Xiran Liu

Email: xiran_liu1@brown.edu | Phone: (+1)314-421-9681
GitHub: github.com/xr-cc | Website: xr-cc.github.io

EDUCATION AND TRAINING

Postdoc	Data Science Institute & Center for Computational Molecular Biology, Brown University	Oct 2023-
	Advisor: Dr. Sohini Ramachandran	
Ph.D.	Computational and Mathematical Engineering, Stanford University	Sep 2023
	Dissertation: "Computational Methods and Mathematical Measures for Population Relationships" *Advisor: Dr. Noah Rosenberg* *Thesis Committee: Dr. Marcus Feldman, Dr. Julia Palacios*	
M.S.	Computational and Mathematical Engineering, Stanford University	Sep 2020
B.S.	Systems Science and Engineering, Washington University, St. Louis, *Summa Cum Laude*	May 2018
B.S.	Computer Science, Washington University, St. Louis, *Summa Cum Laude*	May 2018

AWARDS AND FELLOWSHIPS

Rising Stars in Data Science Workshop 2025 Cohort, hosted by Stanford University in collaboration with the University of California, San Diego, and the University of Chicago; $1,000 travel award (2025)

Gene Golub Dissertation Award, Institute for Computational and Mathematical Engineering, Stanford University; awarded to 1-3 recipients annually (2023)

CEHG Predoctoral Fellowship, Center for Computational, Evolutionary and Human Genomics, Stanford University; one-quarter fellowship including $17,173 total support (2022–2023)

The Enlight Foundation Graduate Fellowship, Stanford University; departmental fellowship supporting first-year doctoral students (2018–2019)

Dean's List, Outstanding Sophomore, Junior, and Senior Awards, and Undergraduate Research Scholarship, Washington University (2014–2018)

RESEARCH INTERESTS

Data science; machine learning; computational biology; bioinformatics; AI for science; development of scalable computational frameworks to extract patterns, model complex relationships, and interpret biological variation; translational method innovations with cross-domain adaptation; data visualization.

PUBLICATIONS

* denotes corresponding author(s).

Peer-Reviewed Journal Articles

Liu X*, Crawford L, Ramachandran S (2025). ML-MAGES enables multivariate genetic association analyses with genes and effect size shrinkage. *Genome Res. (accepted)*

Liu X*, Ahsan Z, Rosenberg NA (2025). Using mathematical constraints to explain narrow ranges for allele-sharing dissimilarities. *Theor. Popul. Biol*. *(in press)*. doi:10.1016/j.tpb.2025.05.002

Liu X*, Kopelman N, Rosenberg NA* (2024). Clumppling: cluster matching and permutation program with integer linear programming. *Bioinformatics* 40(1):btad751. doi:10.1093/bioinformatics/btad751

Liu X*, Ahsan Z, Martheswaran T, Rosenberg NA (2023). When is the allele-sharing dissimilarity between two populations exceeded by the allele-sharing dissimilarity of a population with itself?. *Stat. Appl. Genet. Mol. Biol.* 22(1):20230004. doi:10.1515/sagmb-2023-0004

Liu X, Kopelman N, Rosenberg NA* (2022). A Dirichlet model of alignment cost in mixed-membership unsupervised clustering. *J. Comput. Graph. Stat.* 32(3):1145-1159. doi:10.1080/10618600.2022.2127739

Liu X, Rosenberg NA*, Greenbaum G* (2022). Extracting hierarchical features of cultural variation using network-based clustering. *Evol. Hum. Sci.* 4:e18. doi:10.1017/ehs.2022.15

Parikh VN, Ioannidis AG, Jimenez-Morales D, Gorzynski JE, De Jong HN, Liu X, Roque J, Cepeda-Espinoza VP, *et al.* (2022) Deconvoluting complex correlates of COVID-19 severity with local ancestry inference and viral phylodynamics: Results of a multi-omic pandemic tracking strategy. *Nat. Comm. 13*(1): 1-10. doi:10.1038/s41467-022-32397-8

Liu X, Feldman MW* (2021). Effects of cultural transmission of surnaming decisions on the sex ratio at birth. *Theor. Popul. Biol.* 141:44-53. doi:10.1016/j.tpb.2021.07.001

Ye Z, Price RL, Liu X, *et al.* (2020). Diffusion histology imaging detects and classifies glioblastoma pathology missed by conventional magnetic resonance imaging. *Clin. Cancer Res. 26*(20):5388-5399. doi:10.1158/1078-0432.CCR-20-0736

