study of the evolutionary history of life, and statistics

15 Computational Approaches in Comparative Genomics, 393

Andreas D. Baxevanis
Computational approaches in comparative genomics are used to analyze

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

**PART FIVE
DEVELOPING TOOLS**

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

Appendices, 497

Glossary, 517

Index, 525

PART THREE ANALYSIS AT THE PROTEIN LEVEL <ul style="list-style-type: none"> 8 Peptide Mass Fingerprinting Methods Using Protein Sequencing, 161 Youn Oktay and Gundula Holt 9 Protein Structure Prediction and Analysis, 173 David Wishart 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 10 Interproteomic Interactions and Biological Pathways, 253 Gary D. Bader and Amnon E. Perna 	PART TWO INTERRELATING RELATEDNESS <ul style="list-style-type: none"> 11 Assessing Pairwise Sequence Similarity, 302 Andrea D. Goksøyr 	PART TWO ANALYSIS AT THE NUCLEOTIDE LEVEL <ul style="list-style-type: none"> 12 Combustion Abbocates in Comparative Genomics, 333 Andrea D. Goksøyr 13 Phylogenetic Analysis of Protein Families, 341 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 14 Sequence Analysis Using RNA, 365 Roni Z. L. Glusman 	PART TWO ANALYSIS AT THE PROTEIN LEVEL <ul style="list-style-type: none"> 15 Classification and Analysis of Protein Multiples, 373 Goeffrey L. Gatenby 16 Sequence Assembly and Fingerprinting Methods, 381 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 17 Sequence Polymorphisms, 391 Peter C. Whitten and Stephen T. Strobl
PART THREE ANALYSIS AT THE PROTEIN LEVEL <ul style="list-style-type: none"> 18 Peptide Mass Fingerprinting Methods Using Protein Sequencing, 161 Youn Oktay and Gundula Holt 9 Protein Structure Prediction and Analysis, 173 David Wishart 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 10 Interproteomic Interactions and Biological Pathways, 253 Gary D. Bader and Amnon E. Perna 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 11 Assessing Pairwise Sequence Similarity, 302 Andrea D. Goksøyr 	PART TWO INTERRELATING RELATEDNESS <ul style="list-style-type: none"> 12 Combustion Abbocates in Comparative Genomics, 333 Andrea D. Goksøyr 	PART TWO ANALYSIS AT THE NUCLEOTIDE LEVEL <ul style="list-style-type: none"> 13 Phylogenetic Analysis of Protein Families, 341 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 14 Sequence Analysis Using RNA, 365 Roni Z. L. Glusman
PART THREE ANALYSIS AT THE PROTEIN LEVEL <ul style="list-style-type: none"> 15 Classification and Analysis of Protein Multiples, 373 Goeffrey L. Gatenby 16 Sequence Assembly and Fingerprinting Methods, 381 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 17 Sequence Polymorphisms, 391 Peter C. Whitten and Stephen T. Strobl 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 10 Interproteomic Interactions and Biological Pathways, 253 Gary D. Bader and Amnon E. Perna 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 11 Assessing Pairwise Sequence Similarity, 302 Andrea D. Goksøyr 	PART TWO INTERRELATING RELATEDNESS <ul style="list-style-type: none"> 12 Combustion Abbocates in Comparative Genomics, 333 Andrea D. Goksøyr 	PART TWO ANALYSIS AT THE NUCLEOTIDE LEVEL <ul style="list-style-type: none"> 13 Phylogenetic Analysis of Protein Families, 341 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 14 Sequence Analysis Using RNA, 365 Roni Z. L. Glusman
PART THREE ANALYSIS AT THE PROTEIN LEVEL <ul style="list-style-type: none"> 15 Classification and Analysis of Protein Multiples, 373 Goeffrey L. Gatenby 16 Sequence Assembly and Fingerprinting Methods, 381 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 17 Sequence Polymorphisms, 391 Peter C. Whitten and Stephen T. Strobl 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 10 Interproteomic Interactions and Biological Pathways, 253 Gary D. Bader and Amnon E. Perna 	PART ONE BIOLOGICAL DATABASES <ul style="list-style-type: none"> 11 Assessing Pairwise Sequence Similarity, 302 Andrea D. Goksøyr 	PART TWO INTERRELATING RELATEDNESS <ul style="list-style-type: none"> 12 Combustion Abbocates in Comparative Genomics, 333 Andrea D. Goksøyr 	PART TWO ANALYSIS AT THE NUCLEOTIDE LEVEL <ul style="list-style-type: none"> 13 Phylogenetic Analysis of Protein Families, 341 Nanci E. Holter, Julian Jochum, Jerome J. Journeé and Gordon G. Golub 14 Sequence Analysis Using RNA, 365 Roni Z. L. Glusman