Shiyu Jiang

Last modified: December 8, 2025

Dept. of Quantitative and Computational Biology (QCB)
USC Dornsife College of Letters, Arts and Sciences
Ray R. Irani Hall, 1050 Childs Way
Los Angeles, CA 90089
Google Scholar: citations 182; h-index 7; i10-index 6

Phone: (667) 379-4478
E-mail: shiyujia@usc.edu
Homepage: https://jasonjiangs.github.io/
GitHub: https://github.com/JasonJiangs

Research Interests

- Deep Learning Methods for Structural Biology
- Large Language Model for Life Sciences
- Computational Genomics and Systems Biology
- AI-driven Drug Discovery

Education and Training

- Ph.D. student in Computational Biology and Bioinformatics — 08/2025 – Present
  Department of QCB, **University of Southern California** (USC), Los Angeles, CA, USA.
- MSE in Computer Science, Bioinformatics Track — 08/2022 – 12/2023
  Department of Computer Science, **Johns Hopkins University**, Baltimore, MD, USA.
  Thesis Advisor: Prof. Kimia Ghobadi
  Committee: Kimia Ghobadi, Anton Dahbura, Claus Aranha
  Thesis: Analyzing epidemic spread and emotion contagion with agent-based simulation and system dynamic modeling
- Undergraduate Visiting Student — 01/2022 – 06/2022
  Institute of Automation, **Chinese Academy of Sciences**, Beijing, China
  Advisor: Prof. Zhen Shen
- B.S. in Computer Science; Minor in Mathematical Sciences — 08/2018 – 05/2022
  Department of Computer Science, **Wenzhou-Kean University**, Wenzhou, China
  Advisor: Prof. Aloysius Wong

Peer-reviewed Publications and Patents

(† ‡ = equal contribution; * = co-corresponding author)

Journal articles

1. **Jiang S†**, Taghavi A†, Wang T, Meyer SM, Childs-Disney JL, Li C, Disney MD*, Li Y*. Small Molecule Approach to RNA Targeting Binder Discovery (SMARTBind) Using Deep Learning Without Structural Input. bioRxiv, doi: 10.1101/2025.09.24.678312. (Major revision in Nature research Journal)
2. Cheng L†, Zheng X†, **Jiang S†**, Hu Y, Liu Y, Yang K, Rui J, Ding H, Zhang M, Yuan T, Ye H, Li C, Yang KK, Huang X*, Xiao H*. Sequence Display: Generating Large-Scale Sequence–Activity Datasets to Advance Universal Protein Evolution. (Major revision in Nature research Journal)
3. Yuan T†, Zhang M†, Cheng L, Zheng X, **Jiang S**, Huang X, Xiao H*. Biocatalytic Synthesis of N-protected α-Amino Acids through 1,3-Nitrogen Migration by Nonheme Iron Enzymes. ***Journal of the American Chemical Society***, (2025).
4. Su J, Li Z, Han C, Zhou Y, He Y, Shan J, Zhou X, Chang X, **Jiang S**, Ma D, OPMC, Steinegger M, Ovchinnikov S, Yuan F*. Democratizing protein language model training, sharing and collaboration. ***Nature Biotechnology***, (2025): 1-7.