Conference Paper

Liu X*, Crawford L, Ramachandran S (2025). ML-MAGES: A machine learning framework for multivariate genetic association analyses with genes and effect size shrinkage. In: Sankararaman, S. (eds), *Research in Computational Molecular Biology (RECOMB 2025)*. *LNCS*, Vol. 15647. doi:10.1007/978-3-031-90252-9_31

Works in Progress

Liu X, Singh R, Ramachandran S. Clustering alignment for single-cell clustering analysis facilitates model comparison and characterization of clustering-informative genes contributing to cellular heterogeneity. *(in prep)*

Ahsan Z, Liu X, Rosenberg NA. Combinatorics of a dissimilarity measure for pairs of draws from discrete probability vectors on finite sets of objects. (*under review; preprint on arXiv*) doi:10.48550/arXiv.2410.00221

Liu X, Greenbaum G, Rosenberg NA. Can migration be inferred from pairwise F_{ST}? *(in prep)*

Guo A, Ramachandran S*, Liu X*. *KAlignedoscope*: an interactive visualization tool for aligned clustering results. *(under review)*

Liu X, Rosenberg NA, Ramachandran S. *Clumppling 2.0*: A Clustering Alignment Program for Population Structure Analyses. *(in prep)*

SOFTWARE AND TOOLS

Clumppling: a distribution Python package for clustering alignment for population structure inference and general unsupervised clustering. Link: github.com/ramachandran-lab/KAlignedoscope

KAlignedoscope: a JavaScript-based distribution Python package for interactive visualization of aligned clustering results. Link: github.com/PopGenClustering/Clumppling

ML-MAGES: a Python-based tool for shrinking and categorizing multi-trait genetic association effects. Link: github.com/ramachandran-lab/ML-MAGES

CONFERENCE PRESENTATIONS

Oral Presentations

"Clustering alignment for single cell analyses: streamlining model comparison and revealing informative genes." *Genome Informatics 2025*. Cold Spring Harbor Laboratory, NY, Nov 2025

"*ML-MAGES:* A machine learning framework for multivariate genetic association analyses with genes and effect size shrinkage." *The Research in Computational Molecular Biology (RECOMB) 2025 Conference*. Seoul, Korea, Apr 2025.

"*Clumppling*: a new method for aligning replicate solutions in population structure analysis." *Evolution 2023 Conference*. Albuquerque, NM, Jun 2023.

"A Dirichlet model of alignment cost in mixed-membership clustering results of ancestry inference." *Bay Area Population Genomics Conference*. UC Berkeley, Berkeley, CA, Oct 2022.

Poster Presentations

"Clustering alignment for single cell analysis facilitates model comparison and characterization of informative genes contributing to cellular heterogeneity." *The American Society of Human Genetics 2025 Annual Meeting (ASHG 2025)*. Boston, MA, Oct 2025.

"*ML-MAGES*: a computationally efficient tool for multivariate genetic association analyses with effect size shrinkage via machine learning." *ASHG 2024*. Denver, CO, Nov 2024. (**Reviewers' Choice Abstract** - *top 10%.*)

"Mathematical properties of allele-sharing dissimilarities." *The Allied Genetics Conference (TAGC) 2024*. Metro Washington, DC, Mar 2024.

"Modeling alignment cost in mixed-membership unsupervised genetic clustering." *Population, Evolutionary, and Quantitative Genetics Conference (PEQG) 2022*. Pacific Grove, CA, May 2022.

"Optimizing the quality-cost trade-off of human annotation for labeling webpages to train webpage classifiers." (*Sucar E*, **Liu X***, Khillan S*, Li Z**) *Nebraska Conference for Undergraduate Women in Mathematics* (*NCUWM*). Lincoln, NE, Feb 2017.

"Optimizing the quality-cost trade-off of human annotation for labeling webpages to train webpage classifiers." (*Khillan S*, Sucar E*, **Liu X***, Li Z**) AMS Contributed Paper Session on Undergraduate Research at *Joint Mathematics Meetings*. Atlanta, GA, Feb 2017. *(Poster accepted; unable to attend.)*

ADDITIONAL PROJECT EXPERIENCE

Improving Fundraiser Efficacy Using Large Language Models, Data Scientist Intern
PayPal Global Data Science Group, San Jose, CA   2022
- Built tree-based models on text-derived features and applied explainable AI techniques to interpret results and generate actionable writing suggestions for fundraiser posts.
- Trained transformer-based large language models for text classification and guided text generation to suggest higher-quality headlines.
- Delivered a baseline model improving one-third of fundraisers and an advanced model achieving improvements in over half of the cases.

