Jim Shaw

jshaw@math.toronto.edu — github — google scholar – personal website

EDUCATION

PhD in Mathematics — Expected 2024
University of Toronto
Supervisor: Yun William Yu

MSc in Mathematics — Received 2020
University of Toronto
Supervisor: Yun William Yu

BASc in Engineering Physics — Received 2019
Minor in Honours Mathematics
University of British Columbia

RESEARCH SUMMARY

My research is on developing algorithms, tools, and theory for large-scale analysis of biological sequences. The overall goal is to accelerate fundamental genomics research; specific current applications include metagenomics and microbiome informatics.

PUBLICATIONS, PROCEEDINGS, AND PREPRINTS

Publications are listed in order of preprinting date. Proceedings in computational biology conferences sometimes have partner journals. In this case, only the journal version is listed.

* indicates co-first authors. **Bolded** is me.

1. **J. Shaw***, Y.W. Yu. Fairy: fast approximate coverage for multi-sample metagenomic binning (2024). bioRxiv.
2. **J. Shaw***, J.S. Gounot*, H. Chen, N. Nagarajan, Y.W. Yu. Floria: Fast and accurate strain haplotyping in metagenomes (2024). Accepted to Bioinformatics (ISMB conference associated).
3. **J. Shaw**, Y.W. Yu. Metagenome profiling and containment estimation through abundance-corrected k-mer sketching with sylph (2023). bioRxiv.
4. **J. Shaw**, Y.W. Yu. Fast and robust metagenomic sequence comparison through sparse chaining with skani (2023). Nature Methods.
5. A. Zheng, **J. Shaw**, Y.W. Yu. Mora: abundance aware metagenomic read re-assignment for disentangling similar strains (2022). bioRxiv.
6. **J. Shaw**, Y.W. Yu. Proving sequence aligners can guarantee accuracy in almost O(m log n) time through an average-case analysis of the seed-chain-extend heuristic (2023). Genome Research. (RECOMB conference associated)
7. M.C. Frith, **J. Shaw**, J. Spouge. How to optimally sample a sequence for rapid analysis (2023). Bioinformatics.
8. **J. Shaw**, Y.W. Yu. Theory of local k-mer selection with applications to long-read alignment (2021). Bioinformatics.
9. **J. Shaw**, Y.W. Yu. flopp: Extremely Fast Long-Read Polyploid Haplotype Phasing by Uniform Tree Partitioning (2022). Journal of Computational Biology. (RECOMB conference associated)

10. R. Cotsakis*, **J. Shaw***, J. Tierny, J. Levine. Implementing Persistence-Based Clustering of Point Clouds in the Topology ToolKit (2020). Topological Methods in Visualization: Theory, Software and Applications.
11. S. Hu*, O. Schnetz*, **J. Shaw***, K. Yeats*. Further investigations into the graph theory of ϕ^4-periods and the c_2 invariant. (2020). Annales de l'Institut Henri Poincare D.
12. D. Bertrand, **J. Shaw**, M Narayan, H.Q.A. Ng, S. Kumar, C. Li, M. Dvornicic, J.P. Soldo, J.Y. Kho, O.T. Ng, T. Barkham, B. Young, K. Marimuthu, K.R. Chng, M. Sikic, N. Nagarajan. Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes (2019). Nature Biotechnology.

ACCEPTED CONFERENCE TALKS

This section is for talks that have been submitted, reviewed by a conference committee, and accepted for oral presentation. Some are proceedings associated while others are not.

1. ISMB, Montreal, Canada. 2024. (Proceedings)
2. Great Lakes Bioinformatics Conference, Pittsburgh, USA. 2024. (Talk only)
3. Genome Informatics, Cold Spring Harbor, USA. 2023. (Talk only)
4. ISMB, Lyon, France. 2023. (Talk only)
5. RECOMB, Istanbul, Turkey. 2023. (Proceedings)
6. RECOMB, Padova, Italy (virtual). 2021. (Proceedings)

SEMINARS AND TALKS

1. Indian Institute of Science. Department of Computational and Data Sciences Seminar. Virtual, 2024.
2. University of Waterloo. Computer Science Seminar. Waterloo, Canada, 2024.
3. University of Tokyo. Computational Biology Seminar. Tokyo, Japan. 2023.
4. University of Toronto. Combinatorics student seminar. Toronto, Canada. 2022.

