

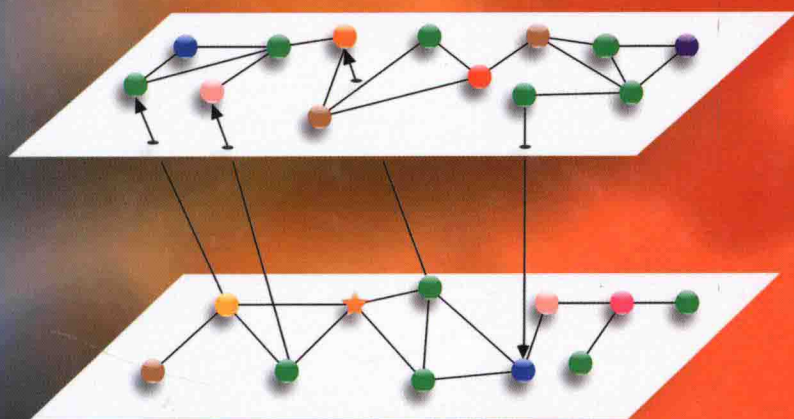
Quantitative and Network Biology  
Series Editors M. Dehmer and F. Emmert-Streib

Volume 1

# Applied Statistics for Network Biology

Methods in Systems Biology

Edited by Matthias Dehmer,  
Frank Emmert-Streib, Armin Graber,  
and Armino Salvador



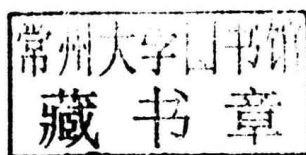
*Quantitative and Network Biology*  
*Series Editors M. Dehmer and F. Emmert-Streib*  
*Volume 1*

## **Applied Statistics for Network Biology**

Methods in Systems Biology

*Edited by*

*Matthias Dehmer, Frank Emmert-Streib, Armin Graber,  
and Armino Salvador*



 **WILEY-BLACKWELL**

## The Editors

### **Matthias Dehmer**

UMIT  
Institute for Bioinformatics  
and Translational Research  
Eduard Wallnöfer Zentrum 1  
6060 Hall, Tyrol  
Austria

### **Frank Emmert-Streib**

Queen's University Belfast  
Center for Cancer Research and Cell Biology  
97, Lisburn Road  
Belfast BT9 7BL  
United Kingdom

### **Armin Graber**

UMIT  
Institute for Bioinformatics  
and Translational Research  
Eduard Wallnöfer Zentrum 1  
6060 Hall, Tyrol  
Austria

and

Novartis Pharmaceuticals Corporation  
Oncology Biomarkers and Imaging  
One Health Plaza  
East Hanover, NJ 07936  
USA

### **Armindo Salvador**

University of Coimbra  
Center for Neuroscience and  
Cell Biology, Department of Chemistry  
3004-535 Coimbra  
Portugal

**Composition** Thomson Digital, Noida, India  
**Printing and Binding** betz-druck GmbH, Darmstadt  
**Cover Design** Adam Design, Weinheim

**Limit of Liability/Disclaimer of Warranty:** While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty can be created or extended by sales representatives or written sales materials. The Advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

**Library of Congress Card No.:** applied for

### **British Library Cataloguing-in-Publication Data**

A catalogue record for this book is available from the British Library.

### **Bibliographic information published by the Deutsche Nationalbibliothek**

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

© 2011 Wiley-VCH Verlag & Co. KGaA,  
Boschstr. 12, 69469 Weinheim, Germany

**Wiley-Blackwell** is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical, and Medical business with Blackwell Publishing.

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany  
Printed on acid-free paper

**ISBN:** 978-3-527-32750-8

*Edited by*  
*Matthias Dehmer,*  
*Frank Emmert-Streib*  
*Armin Graber, and*  
*Armando Salvador*

**Applied Statistics  
for Network Biology**

## **Related Titles**

Emmert-Streib, F., Dehmer, M. (eds.)

### **Medical Biostatistics for Complex Diseases**

2010

ISBN: 978-3-527-32585-6

Dehmer, M., Emmert-Streib, F. (eds.)

