

Subseries of Lecture Notes in Computer Science

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

Editorial Board Members

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 subseries at <https://link.springer.com/bookseries/5381>

Itsik Pe'er (Ed.)

Research in Computational Molecular Biology

26th Annual International Conference, RECOMB 2022
San Diego, CA, USA, May 22–25, 2022
Proceedings

Editor
Itsik Pe'er 
Columbia University
New York, NY, USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Bioinformatics
ISBN 978-3-031-04748-0 ISBN 978-3-031-04749-7 (eBook)
<https://doi.org/10.1007/978-3-031-04749-7>

LNCS Sublibrary: SL8 – Bioinformatics

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2022

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, expressed 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.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains 17 extended abstracts and 23 short abstracts representing a total of 40 proceedings papers presented at the 26th International Conference on Research in Computational Molecular Biology (RECOMB 2022), hosted by the University of California, San Diego. The conference took place at La Jolla, California, USA, during May 22–25, 2022. These 40 contributions were selected by a rigorous peer-review process from 188 submissions to the conference. Each of these 188 submissions received reviews from at least three members of the Program Committee (PC) members or their designated sub-reviewers. Following an initial process of independent reviewing, all submissions were opened for discussion by their reviewers and the conference program chair through the EasyChair Conference Management System. Final decisions were made based on reviewer assessments with some adjustment to ensure the technical diversity of the conference program.

RECOMB 2022 allowed authors an option to publish their full extended papers in the conference proceedings or to provide short abstracts for the proceedings and pursue alternative arrangements for publishing the full paper. In addition, the authors of a select set of accepted papers were invited to submit revised manuscripts for consideration for publication in the partner journals, *Cell Systems* and *Genome Research*. All papers that appear as extended abstracts in the proceedings were invited for submission to the RECOMB 2022 special issue of the *Journal of Computational Biology*.

RECOMB 2022 also featured highlight talks of computational biology papers that were published in journals during the previous 18 months. Of the 38 submissions to the highlights track, ten were selected for oral presentation at RECOMB.

In addition to presentations of these contributed papers, RECOMB 2022 featured seven invited keynote talks given by leading scientists:

- Bing Ren (University of California, San Diego), “Single-cell analysis of epigenome in health and disease”.
- Howard Chang (Stanford University), “Personal regulome navigation”.
- Lenore Cowen (Tufts University), “Pathways for Learning from Structure and Organization of Protein Interaction Networks”.
- John Chodera (Sloan Kettering Institute), “The COVID Moonshot: Open science discovery of a novel oral SARS-CoV-2 antiviral”.
- John Marioni (European Bioinformatics Institute), “Computational challenges in single-cell genomics”.
- Wenyi Wang (MD Anderson Cancer Center), “Deciphering cancer cell evolution and ecology”.
- Regina Barzilai (Massachusetts Institute of Technology), “Infusing Biology into Molecular Models for Drug Discovery”.

RECOMB also featured a special invited workshop on genomic privacy, organized by Gamze Gürsoy (Columbia University).

In addition, five topical RECOMB satellite meetings took place in parallel directly preceding the main RECOMB meeting:

- The 19th RECOMB Satellite Conference on Comparative Genomics (RECOMB-CG) co-chaired by Siavash Mirarab (University of California, San Diego), Lingling Jin (University of Saskatchewan), and Dannie Durand (Carnegie Mellon University).
- The 6th RECOMB Satellite Conference on Bioinformatics Education (RECOMB-BE) chaired by Niema Moshiri (University of California, San Diego).
- The 10th RECOMB Satellite on Computational Methods in Genetics (RECOMB-Genetics) co-chaired by Anna-Sapfo Malaspinas (Swiss Institute of Bioinformatics), Sriram Sankararaman (University of California, Los Angeles), and Gillian Belbin (Mount Sinai Institute for Genomics Health).
- The 12th RECOMB Satellite Workshop on Massively Parallel Sequencing (RECOMB-Seq) co-chaired by Can Alkan (Bilkent University), Benjamin Langmead (Johns Hopkins University), Paul Medvedev (Pennsylvania State University), and Stefano Tonzani (Cell Press)
- The RECOMB Satellite Workshop on Computational Cancer Biology (RECOMB-CCB) co-chaired by Hannah Carter (University of California, San Diego) and Simone Zaccaria (University College London).

