# **Lecture Notes in Bioinformatics**

# 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

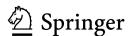
University of Texas at Dallas, Richardson, TX, USA

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

Ignacio Rojas · Francisco Ortuño (Eds.)

# Bioinformatics and Biomedical Engineering

5th International Work-Conference, IWBBIO 2017 Granada, Spain, April 26–28, 2017 Proceedings, Part I



Editors Ignacio Rojas Universidad de Granada Granada Spain

Francisco Ortuño Universidad de Granada Granada Spain

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Bioinformatics ISBN 978-3-319-56147-9 ISBN 978-3-319-56148-6 (eBook) DOI 10.1007/978-3-319-56148-6

Library of Congress Control Number: 2017935843

LNCS Sublibrary: SL8 - Bioinformatics

#### © Springer International Publishing AG 2017

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.

Printed on acid-free paper

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

## **Preface**

We are proud to present the set of final accepted full papers for the third edition of the IWBBIO conference "International Work-Conference on Bioinformatics and Biomedical Engineering" held in Granada (Spain) during April 26–28, 2017.

The IWBBIO 2017 (International Work-Conference on Bioinformatics and Biomedical Engineering) seeks to provide a discussion forum for scientists, engineers, educators, and students about the latest ideas and realizations in the foundations, theory, models, and applications for interdisciplinary and multidisciplinary research encompassing the disciplines of computer science, mathematics, statistics, biology, bioinformatics, and biomedicine.

The aim of IWBBIO is to create a friendly environment that could lead to the establishment or strengthening of scientific collaborations and exchanges among attendees, and, therefore, IWBBIO 2017 solicited high-quality original research papers (including significant work in progress) on any aspect of bioinformatics, biomedicine, and biomedical engineering.

The following topics were especially encouraged: new computational techniques and methods in machine learning; data mining; text analysis; pattern recognition; data integration; genomics and evolution; next-generation sequencing data; protein and RNA structure; protein function and proteomics; medical informatics and translational bioinformatics; computational systems biology; modelling and simulation and their application in the life science domain, biomedicine, and biomedical engineering. The list of topics in the successive call for papers also evolved, resulting in the following list for the present edition:

- 1. **Computational proteomics**. Analysis of protein–protein interactions. Protein structure modelling. Analysis of protein functionality. Quantitative proteomics and PTMs. Clinical proteomics. Protein annotation. Data mining in proteomics.
- Next-generation sequencing and sequence analysis. De novo sequencing, re-sequencing, and assembly. Expression estimation. Alternative splicing discovery. Pathway analysis. Chip-seq and RNA-Seq analysis. Metagenomics. SNPs prediction.
- 3. **High performance in bioinformatics**. Parallelization for biomedical analysis. Biomedical and biological databases. Data mining and biological text processing. Large-scale biomedical data integration. Biological and medical ontologies. Novel architecture and technologies (GPU, P2P, Grid etc.) for bioinformatics.
- 4. **Biomedicine**. Biomedical Computing. Personalized medicine. Nanomedicine. Medical education. Collaborative medicine. Biomedical signal analysis. Biomedicine in industry and society. Electrotherapy and radiotherapy.
- 5. **Biomedical engineering**. Computer-assisted surgery. Therapeutic engineering. Interactive 3D modelling. Clinical engineering. Telemedicine. Biosensors and data acquisition. Intelligent instrumentation. Patient monitoring. Biomedical robotics. Bio-nanotechnology. Genetic engineering.

- 6. Computational systems for modelling biological processes. Inference of biological networks. Machine learning in bioinformatics. Classification for biomedical data. Microarray data analysis. Simulation and visualization of biological systems. Molecular evolution and phylogenetic modelling.
- 7. **Health care and diseases**. Computational support for clinical decisions. Image visualization and signal analysis. Disease control and diagnosis. Genome-phenome analysis. Biomarker identification. Drug design. Computational immunology.
- 8. **E-health**. E-health technology and devices. E-health information processing. Telemedicine/e-health application and services. Medical image processing. Video techniques for medical images. Integration of classical medicine and e-health.

After a careful peer-review and evaluation process (each submission was reviewed by at least two, and on average 3.2, Program Committee members or additional reviewers), 120 papers were accepted for oral, poster, or virtual presentation, according to the recommendations of the reviewers and the authors' preferences, and to be included in the LNBI proceedings.

During IWBBIO 2017, several Special Sessions were carried out. Special Sessions are a very useful tool to complement the regular program with new and emerging topics of particular interest for the participating community. Special Sessions that emphasize multi-disciplinary and transversal aspects as well as cutting-edge topics were especially encouraged and welcomed, and in IWBBIO 2017 they were the following:

#### - SS1: Advances in Computational Intelligence for Critical Care

Decision-making in health care in clinical environments is often made on the basis of multiple parameters and in the context of patient presentation, which includes the setting and the specific conditions related to the reason for admission and the procedures involved. The data used in clinical decision-making may originate from manifold sources and at multiple scales: devices in and around the patient, laboratory, blood tests, omics analyses, medical images, and ancillary information available both prior to and during the hospitalization.

Arguably, one of the most data-dependent clinical environments is the intensive care unit (ICU). The ICU environment cares for acutely ill patients. Many patients within ICU environments, and particularly surgical intensive care units (SICU), are technologically dependent on the life-sustaining devices that surround them. Some of these patients are indeed dependent for their very survival on technologies such as infusion pumps, mechanical ventilators, catheters and so on. Beyond treatment, assessment of prognosis in critical care and patient stratification combining different data sources are extremely important in a patient-centric environment.

With the advent and quick uptake of omics technologies in critical care, the use of data-based approaches for assistance in diagnosis and prognosis becomes paramount. New approaches to data analysis are thus required, and some of the most interesting ones currently stem from the fields of computational intelligence (CI) and machine learning (ML). This session is particularly interested in the proposal of novel CI and ML approaches and in the discussion of the challenges for the application of the existing ones to problems in critical care.