### **Analysis of Complex Networks**

**From Biology to Linguistics**

2009

ISBN: 978-3-527-32345-6

Emmert-Streib, F., Dehmer, M. (eds.)

### **Analysis of Microarray Data**

**A Network-Based Approach**

2008

ISBN: 978-3-527-31822-3

Junker, B. H., Schreiber, F.

### **Analysis of Biological Networks**

2008

ISBN: 978-0-470-04144-4

Stolovitzky, G., Califano, A. (eds.)

### **Reverse Engineering Biological Networks**

**Opportunities and Challenges in Computational Methods for Pathway Inference**

2007

ISBN: 978-1-57331-689-7

## Preface

For the field of systems biology to mature, novel statistical and computational analysis methods are needed to deal with the growing amount of high-throughput data from genomics and genetics experiments. This book presents such methods and applications to data from biological and biomedical problems. Nowadays, it is widely recognized that *networks* form a very fruitful representation for studying problems in systems biology [1, 2]. However, many traditional methods do not make explicit use of a network representation of the data. For this reason, the topics treated in this book explore statistical and computational data analysis aspects of *networks* in systems biology [3–6].

Biological phenotypes are mediated by very intricate networks of interactions among biological components. This book covers extensively what we view as two complementary but strongly interrelated challenges in network biology. The first lies in inferring networks from experimental observations of state variables of a system. Interactions among molecular components are traditionally characterized through equilibrium binding or kinetic experiments *in vitro* with dilute solutions of the purified components. However, such experiments are typically low throughput and unable to properly account for the conditions prevailing *in vivo*, where factors such as molecular crowding, spatial heterogeneity, and the presence of ligands might strongly modify the interactions of interest. The possibility of inferring network connectivity and even quantitative interaction parameters from observations of intact living systems is attracting considerable research interest as a way of escaping such shortcomings. The fact that biological networks are complex, that problems are often poorly constrained, and that data are often high dimensional and noisy makes this challenge daunting. The second and perhaps equally difficult challenge lies in deriving results that are both biologically relevant and reliable from incomplete and uncertain information about biological interaction networks. We hope that the contributions in the subsequent chapters will help the reader understand and meet these challenges.

This book is intended for researches and graduate and advanced undergraduate students in the interdisciplinary fields of computational biology, biostatistics, bioinformatics, and systems biology studying problems in biological and biomedical sciences. The book is organized in four main parts: Part One: Modeling, Simulation, and Meaning of Gene Networks; Part Two: Inference of Gene Networks; Part 3: Analysis of Gene Networks; and Part Four: Systems Approach to Diseases. Each part

consists of chapters that emphasize the topic of the corresponding part, however, without being disconnected from the remainder of the book. Overall, to order the different parts we assumed an intuitive – problem-oriented – perspective moving from *Modeling, Simulation, and Meaning of Gene Networks* to *Inference of Gene Networks* and *Analysis of Gene Networks*. The last part presents biomedical applications of various methods in *Systems Approach to Diseases*.

Each chapter is comprehensively presented, accessible not only to researchers from this field but also to advanced undergraduate or graduate students. For this reason, each chapter not only presents technical results but also provides background knowledge necessary to understand the statistical method or the biological problem under consideration. This allows to use this book as a textbook for an interdisciplinary seminar for advanced students not only because of the comprehensiveness of the chapters but also because of its size allowing to fill a complete semester.

Many colleagues, whether consciously or unconsciously, have provided us with input, help, and support before and during the preparation of this book. In particular, we would like to thank Andreas Albrecht, Gökmen Altay, Subhash Basak, Danail Bonchev, Maria Duca, Dean Fennell, Galina Glazko, Martin Grabner, Beryl Graham, Peter Hamilton, Des Higgins, Puthen Jithesh, Patrick Johnston, Frank Kee, Terry Lappin, Kang Li, D. D. Lozovanu, Dennis McCance, James McCann, Alexander Mehler, Abbe Mowshowitz, Ken Mills, Arcady Mushegian, Katie Orr, Andrei Perjan, Bert Rima, Brigitte Senn-Kircher, Ricardo de Matos Simoes, Francesca Shearer, Fred Sobik, John Storey, Simon Tavaré, Shailesh Tripathi, Kurt Varmuza, Bruce Weir, Pat White, Kathleen Williamson, Shu-Dong Zhang, and Dongxiao Zhu and apologize to all who have not been named mistakenly. We would also like to thank our editors Andreas Sendtko and Gregor Cicchetti from Wiley-VCH who have been always available and helpful.