Two additional RECOMB satellite meetings were held in honor of members of our community: Mike Waterman's 80th birthday was celebrated in a satellite meeting co-chaired by Remo Rohs and Fengzhu Sun (University of Southern California). Benny Chor, who our community lost this year, was commemorated in a satellite meeting co-chaired by Sagi Snir (University of Haifa) and Zohar Yakhini (Israel Institute of Technology). We thank them for organizing these great companion meetings and they and their Program Committees for their hard work in making them possible.

The organization of this conference was the work of many colleagues contributing their time, effort, and expertise. I am especially grateful to the local organizing committee, particularly Conference Chair Vineet Bafna (University of California, San Diego) and co-organizers Vikas Bansal (University of California, San Diego), Jocelyn Bernardo (University of California, San Diego), Melissa Gymrek (University of California, San Diego), Siavash Mirarab (University of California, San Diego), Glenn Tesler (University of California, San Diego), and Kaiyuan Zhu (University of California, San Diego). I am grateful to the many others who volunteered their time and work, including those whose names were not yet known to us at the time of this writing. I also want to thank the Poster Chair, Yaron Orenstein (Ben-Gurion University, Israel), Keynotes Chair, Ewa Szczurek (University of Warsaw), Satellites Chair, Sriram Sankararaman (University of California, Los Angeles), and Highlights Chair, Sushmita Roy (University of Wisconsin, Madison) for their efforts in ensuring a high-quality technical program. I am further grateful to all of those PC members and sub-reviewers who took time out of their busy schedules to review and discuss submissions on a very tight schedule. I also thank the authors of the proceedings papers, the highlights, and the posters for contributing their work to the meeting and for their attendance at the conference.

Final thanks go to all our conference sponsors for their support, who at press time for this volume included Akamai Technologies, Illumina, the University of California, San Diego, and the Department of Computer Science and Engineering (University of California, San Diego), and especially to the sponsors of our student travel awards, the National Science Foundation (NSF) and the International Society for Computational Biology (ISCB).

Beyond the formal details about the conference, and thanking all the direct contributors, RECOMB 2022 represents a pivotal moment for our community, coming together in person for the first time in three years. RECOMB 2020 was canceled in physical form with only a few weeks' notice - the 2020 template I was using for this preface document had no idea a pandemic was coming, squeezing the conference into a Zoom window. We have since all learned the extent to which in-person meetings provide deeper engagement, greater value, and a more immersive sense of community. One of the highlighted talks in RECOMB 2022 even made this point quantitatively. I would thus like to express how honored I am to be ushering back participation in RECOMB in person.

May 2022

Itsik Pe'er

To Benny, to the RECOMB Community in Memory of Benny

This past year, 2021, sadly marked the passing away of Prof. Benny Chor, a scientist, a pillar in the bioinformatics community and in RECOMB, a colleague and a friend. Benny made significant contributions in cryptography, in computational biology, and to the teaching of computer science at all levels.

Benny introduced all his students and many of his colleagues to seeing science in the context of culture, of friendship and human interactions, as a lifestyle more than as a vocation. For Benny, a student or a colleague was first a friend, a person with a full life, maybe with a family, maybe with habits and ideas to explore and to learn from. He was proud, for example, of a PhD student from the Technion who took his complexity course and upon graduation opened a falafel shop. Benny was an avid traveler and sailor, taking his students on challenging adventures. He was an evangelist of high standards and zero compromise morality in science. An outstanding characteristic of Benny was his modesty. Few people know that he won the ACM award for his PhD, due to its seminal founding contributions in cryptography. Only some colleagues in the RECOMB community know of his theoretical computer science career, and his remarkable achievements. To the privileged people who worked with him he conveyed a culture of rigorous science in a pleasant atmosphere. Science with a smile.



