

Lecture Notes in Bioinformatics

6620

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

Ana T. Freitas Arcadi Navarro (Eds.)

Bioinformatics for Personalized Medicine

10th Spanish Symposium, JBI 2010
Torremolinos, Spain, October 27-29, 2010
Revised Selected Papers

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

Ana T. Freitas
INESC-ID/Instituto Superior Técnico
R. Alves Redol 9
1000-029 Lisboa, Portugal
E-mail: atf@kdbio.inesc-id.pt

Arcadi Navarro
Institut de Biologia Evolutiva (UPF-CSIC)
Doctor Aiguader 88
08003 Barcelona, Spain
E-mail: arcadi.navarro@upf.edu

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-28061-0 e-ISBN 978-3-642-28062-7
DOI 10.1007/978-3-642-28062-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011945508

CR Subject Classification (1998): H.3, H.2.8, F.2.1, H.4, C.2, H.5, D.2

LNCS Sublibrary: SL 8 – Bioinformatics

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

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

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains the papers presented at JBI 2010: the 10th Spanish Symposium on Bioinformatics held in Torremolinos, Málaga, Spain, during October 27–29, 2010. After 10 years of conferences (the JBI conference series was started in 2000) JBI 2010 centered on the topic “Bioinformatics for Personalized Medicine” and was hosted by the University of Málaga, co-organized by the National Institute of Bioinformatics, Spain and the Portuguese Bioinformatics Network chaired by Oswaldo Trelles, Arcadi Navarro, Federico Morán and Pedro Fernandes and took place in the Hotel Son Don Pablo, at Paseo Marítimo, Torremolinos.

This year, 18 papers were accepted for presentation out of more than 100 submissions. The papers presented were selected by the Program Committee (PC) assisted by a number of external reviewers. Each paper was reviewed by two members of the PC, or by external reviewers. JBI 2010 also introduced a Highlights Track, in which eight additional presentations were chosen from journal papers published in 2009 or 2010. This made a total of 26 presentations. Authors were invited to submit a revised version of their work to be considered for publication in *Lecture Notes in Bioinformatics* after a second reviewing process by independent reviewers. This volume contains the papers that were finally selected.

JBI 2010 invited several distinguished speakers as keynotes and for special sessions on Genomic Projects Design: Success Cases, Bioinformatics Challenges in Personalized Medicine, Bioinformatics Software Developments in Spain and Beyond, and Options and Priorities in Hands-On Training in Spain and Portugal. A special session to present the status of the ELIXIR ESFRI project also took place. Invited speakers and special sessions organizers included Ivo Gut (National Genome Analysis Center Barcelona), Sara Hunter (EMBL-EBI Hinxton), Julio Saez Rodriguez (EMBL Outstation - Hinxton, European Bioinformatics Institute), Victoriano Valpuesta Fernández (University of Málaga), Alfonso Valencia (Centro Nacional de Investigaciones Oncológicas), José Pereira Leal (Instituto Gulbenkian de Ciência), Fernando Martín-Sánchez (Instituto de Salud Carlos III), Pedro Fernandes (Instituto Gulbenkian de Ciência), Jose María Fernández (Centro Nacional de Investigaciones Oncológicas), Jose Ramon Valverde (Centro Nacional de Biotecnología), Victoria Martín-Requena (Universidad de Málaga), Maximiliano García (Universidad de Málaga) and Allan Orozco (Centro Nacional de Investigaciones Oncológicas).

JBI 2010 was only possible through the dedication and hard work of many individuals and organizations. Special thanks go to the PC and external reviewers for helping to improve the quality of the conference program, and the Organizing Committee, namely its Chair Oswaldo Trelles, for hosting the conference and providing the administrative, logistic, and financial support. We also

thank our sponsors, including the INB (Instituto Nacional de Bioinformática), Bitlab (Bioinformatics and Information Technology Laboratory), Universidad de Málaga, Instituto de Salud Carlos III, Portuguese Bioinformatics Network, Life Sequencing, Integromics, UNIA, SEBIOT and Combiomed. Without them the conference would not have been financially viable. Finally, we thank all the authors who contributed papers and posters, as well as the attendees of the conference for their enthusiastic participation.

