

Contents

- 1 Analysis of Combinatorial Gene Regulation with Thermodynamic Models 1**
Chieh-Chun Chen and Sheng Zhong
 - 1.1 Introduction 1
 - 1.2 Thermodynamic Models for TF–DNA Binding 2
 - 1.2.1 TF–DNA Interactions 2
 - 1.2.2 TF–RNAP–DNA Interactions 3
 - 1.3 Models for Gene Expression 6
 - 1.3.1 Kinetic Model 7
 - 1.3.2 Logistic Model 7
 - 1.4 Reconstruction of Regulatory Networks 10
 - 1.4.1 Interaction-Identifier 10
 - 1.4.2 Network-Identifier 11
 - 1.5 Applications 11
 - 1.5.1 Analysis of Combinatorial *cis*-regulation in Synthetic Promoter in Yeast 12
 - 1.5.2 Predicting Spatial Expression Patterns from Sequence in *Drosophila* Segmentation 13
 - 1.5.3 Inferring Gene Regulatory Networks in Mouse Embryonic Stem Cells 15
 - 1.6 Concluding Remarks 16

- 2 RNA Secondary Structure Prediction and Gene Regulation by Small RNAs 19**
Ye Ding
 - 2.1 Introduction 19
 - 2.2 RNA Secondary Structure Prediction 20
 - 2.2.1 Free Energy Minimization 20
 - 2.2.2 Partition Function Approach 21
 - 2.2.3 Statistical Sampling Approach 22

2.2.4	Cluster and Centroid Representation of Boltzmann Ensemble	23
2.3	Gene Silencing by Small Interfering RNAs	24
2.3.1	Design Rules for Improving Potency	24
2.3.2	Structure Based Assessment of Target Accessibility	25
2.3.3	Specificity and Off-targeting	27
2.4	Posttranscriptional Gene Regulation by MicroRNAs	28
2.4.1	Target Identification Using Sequence Features	28
2.4.2	A Target Structure-Based Model for MicroRNA: Target Hybridization	29
2.5	Concluding Remarks	31
3	Some Critical Data Quality Control Issues of Oligoarrays	39
	Wenjiang J. Fu, Ming Li, Yalu Wen, and Likit Preeyanon	
3.1	Introduction	39
3.2	Quality Control in Microarray Data Analysis	41
3.3	Physico-Chemical Properties in Sequence Duplex Hybridization	43
3.4	The MM Phenomenon: $MM > PM$	46
3.5	Abundance of Gene Expression: Copy Number Versus Probe Intensity	51
3.6	Array Image Quality and Repair Through an Imputation Method with a Mixed Effects Model	54
3.7	Concluding Remarks	55
4	Stochastic-Process Approach to Nonequilibrium Thermodynamics and Biological Signal Transduction	61
	Hao Ge	
4.1	Introduction	61
4.1.1	Nonequilibrium Thermodynamics	61
4.1.2	Biological Signal Transduction	62
4.2	Stochastic-Process Approach: Examples	63
4.2.1	Mesoscopic Description of Biochemical Systems	63
4.2.2	Langevin Systems	65
4.3	Stochastic Thermodynamics	65
4.4	Ultrasensitivity and Temporal Cooperativity of PdPC Module	70
4.4.1	Reversible Kinetic Model for Covalent Modification	70
4.4.2	Reduced Models	71
4.4.3	Deterministic Model	72
4.4.4	Stochastic Model: Chemical Master Equation	73
4.4.5	Simple PdPC Switch: First-Order Approximation	74
4.4.6	Ultrasensitive PdPC Switch: Zero-Order Approximation	75
4.4.7	Mathematical Equivalence to Allosteric Cooperativity	76
4.5	Conclusion and Discussion	78
5	Granger Causality: Theory and Applications	83
	Shuixia Guo, Christophe Ladrone, and Jianfeng Feng	

5.1	Introduction	83
5.2	Partial Granger Causality	86
5.2.1	Time Domain Formulation	87
5.2.2	Numerical Example	89
5.3	Frequency Analysis	90
5.4	Group Interaction: Complex Granger Causality	93
5.4.1	Time Domain Formulation	93
5.4.2	Frequency Domain Formulation	94
5.4.3	Effect of Correlation Between Sources	96
5.5	Harmonic Granger Causality	97
5.5.1	Time Domain Formulation	97
5.5.2	Two Remarks About Harmonic Granger Causality	98
5.5.3	Frequency Domain Formulation	98
5.5.4	A Circadian Circuit	100
5.6	A Comparative Study Between Granger Causality and Bayesian Network	103
5.7	Unified Causal Model (UCM)	105
5.8	Large Networks	106
5.9	Summary	107
	Appendix: Estimating the Error Covariance Matrix	108
6	Transcription Factor Binding Site Identification by Phylogenetic Footprinting	113
	Haiyan Hu and Xiaoman Li	
6.1	Introduction	113
6.2	Current TFBS Identification Based on Alignments	116
6.3	Methods Independent of Alignments	119
6.3.1	Defects of Alignments Especially Genome Alignments	119
6.3.2	TFBS Identification Without Sequence Alignments	121
6.4	Future Direction of Phylogenetic Footprinting	126
7	Learning Network from High-Dimensional Array Data	133
	Li Hsu, Jie Peng, and Pei Wang	
7.1	Introduction	133
7.2	space	134
7.2.1	Model	134
7.2.2	Simulation	137
7.3	LogitNet	139
7.3.1	Model	139
7.3.2	Application to Genomic Instability Data	142
7.3.3	Simulation	143
7.4	remMap	146
7.4.1	Motivation and Model	146
7.4.2	Simulation	149
7.5	Real Application	150
7.6	Concluding Remarks	153

