

# 23rd International Workshop on Algorithms in Bioinformatics

WABI 2023, September 4–6, 2023, Houston, TX, USA

Edited by

Djamal Belazzougui  
Aïda Ouangraoua



*Editors*

**Djamal Belazzougui**

CERIST, Algeria

dbelazzougui@cerist.dz

**Aïda Ouangraoua** 

University of Sherbrooke, Canada

Aida.Ouangraoua@USherbrooke.ca

*ACM Classification 2012*

Applied computing → Bioinformatics; Applied computing → Computational biology; Theory of computation → Design and analysis of algorithms; Mathematics of computing → Discrete mathematics; Mathematics of computing → Information theory

**ISBN 978-3-95977-294-5**

*Published online and open access by*

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <https://www.dagstuhl.de/dagpub/978-3-95977-294-5>.

*Publication date*

August, 2023

*Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <https://portal.dnb.de>.

*License*

This work is licensed under a Creative Commons Attribution 4.0 International license (CC-BY 4.0):

<https://creativecommons.org/licenses/by/4.0/legalcode>.



In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.

The copyright is retained by the corresponding authors.

Digital Object Identifier: 10.4230/LIPIcs.WABI.2023.0

ISBN 978-3-95977-294-5

ISSN 1868-8969

<https://www.dagstuhl.de/lipics>

## 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*, Reykjavik University, IS and Gran Sasso Science Institute, IT)
- Christel Baier (TU Dresden, DE)
- Roberto Di Cosmo (Inria and Université de Paris, FR)
- Faith Ellen (University of Toronto, CA)
- Javier Esparza (TU München, DE)
- Daniel Král' (Masaryk University, Brno, CZ)
- Meena Mahajan (Institute of Mathematical Sciences, Chennai, IN)
- Anca Muscholl (University of Bordeaux, FR)
- Chih-Hao Luke Ong (University of Oxford, GB)
- Phillip Rogaway (University of California, Davis, US)
- Eva Rotenberg (Technical University of Denmark, Lyngby, DK)
- Raimund Seidel (Universität des Saarlandes, Saarbrücken, DE and Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Wadern, DE)
- Pierre Senellart (ENS, Université PSL, Paris, FR)

**ISSN 1868-8969**

**<https://www.dagstuhl.de/lipics>**



## ■ Contents

Preface	
<i>Djamal Belazzougui and Aïda Ouangraoua</i> .....	0:vii–0:viii
WABI 2023 Committees	
.....	0:ix–0:xiii

### Invited Talk

Algorithmic Approaches to Study Mutational Processes in Cancer	
<i>Teresa M. Przytycka</i> .....	1:1–1:2

### Papers

EMMA: Adding Sequences into a Constraint Alignment with High Accuracy and Scalability	
<i>Chengze Shen, Baqiao Liu, Kelly P. Williams, and Tandy Warnow</i> .....	2:1–2:2
BATCH-SCAMPP: Scaling Phylogenetic Placement Methods to Place Many Sequences	
<i>Eleanor Wedell, Chengze Shen, and Tandy Warnow</i> .....	3:1–3:2
Optimal Subtree Prune and Regraft for Quartet Score in Sub-Quadratic Time	
<i>Shayesteh Arasti and Siavash Mirarab</i> .....	4:1–4:20
Leveraging Constraints Plus Dynamic Programming for the Large Dollo Parsimony Problem	
<i>Junyan Dai, Tobias Rubel, Yunheng Han, and Erin K. Molloy</i> .....	5:1–5:23
Simultaneous Reconstruction of Duplication Episodes and Gene-Species Mappings	
<i>Paweł Górecki, Natalia Rutecka, Agnieszka Mykowiecka, and Jarosław Paszek</i> ....	6:1–6:18
Making a Network Orchard by Adding Leaves	
<i>Leo van Iersel, Mark Jones, Esther Julien, and Yukihiro Murakami</i> .....	7:1–7:20
Quartets Enable Statistically Consistent Estimation of Cell Lineage Trees Under an Unbiased Error and Missingness Model	
<i>Yunheng Han and Erin K. Molloy</i> .....	8:1–8:2
Inferring Temporally Consistent Migration Histories	
<i>Mrinmoy Saha Roddur, Sagi Snir, and Mohammed El-Kebir</i> .....	9:1–9:22
Finding Maximal Exact Matches in Graphs	
<i>Nicola Rizzo, Manuel Cáceres, and Veli Mäkinen</i> .....	10:1–10:17
Revisiting the Complexity of and Algorithms for the Graph Traversal Edit Distance and Its Variants	
<i>Yutong Qiu, Yihang Shen, and Carl Kingsford</i> .....	11:1–11:22
Co-Linear Chaining on Pangenome Graphs	
<i>Jyotshna Rajput, Ghanshyam Chandra, and Chirag Jain</i> .....	12:1–12:18

