GENOME RESEARCH
Volume 33 Number 7 July 2023

This Special Issue features Methods papers that were accepted for publication by Genome Research *following selection and review in coordination with the 27th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2023).*

Methods

Minimal positional substring cover is a haplotype threading alternative to Li and Stephens model 1007
Ahsan Sanaullah, Degui Zhi, and Shaojie Zhang

Fast inference of genetic recombination rates in biobank scale data 1015
Ardalan Naseri, William Yue, Shaojie Zhang, and Degui Zhi

Ultrafast genome-wide inference of pairwise coalescence times 1023OA
Regev Schweiger and Richard Durbin

Leveraging family data to design Mendelian randomization that is provably robust to population stratification 1032
Nathan LaPierre, Boyang Fu, Steven Turnbull, Eleazar Eskin, and Sriram Sankararaman

Improving quartet graph construction for scalable and accurate species tree estimation from gene trees 1042
Yunheng Han and Erin K. Molloy

A fast and scalable method for inferring phylogenetic networks from trees by aligning lineage taxon strings 1053OA
Louxin Zhang, Niloufar Abhari, Caroline Colijn, and Yufeng Wu

Deriving confidence intervals for mutation rates across a wide range of evolutionary distances using FracMinHash 1061OA
Mahmudur Rahman Hera, N. Tessa Pierce-Ward, and David Koslicki

Efficient taxa identification using a pangenome index 1069OA
Omar Ahmed, Massimiliano Rossi, Christina Boucher, and Ben Langmead

Modeling and predicting cancer clonal evolution with reinforcement learning 1078
Stefan Ivanovic and Mohammed El-Kebir

Single-cell methylation sequencing data reveal succinct metastatic migration histories and tumor progression models 1089
Yuelin Liu, Xuan Cindy Li, Farid Rashidi Mehrabadi, Alejandro A. Schäffer, Drew Pratt, David R. Crawford, Salem Malikić, Erin K. Molloy, Vishaka Gopalan, Stephen M. Mount, Eytan Ruppin, Kenneth D. Aldape, and S. Cenk Sahinalp

(continued)

Assessing transcriptomic reidentification risks using discriminative sequence models 1101[OA]
Shuvom Sadhuka, Daniel Fridman, Bonnie Berger, and Hyunghoon Cho

Enabling tradeoffs in privacy and utility in genomic data Beacons and summary statistics 1113[OA]
Rajagopal Venkatesaramani, Zhiyu Wan, Bradley A. Malin, and Yevgeniy Vorobeychik

Partial alignment of multislice spatially resolved transcriptomics data 1124
Xinhao Liu, Ron Zeira, and Benjamin J. Raphael

Unsupervised contrastive peak caller for ATAC-seq 1133[OA]
Ha T.H. Vu, Yudi Zhang, Geetu Tuteja, and Karin S. Dorman

Leveraging protein language models for accurate multiple sequence alignments 1145[OA]
Claire D. McWhite, Isabel Armour-Garb, and Mona Singh

Efficient minimizer orders for large values of k using minimum decycling sets 1154[OA]
David Pellow, Lianrong Pu, Bariş Ekim, Lior Kotlar, Bonnie Berger, Ron Shamir, and Yaron Orenstein

Entropy predicts sensitivity of pseudorandom seeds 1162[OA]
Benjamin Dominik Maier and Kristoffer Sahlin

Proving sequence aligners can guarantee accuracy in almost $O(m \log n)$ time through an average-case analysis of the seed-chain-extend heuristic 1175[OA]
Jim Shaw and Yun William Yu

Efficient mapping of accurate long reads in minimizer space with mapquik 1188[OA]
Bariş Ekim, Kristoffer Sahlin, Paul Medvedev, Bonnie Berger, and Rayan Chikhi

Extremely fast construction and querying of compacted and colored de Bruijn graphs with GGCAT 1198[OA]
Andrea Cracco and Alexandru I. Tomescu

Aligning distant sequences to graphs using long seed sketches 1208
Amir Joudaki, Alexandru Meterez, Harun Mustafa, Ragnar Groot Koerkamp, André Kahles, and Gunnar Rätsch

Genomic sketching with multiplicities and locality-sensitive hashing using Dashing 2 1218[OA]
Daniel N. Baker and Ben Langmead

[OA] Open Access paper

Cover "Moiré By the Sea." A moiré pattern is an interference pattern that results from overlapping graphical elements, such as lines or dots. In the cover image, the overlapping lines in the ocean and overlapping dots in the clouds produce moiré patterns. In this issue, the PASTE2 algorithm to align spatial transcriptomics data from adjacent tissue slices is described. Spatial transcriptomics technologies measure gene expression at thousands of discrete locations, or spots, and thus images of aligned slices often yield moiré patterns (e.g., see Fig. 3 in Liu et al.). (Cover illustration by Lillian Raphael, www.lilymraphael.com. The cover art was generated using licensed versions of Procreate and Adobe Illustrator. Positions of points in clouds were computed in Mathematica using code from Stephen Wolfram, "Moiré Patterns," http://demonstrations.wolfram.com/MoirePatterns/, Wolfram Demonstrations Project [published: March 7, 2011]. [For details, see Liu et al., pp. 1124–1132.])