In 1985 Benny and colleagues established the concept of verifiable secret sharing and developed related methods and definitions that are still used in the field. In 1998 he started his work on private information retrieval (PIR), a fundamental topic

of high importance in cryptography. His contribution to cryptography continued in further developing techniques and theories related to PIR and other privacy primitives. Among Benny's most prominent contributions to computational biology is the introduction of highly rigorous approaches from computer science and math to the study of evolution, specifically to phylogenetics. This includes his dynamic programming algorithm for quartet-based phylogenetics, the introduction of analytical, symbolic algebra-based approaches to maximum likelihood phylogenetics, and his proof of the NP-hardness of maximum likelihood phylogenetics. In recent years Benny also worked on analyzing gene expression and genomics data, including studying techniques and complexity results related to order-preserving submatrices, as well as developing methods for analyzing HiC data and producing biologically interesting results. Benny greatly contributed to computer science education. Benny once likened the use of Python in teaching CS to the use of Toyota in driving lessons. One doesn't learn how to drive a Toyota – one learns the art and skills of driving.

We are grateful to have had the privilege to work with Prof. Benny Chor and to have learned so much from him. We are grateful for having spent many fun science hours with Benny on trails, in pubs and cafes, and on sailboats. He will always be remembered for his great science and personality. He will always be remembered with a smile.

Sagi Snir
Zohar Yakhini

Organization

General Chair

Vineet Bafna

University of California, San Diego, USA

Program Chair

Itsik Pe'er

Columbia University, USA

Steering Committee

Vineet Bafna

University of California, San Diego, USA

Bonnie Berger (Chair)

Massachusetts Institute of Technology, USA

Eleazar Eskin

University of California, Los Angeles, USA

Teresa Przytycka

National Institutes of Health, USA

Cenk Sahinalp

National Institutes of Health, USA

Roded Sharan

Tel Aviv University, Israel

Martin Vingron

Max Planck Institute for Molecular Genetics,
Germany

Program Committee

Derek Aguiar

University of Connecticut, USA

Tatsuya Akutsu

Kyoto University, Japan

Can Alkan

Bilkent University, Turkey

Srinivas Aluru

Georgia Institute of Technology, USA

Mukul S. Bansal

University of Connecticut, USA

Niko Beerenwinkel

ETH Zurich, Switzerland

Bonnie Berger

Massachusetts Institute of Technology, USA

Sebastian Böcker

Friedrich Schiller University Jena, Germany

Christina Boucher

University of Florida, USA

Dongbo Bu

Chinese Academy of Sciences, China

Sha Cao

Indiana University School of Medicine, USA

Cedric Chauve

Simon Fraser University, Canada

Brian Chen

Lehigh University, USA

Rayan Chikhi

CNRS, France

Lenore Cowen

Tufts University, USA

Barbara Di Camillo

University of Padua, Italy

Anthony Gitter	University of Wisconsin-Madison, USA
Mohammed El-Kebir	University of Illinois at Urbana-Champaign, USA
Nadia El-Mabrouk	University of Montreal, Canada
Travis Gagie	Diego Portales University, Chile
Faraz Hach	University of British Columbia and Vancouver Prostate Centre, Canada
Xin Gao	King Abdullah University of Science and Technology, Saudi Arabia
Iman Hajirasouliha	Cornell University, USA
Bjarni Halldorsson	deCODE genetics and Reykjavik University, Iceland
Fereydoun Hormozdiani	University of Washington, USA
Tao Jiang	University of California, Riverside, USA
Daisuke Kihara	Purdue University, USA
Carl Kingsford	Carnegie Mellon University, USA
Gunnar W. Klau	Heinrich Heine University Düsseldorf, Germany
Tal Korem	Columbia University, USA
Mehmet Koyuturk	Case Western Reserve University, USA
Gregory Kucherov	CNRS and LIGM, France
Jens Lagergren	KTH Royal Institute of Technology, Sweden
Benjamin Langmead	Johns Hopkins University, USA
Jingyi Jessica Li	University of California, Los Angeles, USA
Wei Vivian Li	Rutgers University, USA
Stefano Lonardi	University of California, Riverside, USA
Jian Ma	Carnegie Mellon University, USA
Jianzhu Ma	Purdue University, USA
Wenxiu Ma	University of California, Riverside, USA
Veli Mäkinen	University of Helsinki, Finland
Salem Malikic	National Institutes of Health, USA
Florian Markowetz	University of Cambridge, UK
Bernard Moret	Ecole Polytechnique Fédérale de Lausanne, Switzerland
Onur Mutlu	ETH Zurich, Switzerland
Luay Nakhleh	Rice University, USA
William Stafford Noble	University of Washington, USA
Ibrahim Numanagic	University of Victoria, Canada
Layla Oesper	Carleton College, USA
Yaron Orenstein	Ben-Gurion University, Israel
Laxmi Parida	IBM, USA
Cinzia Pizzi	University of Padua, Italy
Teresa Przytycka	National Institutes of Health, USA
Ben Raphael	Princeton University, USA