23rd International Workshop on Algorithms in Bioinformatics (WABI 2023).

Editors: Djamal Belazzougui and Aïda Ouangraoua



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Acceleration of FM-Index Queries Through Prefix-Free Parsing <i>Aaron Hong, Marco Oliva, Dominik Köppl, Hideo Bannai, Christina Boucher, and Travis Gagie</i> .....	13:1–13:16
Exact Sketch-Based Read Mapping <i>Tizian Schulz and Paul Medvedev</i> .....	14:1–14:19
Fractional Hitting Sets for Efficient and Lightweight Genomic Data Sketching <i>Timothé Rouzé, Igor Martayan, Camille Marchet, and Antoine Limasset</i> .....	15:1–15:27
Fast, Parallel, and Cache-Friendly Suffix Array Construction <i>Jamshed Khan, Tobias Rubel, Larman Dhulipala, Erin Molloy, and Rob Patro</i> ....	16:1–16:21
Compression Algorithm for Colored de Bruijn Graphs <i>Amatur Rahman, Yoann Dufresne, and Paul Medvedev</i> .....	17:1–17:14
Fulgor: A Fast and Compact $k$ -mer Index for Large-Scale Matching and Color Queries <i>Jason Fan, Noor Pratap Singh, Jamshed Khan, Giulio Ermanno Pibiri, and Rob Patro</i> .....	18:1–18:21
SparseRNAFold: Sparse RNA Pseudoknot-Free Folding Including Dangles <i>Mateo Gray, Sebastian Will, and Hosna Jabbari</i> .....	19:1–19:18
Automatic Exploration of the Natural Variability of RNA Non-Canonical Geometric Patterns with a Parameterized Sampling Technique <i>Théo Boury, Yann Ponty, and Vladimir Reinharz</i> .....	20:1–20:22
Balancing Minimum Free Energy and Codon Adaptation Index for Pareto Optimal RNA Design <i>Xinyu Gu, Yuanyuan Qi, and Mohammed El-Kebir</i> .....	21:1–21:20
Bridging Disparate Views on the DCJ-Indel Model for a Capping-Free Solution to the Natural Distance Problem <i>Leonard Bohnenkämper</i> .....	22:1–22:18
Reinforcement Learning for Robotic Liquid Handler Planning <i>Mohsen Ferdosi, Yuejun Ge, and Carl Kingsford</i> .....	23:1–23:16

## ■ Preface

This proceedings volume contains papers and abstracts presented at the 23rd Workshop on Algorithms in Bioinformatics (WABI 2023), which was held in Houston, Texas, USA, from September 4th to 6th, 2023.

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 conference serves as a forum for the presentation of new insights about discrete algorithms and machine-learning methods that address important problems in biology, particularly those based on molecular data and phenomena. These algorithms are founded on sound models, exhibit computational efficiency, and have been implemented and tested in simulations and on real datasets. The focus of the meeting is on recent research results, including significant work-in-progress, as well as identifying and exploring directions for future research. Over the 23 instances of WABI, computational biology has significantly grown in importance. Computational analysis methods, some of which have been advanced significantly over the years at WABI, have proven crucial for the global response to the COVID-19 pandemic and for vaccine development.

