

Mihir Bafna

mihirb14@mit.edu | mihirbafna.github.io | [LinkedIn](https://www.linkedin.com/in/mihir-bafna/) | [GitHub](https://github.com/mihirbafna) | [Google Scholar](https://scholar.google.com/citations?user=HgkzgQAAAAJ&hl=en)

EDUCATION

Massachusetts Institute of Technology Ph.D. in Electrical Engineering & Computer Science

Cambridge, MA
Sep 2024 - present

- Generative modeling for protein design, molecular dynamics, and single cell omics
- Awarded NSF GRFP Fellowship (2024)

Georgia Institute of Technology B.S. in Computer Science (Machine Learning & Theory)

Atlanta, GA
Aug 2020 - Dec 2023

- GPA: 4.0
- PURA Salary/Travel Awards for research in GNN methods for CCI/GRN prediction (Xiuwei Zhang Lab)
- Head TA for Graduate Deep Learning (CS 4644/7643)
- Undergraduate Research Ambassador

PUBLICATIONS

† indicates equal contribution

- **Learning residue level protein dynamics with multiscale Gaussians**
M. Bafna, B. Jing, B. Berger
ICLR 2026 and MLSB 2025
- **SWITCHCRAFT: Programmatic Design of State-Switching Proteins**
B. Jing†, M. Bafna†, B. Berger
MLSB 2025 (Oral)
- **Sparse autoencoders uncover biologically interpretable features in protein language model representations**
O. Gujral, M. Bafna, E. Alm, B. Berger
Proceedings of the National Academy of Sciences 2025
- **Generating functional and multistate proteins with a multimodal diffusion transformer**
B. Jing†, A. Sappington†, M. Bafna, R. Shah, A. Tang, R. Krishna, A. Klivans, D. Diaz, B. Berger
MLSB 2025 and under review
- **DiffRNAFold: Generating RNA Tertiary Structures with Latent Space Diffusion**
M. Bafna, V. Keerthipati, S. Kanaparthi, R. Zhang.
NeurIPS MLSB 2023
- **CLARIFY: Cell-cell interaction and gene regulatory network refinement from spatially resolved transcriptomics**
M. Bafna, H. Li, X. Zhang
ISMB & Bioinformatics 2023
- **DeepViFi: detecting oncoviral infections in cancer genomes using transformers**
U. Rajkumar, S. Javadzadeh, M. Bafna, D. Wu, R. Yu, J. Shang, V. Bafna
ACM-BCB 2022
- **Computer-implemented methods for quantitation of features of interest in whole slide imaging**
N. Nguyen, L. Mora-Blanco, K. Turner, J. Wiese, J. Christiansen, M. Bafna
Provisional patent PCT/US2021/022308 2021

PREPRINTS

1. DANGO: Predicting higher-order genetic interactions

R. Zhang, M. Bafna, J. Ma, J. Ma

Accepted in principle to Cell Systems

WORK / RESEARCH EXPERIENCE

Bonnie Berger Lab | Massachusetts Institute of Technology

Graduate Research Assistant

- Generative modeling for proteins

Cambridge, MA

Sep 2024 – present

Liquid AI

Machine Learning Scientist

Boston, MA

Jun 2024 – Present

- Developing in house protein language models

Bonnie Berger Lab | Massachusetts Institute of Technology

Undergraduate Research Assistant

Cambridge, MA

May 2023 – Jan 2024

- Diffusion models for genome structure (via scHi-C contact map generation conditioned on methylation data) and protein structure conformational change.
- Accepted into the Broad Summer Research Program (BSRP) starting June

Xiuwei Zhang Lab | Georgia Institute of Technology

Undergraduate Research Assistant

Atlanta, GA

Nov 2021 – Dec 2023

- Developed *Clarify*: a graph autoencoder based tool for refining extracellular AND intracellular interactions (cell-cell inference, gene regulatory network prediction, subnetwork comparison) using spatial transcriptomics
 - * Paper accepted to ISMB
 - * Best Poster Award at AWSOM

Jian Ma Lab | Carnegie Mellon University

Undergraduate Research Assistant

Pittsburgh, PA

Dec 2021 – Aug 2022

- Hypergraph Graph Neural Network model architecture (hyperSAGNN) for understanding genetic interactions, specifically trigenic, and enrichment in protein complex formation
 - * Predicted 400,000 new trigenic interactions in yeast and bounded each prediction with model uncertainty using Gaussian Processes
 - * Using AlphaFold2 and ESMFold2 Protein Sequence Embeddings to guide model in predicting trigenic interactions that are enriched in protein complex formation.

