

# Lecture Notes in Bioinformatics

5676

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

Katia S. Guimarães Anna Panchenko  
Teresa M. Przytycka (Eds.)

# Advances in Bioinformatics and Computational Biology

4th Brazilian Symposium on Bioinformatics, BSB 2009  
Porto Alegre, Brazil, July 29-31, 2009  
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

Katia S. Guimarães

Federal University of Pernambuco, Center of Informatics

Av. Prof. Luiz Freire, s/n, Cidade Universitária, Recife, PE 50740-540, Brazil

E-mail: katiag@cin.ufpe.br

Anna Panchenko

Teresa M. Przytycka

National Center for Biotechnology Information

National Library of Medicine, National Institutes of Health

8600 Rockville Pike, Building 38A, Bethesda, MD 20894, USA

E-mail: {panch, przytyck}@ncbi.nlm.nih.gov

Library of Congress Control Number: 2009930955

CR Subject Classification (1998): H.2.8, F.2.1, I.2, G.2.2, J.3, E.1

LNCS Sublibrary: SL 8 – Bioinformatics

ISSN 0302-9743

ISBN-10 3-642-03222-2 Springer Berlin Heidelberg New York

ISBN-13 978-3-642-03222-6 Springer Berlin Heidelberg New York

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.

springer.com

© Springer-Verlag Berlin Heidelberg 2009

Printed in Germany

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

Printed on acid-free paper SPIN: 12724642 06/3180 5 4 3 2 1 0

## Preface

This volume contains the papers selected for presentation at the 4th Brazilian Symposium on Bioinformatics, BSB 2009, which was held in Porto Alegre, Brazil, during August 29–31, 2009. The BSB symposium had its origins in the Brazilian Workshop on Bioinformatics (WOB). WOB had three editions, in 2002 (Gramado, RS), in 2003 (Macaé, RJ), and in 2004 (Brasília, DF). The change in the designation from workshop to symposium reflects the increase in the quality of the contributions and also in the interest of the scientific community for the meeting. The previous editions of BSB took place in São Leopoldo, RS, in 2005, in Angra dos Reis, RJ, in 2007, and in Santo André, SP, in 2008.

As evidence of the internationalization of the event, BSB 2009 had 55 submissions from seven countries. Of the 55 papers submitted, 36 were full papers, with up to 12 pages each, and 19 were extended abstracts, with up to 4 pages each. The articles submitted were carefully reviewed and selected by an international Program Committee, comprising three chairs and 45 members from around the world, with the help of 21 additional reviewers. The Program Committee Chairs are very thankful to the authors of all submitted papers, and especially to the Program Committee members and the additional reviewers, who helped select the 12 full papers and the six extended abstracts that make up this book.

The editors would also like to thank the six keynote speakers of BSB 2009: Emil Alexov (Clemson University, USA), Dario Grattapaglia (EMBRAPA, Brazil), Arthur Gruber (USP, Brazil), Anna Panchenko (NIH, USA), Teresa M. Przytycka (NIH, USA), and Eytan Ruppin (Tel Aviv University, Israel).

The conference was sponsored by the Brazilian Computer Society (SBC), and counted on the support of Brazilian sponsoring agencies CNPq and CAPES. Special thanks go to Springer, for agreeing to publish this proceedings volume, and to the Steering Committee of BSB 2009. Last but not least, we would like to express our gratitude to the General and Local Organizing Chairs, Osmar Norberto de Souza and Duncan D.D. Ruiz, and the Local Organizing Committee, whose excellent work and dedication made BSB 2009 an enjoyable experience.

July 2009

Katia S. Guimarães  
Anna Panchenko  
Teresa M. Przytycka

# Organization

## Program Committee Chairs

Katia S. Guimarães	Federal University of Pernambuco (UFPE), Brazil
Anna Panchenko	National Institutes of Health (NIH), USA
Teresa M. Przytycka	National Institutes of Health (NIH), USA

