

Jian Zhou, Ph.D.

Lyda Hill Department of Bioinformatics,
UT Southwestern Medical Center
5323 Harry Hines Blvd, Dallas, TX 75390

jian.zhou@utsouthwestern.edu

RESEARCH EXPERIENCE

2019-Present Assistant Professor,
Lupe Murchison Foundation Scholar in Medical Research,
CPRIT Scholar for Cancer Research,
Lyda Hill Department of Bioinformatics, UT Southwestern Medical Center

2017-2019 Flatiron Research Fellow,
Genomics Group, Center for Computational Biology, Simons Foundation

2012-2017 Graduate research assistant,
Lewis-Sigler Institute for Integrative Genomics, Princeton University

EDUCATION

2017 *Ph.D.*, Quantitative and computational biology, **Princeton University**, Princeton, USA

2011 *B.S.*, Biological Sciences, **Peking University**, Beijing, China

GRANTS AND AWARDS

2021 - present **NIH Director's New Innovator Award (DP2),**
Office of the Director, NIH

2019 - present **CPRIT Scholar in Cancer Research**
Cancer Prevention and Research Institute of Texas

2019 - present **Lupe Murchison Foundation Scholar in Medical Research,**
University of Texas Southwestern Medical Center

PUBLICATIONS AND PREPRINTS

Kathleen M Chen, Aaron K Wong, Olga G Troyanskaya[#], Jian Zhou[#] (2022) A sequence-based global map of regulatory activity for deciphering human genetics, *Nature Genetics (in press)* ([#] co-corresponding authors)

Jian Zhou (2022) Sequence-based modeling of three-dimensional genome architecture from kilobase to chromosome-scale, *Nature Genetics*

Pavel Avdeyev, Jian Zhou (2022) Computational Approaches for Understanding Sequence Variation Effects on the 3D Genome Architecture, *Annual Review of Biomedical Data Science*

Jian Zhou[#], Olga G. Troyanskaya[#] (2021) An analytical framework for interpretable and generalizable 'quasilinear' single-cell data analysis, *Nature Methods* (# co-corresponding authors)

Xi Chen, Jian Zhou, Ran Zhang, Aaron K. Wong, Christopher Y. Park, Chandra L. Theesfeld, Olga G. Troyanskaya (2021) Tissue-specific enhancer functional networks for associating distal regulatory regions to disease, *Cell Systems*

Christopher Y. Park, Jian Zhou, Aaron K. Wong, Kathleen M. Chen, Chandra L. Theesfeld, Robert B. Darnell, Olga G. Troyanskaya (2021) Genome-wide landscape of RNA-binding protein target site dysregulation reveals a major impact on psychiatric disorder risk, *Nature genetics*

Felix Richter*, Sarah U Morton*, Seong Won Kim*, Alexander Kitaygorodsky*, Lauren K Wasson*, Kathleen M Chen*, Jian Zhou, ..., Olga G Troyanskaya, Diane E Dickel, Yufeng Shen, Jonathan G Seidman, Christine E Seidman, Bruce D Gelb (2020) Genomic analyses implicate noncoding de novo variants in congenital heart disease, *Nature Genetics*

Jian Zhou*, Christopher Y. Park*, Chandra L. Theesfeld*, Aaron K. Wong, Yuan Yuan, Kirsty Sawicka, Claudia Scheckel, John Fak, Yoko Tajima, Alan Packer, Robert B. Darnell[†], Olga G. Troyanskaya[†] (2019) Whole-genome deep learning analysis identifies contribution of noncoding mutations in autism, *Nature Genetics* (*co-first authors)

Jian Zhou*, Ignacio E. Schor*, Victoria Yao, Chandra L. Theesfeld, Raquel Marco-Ferrerres, Alicja Tadych, Eileen E. M. Furlong, Olga G. Troyanskaya (2019) Accurate genome-wide predictions of spatio-temporal gene expression during embryonic development, *PLOS Genetics* (*co-first authors)

Jennifer Harder*, Rajasree Menon*, Edgar A. Otto*, Jian Zhou*, Sean Eddy, Noel Wys, Viji Nair, Cristina Cebrian, Jason R. Spence, Olga G. Troyanskaya, Jeffrey Hodgkin, Roger Wiggins, Benjamin Freedman, Matthias Kretzler, European Renal cDNA Bank, Nephrotic Syndrome Study Network (2018) Organoid single-cell profiling identifies a transcriptional signature of glomerular disease, *JCI Insight*, (*co-first authors)

Kathleen Chen*, Evan Cofer*, Jian Zhou, Olga Troyanskaya (2018) Selene: a PyTorch-based deep learning library for sequence-level data, *Nature Methods*

Jian Zhou, Chandra Theesfeld, Kevin Yao, Kathleen Chen, Aaron Wong, and Olga Troyanskaya (2018) Deep learning sequence-based *ab initio* prediction of variant effects on expression and disease risk, *Nature Genetics*

