2nd International Conference on Bioinformatics and Computational Biology 2010

(BICoB-2010)

Honolulu, Hawaii, USA
24-26 March 2010

Editors:

H. Al-Mubaid

ISBN: 978-1-61738-111-9

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2010) by the International Society for Computers and Their Applications
All rights reserved.

Printed by Curran Associates, Inc. (2010)

For permission requests, please contact the International Society for Computers and Their Applications at the address below.

International Society for Computers and Their Applications
975 Walnut Street, Suite 132
Cary, NC 27511-4216

Phone: (919) 467-5559
Fax: (919) 467-3430

isca@ipass.net

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

INTERNATIONAL SOCIETY FOR COMPUTERS AND THEIR APPLICATIONS

2nd International Conference on Bioinformatics and Computational Biology

March 24-26, 2010
Sheraton Waikiki Hotel
Honolulu, Hawaii, USA

TECHNICAL PAPER INDEX

MACHINE LEARNING IN BIOINFORMATICS

A GA based Approach for Image Shadow Removal using PCNN
Gh. Rezaei Rad and H. Jahanirad (Iran University of Science and Technology, Iran) 1

Multiple Kernel Learning for Fold Recognition
Huzefa Rangwala (George Mason University, USA) 7

Iterative Split Adjustment for Building Multilabel Decision Trees
Aiyesha Ma and Ishwar Sethi (Oakland University, USA) 13

TEXT/LINGUISTIC APPLICATIONS

Space-efficient Genome Comparisons with Compressed Full-text Indexes
Enno Ohlebusch and Simon Gog (University of Ulm, Germany) 19

Specifications of Word Set in Linguistic Approach for Similarity Estimation
Satoshi Mizuta and Keishin Hanya (Hirosaki University, Japan) 25

Phylogenomic Analysis Using Bayesian Congruence Measuring
Dazhuo Li, Eric C. Rouchka and Patrick Shafto (University of Louisville, USA) 30

BIOLOGICAL AND PPI NETWORKS

GraMoFoNe: a Cytoscape Plugin for Querying Motifs without Topology in Protein-Protein Interactions Networks
Guillaume Blin, Florian Sikora and Stéphane Vialette (Université Paris-Est, France) 38

Biological Relevance Detection via Network Dynamic Analysis
Zheng Zhao, Huan Liu, Jiangxin Wang and Yung Chang (Arizona State University, USA) 44

Metabolic Pathways Modeling using VHDL-AMS
Mary M. Randolph-Gips (University of Houston-Clear Lake, USA) 50

PROTEIN APPLICATIONS

Predicting Subcellular Localizations of Gram-negative Proteins with Improved Performance
Jing Hu (Franklin & Marshall College, USA) 55

Bioinformatics Analysis of Physichemical Properties of Protein Sorting Signals
Fan Zhang and Jianjun Hu (University of South Carolina, USA) 61

Study of Computing Consolidation Techniques in Computational Protein Loop Structure Modeling
Yaohang Li and Douglas Wardell (North Carolina A&T State University, USA) 67

BIOINFORMATICS AND DISEASE RELATED TASKS

Delay Differential Models in Dynamic Diseases
Fathalla A. Rihan (United Arab Emirates University, UAE) 73

Disease Classification using Gene Similarity and GO Annotations
Hisham Al-Mubaid (University of Houston-Clear Lake, USA) 80

Identification of Similarities among MS/MS Spectra from Human Brain
Alan J. Barton, Arsalan S. Haqqani and Julio J. Valdés (National Research Council Canada, Canada) 86

On Building a Quantitative Food-Disease-Gene Network
Hui Yang, Abhishek Sharma, Rajesh Swaminathan and Vilas Ketkar (San Francisco State University, USA) 92

DNA APPLICATIONS

Algebraic Theory of DNA Recombination
Sergei R. Sverchkov (Novosibirsk State University, Russia) 98

Computational and Biological Approach to Understand the 3' UTR-Mediated Posttranscriptional Regulation of Alfalfa Cytosolic Glutamine Synthetase Genes
Bindu Simon, Champa Sengupta-Gopalan, Pavan K. Yelavarthi (New Mexico State University, USA) and Desh Ranjan (Old Dominion University, USA) 103

