19th International Workshop on Algorithms in Bioinformatics

WABI 2019, September 8–10, 2019, Niagara Falls, NY, USA

Edited by
Katharina T. Huber
Dan Gusfield
LIPIcs – Leibniz International Proceedings in Informatics

LIPIcs is a series of high-quality conference proceedings across all fields in informatics. LIPIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

Editorial Board

- Luca Aceto (Chair, Gran Sasso Science Institute and Reykjavik University)
- Christel Baier (TU Dresden)
- Mikolaj Bojanczyk (University of Warsaw)
- Roberto Di Cosmo (INRIA and University Paris Diderot)
- Javier Esparza (TU München)
- Meena Mahajan (Institute of Mathematical Sciences)
- Dieter van Melkebeek (University of Wisconsin-Madison)
- Anca Muscholl (University Bordeaux)
- Luke Ong (University of Oxford)
- Catuscia Palamidessi (INRIA)
- Thomas Schwentick (TU Dortmund)
- Raimund Seidel (Saarland University and Schloss Dagstuhl – Leibniz-Zentrum für Informatik)

ISSN 1868-8969

https://www.dagstuhl.de/lipics
Contents

Preface  
Katharina T. Huber and Dan Gusfield ........................................... 0:vii

Building a Small and Informative Phylogenetic Supertree  
Jesper Jansson, Konstantinos Mampentzidis, and Sandhya T. P. ............ 1:1–1:14

Alignment- and Reference-Free Phylogenomics with Colored de Bruijn Graphs  
Roland Wittler ............................................................... 2:1–2:14

Quantified Uncertainty of Flexible Protein-Protein Docking Algorithms  
Nathan L. Clement ............................................................ 3:1–3:12

TRACTION: Fast Non-Parametric Improvement of Estimated Gene Trees  
Sarah Christensen, Erin K. Molloy, Pranjal Vachaspati, and Tandy Warnow ..... 4:1–4:16

Better Practical Algorithms for rSPR Distance and Hybridization Number  

pClay: A Precise Parallel Algorithm for Comparing Molecular Surfaces  
Georgi D. Georgiev, Kevin F. Dodd, and Brian Y. Chen ........................ 6:1–6:13

Read Mapping on Genome Variation Graphs  
Kavya Vaddadi, Rajgopal Srinivasan, and Naveen Sivadasan .................... 7:1–7:17

Finding All Maximal Perfect Haplotype Blocks in Linear Time  
Jarno Alanko, Hideo Bannai, Bastien Cazaux, Pierre Peterlongo, and Jens Stoye 8:1–8:9

A New Paradigm for Identifying Reconciliation-Scenario Altering Mutations  
Conferring Environmental Adaptation  
Roni Zoller, Meirav Zehavi, and Michal Ziv-Ukelson ............................ 9:1–9:13

Jointly Embedding Multiple Single-Cell Omics Measurements  
Jie Liu, Yuanhao Huang, Ritambhara Singh, Jean-Philippe Vert, and  
William Stafford Noble ..................................................... 10:1–10:13

Inferring Diploid 3D Chromatin Structures from Hi-C Data  
Alexandra Gesine Cauer, Gürkan Yardımcı, Jean-Philippe Vert,  
Nelle Varoquaux, and William Stafford Noble ................................ 11:1–11:13

Consensus Clusters in Robinson-Foulds Reticulation Networks  
Alexey Markin and Oliver Eulenstein ....................................... 12:1–12:12

Weighted Minimum-Length Rearrangement Scenarios  

Fast and Accurate Structure Probability Estimation for Simultaneous Alignment  
and Folding of RNAs  
Milad Miladi, Martin Raden, Sebastian Will, and Rolf Backofen .............. 14:1–14:13

Context-Aware Seeds for Read Mapping  
Hongyi Xin, Mingfu Shao, and Carl Kingsford .................................. 15:1–15:13

Bounded-Length Smith–Waterman Alignment  
Alexander Tiskin .............................................................. 16:1–16:12

19th International Workshop on Algorithms in Bioinformatics (WABI 2019).  
Editors: Katharina T. Huber and Dan Gusfield  
Leibniz International Proceedings in Informatics  
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
Validating Paired-End Read Alignments in Sequence Graphs