PROFESSIONAL EXPERIENCE AND INTERNSHIPS

Visiting PhD student Jan 2024 - June 2024
Carnegie Mellon University, Pittsburgh, USA
Advised by Yun William Yu.
- Continuing Ph.D. research with my advisor after they switched universities.

Visiting computational biology researcher Sept 2023 - Dec 2023
University of Tokyo, Tokyo, Japan
Advised by Martin C. Frith.
- Researched high-performance and statistically sound DNA sequence analysis algorithms.

Data Science Internship May 2019 - August 2019
DeepND, Vancouver, BC, Canada
- Developed data science methods for analyzing employee churn and retention for a startup and presented results to industry clients.

Mathematics Research Intern May 2018 - September 2018
University of Waterloo, Waterloo, ON, Canada
Advised by Karen Yeats.

- Researched algebraic graph theory for graphs arising from quantum field theories and developed high-performance symbolic algebra algorithms in C++/Maple.

Computational Genomics Intern May 2017 - September 2017
Genome Institute of Singapore, Singapore
Advised by Niranjan Nagarajan.

- Researched computational problems arising from microbial community sequencing. Assisted in the development of a metagenomic genome assembler using graphical and statistical techniques.

Software Engineer Intern January 2016 - May 2016
Zaber Technologies, Vancouver, BC, Canada

- Developed graphical user interfaces in C# for interfacing with engineering hardware, subject to engineering best practices such as documentation and testing.

POSTERS

1. RECOMB. Boston, USA. 2024. **(Best poster award: 2/284 selected)**.
2. ISMB. Lyon, France. 2023.
3. RECOMB. San Diego, USA. 2022.

PROFESSIONAL SERVICE

1. Secondary reviewer for ISMB 2023 and RECOMB 2022, 2023, and 2024 conferences.
2. Reviewer for Bioinformatics Advances, Journal of Computational Biology, BMC Bioinformatics.

HONOURS AND AWARDS

This section is roughly sorted by competitiveness/prestige.

1. NSERC-CGS D Scholarship (2022) - Awarded 105,000 over 3 years through a national PhD scholarship competition.
2. RECOMB 2024 Best Poster Award (**2/284 selected**)
3. NSERC Michael Smith Foreign Supplement (2023) - Awarded $6000 to travel to Japan for research.
4. Singapore International Pre-Graduate Award (2017) - $6000 awarded to pursue a research internship in Singapore for four months.
5. PhD Entrance Scholarship (2020) - Awarded $5000 for performance in the MSc of Mathematics program.
6. Trek Excellence Award (2017) - Awarded $1500 for being ranked in the top 5% of students in the Faculty of Applied Science at the University Of British Columbia for the 2016-2017 year.
7. NSERC Undergraduate Summer Research Award (2018) - $4500 awarded to pursue summer research at the University of Waterloo.
8. Donald J. Evans Scholarship in Engineering (2017) - $500 awarded on the recommendation of the Faculty of Applied Science at the University of British Columbia.
9. NSERC Industrial Undergraduate Student Research Award (2016) - $4500 awarded to pursue an industrial internship.

TEACHING All teaching assistant positions involved a combination of leading tutorials, marking, and office hours.

1. MAT 135 Calculus 1 TA: Fall 2019, Winter 2022.
2. MAT 136 Calculus 2 TA: Winter 2020
3. MAT 223 Linear Algebra 1 TA: Fall 2020, Fall 2021, Winter 2022
4. MAT 224 Linear Algebra 2 TA: Winter 2021
5. MAT learning centre TA: Fall 2019, Fall 2022, Winter 2023.