

---

# Contents

<i>Preface</i> . . . . .	<i>v</i>
<i>Contributors</i> . . . . .	<i>xi</i>

## SECTION I: DATABASES

1. Nucleic Acid Sequence and Structure Databases . . . . .	3
<i>Stefan Washietl and Ivo L. Hofacker</i>	
2. Genomic Databases and Resources at the National Center for Biotechnology Information . . . . .	17
<i>Tatiana Tatusova</i>	
3. Protein Sequence Databases . . . . .	45
<i>Michael Rebhan</i>	
4. Protein Structure Databases . . . . .	59
<i>Roman A. Laskowski</i>	
5. Protein Domain Architectures . . . . .	83
<i>Nicola J. Mulder</i>	
6. Thermodynamic Database for Proteins: Features and Applications . . . . .	97
<i>M. Michael Gromiha and Akinori Sarai</i>	
7. Enzyme Databases . . . . .	113
<i>Dietmar Schomburg and Ida Schomburg</i>	
8. Biomolecular Pathway Databases . . . . .	129
<i>Hong Sain Ooi, Georg Schneider, Teng-Ting Lim, Ying-Leong Chan, Birgit Eisenhaber, and Frank Eisenhaber</i>	
9. Databases of Protein–Protein Interactions and Complexes . . . . .	145
<i>Hong Sain Ooi, Georg Schneider, Ying-Leong Chan, Teng-Ting Lim, Birgit Eisenhaber, and Frank Eisenhaber</i>	

## SECTION II: DATA MINING TECHNIQUES

10. Proximity Measures for Cluster Analysis . . . . .	163
<i>Oliviero Carugo</i>	
11. Clustering Criteria and Algorithms . . . . .	175
<i>Oliviero Carugo</i>	
12. Neural Networks . . . . .	197
<i>Zheng Rong Yang</i>	
13. A User’s Guide to Support Vector Machines . . . . .	223
<i>Asa Ben-Hur and Jason Weston</i>	
14. Hidden Markov Models in Biology . . . . .	241
<i>Claus Vogl and Andreas Futschik</i>	

# SECTION III: DATABASE ANNOTATIONS AND PREDICTIONS

15. Integrated Tools for Biomolecular Sequence-Based Function Prediction as Exemplified by the ANNOTATOR Software Environment . . . . .	257
<i>Georg Schneider, Michael Wildpaner, Fernanda L. Sirota, Sebastian Maurer-Stroh, Birgit Eisenhaber, and Frank Eisenhaber</i>	
16. Computational Methods for Ab Initio and Comparative Gene Finding . . . . .	269
<i>Ernesto Picardi and Graziano Pesole</i>	
17. Sequence and Structure Analysis of Noncoding RNAs . . . . .	285
<i>Stefan Washietl</i>	
18. Conformational Disorder . . . . .	307
<i>Sonia Longhi, Philippe Lieutaud, and Bruno Canard</i>	
19. Protein Secondary Structure Prediction . . . . .	327
<i>Walter Pirovano and Jaap Heringa</i>	
20. Analysis and Prediction of Protein Quaternary Structure . . . . .	349
<i>Anne Poupon and Joel Janin</i>	
21. Prediction of Posttranslational Modification of Proteins from Their Amino Acid Sequence. . . . .	365
<i>Birgit Eisenhaber and Frank Eisenhaber</i>	
22. Protein Crystallizability. . . . .	385
<i>Pawel Smialowski and Dmitrij Frishman</i>	
<i>Subject Index</i> . . . . .	401



<http://www.springer.com/978-1-60327-240-7>

Data Mining Techniques for the Life Sciences

Carugo, O.; Eisenhaber, F. (Eds.)

2010, XII, 408 p. 89 illus., Hardcover

ISBN: 978-1-60327-240-7

A product of Humana Press