

## Subseries of Lecture Notes in Computer Science

### LNBI 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*

### LNBI Editorial Board

Søren Brunak

*Technical University of Denmark, Kongens Lyngby, Denmark*

Mikhail S. Gelfand

*IITP, Research and Training Center on Bioinformatics, Moscow, Russia*

Thomas Lengauer

*Max Planck Institute for Informatics, Saarbrücken, Germany*

Satoru Miyano

*University of Tokyo, Tokyo, Japan*

Eugene Myers

*Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany*

Marie-France Sagot

*Université Lyon 1, Villeurbanne, France*

David Sankoff

*University of Ottawa, Ottawa, Canada*

Ron Shamir

*Tel Aviv University, Ramat Aviv, Tel Aviv, Israel*

Terry Speed

*Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia*

Martin Vingron

*Max Planck Institute for Molecular Genetics, Berlin, Germany*

W. Eric Wong

*University of Texas at Dallas, Richardson, TX, USA*


More information about this series at <http://www.springer.com/series/5381>

Massimo Bartoletti · Annalisa Barla  
Andrea Bracciali · Gunnar W. Klau  
Leif Peterson · Alberto Policriti  
Roberto Tagliaferri (Eds.)


# Computational Intelligence Methods for Bioinformatics and Biostatistics

14th International Meeting, CIBB 2017  
Cagliari, Italy, September 7–9, 2017  
Revised Selected Papers

### *Editors*


Massimo Bartoletti   
University of Cagliari  
Cagliari, Italy


Annalisa Barla   
University of Genova  
Genoa, Italy

Andrea Bracciali   
University of Stirling  
Stirling, UK

Gunnar W. Klau   
Heinrich-Heine-University Düsseldorf  
Düsseldorf, Germany

Leif Peterson   
Houston Methodist Research Institute  
Houston, TX, USA

Alberto Policriti   
University of Udine  
Udine, Italy

Roberto Tagliaferri   
University of Salerno  
Pisciano, Italy

ISSN 0302-9743

Lecture Notes in Bioinformatics

ISBN 978-3-030-14159-2

<https://doi.org/10.1007/978-3-030-14160-8>

ISSN 1611-3349 (electronic)

ISBN 978-3-030-14160-8 (eBook)

Library of Congress Control Number: 2019932170

LNCS Sublibrary: SL8 – Bioinformatics

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, 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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

## Preface

The 14th annual edition of the International Meeting on Computational Intelligence methods for Bioinformatics and Biostatistics (CIBB 2017) built upon the tradition of the CIBB conference series and provided a multi-disciplinary forum open to researchers interested in the application of computational intelligence, in a broad sense, to open problems in bioinformatics, biostatistics, systems and synthetic biology, medical informatics, as well as computational approaches to life sciences in general.

In line with the spirit of CIBB, the 2017 meeting brought together researchers from different communities who address problems from different, but connected and often overlapping, perspectives. CIBB 2017 tackled the difficult task of bridging different backgrounds by providing an inclusive venue to discuss advances and future perspectives in different areas. It also fostered interaction between theory and practice, addressing both the theories underpinning the methodologies used to model and analyze biological systems, the practical applications of such theories, and the supporting technologies. Accordingly, participants at CIBB 2017 came from mathematical, computational, and medical backgrounds and institutions, both from academia and the private sector, offering collaboration opportunities and novel results in the areas of computational life sciences.

CIBB 2017 also offered a view on emerging and strongly developing trends and future opportunities at the edge of mathematics as well as computer and life sciences, such as synthetic biology, statistical investigation of genomic data, and applications to the understanding of complex diseases, such as cancer, and therapy opportunities. Along these lines, three keynote speakers, prominent scholars in their fields, presented the latest advances of their research within the context of their area of interest, and provided insights into open problems and future directions of general interest for the field. While papers in the main conference track addressed a rich set of open problems at the forefront of current research, the conference hosted three further special sessions on specific themes: synthesis of artificial cells by combining bio-interfaces engineering and systems biology modeling, modeling and simulation methods for systems biology and systems medicine, and molecular communication. Researchers from Europe, Asia, and America attended the conference.

CIBB 2017 was made possible by the efforts of the Organizing, Program, and Steering Committees and by the support of sponsors and participants. CIBB 2017 was held in Cagliari, Italy, during September 7–9, 2017. Overall, 44 contributions were submitted for consideration to CIBB 2017, amongst which 33 were invited for an oral presentation at the conference, after a first round of reviews (at this stage, each paper received an average of 3.25 reviews from the Program Committee and about 13 additional referees). Following the conference, selected papers were invited for further submission, after feedback and discussion from the conference. This volume collects the papers that were accepted after a further round of reviews (2.5 for each paper, on average).