WABI 2023 was co-located with ACM-BCB 2023. A total of 44 manuscripts were submitted to WABI 2023, out of which 22 were selected for presentation at the conference. Among these, 19 are included in this proceedings volume as full papers, and 3 are included as extended abstracts. The 22 papers chosen for the conference underwent a rigorous peer review process, involving at least three (most frequently four) independent reviewers per submitted paper. Following this, discussions took place among the WABI Program Committee members. The reviews were conducted by both Program Committee members and additional reviewers who were selected based on their expertise relevant to specific papers. The selected papers encompass a wide range of topics, including phylogenetic trees and networks, cancer phylogenetics, sequence alignment and assembly, gene and genomic-level evolution, genome rearrangement, sequence and genome analysis, RNA structure, pattern matching, data compression, and more. These papers are arranged according to the conference program within this volume. The program also featured an invited talk, two highlight talks, and a poster session.

Extended versions of select papers have been invited for publication in a thematic series in the journal “Algorithms for Molecular Biology” (AMB), published by BioMed Central. A total of 16 extended papers will be submitted to this thematic series.

We extend our gratitude to all the authors of the submitted papers, whose contributions made this conference possible. Special thanks are due to all the members of the WABI 2023 Program Committee and their sub-reviewers for their diligent efforts in reviewing the manuscripts and engaging in comprehensive discussions. These discussions informed the decision-making process and resulted in constructive review reports for the authors. We are grateful to the WABI Steering Committee for their availability, assistance, and guidance. Our thanks go out to all the conference participants, session chairs, and speakers who contributed to a highly successful scientific program. In particular, we express our gratitude to the conference’s keynote speaker, Teresa Przytycka (NCBI, NIH), for her presentation titled “Algorithmic Approaches to Study Mutational Processes in Cancer”. We extend our thanks to the highlight presentation speakers, Christina Boucher and Travis Gagie, for graciously agreeing to speak at the conference on the topic of “The Theory and Applications of the r-index in Bioinformatics”. Furthermore, we would like to express our special appreciation

23rd International Workshop on Algorithms in Bioinformatics (WABI 2023).

Editors: Djamel Belazzougui and Aïda Ouangraoua



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

to the ACM-BCB 2023 Organizing Committee for their dedicated efforts and generosity in organizing the event. The EasyChair system was utilized for handling submissions and reviews.

Previous WABI proceedings appeared in LNCS/LNBI volumes as follows: 2149 (WABI 2001, Aarhus), 2452 (WABI 2002, Rome), 2812 (WABI 2003, Budapest), 3240 (WABI 2004, Bergen), 3692 (WABI 2005, Mallorca), 4175 (WABI 2006, Zurich), 4645 (WABI 2007, Philadelphia), 5251 (WABI 2008, Karlsruhe), 5724 (WABI 2009, Philadelphia), 6293 (WABI 2010, Liverpool), 6833 (WABI 2011, Saarbrücken), 7534 (WABI 2012, Ljubljana), 8126 (WABI 2013, Sophia Antipolis), 8701 (WABI 2014, Wrocław), 9289 (WABI 2015, Atlanta), and 9838 (WABI 2016, Aarhus). Beginning in 2017, they appeared in LIPIcs volumes: 88 (WABI 2017, Boston), 113 (WABI 2018, Helsinki), 143 (WABI 2019, Boston), 172 (WABI 2020, held virtually in Pisa), 201 (WABI 2021, held virtually in Chicago), and 242 (WABI 2022, Potsdam).

Djamal Belazzougui and Aïda Ouangraoua



# ■ WABI 2023 Committees

## Steering Committee

Vincent Moulton  
University of East Anglia, UK

Nadia Pisanti  
University of Pisa, Italy

Mona Singh  
Princeton University, USA

Jens Stoye  
Bielefeld University, Germany

## PC Chairs

Djamal Belazzougui  
CERIST, Algeria

Aïda Ouangraoua  
University of Sherbrooke, Canada

## Program Committee

Tatsuya Akutsu  
Kyoto University, Japan

Jasmijn Baaijens  
TU Delft, Netherlands

Anne Bergeron  
Université du Québec à Montréal, Canada

Paola Bonizzoni  
Università di Milano-Bicocca, Italy

Christina Boucher  
University of Florida, USA

Marília Braga  
Bielefeld University, Germany

Cédric Chauve  
Simon Fraser University, Canada

Daniel Doerr  
University of Duesseldorf, Germany

Mohammed El-Kebir  
University of Illinois at Urbana-Champaign, USA

23rd International Workshop on Algorithms in Bioinformatics (WABI 2023).

