

# Martin Jinye Zhang

Updated: 8/30/2026

Phone: (650) 924-5691 | Email: [martinzh@andrew.cmu.edu](mailto:martinzh@andrew.cmu.edu) | Website: <https://mzhanglab.github.io/>

## RESEARCH INTERESTS

---

I focus on integrative analysis of large-scale molecular data such as GWAS and single-cell data. The biological endeavor is on mechanistic interpretation of variants discovered in GWAS. The computational challenges involve analyses of high-dimensional, multimodal, and very-large-scale data. Specific topics include:

**Genetics:** GWAS, gene regulation, scRNA-seq, scATAC-seq, eQTL/pQTL, heritability, disease-critical cellular contexts, Mendelian randomization, foundation models.

**Machine learning and statistics:** graph neural network, multi-agentic systems, multiple hypotheses testing, Monte Carlo methods, empirical Bayes, causal inference, algorithm acceleration, multi-armed bandits.

## ACADEMIC APPOINTMENTS

---

Assistant Professor, Ray and Stephanie Lane Computational Biology Department, School of Computer Science, Carnegie Mellon University 8/2023 - present

Postdoctoral Researcher, Department of Epidemiology, Harvard University 9/2019 – 8/2023  
Advisor: Alkes Price

## EDUCATION

---

Ph.D. Electrical Engineering, Stanford University 9/2014 – 9/2019  
Advisors: David Tse and James Zou

M.S. Electrical Engineering, Stanford University 9/2014 – 7/2017

B.Eng. Electrical Engineering, Tsinghua University 9/2010 – 7/2014

## HONORS AND AWARDS

---

### Major Awards

- Charles J. Epstein Trainee Awards Postdoctoral Semifinalist, ASHG 2021
- 2020 Top 50 Life and Biological Sciences Articles, Nature Communications, “Determining sequencing depth in a single-cell RNA-seq experiment” 2020
- Best Paper Award, RECOMB, “AdaFDR: a fast, powerful and covariate-adaptive approach to multiple hypothesis testing” 2019
- Stanford Graduate Fellowship (SGF), Stanford University 2015
- Numerical Technologies Award in Electrical Engineering, Stanford University 2015
- Outstanding Undergraduate Thesis, Tsinghua University, “Speech diarization based on the determinantal point processes” 2014

### Additional Recognitions

Martin Jinye Zhang

- Reviewers' Choice Award, ASHG, "TusoAI-CB: automatic end-to-end optimization of computational biology methods with agentic AI" 2026
- Reviewers' Choice Award, ASHG, "Discovering disease variants and mechanistic networks for small-cohort GWAS with geometric deep learning" 2024
- Reviewers' Choice Award, ASHG, "Estimating SNP-pair effect correlations across functional annotations" 2023
- Reviewers' Choice Award, ASHG, "Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data" 2021
- Reviewers' Choice Award, ASHG, "Transcriptome-wide association studies and fine-mapping at cell-type resolution" 2021
- Travel Award, PQG (Program in Quantitative Genomics, Harvard School of Public Health) 2020
- Travel Award, RECOMB 2019
- Travel Award, NIPS 2017
- Ranked 2/79 in the Electrical Engineering Ph.D. qualifying exam, Stanford University 2015

## RESEARCH GRANTS

---

1. NIH/NHGRI R01 HG014880 6/2026 – 6/2030  
"Fine-mapping GWAS disease mechanisms via multi-modal genomics and refined computational modeling"  
(PI: Zhang; \$2,255,590)
2. Jones Interdisciplinary Seed Fund 3/2026 – 3/2027  
"Exploring RNA-mediated mechanisms for disease-associated non-coding variants"  
(PI: Henninger, Co-PI: Zhang; \$100,000)
3. The Shurl and Kay Curci Foundation Award 3/2024 – 3/2026  
"Understanding cell type-specific disease mechanisms using GWAS and single-cell data"  
(PI: Zhang; \$500,000)

