

QIN ZHU, PhD

COMPUTATIONAL BIOLOGY · SINGLE-CELL GENOMICS · SCIENTIFIC ML

San Francisco, CA · qin.zhu@ucsf.edu · qinzhu.github.io · GitHub · Google Scholar

PROFILE

Computational and experimental biologist who leads research spanning biological experiments, statistical analysis, and open-source software. **Lead-author contributions published in *Science*, *Nature Biotechnology*, *Nature Communications*, *Genome Biology*, and *Blood*.** Deep expertise in single-cell multi-omics and scientific machine learning, with broad work across development, cancer, immunology, hematopoiesis, and perturbation biology.

EXPERTISE

Machine learning, statistical & bioinformatics methods: first-principles method and model architecture development; self-supervised learning; contrastive representation learning; differential geometry; manifold learning; statistical machine learning; expectation-maximization (EM); generalized linear models; cross-modal alignment; graph-based methods

Single-cell perturbation, multimodal & spatial biology: scRNA-seq, scATAC-seq, CODEX spatial proteomics analysis, spatial transcriptomics analysis, high-throughput drug perturbation, sample demultiplexing and doublet detection, trajectory modeling, enhancer-promoter analysis, regulatory network analysis

Scientific computing: Python (PyTorch, TensorFlow, NumPy, pandas, scikit-learn, Scanpy/AnnData, cuML/cuGraph), R, C/C++, SQL, and Bash; open-source package and model development, evaluation, and benchmarking

Reproducible infrastructure: Git/GitHub, Linux/Unix, Slurm HPC, AWS EC2/S3, Docker, automated testing, versioned and documented analysis pipelines, and researcher-facing software

Experimental skills: single-cell sequencing, MULTI-seq, drug/cytokine perturbation, molecular biology experiments, cell culture, organoid and co-culture models, mouse behavioral studies, experimental design

Agentic scientific development: use of general-purpose coding agents for analysis design, algorithm prototyping, pipeline implementation, testing, code review, documentation, and hypothesis generation

RESEARCH EXPERIENCE

University of California, San Francisco (UCSF)

Gartner Lab | San Francisco, CA

Postdoctoral Scholar · CRI Immuno-Informatics Postdoctoral Fellow (2023–2026)

2021–Present

- **Developed CONCORD**, a unified self-supervised representation-learning framework that simultaneously performs dimensionality reduction, batch integration, and denoising while preserving local geometry and global topology. Designed its model architecture and training strategy, demonstrating that principled probabilistic minibatch sampling enables state-of-the-art performance even with a minimalist architecture; benchmarked the framework against leading methods on representation quality, biological resolution, generalization, and computational efficiency; first and co-corresponding author, *Nature Biotechnology* (2026).
- Designed and carried out high-throughput single-cell time-course studies of drug and cytokine perturbations in human PBMCs; identified epigenetic compounds that differentially modulate T-cell activation and translated response landscapes into mechanistic hypotheses for experimental follow-up; manuscript in preparation.
- **Developed deMULTIplex2**, which uses expectation-maximization to fit negative-binomial GLMs for robust sample assignment and doublet detection in high-throughput multiplexed single-cell experiments, including MULTI-seq; released and documented the open-source R package; first and co-corresponding author, *Genome Biology* (2024).
- Coauthored D-SPIN, a regulatory-network modeling study of scRNA-seq perturbation responses; carried out drug-perturbation experiments in PBMCs from multiple donors and developed an analysis workflow that mapped known drug-protein interactions and classified them into mechanistically distinct categories for biological interpretation and follow-up.
- Contributed computational and experimental work across developmental biology, technology development, cancer, and immunology; led and submitted an NIH K99 application as principal investigator and drafted and co-submitted an NIH R01 application as co-investigator with PI Zev Gartner.

University of Pennsylvania

Tan and Kim Labs | Philadelphia, PA

PhD Researcher / Research Specialist

2015–2021

- Developed a cross-modal scRNA-seq/scATAC-seq integration workflow using a custom gene-activity representation and nearest-neighbor metacell pairing to link chromatin accessibility with gene expression, infer enhancer–promoter relationships, and characterize Runx1-driven mechanisms of the endothelial-to-hematopoietic transition; co-first author, *Blood* (2020).
- Analyzed CODEX multiplexed-imaging data using channel-wise Gaussian-mixture modeling, clustering, cell-type annotation, and spatial-neighborhood analysis; interpreted spatial protein measurements alongside scRNA-seq to map spatial cell–cell interactions in the human colorectal tumor microenvironment; co-first author, *Nature Communications* (2023).
- Co-led construction of a lineage-resolved atlas of *C. elegans* embryogenesis and developed scalable analyses connecting cell identity, developmental trajectories, and regulatory programs; co-first author, *Science* (2019).
- Developed open-source computational methods, analysis pipelines, and interactive software for genomic, transcriptomic, and epigenomic analysis, including PIVOT, VisCello, Kidney Cell Explorer, scATAC-pro, SCRABBLE, and VERSE, with associated methods published across multiple peer-reviewed journals.

