

CONTENTS

Committees	vii
Referees	ix
Preface	xi
Keynote Addresses	
Exploring the Ocean's Microbes: Sequencing the Seven Seas <i>Marvin E. Frazier et al.</i>	1
Don't Know Much About Philosophy: The Confusion Over Bio-Ontologies <i>Mark A. Musen</i>	3
Invited Talks	
Biomedical Informatics Research Network (BIRN): Building a National Collaboratory for BioMedical and Brain Research <i>Mark H. Ellisman</i>	5
Protein Network Comparative Genomics <i>Trey Ideker</i>	7
Systems Biology in Two Dimensions: Understanding and Engineering Membranes as Dynamical Systems <i>Erik Jakobsson</i>	9
Bioinformatics at Microsoft Research <i>Simon Mercer</i>	11
Movie Crunching in Biological Dynamic Imaging <i>Jean-Christophe Olivo-Marin</i>	13
Engineering Nucleic Acid-Based Molecular Sensors for Probing and Programming Cellular Systems <i>Christina D. Smolke</i>	15
Reactome: A Knowledgebase of Biological Pathways <i>Lincoln Stein et al.</i>	17

Structural Bioinformatics

- Effective Optimization Algorithms for Fragment-Assembly based Protein Structure Prediction 19
Kevin W. DeRonne and George Karypis
- Transmembrane Helix and Topology Prediction Using Hierarchical SVM Classifiers and an Alternating Geometric Scoring Function 31
Allan Lo, Hua-Sheng Chiu, Ting-Yi Sung and Wen-Lian Hsu
- Protein Fold Recognition Using the Gradient Boost Algorithm 43
Feng Jiao, Jinbo Xu, Libo Yu and Dale Schuurmans
- A Graph-Based Automated NMR Backbone Resonance Sequential Assignment 55
Xiang Wan and Guohui Lin
- A Data-Driven, Systematic Search Algorithm for Structure Determination of Denatured or Disordered Proteins 67
Lincong Wang and Bruce Randall Donald
- Multiple Structure Alignment by Optimal RMSD Implies that the Average Structure is a Consensus 79
Xueyi Wang and Jack Snoeyink
- Identification of α -Helices from Low Resolution Protein Density Maps 89
Alessandro Dal Palù, Enrico Pontelli, Jing He and Yonggang Lu
- Efficient Annotation of Non-Coding RNA Structures Including Pseudoknots via Automated Filters 99
Chunmei Liu, Yinglei Song, Ping Hu, Russell L. Malmberg and Liming Cai
- Thermodynamic Matchers: Strengthening the Significance of RNA Folding Energies 111
Thomas Höchsmann, Matthias Höchsmann and Robert Giegerich

Microarray Data Analysis and Applications

- PEM: A General Statistical Approach for Identifying Differentially Expressed Genes in Time-Course cDNA Microarray Experiment without Replicate 123
Xu Han, Wing-Kin Sung and Lin Feng
- Efficient Generalized Matrix Approximations for Biomarker Discovery and Visualization in Gene Expression Data 133
Wenyuan Li, Yanxiong Peng, Hung-Chung Huang and Ying Liu

Computational Genomics and Genetics

- Efficient Computation of Minimum Recombination with Genotypes (not Haplotypes) 145
Yufeng Wu and Dan Gusfield
- Sorting Genomes by Translocations and Deletions 157
Xingqin Qi, Guojun Li, Shuguang Li and Ying Xu