Knut Reinert	FU Berlin, Germany
Sushmita Roy	University of Wisconsin – Madison, USA
Cenk Sahinalp	National Cancer Institute, USA
Sriram Sankararaman	University of California, Los Angeles, USA
Michael Schatz	Cold Spring Harbor Laboratory, USA
Alexander Schoenhuth	Bielefeld University, Germany
Russell Schwartz	Carnegie Mellon University, USA
Mingfu Shao	Carnegie Mellon University, USA
Roded Sharan	Tel Aviv University, Israel
Ritambhara Singh	Brown University, USA
Sagi Snir	University of Haifa, Israel
Jens Stoye	Bielefeld University, Germany
Fengzhu Sun	University of Southern California, USA
Wing-Kin Sung	National University of Singapore, Singapore
Krister Swenson	CNRS and Université de Montpellier, France
Haixu Tang	Indiana University Bloomington, USA
Glenn Tesler	University of California, San Diego, USA
Tamir Tuller	Tel Aviv University, Israel
Fabio Vandin	University of Padua, Italy
Martin Vingron	Max Planck Institute for Molecular Genetics, Germany
Jerome Waldispuhl	McGill University, Canada
Sheng Wang	University of Washington, USA
Wei Wang	University of California, Los Angeles, USA
Tandy Warnow	University of Illinois at Urbana-Champaign, USA
Sebastian Will	Ecole Polytechnique, France
Jinbo Xu	Toyota Technological Institute at Chicago, USA
Min Xu	Carnegie Mellon University, USA
Yuedong Yang	Sun Yat-sen University, China
Yun William Yu	University of Toronto, Canada
Jianyang Zeng	Tsinghua University, China
Chi Zhang	Indiana University School of Medicine, USA
Louxin Zhang	National University of Singapore, Singapore
Xiuwei Zhang	Georgia Institute of Technology, USA
Jie Zheng	ShanghaiTech University, China

Additional Reviewers

Gad Abraham	Matthew Aldridge
Alyssa Adams	Mohammed Alquraishi
Ben Adcock	Bayarbaatar Amgalan
Jeffrey Adrion	Ulzee An
Jarno Alanko	Juan Andrade

Sandro Andreotti
 Maria Anisimova
 Yoann Anselmetti
 Hrayr Aprahamian
 Jesus Eduardo Rojo Arias
 Peter Arndt
 Jasmijn Baaijens
 Xin Bai
 Peter Bailey
 Dipankar Ranjan Baisya
 Alexander Baker
 Daniel Baker
 Giacomo Baruzzo
 C. J. Battey
 Brittany Baur
 Fritz Bayer
 Michael Beckstette
 Giora Beller
 Chandrima Bhattacharya
 Jeremy Bigness
 Wout Bittremieux
 Luc Blassel
 Alexander Bockmayr
 Richard Border
 Nico Borgsmüller
 Vincent Brault
 Maria Brbic
 Laurent Brehelin
 Matthew Brendel
 Sharon Browning
 Diyue Bu
 Volker Busskamp
 Patrick Cahan
 Stefan Canzar
 Marco Cappellato
 Marcello Carazzo
 Bastien Cazaux
 Giulia Cesaro
 Mark Chaisson
 Sakshar Chakravarty
 Addie Chambers
 Wennan Chang
 Samrat Chatterjee
 Ke Chen
 Ken Chen

