

# Lecture Notes in Bioinformatics

5488

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

Francesco Masulli Roberto Tagliaferri  
Gennady M. Verkhivker (Eds.)

# Computational Intelligence Methods for Bioinformatics and Biostatistics

5th International Meeting, CIBB 2008  
Vietri sul Mare, Italy, October 3-4, 2008  
Revised Selected Papers



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

Francesco Masulli

Università di Genova

Dipartimento di Informatica e Scienze dell'Informazione (DISI)

Via Dodecaneso 35, 16146 Genova, Italy

E-mail: masulli@disi.unige.it

Roberto Tagliaferri

Università di Salerno

Dipartimento di Matematica ed Informatica (DMI)

Via Ponte don Melillo, 84084 Fisciano (SA), Italy

E-mail: rtagliaferri@unisa.it

Gennady M. Verkhivker

The University of Kansas

School of Pharmacy, Department of Pharmaceutical Chemistry

2095 Constant Avenue, Lawrence, KS 66047, USA

E-mail: verk@ku.edu

Library of Congress Control Number: 2009929551

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 3-642-02503-X Springer Berlin Heidelberg New York

ISBN-13 978-3-642-02503-7 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: 12697106 06/3180 5 4 3 2 1 0

# Preface

This volume contains a selection of the best contributions delivered at the 5th International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics (CIBB 2008) held at IIASS “E. R. Caianiello”, Vietri sul Mare, Salerno, Italy during October 3–4, 2008.

The CIBB meeting series is organized by the Special Interest Group on Bioinformatics of the International Neural Network Society (INNS) to provide a forum open to researchers from different disciplines to present and discuss problems concerning computational techniques in bioinformatics, systems biology and medical informatics with a particular focus on neural networks, machine learning, fuzzy logic, and evolutionary computational methods. Previous CIBB meetings were held with an increasing number of participants within the format of a special session of larger conferences, namely, WIRN 2004 in Perugia, WILF 2005 in Crema, FLINS 2006 in Genoa and WILF 2007 in Camogli. Given the great success of the special session at WILF 2007 that included 26 papers after a strong selection, the 2008 edition of CIBB was organized, at last, as an autonomous conference, governed by its own Steering Committee.

CIBB 2008 attracted 69 paper submissions from all over the world. A rigorous peer-review selection process was applied to ultimately select the papers included in the program of the conference. This volume collects the best contributions presented at the conference. Moreover, the volume also includes three presentations from keynote speakers.

The success of this conference is to be credited to the contribution of many people. In the first place, we would like to thank the organizers of the special sessions for their strong effort in attracting so many good papers that extend and deepen the main topics of CIBB. Moreover, special thanks are due to the Program Committee members and reviewers for their commitment to the task of providing high-quality reviews. Last, but not least, we would like to thank the keynote speakers Mario Lauria (Systems Biology Lab, TIGEM, Telethon Institute of Genetics and Medicine, Naples, Italy), Nicolas Le Novere (Computational Neurobiology, EMBL-EBI, Wellcome-Trust Genome Campus, Hinxton, UK), and Giorgio Valentini (DSI, Dipartimento di Scienze dell’Informazione, Università degli Studi di Milano, Milan, Italy).

While we are concluding this editorial effort, a special thought goes to Tonina Starita who pioneered the applications of computational intelligence to the biomedical field and left us on August 2008. Our tender thanks go to her.

October 2008

Francesco Masulli  
Roberto Tagliaferri  
Gennady M. Verkhivker

# Organization

The 5th CIBB meeting was a joint operation of the Special Interest Group on Bioinformatics of INNS with the International Society for Computational Biology (ISCB), and the collaboration of the Gruppo Nazionale Calcolo Scientifico (GNCS), the Italian Neural Networks Society (SIREN), the Department of Computer and Information Sciences of the University of Genova, Italy (DISI), and the Department of Mathematics and Computer Science of the University Salerno, Italy (DMI).

## Conference Chairs

|                       |                                                     |
|-----------------------|-----------------------------------------------------|
| Francesco Masulli     | University of Genoa, Italy & Temple University, USA |
| Roberto Tagliaferri   | University of Salerno, Italy                        |
| Gennady M. Verkhivker | University of Kansas, USA & UCSD, USA               |

## Program Committee

|                       |                                                     |
|-----------------------|-----------------------------------------------------|
| Klaus-Peter Adlassnig | Medical University of Vienna, Austria               |
| Joaquin Dopazo        | C.I. Príncipe Felipe, Valencia, Spain               |
| Antonio Giordano      | University of Siena, Italy & Temple University, USA |
| Emmanuel Ifeakor      | University of Plymouth, UK                          |
| Samuel Kaski          | Helsinki University of Technology, Finland          |
| Natalio Krasnogor     | University of Nottingham, UK                        |
| Giancarlo Mauri       | University of Milano Bicocca, Italy                 |
| Oleg Okun             | University of Oulu, Finland                         |
| Giulio Pavesi         | University of Milan, Italy                          |
| David Alejandro Pelta | University of Granada, Spain                        |
| Graziano Pesole       | University of Bari, Italy                           |
| Leif E. Peterson      | Baylor College of Medicine, Houston, USA            |
| Giancarlo Raiconi     | University of Salerno, Italy                        |
| Volker Roth           | ETH Zurich, Switzerland                             |
| Giuseppe Russo        | Temple University, Philadelphia, USA                |
| Giorgio Valentini     | University of Milan, Italy                          |
| Giorgio Valle         | University of Padua, Italy                          |
| Jean-Philippe Vert    | Ecole des Mines de Paris, France                    |