Detecting Transcriptomic Structural Variants in Heterogeneous Contexts via the Multiple Compatible Arrangements Problem
   Yutong Qiu, Cong Ma, Han Xie, and Carl Kingsford .................................. 18:1–18:5

Faster Pan-Genome Construction for Efficient Differentiation of Naturally Occurring and Engineered Plasmids with Plaster
   Qi Wang, R. A. Leo Elworth, Tian Rui Liu, and Todd J. Treangen ..................... 19:1–19:12

Rapidly Computing the Phylogenetic Transfer Index
   Jakub Truszkowski, Olivier Gascuel, and Krister M. Swenson ......................... 20:1–20:12

Empirical Performance of Tree-Based Inference of Phylogenetic Networks
   Zhen Cao, Jiafan Zhu, and Luay Nakhleh ............................................... 21:1–21:13

A Combinatorial Approach for Single-cell Variant Detection via Phylogenetic Inference

Topological Data Analysis Reveals Principles of Chromosome Structure in Cellular Differentiation
   Natalie Sauerwald, Yihang Shen, and Carl Kingsford .................................. 23:1–23:16

Synteny Paths for Assembly Graphs Comparison
   Evgeny Poleviov and Mikhail Kolmogorov .............................................. 24:1–24:14
This proceedings volume contains papers presented at the 19th Workshop on Algorithms in Bioinformatics (WABI 2019), which was held in Niagara Falls, New York, USA, September 8–10, 2019. WABI 2019 was held together with the 10th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB).

The Workshop on Algorithms in Bioinformatics is an annual conference established in 2001 to cover all aspects of algorithmic work in bioinformatics, computational biology, and systems biology. The workshop is intended as a forum for discrete algorithms and machine-learning methods that address important problems in biology (particularly problems based on molecular data and phenomena); that are founded on sound models; that are computationally efficient; and that have been implemented and tested in simulations and on real data-sets. The meeting’s focus is on recent research results, including significant work-in-progress, as well as identifying and exploring directions of future research.

In 2019, a total of 47 manuscripts were submitted to WABI from which 24 were selected for presentation at the conference and are included in this proceedings volume as full papers. Extended versions of selected papers have been invited for publication in a thematic series in the journal Algorithms for Molecular Biology (AMB), published by BioMed Central. The 24 papers selected for the conference underwent a thorough peer review, involving at least three (and often four or five) independent reviewers per submitted paper, followed by discussions among the WABI Program Committee members. The selected papers cover a wide range of topics including phylogenetic trees and networks, biological network analysis, sequence alignment and assembly, genomic-level evolution, sequence and genome analysis, RNA and protein structure, topological data analysis, and more.

We thank all the authors of submitted papers and the members of the WABI Program Committee and their reviewers for their efforts that made this conference possible. We are also grateful to the WABI Steering Committee for their help and advice. We thank all the conference participants and speakers who contributed to a great scientific program. In particular, we are indebted to the keynote speaker of the conference, Nadia El-Mabrouk, for her presentation. WABI 2019 is grateful for the support of the sponsors of ACM BCB 2019, The ACM, and SIGBio. We thank Letu Qingge for setting up the WABI webpage, Michael Wagner for his assistance with putting together the WABI conference proceedings, and the ACM BCB Organizing Committee, especially Xinghua (Mindy) Shi for her efforts to coordinate WABI and ACM-BCB 2019. Finally, we also thank Dong Si, Michael Buck and Pierangelo Veltri for their hard work in making all the local arrangements to ensure an exciting and successful WABI and ACM BCB.
## List of Authors

