

Wiley Series on Bioinformatics • Albert Y. Zomaya and Yi Pan, Series Editors

BIOMOLECULAR NETWORKS

Methods and Applications in Systems Biology

LUONAN CHEN
RUI-SHENG WANG
XIANG-SUN ZHANG

 **WILEY**

CONTENTS

PREFACE	xiii
ACKNOWLEDGMENTS	xv
LIST OF ILLUSTRATIONS	xvii
ACRONYMS	xxiii

1 Introduction	1
-----------------------	----------

- 1.1 Basic Concepts in Molecular Biology / 1
 - 1.1.1 Genomes, Genes, and DNA Replication Process / 5
 - 1.1.2 Transcription Process for RNA Synthesis / 6
 - 1.1.3 Translation Process for Protein Synthesis / 7
- 1.2 Biomolecular Networks in Cells / 8
- 1.3 Network Systems Biology / 13
- 1.4 About This Book / 18

I GENE NETWORKS	23
------------------------	-----------

2 Transcription Regulation: Networks and Models	25
--	-----------

- 2.1 Transcription Regulation and Gene Expression / 25
 - 2.1.1 Transcription and Gene Regulation / 25
 - 2.1.2 Microarray Experiments and Databases / 28
 - 2.1.3 ChIP-Chip Technology and Transcription Factor Databases / 30
- 2.2 Networks in Transcription Regulation / 32
- 2.3 Nonlinear Models Based on Biochemical Reactions / 36
- 2.4 Integrated Models for Regulatory Networks / 43
- 2.5 Summary / 44

3	Reconstruction of Gene Regulatory Networks	47
3.1	Mathematical Models of Gene Regulatory Network / 47	
3.1.1	Boolean Networks / 48	
3.1.2	Bayesian Networks / 49	
3.1.3	Markov Networks / 52	
3.1.4	Differential Equations / 53	
3.2	Reconstructing Gene Regulatory Networks / 55	
3.2.1	Singular Value Decomposition / 56	
3.2.2	Model-Based Optimization / 58	
3.3	Inferring Gene Networks from Multiple Datasets / 61	
3.3.1	General Solutions and a Particular Solution of Network Structures for Multiple Datasets / 63	
3.3.2	Decomposition Algorithm / 65	
3.3.3	Numerical Validation / 67	
3.4	Gene Network-Based Drug Target Identification / 72	
3.4.1	Network Identification Methods / 73	
3.4.2	Linear Programming Framework / 77	
3.5	Summary / 87	
4	Inference of Transcriptional Regulatory Networks	89
4.1	Predicting TF Binding Sites and Promoters / 89	
4.2	Inference of Transcriptional Interactions / 92	
4.2.1	Differential Equation Methods / 93	
4.2.2	Bayesian Approaches / 96	
4.2.3	Data Mining and Other Methods / 98	
4.3	Identifying Combinatorial Regulations of TFs / 99	
4.4	Inferring Cooperative Regulatory Networks / 105	
4.4.1	Mathematical Models / 105	
4.4.2	Estimating TF Activity / 106	
4.4.3	Linear Programming Models / 108	
4.4.4	Numerical Validation / 109	
4.5	Prediction of Transcription Factor Activity / 114	
4.5.1	Matrix Factorization / 114	
4.5.2	Nonlinear Models / 117	
4.6	Summary / 118	

II	PROTEIN INTERACTION NETWORKS	119
5	Prediction of Protein–Protein Interactions	121
5.1	Experimental Protein–Protein Interactions /	121
5.2	Prediction of Protein–Protein Interactions /	126
5.2.1	Association Methods /	127
5.2.2	Maximum-Likelihood Estimation /	134
5.2.3	Deterministic Optimization Approaches /	139
5.3	Protein Interaction Prediction Based on Multidomain Pairs /	150
5.3.1	Cooperative Domains, Strongly Cooperative Domains, Superdomains /	152
5.3.2	Inference of Multidomain Interactions /	154
5.3.3	Numerical Validation /	157
5.3.4	Reconstructing Complexes by Multidomain Interactions /	160
5.4	Domain Interaction Prediction Methods /	163
5.4.1	Statistical Method /	163
5.4.2	Domain Pair Exclusion Analysis /	163
5.4.3	Parsimony Explanation Approaches /	164
5.4.4	Integrative Approaches /	165
5.5	Summary /	167
6	Topological Structure of Biomolecular Networks	169
6.1	Statistical Properties of Biomolecular Networks /	169
6.2	Evolution of Protein Interaction Networks /	173
6.3	Hubs, Motifs, and Modularity in Biomolecular Networks /	174
6.3.1	Network Centralities and Hubs /	174
6.3.2	Network Modularity and Motifs /	177
6.4	Explorative Roles of Hubs and Network Motifs /	179
6.4.1	Dynamic Modularity Organized by Hubs and Network Motifs /	180
6.4.2	Network Motifs Acting as Connectors between Pathways /	186
6.5	Modularity Evaluation of Biomolecular Networks /	194
6.5.1	Modularity Density D /	195
6.5.2	Improving Module Resolution Limits by D /	196