Topics that are of interest to this session include (but are not necessarily limited to):

- Novel applications of existing CI and ML and advanced statistical methods in critical care
- Novel CI and ML techniques for critical care
- CI and ML-based methods to improve model interpretability in a critical care context, including data/model visualization techniques
- Novel CI and ML techniques for dealing with non-structured and heterogeneous data formats in critical care

Organizers: Dr. Alfredo Vellido, PhD, Department of Computer Science, Universitat Politécnica de Catalunya, BarcelonaTECH (UPC), Barcelona (Spain). Dr. Vicent Ribas, eHealth Department, EURECAT Technology Centre of Catalonia, Barcelona, Barcelona (Spain).

#### SS2: Time-Lapse Experiments and Multivariate Biostatistics

Biological samples are evolving in time, phases, periods, behavior. To be able to understand the dynamics, we need to perform time lapse experiments. Today's technique and measurement devices allow us to monitor numerous parameters in semi-controled environments during the experiment. The increase of measured data is enormous. The interpretation requires both qualitative and quantitative analysis. There are useful methods of biostatistics, multivariate data analysis, and artificial intelligence, namely, neural networks, genetic algorithms, and agent-based modeling, respectively.

In this special section we will provide a discussion on broad examples from time-lapse experimental design through information and data acquisition, using methods of bioinformatics, biophysics, biostatistics, and artificial inteligence. The aim of this section is to present the possible increase in data interpretation and related methods.

Organizer: Dr. Jan Urban, Laboratory of Signal and Image Processing, Institute of Complex Systems, Faculty of Fisheries and Protection of Waters, University of South Bohemia.

## SS3: Half-Day GATB Tutorial. The Genome Analysis Toolbox with de Bruijn Graph

The GATB programming day is an educational event organized by the GATB team. This free event is open to everyone who is familiar in C++ programming and wants to learn how to create NGS data analysis software.

The tutorial has a focus on the high-performance GATB-Core library and is taught by the developers of this library.

During this half-day tutorial, some of the following topics are explored:

- A theoretical introduction to GATB: the basic concepts.
- GATB-Core practical coding session 1: I/O operations on reads files.
- The GATB de Bruijn graph API.
- GATB-Core practical coding session 2: k-mer and graph APIs in action.
- GATB-Core practical coding session 3: writing a short read corrector tool.
- Q and A session: obtain answers from GATB experts.

Organizers: Dr. Dominique Lavenier, GenScale Team Leader, Inria/IRISA, Campus de Beaulieu, Rennes, France.

Dr. Patrick Durand, Inria, Genscale Team, Campus de Beaulieu, Rennes, France.

## SS4: Medical Planning: Operating Theater Design and Its Impact on Cost, Area, and Workflow

The design of operating rooms is one of the most complicated tasks of hospital design because of its characteristics and requirements. Patients, staff, and tools should have determined passes through the operating suite. Many hospitals assume that the operating suite is the most important unit in the hospital because of its high–revenue. Arch design of these suites is a very critical point in solving an optimized problem in spaces, workflow of clean, dirty, and patient in/out in addition to staff together with their relations with adjacent departments. In this session, we illustrate the most common designs of operating suites and select the most suitable one to satisfy the effectiveness of the operating suite, maximizing throughput, minimizing the costs, and decreasing the required spaces related to available resources/possibilities. The design should comply with country guidelines, infection control rules, occupational safety and health, and satisfy the maximum benefits for patients and staff. A comparative study was performed on 15 hospitals and it recorded that the single input—output technique is the best design.

Motivation and objectives for the session: Operating suite design is a very critical task owing to its impact. Biomedical engineers should participate in the design and review the workflow and available functions.

Organizer: Dr. Khaled El-Sayed, Department of Electrical and Medical Engineering at Benha University, Egypt

#### SS5: Challenges Representing Large-Scale Biological Data

Visualization models have been shown to be remarkably important in the interpretation of datasets across many fields of study. In the context of bioinformatics and computational biology various tools have been proposed to visualize molecular data and help understand how biological systems work. Despite that, several challenges still persist when faced with complex and dynamic data and major advances are required to correctly manage the multiple dimensions of the data.

The aim of this special session is to bring together researchers to present recent and ongoing research activities related to advances in visualization techniques and tools, focused on any major molecular biology problem, with the aim of allowing for the exchange and sharing of proposed ideas and experiences with novel visualization metaphors.

Topics of interest include:

- Genome and sequence data
- Omics data (transcriptomics, proteomics, metabolomics)
- Biological networks and pathways
- · Time-series data
- · Biomedical ontologies
- Macromolecular complexes
- Phylogenetic data

- Biomarker discovery
- · Integration of image and omics data for systems biology
- Modeling and simulation of dynamical processes

Organizer: Prof. Joel P. Arrais FCTUC – University of Coimbra, Portugal.

# SS6: Omics of Space Travelled Microbes – Bioinformatics and Biomedical Aspects

The National Research Council (NRC) Committee for the Decadal Survey on Biological and Physical Sciences in Space reported that "microbial species that are uncommon, or that have significantly increased or decreased in number, can be studied in a "microbial observatory" on the International Space Station (ISS)." As part of the microbial observatory effort the NRC decadal survey committee suggested that NASA should: "(a) capitalize on the technological maturity, low cost, and speed of genomic analyses and the rapid generation time of microbes to monitor the evolution of microbial genomic changes in response to the selective pressures present in the spaceflight environment; (b) study changes in microbial populations from the skin and feces of the astronauts, plant and plant growth media, and environmental samples taken from surfaces and the atmosphere of the ISS; and (c) establish an experimental program targeted at understanding the influence of the spaceflight environment on defined microbial populations."

The proposed session discusses state-of-the-art molecular techniques, bioinformatics tools, and their benefit in answering the astronauts and others who live in closed systems.