5. Hu Y, Wang Y, Cheng L, Wang C, Liu Y, Wang Y, Chen Y, Yang S, Guo Y, **Jiang S**, Yang K, Xiao H*. Engineering unnatural cells with a 21st amino acid as a kiving epigenetic sensor. ***Nature Communications***, 16.1 (2025), 9388.
6. Su J†, He Y†, You S†, **Jiang S**, Zhou X, Zhang X, Wang Y, Su X, Tolstoy I, Chang X*, Lu H*, Yuan F*. A tri-modal protein language model enables advanced protein searches. ***Nature Biotechnology***, (2025): 1-7.
7. Hu Y†, Cheng L†, Liu Y, Liu R, **Jiang S**, Yuan T, Wang Y, Ye H, Xiao H*. Biosynthesis of Unnatural Cyclodipeptides through Genetic Code Expansion and Cyclodipeptide Synthase Evolution. ***Journal of the American Chemical Society***, 2025, 147, 38, 34517–34526.
8. Ye H, **Jiang S**, Yan Y, Zhao B, Grant ER, Kitts DD, Yada RY, Singh AP, Baldelli A, Yang T*. Integrating Metal Phenolic Networks-Mediated Separation and Machine Learning-Aided SERS for High-Precision Quantification and Classification of Nanoplastics. ***ACS Nano***, 18.38 (2024): 26281-26296. **[Cover article]**
9. Guo Y†, Cheng L†, Hu Y, Zhang M, Liu R, Wang Y, **Jiang S**, Xiao H*. Biosynthesis of halogenated tryptophans for protein engineering using genetic code expansion. ***ChemBioChem*** 25.20 (2024): e202400366. **[VIP paper]**
10. **Jiang S**, Abdalla H, Bi C, Zhu Y, Tian X, Yang Y*, Wong A*. HNOXPred: a web tool for the prediction of gas-sensing H-NOX proteins from amino acid sequences. ***Bioinformatics***, 38(19), 4643-4644, 2022.

Conference articles

11. Zheng X, **Jiang S**, Seabra G, Li C, Li Y*. Apo2Mol: 3D Molecule Generation via Dynamic Pocket-Aware Diffusion Models. In ***AAAI 2026***.
12. Ding J†, Lin J†, **Jiang S†**, Wang Y, Mao Z, Fang Z, Tang J*, Li M*, Qiu X*. Tabula: A Tabular Self-Supervised Foundation Model for Single-Cell Transcriptomics. In ***NeurIPS 2025***.
13. Qi X†, Zhang Z†, Handoko AB†, Zheng H, Chen M, Huy TD, Ohan VMH, Zhang L, Cheng L, **Jiang S**, Zhang Z, Liao Z, Zhao Y*, To MS*. ProjectedEx: Enhancing Generation in Explainable AI for Prostate Cancer. In ***2025 IEEE 38th International Symposium on Computer-Based Medical Systems (CBMS)***.
14. Song Z†, Ouyang G†, Fang M†*, Na H, Shi Z, Chen Z, Fu Y, Zhang Z, **Jiang S**, Fang M, Chen L, Chen X*. Hazards in Daily Life? Enabling Robots to Proactively Detect and Resolve Anomalies. In ***NAACL 2025***.
15. **Jiang S***, Kim H, Tanaka F, Aranha C*, Bogdanova A, Ghobadi K, Dahbura A. Simulating Disease Spread during Disaster Scenarios. In ***2023 Conference on Artificial Life (ALIFE)***.
16. Zheng S, Wu Y, **Jiang S**, Lu C, Gupta G*. Debulr-yolo: Real-time object detection with efficient blind motion deblurring. In ***2021 International Joint Conference on Neural Networks (IJCNN)***.

Workshops

17. **Jiang S***, Liu X, Wang JZ*. Predicting function of evolutionarily implausible DNA sequences. In ***ICML 2025 Workshop for Generative AI and Biology***.

Preparing manuscripts

18. Zhou X†, Han C†, Zhang Y‡, Su J‡, Zhuang K‡, **Jiang S**, Yuan Z, Zheng W, Dai F, Zhou Y, Tao Y, Wu D, Yuan F*. Decoding the Molecular Language of Proteins with Evolla. bioRxiv, doi: 10.1101/2025.01.05.630192.
19. Ding J†, Lin J†, **Jiang S†**, Wang Y, Mao Z, Fang Z, Tang J*, Li M*, Qiu X*. Toward a privacy-preserving predictive foundation model of single-cell transcriptomics with federated learning and tabular modeling. bioRxiv, doi: 10.1101/2025.01.06.631427. (Journal version preparation)

Patents

1. Xiao H, Cheng L, **Jiang S**, Rui J, Huang X. “Sequence Display enable ML.” *U.S. Patent Application* (Institutionally issued, pending official publication).