<table>
<thead>
<tr>
<th>Author</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarno Alanko</td>
<td>Department of Computer Science, University of Helsinki, Finland</td>
</tr>
<tr>
<td>Srinivas Aluru</td>
<td>School of Computational Science and Engineering, Georgia Institute of Technology, USA</td>
</tr>
<tr>
<td>Rolf Backofen</td>
<td>Bioinformatics Group, Department of Computer Science, University of Freiburg, Germany; Signalling Research Centres BIOSS and CIBSS, University of Freiburg, Germany</td>
</tr>
<tr>
<td>Hideo Bannai</td>
<td>Department of Informatics, Kyushu University, Japan</td>
</tr>
<tr>
<td>Zhen Cao</td>
<td>Department of Computer Science, Rice University, 6100 Main Street, Houston, TX 77005, USA</td>
</tr>
<tr>
<td>Alexandra Gesine Cauer</td>
<td>Department of Genome Sciences, University of Washington, Seattle, WA, USA</td>
</tr>
<tr>
<td>Bastien Cazaux</td>
<td>Department of Computer Science, University of Helsinki, Finland</td>
</tr>
<tr>
<td>Annie Chateau</td>
<td>LIRMM - Université Montpellier, France</td>
</tr>
<tr>
<td>Brian Y. Chen</td>
<td>Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA</td>
</tr>
<tr>
<td>Zhi-Zhong Chen</td>
<td>Division of Information System Design, Tokyo Denki University, Japan</td>
</tr>
<tr>
<td>Sarah Christensen</td>
<td>University of Illinois at Urbana-Champaign, USA</td>
</tr>
<tr>
<td>Nathan L. Clement</td>
<td>Department of Computer Science, University of Texas at Austin, USA</td>
</tr>
<tr>
<td>Alexander Dilthey</td>
<td>Institute of Medical Microbiology, University Hospital of Düsseldorf, Germany</td>
</tr>
<tr>
<td>Kevin F. Dodd</td>
<td>Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA</td>
</tr>
<tr>
<td>Mohammadamin Edrisi</td>
<td>Department of Computer Science, Rice University, Houston, TX, USA</td>
</tr>
<tr>
<td>R. A. Leo Elworth</td>
<td>Department of Computer Science, Rice University, Houston, TX 77005, USA</td>
</tr>
<tr>
<td>Oliver Eulenstein</td>
<td>Department of Computer Science, Iowa State University, Ames, IA, USA</td>
</tr>
<tr>
<td>Olivier Gascuel</td>
<td>Unité Bioinformatique Evolutive, Département de Biologie Computationnelle, USR 3756, Institut Pasteur et CNRS, Paris, France; LIRMM, CNRS, Université Montpellier, Montpellier, France</td>
</tr>
<tr>
<td>Georgi D. Georgiev</td>
<td>Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA</td>
</tr>
<tr>
<td>Yuanhao Huang</td>
<td>Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA</td>
</tr>
<tr>
<td>Chirag Jain</td>
<td>School of Computational Science and Engineering, Georgia Institute of Technology, USA</td>
</tr>
<tr>
<td>Jesper Jansson</td>
<td>The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong</td>
</tr>
<tr>
<td>Carl Kingsford</td>
<td>Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA</td>
</tr>
<tr>
<td>Mikhail Kolmogorov</td>
<td>Department of Computer Science and Engineering, University of California, San Diego, CA, USA</td>
</tr>
<tr>
<td>Jie Liu</td>
<td>Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA</td>
</tr>
</tbody>
</table>
Qi Wang (19)
Systems, Synthetic, and Physical Biology (SSPB) Graduate Program, Rice University, Houston, TX 77005, USA

Tandy Warnow (4)
University of Illinois at Urbana-Champaign, USA

Sebastian Will (14)
Theoretical Biochemistry Group (TBI), Institute for Theoretical Chemistry, University of Vienna, Austria

Roland Wittler (2)
Genome Informatics, Faculty of Technology, Bielefeld University, Germany; Center for Biotechnology, Bielefeld University, Germany

Han Xie (18)
Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

Hongyi Xin (15)
Computer Science Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

Kohei Yamada (5)
Division of Information System Design, Tokyo Denki University, Japan

Gürkan Yardımcı (11)
Department of Genome Sciences, University of Washington, Seattle, WA, USA

Hamim Zafar (22)
Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

Meirav Zehavi (9)
Ben Gurion University of the Negev, Israel

Haowen Zhang (17)
School of Computational Science and Engineering, Georgia Institute of Technology, USA

Jiafan Zhu (21)
Department of Computer Science, Rice University, 6100 Main Street, Houston, TX 77005, USA

Michal Ziv-Ukelson (9)
Ben Gurion University of the Negev, Israel

Roni Zoller (9)
Ben Gurion University of the Negev, Israel