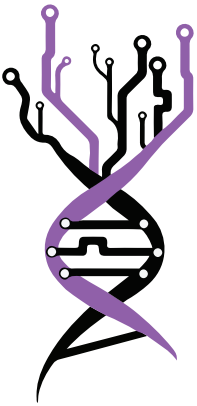
#BelBi2023 • Belgrade, Serbia

BOOK OF ABSTRACTS



4th Belgrade Bioinformatics Conference

HYBRID • 19 - 23 JUNE **2023**

EDITORS
Dr. Ivana Morić
Dr. Valentina Đorđević

ISBN: 978-86-82679-14-1

belbi.bg.ac.rs

Title 4th Belgrade Bioinformatics Conference
BOOK OF ABSTRACTS

Publisher Institute of Molecular Genetics and Genetic Engineering, University of Belgrade
Vojvode Stepe 444a, Belgrade, Serbia
https://www.imgge.bg.ac.rs/

Editors dr. Ivana Morić
dr. Valentina Đorđevic

Technical editor Dušan Radojević

ISBN **978-86-82679-14-1**

Copyright © 2023 Institute of Molecular Genetics and Genetic Engineering, University of Belgrade

BelBi2023 Committees

International Advisory Board

Alessandro Treves,
International School for Advanced Studies, Trieste, Italy

Guanglan Zhang,
Health Informatics Lab, Metropolitan College, Boston University, Boston, USA

Konstantin Severinov,
Waksman Institute for Microbiology, Rutgers University, Piscataway, USA

Lou Chitkushev,
Health Informatics Lab, Metropolitan College, Boston University, Boston, USA

Milena Stevanović,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia

Oxana Galzitskaya,
Institute of Protein Research, Russian Academy of Sciences, Moscow, Russia

Paul Sorba,
Laboratory of Theoretical Physics and CNRS, Annecy, France

Peter Tompa,
VIB Structural Biology Research Center, Flanders Institute for Biotechnology, Brussels, Belgium

Predrag Radivojac,
Khoury College of Computer Sciences, Northeastern University, Boston, SA

Sergey Volkov,
Bogolyubov Institute for Theoretical Physics, National Academy of Sciences, Kiev, Ukraine

Vladimir Brusić,
Li Dak Sum Chair Professor of Computer Science, University of Nottingham, Ningbo, China

Vladimir Uversky,
Department of Molecular Medicine, University of South Florida, Tampa, USA

Yuriy Orlov,
I.M. Sechenov First Moscow State Medical University, Moscow, Russia

Zoran Ognjanović,
Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade, Serbia

International Program Committee

Alexandre de Brevern,
INSERM, Université Paris Cité, Université de la Réunion, Paris, France

Branislava Gemović,
VINCA Institute of Nuclear Sciences, University of Belgrade, Serbia

Branka Zukić,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Serbia

Branko Dragovich,
Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade, Serbia

Dragan Matić,
Faculty of Sciences, Department of Mathematics and Informatics, University of Banja Luka, Bosnia and Herzegovina

George Patrinos,
Department of Pharmacy, University of Patras, Greece

Gordana Pavlović-Lazetić,
Faculty of Mathematics, University of Belgrade, Serbia

Hong-Yu OU,
Shanghai Jiao Tong University, The Microbial Bioinformatics Group, State Key Laboratory of Microbial Metabolism, Shanghai, China

Ivana Morić,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Serbia

Jelena Bojović,
Center for the Fourth Industrial Revolution, Belgrade, Serbia

Marko Đordević,
Faculty of Biology, University of Belgrade, Serbia

Mirjana Maljković,
Faculty of Mathematics,
University of Belgrade, Serbia

Nataša Kovacevic-Grujičić,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Serbia

Nataša Pržulj,
Catalan Institution for Research and Advanced Studies (ICREA), Spain;
Barcelona Supercomputing Center, Spain;
University College London, UK

Nenad Mitić,
Faculty of Mathematics,
University of Belgrade, Serbia

Nevena Veljković,
VINCA Institute of Nuclear Sciences,
University of Belgrade, Serbia

Predrag Radivojac,
Khoury College of Computer Sciences,
Northeastern University, Boston, USA

Saša Malkov,
Faculty of Mathematics,
University of Belgrade, Serbia

Sergei Kozyrev,
Department of Mathematical Physics,
Steklov Mathematical Institute RAS,
Moscow, Russia

Valentina Đorđević,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Serbia

Vladimir Babenko,
Institute of Cytology and Genetics,
Novosibirsk, Russia

Local Organizing Committee

Anđela Rodić,
Faculty of Biology, University of Belgrade

Branislava Gemović,
VINCA Institute of Nuclear Sciences,
University of Belgrade

Branko Dragovich,
Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade

Gordana Pavlović-Lazetić,
Faculty of Mathematics,
University of Belgrade

Ivana Morić,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade

Jovana Kovačević,
Faculty of Mathematics,
University of Belgrade

Marko Đorđević,
Faculty of Biology, University of Belgrade

Marko Živanović,
Bioengineering Research and Development Center

Nenad Filipović,
Faculty of Engineering, University of Kragujevac

Nenad Mitić,
Faculty of Mathematics,
University of Belgrade

Radoslav Davidović,
VINCA Institute of Nuclear Sciences,
University of Belgrade

Saša Malkov,
Faculty of Mathematics,
University of Belgrade

Valentina Đorđević,
Institute of Molecular Genetics and Genetic Engineering, University of Belgrade

Željko D. Popović,
Faculty of Sciences, University of Novi Sad

FOREWORD

Dear colleagues and friends,

The 4th Belgrade Bioinformatics Conference - BelBi2023, where many high-quality scientific contributions were presented, has just ended. With great thanks to all participants, we now proudly present a book of abstracts that both reflects the scientific abundance and diversity of the conference and serves as a reminder of a memorable event.

