Lecture Notes in Bioinformatics 6674
Edited by S. Istrail, P. Pevzner, and M. Waterman

Editorial Board: A. Apostolico S. Brunak M. Gelfand

T. Lengauer S. Miyano G. Myers M.-F. Sagot D. Sankoff

R. Shamir T. Speed M. Vingron W. Wong

Subseries of Lecture Notes in Computer Science

Jianer Chen Jianxin Wang
Alexander Zelikovsky (Eds.)

Bioinformatics Research and Applications

7th International Symposium, ISBRA 2011
Changsha, China, May 27-29, 2011
Proceedings

Springer

Series Editors

Sorin Istrail, Brown University, Providence, RI, USA
Pavel Pevzner, University of California, San Diego, CA, USA
Michael Waterman, University of Southern California, Los Angeles, CA, USA

Volume Editors

Jianer Chen
Texas A&M University, Department of Computer Science
College Station, TX 77843-3112, USA
E-mail: chen@cs.tamu.edu

Jianxin Wang
Central South University, School of Information Science and Engineering
Changsha, 410083, China
E-mail: jxwang@mail.csu.edu.cn

Alexander Zelikovsky
Georgia State University, Department of Computer Science
Atlanta, GA 30303, USA
E-mail: alexz@cs.gsu.edu

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-21259-8 e-ISBN 978-3-642-21260-4
DOI 10.1007/978-3-642-21260-4
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: Applied for

CR Subject Classification (1998): J.3, H.2.8, H.3-4, F.1, I.5

LNCS Sublibrary: SL 8 – Bioinformatics

© Springer-Verlag Berlin Heidelberg 2011
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The seventh edition of the International Symposium on Bioinformatics Research and Applications (ISBRA 2011) was held during May 27-29, 2011 at Central South University in Changsha, China. The symposium provides a forum for the exchange of ideas and results among researchers, developers, and practitioners working on all aspects of bioinformatics and computational biology and their applications.

The technical program of the symposium included 36 contributed papers, selected by the Program Committee from 92 full submissions received in response to the call for papers. Additionally, the symposium included poster sessions and featured invited keynote talks by four distinguished speakers: Bernard Moret from the Swiss Federal Institute of Technology in Lausanne, Switzerland, spoke on phylogenetic analysis of whole genomes, David Sankoff from the University of Ottawa spoke on competing formulations of orthologs for multiple genomes, Russell Schwartz from Carnegie Mellon University spoke on phylogenetics of heterogeneous samples, Liping Wei from Peking University spoke on the critical role of bioinformatics in genetic and pharmacogenetic studies of neuropsychiatric disorders, and Eric Xing from Carnegie Mellon University spoke on the a structured sparse regression approach to diseases.

We would like to thank the Program Committee members and external reviewers for volunteering their time to review and discuss symposium papers. We would like to extend special thanks to the Steering and General Chairs of the symposium for their leadership, and to the Finance, Publication, Publicity, and Local Organization Chairs for their hard work in making ISBRA 2011 a successful event. Last but not least we would like to thank all authors for presenting their work at the symposium.

May 2011

Jianer Chen
Jianxin Wang
Alex Zelikovsky

Symposium Organization

Steering Chairs

Dan Gusfield	University of California, Davis, USA
Yi Pan	Georgia State University, USA
Marie-France Sagot	INRIA, France

General Chairs

Mona Singh	Princeton University, USA
Lijian Tao	Central South University, China
Albert Y. Zomaya	The University of Sydney, Australia

Program Chairs

Jianer Chen	Texas A&M University, USA and Central South University, China
Jianxin Wang	Central South University, China
Alex Zelikovsky	Georgia State University, USA

Publicity Chairs

Ion Mandoiu	University of Connecticut, USA
Yanqing Zhang	Georgia State University, USA

Publication Chair

Raj Sunderraman	Georgia State University, USA

Finance Chair

Yu Sheng	Central South University, China

Local Organization Chairs

Min Wu	Central South University, China
Zhiming Yu	Central South University, China