May 2011

Anna Teresa Freitas
Arcadi Navarro

Conference Organization

Honorary Committee

José Antonio Griñán Martínez (President of Andalucía)

Francisco de la Torre (Mayor of Málaga, Spain)

Pedro Fernández Montes (Mayor of Torremolinos, Spain)

Adelaida de la Calle Martín (Chancellor, University of Malaga, Spain)

Scientific Committee

Chairs

Alfonso Valencia

Centro Nacional de Investigaciones Oncológicas

Ana Teresa Freitas

INESC-ID/Instituto Superior Técnico

Committee

M. Mar Albà

Universitat Pompeu Fabra

Patrick Aloy

Institute for Research in Biomedicine

Jose Antonio Álvarez

Universidad de Santiago de Compostela

Julio R. Banga

Instituto de investigaciones marinas de Vigo

Ugo Bastolla

Instituto de Investigaciones Biomédicas

Pedro Bernaola Galván

Universidad de Málaga

Miguel Blanca

Hospital Carlos Haya

Christian Blaschke

BioAlma

José Luis Caballero

Universidad de Córdoba

José María Carazo

Centro Nacional de Biotecnología

Pedro Carpena

Universidad de Málaga

João Carriço

Universidade de Lisboa

Robert Castelo

Universitat Pompeu Fabra

Pablo Chacón

Centro de Investigaciones Biológicas

Julio Collados Vides

Universidad Nacional Autónoma de México

Ana Conesa

Centro de Investigaciones Príncipe Felipe

Francisco Couto

Universidade de Lisboa

Javier De Las Rivas

Universidad de Salamanca

Ramon Diaz-Uriarte

Spanish National Cancer Centre

Joaquin Dopazo

Centro de Investigación Príncipe Felipe

Hernán J. Dopazo

Centro de Investigación Príncipe Felipe

Julián Dorado	Facultad de Humanidades Universidade da Coruña
Eduardo Eyras	ICREA and Universitat Pompeu Fabra
André Falcão	Universidade de Lisboa
Toni Gabaldon	Centre for Genomic Regulation
Gian Gaetano Tartaglia	Center for Genomic Regulation
Juan Antonio García Ranea	Universidad de Málaga
Francesco Gervasio	Spanish National Cancer Research Centre
Manuel J. Gómez	Centro de Astrobiología INTA-CSIC
M. Gonzalo Claros	Universidad de Málaga
Antonio Granell	Instituto de Biología Molecular y Celular de Plantas de Valencia
Hugo Gutierrez de Teran	Fundación Pública Galega de Medicina Xenómica
Michael Hackenberg	Universidad de Granada
Fyodor Kondrashov	Center for Genomic Regulation
Josep Lluís Gelpi	Barcelona Supercomputing Center
Paulino Gómez-Puertas	Centro de Biología Molecular “Severo Ochoa”
Nuria Lopez-Bigas	Universitat Pompeu Fabra
Juan Fernandez-Recio	Barcelona Supercomputing Center
Antonio Figueras	Instituto de investigaciones Marinas
Roderic Guigó	Center for Genomic Regulation
Roberto Marabini	Universidad Autónoma de Madrid
Antonio Marin	Universidad de Sevilla
Miguel Angel Medina Torres	University of Malaga
Jordi Mestres	Instituto Municipal de Investigación Médica
Francisco Montero	Universidad Complutense de Madrid
Antonio Morreale	Universidad Autónoma de Madrid
Andrés Moya	University of Valencia
Juan Muñoz	Universidad de Cordoba
Arcadi Navarro	Universitat Pompeu Fabra
Cedric Notredame	Centre de Regulacio Genomica
José L. Oliver	Universidad de Granada
José Luís Oliveira	Universidade de Aveiro
Arlindo Oliveira	INESC-ID/Instituto Superior Técnico
Modesto Orozco	Barcelona Supercomputing Center
Alejandro Pazos	Universidad de A Coruña
José Pereira-Leal	Instituto Gulbenkian de Ciência
Alberto Pascual Montano	National Center for Biotechnology
Florencio Pazos	National Center for Biotechnology
Antonio Pérez Pulido	Universidad Pablo de Olávide
Francisco Pinto	Universidade de Lisboa

David G. Pisano	Spanish National Cancer Research Centre
David Posada	Universidad de Vigo
Raul Rabadan	Columbia University
Pedro Reche	Universidad Complutense de Madrid
Cristian Robert Munteanu	Universidade da Coruña
Isabel Rocha	Universidade do Minho
Ana M. Rojas	Institute for Predictive and Personalized Medicine of Cancer
Armando Salvador	Universidade de Coimbra
Kika Sánchez	University of Malaga
Javier Sancho	Universidad de Zaragoza
Ferran Sanz	Institut Municipal d'Investigació Mèdica
Cláudio M. Soares	Instituto de Tecnologia Química e Biológica
Javier Tamames	National Center for Biotechnology
Oswaldo Trelles	Universidad de Málaga
Ana Tereza Vasconcelos	Universidad Federal do Río de Janeiro
Enrique Viguera-Mínguez	Universidad de Málaga
ordi Villà-Freixa	Universitat Pompeu Fabra

