

Ion Măndoiu Alexander Zelikovsky (Eds.)

# Bioinformatics Research and Applications

Third International Symposium, ISBRA 2007  
Atlanta, GA, USA, May 7-10, 2007  
Proceedings



Springer

# Table of Contents

GFBA: A Biclustering Algorithm for Discovering Value-Coherent Biclusters . . . . .	1
<i>Xubo Fei, Shiyong Lu, Horia F. Pop, and Lily R. Liang</i>	
Significance Analysis of Time-Course Gene Expression Profiles . . . . .	13
<i>Fang-Xiang Wu</i>	
Data-Driven Smoothness Enhanced Variance Ratio Test to Unearth Responsive Genes in 0-Time Normalized Time-Course Microarray Data . . . . .	25
<i>Juntao Li, Jianhua Liu, and R. Krishna Murthy Karuturi</i>	
Efficiently Finding the Most Parsimonious Phylogenetic Tree Via Linear Programming . . . . .	37
<i>Srinath Sridhar, Fumei Lam, Guy E. Blesloch, R. Ravi, and Russell Schwartz</i>	
A Multi-Stack Based Phylogenetic Tree Building Method . . . . .	49
<i>Róbert Busa-Fekete, András Kocsor, and Csaba Bagyinka</i>	
A New Linear-Time Heuristic Algorithm for Computing the Parsimony Score of Phylogenetic Networks: Theoretical Bounds and Empirical Performance . . . . .	61
<i>Guohua Jin, Luay Nakhleh, Sagi Snir, and Tamir Tuller</i>	
A Bootstrap Correspondence Analysis for Factorial Microarray Experiments with Replications . . . . .	73
<i>Qihua Tan, Jesper Dahlgaard, Basem M. Abdallah, Werner Vach, Moustapha Kassem, and Torben A. Kruse</i>	
Clustering Algorithms Optimizer: A Framework for Large Datasets . . . . .	85
<i>Roy Varshavsky, David Horn, and Michal Linial</i>	
Ranking Function Based on Higher Order Statistics (RF-HOS) for Two-Sample Microarray Experiments . . . . .	97
<i>Jahangheer Shaik and Mohammed Yasin</i>	
Searching for Recombinant Donors in a Phylogenetic Network of Serial Samples . . . . .	109
<i>Patricia Buendia and Giri Narasimhan</i>	
Algorithm for Haplotype Inferring Via Galled-Tree Networks with Simple Galls . . . . .	121
<i>Arvind Gupta, Ján Maňuch, Ladislav Stacho, and Xiaohong Zhao</i>	

Estimating Bacterial Diversity from Environmental DNA: A Maximum Likelihood Approach . . . . .	133
<i>Frederick Cohan, Danny Krizanc, and Yun Lu</i>	
Invited Talk: Modern Homology Search . . . . .	145
<i>Ming Li</i>	
Statistical Absolute Evaluation of Gene Ontology Terms with Gene Expression Data . . . . .	146
<i>Pramod K. Gupta, Ryo Yoshida, Seiya Imoto, Rui Yamaguchi, and Satoru Miyano</i>	
Discovering Relations Among GO-Annotated Clusters by Graph Kernel Methods . . . . .	158
<i>Italo Zoppis, Daniele Merico, Marco Antoniotti, Bud Mishra, and Giancarlo Mauri</i>	
An Empirical Comparison of Dimensionality Reduction Methods for Classifying Gene and Protein Expression Datasets . . . . .	170
<i>George Lee, Carlos Rodriguez, and Anant Madabhushi</i>	
NEURONgrid: A Toolkit for Generating Parameter-Space Maps Using NEURON in a Grid Environment . . . . .	182
<i>Robert J. Calin-Jageman, Chao Xie, Yi Pan, Art Vandenberg, and Paul S. Katz</i>	
An Adaptive Resolution Tree Visualization of Large Influenza Virus Sequence Datasets . . . . .	192
<i>Leonid Zaslavsky, Yiming Bao, and Tatiana A. Tatusova</i>	
Wavelet Image Interpolation (WII): A Wavelet-Based Approach to Enhancement of Digital Mammography Images . . . . .	203
<i>Gordana Derado, F. DuBois Bowman, Rajan Patel, Mary Newell, and Brani Vidakovic</i>	
High Level Programming Environment System for Protein Structure Data . . . . .	215
<i>Yanchao Wang, Rajshekhar Sunderraman, and Piyaphol Phoungphol</i>	
Finding Minimal Sets of Informative Genes in Microarray Data . . . . .	227
<i>Kung-Hua Chang, Yong Kyun Kwon, and D. Stott Parker</i>	
Noise-Based Feature Perturbation as a Selection Method for Microarray Data . . . . .	237
<i>Li Chen, Dmitry B. Goldgof, Lawrence O. Hall, and Steven A. Eschrich</i>	
Efficient Generation of Biologically Relevant Enriched Gene Sets . . . . .	248
<i>Igor Trajkovski and Nada Lavrač</i>	