From 2004 to 2007, CIBB had 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 strongly rated papers, the Steering Committee decided to turn CIBB into an autonomous conference starting with the 2008 edition in Vietri. The following editions in Italian venues were held in Genoa (2009), Palermo (2010), and Gargnano (2011). Until 2012, CIBB meetings were held annually in Italy with an increasing number of participants. CIBB 2012 was the first edition organized outside Italy, in Houston, then in Nice, France (2013), Cambridge, UK (2014), Naples, Italy (2015), and Stirling, UK (2016).

A rigorous peer-review selection process is applied every time to ultimately select the papers included in the program of the conference, in the conference proceedings published in the LNBI-LNCS book series by Springer, and in some cases, selected papers were published in special issues of well-qualified international journals, such as *BMC Bioinformatics*.

December 2018

Annalisa Barla  
Massimo Bartoletti  
Andrea Bracciali  
Gunnar W. Klau  
Leif Peterson  
Alberto Policriti  
Roberto Tagliaferri

# Organization

## Program Committee

Antonino Abbruzzo	University of Palermo, Italy
Marco Antoniotti	University of Milano-Bicocca, Italy
Sansanee Auephanwiriyakul	Chiang Mai University, Thailand
Massimo Bartoletti	University of Cagliari, Italy
Gilles Bernot	University of Nice Sophia Antipolis, France
Daniela Besozzi	University of Milano-Bicocca, Italy
Andrea Bracciali	University of Stirling, UK
Stefan Canzar	Gene Center, LMU, Germany
Giulio Caravagna	Institute of Cancer Research, UK
Paolo Cazzaniga	University of Bergamo, Italy
Davide Chicco	Peter Munk Cardiac Centre, Canada
Angelo Ciaramella	University of Naples Parthenope, Italy
Luisa Cutillo	SITraN, University of Sheffield, UK
Chiara Damiani	University of Milano-Bicocca, Italy
Angelo Facchiano	CNR, Istituto di Scienze dell'Alimentazione, Italy
Alessandro Filisetti	Explora Biotech Srl, Italy
Enrico Formenti	University of Nice Sophia Antipolis, France
Christoph M. Friedrich	University of Applied Science and Arts Dortmund, Germany
Yair Goldberg	University of Haifa, Israel
Alex Graudenzi	University of Milano-Bicocca, Italy
Sean Holden	University of Cambridge, UK
Gunnar W. Klau	Heinrich-Heine-University Düsseldorf, Germany
Johannes Köster	University of Duisburg-Essen, Germany
Paola Lecca	University of Trento, Italy
Hassan Mahmoud	University of Genoa, Italy
Anna Marabotti	University of Salerno, Italy
Roberto Marangoni	University of Pisa, Italy
Tobias Marschall	Saarland University/Max Planck Institute for Informatics, Germany
Giancarlo Mauri	University of Milano-Bicocca, Italy
Fabio Mavelli	University of Bari "Aldo Moro", Italy
Pasquale Palumbo	IASI-CNR, Italy
Dario Pescini	University of Milano-Bicocca, Italy
Leif Peterson	Houston Methodist Research Institute, USA
Alberto Policriti	University of Udine, Italy
Vilda Purutçuoğlu	Middle East Technical University, Turkey

Davide Riso	Cornell University, USA
Riccardo Rizzo	ICAR-CNR, Italy
Andrea Roli	University of Bologna, Italy
Federico Rossi	University of Salerno, Italy
Stefano Rovetta	University of Genova, Italy
Antonino Staiano	University of Naples Parthenope, Italy
Pasquale Stano	University of Salento, Italy
Francesco Stingo	MD Anderson Cancer Center, USA
Roberto Tagliaferri	University of Salerno, Italy
Paolo Tieri	Consiglio Nazionale delle Ricerche, Italy
Filippo Utro	IBM T. J. Watson Research, USA
Alfredo Vellido	Universitat Politècnica de Catalunya, Spain
Paweł P. Łabaj	Boku University Vienna, Austria

## **Additional Reviewers**

Alqassem, Israa	Nouri, Nima
Angaroni, Fabrizio	Pucciarelli, Sandra
Femminella, Mauro	Ramazzotti, Daniele
Graudenzi, Alex	Rautiainen, Mikko
Kohvaei, Parastou	Rossi, Nicolò