SELECTED RESEARCH SOFTWARE

CONCORD (*Python · PyTorch*) — self-supervised contrastive learning for batch-integrated, denoised, topology-preserving representations of single-cell data

deMULTIplex2 (*R · EM · negative-binomial GLMs*) — probabilistic sample assignment and doublet detection for multiplexed scRNA-seq

VisCello (*R · Shiny · JavaScript*) — interactive software for large-scale single-cell and CODEX data visualization

PIVOT (*R · Shiny*) — interactive analysis and visualization platform for transcriptomics data

Kidney Cell Explorer (*R · Shiny · VisCello*) — interactive map of adult mouse kidney cell diversity, gene expression, and ontology-aware metacells

SCRABBLE (*R · MATLAB*) — single-cell RNA-seq imputation constrained by matched bulk RNA-seq data

scATAC-pro (*Bash · R*) — comprehensive workbench for single-cell chromatin-accessibility data processing

VERSE (*C*) — versatile and efficient RNA-seq read assignment tool

SELECTED PUBLICATIONS

Zhu Q, Jiang Z, Zuckerman B, Weinberger L, Thomson M, Gartner Z. Revealing a coherent cell-state landscape across single-cell datasets with CONCORD. *Nature Biotechnology* (2026). First and co-corresponding author.

Jiang J, Chen S, Tsou T, McGinnis CS, Khazaei T, **Zhu Q**, et al. D-SPIN constructs regulatory network models from scRNA-seq that reveal organizing principles of perturbation response. *Cell* (2026).

Zhu Q, Conrad DN, Gartner ZJ. deMULTIplex2: robust sample demultiplexing for scRNA-seq. *Genome Biology* (2024). First and co-corresponding author.

Zhu Q*, Gao P*, Tober J*, et al. Developmental trajectory of pre-hematopoietic stem cell formation from endothelium. *Blood* (2020). Co-first author.

Li N*, **Zhu Q***, Tian Y, et al. Mapping and modeling human colorectal carcinoma interactions with the tumor microenvironment. *Nature Communications* (2023). Co-first author.

Packer JS*, **Zhu Q***, Huynh C, et al. A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* (2019). Co-first author.

Yu W, Uzun Y, **Zhu Q**, Chen C, Tan K. scATAC-pro: a comprehensive workbench for single-cell chromatin-accessibility sequencing data. *Genome Biology* (2020).

Zhu Q, Fisher SA, Dueck H, et al. PIVOT: platform for interactive analysis and visualization of transcriptomics data. *BMC Bioinformatics* (2018). First author.

* Equal contribution. Selected publications; full list on [Google Scholar](#).

EDUCATION

University of Pennsylvania — PhD, Genomics & Computational Biology, 2021; Master of Biotechnology, 2015

Nanjing University — BS, Biological Science, 2013

LEADERSHIP & SCIENTIFIC COMMUNICATION

- Mentored a PhD student from rotation through candidacy and two rotation researchers across UCSF's Tetrad and UC Berkeley-UCSF Bioengineering graduate programs.
- Taught computational biology and single-cell analysis through graduate teaching, guest lectures, and hands-on workshops; developed materials for audiences spanning experimental and computational biology.
- **Selected talks & presentations:** Boston Children's Hospital & Harvard Medical School Digital Science Seminar (2026); Genentech AI Biology & Translation Seminar (2025); Janelia AI Revolution Meets 4D Cellular Physiology (2025); Chan Zuckerberg Initiative Single Cell Group Meeting (2024).

FELLOWSHIPS, HONORS & RESEARCH SUPPORT

Anthropic AI for Science Program, 2026 — Selected participant; received API credits supporting AI-enabled biological research.

Cancer Research Institute Immuno-Informatics Postdoctoral Fellowship

2023–2026

Tom Kadesch Prize in Genetics Research

2021

Blavatnik Family Fellowship in Biomedical Research

2020