Mandi Chen
 Runsheng Chen
 Xing Chen
 Liang Cheng
 Benjamin Chidester
 Leonid Chindelevitch
 Julien Chiquet
 Philippe Chlenski
 Kil To Chong
 Julie Chow
 Simone Ciccolella
 Phillip Clarke
 Alejandro Cohen
 Matteo Comin
 Helen Cook
 Raghuram Dandinasivara
 Anjali Das
 Bu Dechao
 James Degnan
 Mattéo Delabre
 Pinar Demetci
 Emilie Devijver
 Meleshko Dmitrii
 Norbert Dojer
 Monica Dragan
 Ruben Drews
 Yuxuan Du
 Timothy Durham
 J. N. Eberhardt
 Ghazal Ebrahimi
 Gudmundur Einarsson
 Baris Ekim
 Michael Epstein
 Michael Ford
 Anat Fuchs
 Philippe Gambette
 Shan Gao
 Lana Garmire
 Mathieu Gascon
 Alexander Gawronski
 Paul Geeleher
 Sam Gelman
 Vanesa Getseva
 Ali Ghaffari
 Maryam Ghareghani

Parham Ghasemloo Gheidari
Mazyar Ghezelji
Manoj Gopalkrishnan
Olivier Gossner
Ilan Gronau
Barak Gross
Mario Guarracino
Alessandro Guazzo
Pawel Górecki
Georg Hahn
Spencer Halberg
Gisli Halldorsson
Michiaki Hamada
Renmin Han
Wenkai Han
Marteinn Hardarson
Ananth Hari
Shabbeer Hassan
Verena Heinrich
Hadas Hezroni
Minh Hoang
Ermin Hodzic
Christian Hoener zu Siederdisen
Steve Hoffmann
Jan Hoinka
Jill Hollenbach
Yuhui Hong
Farhad Hormozdiari
Borislav Hristov
Jian Hu
Jiawei Huang
Xingfan Huang
Michael Hudgens
Pelin Icer
Anastasia Ignatieva
Elham Jafari
Mansoureh Jalilkhany
Joseph Janizek
Aleksander Jankowski
Joulien Jarroux
Cangzhi Jia
Xiaofang Jiang
Suraj Joshi
Ben Kaufman
Yasin Kaymaz

David Kelley
Raiyan Khan
Parsoa Khorsand
Kengo Kinoshita
Florian Klimm
Maren Knop
Liang Kong
Alexey Kozlov
Sandeep Krishna
Semih Kurt
Chee Keong Kwoh
Lukas Käll
Da-Inn Lee
Heewook Lee
Peter Lee
Seokho Lee
Jure Leskovec
Haipeng Li
Xiang Li
Qi Liao
Xingyu Liao
Ran Libeskind-Hadas
Antoine Limasset
Jie Liu
Kaiyuan Liu
Yining Liu
Yuelin Liu
Yu-Chen Lo
Enrico Longato
Xiaoyu Lu
Yang Lu
Shishi Luo
Xiang Ge Luo
Xiao Luo
Yunan Luo
Cong Ma
Feiyang Ma
Jianzhu Ma
Chris Magnano
Lauren Mak
Bastien Mallein
Ion Mandoiu
Camille Marchet
Parmita Mehta
Cassidy Mentus