Organizer: Dr. Kasthuri Venkateswaran, Senior Research Scientist, California Institute of Technology, Jet Propulsion Laboratory, Biotechnology and Planetary Protection Group, Pasadena, CA

#### SS7: Data-Driven Biology – New Tools, Techniques, and Resources

Advances in sequencing techniques have accelerated data generation at diverse regulatory levels in an unprecedented way. The challenge now is to integrate these data to understand regulation at a systems level. As the sequencing technologies evolve, new tools and resources follow, revealing new aspects of complex biological systems.

This special session brings together experts from computational biology and machine learning to present recent advances in the development of new tools and resources using next-generation sequencing data including novel emerging fields such as single-cell transcriptomics. The session features an invited speaker and three/four short talks. To promote emerging leaders of the field, we select invited speakers who have gained their independence in recent years.

Organizer: Dr. Joshi Anagha, Division of Developmental Biology at the Roslin Institute, University of Edinburgh

#### SS8: Smart Sensor and Sensor-Network Architectures

There is a significant demand for tools and services supporting rehabilitation, well-being and healthy life styles while reducing the level of intrusiveness as well as increasing real-time available and reliable results. For example, self-monitoring

applications need to be improved to move beyond tracking exercise habits and capture a more comprehensive digital footprint of human behavior. This session focuses on primary parameter capturing devices and networks demonstrating advances in sensor development including a customized algorithmic shell research to support diagnostic decisions. Target domains are, for example, continuous differentiating between mental and physical stress, blood pressure monitoring, sleep quality monitoring, HRV etc.

Organizers: Prof. Dr. Natividad Martinez, Internet of Things Laboratory, Reutlingen University, Germany.

Prof. Dr. Juan Antonio Ortega, University of Seville, Spain.

Prof. Dr. Ralf Seepold, Ubiquitous Computing Lab, HTWG Konstanz, Germany.

#### SS9: High-Throughput Bioinformatic Tools for Genomics

Genomics is concerned with the sequencing and analysis of an organism's genome. It is involved in the understanding of how every single gene can affect the entire genome. This goal is mainly afforded using the current, cost-effective, high-throughput sequencing technologies. These technologies produce a huge amount of data that usually require high-performance computing solutions and opens new ways for the study of genomics, but also transcriptomics, gene expression, and systems biology, among others. The continuous improvements and broader applications on sequencing technologies is producing an ongoing demand for improved high-throughput bioinformatics tools.

Therefore, we invite authors to submit original research, new tools or pipelines, or their update, and review articles on topics helping in the study of genomics in the wider sense, such as (but not limited to):

- Tools for data pre-processing (quality control and filtering)
- Tools for sequence mapping
- Tools for de novo assembly
- Tools for quality check of sequence assembling
- Tools for the comparison of two read libraries without an external reference
- Tools for genomic variants (such as variant calling or variant annotation)
- Tools for functional annotation: identification of domains, orthologues, genetic markers, controlled vocabulary (GO, KEGG, InterPro)
- Tools for biological enrichment in non-model organisms
- Tools for gene expression studies
- Tools for Chip-Seq data
- Tools for "big-data" analyses
- Tools for handling and editing complex workflows and pipelines
- Databases for bioinformatics

Organizers: Prof. M. Gonzalo Claros, Department of Molecular Biology and Biochemistry, University of Málaga, Spain.

Dr. Javier Pérez Florido, Bioinformatics Research Area, Fundación Progreso y Salud, Seville, Spain.

#### SS10: Systems Biology Approaches to Decipher Long Noncoding RNA-Protein Associations

Long noncoding RNAs (lncRNAs) make up large amounts of the RNA and total genomic repertoire. Studies on the functional characterization of lncRNAs have resulted in data on interactions with their RNA peers, DNA, or proteins. Although there has been an increase in evidence on the link between lncRNAs and diverse human diseases, there is a dearth of lncRNA-protein association studies. Additionally, existing methods do not provide theories about the possible molecular causes of such associations linking to diseases. How such regulatory interactions between classes of lncRNAs and proteins would have a significant influence on the organism and disease remains a challenge. A good number of bioinformatics approaches have arisen in the recent past exploring these challenges. The idea behind this session is to bring together the wide gamut of researchers who have worked on these methods across different organisms.

The following are the sub-topics of the proposed session, which we would like to call for papers.

- LncRNAs in genomes: annotation and curation
- LncRNA-protein interactions leading to important diseases: systems Biology approaches
- Identifying lncRNAs with respect to their mechanism and dysregulation in diseases
- LncRNA databases and webservers
- Machine learning approaches and prediction servers

Organizer: Prashanth Suravajhala, PhD, Department of Biotechnology and Bioinformatics, Birla Institute of Scientific Research, India

#### - SS11: Gamified Rehabilitation for Disabled People

Gamification is a hot topic in many areas as it aims at motivating people to do things driven by different innate needs like the wish to accomplish tasks, to compete against others, or to gain something. These and other motivators are efficiently applied in computer games and could be extraordinarily useful in ensuring that patients perform their daily exercises regularly and have fun.

The idea of exergames (exercise games) is not new; the literature reveals that much effort has already been made and with great success. Nevertheless, most applications have been developed for special problems or diseases (i.e., stroke, parkinson, cerebral palsy etc.) and are not generally applicable. In general, people suffering from severe disabilities and chronic diseases are rarely addressed as a target group. Also, the focus is generally set on the medical achievements, which is correct, but the next step would be to enhance the fun factor because no tool is of much use if the patient is not using it because of boredom or demotivation.

The objective of this special session is to gather new ideas about the combination of need and fun, i.e., find ways to create exercise platforms that fit everybody's needs, provide access to the therapist for monitoring and configuration, while the patients benefit physically and mentally when having a good time.

Target groups would be people of:

- All ages, while focusing on younger people, who can be involved more easily but are less addressed in the literature
- All diseases, while focusing on chronic illness and severe disabilities (e.g., muscle dystrophies and atrophies)

The contributions should show advances in at least one of the following areas:

- Adaptability to users with all kinds of problems (e.g., possibility to configure the limbs used to play or playing with facial movements, wheelchair and standing modes, coping with muscle weaknesses etc.)
- Implementation of gaming techniques and special motivators
- Physical or mental exercises, aimed at rehabilitation or daily practice
- Understanding the users, awareness of their level of motivation, fatigue or progress and react accordingly

Organizer: Dr. Martina Eckert, Associate Professor at University of Madrid, Spain.