Space and Time Efficient Algorithms to Discover Endogenous RNAi Patterns in Complete Genome Data . . . . .	260
<i>Sudha Balla and Sanguthevar Rajasekaran</i>	
A Fast Approximate Covariance-Model-Based Database Search Method for Non-coding RNA . . . . .	270
<i>Scott F. Smith</i>	
Extensions of Naive Bayes and Their Applications to Bioinformatics . . . . .	282
<i>Raja Loganantharaj</i>	
The Solution Space of Sorting by Reversals . . . . .	293
<i>Marília D.V. Braga, Marie-France Sagot, Celine Scornavacca, and Eric Tannier</i>	
A Fast and Exact Algorithm for the Perfect Reversal Median Problem . . . . .	305
<i>Matthias Bernt, Daniel Merkle, and Martin Middendorf</i>	
Genomic Signatures from DNA Word Graphs . . . . .	317
<i>Lenwood S. Heath and Amrita Pati</i>	
Enhancing Motif Refinement by Incorporating Comparative Genomics Data . . . . .	329
<i>Erliang Zeng and Giri Narasimhan</i>	
Mining Discriminative Distance Context of Transcription Factor Binding Sites on ChIP Enriched Regions . . . . .	338
<i>Hyunmin Kim, Katherina J. Kechris, and Lawrence Hunter</i>	
Enhanced Prediction of Cleavage in Bovine Precursor Sequences . . . . .	350
<i>Allison N. Tegge, Sandra L. Rodriguez-Zas, J.V. Sweedler, and Bruce R. Southey</i>	
Invited Talk: A Computational Study of Bidirectional Promoters in the Human Genome . . . . .	361
<i>Mary Qu Yang and Laura L. Elnitski</i>	
The Identification of Antisense Gene Pairs Through Available Software . . . . .	372
<i>Mark J. Lawson and Liqing Zhang</i>	
Inferring Weak Adaptations and Selection Biases in Proteins from Composition and Substitution Matrices . . . . .	382
<i>Steinar Thorvaldsen, Elinor Ytterstad, and Tor Flå</i>	
Markov Model Variants for Appraisal of Coding Potential in Plant DNA . . . . .	394
<i>Michael E. Sparks, Volker Brendel, and Karin S. Dorman</i>	

