Zhen Miao

1201 Eastlake Ave E, Seattle, WA 98109 | zmiao2@fredhutch.org | Home page [link]

Last updated: Oct 28, 2026

ACADEMIC APPOINTMENTS

Postdoctoral Researcher, Fred Hutch Cancer Center and HHMI **Feb 2026–present**
Computational Biology Program; Howard Hughes Medical Institute; advisor: Dr. Erick Matsen

Postdoctoral Researcher, University of Pennsylvania **Dec 2024–Dec 2025**
Advisors: Dr. Junhyong Kim and Dr. Nancy Zhang

EDUCATION

Ph.D. in Genomics and Computational Biology, University of Pennsylvania **Dec 2024**
Thesis: Statistical foundations of single cell open chromatin assays
Advisor: Dr. Junhyong Kim

M.A. in Statistics and Data Science, University of Pennsylvania **May 2024**
Thesis: Charting the coordinated progression of cell states in spatial transcriptomic data
Advisor: Dr. Nancy Zhang

B.S. in Biological Sciences, Nankai University **May 2019**
Minor in Applied Mathematics

RESEARCH INTERESTS

Developing statistical and machine learning models grounded in biological processes.

- Modeling coordinated cell-state changes in spatial transcriptomics and integrating single-cell multiomics
- Separating latent chromatin accessibility from measurement processes in single-cell epigenomics
- Language models for biological sequences that disentangle mutation, codon preference, and amino-acid selection in B-cell receptor (BCR) evolution

SELECTED PUBLICATIONS

* indicates co-first authorship

First or co-first author:

1. Miao Z, Qu Y, and Zhang NR. Spatial Kernel Restricted Canonical Correlation Analysis: Population Targets, Identifiability, and Stability. *arXiv* (2026)
2. **Miao Z**, Qu Y, Huang S, Laux L, Peters S, Aristel A, Zhang Z, Niedernhofer L, McMahon A, Kim J and Zhang NR. Dissecting the coordinated progression of cell states in spatial transcriptomics with CoPro. *bioRxiv* (2026)
3. **Miao Z**, Wang J, Park K, Kuang D, and Kim J. Depth-corrected multi-factor dissection of chromatin accessibility for scATAC-seq data with PACS. *Nature Communications* (2025)
4. **Miao Z** and Kim J. Uniform quantification of single-nucleus ATAC-seq data with Paired-Insertion Counting (PIC) and a model-based insertion rate estimator. *Nature Methods* (2024)
5. **Miao Z**, Humphreys BD, McMahon AP, and Kim J. Multiomics integration in the age of million single cell data. *Nature Reviews Nephrology* (2021)
6. **Miao Z***, Balzer MS*, Ma Z*, …, and Susztak K. Single cell regulatory landscape of the mouse kidney highlights cellular differentiation programs and disease targets. *Nature Communications* (2021)

***Contributing author*:**

7. Schnell J, **Miao Z**, Achieng M, …, and Lindstrom N. Controlling nephron precursor differentiation to generate proximal-biased kidney organoids with emerging maturity. *Nature Communications* (2025)

8. Lu Z, Verginadis I, …, **Miao Z**, …, and Fuchs S. Modified C-type natriuretic peptide normalizes tumor vasculature, reinvigorates antitumor immunity, and improves solid tumor therapies. *Science Translational Medicine* (2024)
9. Xiong L*, Liu J*, …, **Miao Z**, …, and McMahon A. Direct androgen receptor control of sexually dimorphic gene expression in the mammalian kidney. *Developmental Cell* (2023)
10. Liu H, Doke T, …, **Miao Z**, …, and Susztak K. Epigenomic and transcriptomic analyses define core cell types, genes and targetable mechanisms for kidney disease. *Nature Genetics* (2022)
11. Lu Z, McBrearty N, …, **Miao Z**, …, and Fuchs S. ATF3 and CH25H regulate effector trogocytosis and anti-tumor activities of endogenous and immunotherapeutic cytotoxic T lymphocytes. *Cell Metabolism* (2022)
12. Govek KW*, Troisi EC*, **Miao Z**, …, and Camara P. Single-Cell Transcriptomic Analysis of mIHC Images via Antigen Mapping. *Science Advances* (2021)
13. Zhao Q, Wang J, **Miao Z**, Zhang N, Hennessy S, Small D. S, & Rader D. J. "The role of lipoprotein subfractions in coronary artery disease: A Mendelian randomization study." *eLife* (2021)
14. Sheng X*, Guan Y*, …, **Miao Z**, …, and Susztak K. Mapping the genetic architecture of human traits to cell types in the kidney identifies mechanisms of disease and potential treatments. *Nature genetics* (2021)
15. Nguyen A*, Wang K*, Hu G*, Wang X, **Miao Z**, …, and Susztak K. APOE and TREM2 regulate amyloid-responsive microglia in Alzheimer's disease. *Acta Neuropathologica* (2020)
16. Zhong L, Yao L, Tower RJ, Wei Y, **Miao Z**, …, and Qin L. Single cell transcriptomics identifies a unique adipose lineage cell population that regulates bone marrow environment. *eLife* (2020)
17. Gao J, **Miao Z**, Zhang Z, Wei H, and Kurgan L. "Prediction of Ion Channels and Their Types from Protein Sequences: Comprehensive Review and Comparative Assessment." *Current Drug Targets* (2018)

