

Jim Shaw

jshaw@ds.dfci.harvard.edu — <https://jim-shaw-bluenote.github.io>

ACADEMIC POSITIONS	Postdoctoral researcher Harvard Medical School (Biomedical Informatics) & Dana-Farber Cancer Institute (Data Science)	September 2024 - Now Advisor: Heng Li
--------------------	---	--

EDUCATION	PhD in Mathematics University of Toronto	Received 2024 Advisor: Yun William Yu
	Master's in Mathematics University of Toronto	Received 2020 Advisor: Yun William Yu
	BASc in Engineering Physics <i>Minor in Honours Mathematics</i> University of British Columbia	Received 2019

RESEARCH SUMMARY	I develop theoretically-sound algorithms and software for large-scale analysis of biological sequences. My applied interests lie in understanding the latent complexity of microbiomes through the application of novel genomics analysis algorithms.
------------------	---

PUBLICATIONS, PROCEEDINGS, AND PREPRINTS	Proceedings in computational biology conferences can have associated journals. In this case, the published journal version is listed with the associated conference.  * indicates co-first authors.
--	---

1. **J. Shaw**, M.G. Marin, H. Li. High-resolution metagenome assembly for modern long reads with myloasm (2025). bioRxiv.
2. **J. Shaw**, C. Boucher, Y.W. Yu, N. Noyes, H. Li. devider: long-read reconstruction of many diverse haplotypes (2024). Accepted to RECOMB 2025.
3. X.B. Zhang, G. Oualline, **J. Shaw**, Y.W. Yu. skandiver: a divergence-based analysis tool for identifying intercellular mobile genetic elements (2024). Bioinformatics (ECCB conference associated).
4. **J. Shaw**, Y.W. Yu. Fairy: fast approximate coverage for multi-sample metagenomic binning (2024). Microbiome.
5. **J. Shaw\***, J.S. Gounot\*, H. Chen, N. Nagarajan, Y.W. Yu. Floria: Fast and accurate strain haplotyping in metagenomes (2024). Bioinformatics (ISMB conference associated).
6. **J. Shaw**, Y.W. Yu. Rapid species-level metagenome profiling and containment estimation with sylph (2024). Nature Biotechnology.
7. **J. Shaw**, Y.W. Yu. Fast and robust metagenomic sequence comparison through sparse chaining with skani (2023). Nature Methods.
8. A. Zheng, **J. Shaw**, Y.W. Yu. Mora: abundance aware metagenomic read re-assignment for disentangling similar strains (2024). BMC Bioinformatics.
9. **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)

10. M.C. Frith, **J. Shaw**, J. Spouge. How to optimally sample a sequence for rapid analysis (2023). Bioinformatics (GIW/ISCB Asia conference associated).
11. **J. Shaw**, Y.W. Yu. Theory of local k-mer selection with applications to long-read alignment (2021). Bioinformatics.
12. **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)
13. 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 (Conference proceedings).
14. S. Hu\*, O. Schnetz\*, **J. Shaw\***, K. Yeats\*. Further investigations into the graph theory of  $\phi^4$ -periods and the  $c_2$  invariant. (2020). Annales de l'Institut Henri Poincaré D.
15. 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

These talks were reviewed by a conference committee and accepted for oral presentation. Some are associated with a proceedings paper while others are not.

1. RECOMB-Microbiome, Seoul, South Korea. 2025. (Highlights talk)
2. RECOMB, Seoul, South Korea. 2025. (Proceedings)
3. ISMB, Montreal, Canada. 2024. (Proceedings)
4. Great Lakes Bioinformatics Conference, Pittsburgh, USA. 2024. (Highlights talk)
5. Genome Informatics, Cold Spring Harbor, USA. 2023. (Highlights talk)
6. ISMB, Lyon, France. 2023. (Highlights talk)
7. RECOMB, Istanbul, Turkey. 2023. (Proceedings)
8. RECOMB, Padova, Italy (virtual). 2021. (Proceedings)

