

# Jessica L. Zhou

Website: <https://zrcjessica.github.io/>

Email: [jessicazhou96@gmail.com](mailto:jessicazhou96@gmail.com)

LinkedIn: [jessica-l-zhou](#)

GitHub: [github.com/zrcjessica](https://github.com/zrcjessica)

## EDUCATION

---

### University of California, San Diego

Ph.D., Bioinformatics and Systems Biology

La Jolla, CA

Sep. 2017–Jul. 2023

### University of Southern California

B.S., Biomedical Engineering

Los Angeles, CA

Aug. 2013–May 2017

## EXPERIENCE

---

- **Stripe** New York, NY  
*Data scientist* Jun. 2025 - Oct. 2025
  - Built unsupervised models to segment consumers for Link product marketing
  - Developed autoencoder to generate meaningful representations of multimodal consumer feature data
  - Deployed SQL pipelines with Apache Airflow to automate multi-step ETL workflows
- **Cold Spring Harbor Laboratory, Koo Lab** Cold Spring Harbor, NY  
*Postdoctoral researcher* Oct. 2023 - Jun. 2025
  - Pioneered a knowledge distillation framework for uncertainty quantification in deep neural networks
  - Applied explainable AI techniques (SHAP) to evaluate model interpretability and identify learned biological sequence patterns governing gene regulation
  - Benchmarked SOTA genomics foundation model Evo2 by training downstream CNNs on embeddings and evaluating variant effect prediction
  - Worked on a Retrieval Augmented Generation workflow with LangChain using GPT-4o Mini and FAISS to query a database of scientific talk transcripts
- **UC San Diego & Salk Institute for Biological Studies, McVicker Lab** La Jolla, CA  
*Ph.D. Researcher* Sep. 2017 - Jul. 2023
  - Developed generalized linear models to measure DNA sequence activity from single-cell CRISPR perturbation data
  - Used genomic foundation models to validate biological hypotheses through counterfactual analysis
  - Led an interdisciplinary team spanning experimental biologists, behavioral geneticists, and neuroscientists
- **Hasso-Plattner Institute & Berlin Institute for Medical Systems Biology** Berlin, Germany  
*Fulbright Scholar* Sep. 2019 - Mar. 2020
  - Visiting researcher funded by Fulbright Research Award working with Drs. Roland Schwarz and Christoph Lippert.
- **University of Southern California** Los Angeles, CA  
*Undergraduate Researcher* Feb. 2015 - Jun. 2017
  - Performed integrative statistical analysis of multi-omics datasets to study cancer metabolism
- **Harvard Medical School** Boston, MA  
*Research Intern* Jun. 2016 - Aug. 2016
  - Participant in Harvard-MIT Health Sciences and Technology (HST) Summer Bioinformatics Program
  - Developed algorithm for identifying enrichment of transcription factor binding motifs in ChIP-seq data

## SKILLS

---

- Programming languages: Python, R, MATLAB, SQL, bash
- Tools: TensorFlow, PyTorch, scikit-learn, scipy, numpy, pandas, git, HuggingFace, LangChain, Databricks, Apache Airflow, PySpark
- Machine learning & statistics: neural networks, deep learning, uncertainty quantification, explainable AI, knowledge distillation, supervised & unsupervised learning, representation learning, statistical modeling, NLP, LLMs, MC methods, generalized linear models, maximum likelihood estimation, hypothesis testing
- High performance computing: Microsoft Azure, AWS, SGE, slurm

## PUBLICATIONS

---

1. A. Sarkar, Y. Kang, N. Somia, P. M. Puccetti, **J. L. Zhou**, M. Nagai, Z. Tang, C. Zhao & P. K. Koo. Designing DNA With Tunable Regulatory Activity Using Score-Entropy Discrete Diffusion. *bioRxiv* (2025).
2. **J. L. Zhou**, K. Guruvayurappan, S. Toneyan, H. V. Chen, A. R. Chen, P. Koo & G. McVicker. Analysis of single-cell CRISPR perturbations indicates that enhancers predominantly act multiplicatively. *Cell Genomics* **0**. ISSN: 2666-979X (Oct. 2024).
3. **J. L. Zhou**, K. Rizzo, Z. Tang & P. K. Koo. Uncertainty-aware genomic deep learning with knowledge distillation. *bioRxiv* (Nov. 2024).
4. **J. L. Zhou**, G. de Guglielmo, A. J. Ho, M. Kallupi, N. Pokhrel, H.-R. Li, A. S. Chitre, D. Munro, P. Mohammadi, L. L. G. Carrette, O. George, A. A. Palmer, G. McVicker & F. Telese. Single-nucleus genomics in outbred rats with divergent cocaine addiction-like behaviors reveals changes in gene amygdala GABAergic inhibition. *Nat Neurosci*, 1–12. ISSN: 1546-1726 (Oct. 2023).
5. A. Flores, J. Schell, A. S. Krall, D. Jelinek, M. Miranda, M. Grigorian, D. Braas, A. C. White, **J. L. Zhou**, N. A. Graham, T. Graeber, P. Seth, D. Evseenko, H. A. Collier, J. Rutter, H. R. Christofk & W. E. Lowry. Lactate dehydrogenase activity drives hair follicle stem cell activation. *Nat Cell Biol* **19**, 1017–1026. ISSN: 1476-4679 (Sept. 2017).