DNADigest: In silico Tool for Customized Digestion of Nucleic Acid Sequences
Sina Vivekanandan, Avinash Panigrahi, Sunil Kumar and Randeep Singh (Philips Research Asia-Bangalore, India) 109

ALGORITHMIC TECHNIQUES WITH BIOINFORMATICS APPLICATIONS

A New Relational KMR Based Algorithm for Microsatellite Identification
Nahla El Zant El Kadhi (Ahlia University, Bahrain) 115

SMS-Forbid: An Efficient Agorithm for Simple Motif Problem
Tarek El Falah (University of Tunis, Tunisia and University of Rouen, France), Thierry Lecroq (University of Rouen, France) and Mourad Elloumi (University of Tunis, Tunisia) 121

Triplet-Supertrees constructed from Minimum Triplet Presentations
Max Homilius (Max-Planck-Institute for Mol. Genetics, Germany), J. Gordon Burleigh (University of Florida, USA) and Oliver Eulenstein (Iowa State University, USA) 127

GENE AND PROTEIN APPLICATIONS

Scaling the Gene Duplication Problem Towards the Tree of Life
André Wehe (Iowa State University, USA) and J. Gordon Burleigh (University of Florida, Gainesville, USA) 133

Nonlinear Retention Time Shifts in nanoLC-MS Data: A Fuzzy-Evolutionary Computational Proteomics Approach
Alan J. Barton and Arsalan S. Haqqani (National Research Council Canada, Canada) 139

Folding-aware and Structure-conscious 3D Substructures in Folding Data: Identification and Applications
Hui Yang, Saurabh Gupte, Runtang Wang and Rahul Parte (San Francisco State University, USA) 145

RNA SEQUENCES and CLUSTERING

Clustering Micro-RNA Array Data Using an Information Fusion-Based Approach with Multiple Types of Input Data
Jane Synnergren, Alper Özdogan and Björn Olsson (University of Skövde, Sweden) and Peter Sartipy (Cellartis AB, Sweden) 151

miRSAT & miRCDB: An Integrated MicroRNA Sequesce Analysis Tool and a Cancer-Associated MicroRNA Database
Boya Xie, Robert Hochberg, Qin Ding and Di Wu (East Carolina University, USA) 159

A Dynamic Programming Algorithm for Finding the Optimal Segmentation of an RNA Sequence in Secondary Structure Predictions
Abel Licon, Michela Taufer (University of Delaware, USA) and Ming-Ying Leung, Kyle L. Johnson (The University of Texas at El Paso, USA) 165

DATA MINING AND MACHINE LEARNING IN BIOINFORMATICS

Integrative Biomarker Discovery for Breast Cancer Metastasis from Gene Expression and Protein Interaction Data Using Error-tolerant Pattern Mining
Rohit Gupta, Smita Agrawal, Navneet Rao, Ze Tian, Rui Kuang, and Vipin Kumar (University of Minnesota, USA) 171

A Method to Identify Protein Names with Iterative Extension of Training Data Set
Kazunori Miyanishi,Tomonobu Ozaki and Takenao Ohkawa (Kobe University, Japan) 178

Novel Morphological Phenotypes Discovery in High-Content Screens Using Underused Features
Chen Lin and Pengyu Hong (Brandeis University, USA), Chris Bakal (The Institute of Cancer Research London, UK), and Norbert Perrimon (Harvard Medical School, USA) 184

An Improved Model for Gene Cluster Inference
Xiao Yang and Srinivas Aluru (Iowa State University, USA) 190

TREES AND GRIDS WITH BIOINFORMATICS APPLICATIONS

Corrected Statistical Distance and Probability Constrained Optimization in Reconstruction of Phylogenetic Trees
D. A. Thomas, J. F. Weng and I. Mareels (The University of Melbourne, Australia) 196

A Grid-Enabled Problem Solving Environment for QTL Analysis in R
Mahen Jayawardena, Carl Nettelblad, Salman Zubair Toor (Uppsala University, Sweden), Per-Olov Östberg, Erik Elmroth (Umeä University, Sweden), and Sverker Holmgren (Uppsala University, Sweden) 202

A More Efficient Closest String Problem
Ruixuan Zhao (Sun Yat-Sen University, China) and Ning Zhang (University of Waterloo, Canada) 210

Author Index