#### SS12: Modelling of Glucose Dynamics for Diabetes

Diabetes is the eighth most common cause of death, while its treatment relies on technology to process continuously measured glucose levels.

Organizer: Dr. Tomas Koutny, Faculty of Applied Sciences, University of West Bohemia

#### SS13: Biological Network Analysis in Multi-omics Data Integration

In many biological applications, multiple data types may be produced to determine the genetics, epigenetics, and microbiome affecting gene regulation and metabolism. Although producing multiple data types should provide a more complete description of the processes under study due to multiple factors such as study design (synchronization of data production, number of samples, varying conditions), the analysis may leave more unfulfilled promises than synergy expected from the wealth of data.

In this session, some of the following challenges are addressed:

- How to conduct meaningful meta-analysis on historical data.
- How to use biological knowledge (represented in reproducible and interoperable manner) in the analysis of large and sparse data sets more effectively.
- How to fill the gap between hypothesis-driven mechanistic studies, e.g., applying modelling to very well studied biochemical processes and data-driven hypothesis-free approaches. How omics data can help.
- Beyond meta-transcriptomics and metagenomics: integration and interpretation of microbiome and host data.

We would like to bring together communities concerned with these topics to present state-of-the-art and current cutting-edge developments, preferably work under construction or published within the past year.

- Objective 1: presentation and discussion of newest methods
- Objective 2: round-table discussions on the topics highlighted above and other related topics suggested by session participants

An additional topic that does not fit the proposed session but that I would love to see addressed is: How to improve open access to the data that is not next-generation sequencing (e.g., metabolomics, proteomics, plant phenotyping). For this an active participation of journal editors would be necessary to discuss opportunities to change journal publication policies.

Organizer: Dr. Wiktor Jurkowski, Jurkowski Group, Earlham Institute, Norwich Research Park, UK

#### SS14: Oncological Big Data and New Mathematical Tools

Current scientific methods produce various omics data sets covering many cellular functions. However, these data sets are commonly processed separately owing to limited ways in how to connect different omics data together for a meaningful analysis. Moreover, it is currently a problem to integrate such data into mathematical models. We are entering the new era of biological research where the main problem is not to obtain the data but to process and analyze them. In this regard, a strong mathematical approach can be very effective (see J. Gunawardena's essay "Models in biology: accurate descriptions of our pathetic thinking," BMC Biology 2014, 12:29).

In this Special Session we focus on big data (omics and biological pathways) related to oncological research. General biological processes that are relevant to cancer can also be studied. Mathematical tools basically mean statistical learning (data mining, inference, prediction), modeling, and simulation. We want to place a special emphasis on causality. Closely tied to mathematical tools, efficient computational tools can be considered.

Organizers: Dr. Gregorio Rubio, Instituto de Matematica Multidisciplinar, Universitat Politecnica de Valencia, Valencia, Spain. Dr. Rafael Villanueva, Instituto de Matematica Multidisciplinar, Universitat Politecnica de Valencia, Valencia, Spain.

In this edition of IWBBIO, we were honored to have the following invited speakers:

- Prof. Roderic Guigo, Coordinator of Bioinformatics and Genomics at Centre de Regulacio Genomica (CRG). Head of the Computational Biology of RNA Processing Group. Universitat Pompeu Fabra, Barcelona, Spain
- 2. Prof. Joaquin Dopazo, Director of the Computational Genomics Department, Centro de Investigación Príncipe Felipe- CIPF, Valencia, Spain
- Prof. Jose Antonio Lorente, Director of Centre for Genomics and Oncological Research (GENYO). Professor of Legal and Forensic Medicine, University of Granada, Spain

It is important to note, that for the sake of consistency and readability of the book, the presented papers are classified under 16 chapters. The organization of the papers is in two volumes arranged following the topics list included in the call for papers. The first

volume (LNBI 10208), entitled "Advances in Computational Intelligence: Part I" is divided into seven main parts and includes the contributions on:

- 1. Advances in computational intelligence for critical care
- 2. Bioinformatics for health care and diseases
- 3. Biomedical engineering
- 4. Biomedical image analysis
- 5. Biomedical signal analysis
- 6. Biomedicine
- 7. Challenges representing large-scale biological data

The second volume (LNBI 10209), entitled "Advances in Computational Intelligence: Part II" is divided into nine main parts and includes the contributions on:

- 1. Computational genomics
- 2. Computational proteomics
- 3. Computational systems for modelling biological processes
- 4. Data-driven biology: new tools, techniques, and resources
- 5. E-health
- 6. High-throughput bioinformatic tools for genomics
- 7. Oncological big data and new mathematical tools
- 8. Smart Sensor and sensor-network architectures
- 9. Time-lapse experiments and multivariate biostatistics

This fifth edition of IWBBIO was organized by the Universidad de Granada together with the Spanish Chapter of the IEEE Computational Intelligence Society. We wish to thank to our main sponsor and the Faculty of Science, Department of Computer Architecture and Computer Technology and CITIC-UGR, from the University of Granada, for their support and grants. We also wish to thank the Editors-in-Chief of different international journals for their interest in having special issues with the best papers of IWBBIO.

We would also like to express our gratitude to the members of the different committees for their support, collaboration, and good work. We especially thank the local Organizing Committee, Program Committee, the reviewers, and special session organizers. Finally, we want to thank Springer, and especially Alfred Hofmann and Anna Kramer, for their continuous support and cooperation.