Turning Repeats to Advantage: Scaffolding Genomic Contigs Using LTR Retrotransposons <i>Ananth Kalyanaraman, Srinivas Aluru and Patrick S. Schnable</i>	167
Whole Genome Composition Distance for HIV-1 Genotyping <i>Xiaomeng Wu, Randy Goebel, Xiu-Feng Wan and Guohui Lin</i>	179
Efficient Recursive Linking Algorithm for Computing the Likelihood of an Order of a Large Number of Genetic Markers <i>S. Tewari, S. M. Bhandarkar and J. Arnold</i>	191
Optimal Imperfect Phylogeny Reconstruction and Haplotyping (IPPH) <i>Srinath Sridhar, Guy E. Blelloch, R. Ravi and Russell Schwartz</i>	199
Toward an Algebraic Understanding of Haplotype Inference by Pure Parsimony <i>Daniel G. Brown and Ian M. Harrower</i>	211
Global Correlation Analysis Between Redundant Probe Sets Using a Large Collection of Arabidopsis ATH1 Expression Profiling Data <i>Xiangqin Cui and Ann Loraine</i>	223
Motif Sequence Identification	
Distance-Based Identification of Structure Motifs in Proteins Using Constrained Frequent Subgraph Mining <i>Jun Huan, Deepak Bandyopadhyay, Jan Prins, Jack Snoeyink, Alexander Tropsha and Wei Wang</i>	227
An Improved Gibbs Sampling Method for Motif Discovery via Sequence Weighting <i>Xin Chen and Tao Jiang</i>	239
Detection of Cleavage Sites for HIV-1 Protease in Native Proteins <i>Liwen You</i>	249
A Methodology for Motif Discovery Employing Iterated Cluster Re-Assignment <i>Osman Abul, Finn Drabløs and Geir Kjetil Sandve</i>	257
Biological Pathways and Systems	
Identifying Biological Pathways via Phase Decomposition and Profile Extraction <i>Yi Zhang and Zhidong Deng</i>	269
Expectation-Maximization Algorithms for Fuzzy Assignment of Genes to Cellular Pathways <i>Liviu Popescu and Golan Yona</i>	281
Classification of Drosophila Embryonic Developmental Stage Range Based on Gene Expression Pattern Images <i>Jieping Ye, Jianhui Chen, Qi Li and Sudhir Kumar</i>	293
Evolution versus “Intelligent Design”: Comparing the Topology of Protein-Protein Interaction Networks to the Internet <i>Qiaofeng Yang, Georgos Siganos, Michalis Faloutsos and Stefano Lonardi</i>	299

Protein Functions and Computational Proteomics

Cavity-Aware Motifs Reduce False Positives in Protein Function Prediction <i>Brian Y. Chen, Drew H. Bryant, Viacheslav Y. Fofanov, David M. Kristensen, Amanda E. Cruess, Marek Kimmel, Olivier Lichtarge and Lydia E. Kavraki</i>	311
Protein Subcellular Localization Prediction Based on Compartment-Specific Biological Features <i>Chia-Yu Su, Allan Lo, Hua-Sheng Chiu, Ting-Yi Sung and Wen-Lian Hsu</i>	325
Predicting the Binding Affinity of MHC Class II Peptides <i>Fatih Altiparmak, Altuna Akalin and Hakan Ferhatosmanoglu</i>	331
Codon-Based Detection of Positive Selection Can be Biased by Heterogeneous Distribution of Polar Amino Acids Along Protein Sequences <i>Xuhua Xia and Sudhir Kumar</i>	335
Bayesian Data Integration: A Functional Perspective <i>Curtis Huttenhower and Olga G. Troyanskaya</i>	341
An Iterative Algorithm to Quantify the Factors Influencing Peptide Fragmentation for MS/MS Spectrum <i>Chungong Yu, Yu Lin, Shiwei Sun, Jinjin Cai, Jingfen Zhang, Zhuo Zhang, Runsheng Chen and Dongbo Bu</i>	353
Complexity and Scoring Function of MS/MS Peptide De Novo Sequencing <i>Changjiang Xu and Bin Ma</i>	361
Biomedical Applications	
Expectation-Maximization Method for Reconstructing Tumor Phylogenies from Single-Cell Data <i>Gregory Pennington, Charles A. Smith, Stanley Shackney and Russell Schwartz</i>	371
Simulating <i>In Vitro</i> Epithelial Morphogenesis in Multiple Environments <i>Mark R. Grant, Sean H. J. Kim and C. Anthony Hunt</i>	381
A Combined Data Mining Approach for Infrequent Events: Analyzing HIV Mutation Changes Based on Treatment History <i>Ray S. Lin, Soo-Yon Rhee, Robert W. Shafer and Amar K. Das</i>	385
A Systems Biology Case Study of Ovarian Cancer Drug Resistance <i>Jake Y. Chen, Changyu Shen, Zhong Yan, Dawn P. G. Brown and Mu Wang</i>	389
Author Index	399