## CIBB Steering Committee

|                   |                                            |
|-------------------|--------------------------------------------|
| Pierre Baldi      | University of California, Irvine, USA      |
| Alexandru Floares | Oncological Institute Cluj-Napoca, Romania |

## VIII Organization

|                     |                                                         |
|---------------------|---------------------------------------------------------|
| Jon Garibaldi       | University of Nottingham, UK                            |
| Francesco Masulli   | University of Genova, Italy & Temple<br>University, USA |
| Roberto Tagliaferri | University of Salerno, Italy                            |

## Special Session Organizers

|                                      |                                                                            |
|--------------------------------------|----------------------------------------------------------------------------|
| J. DeLeo, A. Floares                 | <i>Intelligent Systems for Medical Decisions<br/>Support</i>               |
| L. Milanese, R. Rizzo, A. Urso       | <i>Computational Intelligence for Biological<br/>Data Visualization</i>    |
| R. Battiti, M. Brunato, A. Passerini | <i>Machine Learning and Intelligent<br/>Optimization in Bioinformatics</i> |

## Referees

(in addition to previous committees)

|                |                   |                |
|----------------|-------------------|----------------|
| A. Avogadri    | V. Fabry-Asztalos | F. Pappalardo  |
| J. Bacardit    | B. Fischer        | A. Passerini   |
| M. Biba        | H. Fröhlich       | G. Pavesi      |
| F. Camastra    | S. Kaski          | D. Pelta       |
| R. Casadio     | D. Liberati       | C. Pizzuti     |
| M. Cannataro   | A. Maratea        | V. Plagianakos |
| M. Ceccarelli  | P.L. Martelli     | S. Rovetta     |
| J. Chen        | A. Micheli        | D. Saccà       |
| L. Coelho      | S. Motta          | F.M. Schleif   |
| M. Cristani    | O. Okun           | A. Zell        |
| D. di Bernardo | L. Palopoli       |                |

## Local Scientific Committee

|                      |                                                   |
|----------------------|---------------------------------------------------|
| Ida Bifulco          | NEuRoNe Lab, DMI, University of Salerno,<br>Italy |
| Loredana Murino      | NEuRoNe Lab, DMI, University of Salerno,<br>Italy |
| Francesco Napolitano | NEuRoNe Lab, DMI, University of Salerno,<br>Italy |

## Financing Institutions

GNCS, Gruppo Nazionale Calcolo Scientifico, Italy  
DMI, University of Salerno, Italy

# Table of Contents

## Invited Papers

|                                                                                                                                                                          |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Coarse-Grained Modeling of the HIV-1 Protease Binding Mechanisms:<br>I. Targeting Structural Flexibility of the Protease Flaps and<br>Implications for Drug Design ..... | 1  |
| <i>Gennady M. Verkhivker</i>                                                                                                                                             |    |
| Coarse-Grained Modeling of the HIV-1 Protease Binding Mechanisms:<br>II. Folding Inhibition .....                                                                        | 13 |
| <i>Gennady M. Verkhivker</i>                                                                                                                                             |    |
| Unsupervised Stability-Based Ensembles to Discover Reliable<br>Structures in Complex Bio-molecular Data .....                                                            | 25 |
| <i>Alberto Bertoni and Giorgio Valentini</i>                                                                                                                             |    |

## CIBB Regular Session

|                                                                                                                                                                         |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Comparative <i>In Silico</i> Evaluation of MYB Transcription Factors in<br>Eucalyptus, Sugarcane and Rice Transcriptomes .....                                          | 44  |
| <i>Nina M. Soares-Cavalcanti, Ana C. Wanderley-Nogueira,<br/>Luis C. Belarmino, Petra dos Santos Barros, and<br/>Ana M. Benko-Iseppon</i>                               |     |
| Building Maps of Drugs Mode-of-Action from Gene Expression Data ...                                                                                                     | 56  |
| <i>Francesco Iorio, Roberto Tagliaferri, and Diego di Bernardo</i>                                                                                                      |     |
| <i>In Silico</i> Evaluation of Osmoprotectants in <i>Eucalyptus</i><br>Transcriptome .....                                                                              | 66  |
| <i>Petra dos Santos Barros, Nina M. Soares-Cavalcanti,<br/>Gabriela S. Vieira-Mello, Ana C. Wanderley-Nogueira,<br/>Tercilio Calsa-Junior, and Ana M. Benko-Iseppon</i> |     |
| Mining Association Rule Bases from Integrated Genomic Data and<br>Annotations .....                                                                                     | 78  |
| <i>Ricardo Martinez, Nicolas Pasquier, and Claude Pasquier</i>                                                                                                          |     |
| Stability and Performances in Biclustering Algorithms .....                                                                                                             | 91  |
| <i>Maurizio Filippone, Francesco Masulli, and Stefano Rovetta</i>                                                                                                       |     |
| Splice Site Prediction Using Artificial Neural Networks .....                                                                                                           | 102 |
| <i>Øystein Johansen, Tom Ryen, Trygve Eftesøl,<br/>Thomas Kjosmoen, and Peter Ruoff</i>                                                                                 |     |

