

Lecture Notes in Bioinformatics

6398

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

Eric Tannier (Ed.)

Comparative Genomics

International Workshop, RECOMB-CG 2010
Ottawa, Canada, October 9-11, 2010
Proceedings

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 Editor

Eric Tannier

INRIA Rhône-Alpes

Laboratoire de Biométrie et Biologie Evolutive

Université de Lyon 1

43, boulevard du 11 novembre 1918

69622 Villeurbanne, France

E-mail: Eric.Tannier@inria.fr

Library of Congress Control Number: 2010935849

CR Subject Classification (1998): F.2, G.3, E.1, H.2.8, J.3

LNCS Sublibrary: SL 8 – Bioinformatics

ISSN 0302-9743

ISBN-10 3-642-16180-4 Springer Berlin Heidelberg New York

ISBN-13 978-3-642-16180-3 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 2010

Printed in Germany

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

Printed on acid-free paper 06/3180

Preface

The complexity of genome evolution has given birth to exciting challenges for computational biologists. A various range of algorithmic, statistical, mathematical techniques to elucidate the histories of molecules are developed each year and many are presented at the RECOMB satellite workshop on Comparative Genomics. It is a place where scientists working on all aspects of comparative genomics can share ideas on the development of tools and their application to relevant questions.

This volume contains the papers presented at RECOMB-CG 2010, held on October 9–11 in Ottawa. The field is still flourishing as seen from the papers presented this year: many developments enrich the combinatorics of genome rearrangements, while gene order phylogenies are becoming more and more accurate, thanks to a mixing of combinatorial and statistical principles, associated with rapid and thoughtful heuristics. Several papers tend to refine the models of genome evolution, and more and more genomic events can be modeled, from single nucleotide substitutions in whole genome alignments to large structural mutations or horizontal gene transfers.

There were 35 submissions. Each submission was reviewed by at least 2, and on average 2.9, program committee members. The committee decided to accept 24 papers. The program also included 6 invited talks:

Brenda Andrews, University of Toronto
Andrew G. Clark, Cornell University
Nicolas Corradi, University of Ottawa
Jan Dvorak, University of California at Davis
Aoife McLysaght, University of Dublin
Nicholas Putnam, Rice University

I would like to thank all the participants of the conference, and in addition all the people who submitted a paper or a poster. Thanks also to the program committee members and the other reviewers, who did a great amount of work for the conference in a limited amount of time.

The work of the Program Committee was considerably facilitated by the EasyChair website. RECOMB CG 2010 was supported in part by grants from the Canadian Institute for Advanced Research, the Fields Institute for Research in Mathematical Sciences, the Mathematics of Information Technology and Complex Systems (MITACS) Network of Centres of Excellence, and the Vice-President of Research of the University of Ottawa.

Thanks are due to Chunfang Zheng for website design and management, and for implementation of the online registration process.

Organization

Committees and Additional Reviewers

Conference Chair

David Sankoff University
of Ottawa, Canada

Steering Committee

Aoife McLysaght	University of Dublin, Trinity College, Ireland
Jens Lagergren	Royal Institute of Technology (KTH), Sweden
David Sankoff	University of Ottawa, Canada

Local Organizing Committee

Anne Bergeron	Université du Québec à Montréal
Frank Dehne	Carleton University
Guy Drouin	University of Ottawa
Nadia El-Mabrouk	Université de Montréal
Evangelos Kranakis	Carleton University
Marcel Turcotte	University of Ottawa

Program Committee Chair

Eric Tannier	INRIA, Université de Lyon, France
--------------	-----------------------------------

Program Committee

Lars Arvestad	Royal Institute of Technology (KTH), Sweden
Anne Bergeron	Université du Québec à Montréal, Canada
Mathieu Blanchette	McGill University, Montreal, Canada
Guillaume Blin	Université Paris-Est, France
Guillaume Bourque	Genome Institute of Singapore, Singapore
Marilia Braga	Universität Bielefeld, Germany
Michael Brudno	University of Toronto, Canada
Jeremy Buhler	Washington University in Saint Louis, USA
Sèverine Bérard	Université de Montpellier, France
Cedric Chauve	Simon Fraser University, Canada
Avril Coghlan	University College Cork, Ireland
Aaron Darling	University of California-Davis, USA
Dannie Durand	Carnegie Mellon University, USA
Nadia El-Mabrouk	University of Montreal, Canada
Niklas Eriksen	University of Gothenburg, Sweden

VIII Organization

Patricia Evans	University of New Brunswick, Canada
Guillaume Fertin	Université de Nantes, France
Matthew Hahn	Indiana University, USA
Joao Meidanis	University of Campinas / Scylla Bioinformatics, Brazil
Bernard Moret	Ecole Polytechnique Fédérale de Lausanne, Switzerland
Craig Nelson	University of Connecticut, USA
Aida Ouangraoua	INRIA - Lille Nord Europe, France
Michal Ozery-Flato	Tel-Aviv University, Israel
Teresa Przytycka	NIH, USA
Eric Rivals, CNRS	Université de Montpellier, France
Eduardo Rocha	CNRS, Université Paris 6, France
Hugues Roest Crollius	CNRS, Ecole Normale Supérieure Paris, France
Jens Stoye	Universität Bielefeld, Germany
Krister Swenson	Ecole Polytechnique Fédérale de Lausanne, Switzerland
Glenn Tesler	University of California, San Diego, USA
Elisabeth Tillier	Cancer Institute of Ontario, University Health Network, Canada
Stéphane Vialette	CNRS, Université Paris-Est, France
Tiffani Williams	Texas A&M University, USA
Sophia Yancopoulos	Feinstein Institute for Medical Research, USA
Louxin Zhang	National University of Singapore, Singapore