Organizing Committee

Chairs

Pedro Fernandes	Instituto Gulbenkian de Ciência
Federico Morán	Universidad Complutense Madrid
Arcadi Navarro	Universitat Pompeu Fabra
Oswaldo Trelles	Universidad de Málaga

Committee

Maximiliano García	Universidad de Málaga
Johan Karlsson	Universidad de Málaga
Victoria Martín-Requena	Universidad de Málaga
Alfredo Martínez-López	Universidad de Málaga
José Manuel Mateos	Universidad de Málaga
Antonio Muñoz-Mérida	Universidad de Málaga
Javier Ríos Pérez	Universidad de Málaga
Manuel Gonzalo Claros	Universidad de Málaga
Juan Falgueras Cano	Universidad de Málaga
Enrique Viguera Mínguez	Universidad de Málaga
Allan Orozco	Centro Nacional de Investigaciones Oncológicas

JBI Past Editions

Edition	Date	Organizing Committee
I	June 14-16, 2000	Universidad Politécnica de Cartagena
II	30 May – 1 June, 2001	Universidad de Málaga
III	September 18-20, 2002	Centro de Investigación del Cancer, Universidad de Salamanca
IV	September 13-15, 2003	Facultad de Informática de la Universidad de Coruña
V	29 November- 2 December, 2004	Universitat Politecnica de Catalunya
VI	28 September-2 October, 2005	Centro Nacional de Investigaciones Oncológicas
VII	November 20-22, 2006	Universidad de Zaragoza
VIII	February 13-15, 2008	Centro de Investigaciones Príncipe Felipe
IX	November 2-6, 2009	Instituto Gulbenkian de Lisboa

Table of Contents

Statistical Significance for NGS Reads Similarities	1
<i>Antonio Muñoz-Mérida, Javier Ríos, Hicham Benzekri, and Oswaldo Trelles</i>	
Application of Array-Oriented Scientific Data Formats (NetCDF) to Genotype Data, GWASpi as an Example	8
<i>Fernando Muñoz Fernandez, Angel Carreño Torres, Carlos Morcillo-Suarez, and Arcadi Navarro</i>	
An Ontology and a REST API for Sequence Based Microbial Typing Data	21
<i>João Almeida, João Tiple, Mário Ramirez, José Melo-Cristino, Cátia Vaz, Alexandre P. Francisco, and João A. Carriço</i>	
Multivariate Methods for the Integration and Visualization of Omics Data	29
<i>Alex Sánchez, José Fernández-Real, Esteban Vegas, Francesc Carmona, Jacques Amar, Remy Burcelin, Matteo Serino, Francisco Tinahones, M. Carmen Ruíz de Villa, Antonio Minárrro, and Ferran Reverter</i>	
A Holistic Approach for Integrating Genomic Variation Information	42
<i>Pedro Lopes and José Luís Oliveira</i>	
SB-KOM: Integration of Pathway Information with BioPax	50
<i>María Jesús García-Godoy, Ismael Navas-Delgado, and José F. Aldana-Montes</i>	
GENote v.β: A Web Tool Prototype for Annotation of Unfinished Sequences in Non-model Eukaryotes	66
<i>Noé Fernández-Pozo, Darío Guerrero-Fernández, Rocío Bautista, Josefa Gómez-Maldonado, Concepción Avila, Francisco M. Cánovas, and M. Gonzalo Claros</i>	
Interface Similarity Improves Comparison of DNA-Binding Proteins: The Homeobox Example	72
<i>Álvaro Sebastián, Carlos P. Cantalapiedra, and Bruno Contreras-Moreira</i>	
Pyicos: A Flexible Tool Library for Analyzing Protein-Nucleotide Interactions with Mapped Reads from Deep Sequencing	83
<i>Juan Gonzalez-Vallinas, Sonja Althammer, and Eduardo Eyras</i>	

Novel Angiogenic Functional Targets Predicted through “ <i>Dark Matter</i> ” Assessment in Protein Networks	89
<i>Ian Morilla, Miguel A. Medina, and Juan A.G. Ranea</i>	
jORCA and Magallanes Sailing Together towards Integration of Web Services.....	94
<i>Johan Karlsson and Oswaldo Trelles</i>	
iHOP Web Services Family	102
<i>José M. Fernández, Robert Hoffmann, and Alfonso Valencia</i>	
Bioinformatic Software Developments in Spain	108
<i>José M. Fernández and Alfonso Valencia</i>	
Author Index	121