

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>

Joao Meidanis · Luay Nakhleh (Eds.)

Comparative Genomics

15th International Workshop, RECOMB CG 2017
Barcelona, Spain, October 4–6, 2017
Proceedings

Editors
Joao Meidanis
University of Campinas
Campinas, São Paulo
Brazil

Luay Nakhleh
Rice University
Houston, TX
USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Bioinformatics
ISBN 978-3-319-67978-5 ISBN 978-3-319-67979-2 (eBook)
DOI 10.1007/978-3-319-67979-2

Library of Congress Control Number: 2017953424

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

This volume contains the papers presented at RECOMBCG-17: the 15th RECOMB Comparative Genomics Satellite Workshop, held during October 4–6, 2017, in Barcelona, Spain.

There were 32 submissions. Each submission was reviewed by at least three reviewers, and some received up to five reviews. The reviews were conducted by members of the Program Committee as well as by additional reviewers who were sought based on their expertise for specific papers. The committee decided to accept 16 papers. The program also included four invited talks and a poster session.

We thank the members of the Program Committee as well as the additional reviewers for their diligent work reviewing the manuscripts and conducting thorough discussions on the manuscripts and their reviews that informed the decision process. We also thank the members of the Steering Committee, Jens Lagergren, Aoife McLysaght, and David Sankoff, for their guidance. Special thanks go to the local organizers, Laura Prat Busquets, Tomàs Marquès-Bonet, and Roderic Guigó Serra. Last but not least, we would like to thank the four keynote speakers who agreed to speak at the conference despite their very busy schedules: Iria Fernández Silva, Toni Gabaldón, Iñaki Ruiz-Trillo, and Wes Warren.

The workshop would not have been possible without the generous contribution of our sponsors, whose support is greatly appreciated:

- FEDER - Fondo Europeo de Desarrollo Regional
- Gobierno de España - Ministerio de Economía, Industria y Competitividad
- Instituto de Salud Carlos III
- PRB2 - Plataforma de Recursos Biomoleculares y Bioinformáticos
- ELIXIR Spain
- INB - Instituto Nacional de Bioinformática
- Illumina
- The Company of Biologists

We used the EasyChair system for submissions, reviews, and proceedings formatting.

August 2017

Joao Meidanis
Luay Nakhleh

Organization

Program Committee

Max Alekseyev	George Washington University, USA
Lars Arvestad	Stockholm University, Sweden
Anne Bergeron	Université du Québec a Montreal, Canada
Marília Braga	Bielefeld University, Germany
Cedric Chauve	Simon Fraser University, Canada
Leonid Chindelevitch	Simon Fraser University, Canada
Miklos Csuros	Université de Montreal, Canada
Ingo Ebersberger	Goethe University, Germany
Nadia El-Mabrouk	University of Montreal, Canada
Guillaume Fertin	University of Nantes, France
Katharina Jahn	ETH Zurich, Switzerland
Asif Javed	GIS A-STAR, Singapore
Kevin Liu	Michigan State University, USA
Ketil Malde	Institute of Marine Research
Joao Meidanis	University of Campinas/Scylla Bioinformatics, Brazil
István Miklós	Rnyi Institute, Budapest, Hungary
Siavash Mirarab	University of Texas at Austin, USA
Luay Nakhleh	Rice University, USA
Aida Ouangraoua	Université de Sherbrooke, Canada
Marie-France Sagot	Inria, France
Michael Sammeth	Centre d'Anàlisi Genòmica (CNAG)
Jens Stoye	Bielefeld University, Germany
Krister Swenson	CNRS, Université de Montpellier, France
Eric Tannier	Inria, France
Glenn Tesler	University of California, San Diego, USA
Tamir Tuller	Tel Aviv University, Israel
Lusheng Wang	City University of Hong Kong, SAR China
Tandy Warnow	The University of Illinois at Urbana-Champaign, USA
Louxin Zhang	National University of Singapore, Singapore
Jie Zheng	Nanyang Technological University, Singapore

Additional Reviewers

Aganezov, Sergey	Durand, Dannie	Sinaimeri, Blerina
Alexeev, Nikita	Fransiskus Xavier, Ivan	Viduani Martinez,
Avdeyev, Pavel	Labarre, Anthony	Fábio Henrique
Bansal, Mukul S.	Moret, Bernard	Willing, Eyla
Bulteau, Laurent	Peterlongo, Pierre	Wittler, Roland
Coelho, Vítor	Sankoff, David	Zhou, Xinrui
Doerr, Daniel	Schulz, Tizian	

Contents

The Similarity Distribution of Paralogous Gene Pairs Created by Recurrent Alternation of Polyploidization and Fractionation	1
<i>Yue Zhang and David Sankoff</i>	
A Tractable Variant of the Single Cut or Join Distance with Duplicated Genes.	14
<i>Pedro Feijão, Aniket Mane, and Cedric Chauve</i>	
Generation and Comparative Genomics of Synthetic Dengue Viruses	31
<i>Eli Goz, Yael Tsalenchuck, Rony Oren Benaroya, Shimshi Atar, Tahel Altman, Justin Julander, and Tamir Tuller</i>	
ASTRAL-III: Increased Scalability and Impacts of Contracting Low Support Branches.	53
<i>Chao Zhang, Erfan Sayyari, and Siavash Mirarab</i>	
Algorithms for Computing the Family-Free Genomic Similarity Under DCJ.	76
<i>Diego P. Rubert, Gabriel L. Medeiros, Edna A. Hoshino, Marília D.V. Braga, Jens Stoye, and Fábio V. Martinez</i>	
New Algorithms for the Genomic Duplication Problem	101
<i>Jarosław Paszek and Paweł Górecki</i>	
TreeShrink: Efficient Detection of Outlier Tree Leaves	116
<i>Uyen Mai and Siavash Mirarab</i>	
Rearrangement Scenarios Guided by Chromatin Structure	141
<i>Sylvain Pulicani, Pijus Simonaitis, Eric Rivals, and Krister M. Swenson</i>	
A Unified ILP Framework for Genome Median, Halving, and Aliquoting Problems Under DCJ.	156
<i>Pavel Avdeyev, Nikita Alexeev, Yongwu Rong, and Max A. Alekseyev</i>	
Orientation of Ordered Scaffolds.	179
<i>Sergey Aganezov and Max A. Alekseyev</i>	
Finding Teams in Graphs and Its Application to Spatial Gene Cluster Discovery	197
<i>Tizian Schulz, Jens Stoye, and Daniel Doerr</i>	
Inferring Local Genealogies on Closely Related Genomes	213
<i>Ryan A. Leo Elworth and Luay Nakhleh</i>	

Enhancing Searches for Optimal Trees Using SIESTA.	232
<i>Pranjal Vachaspati and Tandy Warnow</i>	
On the Rank-Distance Median of 3 Permutations	256
<i>Leonid Chindelevitch and João Meidanis</i>	
Statistical Consistency of Coalescent-Based Species Tree Methods Under Models of Missing Data.	277
<i>Michael Nute and Jed Chou</i>	
Fast Heuristics for Resolving Weakly Supported Branches Using Duplication, Transfers, and Losses	298
<i>Han Lai, Maureen Stolzer, and Dannie Durand</i>	
Author Index	321