David Merrell	Julie Rodor
Ivana Mikocziova	Caroline Ross
Ryan Mills	Oscar Rueda
Hossein Moeinzadeh	Jean-Francois Rupprecht
Mohammadhossein Moeinzadeh	Timothy Sackton
Mohammadreza Mohaghegh Neyshabouri	Armita Safa
Amirsadra Mohseni	Negar Safinianaini
Matthieu Muffato	Yana Safonova
Ghulam Murtaza	Tiziana Sanavia
Veli Mäkinen	Rebecca Santorella
Sheida Nabavi	Diego Santoro
Ozkan Nalbantoglu	Roman Sarrazin-Gendron
Rami Nasser	Itay Sason
Wilfred Ndifon	Gryte Satas
Tin Nguyen	Johannes Schlüter
Pavlos Nikolopoulos	Sebastian Schmeier
Macha Nikolski	Jacob Schreiber
Mor Nitzan	Tizian Schulz
Eirini Ntoutsis	Celine Scornavacca
Ziad Obermeyer	Bengt Sennblad
Baraa Orabi	Ekaterina Seregina
Soumitra Pal	Saleh Sereshki
Xiaoyong Pan	Bar Shalem
Enrico Paolini	Tal Shay
Luca Parmigiani	Ron Sheinin
Ali Pazokitoroudi	Hong-Bin Shen
Leonardo Pellegrina	Yihang Shen
Matteo Pellegrini	Noam Shental
David Pellow	Qian Shi
Sitara Persad	Shinsuke Shigeto
Johan Persson	Ilan Shomorony
Nico Pfeifer	Tomer Sidi
Marina Pinskaya	Brynja Sigurpalsdottir
Daniel Pirak	Nasa Sinnott-Armstrong
Petr Popov	Pavel Skums
Gunnar Pálsson	Haris Smajlovic
Yutong Qiu	Ruslan Soldatov
Yoav Ram	Reza Soltani
Vincent Ranwez	Yan Song
Nadav Rappoport	Yun Song
Farid Rashidi Mehrabadi	Vsevolod Sourkov
Anat Reiner-Benaim	Alexandros Stamatakis
Vladimir Reinharz	Leen Stougie
Stefano Rensi	Alexander Strzalkowski
Mina Rho	Pascal Sturmfels

Shiwei Sun
 Yuqi Tan
 Jing Tang
 Tianqi Tang
 Eric Tannier
 Zahin Tasfia
 Erica Tavazzi
 Fabian Theis
 Marketa Tomkova
 Cole Trapnell
 Isotta Trescato
 Xinming Tu
 Laura Tung
 Eline van Mantgem
 Jan H. van Schuppen
 Thomas Varley
 Isana Veksler
 M. P. Verzi
 Davide Verzotto
 Riccardo Vicedomini
 Gian Marco Visani
 Pieter-Jan Volders
 Dong Wang
 Jihua Wang
 Liangjiang Wang
 Tianyang Wang
 Weixu Wang
 Xuran Wang
 Ying Wang
 Leah Weber
 April Wei
 Ethan Weinberger
 Mathias Weller
 Daniel Westreich
 J. White Bear
 Roland Wittler
 Jochen Wolf
 Aiping Wu
 Chih Hao Wu
 Lani Wu
 Yi-Chieh Wu
 Yufeng Wu
 Chencheng Xu

Hanwen Xu
 Jie Xu
 Junyan Xu
 Xiaopeng Xu
 Muyu Yang
 Yang Yang
 Moran Yassour
 Esti Yeger-Lotem
 Melih Yilmaz
 Keren Yizhak
 Kaan Yorgancioglu
 Serhan Yilmaz
 Tristan Zaborniak
 Xiaofei Zang
 Ron Zeira
 Chao Zeng
 Xiangrui Zeng
 Qing Zhan
 Bin Zhang
 Chuanyi Zhang
 Louxin Zhang
 Michael Zhang
 Qimin Zhang
 Ran Zhang
 Ruochi Zhang
 Sai Zhang
 Shilu Zhang
 Wen Zhang
 Xiang Zhang
 Xining Zhang
 Xiuwei Zhang
 Yang Zhang
 Zidong Zhang
 Zijun Zhang
 Ziqi Zhang
 Yi Zhao
 Fan Zheng
 Hongyu Zheng
 Qinghui Zhou
 Tianming Zhou
 Kaiyuan Zhu
 Zifan Zhu