|                                                                                                                   |     |
|-------------------------------------------------------------------------------------------------------------------|-----|
| Interval Length Analysis in Multi Layer Model .....                                                               | 114 |
| <i>Vito Di Gesù, Giosuè Lo Bosco, and Luca Pinello</i>                                                            |     |
| A Multivariate Algorithm for Gene Selection Based on the Nearest Neighbor Probability .....                       | 123 |
| <i>Enrico Ferrari and Marco Muselli</i>                                                                           |     |
| Control of Cellular Glycolysis by Perturbations in the Glucose Influx .....                                       | 132 |
| <i>Jaime Arturo de la Torre, Maria Carmen Lemos, and Antonio Córdoba</i>                                          |     |
| Curating a Large-Scale Regulatory Network by Evaluating Its Consistency with Expression Datasets .....            | 144 |
| <i>Carito Guziolowski, Jeremy Gruel, Ovidiu Radulescu, and Anne Siegel</i>                                        |     |
| Spatial Clustering of Molecular Dynamics Trajectories in Protein Unfolding Simulations .....                      | 156 |
| <i>Pedro Gabriel Ferreira, Cândida G. Silva, Paulo J. Azevedo, and Rui M.M. Brito</i>                             |     |
| Clustering Bacteria Species Using Neural Gas: Preliminary Study .....                                             | 167 |
| <i>Claudio Mirabello, Riccardo Rizzo, Alfonso Urso, and Salvatore Gaglio</i>                                      |     |
| A New Linear Initialization in SOM for Biomolecular Data .....                                                    | 177 |
| <i>Antonino Fiannaca, Giuseppe Di Fatta, Alfonso Urso, Riccardo Rizzo, and Salvatore Gaglio</i>                   |     |
| 3D Volume Reconstruction and Biometric Analysis of Fetal Brain from MR Images .....                               | 188 |
| <i>Paola Campadelli, Elena Casiraghi, Gabriele Lombardi, and Graziano Serrao</i>                                  |     |
| Searching for Glycomics Role in Stem Cell Development .....                                                       | 198 |
| <i>Anil Sorathiya, Tadas Jucikas, Stephanie Piecewicz, Shiladitya Sengupta, and Pietro Liò</i>                    |     |
| A New Protein Representation Based on Fragment Contacts: Towards an Improvement of Contact Maps Predictions ..... | 210 |
| <i>Pietro Di Lena, Luciano Margara, Marco Vassura, Piero Fariselli, and Rita Casadio</i>                          |     |
| Analysis of Kernel Based Protein Classification Strategies Using Pairwise Sequence Alignment Measures .....       | 222 |
| <i>Dino Franklin, Somdutta Dhir, and Sándor Pongor</i>                                                            |     |

## Special Session: ISMDS - Intelligent Systems for Medical Decisions Support

|                                                                                                                          |     |
|--------------------------------------------------------------------------------------------------------------------------|-----|
| Topology Preserving Neural Networks for Peptide Design in Drug Discovery .....                                           | 232 |
| <i>Jörg D. Wichard, Sebastian Bandholtz, Carsten Grötzinger, and Ronald Kühne</i>                                        |     |
| A Machine Learning Approach to Mass Spectra Classification with Unsupervised Feature Selection .....                     | 242 |
| <i>Michele Ceccarelli, Antonio d’Acierno, and Angelo Facchiano</i>                                                       |     |
| Liver i-Biopsy <sup>TM</sup> and the Corresponding Intelligent Fibrosis Scoring Systems: i-Metavir F and i-Ishak F ..... | 253 |
| <i>Alexandru George Floares</i>                                                                                          |     |

## Special Session: Computational Intelligence for Biological Data Visualization

|                                                                                                                                                              |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| An Extension of the TIGR M4 Suite to Preprocess and Visualize Affymetrix Binary Files .....                                                                  | 265 |
| <i>Mario Cannataro, Maria Teresa Di Martino, Pietro Hiram Guzzi, Pierosandro Tagliaferri, Pierfrancesco Tassone, Giuseppe Tradigo, and Pierangelo Veltri</i> |     |
| A Supervised Learning Technique and Its Applications to Computational Biology .....                                                                          | 275 |
| <i>Mario R. Guarracino, Altannar Chinchuluun, and Panos M. Pardalos</i>                                                                                      |     |
| A Visualization ToolKit Based Application for Representing Macromolecular Surfaces .....                                                                     | 284 |
| <i>Paolo Cozzi, Ivan Merelli, and Luciano Milanese</i>                                                                                                       |     |
| <b>Author Index</b> .....                                                                                                                                    | 293 |