Spatial Analysis of Relationships Between Environmental and Socioeconomic Factors, Team Project
Data Science for All (DS4A): Women's Summit, Correlation One   2020
- Performed geographically weighted regression and spatial cross-correlation analysis to analyze spatial relationships between public health risks, socioeconomic disparities, and seismic hazards across Bay Area.
- Visualized correlation between various factors and provided data-driven insights for local policy making.

TEACHING EXPERIENCE

Instructor, Data Science, Computation, and Visualization (DSCoV) Workshop, *Brown University*   2025
- Prepare materials and deliver a workshop on topic modeling and its applications within and beyond NLP.

Instructor, Advanced MATLAB for Scientific Computing, *Stanford University*   Winter 2022, 2023
- A 4-week short course for students of all levels interested in advanced MATLAB features and toolboxes.

- Designed course syllables, prepared materials, and delivered lectures both over Zoom and in class.
- Open-sourced teaching materials on MATLAB File Exchange with >1.3K downloads (as of Oct 2025).

Graduate Teaching Assistant, *Stanford University*
Linear Algebra with Application to Engineering Computations (CME 200/ME 300A)
 Dr. Parviz Moin, Department of Mechanical Engineering, Stanford University   Fall 2021
 Dr. Margot Gerritsen, ICME, Stanford University   Fall 2020
 <u>Overall effectiveness rating in course evaluation</u>: average of 4.31/5 from 161/173 responses over 2 offerings
- A lecture-style course for approximately 90 graduate students in engineering.
- Held office hours, led recitation sessions, designed exercise problems and graded problem sets and exams.

Ordinary Differential Equations for Engineers (CME 102)   2020-22
 Dr. Hung Le, ICME, Stanford University
 <u>Overall effectiveness rating in course evaluation</u>: average of 4.26/5 from 97/133 responses over 3 offerings
- A lecture-style course for approximately 50 undergraduate students in engineering.
- Held office hours, led recitation sessions, designed exercise problems and graded problem sets and exams.

Summer Workshop on Intro to Machine Learning, *ICME*   2021

Peer Mentor, *School of Engineering, Washington University in St. Louis*   2015-17

Undergraduate Teaching Assistant, *School of Engineering, Washington University in St. Louis* 2015-17
 Engineering Freshman Seminar (EN 120, Course Assistant)
 Algorithms and Data Structure (CSE 241/247)
 Introduction to Electrical Engineering (ESE 103)
 Signals and Systems (ESE 351)

MENTORING EXPERIENCE

Undergraduate Research Mentor
- Mentor undergraduate students in project design, methodological reasoning, and scientific communications.
- Foster independent learning and professional growth, including the acquisition of computational and research skills.

Ramachandran Lab, Brown University   *2025-*
- Co-supervised an undergraduate student (Applied Math and Design) in a summer research project on advancing interactive visualization for clustering alignment in population genetics.
- Led the project toward a manuscript currently in preparation.

Rosenberg Lab, Stanford University   *2022-23*
- Co-supervised two undergraduate students (Math; CS) in summer and academic-year research projects on understanding the mathematical properties of genetic diversity measures and demonstrating the theoretical findings with empirical data analysis.
- Led the projects toward two published manuscripts and one preprint.

LEADERSHIP AND SERVICE

Ad hoc Reviewer: *PNAS*, *Theor. Popul. Biol.*

Reviewer (GSA Journals Peer Review Training Program), *Genetics*   2021-23

Ambassador and Organizing Committee Member, *Women in Data Science (WiDS) at Brown University* 2024-
- Co-organized the Women in Data Science (WiDS) Providence 2024 Conference.
- Planed the event agenda and invited speakers for talks and panels.
- Brought over 100 students and community members for a day of panels, seminars, and networking events.

International Chair, *Brown Postdoc Council*   2024-
  - Organized social and networking events for postdoctoral researchers.

Reviewer, 2024 *SACNAS* NDiSTEM Conference   2024

Member, *ICME Student Action Group, Stanford University*   2021-23
  - Proposed improvements for the ICME program, discussed proposals with faculty and staff, and monitored the progress of proposal implementation.
  - Resolved concerns in the general student body and organized student community events.

Member and Secretary, *Alpha Omega Epsilon Sorority Beta Xi Chapter, Washington University*   2016-18

Program Assistant, *International Summer Study Programs, Washington University*   2015, 2017

Volunteer, *Campus Y, Washington University*   2015-16

SKILLS

Programming Languages: Python, R, MATLAB, C++, MySQL, Mathematica, JavaScript, HTML

Software, Packages, and Tools: Git, Slurm, PyTorch, TensorFlow, Hugging Face, PLINK, BEAST 2

Languages: Chinese, English, German (basic)