Finally, we hope that this book will help to spread out the enthusiasm and joy we have for this field and inspire people regarding their own practical or theoretical research problems.

March 2011

Belfast, Hall/Tyrol, and Coimbra

Matthias Dehmer,  
Frank Emmert-Streib,  
Armin Graber,  
and Armindo Salvador

## References

- 1 Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.*, **5**, 101–113.
- 2 Emmert-Streib, F. and Glazko, G. (2011) Network biology: a direct approach to study biological function. *WIREs Syst. Biol. Med.*, in press.
- 3 Alon, U. (2006) *An Introduction to Systems Biology: Design Principles of Biological Circuits*, Chapman & Hall/CRC.
- 4 Bertalanffy, L. von. (1950) An outline of general systems theory. *Br. J. Philos. Sci.*, **1** (2)
- 5 Kitano, H. (ed.) (2001) *Foundations of Systems Biology*, MIT Press.
- 6 Palsson, B.O. (2006) *Systems Biology: Properties of Reconstructed Networks*, Cambridge University Press.

## List of Contributors

**Andreas Bernthaler**

Vienna University of Technology  
Institute of Computer Languages  
Theory and Logics Group  
Favoritenstrasse 9  
1040 Vienna  
Austria

**Marina Bessarabova**

Thomson Reuters  
Healthcare & Life Sciences  
169 Saxony Road  
Encinitas, CA 92024  
USA

**Mariano Bizzarri**

Sapienza University  
Department of Experimental Medicine  
Viale Regina Elena 324  
00161 Rome  
Italy

**Gianluca Bontempi**

Université Libre de Bruxelles  
Computer Science Department  
Machine Learning Group  
Boulevard du Triomphe  
1050 Brussels  
Belgium

**Gastone Castellani**

Università di Bologna  
Department of Physics  
INFN Bologna Section and  
Galvani Center for Biocomplexity  
40127 Bologna  
Italy

**Jing Chen**

University of Cincinnati  
Department of Environmental Health  
Cincinnati, OH 45229  
USA

**Xi Chen**

The University of Hong Kong  
Department of Mathematics  
Pok Fu Lam Road  
Hong Kong  
China

**Wai-Ki Ching**

The University of Hong Kong  
Department of Mathematics  
Pok Fu Lam Road  
Hong Kong  
China

**Zoltan Dezso**

Thomson Reuters  
Healthcare & Life Sciences  
169 Saxony Road  
Encinitas, CA 92024  
USA



**Cathy S. J. Fann**

Academia Sinica  
Institute of Biomedical Sciences  
Academia Road, Nankang  
115 Taipei  
Taiwan

**Raul Fechete**

Emergentec Biodevelopment GmbH  
Gersthofer Strasse 29-31  
1180 Vienna  
Austria

**Rudolf Freund**

Vienna University of Technology  
Institute of Computer Languages  
Theory and Logics Group  
Favoritenstrasse 9  
1040 Vienna  
Austria

**Alessandro Giuliani**

Istituto Superiore di Sanità  
Department of Environment  
and Health  
Viale Regina Elena 299  
00161 Rome  
Italy

**Erich Gombocz**

IO Informatics Inc.  
2550 Ninth Street  
Berkeley, CA 94710-2549  
USA

**Jing-Dong J. Han**

Chinese Academy of Sciences  
Institute of Genetics and  
Developmental Biology  
Center for Molecular Systems Biology  
Key Laboratory of  
Molecular Developmental Biology  
Lincui East Road  
100101 Beijing  
China

and

Chinese Academy of Sciences—  
Max Planck Partner Institute for  
Computational Biology  
Shanghai Institutes for  
Biological Sciences  
Chinese Academy of Sciences  
320 Yue Yang Road  
200031 Shanghai  
China