## PUBLICATIONS

---

\*equal contribution, †co-corresponding author

Full list at <https://scholar.google.com/citations?user=zjr6n-QAAAAJ&hl=en>

### *Selected main publications (first/senior author)*

1. A. Turcan, K. Huang, L. Li, **M. J. Zhang**. "TusoAI: Agentic Optimization for Scientific Methods", *ICLR*, 2026.
2. **M. J. Zhang**<sup>†</sup>, A. Durvasula<sup>\*†</sup>, C. Chiang<sup>\*†</sup>, ..., S. Sunyaev<sup>†</sup>, A. L. Price<sup>†</sup>. "Pervasive correlations between causal disease effects of proximal SNPs vary with functional annotations and implicate stabilizing selection", *Nature Genetics*, 2026.
3. M. Tiwari, R. Kang<sup>\*</sup>, J. Lee<sup>\*</sup>, D. Lee<sup>\*</sup>, ..., I. Shomorony<sup>†</sup>, **M. J. Zhang**<sup>†</sup>. "Faster Maximum Inner Product Search in High Dimensions", *ICML*, 2024.
4. M. Tiwari, R. Kang<sup>\*</sup>, D. Lee<sup>\*</sup>, S. Thrun, C. Piech, I. Shomorony<sup>†</sup>, **M. J. Zhang**<sup>†</sup>. "BanditPAM++: Faster k-medoids Clustering", *NeurIPS*, 2023.
5. **M. J. Zhang**<sup>\*†</sup>, K. Hou<sup>\*†</sup>, ..., B. Pasaniuc<sup>†</sup>, A. L. Price<sup>†</sup>. "Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data", *Nature Genetics*, 2022. **ASHG 2021 Charles J. Charles J. Epstein Trainee Awards Postdoctoral Semifinalist.**
6. M. Tiwari, R. Kang<sup>\*</sup>, J. Lee<sup>\*</sup>, ..., I. Shomorony<sup>†</sup>, **M. J. Zhang**<sup>†</sup>. "MABSsplit: Faster Forest Training Using Multi-Armed Bandits", *NeurIPS*, 2022.
7. **M. J. Zhang**<sup>†</sup>, A. O. Pisco<sup>†</sup>, S. Darmanis, J. Zou<sup>†</sup>. "Mouse aging cell atlas analysis reveals global and cell type-specific aging signatures", *eLife*, 2021.

8. **M. J. Zhang\***, V. Ntranos\*, D. Tse. “Determining sequencing depth in a single-cell RNA-seq experiment”, *Nature Communications*, 2020. **Top 50 Life and Biological Sciences Articles**.
9. **M. J. Zhang**, J. Zou, D. Tse. “Adaptive Monte Carlo Multiple Testing via Multi-armed Bandits”, *ICML*, 2019.
10. **M. J. Zhang**, F. Xia, J. Zou. “Fast and covariate-adaptive method amplifies detection power in large-scale multiple hypothesis testing”, *Nature Communications*, 2019. **RECOMB 2019 Best Paper Award**.
11. A. Abid\*, **M. J. Zhang\***, V. K. Bagaria, J. Zou, “Exploring patterns enriched in a dataset with contrastive principal component analysis”, *Nature Communications*, 2018.
12. V. Bagaria\*, G. Kamath\*, V. Ntranos\*, **M. J. Zhang\***, D. Tse, “Medoids in Almost Linear Time via Multi-armed Bandits”, *AISTATS*, 2018.
13. F. Xia\*, **M. J. Zhang\***, J. Zou<sup>†</sup>, and D. Tse<sup>†</sup>, “NeuralFDR: Learning Discovery Thresholds from Hypothesis Features”, *NIPS*, 2017.