Contents

Extended Abstracts

Unsupervised Integration of Single-Cell Multi-omics Datasets with Disproportionate Cell-Type Representation	3
<i>Pınar Demetçi, Rebecca Santorella, Björn Sandstede, and Ritambhara Singh</i>	
Semi-supervised Single-Cell Cross-modality Translation Using Polarbear	20
<i>Ran Zhang, Laetitia Meng-Papaxanthos, Jean-Philippe Vert, and William Stafford Noble</i>	
Transcription Factor-Centric Approach to Identify Non-recurring Putative Regulatory Drivers in Cancer	36
<i>Jingkang Zhao, Vincentius Martin, and Raluca Gordân</i>	
DeepMinimizer: A Differentiable Framework for Optimizing Sequence-Specific Minimizer Schemes	52
<i>Minh Hoang, Hongyu Zheng, and Carl Kingsford</i>	
MetaCoAG: Binning Metagenomic Contigs via Composition, Coverage and Assembly Graphs	70
<i>Vijini Mallawaarachchi and Yu Lin</i>	
A Fast, Provably Accurate Approximation Algorithm for Sparse Principal Component Analysis Reveals Human Genetic Variation Across the World	86
<i>Agniva Chowdhury, Aritra Bose, Samson Zhou, David P. Woodruff, and Petros Drineas</i>	
Gene Set Priorization Guided by Regulatory Networks with p-values through Kernel Mixed Model	107
<i>Haohan Wang, Oscar L. Lopez, Wei Wu, and Eric P. Xing</i>	
Real-Valued Group Testing for Quantitative Molecular Assays	126
<i>Seyran Saeedi, Myrna Serrano, Dennis G. Yang, J. Paul Brooks, Gregory A. Buck, and Tomasz Arodz</i>	
On the Effect of Intralocus Recombination on Triplet-Based Species Tree Estimation	143
<i>Max Hill and Sebastien Roch</i>	

QT-GILD: Quartet Based Gene Tree Imputation Using Deep Learning Improves Phylogenomic Analyses Despite Missing Data	159
<i>Sazan Mahbub, Shashata Sawmya, Arpita Saha, Rezwana Reaz, M. Sohel Rahman, and Md. Shamsuzzoha Bayzid</i>	
Safety and Completeness in Flow Decompositions for RNA Assembly	177
<i>Shahbaz Khan, Milla Kortelainen, Manuel Cáceres, Lucia Williams, and Alexandru I. Tomescu</i>	
NetMix2: Unifying Network Propagation and Altered Subnetworks	193
<i>Uthsav Chitra, Tae Yoon Park, and Benjamin J. Raphael</i>	
Multi-modal Genotype and Phenotype Mutual Learning to Enhance Single-Modal Input Based Longitudinal Outcome Prediction	209
<i>Alireza Ganjdanesh, Jipeng Zhang, Wei Chen, and Heng Huang</i>	
Fast, Flexible, and Exact Minimum Flow Decompositions via ILP	230
<i>Fernando H. C. Dias, Lucia Williams, Brendan Mumey, and Alexandru I. Tomescu</i>	
Co-linear Chaining with Overlaps and Gap Costs	246
<i>Chirag Jain, Daniel Gibney, and Sharma V. Thankachan</i>	
The Complexity of Approximate Pattern Matching on de Bruijn Graphs	263
<i>Daniel Gibney, Sharma V. Thankachan, and Srinivas Aluru</i>	
ProTranslator: Zero-Shot Protein Function Prediction Using Textual Description	279
<i>Hanwen Xu and Sheng Wang</i>	
Short Papers	
Single-Cell Multi-omic Velocity Infers Dynamic and Decoupled Gene Regulation	297
<i>Chen Li, Maria Virgilio, Kathleen L. Collins, and Joshua D. Welch</i>	
Ultrafast and Interpretable Single-Cell 3D Genome Analysis with Fast-Higashi	300
<i>Ruochi Zhang, Tianming Zhou, and Jian Ma</i>	
DiffDomain Enables Identification of Structurally Reorganized Topologically Associating Domains	302
<i>Dunming Hua, Ming Gu, Yanyi Du, Li Qi, Xiangjun Du, Zhidong Bai, Xiaopeng Zhu, and Dechao Tian</i>	