8	Computational Methods for Predicting Domain–Domain Interactions	157
	Hyunju Lee, Ting Chen, and Fengzhu Sun	
8.1	Introduction	157
8.1.1	Data Sources	159
8.1.2	Assessing the Accuracy of Predicted Domain–Domain Interactions	159
8.2	Computational Methods for Predicting Domain Interactions Using Protein–Protein Interactions	160
8.2.1	Predicting Domain Interactions Based on Over-represented Domain Pairs	160
8.2.2	Maximum Likelihood Estimation (MLE) Method	162
8.2.3	A Bayesian Method for Predicting Domain Interactions	164
8.2.4	A Likelihood-Ratio-Based Method: Domain Pair Exclusion Analysis (DPEA)	164
8.2.5	Maximum-Parsimony-Based Method–Linear Programming Optimization	166
8.3	Integrated Approaches for Predicting Domain Interactions	167
8.3.1	An Extended Likelihood Approach for Predicting Domain Interactions Based on Protein Interactions from Multiple Species	168
8.3.2	Predicting Domain Interactions from Multiple Data Sources	169
8.4	Discussion	170
9	Irreversible Stochastic Processes, Coupled Diffusions and Systems Biochemistry	175
	Pei-Zhe Shi and Hong Qian	
9.1	Introduction	175
9.2	Single-Molecule Michaelis–Menten Enzyme Kinetics and Irreversible Markov Processes	176
9.3	Coupled Diffusion	180
9.3.1	Fluctuating Enzymes	181
9.3.2	Motor Proteins	183
9.3.3	Self-regulating Genes	184
9.3.4	General Form	185
9.4	Limit Cases of Coupled Diffusion Processes	187
9.4.1	Limit Case: Fast Jump Process	187
9.4.2	Limit Case: Fast Diffusion	188
9.4.3	NESS Flux	189
9.4.4	Entropy Production	190
9.5	Stochastic Bifurcation	190
9.6	Numerical Methods	192
9.7	Discussion	194
9.8	Mathematical Methods	195
9.8.1	Proof of Sturm–Liouville Operator	195

9.8.2	Asymptotic Solution in the Limit of Fast Jump Process . . .	196
9.8.3	Asymptotic Solution in the Limit of Fast Diffusion	198
9.8.4	Bifurcation of the Toy Model	199
10	Probability Modeling and Statistical Inference in Periodic Cancer Screening	203
	Dongfeng Wu and Gary L. Rosner	
10.1	Background	203
10.2	Current Methods in Periodic Cancer Screening	205
10.2.1	MLE and Bayesian Inference of Age-Dependent Sensitivity and Transition Probability in Periodic Screening	206
10.2.2	Bayesian Inference for the Lead Time in Periodic Cancer Screening	208
10.2.3	Testing the Dependence of Two Screening Modalities	211
10.3	Future Developments in Cancer Screening	213
10.3.1	Evaluate Long Term Benefits of Periodic Cancer Screening	213
10.3.2	Sensitivity as a Function of Age, Time Spent in S_p and Sojourn Time	214
10.3.3	Optimal Scheduling for the Next Exam	215
10.3.4	Survival Benefit due to Periodic Screening	216
11	On Construction of the Smallest One-sided Confidence Intervals and Its Application in Identifying the Minimum Effective Dose . . .	219
	Weizhen Wang	
11.1	Introduction	219
11.2	The Smallest One-sided Confidence Interval	222
11.3	A Smallest Interval for the Difference of Two Independent Proportions	224
11.4	Identifying the Minimum Effective Dose	227
11.5	Discussion	228
	References	229
12	Group Variable Selection Methods and Their Applications in Analysis of Genomic Data	231
	Jun Xie and Lingmin Zeng	
12.1	Introduction	231
12.2	Background	232
12.2.1	Existing Variable Selection Methods	232
12.2.2	Large Scale Genomic Data	234
12.3	gLars and gRidge Algorithms	235
12.3.1	Simulation Studies	238
12.4	Unbiased Variable Selection via SCAD_ℓ2	241
12.5	Applications in Genomic Data Analysis	243
12.5.1	SNP Data Analysis	243
12.5.2	Gene Expression Data Analysis	245
12.6	Discussion	246