**Katsuhisa Horimoto**

National Institute of Advanced  
Industrial Science Technology  
Computational Biology Research Center  
2-4-7, Aomi, Koto-ku  
135-0064 Tokyo  
Japan

**Ching-Lin Hsiao**

Academia Sinica  
Institute of Biomedical Sciences  
Academia Road, Nankang  
115 Taipei  
Taiwan

**Jialiang Huang**

Chinese Academy of Sciences  
Institute of Genetics and  
Developmental Biology  
Center for Molecular Systems Biology  
Key Laboratory of  
Molecular Developmental Biology  
Lincui East Road  
100101 Beijing  
China

**Anil G. Jegga**

Cincinnati Children's Hospital  
Medical Center  
Division of Biomedical Informatics  
Cincinnati, OH 45229  
USA

and

University of Cincinnati  
Department of Biomedical Engineering  
Cincinnati, OH 45229  
USA

and

University of Cincinnati  
College of Medicine  
Department of Pediatrics  
Cincinnati, OH 45229  
USA

**Eugene Kirillov**

Thomson Reuters  
Healthcare & Life Sciences  
169 Saxony Road  
Encinitas, CA 92024  
USA

**Younhee Ko**

University of Illinois at  
Urbana-Champaign  
Department of Animal Sciences  
1207 W. Gregory Dr.  
Urbana, IL 61801  
USA

and

University of Illinois at  
Urbana-Champaign  
Institute for Genomic Biology  
1205 W. Gregory Drive  
Urbana, IL 61801  
USA

**Rainer König**

University of Heidelberg  
Institute of Pharmacy and Molecular  
Biotechnology  
Bioquant  
Im Neuenheimer Feld 267  
69120 Heidelberg  
Germany

**Xiangfang Li**

Texas A&M University  
Genomic Signal Processing Laboratory  
TAMU 3128  
College Station, TX 77843  
USA

**Arno Lukas**

Emergentec Biodevelopment GmbH  
Gersthofer Strasse 29-31  
1180 Vienna  
Austria

**Bernd Mayer**

Emergentec Biodevelopment GmbH  
Gersthofer Strasse 29-31  
1180 Vienna  
Austria

**Patrick E. Meyer**

Université Libre de Bruxelles  
Computer Science Department  
Machine Learning Group  
Boulevard du Triomphe  
1050 Brussels  
Belgium

**Konrad Mönks**

Vienna University of Technology  
Institute of Computer Languages  
Theory and Logics Group  
Favoritenstrasse 9  
1040 Vienna  
Austria

and

Emergentec Biodevelopment GmbH  
Gersthofer Strasse 29-31  
1180 Vienna  
Austria

**Irmgard Mühlberger**

Emergentec Biodevelopment GmbH  
Gersthofer Strasse 29-31  
1180 Vienna  
Austria

**Tatiana Nikolskaya**

Thomson Reuters  
Healthcare & Life Sciences  
169 Saxony RD  
Encinitas, CA 92024  
USA

**Yuri Nikolsky**

Thomson Reuters  
Healthcare & Life Sciences  
169 Saxony Road  
Encinitas, CA 92024  
USA

**Catharina Olsen**

Université Libre de Bruxelles  
Computer Science Department  
Machine Learning Group  
Boulevard du Triomphe  
1050 Brussels  
Belgium

**Paul Perco**

Emergentec Biodevelopment GmbH  
Gersthofer Strasse 29-31  
1180 Vienna  
Austria

**Kitiporn Plaimas**

University of Heidelberg  
Institute of Pharmacy and Molecular  
Biotechnology  
Bioquant  
Im Neuenheimer Feld 267  
69120 Heidelberg  
Germany

**Thomas N. Plasterer**

Northeastern University  
Department of Chemistry and  
Chemical Biology  
360 Huntington Ave.  
Boston, MA 02115  
USA

and

Pharmacogenetics Clinical Advisory  
Board  
2000 Commonwealth Avenue, Suite 200  
Auburndale, MA 02466  
USA

**Lijun Qian**

Texas A&M University System  
Prairie View A&M University  
Department of Electrical and  
Computer Engineering  
MS2520, POB 519  
Prairie View, TX 77446  
USA