Several research institutions, faculties, and scientific societies from Serbia joined forces in organizing this international conference, which covered numerous topics in computational biology, bioinformatics, and biomedical and health informatics. The main goal of BelBi2023 was to foster contact between scientists, both early stage career and senior researchers, allowing them to share experiences and latest advances in their fields. We sincerely hope that BelBi2023 has served as a platform for researchers from around the world to meet, initiate new collaborations, and expand professional contacts, and that all of you would become a part of the growing BelBi community.

We are grateful and proud to have welcomed more than 250 researchers from 21 countries. We have had 28 scientific sessions, consisting of more than 60 lectures (including eight Keynote talks), 47 presented posters, as well as three workshops and one satellite event – COST action. We have also organized seven industry lectures, including the NGS Challenge, two Meet the Expert Sessions, and one Business Coffee Break where ten start-up companies took part. And finally, the future BIO4 campus was presented and first panel on Serbia's resources for storage and analyses of genetic data was organized.

We would like to thank all the members of the International Advisory Board and the International Program Committee for their efforts and help in making this event a success. We are very grateful to the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, SAIGE project, and UNDP-Serbia for their support. Finally, the Local Organizing Committee is very grateful to all the sponsors of the conference - BGI, Illumina & Elta'90MS, PacBio & East Diagnostics, ThermoFisher Scientific & Vivogen, Huawei, Labena, DSP Chromatography, RNIDS, Telekom Srbija, Alfa Genetics, Kefo and Superlab, hoping that they will stay with us for many years to come.

Looking forward to seeing you again at the 5th Belgrade Bioinformatics Conference.

Belgrade, July 2023

Dr. Valentina Đorđević
& Dr. Ivana Morić,
On behalf of BelBi2023
Organizing Committee

ORGANIZER

Institute of Molecular Genetics and Genetic Engineering, University of Belgrade

MAIN CO-ORGANIZERS

Faculty of Biology, University of Belgrade

Faculty of Mathematics, University of Belgrade

Mathematical Institute of SASA, Belgrade

Vinča Institute of Nuclear Sciences, University of Belgrade

Serbian Society for Bioinformatics and Computational Biology

CO-ORGANIZERS

Faculty of Sciences, University of Novi Sad

BioIRC - Bioengineering Research and Development Center

Faculty of Engineering, University of Kragujevac

Centre for the Fourth Industrial Revolution in Serbia

Serbian Society for Molecular Biology

Biophysical Society of Serbia

TABLE OF CONTENT

KEYNOTE LECTURES

INVITED LECTURES

ORAL PRESENTATION

POSTER PRESENTATION

Pangenomic Alignment: Strings plus Graphs

Travis Gagie[1]

[1] Faculty of Computer Science, Dalhousie University,
6050 University Avenue, Halifax, Nova Scotia, B3H 1W5, Canada
travis.gagie@dal.ca

The use of only one or a few reference genomes for DNA alignment is known to bias research results and medical diagnoses, but aligning against many reference genomes has been problematic. If we represent such a pangenomic reference as a set of strings, then each seed we find in a DNA read may occur in many of the genomes, so even reporting all those occurrences can be slow, and extending and chaining seeds can be infeasible. On the other hand, if we represent them as a graph then --- even apart from the significant technical challenges of indexing graphs --- we may find many chimeric matches. The more of humanity's genetic diversity we try to represent in the graph, the fuzzier it becomes, and the greater the probability of spurious results.

Most research on pangenomic alignment uses either a string representation or a graph representation, but not both. In this talk we first describe how a tool called MONI indexes a pangenomic reference as a set of strings in small space such that later, for each maximal exact match in a given read, we can quickly find that match's length, the position of one of its occurrences in the set of strings, and the lexicographic rank of the suffix starting with that occurrence. We then describe how a tool called MARIA will, when fully implemented, store a pangenomic reference as a graph in small space such that, given MONI's output about a maximal exact match, we can quickly report all the non-chimeric occurrences of that match in the graph.

Combining MONI and MARIA will give us the advantages of working with both strings and graphs: we index the set of reference genomes, the whole set of reference genomes, and nothing but the set of reference genomes, but for each maximal exact match we output relatively few occurrences in the graph, which are easy to use later in a pipeline.

Keywords: pangenomic alignment, reference genomes, data structures, indexing

Acknowledgement: This talk covers results obtained in collaboration with many other researchers, in particular Christina Boucher and Marco Oliva at the University of Florida, Ben Langmead at Johns Hopkins University and Massimiliano Rossi at Illumina, for MONI; and Andrej Baláž, Adrián Goga and Alessia Petescia at Comenius University, Simon Heumos at the University of Tübingen and Jouni at the UCSC Genomics Institute, for MARIA. The author was funded by NSERC grant RGPIN-07185-2020, NSF/BIO grant DBI-2029552 to Christina Boucher, and NIH/NHGRI grant R01HG011392 to Ben Langmead.

ISBN: 978-86-82679-14-1