Local Organizing Committee

Jianxin Wang	Central South University, China
Guojun Wang	Central South University, China
Dongjun Huang	Central South University, China
Fei Li	Central South University, China
Min Li	Central South University, China
Gang Chen	Central South University, China

Program Committee

Srinivas Aluru	Iowa State University, USA
Danny Barash	Ben-Gurion University, Israel
Robert Beiko	Dalhousie University, Canada
Anne Bergeron	Université du Québec à Montréal, Canada
Daniel Berrar	University of Ulster, UK
Paola Bonizzoni	Universitá de Studi di Milano-Bicocca, Italy
Daniel Brown	University of Waterloo, Canada
Tien-Hao Chang	National Cheng Kung University, Taiwan
Chien-Yu Chen	National Taiwan University, Taiwan
Jianer Chen	Texas A&M University, USA
Bhaskar Dasgupta	University of Illinois at Chicago, USA
Amitava Datta	University of Western Australia, Australia
Guillaume Fertin	Université de Nantes, France
Vladimir Filkov	University of California Davis, USA
Jean Gao	University of Texas at Arlington, USA
Katia Guimarães	Universidade Federal de Pernambuco, Brazil
Jieyue He	Southeast University, China
Jinling Huang	East Carolina University, USA
Lars Kaderali	University of Heidelberg, Germany
Iyad Kanj	DePaul University, USA
Ming-Yang Kao	Northwestern University, USA
Yury Khudyakov	Centers for Disease Control and Prevention, USA
Danny Krizanc	Wesleyan University, USA
Jing Li	Case Western Reserve University, USA
Zhiyong Liu	Chinese Academy of Science, China
Ion Mandoiu	University of Connecticut, USA
Fenglou Mao	University of Georgia, USA
Osamu Maruyama	Kyushu University, Japan
Ion Moraru	University of Connecticut Health Center, USA
Craig Nelson	University of Connecticut, USA
Andrei Paun	Louisiana Tech University, USA
Maria Poptsova	University of Connecticut, USA
Sven Rahmann	Technical University of Dortmund, Germany

Shoba Ranganathan	Macquarie University, Australia
Isidore Rigoutsos	IBM Research, USA
Cenk Sahinalp	Simon Fraser, Canada
Russell Schwartz	Carnegie Mellon University, USA
João Carlos Setubal	Virginia Polytechnic Institute and State University, USA
Jens Stoye	Universität Bielefeld, Germany
Raj Sunderraman	Georgia State University, USA
Wing-Kin Sung	National University of Singapore, Singapore
Sing-Hoi Sze	Texas A&M University, USA
Haixu Tang	Indiana University, USA
Gabriel Valiente	Technical University of Catalonia, Spain
Stéphane Vialette	Université Paris-Est Marne-la-Vallée, France
Jianxin Wang	Central South University, China
Li-San Wang	University of Pennsylvania, USA
Lusheng Wang	City University of Hong Kong, China
Carsten Wiuf	University of Aarhus, Denmark
Richard Wong	Kanazawa University, Japan
Yufeng Wu	University of Connecticut, USA
Alex Zelikovsky	Georgia State University, USA
Fa Zhang	Chinese Academy of Science, China
Yanqing Zhang	Georgia State University, USA
Leming Zhou	University of Pittsburgh, USA

External Reviewers

Aggarwala, Varun
Al Seesi, Sahar
Amberkar, Sandeep
Andres, Stephan Dominique
Astrovskaya, Irina
Berman, Piotr
Bernauer, Julie
Blin, Guillaume
Cao, Kajia
Cliquet, Freddy
Cohen-Boulakia, Sarah
Della Vedova, Gianluca
Dondi, Riccardo
Dörr, Daniel
Evans, Patricia
Fang, Ming
Gherman, Marius
Grudinin, Sergei
Hwang, Yih-Chii
Jahn, Katharina
Jiang, Minghui
Jiao, Dazhi
Kang, Min Gon
Kiani, Narsis Aftab
Kim, Dong-Chul
Kim, Wooyoung
Kopczynski, Dominik
Köster, Johannes
Lara, James
Lee, Byoungkoo
Leung, Fanny
Li, Fan
Li, Shuo
Li, Weiling
Lin, Chiao-Feng
Macdonald, Norman
Malhotra, Samta
Mao, Xizeng

