

Lecture Notes in Bioinformatics

7292

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

Leonidas Bleris Ion Măndoiu
Russell Schwartz Jianxin Wang (Eds.)

Bioinformatics Research and Applications

8th International Symposium, ISBRA 2012
Dallas, TX, USA, May 21-23, 2012
Proceedings

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

Leonidas Bleris

University of Texas at Dallas

Richardson, TX 75080, USA

E-mail: bleris@utdallas.edu

Ion Măndoiu

University of Connecticut

Storrs, CT 06269, USA

E-mail: ion@engr.uconn.edu

Russell Schwartz

Carnegie Mellon University

Pittsburgh, PA 15213, USA

E-mail: russells@andrew.cmu.edu

Jianxin Wang

Central South University

Changsha 410083, China

E-mail: jxwang@mail.csu.edu.cn

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-642-30190-2

e-ISBN 978-3-642-30191-9

DOI 10.1007/978-3-642-30191-9

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012937075

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

LNCS Sublibrary: SL 8 – Bioinformatics

© Springer-Verlag Berlin Heidelberg 2012

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 8th International Symposium on Bioinformatics Research and Applications (ISBRA 2012) was held during May 21–23, 2012, in Dallas, Texas. 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 26 extended abstracts, selected by the Program Committee from a number of 66 full submissions received in response to the call for papers. Additionally, the symposium included poster sessions and featured invited keynote talks by five distinguished speakers: Ambuj Singh from the University of California at Santa Barbara spoke on learning discriminative graph fragments, Bhaskar Dasgupta from the University of Illinois at Chicago spoke on models and algorithmic tools for computational processes in cellular biology, Cynthia Gibas from the University of North Carolina at Charlotte spoke on analytics approaches for the era of 10,000 genomes, Dong Xu from the University of Missouri spoke on protein structure prediction and clustering, and Michael Zhang from the University of Texas at Dallas and Tsinghua University spoke on computational modeling of mammalian promoters.

We would like to thank the Program Committee members and external reviewers for volunteering their time to review ISBRA submissions. We would like to extend a special thanks to the Steering and General Chairs of the symposium for their leadership, and the ISBRA 2012 Publicity Chair, webmaster, and local organizers for their hard work in making ISBRA 2012 a successful event. Last but not least we would like to thank all authors for presenting their work at the symposium.

May 2012

Leonidas Bleris
Ion Măndoiu
Russell Schwartz
Jianxin Wang

Organization

Steering Chairs

Dan Gusfield	University of California Davis, USA
Ion Măndoiu	University of Connecticut, USA
Yi Pan	Georgia State University, USA
Marie-France Sagot	INRIA, France
Alexander Zelikovsky	Georgia State University, USA

General Chairs

Ovidiu Daescu	University of Texas at Dallas, USA
Raj Sunderraman	Georgia State University, USA

Program Chairs

Leonidas Bleris	University of Texas at Dallas, USA
Ion Măndoiu	University of Connecticut, USA
Russell Schwartz	Carnegie Mellon University, USA
Jianxin Wang	Central South University, China

Publicity Chair

Sahar Al Seesi	University of Connecticut, USA
----------------	--------------------------------

Webmaster

J. Steven Kirtzic	University of Texas at Dallas, USA
-------------------	------------------------------------

Program Committee

Srinivas Aluru	IIT Bombay, India and 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	Tokyo Institute of Technology, USA
Paola Bonizzoni	Università di Studi di Milano-Bicocca, Italy
Daniel Brown	University of Waterloo, Canada
Doina Caragea	Kansas State University, USA

Tien-Hao Chang	National Cheng Kung University, Taiwan
Chien-Yu Chen	National Taiwan University, Taiwan
Matteo Comin	Università degli Studi di Padova, Italy
Ovidiu Daescu	University of Texas at Dallas, USA
Bhaskar Dasgupta	University of Illinois at Chicago, USA
Douglas Densmore	Boston University, USA
Jorge Duitama	International Center for Tropical Agriculture, Colombia
Oliver Eulenstein	Iowa State University, USA
Guillaume Fertin	University of Nantes, France
Vladimir Filkov	University of California Davis, USA
Jean Gao	University of Texas at Arlington, USA
Katia Guimaraes	Universidade Federal de Pernambuco, Brazil
Jiong Guo	Universität des Saarlandes, Germany
Robert Harrison	Georgia State University, USA
Jieyue He	Southeast University, China
Steffen Heber	North Carolina State University, USA
Allen Holder	Rose-Hulman Institute of Technology, USA
Jinling Huang	East Carolina University, USA
Lars Kaderali	University of Technology Dresden, 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
Yanchun Liang	Jilin University, China
Zhiyong Liu	Chinese Academy of Science, China
Fenglou Mao	University of Georgia, USA
Osamu Maruyama	Kyushu University, Japan
Li Min	Central South University, China
Ion Moraru	University of Connecticut Health Center, USA
Axel Mosig	CAS-MPG Partner Institute and Key Laboratory for Computational Biology, China and Ruhr University Bochum, Germany
Giri Narasimhan	Florida International University, USA
Yi Pan	Georgia State University, USA
Laxmi Parida	IBM T.J. Watson Research Center, USA
Bogdan Pasaniuc	Harvard School of Public Health, USA
Andrei Paun	Louisiana Tech University, USA
Itsik Pe'Er	Columbia University, USA
Weiqun Peng	George Washington University, USA
Nadia Pisanti	Università di Pisa, Italy