Editors: Djamal Belazzougui and Aïda Ouangraoua



Leibniz International Proceedings in Informatics  
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

**0:x WABI 2023 Committees**

Nadia El-Mabrouk  
University of Montreal, Canada

Anna Gambin  
Warsaw University, Poland

Bjarni Halldorsson  
deCODE genetics and Reykjavik University, Iceland

Katharina Huber  
University of East Anglia, UK

Gregory Kucherov  
University of Paris Est, France

Manuel Lafond  
Université de Sherbrooke, Canada

Elodie Laine  
Sorbonne Université, France

Veli Makinen  
University of Helsinki, Finland

Erin Molloy  
University of Maryland College Park, USA

Vincent Moulton  
University of East Anglia, UK

Francesca Nadalin  
EMBL-EBI / ITT, Italy, England

Gonzalo Navarro  
University of Chile, Chile

Yann Ponty  
CNRS & École poly technique, France

Teresa Przytycka  
NCBI, NIH, USA

Nadia Pisanti  
University of Pisa, Italy

Solon Pissis  
CWI Amsterdam, the Netherlands

Alberto Policriti  
University of Udine, Italy

Sven Rahmann  
Saarland University, Germany

Knut Reinert  
FU Berlin, Germany

Eric Rivals  
LIRMM & Université de Montpellier, France

Giovanna Rosone  
University of Pisa, Italy

Mikaël Salson  
Université de Lille, France

Alexander Schoenhuth  
Bielefeld University, Germany

Mingfu Shao  
Penn State, USA

Blerina Sinimeri  
Université Lyon I & INRIA, France

Krister Swenson  
CNRS & Université de Montpellier, France

Sharma Thankachan  
University of Central Florida, USA

Alexandru Tomescu  
University of Helsinki, Finland

Hélène Touzet  
CNRS & Université de Lille, France

Gianluca Della Vedova  
University of Milano-Bicocca, Italy

Tomas Vinar  
Comenius University, Slovakia

Jérôme Waldispühl  
McGill University, Canada

Tandy Warnow  
University of Illinois at Urbana-Champaign, USA

Prudence W. H. Wong  
University of Liverpool, UK

Louxin Zhang  
National University of Singapore, Singapore

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

### **Subreviewers**

Paniz Abedin

Metin Balaban

Leonard Bohnenkämper

Vladimir Boza

Broňa Brejová

Hannes P. Eggertsson Manuel Caceres

Bastien Cazaux

Isaure Chauvot de Beauchene

Ke Chen

Sriram P. Chockalingam

Nicola Cotumaccio

Mitra Darvish

Mattéo Delabre

Diego Diaz

Travis Gagie

Mathieu Gascon

Nimisha Ghosh

Daniel Gibney

Mauricio Soto Gomez

Simon Gottlieb

Xinyu Gu

Pina Krell

Xiang Li

Ronny Lorenz

Davide Martincigh

Fábio Henrique Viduani Martinez

Roberto Pagliarini

Murray Patterson

Anna Vathrakokoili Pournara

Rangaram D Raghuram

Vaibhav Rajan

Hugues Richard

Diego P. Rubert

Cenk Sahinalp

Johannes Schlüter

Sebastian Schmidt

Johanna Schmitz

Grzegorz Skoraczyński

Francesco Stranieri

Riccardo Vicedomini

Tasfia Zahin

Diego Zea

**ACM-BCB Organizing Committee**

May D. Wang (general chair)  
Georgia Institute of Technology, USA

Byung-Jun Yoon (general chair)  
Texas A&M University, USA

Peter L. Elkin (technical program chair)  
University at Buffalo, USA

Xiaoning Qian (technical program chair and local chair)  
Texas A&M University, USA

Wenyi Wang (technical program chair)  
MD Anderson Cancer Center, USA

Yijie Wang (poster chair)  
Indiana University, USA

Vicky Yao (local chair)  
Rice University, USA

Ananth Kalyanaraman (website)  
Washington State University, USA

(and all the ACM-BCB Organizing Committee)