**Daniel Remondini**

Università di Bologna  
Department of Physics  
INFN Bologna Section and  
Galvani Center for Biocomplexity  
40127 Bologna  
Italy

**Sandra Rodriguez-Zas**

University of Illinois at  
Urbana-Champaign  
Department of Animal Sciences  
1207 W. Gregory Drive  
Urbana, IL 61801  
USA

**Debashis Sahoo**

Instructor of Pathology and Siebel  
Fellow at Institute of Stem Cell Biology  
and Regenerative Medicine  
Lorry I. Lokey Stem Cell Research  
Building  
265 Campus Drive, Rm G3101B  
Stanford, CA 94305  
USA

**Shigeru Saito**

Infocom Corp.  
Chem & Bio Informatics Department  
Sumitomo Fudosan Harajuku Building  
2-34-17, Jingumae, Shibuya-ku  
150-0001 Tokyo  
Japan

**Robert Stanley**

IO Informatics Inc.  
2550 Ninth Street  
Berkeley, CA 94710-2549  
USA

**Gautam S. Thakur**

University of Florida  
Department of Computer and  
Information Science and Engineering  
Science  
PO Box 116120  
Gainesville, FL 32611-6120  
USA

**Tianhai Tian**

University of Glasgow  
Department of Mathematics  
University Gardens  
Glasgow G12 8QW  
UK

**Nam-Kiu Tsing**

The University of Hong Kong  
Department of Mathematics  
Pok Fu Lam Road  
Hong Kong  
China

**Eberhard O. Voit**

Georgia Tech and Emory University  
The Wallace H. Coulter Department  
of Biomedical Engineering  
313 Ferst Drive  
Atlanta, GA 30332  
USA

**Haixin Wang**

Fort Valley State University  
Department of Mathematics and  
Computer Science  
CTM 101A  
Fort Valley, GA 31030  
USA

**Matthew Weirauch**

University of Toronto  
Banting and Best Department  
of Medical Research and  
Donnelly Centre for  
Cellular and Biomolecular Research  
160 College Street  
Toronto, ON, M5S 3E1  
Canada

**Hong Yu**

Chinese Academy of Sciences  
Institute of Genetics and  
Developmental Biology  
Center for Molecular Systems Biology  
Key Laboratory of  
Molecular Developmental Biology  
Lincui East Road  
100101 Beijing  
China

**Wei Zhang**

Chinese Academy of Sciences  
Institute of Genetics and  
Developmental Biology  
Center for Molecular Systems Biology  
Key Laboratory of  
Molecular Developmental Biology  
Lincui East Road  
100101 Beijing  
China