### ***Selected other publications***

1. B. J. Strober, **M. J. Zhang**, T. Amariuta, J. Rossen, A. L. Price. “Fine-mapping causal tissues and genes at disease-associated loci”, *Nature Genetics*, 2025.
2. X. Jiang, **M. J. Zhang\***, Y. Zhang\*, M. Inouye, C. Holmes, A. L. Price<sup>†</sup>, G. McVean<sup>†</sup>. “Age-dependent topic modeling of comorbidities in UK Biobank identifies disease subtypes with differential genetic risk”. *Nature Genetics*, 2022.
3. A. Ginart, **M. J. Zhang**, J. Zou. “MLDemon: Deployment Monitoring for Machine Learning Systems”, *AISTATS*, 2021.
4. **The Tabula Muris Consortium**. “A single-cell transcriptomic atlas characterizes ageing tissues in the mouse”, *Nature*, 2020.
5. M. Tiwari, **M. J. Zhang**, J. Mayclin, S. Thrun, C. Piech, I. Shomorony. “Bandit-PAM: Almost Linear Time k-Medoids Clustering via Multi-Armed Bandits”, *NeurIPS*, 2020.
6. W. Zhou\*, M. R. Sailani\*, K. Contrepois\*, Y. Zhou\*, S. Ahadi\*, S. Leopold, **M. J. Zhang**, ..., G. M. Weinstock<sup>†</sup>, M. Snyder<sup>†</sup>, “Longitudinal multi-omics of host-microbe dynamics in prediabetes”, *Nature*, 2019.

### ***Preprints / papers in preparation***

1. A. Turcan, K. Hou, K. Z. Lin, A. Pfenning, S. Sakaue, **M. J. Zhang**. “Conditional polygenic enrichment distinguishes causal from tagging disease-critical cell populations in single-cell RNA-seq”, *bioRxiv*, 2026.
2. R. Lyu\*, A. Turcan\*, **M. J. Zhang<sup>†</sup>**, B. Wilder<sup>†</sup>. “Improving constraint-based discovery with robust propagation and reliable LLM priors”, *arXiv*, 2025.
3. M. Hu, **M. J. Zhang**, M. Chikina. “DALE-Eval: A comprehensive cell type-specific expression deconvolution benchmark for transcriptomics data”, *bioRxiv*, 2025.
4. K. Huang, ..., **M. J. Zhang<sup>†</sup>**, J. Leskovec<sup>†</sup>. “Small-cohort GWAS discovery with AI over massive functional genomics knowledge graph”, *medRxiv*, 2024, *in revision at Nature Genetics*.

## **RESEARCH SUPERVISION AND MENTORING**

---

### **Ph.D. Student**

- |                                        |                       |
|----------------------------------------|-----------------------|
| • Jeongho Chae (CMU CPCB)              | Fall 2024 - Present   |
| • Eric Chen (CMU Statistics)           | Fall 2025 - Present   |
| • Boshen Yan (CMU CPCB)                | Fall 2024 - Present   |
| • Jiwon Lee (CMU CPCB) (Co-advisor)    | Spring 2024 - Present |
| • Ana Prieto (CMU Biological Sciences) | Fall 2023 - Present   |
| • Alistair Turcan (CMU CPCB)           | Fall 2023 - Present   |

### **M.S. Student**

Martin JinYE Zhang

- Kate Zhang (CMU MSCB) Spring 2026 - Present
- Ray Liu (CMU MSCB) Spring 2025 - Present
- Dongbo Sun (CMU MSCB) Spring 2024 - Present
- David Sun (CMU) Fall 2023 - Present

### Undergraduate Student

- Robin Sayer (Charité and Freie Universität Berlin) Summer 2026 - Present
- Kangsheng Qi (ETH) Summer 2026 - Present
- Nicole Sorensen (CMU Statistics) Fall 2025 - Present
- Maxwell Soh (CMU) Spring 2024 - Present

### Past Trainee

- Yuxin Lu (CMU CPCB; CMU MSAS) Spring 2024 – Fall 2025
- Grace Du (CMU MSCB) Fall 2023 – Spring 2025
- Tanush Paradeshi (CMU) Spring 2025 – Fall 2025
- Johnson Huang (CMU MS) Fall 2023 - Spring 2025
- Mo Tiwari (OpenAI; Stanford CS PhD) Spring 2020 - Spring 2024