13	Modeling Protein-Signaling Networks with Granger Causality	
	Test	249
	Wenqiang Yang and Qiang Luo	
	13.1 Introduction	249
	13.2 Granger Causality and Approach	250
	13.3 Data and Results	252
	13.4 Conclusion	256
	References	256
14	DNA Copy Number Profiling in Normal and Tumor Genomes	259
	Nancy R. Zhang	
	14.1 Introduction	259
	14.2 Total Copy Number Estimation for One Sample	260
	14.3 Parent Specific Copy Number Estimation	263
	14.4 Integration of Multiple Array Platforms	267
	14.5 Modeling Recurrence Across Samples	270
	14.5.1 Post-Segmentation Procedures	271
	14.5.2 Cross-Sample Detection of Inherited Variants	273
	14.5.3 Obtaining a Cross-Sample Signature	276
	14.6 Concluding Remarks	277
15	Spatial Disease Surveillance: Methods and Applications	283
	Tonglin Zhang	
	15.1 Introduction	283
	15.2 Review of Cluster Detection Approaches	285
	15.2.1 Scan Statistics	286
	15.2.2 Permutation Testing Methods	289
	15.2.3 Other Methods	292
	15.3 Review of Disease Mapping Approaches	292
	15.4 Simulation and Case Study	293
	15.4.1 Simulation	294
	15.4.2 Case Study	296
	15.5 Concluding Remarks	298
16	From QTL Mapping to eQTL Analysis	301
	Wei Zhang and Jun S. Liu	
	16.1 Introduction	301
	16.2 Biological Background	302
	16.2.1 Genetic Experiments for eQTL Studies	302
	16.2.2 eQTL Hot Spots	303
	16.2.3 eQTL and cQTL	304
	16.3 Methods for QTL and eQTL Mappings	305
	16.3.1 Single QTL Model	305
	16.3.2 Multiple QTL Model	307
	16.3.3 Thresholding	309
	16.3.4 Multiple Trait Mapping	311

16.3.5	Regression Based Methods for eQTL Mapping	313
16.3.6	Bayesian Methods for Studying eQTLs	315
16.3.7	Bayesian Networks	316
16.3.8	Integrative Analysis	317
16.4	A Bayesian Partition Model for eQTL Mapping	318
16.5	Simulation Results	320
16.5.1	Simulation I	320
16.5.2	Simulation II	323
16.6	Discussion	325
17	An Evaluation of Gene Module Concepts in the Interpretation of Gene Expression Data	331
	Xianghua Zhang and Hongyu Zhao	
17.1	Introduction	331
17.2	Methods and Materials	333
17.2.1	WGCN Construction	333
17.2.2	Module Identification from WGCN	335
17.2.3	Enrichment Analysis	335
17.2.4	eQTL Analysis	336
17.2.5	Data Sets	336
17.3	Results	336
17.3.1	Identifying Modules from WGCN	336
17.3.2	Biological Interpretation of Gene Modules	339
17.3.3	Comparison Between Pearson Correlation and Topological Overlap	340
17.3.4	Consistency of Gene Modules	343
17.3.5	Genetic Basis of Gene Modules	344
17.4	Conclusions	345
18	Readout of Spike Waves in a Microcolumn	351
	Xuejuan Zhang	
18.1	Introduction	351
18.2	Theoretical Results	352
18.2.1	Distribution of Interspike Interval	352
18.2.2	MLE Decoding Strategy	355
18.2.3	Comparing with Rate Decoding	357
18.3	Applications	357
18.3.1	Decode Excitatory and Inhibitory Ratio in a Single Neuron with Stationary Input	357
18.3.2	Decode Dynamical Inputs in Networks Without Interactions	359
18.3.3	Decode Input Information in Networks with Interactions	362
18.4	Discussion	366
19	False Positive Control for Genome-Wide ChIP-Chip Tiling Arrays	371
	Yu Zhang	

19.1	Introduction	371
19.2	Methods	372
19.2.1	Poisson Approximation	373
19.2.2	Varying Window Sizes	374
19.3	Results	375
19.3.1	Simulation Study	375
19.3.2	Power of Various Window Sizes	376
19.3.3	FDR Control Accounting for Positive Correlations	377
19.4	Discussion	379
	Index	383
	Color Plates	385

Frontiers in Computational and Systems Biology

Feng, J.; Fu, W.; Sun, F. (Eds.)

2010, XXV, 24 p., Hardcover

ISBN: 978-1-84996-195-0