## Program Committee

Emil Alexov	Clemson University, USA
Naivo F. Almeida Jr.	Federal University of Mato Grosso do Sul (UFMS), Brazil
Ana Lucia C. Bazzan	Federal University of Rio Grande do Sul (UFRGS), Brazil
Ana Maria Benko-Iseppon	Federal University of Pernambuco (UFPE), Brazil
Igor Berezovsky	University of Bergen, Norway
Ulisses Braga-Neto	Texas A&M University, USA
Marcelo M. Brigido	National University of Brasília (UNB), Brazil
Helaine Carrer	University of São Paulo (USP) /Piracicaba, Brazil
Andre C.P.L.F. Carvalho	University of São Paulo (USP) /São Carlos, Brazil
Saikat Chakrabarti	National Institutes of Health (NIH), USA
Alberto M.R. Dávila	Oswaldo Cruz Institute, Brazil
Frank Dehne	Carleton University, Canada
Zanoni Dias	State University of Campinas (UNICAMP), Brazil
Carlos E. Ferreira	University of São Paulo (USP) /São Paulo, Brazil
Jessica Fong	National Institutes of Health (NIH), USA
Oxana Galzitskaya	Institute of Protein Science, Russia
Ronaldo F. Hashimoto	University of São Paulo (USP) /São Paulo, Brazil
Maricel Kann	University of Maryland, Baltimore County (UMBC), USA
Carl Kingsford	University of Maryland, College Park (UMD), USA
Eugene Krissinel	European Bioinformatics Institute (EBI), UK
Ney Lemke	São Paulo State University (UNESP), Brazil
Thomas Madej	National Institutes of Health (NIH), USA
Wojtek Makalowski	University of Münster, Germany
Ion Mandoiu	University of Connecticut, USA
Natalia F. Martins	EMBRAPA, Brazil
Wellington S. Martins	Federal University of Goiás (UFG), Brazil
Jose Carlos M. Mombach	Federal University of Santa Maria (UFSM), Brazil
Osmar Norberto-de-Souza	Pontifical Catholic University of RS (PUCRS), Brazil
Arlindo Oliveira	INESC, Portugal
Carlos A. B. Pereira	University of São Paulo (USP) /São Paulo, Brazil

Ron Y. Pinter	Technion, Israel
Carlos H.I. Ramos	UNICAMP, Brazil
Cenk Sahinalp	Simon Fraser University, Canada
Francisco M. Salzano	Federal University of Rio Grande do Sul (UFRGS), Brazil
João Carlos Setubal	Virginia Bioinformatics Institute (VBI), USA
Roded Sharan	Tel Aviv University, Israel
Hagit Shatkay	Queen's University, Canada
Benjamin Shoemaker	National Institutes of Health (NIH), USA
Mona Singh	Princeton University, USA
Marcílio M.C.P. de Souto	Federal University of Rio Grande do Norte (UFRN), Brazil
Jerzy Tiuryn	University of Warsaw, Poland
Ana Teresa Vasconcelos	National Lab. of Scientific Computation (LNCC), Brazil
Maria Emília M.T. Walter	National University of Brasília (UNB), Brazil
Alex Zelikovsky	Georgia State University, USA
Jie Zheng	National Institutes of Health (NIH), USA

### **Additional Reviewers**

Gulsah Altun	Scripps Res. Institute, USA
Irina Astrovskaya	Georgia State University, USA
Ananth Bommakanti	University of Maryland, Baltimore County (UMBC), USA
Janusz Dutkowski	University of Warsaw, Poland
Katti Faceli	Federal University of São Carlos (UFSCar), Brazil
André Fujita	University of Tokyo, Japan
Anna Gambin	University of Warsaw, Poland
Yang Huang	National Institutes of Health (NIH), USA
Andre Kashiwabara	University of São Paulo (USP) /São Paulo, Brazil
Attila Kertesz-Farkas	University of Maryland, Baltimore County (UMBC), USA
Boguslaw Kluge	University of Warsaw, Poland
Giovani R. Librelotto	Federal University of Santa Maria (UFSM), Brazil
Fabrcio Lopes	Federal University of Technology, Paraná (UTFPR), Brazil
Carlos Maciel	University of São Paulo (USP) /São Paulo, Brazil
Silvio B. de Melo	Federal University of Pernambuco (UFPE), Brazil
Dilvan de A. Moreira	University of São Paulo (USP) /São Paulo, Brazil
Carlos Santos	University of São Paulo (USP) /São Paulo, Brazil
José Soares	University of São Paulo (USP) /São Paulo, Brazil
Rajesh Thangudu	National Institutes of Health (NIH), USA
Manoj Tyagi	National Institutes of Health (NIH), USA
Nir Yosef	Tel Aviv University, Israel

## General Chairs and Local Organizing Chairs

Osmar Norberto de Souza	Pontifical Catholic University of RS (PUCRS)
Duncan D.D. Ruiz	Pontifical Catholic University of RS (PUCRS)

## Local Organizing Committee

Ana Winck	Pontifical Catholic University of RS (PUCRS)
Carla Aguiar	Pontifical Catholic University of RS (PUCRS)
Christian Quevedo	Pontifical Catholic University of RS (PUCRS)
Danieli Forgiarini	Pontifical Catholic University of RS (PUCRS)
Elisa Cohen	Pontifical Catholic University of RS (PUCRS)
Karina Machado	Pontifical Catholic University of RS (PUCRS)