Full list of publications available at Google Scholar Page [link]

RESEARCH EXPERIENCE

Postdoctoral Research, Fred Hutch Cancer Center **Feb 2026–present**

- Leading development of models that combine sequence transformers with explicit mutation, codon-preference, and amino-acid selection components to disentangle molecular processes in BCR evolution
- Derived identifiability results using synonymous-codon contrasts to establish which model components coding sequences can distinguish
- Implemented model fitting and neural-network training in PyTorch; evaluated codon continuous-time Markov chain models using antibody repertoires, Ensembl Compara data, and simulations

Postdoctoral Research, University of Pennsylvania **Dec 2024–Dec 2025**

- Developed CoPro to identify coordinated progression of cell states across cell types using spatial transcriptomics
- Implemented supervised and unsupervised analyses with interpretable gene and cell scores
- Applied CoPro to study coordination among cell populations in colon, kidney, and brain tissues

Zhen Miao
1201 Eastlake Ave E, Seattle, WA 98109 | zmiao2@fredhutch.org | Home page [link]
Last updated: Oct 28, 2026

Doctoral Research, University of Pennsylvania **Jul 2019–Dec 2024**
Advisor: Dr. Junhyong Kim

- Developed Paired-Insertion Counting (PIC) and a probability model of transposition and fragment recovery to infer latent insertion rates from snATAC-seq counts
- Co-developed the PACS statistical framework and implemented software to separate latent accessibility from cell-specific capture and jointly test multiple biological factors
- Demonstrated type I error control and improved power for differential accessibility testing in simulation and data-based benchmarks
- Reviewed methods for integrating matched and unmatched single-cell multiomics data

Research rotations

- Camara lab: Contributed to STvEA for spatial proteomics–transcriptomics integration and implemented feature selection
- Susztak lab: Mapped chromatin accessibility and regulatory programs during mouse kidney differentiation

Research Intern, University of Pennsylvania **Jul 2018–Jun 2019**
Mingyao Li lab: Analyzed microglial expression and ligand–receptor networks in single-nucleus RNA-seq data from 16 human brains with or without Alzheimer's disease

Undergraduate Researcher, Nankai University **Oct 2017–Jul 2018**
Jianzhao Gao lab: Developed a sequence-based m6A modification predictor using feature selection and support vector machines

RESEARCH SOFTWARE

Developer of the following packages:

- **PIC (PICsnATAC)** — Corrects insertion-counting artifacts in snATAC-seq quantification.
- **PACS** — Depth-corrected multifactor testing of chromatin accessibility in scATAC-seq.
- **CoPro** — Detects coordinated cell-state progression in spatial transcriptomics.

TEACHING AND MENTORING

Graduate Teaching Assistant, University of Pennsylvania **2021–2024**
Held weekly office hours for both graduate-level courses below.
BIOM/CIS/MTR-5350: Introduction to Bioinformatics

- Designed homework questions and course materials for advanced topics, including end-to-end analysis of snATAC-seq data

GCB-5330: Statistics for Genomics and Biomedical Informatics

- Reviewed homework and provided personalized answers to homework problems

Guest Lecturer, University of Pennsylvania **2024**
STAT-9915: Advanced Topics in Biomedical Data Science

- Lectured on statistical methods for single-cell chromatin accessibility data, including PIC and PACS

Research mentoring
Kernyu Park, Research Assistant **2024**
Project: Benchmarking statistical tests for scATAC-seq data

Elena Theus, Undergraduate Student **2024**
Project: Visualization tools for single-cell multiomics

Zhen Miao
1201 Eastlake Ave E, Seattle, WA 98109 | zmiao2@fredhutch.org | Home page [link]
Last updated: Oct 28, 2026

AWARDS AND HONORS

Selected Fellowships & Awards:

- *Blavatnik Family Fellowship in Biomedical Research*, awarded to six students across Penn Medicine Ph.D. candidates (2022-2023) [link]
- *Semifinalist, Charles J. Epstein Trainee Awards for Excellence in Human Genetics Research, ASHG (2019)* [link]
- *Third Place Poster Award*, American Statistical Association 2026 Meeting, SSGG Section (2026)
- *ACGA Trainee Awards*, awarded to the top 3 predoctoral students' platform presentations (2019)
- *National Scholarship*, the highest honor for undergraduate students in China (2019, 2017)

Institutional Honors:

- *University of Pennsylvania Nominee*, Regeneron Prize for Creative Innovation, selected as Penn's sole predoctoral nominee (2025)
- *NIH-CZI Junior Atlas Builder Travel Grant,* competitively selected for travel support to the Junior Atlas Builder's Meeting (2024)
- Outstanding Graduate (2019) and Merit Student (all undergraduate years), Nankai University
- *Poling Scholarship*, awarded to top 10 students in the Poling Program for academic performance (2018)
- *First Place Award,* The Fifth Life Sciences Students' Experimental Skills Competition in Tianjin (2017)

CONFERENCE PRESENTATIONS

Platform Presentations:

- American Statistical Association 2024 meeting, SSGG section (15 min)
  - Title: Model-based compound hypothesis testing for snATAC-seq data with PACS
- CSHL Single cell Analyses 2023 meeting (15 min)
  - Title: Model-based compound hypothesis testing for snATAC-seq data with PACS
- ReBuilding A Kidney (RBK) 2021 semi-annual meeting (*25 min*)
  - Title: Best practices of quantification for single cell epigenomics data
- Penn IRM Stem Cell Club 2021 (*25 min*)
  - Title: Charting kidney development through single cell multiomics study
- The 69th American Society of Human Genetics (ASHG) 2019 annual meeting (*15 min*)
  - Title: Characterizing cellular communication in the human brain by single-cell RNA-seq

Lightning Talks:

- American Statistical Association 2026 meeting, SSGG section (*5 min lightning talk*)
  - Title: Modeling coordinated progression in spatial transcriptomics with CoPro
- NIH Junior Atlas Builder 2023 meeting (*5 min lightning talk*)
  - Title: Atlas-scale integrative analysis of snATAC-seq data with PACS
- MidAtlantic Bioinformatics 2019 Conference (*5 min lightning talk*)
  - Title: Charting cellular communication network in Alzheimer's Disease

Poster Presentations:

- American Statistical Association (ASA) meeting, SSGG section (2026), Third Place Poster Award
- The American Society of Human Genetics (ASHG) annual meeting (2022, 2020)
- The American Society of Nephrology (ASN) annual meeting (2022, 2020)
- Fox Chase Epigenetics Symposium (2022, 2019)

- Mid-Atlantic Bioinformatics Conference (2020)

SERVICE

Journal and Conference Reviewer: Nature Machine Intelligence, Genome Biology, PLOS Computational Biology, Communications Biology, BMC Bioinformatics, Briefings in Functional Genomics, Cell Regeneration, PeerJ Computer Science, Functional & Integrative Genomics, Biomedicine & Pharmacotherapy, Conference on Intelligent Systems for Molecular Biology (ISMB), Frontiers in Endocrinology, Frontiers in Cell and Developmental Biology, PLOS One, International Conference on Learning Representations (ICLR), Frontiers in Medicine, Gene Reports, Heliyon.

GCB Advising Committee, University of Pennsylvania, Philadelphia, PA **2021–2023**
Advised first-year Ph.D. students on course selection based on their backgrounds and research interests and guided second-year students preparing for candidacy exams

TECHNICAL SKILLS

- *Programming [GitHub page link]*: Proficient in R, Python, C++, shell scripting
- *Machine learning*: PyTorch, transformer-based models for biological sequences, neural-network training, and model evaluation
- *Statistical Modeling*: categorical data analysis, Bayesian data modeling, time-series data analysis, survival analysis, spatial statistics, non-parametric statistics
- *Evolutionary modeling*: mutation–selection models, continuous-time Markov chains, parameter identifiability, and likelihood-based inference

INTERNSHIP

Computational Biology Group, 23andMe, Sunnyvale, CA **Jun–Aug 2022**
Graduate Student Intern, Supervisor: Robert Tunney, Ph.D.

- Optimized parallel computation and data handling in TACTICAL, a framework for identifying tissue of action, achieving an approximately 200-fold speedup
- Extended TACTICAL to single-cell ATAC-seq data for finer-resolution identification of disease-associated cell types