Joint Inference of Repeated Evolutionary Trajectories and Patterns of Clonal Exclusivity or Co-occurrence from Tumor Mutation Trees	304
<i>Xiang Ge Luo, Jack Kuipers, and Niko Beerenwinkel</i>	
Fast and Optimal Sequence-to-Graph Alignment Guided by Seeds	306
<i>Pesho Ivanov, Benjamin Bichsel, and Martin Vechev</i>	
CLMB: Deep Contrastive Learning for Robust Metagenomic Binning	326
<i>Pengfei Zhang, Zhengyuan Jiang, Yixuan Wang, and Yu Li</i>	
Unsupervised Cell Functional Annotation for Single-Cell RNA-Seq	349
<i>Dongshunyi Li, Jun Ding, and Ziv Bar-Joseph</i>	
A Novel Matrix Factorization Model for Interpreting Single-Cell Gene Expression from Biologically Heterogeneous Data	353
<i>Kun Qian, Shiwei Fu, Hongwei Li, and Wei Vivian Li</i>	
Tractable and Expressive Generative Models of Genetic Variation Data	356
<i>Meihua Dang, Anji Liu, Xinzhu Wei, Sriram Sankararaman, and Guy Van den Broeck</i>	
CONCERT: Genome-Wide Prediction of Sequence Elements That Modulate DNA Replication Timing	358
<i>Yang Yang, Yuchuan Wang, Yang Zhang, and Jian Ma</i>	
CORSID Enables <i>de novo</i> Identification of Transcription Regulatory Sequences and Genes in Coronaviruses	360
<i>Chuanyi Zhang, Palash Sashittal, and Mohammed El-Kebir</i>	
Learning Probabilistic Protein-DNA Recognition Codes from DNA-Binding Specificities Using Structural Mappings	363
<i>Joshua L. Wetzel, Kaiqian Zhang, and Mona Singh</i>	
Uncertainty Quantification Using Subsampling for Assembly-Free Estimates of Genomic Distance and Phylogenetic Relationships	366
<i>Eleonora Rachtman, Shahab Sarmashghi, Vineet Bafna, and Siavash Mirarab</i>	
SOPHIE: Viral Outbreak Investigation and Transmission History Reconstruction in a Joint Phylogenetic and Network Theory Framework	369
<i>Pavel Skums, Fatemeh Mohebbi, Vyacheslav Tsyvina, Pelin Icer, Sumathi Ramachandran, and Yuri Khudyakov</i>	

Identifying Systematic Variation at the Single-Cell Level by Leveraging Low-Resolution Population-Level Data	371
<i>Elior Rahmani, Michael I. Jordan, and Nir Yosef</i>	
Belayer: Modeling Discrete and Continuous Spatial Variation in Gene Expression from Spatially Resolved Transcriptomics	372
<i>Cong Ma, Uthsav Chitra, Shirley Zhang, and Benjamin J. Raphael</i>	
Lossless Indexing with Counting de Bruijn Graphs	374
<i>Mikhail Karasikov, Harun Mustafa, Gunnar Rätsch, and André Kahles</i>	
Uncovering Hidden Assembly Artifacts: When Unitigs are not Safe and Bidirected Graphs are not Helpful (ABSTRACT)	377
<i>Amatur Rahman and Paul Medvedev</i>	
Mapping Single-Cell Transcriptomes to Copy Number Evolutionary Trees	380
<i>Pedro F. Ferreira, Jack Kuipers, and Niko Beerenwinkel</i>	
ImmunoTyper-SR: A Novel Computational Approach for Genotyping Immunoglobulin Heavy Chain Variable Genes Using Short Read Data	382
<i>Michael Ford, Ananth Hari, Oscar Rodriguez, Junyan Xu, Justin Lack, Cihan Oguz, Yu Zhang, Sarah Weber, Mary Magliocco, Jason Barnett, Sandhya Xirasagar, Smilee Samuel, Luisa Imberti, Paolo Bonfanti, Andrea Biondi, Clifton L. Dalgard, Stephen Chanock, Lindsey Rosen, Steven Holland, Helen Su, Luigi Notarangelo, SPS NIAID COVID Consortium, Uzi Vishkin, Corey Watson, and S. Cenk Sahinalp</i>	
AutoComplete: Deep Learning-Based Phenotype Imputation for Large-Scale Biomedical Data	385
<i>Ulzee An, Na Cai, Andy Dahl, and Sriram Sankararaman</i>	
Resistor: An Algorithm for Predicting Resistance Mutations Using Pareto Optimization over Multistate Protein Design and Mutational Signatures	387
<i>Nathan Guerin, Teresa Kaserer, and Bruce R. Donald</i>	
Ultra High Diversity Factorizable Libraries for Efficient Therapeutic Discovery	390
<i>Zheng Dai, Sachit D. Saksena, Geraldine Horny, Christine Banholzer, Stefan Ewert, and David K. Gifford</i>	
Author Index	393