- 6.5.3 Equivalence between D and Kernel k Means / 198
- 6.5.4 Extension of D to General Criteria: D_λ and D_w / 199
- 6.5.5 Numerical Validation / 200
- 6.6 Summary / 204

7 Alignment of Biomolecular Networks 205

- 7.1 Biomolecular Networks from Multiple Species / 205
- 7.2 Pairwise Alignment of Biomolecular Networks / 207
 - 7.2.1 Score-Based Algorithms / 208
 - 7.2.2 Evolution-Guided Method / 211
 - 7.2.3 Graph Matching Algorithm / 212
- 7.3 Network Alignment by Mathematical Programming / 213
 - 7.3.1 Integer Programming Formulation / 214
 - 7.3.2 Components of the Integer Quadratic Programming Approach / 216
 - 7.3.3 Numerical Validation / 217
- 7.4 Multiple Alignment of Biomolecular Networks / 223
- 7.5 Subnetwork and Pathway Querying / 225
- 7.6 Summary / 228

8 Network-Based Prediction of Protein Function 231

- 8.1 Protein Function and Annotation / 231
- 8.2 Protein Functional Module Detection / 234
 - 8.2.1 Distance-Based Clustering Methods / 235
 - 8.2.2 Graph Clustering Methods / 236
 - 8.2.3 Validation of Module Detection / 238
- 8.3 Functional Linkage for Protein Function Annotation / 239
 - 8.3.1 Bayesian Approach / 239
 - 8.3.2 Hopfield Network Method / 241
 - 8.3.3 p -Value Method / 242
 - 8.3.4 Statistical Framework / 243
- 8.4 Protein Function Prediction from High-Throughput Data / 249
 - 8.4.1 Neighborhood Approaches / 250
 - 8.4.2 Optimization Methods / 251
 - 8.4.3 Probabilistic Methods / 254
 - 8.4.4 Machine Learning Techniques / 256

8.5	Function Annotation Methods for Domains /	265
8.5.1	Domain Sources /	267
8.5.2	Integration of Heterogeneous Data /	268
8.5.3	Domain Function Prediction /	270
8.5.4	Numerical Validation /	271
8.6	Summary /	277
III	METABOLIC NETWORKS AND SIGNALING NETWORKS	279
9	Metabolic Networks: Analysis, Reconstruction, and Application	281
9.1	Cellular Metabolism and Metabolic Pathways /	281
9.2	Metabolic Network Analysis and Modeling /	286
9.2.1	Flux Balance Analysis /	286
9.2.2	Elementary Mode and Extreme Pathway Analysis /	288
9.2.3	Modeling Metabolic Networks /	292
9.3	Reconstruction of Metabolic Networks /	294
9.3.1	Pathfinding Based on Reactions and Compounds /	294
9.3.2	Stoichiometric Approaches Based on Flux Profiles /	297
9.3.3	Inferring Biochemical Networks from Timecourse Data /	298
9.4	Drug Target Detection in Metabolic Networks /	300
9.4.1	Drug Target Detection Problem /	301
9.4.2	Integer Linear Programming Model /	302
9.4.3	Numerical Validation /	305
9.5	Summary /	311
10	Signaling Networks: Modeling and Inference	313
10.1	Signal Transduction in Cellular Systems /	313
10.2	Modeling of Signal Transduction Pathways /	316
10.2.1	Differential Equation Models /	317
10.2.2	Petri Net Models /	319
10.3	Inferring Signaling Networks from High-Throughput Data /	321
10.3.1	NetSearch Method /	322

10.3.2	Ordering Signaling Components /	323
10.3.3	Color-Coding Methods /	324
10.4	Inferring Signaling Networks by Linear Programming /	326
10.4.1	Integer Linear Programming Model /	327
10.4.2	Significance Measures /	329
10.4.3	Numerical Validation /	329
10.4.4	Inferring Signaling Networks by Network Flow Models /	338
10.5	Inferring Signaling Networks from Experimental Evidence /	341
10.6	Summary /	343

11 Other Topics and New Trends 345

11.1	Network-Based Protein Structural Analysis /	345
11.2	Integration of Biomolecular Networks /	347
11.3	Posttranscriptional Regulation of Noncoding RNAs /	349
11.4	Biomolecular Interactions and Human Diseases /	350
11.5	Summary /	352

REFERENCES 353

INDEX 381