## Contents

**Preface** *XVII*

**List of Contributors** *XIX*

### **Part One Modeling, Simulation, and Meaning of Gene Networks** *1*

#### **1 Network Analysis to Interpret Complex Phenotypes** *3*

*Hong Yu, Jialiang Huang, Wei Zhang, and Jing-Dong J. Han*

##### 1.1 Introduction *3*

##### 1.2 Identification of Important Genes based on Network Topologies *5*

##### 1.2.1 Degree *5*

##### 1.2.2 Betweenness *6*

##### 1.2.3 Network Motifs *6*

##### 1.2.4 Hierarchical Structure *7*

##### 1.3 Inferring Information from Known Networks *8*

##### 1.3.1 Understanding Biological Functions based on Network Modularity *8*

##### 1.3.2 Inferring Functional Relationships and Novel Functional Genes Through Networks *8*

##### 1.3.3 Unraveling Transcriptional Regulations from Expression Data through Transcriptional Networks *9*

##### 1.3.4 Extracting the Pathway-Linked Regulators and Effectors based on Network Flows *10*

##### 1.4 Conclusions *10*

References *11*

#### **2 Stochastic Modeling of Gene Regulatory Networks** *13*

*Tianhai Tian*

##### 2.1 Introduction *13*

##### 2.2 Discrete Stochastic Simulation Methods *14*

##### 2.2.1 SSA *15*

##### 2.2.2 Accelerating $\tau$ -Leap Methods *16*

##### 2.2.3 Langevin Approach *19*

2.3	Discrete Stochastic Modeling	20
2.3.1	Stochastic Modeling Method	20
2.3.2	Toggle Switch with the SOS Pathway	22
2.3.3	Other Models for the Genetic Toggle Switch	24
2.4	Continuous Stochastic Modeling	26
2.4.1	Deterministic Models for the $\lambda$ Phage Network	26
2.4.2	Stochastic Models for External Noise	28
2.4.3	Deterministic Models with Threshold Values	29
2.4.4	Stochastic Switching	30
2.5	Stochastic Models for Both Internal and External Noise	31
2.5.1	Stochastic Models for Microarray Gene Expression Data	33
2.6	Conclusions	34
	References	35
<b>3</b>	<b>Modeling Expression Quantitative Trait Loci in Multiple Populations</b>	<b>39</b>
	<i>Ching-Lin Hsiao and Cathy S. J. Fann</i>	
3.1	Introduction	39
3.1.1	Data Structure in eQTL Studies	39
3.1.2	Current eQTL Studies	40
3.1.2.1	eQTL Studies in a Single Human Population	40
3.1.2.2	eQTL Studies in Multiple Human Populations	43
3.1.3	An Illustrated Example	45
3.1.4	Notations	46
3.2	IGM Method	47
3.2.1	Modeling SNP-GE Association in a Single Population	47
3.2.2	Integrating Hypotheses to Identify Common eQTL	48
3.2.3	Applying the IGM Method to HapMap Data	48
3.2.3.1	Characterizing Putative eQTL Identified by the IGM	49
3.3	CTWM	50
3.3.1	Modeling SNP-GE Association in Pooled Data by CTWM	50
3.3.2	Applying CTWM to HapMap Data	52
3.3.2.1	Characterizing Putative eQTL Identified by CTWM	52
3.3.2.2	Justification of Model Assumptions	53
3.4	CTWM-GS Method	53
3.4.1	Solving Normal Equations in CTWM	54
3.4.2	Estimators of BD and GS	55
3.4.3	Testing BD and GS	56
3.4.4	Applying CTWM-GS to HapMap Data	56
3.4.4.1	Applying the GS to Population Studies	57
3.5	Discussion	60
	References	61

**Part Two Inference of Gene Networks 67**

<b>4</b>	<b>Transcriptional Network Inference Based on Information Theory 69</b>
	<i>Patrick E. Meyer, Catharina Olsen, and Gianluca Bontempi</i>
4.1	Introduction 69
4.1.1	Notation 69
4.1.2	Formalization 71
4.1.3	Performance Measures in Undirected Network Inference 72
4.1.3.1	Precision–Recall (PR) Curves 73
4.1.3.2	F-Scores 74
4.1.4	Causal Subset Selection 74
4.2	Inference Based on Conditional Mutual Information 76
4.2.1	Constraint-Based Methods 77
4.2.2	Approximated Conditional Mutual Information 78
4.2.3	Variable Selection Algorithms 78
4.3	Inference Based on Pairwise Mutual Information 80
4.3.1	Relevance Network (RELNET) 80
4.3.2	Context Likelihood of Relatedness (CLR) 81
4.3.3	Chow–Liu Tree 81
4.3.4	Algorithm for the Reconstruction of Accurate Cellular Networks (ARACNE) 82
4.3.5	Minimum Redundancy Networks (MRNET) 83
4.4	Arc Orientation 84
4.4.1	Assessing Arc Orientation Methods 86
4.5	Conclusions 87
	References 87
<b>5</b>	<b>Elucidation of General and Condition-Dependent Gene Pathways Using Mixture Models and Bayesian Networks 91</b>
	<i>Sandra Rodriguez-Zas and Younhee Ko</i>
5.1	Introduction 91
5.2	Methodology 92
5.2.1	Network Learning Algorithms: Frequentist- and Bayesian MCMC-based algorithms 93
5.3	Applications 95
5.3.1	Elucidation of Gene Networks 95
5.3.2	Discovery of Condition-Dependent Gene Relationships 96
5.3.3	MCMC Mixture Bayesian Network 99
5.3.4	Computational Considerations 101
5.4	Conclusions 101
	References 102