Professional Experience

- Graduate Research Assistant, **University of Southern California** 08/2025 – Present

- Remote Research Assistant, **Rice University** 03/2024 – 08/2025
  - Advisor: Prof. Han Xiao
  - Collaborator: Prof. Xiongyi Huang (Johns Hopkins Univ.), Dr. Kevin K. Yang (Microsoft Research)
  - Sequence display enables machine learning for protein evolution.
- Remote Research Assistant, **University of Florida** 08/2023 – 07/2025
  - Advisor: Prof. Yanjun Li, Prof. Chenglong Li
  - Collaborator: Prof. Matthew D. Disney (Scripps Research), Prof. Gustavo Seabra (Univ. of Florida)
  - RNA-ligand interaction modeling; dynamic molecule generation.
- Research Associate, **Westlake University** 08/2024 – 07/2025
  - Advisor: Prof. Fajie Yuan, Dr. Zitong Jerry Wang
  - Development and evaluation of protein language model and genomic language model.
- Visiting Master Student, **Stanford University School of Medicine** 08/2023 – 07/2025
  - Advisor: Prof. Xiaojie Qiu
  - Collaborator: Prof. Wei Ouyang (KTH), Dr. Ryan Lu (MIT)
  - Single-cell foundation model development, genetic perturbation, gene regulation modeling.
- Lab Specialist, **University of Virginia School of Medicine** 01/2024 – 07/2024
  - Advisor: Prof. Chongzhi Zang, Prof. Sheng'en Shawn Hu
  - ChIP-seq broach peak calling tool (SICER) development.
- Master Research Assistant, **Johns Hopkins University** 10/2022 – 07/2023
  - Advisor: Prof. Kimia Ghobadi, Prof. Anton Dahbura
  - Collaborator: Prof. Claus Aranha (University of Tsukuba)
  - Agent-based modular simulator for small-community disease spread.
- Software Engineer Intern, **Alibaba Cloud** 06/2022 – 08/2022
  - Mentor: Mr. Jiabang Pan
  - PolarDB database development (Java and MySQL).
- Undergraduate Research Assistant, **Wenzhou-Kean University** 10/2021 – 05/2022
  - Advisor: Prof. Aloysius Wong
  - Bioinformatics webserver development for protein motif recognition.

Honors & Awards

- USC Dana and David Dornsife College Graduate School Fellowship 2025

Posters & Presentations

- Small Molecule Approach to RNA Targeting Binder Discovery (SMARTBind) Using Deep Learning Without Structural Input 01/29/2026
  Poster presentation. **UF Health Cancer Institute Annual Research Showcase 2026**, University of Florida, Gainesville, FL, USA
- Tabula: A Tabular Self-Supervised Foundation Model for Single-Cell Transcriptomics 12/02/2025
  Poster presentation. **NeurIPS 2025**, Hilton Reforma & San Diego Convention Center, San Deigo, CA, USA
- Predicting function of evolutionarily implausible DNA sequences 07/23/2025
  Poster presentation. **Q-BIO 2025 Conference: Emergent Orders in Living Systems Across Scales**, Center of Quantitative Bilogy (CQB), Peking University (PKU), Beijing, China
- Predicting function of evolutionarily implausible DNA sequences 07/18/2025
  Poster presentation. **ICML 2025, Workshop for Generative AI and Biology**, Vancouver, Canada
- Sequence Display-Enable Machine Learning for Protein Evolution 06/23/2025
  Poster presentation. **2025 Synthetic Biology: Engineering, Evolution, & Design**, Houston, TX
- Simulating Disease Spread during Disaster Scenarios 07/23/2023
  Talk presentation. **The 2023 Conference on Artificial Life**, Sapporo, Japan

Professional Service

- Journal Reviewer: IEEE Transactions on Computational Biology and Bioinformatics (1 paper).
- Conference Reviewer: AAAI 2026 (4 papers).