

Contents

Preface	v
List of Reviewers	vii

Section 1: Genes

1 A Survey of the Computational Methods for Enhancers and Enhancer-target Predictions.....	3
1. Introduction.....	3
2. Computational methods for enhancer prediction	7
3. Computational methods for enhancer target prediction.....	15
4. Databases useful for enhancer and enhancer-promoter association prediction	20
5. Conclusions and discussions.....	21
2. Cormotif: An R Package for Jointly Detecting Differential Gene Expression in Multiple Studies	28
1. Introduction.....	29
2. Methods.....	32
3. Simulations.....	35
4. Discussion	41
5. Software	43
3. Granger Causality for Time Series Gene Expression Data.....	48
1. Introduction.....	49
2. Granger Causality for Sets of Time Series.....	50
3. Canonical Correlation Analysis and Granger Causality.....	51
4. Functional Clustering in Terms of Granger Causality.....	58
5. Network Construction from Large Datasets	63
6. Software	64
Appendix.....	64
Acknowledgments.....	65

Section 2: RNAs

4. RNA Sequencing and Gene Expression Regulation.....	71
1. The Fundamentals of DNA, RNA and Gene	72
2. The Fundamentals of Gene Expression and Regulation	84
5. Modern Technologies and Approaches for Decoding Non-Coding Rna-Mediated Biological Networks in Systems Biology and Their Applications	106
1. Introduction.....	106

2. ncRNA-mediated regulatory networks generation and visualization	109
3. Interactomics.....	111
4. Network visualization tools.....	116
5. Biological Network Architecture	117
6. Analyzing ncRNA-mediated regulatory network	120
7. Application of non-coding RNA-mediated biological network	123
8. Challenges and Future Directions	124

Section 3: Proteins

6. Annotation of Hypothetical Proteins- a Functional Genomics	
Approach.....	135
1. Introduction.....	136
2. Methodology for functional annotation of hypothetical proteins	142
3. Conclusion.....	153
7. Protein-Protein Functional Linkage Predictions: Bringing	
Regulation to Context.....	159
1. Protein Functions in the Post-Genomic Era	160
2. Computational Methods for Predicting Functional Linkages between Proteins.....	161
3. Phylogenetic Profiling	162
4. Analysis of Correlated Mutations in Protein Families by Mirrortree Approach: Indicator of Protein-Protein Interaction.....	165
5. Chromosomal Proximity of Genes Reflects their Functional Links.....	166
6. Expression Similarity of Genes as an Indicator of Functional Linkage ..	169
7. From Transcriptional Regulation to Predicting Protein-Protein Functional Linkages: A Novel Approach.....	170
8. Measuring Prediction Performance of Methods	171
9. Perspectives.....	173
Acknowledgements	174

Section 4: Epigenetics

8. Epigenomic Analysis of Chromatin Organization and	
DNA Methylation	181
1. Introduction.....	181
2. Chromatin organization	182
4. Bioinformatics databases and resources.....	197
5. Applications of epigenomics.....	199
6. Emerging fields in epigenetics	201
Acknowledgements	203
9. Gene Body Methylation and Transcriptional Regulation: Statistical	
Modelling and More	212
1. Introduction.....	212
2. Methylation Data and global patterns.....	215
3. Global correlation and patterns between gene methylation and expression	217

4. Statistical modeling and data mining	219
5. Gene body methylation and splicing	222
6. Discussion	223
7. Methods	224
8. Acknowledgements	227

Section 5: Case Study

III. Computational Characterization of Non-small-cell Lung Cancer with EGFR Gene Mutations and its Applications to Drug Resistance	
Prediction	233
1. Introduction	234
2. Computational Modeling of Interaction between an EGFR Tyrosine Kinase and an Inhibitor	238
3. Characterization of EGFR or ErbB-3 Heterodimerization Using Computer Simulations	246
4. New-generation Irreversible EGFR TKIs	250
Acknowledgements	252

Section 6: Advanced Topics

11. Quality Assurance in Genome-Scale Bioinformatics Analyses	259
1. Introduction	259
2. Whole genome sequencing analysis for genomic medicine	260
3. The problem of quality assurance	264
4. Standard validation and QC approaches used in diagnostic laboratories	264
5. Introduction to a software-testing framework	267
6. Software testing approaches, applications and evaluations	268
7. Cloud-based testing as a service (TaaS) for bioinformatics	273
8. Future Directions	275
Acknowledgements	275
12. Recent Computational Trends in Biological Sequence Alignment	279
1. Introduction	279
2. Pairwise Sequence Alignment	280
3. Acceleration of Sequence Alignment Algorithms:	291
4. Gene Tracer Application	296
Acknowledgement	302
13. State Estimation and Process Monitoring of Nonlinear Biological Phenomena Modeled by S-systems	305
1. Introduction	305
2. State Estimation in Non-linear Biological Systems	308
3. Description of State Estimation Techniques	309
4. Faults Detection of Biological Systems Representing Continuously Stirred Tank Reactor Model	313
5. Simulation Results Analysis	318
6. Conclusions	326
Acknowledgment	327

14. Next-Generation Sequencing and Metagenomics.....	331
1. Introduction.....	331
Acknowledgements	346
15. Metabolic Engineering: Dimensions and Applications.....	352
1. Introduction.....	352
2. Metabolic flux balance analysis.....	354
3. Metabolic engineering in microorganisms.....	356
4. Metabolic engineering in plants	360
5. Transcription factors vs. enzymes.....	363
6. Metabolic trafficking and sequestration	364
7. Genome editing.....	364
8. Metabolic engineering in human disorders.....	365
9. Principles and techniques of metabolic engineering in human diseases.....	366
10. Heart Models.....	366
11. Computational Biology has come to aid metabolic engineering	369
13. Computational Biology in identify metabolic pathways.....	374
14. Bioinformatics in identifying genome editing elements such as CRISPRs	375
15. Concluding remarks	376
16. Methods to Identify Evolutionary Conserved Regulatory Elements Using Molecular Phylogenetics in Microbes	381
1. Functional annotation of regulatory proteins.....	381
2. Structure of phylogenetic tree.....	382
3. Methods of phylogenetics inference	383
4. Computational methods to identify conserved regulatory elements in microbes	384
5. Phylogenetic-based methods to identify regulatory proteins.....	386
6. Tools and web-servers for phylogenetic analysis	389
7. Protocol for phylogenetics analysis with Phylogeny.fr.....	394
8. Protocol of phylogenetics analysis using PhyML.....	398
9. Protocol of phylogenetics analysis using BioNJ.....	400
Acknowledgements	401
17. Improved Protein Model Ranking through Topological Assessment	406
1. Introduction.....	407
2. Methodology	412
3. Results	416
4. Discussion	418
5. Conclusion.....	422
6. Limitations and Further Research Possibilities.....	422
Acknowledgements	423
Index	425