# Contents

An Open-Source Tool for Managing Time-Evolving Variant Annotation . . . .	1
<i>Ilio Catallo, Eleonora Ciceri, Stefania Stenirri, Stefania Merella, Alberto Sanna, Maurizio Ferrari, Paola Carrera, and Sauro Vicini</i>	
Extracting Few Representative Reconciliations with Host Switches . . . . .	9
<i>Mattia Gastaldello, Tiziana Calamoneri, and Marie-France Sagot</i>	
A Quantitative and Qualitative Characterization of $k$ -mer Based Alignment-Free Phylogeny Construction . . . . .	19
<i>Filippo Utro, Daniel E. Platt, and Laxmi Parida</i>	
Cancer Mutational Signatures Identification with Sparse Dictionary Learning. . . . .	32
<i>Veronica Tozzo and Annalisa Barla</i>	
ICING: Large-Scale Inference of Immunoglobulin Clonotypes. . . . .	42
<i>Federico Tomasi, Margherita Squillario, Alessandro Verri, Davide Bagnara, and Annalisa Barla</i>	
ADENINE: A HPC-Oriented Tool for Biological Data Exploration . . . . .	51
<i>Samuele Fiorini, Federico Tomasi, Margherita Squillario, and Annalisa Barla</i>	
Disease–Genes Must Guide Data Source Integration in the Gene Prioritization Process . . . . .	60
<i>Marco Frasca, Jean Fred Fontaine, Giorgio Valentini, Marco Mesiti, Marco Notaro, Dario Malchiodi, and Miguel A. Andrade-Navarro</i>	
Ensembling Descendant Term Classifiers to Improve Gene - Abnormal Phenotype Predictions . . . . .	70
<i>Marco Notaro, Max Schubach, Marco Frasca, Marco Mesiti, Peter N. Robinson, and Giorgio Valentini</i>	
GP-Based Grammatical Inference for Classification of Amyloidogenic Sequences . . . . .	81
<i>Wojciech Wieczorek and Olgierd Unold</i>	
Estimation of Kinetic Reaction Constants: Exploiting Reboot Strategies to Improve PSO’s Performance. . . . .	92
<i>Simone Spolaor, Andrea Tangherloni, Leonardo Rundo, Paolo Cazzaniga, and Marco S. Nobile</i>	

Haplotype and Repeat Separation in Long Reads . . . . .	103
<i>German Tischler-Höhle</i>	
Tumor Subclonal Progression Model for Cancer Hallmark Acquisition . . . . .	115
<i>Yusuke Matsui, Satoru Miyano, and Teppei Shimamura</i>	
GIMLET: Identifying Biological Modulators in Context-Specific Gene Regulation Using Local Energy Statistics . . . . .	124
<i>Teppei Shimamura, Yusuke Matsui, Taisuke Kajino, Satoshi Ito, Takashi Takahashi, and Satoru Miyano</i>	
Structural Features of a DPPG Liposome Layer Adsorbed on a Rough Surface. . . . .	138
<i>Maria Raposo, Andreia A. Duarte, Paulo J. Gomes, Paulo A. Ribeiro, Marli L. Moraes, and Roland Steitz</i>	
Chemical Exchanges and Actuation in Liposome-Based Synthetic Cells: Interaction with Biological Cells . . . . .	145
<i>Giordano Rampioni, Francesca D'Angelo, Alessandro Zennaro, Livia Leoni, and Pasquale Stano</i>	
A Nano Communication System for CTC Detection in Blood Vessels . . . . .	159
<i>Luca Felicetti, Mauro Femminella, and Gianluca Reali</i>	
Experimental Evidences Suggest High Between-Vesicle Diversity of Artificial Vesicle Populations: Results, Models and Implications . . . . .	171
<i>Pasquale Stano, Roberto Marangoni, and Fabio Mavelli</i>	
Towards the Synthesis of Photo-Autotrophic Protocells . . . . .	186
<i>Emiliano Altamura, Paola Albanese, Roberto Marotta, Pasquale Stano, Francesco Milano, Massimo Trotta, and Fabio Mavelli</i>	
Hierarchical Block Matrix Approach for Multi-view Clustering. . . . .	200
<i>Angela Serra, Maria Domenica Guida, Pietro Lió, and Roberto Tagliaferri</i>	
Author Index . . . . .	213