Weixiang Fang, PhD

Department of Biomedical Engineering

Johns Hopkins University · 855 N. Wolfe St., Rangos 570 · Baltimore, MD 21205

wfang58@gmail.com | (631) 703-1388 | wefang.github.io | github.com/wefang | LinkedIn | ORCID | Google Scholar

Research Focus

I develop **statistical and machine learning methods** for data produced by emerging genomics technologies, beginning with genomic records, the edits an engineered recorder writes into a cell's own DNA. I use them to infer the dynamics of cell fate determination, and to transfer what a model organism reveals about the regulatory control of fate into human.

Research program

- **Model-based and scalable lineage inference from genomic recorders.** Recorders now write a cell's history into its DNA and the readout reaches millions of cells, while the methods that infer the lineage from those readouts remain largely heuristic. I write the generative process for each class of recorder, evolving, static, sequential and conditional, derive likelihood-based inference from it, and train learned proposals so that whole trees can be optimized at that scale. The same treatment extends to somatic mutations, the recorder every human already carries.
- **Joint inference of the intrinsic and extrinsic control of cell fate.** A branching process model in which extrinsic signals, alongside a cell's own state, decide its fate over generations, with a transition kernel, a migration model and a signal model parameterized by neural networks and fitted in one joint likelihood to time-course, spatial and lineage data. The fitted model gives counterfactual estimates of the plasticity of each cell state and of the effect of individual signals on fate, which is the design problem behind differentiation protocols and cell therapy.
- **Cross-species transfer learning for the regulatory control of fate decisions.** One model of regulatory element activity across species, developmental stage and assay, with a species-invariant cell embedding and an explicit developmental stage, which names the human cell states matching an animal fate decision and the human element whose predicted activity agrees at each of them, and a database of animal evidence transferred to human, naming for each the human sequence and cell state where the same effect is expected, and flagging where no human counterpart exists.

Selected Works

- **Quantitative fate mapping: A general framework for analyzing progenitor state dynamics via retrospective lineage barcoding**. *Cell*, 2022.
- **Quantifying Functional Conservation of Human and Mouse Regulatory Elements via FUNCODE**. *bioRxiv preprint*, 2024. (In revision *Genome Biology*)
- **A branching process model of cell lineage links positional signaling to the canalization of clonal fate**. *Manuscript in preparation*, 2026.

Academic Appointments

Postdoctoral Fellow in Biomedical Engineering 2022–Present
Johns Hopkins University · Baltimore, MD · PI: **Dr. Reza Kalhor**

Education

PhD in Biostatistics 2016–2022
Johns Hopkins University · Baltimore, MD
Dissertation: "**Quantitative Methods for Genomics and Lineage Tracing Data**" · Advisor: **Dr. Hongkai Ji**

Master's program in Biostatistics 2014–2016
Johns Hopkins University · Baltimore, MD · Transferred to PhD program

B.Eng in Engineering Science 2010–2014
Stony Brook University · Stony Brook, NY

Major in Psychology 2009–2010
Fudan University · Shanghai, China · Transferred to Stony Brook University

Publications

Peer-reviewed journal articles

1. **W. Fang**, C. M. Bell, A. Sapirstein, S. Asami, K. Leeper, D. J. Zack, H. Ji, and R. Kalhor, 2022. "Quantitative fate mapping: A general framework for analyzing progenitor state dynamics via retrospective lineage barcoding". *Cell*
2. W. Zhou, Z. Ji, **W. Fang**, and H. Ji, 2019. "Global prediction of chromatin accessibility using small-cell-number and single-cell RNA-seq". *Nucleic Acids Research*
3. X. Chamling, A. Kallman, **W. Fang**, C. A. Berlinicke, J. L. Mertz, P. Devkota, I. E. M. Pantoja, M. D. Smith, Z. Ji, C. Chang, A. Kaushik, L. Chen, K. A. Whartenby, P. A. Calabresi, H.-Q. Mao, H. Ji, T.-H. Wang, and D. J. Zack, 2021. "Single-cell transcriptomic reveals molecular diversity and developmental heterogeneity of human stem cell-derived oligodendrocyte lineage cells". *Nature Communications*
4. J. S. Handler, Z. Li, R. K. Dveirin, **W. Fang**, H. Goodarzi, E. J. Fertig, and R. Kalhor, 2025. "A cell-state axis underlying colonization in carcinomas with implications for metastasis risk prediction and interception". *Cell Reports*
5. W. Li, C. Berlinicke, Y. Huang, S. Giera, A. G. McGrath, **W. Fang**, C. Chen, F. Takaesu, X. Chang, Y. Duan, D. Kumar, C. Chang, H.-Q. Mao, G. Sheng, J. C. Dodge, H. Ji, S. Madden, D. J. Zack, and X. Chamling, 2023. "High-throughput screening for myelination promoting compounds using human stem cell-derived oligodendrocyte progenitor cells". *iScience*
6. C. J. Keuthan, J. A. Schaub, M. Wei, **W. Fang**, S. Quillen, E. Kimball, T. V. Johnson, H. Ji, D. J. Zack, and H. A. Quigley, 2023. "Regional gene expression in the retina, optic nerve head, and optic nerve of mice with optic nerve crush and experimental glaucoma". *International Journal of Molecular Sciences*