April 2017 Ignacio Rojas Francisco Ortuño

# **Organization**

# **Steering Committee**

Federico Moran

Miguel A. Andrade University of Mainz, Germany Hesham H. Ali University of Nebraska, USA

University of Twente, the Netherlands Oresti Baños

Politecnico di Torino, Italy Alfredo Benso Superior School Sant'Anna, Italy Giorgio Buttazzo

Mario Cannataro University Magna Graecia of Catanzaro, Italy

Spanish National Center for Biotechnology (CNB), Spain Jose María Carazo Jose M. Cecilia Universidad Católica San Antonio de Murcia, Spain

M. Gonzalo Claros University of Malaga, Spain

Research Center Principe Felipe, Spain Joaquin Dopazo

Werner Dubitzky University of Ulster, UK

Universidad Católica San Antonio de Murcia, Spain Afshin Fassihi

Jean-Fred Fontaine University of Mainz, Germany

Humberto Gonzalez University of the Basque Country, Spain Concettina Guerra College of Computing, Georgia Tech, USA

Andy Jenkinson Karolinska Institute, Sweden Craig E. Kapfer Reutlingen University, Germany European Bioinformatics Institute, UK Narsis Aftab Kiani Natividad Martinez Reutlingen University, Germany Marco Masseroli Politechnical University of Milan, Italy Complutense University of Madrid, Spain

Cristian R. Munteanu University of A Coruña, Spain New York University, Abu Dhabi Jorge A. Naranjo

Michael Ng Hong Kong Baptist University, SAR China

Jose L. Oliver University of Granada, Spain Juan Antonio Ortega University of Seville, Spain Julio Ortega University of Granada, Spain Aleiandro Pazos University of A Coruña, Spain

Javier Perez Florido Genomics and Bioinformatics Platform of Andalusia,

Spain

Violeta I. Pérez Nueno Inria Nancy Grand Est, LORIA, France

Horacio Pérez-Sánchez Universidad Católica San Antonio de Murcia, Spain

Alberto Policriti University of Udine, Italy Omer F. Rana Cardiff University, UK

M. Francesca Romano Superior School Sant'Anna, Italy VIB - Ghent University, Belgium Yvan Saevs Vicky Schneider The Genome Analysis Centre, UK

Ralf Seepold HTWG Konstanz, Germany

#### XVI Organization

Mohammad Soruri University of Birjand, Iran

Yoshiyuki Suzuki Tokyo Metropolitan Institute of Medical Science, Japan

Oswaldo Trelles University of Malaga, Spain

Renato Umeton CytoSolve Inc., USA

Jan Urban University of South Bohemia, Czech Republic Alfredo Vellido Polytechnical University of Catalonia, Spain

## **Program Committee**

Jesus S. Aguilar Carlos Cano Carlos Alberola Rita Casadio Hisham Al-Mubaid Daniel Castillo Rui Carlos Alves Ting-Fung Chan Nagasuma Chandra Yuan An Georgios Anagnostopoulos Kun-Mao Chao Bernard Chen Eduardo Andrés León Antonia Aránega Bolin Chen Saúl Ares Brian Chen Masanori Arita Chuming Chen Ruben Armañanzas Jie Chen Joel P. Arrais Yuehui Chen Patrizio Arrigo Jianlin Cheng Shuai Cheng I-Jen Chiang

O. Bamidele Awojoyogbe
Shuai Cheng
Jaume Bacardit
Hazem Bahig
Pedro Ballester
Graham Balls
Ugo Bastolla
Sidahmed Benabderrahmane
Steffanny A. Bennett

Shuai Cheng
Jung-Hsien Chiang
Young-Rae Cho
Justin Choi
Darrell Conklin
Clare Coveney
Aedin Culhane

Steffanny A. Bennett
Daniel Berrar
Miguel Damas
Mahua Bhattcharya
Concha Bielza
Armando Blanco
Ignacio Blanquer
Paola Bonizzoni
Aedin Culhane
Miguel Damas
Antoine Danchin
Bhaskar DasGupta
Ricardo De Matos
Guillermo de la Calle
Javier De Las Rivas

Christina Boucher Fei Deng

Hacene Boukari Marie-Dominique Devignes

David Breen Ramón Diaz-Uriarte
Fiona Browne Julie Dickerson
Dongbo Bu Ye Duan
Jeremy Buhler Beatrice Duval
Keith C.C. Saso Dzeroski
Gabriel Caffarena Khaled El-Sayed

Anna Cai Mamdoh Elsheshengy

Christian Exposito Sun Kim

Weixing Feng Kengo Kinoshita Jose Jesús Fernandez Ekaterina Kldiashvili

Gionata Fragomeni Jun Kong Xiaoyong Fu Tomas Koutny Juan Manuel Galvez Natalio Krasnogor Alexandre G.de Brevern Abhay Krishan

Marija Krstic-Demonacos Pugalenthi Ganesan

Jean Gao Sajeesh Kumar Rodolfo Garcia Lukasz Kurgan Mark Gerstein Stephen Kwok-Wing Razvan Ghinea Istvan Ladunga Daniel Gonzalez Peña T.W. Lam Dianjing Guo Pedro Larrañaga Jun-tao Guo Dominique Lavenier Maozu Guo Jose Luis Lavin Christophe Guyeux Doheon Lee Michael Hackenberg Xiujuan Lei Michiaki Hamada André Leier

Xiyi Hang Kwong-Sak Leung

Jin-Kao Hao Chen Li Nurit Haspel Dingcheng Li Morihiro Hayashida Jing Li Jieyue He Jinyan Li Min Li Pheng-Ann Heng Luis Javier Herrera Xiaoli Li Pietro Hiram Yanpeng Li Lynette Hirschman Li Liao Ralf Hofestadt Hongfei Lin Vasant Honavar Hongfang Liu Wen-Lian Hsu Jinze Liu Jun Hu Xiaowen Liu Xiaohua Hu Xiong Liu Jun Huan Zhenqiu Liu Chun-Hsi Huang Zhi-Ping Liu

Heng Huang Rémi Longuespée Jimmy Huang Miguel Angel Lopez Gordo

Jingshan Huang Ernesto Lowy Seiya Imoto Jose Luis

Anthony J. Kusalik Suryani Lukman Yanqing Ji Feng luo Xingpeng Jiang Bernard M.E. Mingon Kang Bin Ma Dong-Chul Kim Oin Ma Dongsup Kim Malika Mahoui

Hyunsoo Kim Alberto Maria

#### XVIII Organization

Tatiana Marquez-Lago

Keith Marsolo

Francisco Martinez Alvarez

Osamu Maruyama Tatiana Maximova Roderik Melnik

Jordi Mestres Hussain Michelle Ananda Mondal Antonio Morreale

Walter N. Moss Maurice Mulvenna Enrique Muro

Radhakrishnan Nagarajan

Vijayaraj Nagarajan Kenta Nakai

Isabel A. Nepomuceno Mohammad Nezami

Anja Nohe
Michael Ochs
Baldomero Oliva
Jose Luis Oliveira
Motonori Ota
David P.