Maria Poptsova	University of Connecticut, USA
Teresa Przytycka	NCBI, USA
Sven Rahmann	University of Duisburg-Essen, Germany
Shoba Ranganathan	Macquarie University, Australia
S. Cenk Sahinalp	Simon Fraser University, Canada
David Sankoff	University of Ottawa, Canada
Daniel Schwartz	University of Connecticut, USA
Joao Setubal	Virginia Bioinformatics Institute, USA
Mona Singh	Princeton University, USA
Ileana Streinu	Smith College, USA
Raj Sunderraman	Georgia State University, USA
Wing-Kin Sung	National University of Singapore, Singapore
Sing-Hoi Sze	Texas A&M University, USA
Ilias Tagkopoulos	University of California Davis, USA
Marcel Turcotte	University of Ottawa, Canada
Gabriel Valiente	Technical University of Catalonia, Spain
Stéphane Vialette	Université Paris-Est Marne-la-Vallée, France
Panagiotis Vouzis	Carnegie Mellon University, USA
Jianxin Wang	Central South University, China
Li-San Wang	University of Pennsylvania, USA
Lusheng Wang	City University of Hong Kong, Hong Kong
Xiaowo Wang	Tsinghua University, China
Fangxiang Wu	University of Saskatchewan, Canada
Yufeng Wu	University of Connecticut, USA
Zhen Xie	Massachusetts Institute of Technology, USA
Jinbo Xu	Toyota Technical Institute at Chicago, USA
Zhenyu Xuan	University of Texas at Dallas, USA
Zuguo Yu	Xiangtan University, China and Queensland University of Technology, Australia
Alex Zelikovskiy	Georgia State University, USA
Fa Zhang	Institute of Computing Technology, China
Yanqing Zhang	Georgia State University, USA
Leming Zhou	University of Pittsburgh, USA

Additional Reviewers

Bacciu, Davide	Della Vedova, Gianluca
Badr, Ghada	Dondi, Riccardo
Belcaid, Mahdi	Du, Wei
Bernauer, Julie	Fang, Ming
Biswas, Ashis Kumer	Fox, Naomi
Bulteau, Laurent	Guan, Renchu
Caldas, José	Guillemot, Sylvain
Campo, David S.	Haiminen, Niina
Cliquet, Freddy	Hayes, Matthew

Huang, Fangping	Pyon, Yoon Soo
Huang, Yang	Rizzi, Raffaella
Hwang, Yih-Chii	Ryvkin, Paul
Kang, Mingon	Scornavacca, Celine
Kim, Yoo-Ah	Skums, Pavel
Kopczynski, Dominik	Srichandan, Bismita
Köster, Johannes	Thorne, Jeffrey
Labarre, Anthony	Tsirigos, Aristotelis
Lacroix, Vincent	Utro, Filippo
Li, Fan	Verzotto, Davide
Li, Shuo	Wang, Juexin
Liu, Bingqiang	Wang, Lin
Mendes, Nuno	Wang, Yan
Messina, Enza	Wojtowicz, Damian
Mina, Marco	Yang, Sen
Missirian, Victor	Zhang, Jin
Mozhayskiy, Vadim	Zhou, Chan
Nefedov, Alexey	Zola, Jaroslaw