Preprints

7. **W. Fang**, C. Chen, B. Zhang, Y. Wang, R. Zhao, W. Zhou, and H. Ji, 2024. "Quantifying Functional Conservation of Human and Mouse Regulatory Elements via FUNCODE". *Preprint; in revision at Genome Biology*. doi.org/10.1101/2024.10.31.620766
8. T. Fabiha, I. Evergreen, S. Kundu, A. Pampari, S. Abramov, A. Boytsov, K. Strouse, K. Dura, **W. Fang**, G. Kerner, J. Butts, T. Ali, A. Gschwind, K. S. Mualim, J. E. Moore, Z. Weng, J. Ulirsch, H. E. Ji, J. Vierstra, T. E. Reddy, S. B. Montgomery, J. Engreitz, A. Kundaje, R. Tewhey, A. Price, and K. Dey, 2024. "A consensus variant-to-function score to functionally prioritize variants for disease". *Preprint*. doi.org/10.1101/2024.11.07.622307

Submitted

9. J. E. Forsmo, **W. Fang**, J. D. Lin, M. Vazquez-Arroyo, L. Fang, J. Jun, W. Lin, and R. Kalhor, 2026. "Genomic recording of gene-expression history along lineage trees". *In submission*

In preparation

10. **W. Fang**, Y. Yang, J. D. Lin, and R. Kalhor, 2026. "A branching process model of cell lineage links positional signaling to the canalization of clonal fate". *Manuscript in preparation*

Book chapter

11. **W. Fang**, Y. Yang, H. Ji, and R. Kalhor, 2025. "Reconstructing Progenitor State Hierarchy and Dynamics Using Lineage Barcoding Data". *Lineage Tracing: Methods and Protocols*

Software

- **qfm** (R package). Phylotime and ICE-FASE for quantitative fate mapping, with simulation of cell phylogeny and lineage barcodes. github.com/Kalhor-Lab/QFM, archived at Zenodo (record 7114804).
- **FUNCODE scores**. Genome-wide functional conservation scores for human and mouse regulatory elements, released as UCSC Genome Browser track hubs for hg38 and mm10.

Research Experience

Postdoctoral research

Advisor: Dr. Reza Kalhor 2022–Present

- Developed **Phylotime** and **ICE-FASE**, the methods of **Quantitative Fate Mapping** (*Cell*, 2022). Phylotime is likelihood-based inference of time-scaled cell phylogeny from CRISPR **lineage barcodes**, modeling the editing process and dropout; ICE-FASE estimates progenitor state hierarchy, commitment times, population sizes and fate biases from the tree.
- Developed the inference method for the ihgRNA conditional recorder that produced the first **decorated lineage tree**, recovering the history of Pax6 expression across the developing retina (submitted).
- Building scalable **single cell phylogenetic inference** from engineered and somatic lineage barcodes, with explicit models of editing dynamics and dropout, including **PhyloTime+**, a C library that optimizes whole trees by subtree prune-and-regraft moves.
- Building **LINST**, a **branching process model** of development linking a cell's transcriptional state and the extrinsic signals it receives to the fates of its descendants (manuscript in preparation).

Doctoral research

Advisor: Dr. Hongkai Ji 2014–2022

- **FUNCODE**: genome-wide scores for **functional conservation of regulatory DNA** between human and mouse, built on the ENCODE DNase, ATAC and histone ChIP compendium and released as UCSC track hubs for hg38 and mm10.
- **Statistical learning of local patterns of cis-regulation**: a hierarchical mixture model for the joint activity state of regulatory elements within a locus, fitted to ENCODE DNase hypersensitivity across 123 cell types within Hi-C topological domains, and applied to denoise low-input and single-cell **ATAC-seq**.
- **Deep learning methods** for prediction of regulome signal using RNA-seq with extension to **single cell RNA-seq**.
- **Functional data analysis of regulatory activity across cell types**: penalized spline models fitted along a cross-species ordering of ENCODE and Roadmap samples, validated on 11,938 orthologous gene pairs, the approach that became the conservation scores in FUNCODE.