### INVITED TALKS

---

- “Small-cohort GWAS discovery with AI over massive functional genomics knowledge graph”, STATGEN, University of Minnesota 5/22/2025
- “Small-cohort GWAS discovery with AI over massive functional genomics knowledge graph”, seminar talk, Statistical Genetics Wrkg Grp Seminar, Johns Hopkins University 12/5/2024
- “Small-cohort GWAS discovery with AI over massive functional genomics knowledge graph”, seminar talk, Biostatistics, University of Pittsburgh 10/3/2024
- “Leveraging causal inference to elucidate the relationship between genes and disease”, one-hour keynote presentation, UCSD Causality Workshop, UCSD 4/4/2024
- “Estimating SNP-pair effect correlations across functional annotations”, Technical session, MCBIOS, Emory 3/23/2024
- “Estimating SNP-pair effect correlations across functional annotations”, CPCB seminar, CMU 9/8/2023
- “Computational Methods to Decipher Human Disease Using GWAS and Functional Genomics”, Computer Science, UCI 3/8/2023
- “Computational Methods to Decipher Human Disease Using GWAS and Functional Genomics”, SCS, CMU 2/16/2023
- “Computational Methods to Decipher Human Disease Using GWAS and Functional Genomics”, Department of Genetics Symposium I, Yale 2/13/2023
- “Computational Methods to Decipher Human Disease Using GWAS and Functional Genomics”, Special Seminar, UIUC 2/2/2023
- “Computational Methods to Decipher Human Disease Using GWAS and Functional Genomics”, Special Biostatistics Seminars, UNC 1/19/2023

### TEACHING

---

#### Carnegie Mellon University

- Co-instructor, 02-744, Introduction to Statistical Genetics Fall 2026

#### Carnegie Mellon University

- Instructor, 02-620, Machine Learning for Scientists Spring 2026  
Martin JinYE Zhang

**Carnegie Mellon University**

- Co-instructor, 02-704, Introduction to Statistical Genetics

Fall 2025

**Carnegie Mellon University**

- Co-instructor, 02-620, Machine Learning for Scientists

Spring 2025

**Carnegie Mellon University**

- Instructor, 02-704, Introduction to Statistical Genetics

Fall 2024

**Carnegie Mellon University**

- Co-instructor, 02-620, Machine Learning for Scientists

Spring 2024

**Stanford University**

- TA, EE 278, Introduction to Statistical Signal Processing

Spring 2017

**DEPARTMENTAL AND UNIVERSITY SERVICE**

---

1. CPCB Curriculum Committee, 2026
2. CPCB Admissions Committee, 2026.
3. CBD Retreat Organizing Chair, 2025.
4. CPCB Admissions Committee, 2025.
5. Organizing CBD student and faculty retreat, 2024
6. CPCB Admissions Committee, 2024.

**PROFESSIONAL SERVICE**

---

1. Journal paper reviewer for *Cell*, *Nature Genetics*, *Cell Genomics*, *Nature Methods*, *Nature Aging*, *Nature Microbiology*, *Nature Communications*, *Genome Biology*, *Bioinformatics*, *PLoS Genetics*, *Genetics*, *Human Genetics and Genomics Advances*, *Communications Biology*, *IEEE Transactions on Signal Processing*, *Journal of Machine Learning Research*, *Annals of Statistics*, *Annals of Applied Statistics*, *Biometrics*, etc.
2. Conference paper reviewer for *RECOMB* (2023-26), *ISMB* (2022-26), *GLBIO* (2024-25), *ICML* (2020-22), *NeurIPS* (2016, 2019-22), *IJCAI* (2021-22), *ICLR* (2026,2021).
3. ASHG 2025 Moderator for the session “Single-Cell Multiomic Dissection of Gene Regulation Across Disease States”
4. Conference PC member for *RECOMB* (2024-26), *GLBIO* (2024-25), *ISMB* (2023-26), *ISMB/ECCB* (2023), *AAAI Student Program* (2024).
5. Co-organizer for RECOMB-Genetics 2025 Workshop.
6. Frontiers in Genetics topic editor for the research topic “molecular and genetic heterogeneity of human diseases”, 2022.
7. Organizer of Information Systems Laboratory Colloquium, EE, Stanford, 2015-19.