

## CONTENTS

Preface	v
APBC 2007 Organization	vii
<b>Keynote Papers</b>	
Exploring Genomes of Distantly Related Mammals <i>J.A. Marshall Graves</i>	1
Bugs, Guts and Fat - A Systems Approach to the Metabolic 'Axis of Evil' <i>J. Nadeau</i>	3
Protein Identification via Spectral Networks Analysis <i>P. Pevzner</i>	5
<b>Contributed Papers</b>	
Metagenome Analysis using MEGAN <i>D.H. Huson, A.F. Auch, J. Qi, and S.C. Schuster</i>	7
Algorithmic Approaches to Selecting Control Clones in DNA Array Hybridization Experiments <i>Q. Fu, E. Bent, J. Borneman, M. Chrobak, and N. Young</i>	17
Subtle Motif Discovery for Detection of DNA Regulatory Sites <i>M. Comin, and L. Parida</i>	27
An Effective Promoter Detection Method using the Adaboost Algorithm <i>X. Xie, S. Wu, K.-M. Lam, and H. Yan</i>	37
A New Strategy of Geometrical Biclustering for Microarray Data Analysis <i>H. Zhao, A.W.C. Liew, and H. Yan</i>	47

Using Formal Concept Analysis for Microarray Data Comparison <i>V. Choi, Y. Huang, V. Lam, D. Potter, R. Laubenbacher, and K. Duca</i>	57
An Efficient Biclustering Algorithm for Finding Genes with Similar Patterns in Time-series Expression Data <i>S.C. Madeira, and A.L. Oliveira</i>	67
Selecting Genes with Dissimilar Discrimination Strength for Sample Class Prediction <i>Z. Cai, R. Goebel, M.R. Salavatipour, Y. Shi, L. Xu, and G. Lin</i>	81
Computing the All-Pairs Quartet Distance on a Set of Evolutionary Trees <i>M. Stissing, T. Mailund, C.N.S. Pedersen, G.S. Brodal, and R. Fagerberg</i>	91
Computing the Quartet Distance Between Evolutionary Trees of Bounded Degree <i>M. Stissing, C.N.S. Pedersen, T. Mailund, G.S. Brodal, and R. Fagerberg</i>	101
A Global Maximum Likelihood Super-Quartet Phylogeny Method <i>P. Wang, B.B. Zhou, M. Taraeneh, D. Chu, C. Wang, A.Y. Zomaya, and R.P. Brent</i>	111
A Randomized Algorithm for Comparing Sets of Phylogenetic Trees <i>S.-J. Sul, and T.L. Williams</i>	121
Protein Structure-Structure Alignment with Discrete Fréchet Distance <i>M. Jiang, Y. Xu, and B. Zhu</i>	131
Deriving Protein Structure Topology from the Helix Skeleton in Low Resolution Density Map using Rosetta <i>Y. Lu, J. He, and C.E.M. Strauss</i>	143
Fitting Protein Chains to Cubic Lattice is NP-Complete <i>J. Mañuch, and D.R. Gaur</i>	153
Inferring a Chemical Structure from a Feature Vector Based on Frequency of Labeled Paths and Small Fragments <i>T. Akutsu, and D. Fukagawa</i>	165

Exact and Heuristic Approaches for Identifying Disease-Associated SNP Motifs	175
<i>G. Huang, P. Jeavons, and D. Kwiatkowski</i>	
Genotype-Based Case-Control Analysis, Violation of Hardy-Weinberg Equilibrium, and Phase Diagrams	185
<i>Y.J. Suh, and W. Li</i>	
A Probabilistic Method to Identify Compensatory Substitutions for Pathogenic Mutations	195
<i>B.C. Easton, A.V. Isaev, G.A. Huttley, and P. Maxwell</i>	
Exploring Genome Rearrangements using Virtual Hybridization	205
<i>M. Belcaid, A. Bergeron, A. Chateau, C. Chauve, Y. Gingras, G. Poisson, and M. Vendette</i>	
Two Plus Two Does not Equal Three: Statistical Tests for Multiple Genome Comparison	215
<i>N. Raghupathy, R. Hoberman, and D. Durand</i>	
The Distance Between Randomly Constructed Genomes	227
<i>W. Xu</i>	
Computing the Breakpoint Distance between Partially Ordered Genomes	237
<i>Z. Fu, and T. Jiang</i>	
Inferring Gene Regulatory Networks by Machine Learning Methods	247
<i>J. Supper, H. Fröhlich, C. Spieth, A. Dräger, and A. Zell</i>	
A Novel Clustering Method for Analysis of Biological Networks using Maximal Components of Graphs	257
<i>M. Hayashida, T. Akutsu, and H. Nagamochi</i>	
Gene Regulatory Network Inference via Regression Based Topological Refinement	267
<i>J. Supper, H. Fröhlich, and A. Zell</i>	
Algorithm Engineering for Color-Coding to Facilitate Signaling Pathway Detection	277
<i>F. Hüffner, S. Wernicke, and T. Zichner</i>	

De Novo Peptide Sequencing for Mass Spectra Based on Multi-Charge Strong Tags	287
<i>K. Ning, K.F. Chong, and H.W. Leong</i>	
Complexities and Algorithms for Glycan Structure Sequencing using Tandem Mass Spectrometry	297
<i>B. Shan, B. Ma, K. Zhang, and G. Lajoie</i>	
Semi-supervised Pattern Learning for Extracting Relations from Bioscience Texts	307
<i>S. Ding, M. Huang, and X. Zhu</i>	
Flow Model of the Protein-protein Interaction Network for Finding Credible Interactions	317
<i>K. Okada, K. Asai, and M. Arita</i>	
All Hits All The Time: Parameter Free Calculation of Seed Sensitivity	327
<i>D.Y.F. Mak, and G. Benson</i>	
Fast Structural Similarity Search Based on Topology String Matching	341
<i>S.-H. Park, D. Gilbert, and K.H. Ryu</i>	
Simple and Fast Alignment of Metabolic Pathways by Exploiting Local Diversity	353
<i>S. Wernicke, and F. Rasche</i>	
Combining N-grams and Alignment in G-protein Coupling Specificity Prediction	363
<i>B.Y.M. Chen, and J.G. Carbonell</i>	
Author Index	373