Marschall, Tobias
Martin, Marcel
Mayampurath, Anoop
Mi, Qi
Missirian, Victor
Nakhleh, Luay
Nguyen, Minh Q.
Olman, Victor
Rizzi, Raffaella
Rizzi, Romeo
Skums, Pavel
Srichandan, Bismita
Tan, Cheemeng
Tang, Xiaojia
Tannier, Eric
Wang, Wenhui
Willing, Eyla
Wittler, Roland
Wu, Yin
Yang, Xiao
Yu, Chuan-Yih
Zhang, Chi
Zhang, Xinjun
Zhang, Yanqing
Zhou, Fengfeng
Zola, Jaroslaw

Table of Contents

Phylogenetics of Heterogeneous Samples 1
Russell Schwartz

OMG! Orthologs for Multiple Genomes - Competing Formulations 2
David Sankoff

Phylogenetic Analysis of Whole Genomes 4
Bernard M.E. Moret

Genetic and Pharmacogenetic Studies of Neuropsychiatric Disorders: Increasingly Critical Roles of Bioinformatics Research and Applications 8
Liping Wei

Genome-Phenome Association Analysis of Complex Diseases a Structured Sparse Regression Approach 11
Eric Xing

Prediction of Essential Proteins by Integration of PPI Network Topology and Protein Complexes Information 12
Jun Ren, Jianxin Wang, Min Li, Huan Wang, and Binbin Liu

Computing the Protein Binding Sites 25
Fei Guo and Lusheng Wang

SETTER - RNA SEcondary sTructure-based TERtiary Structure Similarity Algorithm 37
David Hoksza and Daniel Svozil

Prediction of Essential Genes by Mining Gene Ontology Semantics 49
Yu-Cheng Liu, Po-I Chiu, Hsuan-Cheng Huang, and Vincent S. Tseng

High-Performance Blob-Based Iterative Reconstruction of Electron Tomography on Multi-GPUs 61
Xiaohua Wan, Fa Zhang, Qi Chu, and Zhiyong Liu

Component-Based Matching for Multiple Interacting RNA Sequences . . . 73
Ghada Badr and Marcel Turcotte

A New Method for Identifying Essential Proteins Based on Edge Clustering Coefficient 87
Huan Wang, Min Li, Jianxin Wang, and Yi Pan

Gene Order in Rosid Phylogeny, Inferred from Pairwise Syntenies among Extant Genomes 99
Chunfang Zheng and David Sankoff

Algorithms to Detect Multiprotein Modularity Conserved during Evolution 111
Luqman Hodgkinson and Richard M. Karp

The Kernel of Maximum Agreement Subtrees 123
Krister M. Swenson, Eric Chen, Nicholas D. Pattengale, and David Sankoff

A Consensus Approach to Predicting Protein Contact Map via Logistic Regression 136
Jian-Yi Yang and Xin Chen

A Linear Time Algorithm for Error-Corrected Reconciliation of Unrooted Gene Trees 148
Paweł Górecki and Oliver Eulenstein

Comprehensive Pharmacogenomic Pathway Screening by Data Assimilation 160
Takanori Hasegawa, Rui Yamaguchi, Masao Nagasaki, Seiya Imoto, and Satoru Miyano

The Deep Coalescence Consensus Tree Problem is Pareto on Clusters . . . 172
Harris T. Lin, J. Gordon Burleigh, and Oliver Eulenstein