Undergraduate research

Advisor: Dr. Juhyuk Moon, Multi-Scale Structural Materials Lab 2013–2014

- Resolved the crystal structure of **yeelimite** ($Ca_4Al_6O_{12}SO_4$), the principal phase of calcium sulfoaluminate cement, by **first-principles DFT** energy minimization across competing cubic, tetragonal and orthorhombic space groups reported in the diffraction literature
- Built **thermodynamic models** of Portland cement hydration by Gibbs energy minimization (GEMS, CEMDATA), covering phase assemblage, pore solution chemistry and temperature dependence in ternary blends with limestone and fly ash
- Measured the thermal behavior of elemental sulfur and **sulfur polymer concrete** by **differential scanning calorimetry**, the material's main application being hazardous waste stabilization

Presentations

Invited talks

- "A scalable approach for characterizing functional conservation of regulatory elements between human and mouse". **ENCODE Consortium Meeting 2021**, comparative genomics section. [Online due to COVID-19 pandemic]
- "Mapping complex cell fate dynamics using cell phylogeny". **2022 4DN annual meeting**, San Diego, 2022
- "Mapping complex cell fate dynamics using cell phylogeny". **4DN Scientific Webinar Series**, Delivered Online, July 31, 2023

Poster presentations

- "Single cell lineage model bridges cell fate and state in development". **CSHL Biological Data Science**, Cold Spring Harbor, NY, November 2024.
- "Understanding developmental variabilities by lineage-resolved single cell model". **EMBO Workshop: Lineage tracing: Dynamics, cellular memory, and somatic evolution**, Sant Feliu de Guíxols, Spain, September 2025. [Delivered remotely]
- "Understanding developmental variabilities by lineage-resolved single cell model". **2026 Lake Conference: Modeling Life from Cells to Tissues: Emergence and Self-Organization**, Allen Institute, Seattle, WA, April 2026.

Presentations by co-authors

- Cell lineages across scales, space and time. Short Talk: "A general framework for inferring progenitor state hierarchy and dynamics from retrospective lineage barcoding" * Presented by **Reza Kalhor**
- Epigenome isoform analysis with applications. JSM, Chicago, IL, Aug 3, 2016 *Presented by **Hongkai Ji**

Honors and Awards

Outstanding Performance on the Biostatistics PhD Comprehensive Exam 2017
Johns Hopkins University

Kocherlakota Award for Best Performance on the Biostatistics Comprehensive Examination by a Master's Student 2015
Johns Hopkins University

Academic Excellence and Outstanding Achievement Award 2012–2013
Stony Brook University

Fudan University First-class Renmin Scholarship 2010
Fudan University

Affiliations

- **The Encyclopedia of DNA Elements (ENCODE) Consortium**, leading analysis on comparative functional genomics
- **4D Nucleome (4DN) program**

Mentoring Experience

Yi Yang · Master's student in Biomedical Engineering 2022–2024
Johns Hopkins University · Now **PhD candidate at Johns Hopkins Biomedical Engineering**
Research projects: Dynamics inference from time-course cell state composition data; web interface for exploring cell phylogeny

Smriti Srikanth · Undergraduate student in Biomedical Engineering 2022–2024
Johns Hopkins University · Now **Computation Associate at Broad Institute**
Research projects: Mosaic fraction analysis to analyze bulk lineage tracing data; model for simulating gene expression on single cell phylogeny in development

Ashwin Rajendran · Master's student in Biomedical Engineering 2023–2025
Johns Hopkins University · Now **PhD candidate at Florida State University**
Research project: Scalable phylogeny inference using diploid lineage barcodes with heritable and random missing data

Teaching Experience

Teaching Assistant 2015–2021
Johns Hopkins University · Baltimore, MD

- **JHSPH 140.688** Statistics for Genomics (Spring 2020)
- **JHSPH 140.776** Statistical Computing (Fall 2017, Fall 2018)
- **JHSPH 140.621-624** Statistical Methods in Public Health (Fall 2016, Spring 2017, Fall 2017, Spring 2018, Spring 2020, Fall 2020, Spring 2021)
- **JHSPH 140.615** Statistics for Laboratory Scientists (Spring 2016, Spring 2019)
- **AS.280.345** Public Health Biostatistics (Fall 2015, Fall 2019)

Undergraduate Teaching Assistant 2013–2014
Stony Brook University · Stony Brook, NY

- Fundamentals of Engineering Chemistry
- Engineering Laboratory

Technical Skills

Statistics. **Likelihood-based inference** for stochastic processes and **phylogenetics**, **Bayesian inference** and MCMC, generalized linear and **mixed-effects models**, **survival analysis**, **functional regression**, **high-dimensional statistics**, experimental design

Languages. **R**, **Python**, **C/C++** (CMake), Bash/shell, SQL

Single cell. **Seurat**, **Scanpy**

Python. NumPy, pandas, SciPy, scikit-learn, matplotlib

R. tidyverse, **Bioconductor**, ggplot2

Machine learning. **TensorFlow/Keras**

HPC and cloud. **SLURM** clusters, **AWS**, **Docker** and **Singularity** containers

Tools. **Git**, LaTeX

Wet Lab Experience

Cell and molecular biology. Cell Culture, Cell Transfection, Cryopreservation, Cell Sorting (FACS), Optimizing DNA Amplification (PCR), NGS Library Preparation

Tissue preparation. Murine Embryo Microdissection, Fresh Frozen Tissue Preparation, Cryosectioning, Laser Capture Microdissection

Imaging. Immunohistochemistry, Fluorescence Microscopy