Rajasree Menon, Edgar A. Otto, Austin Kokoruda, Jian Zhou, Zidong Zhang, Euisik Yoon, Yu-Chih Chen, Olga Troyanskaya, Jason R. Spence, Matthias Kretzler, Cristina Cebrián (2018) Single-cell analysis of progenitor cell dynamics and lineage specification in the human fetal kidney, *Development*

Jian Zhou, Olga Troyanskaya (2016) Probabilistic modeling of chromatin code landscape reveals functional diversity of enhancer-like chromatin states, *Nature Communications*

Jian Zhou, Olga Troyanskaya (2015) Predicting effects of noncoding variants with deep learning-based sequence model, *Nature Methods*

Jian Zhou, Olga Troyanskaya (2014) Global Quantitative Modeling of Chromatin Factor Interactions, *PLOS Computational Biology*

Jian Zhou, Olga Troyanskaya (2014) Deep Supervised and Convolutional Generative Stochastic Network for Protein Secondary Structure Prediction, *International Conference of Machine Learning (ICML 2014)*, *JMLR W&CP*

Xiaoxi Wang*, Lei Pan*, Su Wang*, Jian Zhou, William McDowell, Jungeun Park, Jeff Haug, Karen Staehling, Hong Tang, Ting Xie, (2011) Histone H3K9 trimethylase Egless controls germline stem cell maintenance and differentiation, *PLOS Genetics* (*co-first authors)

PRESENTATIONS

- 2022 Boston Children's Hospital Seminar, Title: Sequence-based modeling of genome 3D architecture from kilobase to chromosome-scale
- 2022 BIRS workshop Deep Learning for Genetics, Genomics and Metagenomics: Latest developments and New Directions, Title: Sequence-based modeling of genome 3D architecture from kilobase to chromosome-scale
- 2022 ISCB Academy Webinar, Title: Sequence-based modeling of genome 3D architecture from kilobase to chromosome-scale
- 2022 Winter Q-Bio 2022, Title: Sequence-based modeling of genome 3D architecture from kilobase to chromosome-scale
- 2021 ISMB/ECCB 2021, Title: Sequence-based modeling of genome 3D architecture from kilobase to chromosome-scale
- 2021 NVIDIA GTC 2021, Title: Understanding Genome Sequences with Deep Learning for Health and Diseases
- 2019 Winter Q-Bio 2020, Title: An analytical framework for interpretable 'quasilinear' single-cell omics analysis
- 2019 NeurIPS 2020 LMRL workshop, Title: A quasilinear framework for interpretable exploratory analysis of single-cell omics data
- 2019 Children's Hospital of Philadelphia, Title: Learning the regulatory program of the genome: from sequence to chromatin, expression, and evolution
- 2019 Broad Institute of MIT and Harvard and Department of Biology, MIT, Title: Learning the regulatory program of the genome: from sequence to chromatin, expression, and evolution
- 2019 Department of Bioinformatics, University of Texas Southwestern Medical Center, Title: Learning the regulatory program of the genome: from sequence to chromatin, expression, and evolution
- 2018 Society of Biological Psychiatry (SoBP) Annual Meeting, Title: Decoding the Role of Noncoding Genome in Neurological Disease With Deep Learning
- 2017 CSHL NY Area Meeting in Quantitative Biology, Title: A unifying view of single cell data exploratory analysis with uncertainty estimation - statistical inference for clusters, trajectories, and surfaces
- 2016 CSHL NY Area Meeting in Quantitative Biology, Title: Predicting effects of noncoding variants with deep learning-based sequence model
- 2016 ICML workshop for computational biology, Title: On multitask learning by linear combination (co-first author)
- 2016 The Allied Genetics Conference (TAGC), Title: Genome-wide spatial-temporal gene expression pattern prediction in *Drosophila melanogaster* embryonic development
- 2016 ISMB, Title: Predicting effects of noncoding variants with deep learning-based sequence model
- 2014 ICML, Title: Deep Supervised and Convolutional Generative Stochastic Network for Protein Secondary Structure Prediction
- 2013 NIPS deep learning workshop, Title: Deep Supervised and Convolutional Generative Stochastic Network for Protein Secondary Structure Prediction

PROFESSIONAL SERVICES

Invited referee: Science, Nature Biotechnology, Nature Methods, Nature Genetics, Nature Communications, PLOS Computational Biology, PLOS Genetics, Genome Biology, iScience, Nucleic Acids Research, Bioinformatics, PSB, Algorithms for Molecular Biology, Journal of Theoretical Biology, Journal of Bioinformatics and Computational Biology, Engineering Applications of Artificial Intelligence. **Posters Committee:** ISMB, PSB. **Conference Session Chair:** Winter Q-Bio 2020. **Program Committee:** ISMB 2022