Tun-Wen Pai Paolo Paradisi

Hyun-Seok Park Kunsoo Park Taesung Park

David Pelta Alexandre Perera

María Del Mar Pérez Gómez Vinhthuy Phan

Antonio Pinti
Héctor Pomares
Mihail Popescu
Benjarath Pupacdi
Sanguthevar Rajasekaran
Shoba Ranganathan

Jairo Rocha Fernando Rojas Jianhua Ruan Gregorio Rubio

Antonio Rueda Irena Rusu

Vincent Shin-Mu Tseng

Michael Sadovsky Belen San Roman

Kunihiko Sadakane

Maria Jose Saez Hiroto Saigo José Salavert

Carla Sancho Mestre Emmanuel Sapin Kengo Sato

Jean-Marc Schwartz Russell Schwartz Jose Antonio Seoane

Xuequn Shang

Piramanayagam Shanmughavel

Xiaoman Shawn Xinghua Shi Tetsuo Shibuya Dong-Guk Shin Amandeep Sidhu Istvan Simon Richard Sinnott Jiangning Song

Joe Song Zhengchang Su Joakim Sundnes Wing-Kin Sung Prashanth Surayajhala

Martin Swain

Sing-Hoi Sze
Mehmet Tan
Xing Tan
Li Teng
Tianhai Tian
Pedro Tomas
Carlos Toro
Carolina Torres
Paolo Trunfio
Esko Ukkonen
Lucia Vaira
Paola Velardi

Julio Vera Konstantinos Votis Slobodan Vucetic Ying-Wooi Wan Chong Wang

Haiying Wang

Jason Wang Jingkai Yu Jian Wang Hong Yue Erliang Zeng Jianxin Wang Jiayin Wang Xue-Qiang Zeng Junbai Wang Aidong Zhang Junwen Wang Chi Zhang Lipo Wang Jin Zhang Li-San Wang Jingfen Zhang Kaizhong Zhang Lusheng Wang Yadong Wang Shao-Wu Zhang Yong Wang Xingan Zhang Ka-Chun Wong Zhongming Zhao Ling-Yun Wu Huiru Zheng Xintao Wu Bin Zhou Zhonghang Xia Shuigeng Zhou Xuezhong Zhou Fang Xiang Lei Xu Daming Zhu Zhong Xue Dongxiao Zhu Patrick Xuechun Shanfeng Zhu Hui Yang Xiaoqin Zou Xiufen Zou Zhihao Yang

#### **Additional Reviewers**

Alquezar, Rene Moulin, Serge Amaya Vazquez, Ivan Olier, Iván

Belanche, Lluís Ortega-Martorell, Sandra

Browne, Fiona Ribas, Vicent

Calabrese, Barbara Scherz, Wilhelm Daniel

Cho, Ken Sebastiani, Laura
Coelho, Edgar Seong, Giun
Datko, Patrick Smith, Peter
Gaiduk, Maksym Treepong, Panisa
Gonzalez-Abril, Luis Tu, Shikui
Li Guomin Valenzuela Olga

Gonzalez-Abril, Luis
Tu, Shikui
Valenzuela, Olga
Kang, Hong
Vizza, Patrizia
Langa, Jorge
Wu, Yonghui
Liang, Zhewei
Yan, Shankai

Martin-Guerrero, Jose D. Zucco, Chiara

# **Contents – Part I**

Advances in Computational Intelligence for Critical Care	
Health Monitoring System Based on Parallel-APPROX SVM	3
Machine Learning for Critical Care: An Overview and a Sepsis Case Study Alfredo Vellido, Vicent Ribas, Carles Morales, Adolfo Ruiz Sanmartín, and Juan Carlos Ruiz-Rodríguez	15
Bioinformatics for Healthcare and Diseases	
A Meta-Review of Feature Selection Techniques in the Context of Microarray Data	33
Immune Network Technology on the Basis of Random Forest Algorithm for Computer-Aided Drug Design	50
H-RACER: Hybrid RACER to Correct Substitution, Insertion, and Deletion Errors	62
A Multi-sensor Approach for Biomimetic Control of a Robotic  Prosthetic Hand	74
Feature Selection in Multiple Linear Regression Problems with Fewer Samples Than Features	85
Colormetric Experiments on Aquatic Organisms	96
Predicting Comprehensive Drug-Drug Interactions for New Drugs via Triple Matrix Factorization	108

Diagnosis of Auditory Pathologies with Hidden Markov Models  Lilia Lazli, Mounir Boukadoum, Mohamed-Tayeb Laskri, and Otmane Ait-Mohamed	118
Biclustering Based on Collinear Patterns	134
A Data- and Model-Driven Analysis Reveals the Multi-omic Landscape of Ageing	145
Experimental Investigation of Frequency Chaos Game Representation for in Silico and Accurate Classification of Viral Pathogens from Genomic Sequences	155
Gamified Mobile Blood Donation Applications	165
Medical Entity Recognition and Negation Extraction: Assessment of NegEx on Health Records in Spanish	177
RISK: A Random Optimization Interactive System Based on Kernel Learning for Predicting Breast Cancer Disease Progression	189
Revealing the Relationship Between Human Genome Regions and Pathological Phenotypes Through Network Analysis	197
Variety Behavior in the Piece-Wise Linear Model of the p53-Regulatory Module	208
Multi-omic Data Integration Elucidates Synechococcus Adaptation  Mechanisms to Fluctuations in Light Intensity and Salinity	220
MASS Studio: A Novel Software Utility to Simplify LC-MS Analyses of Large Sets of Samples for Metabolomics	230