#### INVITED SEMINARS AND TALKS

1. Brigham and Women's Hospital. Internal Microbiome Seminar. Boston, U.S., 2025.
2. Microbiome Virtual International Forum. Short talk. Virtual, 2024. (**Best short talk award: \$200 prize**).
3. Indian Institute of Science. Department of Computational and Data Sciences Seminar. Virtual, 2024.
4. University of Waterloo. Computer Science Seminar. Waterloo, Canada, 2024.
5. University of Tokyo. Computational Biology Seminar. Tokyo, Japan, 2023.
6. University of Toronto. Combinatorics student seminar. Toronto, Canada, 2022.

<b>PROFESSIONAL EXPERIENCE AND INTERNSHIPS</b>	<b>Visiting PhD Student</b>	Jan 2024 - June 2024
	<i>Carnegie Mellon University</i> , Pittsburgh, USA Advised by Yun William Yu.	
	<b>Visiting Computational Biology Researcher</b>	Sept 2023 - Dec 2023
	<i>University of Tokyo</i> , Tokyo, Japan Advised by Martin C. Frith.	
	<b>Data Science Internship</b>	May 2019 - August 2019
	<i>DeepND</i> , Vancouver, BC, Canada	
	<b>Mathematics Research Intern</b>	May 2018 - September 2018
	<i>University of Waterloo</i> , Waterloo, ON, Canada Advised by Karen Yeats.	
	<b>Computational Genomics Intern</b>	May 2017 - September 2017
	<i>Genome Institute of Singapore</i> , Singapore Advised by Niranjan Nagarajan.	
	<b>Software Engineer Intern</b>	January 2016 - May 2016
	<i>Zaber Technologies</i> , Vancouver, BC, Canada	

## POSTERS

1. CSHL Microbiome. Cold Spring Harbor Laboratory, USA. 2024
2. RECOMB. Boston, USA. 2024. (**Best poster award: 2/284 selected**).
3. ISMB. Lyon, France. 2023.
4. RECOMB. San Diego, USA. 2022.

## ACADEMIC SERVICE

### Journal reviewing:

- Cell Genomics
- Genome Biology
- Bioinformatics Advances
- PeerJ
- BMC Bioinformatics
- Briefings in Bioinformatics
- GigaScience
- Microbial Genomics
- Communications Biology

### Conference reviewing:

- ACM-BCB 2025 (primary reviewer)
- ISMB 2023 (secondary reviewer)
- RECOMB 2022, 2023, 2024, 2025 (secondary reviewer)

## HONOURS AND AWARDS

This section is roughly sorted by competitiveness/prestige.

1. NSERC Postdoctoral Fellowship (2024) - \$140,000 awarded over 2 years through a national fellowship competition.
2. NSERC CGS-D Scholarship (2022) - \$105,000 awarded over 3 years through a national PhD scholarship competition.
3. Malcolm Slingsby Robertson Prize (2025) - \$600 awarded to the graduating student who best demonstrated research excellence within the University of Toronto's Mathematics Department.
4. RECOMB 2024 Best Poster Award (**2/284 selected**)
5. Microbiome Virtual International Forum Best Short Talk Award - \$200 awarded.
6. NSERC Michael Smith Foreign Supplement (2023) - \$6000 awarded to travel to Japan for research.
7. PhD Entrance Scholarship (2020) - \$5000 awarded for performance in the MSc of Mathematics program.
8. Trek Excellence Award (2017) - \$1500 awarded 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.
9. Singapore International Pre-Graduate Award (2017) - \$6000 awarded to pursue a research internship in Singapore for four months.
10. NSERC Undergraduate Summer Research Award (2018) - \$4500 awarded to pursue summer research at the University of Waterloo.
11. Donald J. Evans Scholarship in Engineering (2017) - \$500 awarded on the recommendation of the Faculty of Applied Science at the University of British Columbia.
12. 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 and recitations, marking, and office hours.

### University of Toronto

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.