<b>6</b>	<b>Multiscale Network Reconstruction from Gene Expression Measurements: Correlations, Perturbations, and “A Priori Biological Knowledge”</b>	<b>105</b>
	<i>Daniel Remondini and Gastone Castellani</i>	
6.1	Introduction	105
6.1.1	Complex Networks	105
6.1.2	Gene Interaction Networks from Gene Expression Measurements	107
6.2	“Perturbation Method”	108
6.3	Network Reconstruction by the Correlation Method from Time-Series Gene Expression Data	109
6.4	Network Reconstruction from Gene Expression Data by A Priori Biological Knowledge	110
6.5	Examples and Methods of Correlation Network Analysis on Time-Series Data	112
6.6	Examples and Methods for Pathway Network Analysis	117
6.6.1	Gene Selection and Pathway Grouping	118
6.6.2	Pathway Significance and Pathway Network	118
6.6.3	Results	121
6.7	Discussion	126
6.8	Conclusions	127
	References	128
 <b>7</b>	 <b>Gene Regulatory Networks Inference: Combining a Genetic Programming and <math>H_\infty</math> Filtering Approach</b>	 <b>133</b>
	<i>Lijun Qian, Haixin Wang, and Xiangfang Li</i>	
7.1	Introduction	133
7.2	Background	134
7.2.1	Noise in Gene Expression	134
7.2.2	Modeling of Gene Regulatory Networks with Noise	136
7.2.2.1	Boolean Networks Model with Noise	136
7.2.2.2	Bayesian Networks Model with Noise	136
7.2.2.3	Linear Additive Regulation Model with Noise	137
7.2.2.4	Neural Networks Model with Noise	137
7.2.3	Proposed Nonlinear ODE Model with Noise	138
7.3	Methodology for Identification and Algorithm Description	139
7.3.1	GP	140
7.3.2	$H_\infty$ Filter	142
7.4	Simulation Evaluation	144
7.4.1	Synthetic Data	144
7.4.2	Microarray Data	145
7.5	Conclusions	146
	References	150



<b>8</b>	<b>Computational Reconstruction of Protein Interaction Networks</b>	<b>155</b>
	<i>Konrad Mönks, Irmgard Mühlberger, Andreas Bernthaler, Raul Fechete, Paul Perco, Rudolf Freund, Arno Lukas, and Bernd Mayer</i>	
8.1	Introduction	155
8.1.1	Selecting Relevant Features from Omics Profiles	156
8.1.2	Analyzing Omics Data on a Network Level	157
8.2	Protein Interaction Networks	159
8.2.1	Network Categories	159
8.2.1.1	Metabolic Networks	159
8.2.1.2	Paralog Networks	160
8.2.1.3	Physical Interaction Networks	160
8.2.2	Parameters for Protein Annotation	161
8.2.2.1	Gene Expression Profiles	161
8.2.2.2	Subcellular Location	161
8.2.2.3	Gene Annotation	161
8.2.2.4	Transcription Factors	162
8.2.2.5	MicroRNA	162
8.2.2.6	Pathways	162
8.2.3	Data Preparation	163
8.2.3.1	Integration of Data Sources	163
8.2.3.2	Obtaining Edge Weights	164
8.2.3.3	Data Completeness	166
8.2.3.4	Data Normalization	166
8.2.4	Deriving Models	166
8.2.4.1	Basic Considerations	167
8.2.4.2	Choosing an Algorithm	169
8.2.5	Validation Procedures	169
8.2.5.1	Model Performance Evaluation	169
8.2.5.2	Network Structure Assessment	170
8.3	Characterization of Computed Networks	171
8.3.1	Evaluation of the Specific Protein–Protein Interactions	171
8.3.2	Application of the Specific Protein–Protein Interactions	175
8.4	Conclusions	177
	References	178

### Part Three Analysis of Gene Networks 181

<b>9</b>	<b>What if the Fit is Unfit? Criteria for Biological Systems Estimation Beyond Residual Errors</b>	<b>183</b>
	<i>Eberhard O. Voit</i>	
9.1	Introduction	183
9.2	Model Design	184
9.3	Concepts and Challenges of Parameter Estimation	187
9.3.1	Typical Parameter Estimation Problems	190
9.3.1.1	Data Fit is Unacceptable	190