Boundless Bio

ML/Bioinformatics Research Intern

San Diego, CA

Aug 2020 – Dec 2021

- **Provisional patent** for creating metaDetect: a computer vision algorithm for identifying metaphase spreads in stained whole slide images of cancer cells using image analysis filtering/techniques.
- Created automated pipeline for whole slide imaging → metaDetect (CV) → ecDNA quantification (CNN)

AWARDS/HONORS

NSF GRFP Fellowship | (Graduate Research Fellowship Award)

April 2023

Georgia Tech Faculty Honors | (4.0 GPA)

Awarded Every Semester

Atlanta Workshop for Single-cell Omics (AWSOM) | Best Poster Award

April 2023

Broad Summer Research Program (BSRP) | Accepted

March 2023

Harvard Summer Institute for Biomedical Informatics (SBMI) | Accepted

March 2023

CS 7643 Best Project Award | First Place Winner

Dec 2022

Georgia Tech PURA | Presidential Undergraduate Research Award (\$1500 Stipend)

May 2022

Georgia Tech PURA | PURA Travel Award (\$1000)

June 2023

HealthTech Hacks @ GT | First Place Winner

March 2021

Grand Challenges @ GT | First Place Winner for best project idea

May 2021

Westview HS Honors | "Top Math Student" in HS Graduating Class (600)

June 2020

Westview HS Honor Chain | Top 3% GPA in HS Graduating Class (600)

June 2020

TALKS

- **ISMB Oral** CLARIFY: Cell-cell interaction and gene regulatory network refinement from spatially resolved transcriptomics. July 2023.
- **Zhang Lab Group Meeting** Overview of Generative Modeling (VAEs, Energy Based Models, Score Based Generative Modeling, Diffusion). Dec 2023.
- **GT Single Cell Genomics Journal Club** How can we learn from spatial transcriptomics data? (SPICEMIX, COMMOT methodology overview). March 2023.

POSTERS/PRESENTATIONS

- **Mihir Bafna**, Hechen Li, Xiuwei Zhang. CLARIFY: Cell-cell interaction and gene regulatory network refinement from spatially resolved transcriptomics. *Single Cell Analyses Conference at Cold Spring Harbor Laboratory* (2023).
- **Mihir Bafna**, Ruochi Zhang, Bonnie Berger. Daifuku: Elucidating 3D Genome Structure via Diffusion-based scHi-C contact map imputation. *Broad Summer Research Program Final Symposium* (2023).
- **Mihir Bafna**, Vikranth Keerthipati, Subhash Kanaparthi. DiffRNAFold: Generating RNA Structures & Conformations with Latent Space Diffusion. *Intelligent Systems for Molecular Biology (ISMB) 2023*. **PURA Travel Award**.
- **Mihir Bafna**, Hechen Li, Xiuwei Zhang. CLARIFY: Cell-cell interaction and gene regulatory network refinement from spatially resolved transcriptomics. *Atlanta Workshop for Single-cell Omics (AWSOM 2023)*. **Best Poster Award**.
- **Mihir Bafna**, Vikranth Keerthipati, Subhash Kanaparthi. DiffRNAFold: Generating RNA Structures & Conformations with Latent Space Diffusion. *Deep Learning (CS 7643) Final Project Symposium & Poster Presentation*. **Best Overall Project & Poster (1st out of 200)**.
- **Mihir Bafna**, Xiuwei Zhang. Benchmarking and Refining Cell-Cell Interactions with Spatial Transcriptomics and Graph Neural Networks. *St. Jude's National Symposium for Undergraduate Research (NSUR 2022)*. **Poster Presentation (52/352 selected)**.