Additional Reviewers

Sébastien Angibaud	Université de Nantes, France
Robert Beiko	Dalhousie University, Halifax, Canada
Camille Berthelot	Ecole Normale Supérieure Paris, France
Annie Chateau	Université de Montpellier, France
Pedro Feijao	University of Campinas, Brazil
Katharina Jahn	Universität Bielefeld, Germany
Irena Rusu	Université de Nantes, France
Florian Sikora	Université Paris-Est, France
Maureen Stolzer	Carnegie Mellon University, USA
Todd Treangen	Institut Pasteur, CNRS and UPMC University, France
Roland Wittler	Simon Fraser University, Canada
Damian Wojtowicz	NIH, USA
Jie Zheng	NIH, USA

Previous Meetings

2003, Minneapolis, USA

2004, Bertinoro, Italy

2005, Dublin, Ireland

2006, Montreal, Canada

2007, San Diego, USA

2008, Paris, France

2009, Budapest, Hungary

Table of Contents

Genome Aliquoting Revisited	1
<i>Robert Warren and David Sankoff</i>	
The Problem of Chromosome Reincorporation in DCJ Sorting and Halving	13
<i>Jakub Kováč, Marília D.V. Braga, and Jens Stoye</i>	
Advances on Genome Duplication Distances	25
<i>Yves Gagnon, Olivier Tremblay Savard, Denis Bertrand, and Nadia El-Mabrouk</i>	
Listing All Parsimonious Reversal Sequences: New Algorithms and Perspectives	39
<i>Ghada Badr, Krister M. Swenson, and David Sankoff</i>	
Ultra-Perfect Sorting Scenarios	50
<i>Aïda Ouangraoua, Anne Bergeron, and Krister M. Swenson</i>	
On Sorting Genomes with DCJ and Indels	62
<i>Marília D.V. Braga</i>	
The Zero Exemplar Distance Problem	74
<i>Minghui Jiang</i>	
Scaffold Filling under the Breakpoint Distance	83
<i>Haitao Jiang, Chunfang Zheng, David Sankoff, and Binhai Zhu</i>	
An Efficient Algorithm for Gene/Species Trees Parsimonious Reconciliation with Losses, Duplications and Transfers	93
<i>Jean-Philippe Doyon, Celine Scornavacca, K.Yu. Gorbunov, Gergely J. Szöllősi, Vincent Ranwez, and Vincent Berry</i>	
Detecting Highways of Horizontal Gene Transfer	109
<i>Mukul S. Bansal, J. Peter Gogarten, and Ron Shamir</i>	
On Exploring Genome Rearrangement Phylogenetic Patterns	121
<i>Andrew Wei Xu</i>	
Fast and Accurate Phylogenetic Reconstruction from High-Resolution Whole-Genome Data and a Novel Robustness Estimator	137
<i>Yu Lin, Vaibhav Rajan, and Bernard M.E. Moret</i>	
A Simple Measure of the Dynamics of Segmented Genomes: An Application to Influenza	149
<i>Stéphane Aris-Brosou</i>	

Novel Definition and Algorithm for Chaining Fragments with Proportional Overlaps	161
<i>Raluca Uricaru, Alban Mancheron, and Eric Rivals</i>	
Assessing the Robustness of Complete Bacterial Genome Segmentations	173
<i>Hugo Devillers, Hélène Chiapello, Sophie Schbath, and Meriem El Karoui</i>	
An Algorithm to Solve the Motif Alignment Problem for Approximate Nested Tandem Repeats	188
<i>Atheer A. Matroud, Michael D. Hendy, and Christopher P. Tuffley</i>	
Limited Lifespan of Fragile Regions in Mammalian Evolution	198
<i>Max A. Alekseyev and Pavel A. Pevzner</i>	
Mapping Association between Long-Range <i>Cis</i> -Regulatory Regions and Their Target Genes Using Comparative Genomics	216
<i>Emmanuel Mongin, Ken Dewar, and Mathieu Blanchette</i>	
A New Genomic Evolutionary Model for Rearrangements, Duplications, and Losses That Applies across Eukaryotes and Prokaryotes	228
<i>Yu Lin and Bernard M.E. Moret</i>	
A Simulation Tool for the Study of Symmetric Inversions in Bacterial Genomes	240
<i>Ulisses Dias, Zanoni Dias, and João C. Setubal</i>	
Consistency of Sequence-Based Gene Clusters	252
<i>Roland Wittler and Jens Stoye</i>	
Efficient Computation of Approximate Gene Clusters Based on Reference Occurrences	264
<i>Katharina Jahn</i>	
The Complexity of the Gapped Consecutive-Ones Property Problem for Matrices of Bounded Maximum Degree	278
<i>Ján Maňuch and Murray Patterson</i>	
An Approximation Algorithm for Computing a Parsimonious First Speciation in the Gene Duplication Model	290
<i>Aïda Ouangraoua, Krister M. Swenson, and Cedric Chauve</i>	
Author Index	303