## PRESENTATIONS

---

- *Learning to predict model uncertainty in genomic deep learning through knowledge distillation*. Poster presentation at MLGenX Workshop at ICLR 2025. April 27, 2025.
- *Learning to predict model uncertainty in genomic deep learning through knowledge distillation*. Poster presentation at Intelligent Systems for Molecular Biology. July 12-16, 2024.
- *Predicting model uncertainty in genomic deep learning with knowledge distillation*. Poster presentation at Biology of Genomes. May 7-11, 2024.
- *Genome-wide analysis of CRISPR perturbations indicates that enhancers act multiplicatively, but provides no evidence for epistatic-like enhancer interactions*. Selected talk at the International Conference for Intelligent Systems for Molecular Biology. July 23-27, 2023.
- *Single-cell sequencing links reduced glucose metabolism to cocaine addiction-like behavior in rats*. Poster presentation at the American Society of Human Genetics Annual Meeting. Oct. 25-29, 2022. Recipient of Reviewer's Choice Award.

- *Cocaine addiction is associated with long-term changes in gene regulation, metabolic pathways, and GABAergic inhibition within the amygdala.* Selected talk at the Complex Traits Community conference. September 29-30, 2022.
- *snRNA-seq reveals that reduced glucose metabolism is linked in cocaine addiction in HS rats.* Selected talk at the NIDA P50 Center for GWAS in Outbred Rats Retreat. Nov. 4, 2022.
- *Cell type-specific multiomic analysis of substance use disorders in outbred rats.* Selected talk at the Probabilistic Modeling in Genomics conference. May 14-16, 2021.
- *A novel statistical method for identifying combinatorial regulatory elements from multiplexed single-cell CRISPR regulatory screens.* Poster presentation at the Probabilistic Modeling in Genomics conference. October 6-9, 2019.
- *Enhanced breast cancer risk prediction from imputed gene expression.* Poster presentation at the American Society of Human Genetics Annual Meeting. October 17, 2018.

## AWARDS

---

NIH Ruth L. Kirschstein Predoctoral Individual National Research Service Award (F31)  
 UCSD Interdisciplinary Research Award  
 Fulbright Study/Research Award  
 National Science Foundation Graduate Research Fellowship Program Honorable Mention  
 University of California Trustee Scholar  
 University of Southern California Provost's Research Fellowship  
 Women in Science and Engineering Research Fellowship  
 Rose Hills Foundation Research Fellowship

## TEACHING & OUTREACH

---

- **American Society of Human Genetics** USA  
*Public Education and Awareness Committee Member* *Jan. 2020 - Dec. 2022*
  - Brainstorm subjects for public-facing online Discover Genetics fact sheets and advise on content
  - Run and judge annual ASHG DNA Day Essay Contest for high schoolers
  - Participate in AskScience Reddit panels
- **Fleet Science Center** San Diego, CA  
*Volunteer* *2017-2023*
  - Volunteer with Two Scientists Walk Into a Bar program, where pairs of local scientists visit local breweries and invite the public to ask questions about science
  - Volunteer with BE WiSE program, which engages local young women (grades 7-12) in STEM
- **UC San Diego** La Jolla, CA  
*Bioinformatics Bootcamp Instructor* *Sep. 2018 - Sep. 2019*
  - Designed and taught lectures on statistical concepts (e.g. p-values, statistical distributions, OLS, MLE) paired with interactive Jupyter notebooks to incoming first year students in the Bioinformatics PhD program
  - Planned and organized welcome week events for the Bioinformatics PhD program

**University of Southern California**

Los Angeles, CA

- *MATH 226 Supplemental Instructor*

*Aug. 2015 - May 2017*

- Designed and taught weekly lectures (3x/week) to supplement undergraduate vector calculus course