Predicting Palmitoylation Sites Using a Regularised Bio-basis Function Neural Network . . . . .	406
<i>Zheng Rong Yang</i>	
A Novel Kernel-Based Approach for Predicting Binding Peptides for HLA Class II Molecules . . . . .	418
<i>Hao Yu, Minlie Huang, Xiaoyan Zhu, and Yabin Guo</i>	
A Database for Prediction of Unique Peptide Motifs as Linear Epitopes . . . . .	430
<i>Margaret Dah-Tsyng Chang, Hao-Teng Chang, Rong-Yuan Huang, Wen-Shyong Tzou, Chih-Hong Liu, Wei-Jun Zhung, Hsien-Wei Wang, Chun-Tien Chang, and Tun-Wen Pai</i>	
A Novel Greedy Algorithm for the Minimum Common String Partition Problem . . . . .	441
<i>Dan He</i>	
An Efficient Algorithm for Finding Gene-Specific Probes for DNA Microarrays . . . . .	453
<i>Mun-Ho Choi, In-Seon Jeong, Seung-Ho Kang, and Hyeong-Seok Lim</i>	
Multiple Sequence Local Alignment Using Monte Carlo EM Algorithm . . . . .	465
<i>Chengpeng Bi</i>	
Cancer Class Discovery Using Non-negative Matrix Factorization Based on Alternating Non-negativity-Constrained Least Squares . . . . .	477
<i>Hyunsoo Kim and Haesun Park</i>	
A Support Vector Machine Ensemble for Cancer Classification Using Gene Expression Data . . . . .	488
<i>Chen Liao and Shutao Li</i>	
Combining SVM Classifiers Using Genetic Fuzzy Systems Based on AUC for Gene Expression Data Analysis . . . . .	496
<i>Xiujuan Chen, Yichuan Zhao, Yan-Qing Zhang, and Robert Harrison</i>	
A BP-SCFG Based Approach for RNA Secondary Structure Prediction with Consecutive Bases Dependency and Their Relative Positions Information . . . . .	506
<i>Dandan Song and Zhidong Deng</i>	
Delta: A Toolset for the Structural Analysis of Biological Sequences on a 3D Triangular Lattice . . . . .	518
<i>Minghui Jiang, Martin Mayne, and Joel Gillespie</i>	
Statistical Estimate for the Size of the Protein Structural Vocabulary . . . . .	530
<i>Xuezheng Fu, Bernard Chen, Yi Pan, and Robert W. Harrison</i>	

Coclustering Based Parcellation of Human Brain Cortex Using Diffusion Tensor MRI .....	539
<i>Cui Lin, Shiyong Lu, Danqing Wu, Jing Hua, and Otto Muzik</i>	
An Algorithm for Hierarchical Classification of Genes of Prokaryotic Genomes .....	551
<i>Hongwei Wu, Fenglou Mao, Victor Olman, and Ying Xu</i>	
Using Multi Level Nearest Neighbor Classifiers for G-Protein Coupled Receptor Sub-families Prediction .....	564
<i>Mudassir Fayyaz, Asifullah Khan, Adnan Mujahid, and Alex Kavokin</i>	
Invited Talk: Ab Initio Gene Finding Engines: What Is Under the Hood .....	577
<i>Mark Borodovsky</i>	
Reconstruction of 3D Structures from Protein Contact Maps .....	578
<i>Marco Vassura, Luciano Margara, Filippo Medri, Pietro di Lena, Piero Fariselli, and Rita Casadio</i>	
A Feature Selection Algorithm Based on Graph Theory and Random Forests for Protein Secondary Structure Prediction .....	590
<i>Gulsah Altun, Hae-Jin Hu, Stefan Gremalschi, Robert W. Harrison, and Yi Pan</i>	
DNA Sites Buried in Nucleosome Become Accessible at Room Temperature: A Discrete-Event-Simulation Based Modeling Approach .....	601
<i>Amin R. Mazloom, Kalyan Basu, Subhrangsu S. Mandal, Mehran Sorourian, and Sajal Das</i>	
Comparative Analysis of Gene-Coexpression Networks Across Species .....	615
<i>Shiquan Wu and Jing Li</i>	
Comparative Pathway Prediction Via Unified Graph Modeling of Genomic Structure Information .....	627
<i>Jizhen Zhao, Dongsheng Che, and Liming Cai</i>	
Extending the Calculus of Looping Sequences to Model Protein Interaction at the Domain Level .....	638
<i>Roberto Barbuti, Andrea Maggiolo-Schettini, and Paolo Milazzo</i>	
<b>Author Index</b> .....	651