## Steering Committee

Maria Emília M.T. Walter (Chair)	National University of Brasília (UNB)
Ana L.C. Bazzan	Federal University of Rio Grande do Sul (UFRGS)
André C.P.L.F. de Carvalho	University of São Paulo (USP) /São Carlos
Katia S. Guimarães	Federal University of Pernambuco (UFPE)
Osmar Norberto-de-Souza	Pontifical Catholic University of RS (PUCRS)
Francisco M. Salzano	Federal University of Rio Grande do Sul (UFRGS)
João Carlos Setubal	Virginia Bioinformatics Institute (VBI, USA)

# Table of Contents

## Algorithmic Approaches for Molecular Biology Problems

Genotype Tagging with Limited Overfitting . . . . .	1
<i>Irina Astrovskaya and Alex Zelikovskiy</i>	
Constraint Programming Models for Transposition Distance Problem . . .	13
<i>Ulisses Dias and Zanoni Dias</i>	
Comparison of Spectra in Unsequenced Species . . . . .	24
<i>Freddy Cliquet, Guillaume Fertin, Irena Rusu, and Dominique Tessier</i>	

## Micro-array Analysis

BiHEA: A Hybrid Evolutionary Approach for Microarray Biclustering . . . . .	36
<i>Cristian Andrés Gallo, Jessica Andrea Carballido, and Ignacio Ponzoni</i>	
Using Supervised Complexity Measures in the Analysis of Cancer Gene Expression Data Sets . . . . .	48
<i>Ivan G. Costa, Ana C. Lorena, Lician R.M.P. y Peres, and Marcilio C.P. de Souto</i>	
Quantitative Improvements in cDNA Microarray Spot Segmentation . . . .	60
<i>Mónica G. Larese and Juan Carlos Gómez</i>	

## Machine Learning Methods for Classification

SOM-PORTRAIT: Identifying Non-coding RNAs Using Self-Organizing Maps . . . . .	73
<i>Tulio C. Silva, Pedro A. Berger, Roberto T. Arrial, Roberto C. Togawa, Marcelo M. Brigido, and Maria Emilia M.T. Walter</i>	
Automatic Classification of Enzyme Family in Protein Annotation . . . . .	86
<i>Cássia T. dos Santos, Ana L.C. Bazzan, and Ney Lemke</i>	
Representations for Evolutionary Algorithms Applied to Protein Structure Prediction Problem Using HP Model . . . . .	97
<i>Paulo H.R. Gabriel and Alexandre C.B. Delbem</i>	



Comparing Methods for Multilabel Classification of Proteins Using Machine Learning Techniques . . . . . 109  
*Ricardo Cerri, Renato R.O. da Silva, and André C.P.L.F. de Carvalho*

Comparative Study of Classification Algorithms Using Molecular Descriptors in Toxicological DataBases . . . . . 121  
*Max Pereira, Vítor Santos Costa, Rui Camacho, Nuno A. Fonseca, Carlos Simões, and Rui M.M. Brito*

**In Silico Simulation**

Influence of Antigenic Mutations in Time Evolution of the Immune Memory – A Dynamic Modeling . . . . . 133  
*Alexandre de Castro, Carlos Frederico Fronza, Poliana Fernanda Giachetto, and Domingos Alves*

**Short Papers**

FReDD: Supporting Mining Strategies through a Flexible-Receptor Docking Database . . . . . 143  
*Ana T. Winck, Karina S. Machado, Osmar Norberto-de-Souza, and Duncan D.D. Ruiz*

A Wide Antimicrobial Peptides Search Method Using Fuzzy Modeling . . . . . 147  
*Fabiano C. Fernandes, William F. Porto, and Octavio L. Franco*

Identification of Proteins from Tuberculin Purified Protein Derivative (PPD) with Potential for TB Diagnosis Using Bioinformatics Analysis . . . . . 151  
*Sibele Borsuk, Fabiana Kommling Seixas, Daniela Fernandes Ramos, Caroline Rizzi, and Odir Antonio Dellagostin*

Mapping HIV-1 Subtype C gp120 Epitopes Using a Bioinformatic Approach . . . . . 156  
*Dennis Maletich Junqueira, Rúbia Marília de Medeiros, Sabrina Esteves de Matos Almeida, Vanessa Rodrigues Paixão-Cortez, Paulo Michel Rohe, and Fernando Rosado Spilki*

MHC: Peptide Analysis: Implications on the Immunogenicity of Hantaviruses’ N protein . . . . . 160  
*Maurício Menegatti Rigo, Dinler Amaral Antunes, Gustavo Fioravanti Vieira, and José Artur Bogo Chies*

An Ontology to Integrate Transcriptomics and Interatomics Data  
Involved in Gene Pathways of Genome Stability..... 164  
*Giovani Rubert Librelotto, José Carlos Mombach,  
Marialva Sinigaglia, Éder Simão,  
Heleno Borges Cabral, and Mauro A.A. Castro*

**Author Index** ..... 169