Table of Contents

Trie-based Apriori Motif Discovery Approach	1
<i>Isra Al-Turaiki, Ghada Badr, and Hassan Mathkour</i>	
Inapproximability of (1, 2)-Exemplar Distance	13
<i>Laurent Bulteau and Minghui Jiang</i>	
A Mixed Integer Programming Model for the Parsimonious Loss of Heterozygosity Problem	24
<i>Daniele Catanzaro, Martine Labbé, and Bjarni V. Halldórsson</i>	
Reconstruction of Transcription Regulatory Networks by Stability-Based Network Component Analysis	36
<i>Xi Chen, Chen Wang, Ayesha N. Shajahan, Rebecca B. Riggins, Robert Clarke, and Jianhua Xuan</i>	
A Transcript Perspective on Evolution	48
<i>Yann Christinat and Bernard M.E. Moret</i>	
A Fast Algorithm for Computing the Quartet Distance for Large Sets of Evolutionary Trees	60
<i>Ralph W. Crosby and Tiffani L. Williams</i>	
Error Propagation in Sparse Linear Systems with Peptide-Protein Incidence Matrices	72
<i>Peter Damaschke and Leonid Molokov</i>	
Models and Algorithmic Tools for Computational Processes in Cellular Biology: Recent Developments and Future Directions (Invited Keynote Talk)	84
<i>Bhaskar DasGupta</i>	
Identifying Rogue Taxa through Reduced Consensus: NP-Hardness and Exact Algorithms	87
<i>Akshay Deepak, Jianrong Dong, and David Fernández-Baca</i>	
Analytics Approaches for the Era of 10,000 Genomes (Invited Keynote Talk)	99
<i>Cynthia J. Gibas</i>	
GTP Supertrees from Unrooted Gene Trees: Linear Time Algorithms for NNI Based Local Searches	102
<i>Paweł Górecki, J. Gordon Burleigh, and Oliver Eulenstein</i>	

A Robinson-Foulds Measure to Compare Unrooted Trees with Rooted Trees	115
<i>Paweł Górecki and Oliver Eulenstein</i>	
P-Binder: A System for the Protein-Protein Binding Sites Identification	127
<i>Fei Guo, Shuai Cheng Li, and Lusheng Wang</i>	
Non-identifiable Pedigrees and a Bayesian Solution	139
<i>Bonnie Kirkpatrick</i>	
Iterative Piecewise Linear Regression to Accurately Assess Statistical Significance in Batch Confounded Differential Expression Analysis	153
<i>Juntao Li, Kwok Pui Choi, and R. Krishna Murthy Karuturi</i>	
Reconstruction of Network Evolutionary History from Extant Network Topology and Duplication History	165
<i>Si Li, Kwok Pui Choi, Taoyang Wu, and Louxin Zhang</i>	
POPE: Pipeline of Parentally-Biased Expression	177
<i>Victor Missirian, Isabelle Henry, Luca Comai, and Vladimir Filkov</i>	
On Optimizing the Non-metric Similarity Search in Tandem Mass Spectra by Clustering	189
<i>Jiří Novák, David Hoksza, Jakub Lokoč, and Tomáš Skopal</i>	
On the Comparison of Sets of Alternative Transcripts	201
<i>Aïda Ouangraoua, Krister M. Swenson, and Anne Bergeron</i>	
MURPAR: A Fast Heuristic for Inferring Parsimonious Phylogenetic Networks from Multiple Gene Trees	213
<i>Hyun Jung Park and Luay Nakhleh</i>	
Large Scale Ranking and Repositioning of Drugs with Respect to DrugBank Therapeutic Categories	225
<i>Matteo Re and Giorgio Valentini</i>	
Score Based Aggregation of microRNA Target Orderings	237
<i>Debarka Sengupta, Ujjwal Maulik, and Sanghamitra Bandyopadhyay</i>	
Modeling Complex Diseases Using Discriminative Network Fragments (Invited Keynote Talk)	249
<i>Ambuj K. Singh</i>	
Novel Multi-sample Scheme for Inferring Phylogenetic Markers from Whole Genome Tumor Profiles	250
<i>Ayshwarya Subramanian, Stanley Shackney, and Russell Schwartz</i>	
Algorithms for Knowledge-Enhanced Supertrees	263
<i>André Wehe, J. Gordon Burleigh, and Oliver Eulenstein</i>	

Improvement of BLASTp on the FPGA-Based High-Performance Computer RIVYERA.....	275
<i>Lars Wienbrandt, Daniel Siebert, and Manfred Schimmeler</i>	
A Polynomial Time Solution for Protein Chain Pair Simplification under the Discrete Fréchet Distance	287
<i>Tim Wylie and Binhai Zhu</i>	
Designing RNA Secondary Structures in Coding Regions	299
<i>Rukhsana Yeasmin and Steven Skiena</i>	
Phylogenetic Tree Reconstruction with Protein Linkage	315
<i>Junjie Yu, Henry Chi Ming Leung, Siu Ming Yiu, Yong Zhang, Francis Y.L. Chin, Nathan Hobbs, and Amy Y.X. Wang</i>	
Protein Structure Prediction and Clustering Using MUFOLD (Invited Keynote Talk)	328
<i>Jingfen Zhang and Dong Xu</i>	
Computational Modeling of Mammalian Promoters (Invited Keynote Talk)	330
<i>Michael Q. Zhang</i>	
Author Index	331