# **Biomedical Engineering** Adsorption of Bilirubin Toxin in Liver by Chitosan Coated Activated 247 Asel Mwafy, Ameereh Seyedzadeh, Waleed Khalil Ahmed, Basel Alsayyed Ahmad, Betty Mathew, Kamala Pandurangan, Abdel-Hamid Ismail Mourad, and Ali Hilal-Alnagbi Investigation of the Feasibility of Strain Gages as Pressure Sensors 261 Him Wai Ng, Xianta Jiang, Lukas-Karim Merhi, and Carlo Menon Geometric Modelling of the Human Cornea: A New Approach for the Study of Corneal Ectatic Disease. A Pilot Investigation . . . . . . . . . 271 Francisco Cavas-Martínez, Daniel G. Fernández-Pacheco, Dolores Parras, Francisco J.F. Cañavate, Laurent Bataille, and Jorge L. Alio Virtual Surgical Planning for Mandibular Reconstruction: Improving the Fibula Bone Flap...... 282 Dolores Parras, Benito Ramos, Juan José Haro, Manuel Acosta, Francisco Cavas-Martínez, Francisco J.F. Cañavate. and Daniel G. Fernández-Pacheco Elbow Orthosis for Tremor Suppression – A Torque Based Input Case . . . . . 292 Gil Herrnstadt and Carlo Menon Gremlin Language for Querying the BiographDB Integrated Biological Database..... 303 Antonino Fiannaca, Laura La Paglia, Massimo La Rosa, Antonio Messina, Riccardo Rizzo, Dario Stabile, and Alfonso Urso Modelling of Glucose Dynamics for Diabetes....... 314 Tomas Koutny Medical Planning: Operating Theatre Design and Its Impact on Cost, Area and Workflow..... 325 Khaled S. Ahmed **Biomedical Image Analysis** Recognition of Stages in the Belousov-Zhabotinsky Reaction Using Information Entropy: Implications to Cell Biology..... 335 Anna Zhyrova, Renata Rychtáriková, and Dalibor Štys Back Pain During Pregnancy and Its Relationship to Anthropometric 347 Julien A. Leboucher, Antonio Pinti, and Geneviève A. Dumas

Pavla Urbanová, Jan Vaněk, Pavel Souček, Dalibor Štys, Petr Císař, and Miloš Železný	338
Maximal Oxygen Consumption and Composite Indices of Femoral Neck Strength in a Group of Young Women	369
Microaneurysm Candidate Extraction Methodology in Retinal Images for the Integration into Classification-Based Detection Systems  Estefanía Cortés-Ancos, Manuel Emilio Gegúndez-Arias, and Diego Marin	376
A Bio-inspired Algorithm for the Quantitative Analysis of Hind Limb Locomotion Kinematics of Laboratory Rats	385
Quantization and Equalization of Pseudocolor Images in Hand Thermography	397
Superficial Dorsal Hand Vein Estimation	408
Automatic Removal of Mechanical Fixations from CT Imagery with Particle Swarm Optimisation	419
Augmented Visualization as Surgical Support in the Treatment of Tumors Lucio Tommaso De Paolis	432
Salient Networks: A Novel Application to Study Brain Connectivity Nicola Amoroso, Roberto Bellotti, Domenico Diacono, Marianna La Rocca, and Sabina Tangaro	444
Clustering of Food Intake Images into Food and Non-food Categories  Abul Doulah and Edward Sazonov	454
Feature Extraction Using Deep Learning for Food Type Recognition  Muhammad Farooq and Edward Sazonov	464
Biomedical Signal Analysis	
Degrees of Freedom in a Vocal Fold Inverse Problem	475

Contents – Part I	XXV
Information Limits of Optical Microscopy: Application to Fluorescent Labelled Tissue Section	485
Hands-Free EEG-Based Control of a Computer Interface Based on Online Detection of Clenching of Jaw	497
Wavelet Decomposition Based Automatic Sleep Stage Classification Using EEG	508
An Impact of Severe Preeclampsia on Cardiovascular System Adaptation of Newborns in Early Neonatal Period	517
Classification Algorithms for Fetal QRS Extraction in Abdominal ECG Signals	524
Evaluation of Algorithms for Automatic Classification of Heart Sound Signals	536
Automatic Glissade Determination Through a Mathematical Model in Electrooculographic Records	546
Evaluation of the Differentiation of Noisy Electrooculographic Records Using Continuous Wavelet Transform	557
Biomedicine	
Scores of Intestinal Fibrosis from Wavelet-Based Magnetic Resonance Imaging Models	569

# XXVI Contents - Part I

L <sub>1</sub> -regularization Model Enriched with Biological Knowledge	579
Secret Life of Tiny Blood Vessels: Lactate, Scaffold and Beyond Vladimir Salmin, Andrey Morgun, Elena Khilazheva, Natalia Pisareva, Elizaveta Boitsova, Pavel Lavrentiev, Michael Sadovsky, and Alla Salmina	591
Increasing of Data Security and Workflow Optimization in Information and Management System for Laboratory	602
Challenges Representing Large-Scale Biological Data	
GIS-Aided Modelling of Two Siberian Reservation Sites	617
What Can the Big Data Eco-System and Data Analytics Do for E-Health?  A Smooth Review Study	629
On the Ability to Reconstruct Ancestral Genomes from Mycobacterium Genus	642
Representativeness of a Set of Metabolic Pathways	659
Author Index	669