Fast Local Search for Unrooted Robinson-Foulds Supertrees 184
Ruchi Chaudhary, J. Gordon Burleigh, and David Fernández-Baca

A Metric for Phylogenetic Trees Based on Matching 197
Yu Lin, Vaibhav Rajan, and Bernard M.E. Moret

Describing the Orthology Signal in a PPI Network at a Functional, Complex Level 209
Pavol Jancura, Eleftheria Mavridou, Beatriz Pontes, and Elena Marchiori

Algorithms for Rapid Error Correction for the Gene Duplication Problem 227
Ruchi Chaudhary, J. Gordon Burleigh, and Oliver Eulenstein

TransDomain: A Transitive Domain-Based Method in Protein–Protein Interaction Prediction 240
Yi-Tsung Tang and Hung-Yu Kao

Rapid and Accurate Generation of Peptide Sequence Tags with a Graph Search Approach 253
Hui Li, Lauren Scott, Chunmei Liu, Mugizi Rwebangira, Legand Burge, and William Southerland

In Silico Evolution of Multi-scale Microbial Systems in the Presence of Mobile Genetic Elements and Horizontal Gene Transfer 262
Vadim Mozhayskiy and Ilias Tagkopoulos

Comparative Evaluation of Set-Level Techniques in Microarray Classification 274
Jiri Klema, Matej Holec, Filip Zelezny, and Jakub Tolar

Gene Network Modules-Based Liner Discriminant Analysis of Microarray Gene Expression Data 286
Pingzhao Hu, Shelley Bull, and Hui Jiang

A Polynomial Algebra Method for Computing Exemplar Breakpoint Distance 297
Bin Fu and Louxin Zhang

The Maximum Clique Enumeration Problem: Algorithms, Applications and Implementations 306
John D. Eblen, Charles A. Phillips, Gary L. Rogers, and Michael A. Langston

Query-Adaptive Ranking with Support Vector Machines for Protein Homology Prediction 320
Yan Fu, Rong Pan, Qiang Yang, and Wen Gao

A Novel Core-Attachment Based Greedy Search Method for Mining Functional Modules in Protein Interaction Networks 332
Chaojun Li, Jieyue He, Baoliu Ye, and Wei Zhong

ProPhyC: A Probabilistic Phylogenetic Model for Refining Regulatory Networks 344
Xiuwei Zhang and Bernard M.E. Moret

Prediction of DNA-Binding Propensity of Proteins by the Ball-Histogram Method 358
Andrea Szabóová, Ondřej Kuželka, Sergio Morales E., Filip Železný, and Jakub Tolar

Multi-label Correlated Semi-supervised Learning for Protein Function Prediction 368
Jonathan Q. Jiang

Regene: Automatic Construction of a Multiple Component Dirichlet Mixture Priors Covariance Model to Identify Non-coding RNA 380
Felipe Lessa, Daniele Martins Neto, Kátia Guimarães, Marcelo Brigido, and Maria Emilia Walter

Accurate Estimation of Gene Expression Levels from DGE Sequencing Data 392
Marius Nicolae and Ion Măndoiu

An Integrative Approach for Genomic Island Prediction in Prokaryotic Genomes 404
Han Wang, John Fazekas, Matthew Booth, Qi Liu, and Dongsheng Che

A Systematic Comparison of Genome Scale Clustering Algorithms 416
Jeremy J. Jay, John D. Eblen, Yun Zhang, Mikael Benson, Andy D. Perkins, Arnold M. Saxton, Brynn H. Voy, Elissa J. Chesler, and Michael A. Langston

Mining Biological Interaction Networks Using Weighted Quasi-Bicliques 428
Wen-Chieh Chang, Sudheer Vakati, Roland Krause, and Oliver Eulenstein

Towards a Characterisation of the Generalised Cladistic Character Compatibility Problem for Non-branching Character Trees 440
Ján Maňuch, Murray Patterson, and Arvind Gupta

Author Index 453