# **Contents – Part II**

# **Computational Genomics**

Nanopore Structures	3
Uropathogenic Escherichia coli: An Ideal Resource for DNA Microarray Probe Designing	12
Lost Strings in Genomes: What Sense Do They Make?	20
Comprehensive Study of Instable Regions in <i>Pseudomonas Aeruginosa</i> Dan Wang, Jingyu Li, and Lusheng Wang	30
Breathogenomics: A Computational Architecture for Screening, Early Diagnosis and Genotyping of Lung Cancer	41
Mixed-Integer Programming Model for Profiling Disease Biomarkers from Gene Expression Studies	50
Assembly of Gene Expression Networks Based on a Breast Cancer Signature	62
Pairwise and Incremental Multi-stage Alignment of Metagenomes:  A New Proposal	74
An Expanded Association Approach for Rare Germline Variants with Copy-Number Alternation	81
Parallelizing Partial Digest Problem on Multicore System	95

# **Computational Proteomics**

Prediction of Calmodulin-Binding Proteins Using Short-Linear Motifs  Yixun Li, Mina Maleki, Nicholas J. Carruthers, Luis Rueda, Paul M. Stemmer, and Alioune Ngom	107
Data Mining the Protein Data Bank to Identify and Characterise Chameleon Coil Sequences that Form Symmetric Homodimer β-Sheet Interfaces Johanna Laibe, Melanie Broutin, Aaron Caffrey, Barbara Pierscionek, and Jean-Christophe Nebel	118
Associating Gene Ontology Terms with Pfam Protein Domains	127
3D Protein-Structure-Oriented Discovery of Clinical Relation Across Chronic Lymphocytic Leukemia Patients	139
Application of a Membrane Protein Structure Prediction Web Service GPCRM to a Gastric Inhibitory Polypeptide Receptor Model	151
In Silico Prediction of 3D Structure of Anopheles Gambiae  ABCC12 Protein	163
Exploring Symmetric Substructures in Protein Interaction Networks for Pairwise Alignment	173
Identification and in silico Analysis of Glutathione Reductase Transcripts  Expressed in Olive (Olea europaea L.) Pollen and Pistil	185
Computational Systems for Modelling Biological Processes	
Cyber Immunity: A Bio-Inspired Cyber Defense System	199

Management of Data Structures Generated During Simulations

Andreea Robu, Mihaela Crisan-Vida, Nicolae Robu, and Adrian Neagu

325

Finding Transcripts Associated with Prostate Cancer Gleason Stages Using Next Generation Sequencing and Machine Learning Techniques	337
Bioinformatics from a Big Data Perspective: Meeting the Challenge Francisco Gomez-Vela, Aurelio López, José A. Lagares, Domingo S. Baena, Carlos D. Barranco, Miguel García-Torres, and Federico Divina	349
Computational Prediction of Host-Pathogen Interactions Through Omics  Data Analysis and Machine Learning	360
Simultaneous Gene Selection and Weighting in Nearest Neighbor Classifier for Gene Expression Data	372
Machine Learning Approaches for Predicting High Utilizers in Health Care	382
eHealth	
A Device Supporting the Self Management of Tinnitus	399
A Case Study on the Integration of Heterogeneous Data Sources in Public Health	411
QSAR Classification Models for Predicting Affinity to Blood or Liver of Volatile Organic Compounds in e-Health	424
Usage of VR Headsets for Rehabilitation Exergames	434

High-Throughput Bioinformatic Tools for Genomics	
Search of Regions with Periodicity Using Random Position Weight Matrices in the Genome of <i>C. elegans</i>	445
RNA Sequencing Analysis of Neural Cell Lines: Impact of Normalization and Technical Replication	457
Towards a Universal Genomic Positioning System: Phylogenetics and Species IDentification	469
A High Performance Storage Appliance for Genomic Data	480
Obtaining the Most Accurate de novo Transcriptomes for Non-model Organisms: The Case of Castanea sativa	489
Accelerating Smith-Waterman Alignment of Long DNA Sequences with OpenCL on FPGA	500
Smith-Waterman Acceleration in Multi-GPUs:  A Performance per Watt Analysis	512
A Deep Learning Network for Exploiting Positional Information in Nucleosome Related Sequences	524
Oncological Big Data and New Mathematical Tools	
A Clinical Tool for Automated Flow Cytometry Based on Machine Learning Methods	537

Prognostic Modeling and Analysis of Tumor Response to Fractionated Radiotherapy for Patients with Squamous Cell Lung Cancer	549
Uncertainty Quantification for Meningococcus B Carriers Prediction Luis Acedo, Clara Burgos, Juan-Carlos Cortés, and Rafael J. Villanueva	560
Smart Sensor and Sensor-Network Architectures	
A Novel Wearable for Rehabilitation Using Infrared Sensors:  A Preliminary Investigation	573
Real Time Localization Using Bluetooth Low Energy	584
A Sensor Grid for Pressure and Movement Detection Supporting Sleep Phase Analysis	596
A Portable Wireless sEMG and Inertial Acquisition System for Human Activity Monitoring	608
An Automatic and Intelligent System for Integrated Healthcare Processes Management	621
A Microcontroller Based System for Controlling Patient Respiratory Guidelines	631
Requirements Analysis for User Interfaces in Mobile eHealth Applications Armando Statti and Natividad Martinez Madrid	642
Mobile Health System for Evaluation of Breast Cancer Patients During Treatment and Recovery Phases	653

Contents – Part II	XXXIII
Key Factors for Innovative Developments on Health Sensor-Based System	665
Time Lapse Experiments and Multivariate Biostatistics	
Towards Integration of CFD and Photosynthetic Reaction Kinetics in Modeling of Microalgae Culture Systems	679
Kinetic Modelling of Processes Behind S <sub>2,3</sub> -states Deactivation in Photosynthetic Oxygen Evolution	691
Observation of Dynamics Inside an Unlabeled Live Cell Using a Bright-Field Photon Microscopy: Evaluation of Organelles' Trajectories	700
Automatic Multiparameter Acquisition in Aquaponics Systems	712
High Definition Method for Imaging Bacteria in Microconfined Environments on Solid Media	726
Erratum to: Bioinformatics and Biomedical Engineering	E1