

Yusuf Roohani
yroohani@alumni.stanford.edu
www.yusufroohani.com
Palo Alto, CA 94306

Education

2020–24 **Ph.D.**, Biomedical Data Science, Stanford University GPA: 4.10/4.25
Dissertation Title: *Engineering Cells using Artificial Intelligence*
Advisors: Jure Leskovec, Stephen Quake
Committee: Stanley Qi, Barbara Engelhardt, Daphne Koller, Silvana Konermann

2014-15 **M.S.**, Mechanical Engineering, Carnegie Mellon University GPA: 4.00/4.00
2009-13 **B.Tech.** Mechanical Engineering, Vellore Institute of Technology GPA: 8.81/10.00

Work Experience

2026-Present **Arc Institute**, Palo Alto, CA
Associate Director, Machine Learning

2024-2025 **Group Lead, Machine Learning**
Building an AI model of cell state to predict the outcomes of experiments in-silico.
Leading model development as well as AI-guided data generation.
3 papers published/in review at Cell.
Built a new interdisciplinary team of 10+ ML research scientists, engineers
Led development of the largest AI agent-curated biological data repository to date
Member of Arc Institute AI Advisory Board

2024-2025 **Stanford University**, Stanford, CA
Visiting Scholar: Department of Computer Science

2020-2024 **PhD Student: Department of Biomedical Data Science**
Foundation models for cell biology + LLM-driven agents for experiment design.


2019-2020 **GlaxoSmithKline**, Cambridge, MA
Manager, Machine Learning Engineer

2017-2019 **Investigator**

2016-2017 **Data Scientist**
Led a cross-disciplinary team to biologically profile a 2M+ compound collection using multi-modal datasets and high-throughput screening.

2015-2016 **Merrimack Pharmaceuticals**, Cambridge, MA
Computational Modeler Intern
Developed dynamic system models to simulate signaling networks in cancer cells.

Publications

 [Google Scholar Page](#) † → Equal contribution

Journal Articles

- J1. Adduri, A., Gautam, D., Bevilacqua, B., ... & **Roohani, Y.** Predicting cellular responses to perturbation across diverse contexts with STATE. *Cell (in press)* (2026).
- J2. Rosen Y.[†], **Roohani Y.[†]**, Agarwal, A., Samotorčan, L., ... Quake, S. R. & Leskovec, J. Universal cell embeddings: A foundation model for cell biology. *Nature (in press)* (2026).
- J3. Chevalley, M., **Roohani, Y.**, ..., Leskovec, J. & Schwab, P. A large-scale benchmark for network inference from single-cell perturbation data. *Communications Biology* **8**, 412 (2025).
- J4. **Roohani, Y.**, Hua, T. J., Tung, P.-Y. & ... Burke, D. Virtual Cell Challenge: Toward a Turing test for the virtual cell. *Cell* **188**, 3370–3374 (2025).
- J5. Bunne C.[†], **Roohani Y.[†]**, Rosen Y.[†], Gupta, A., *et al.* How to build the virtual cell with artificial intelligence: Priorities and opportunities. *Cell* (2024).
- J6. **Roohani, Y.**, Huang, K. & Leskovec, J. Predicting transcriptional outcomes of novel multi-gene perturbations with GEARS. *Nature Biotechnology* (2024).
- J7. Rosen, Y.[†], Brbić, M.[†], **Roohani Y.[†]**, Swanson, K., Li, Z. & Leskovec, J. Toward universal cell embeddings: integrating single-cell RNA-seq with SATURN. *Nature Methods* (2024).
- J8. Huang, K., Fu, T., Gao, W., Zhao, Y., **Roohani, Y.**, Leskovec, J. & ... Zitnik, M. Artificial intelligence foundation for therapeutic science. *Nature Chemical Biology* (2022).
- J9. **Roohani, Y.**, ..., Robinson, A. L. & Adams, P. J. Impact of natural gas development in the Marcellus and Utica shales on regional ozone and PM_{2.5}. *Atmospheric Environment* (2017).

Peer-reviewed Conference Proceedings

- C1. **Roohani Y.[†]**, Lee A.[†], Huang Q.[†], ..., Marson, A., Liang, P. & Leskovec, J. BioDiscoveryAgent: An AI Agent for Designing Genetic Perturbation Experiments. *ICLR* (2025).
- C2. Nilforoshan, H.[†], Moor, M.[†], **Roohani, Y.**, Chen, Y., Šurina, A., Yasunaga, M., Oblak, S. & Leskovec, J. Zero-shot causal learning. *NeurIPS* (2023).
- C3. Huang, K., Fu, T., Gao, W., Zhao, Y., **Roohani, Y.**, ... & Zitnik, M. Therapeutics data commons: Machine learning datasets and tasks for drug discovery. *NeurIPS* (2021).
- C4. **Roohani, Y.**, Sajid, N., Madhyastha, P., Price, C. J. & Hope, T. M. Predicting language recovery after stroke with convolutional networks on stitched mri. *NeurIPS ML4H* (2018).

Preprints

- W1. Dong, M., Adduri, A., Gautam, D., Carpenter, C., Shah, R., Ricci-Tam, C., ... & **Roohani, Y.** STACK: In-Context Learning of Single-Cell Biology. *Under review at Nature* (2025).
- W2. Huang, K., Zhang, S., Wang, H., Qu, Y., Lu, Y., **Roohani, Y.**, Li, R., Qiu, L., Li, G., Zhang, J., *et al.* Biomni: A general-purpose biomedical ai agent. *Under review at Science* (2025).
- W3. Youngblut, N., Carpenter, C., ..., Goodarzi, H.[†] & **Roohani, Y.[†]**. scBaseCount: an AI agent-curated, uniformly processed single cell data repository. *Under review at Cell* (2025).
- W4. Magnusson J.[†], **Roohani Y.[†]**, ... & Qi, L. S. PreciCE: Precision engineering of cell fates via multi-gene control of transcriptional networks. *Under review at Mol. Cell* (2024).
- W5. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., **Roohani, Y.**, ., ... & Liang, P. On the opportunities and risks of foundation models. *arXiv* (2021).

Patents

- P1. Adduri, A., Roohani, Y., Goodarzi, H. & Burke, D. A Scalable Approach for Learning Perturbation Effects Over Sets of Cells 63/823,201. *U.S. Patent Application* (Filed June 2025).

Invited Talks

- 2026 Generative and Agentic AI for Biology, ICML, Seoul, South Korea
2026 Innovations in Single Cell Omics, Barcelona, Spain
2026 Google DeepMind, London, UK
2026 Gordon Research Conference, Single Cell Genomics, Les Diablerets, Switzerland
2026 Fred Hutchinson Cancer Center, Seattle, WA
2026 ISCB-China Workshop on the Virtual Cell, Sanya, China
2026 Models, Inference and Algorithms Seminar, Broad Institute, Cambridge, MA
2026 Joint Statistical Meeting, Boston, MA
2026 International Plant and Genome Conference, San Diego, CA
2026 NIH Brain Cell Atlas: From Data to Knowledge Workshop, Bethesda, MD
2025 Common Fund Faculty Symposium, Jackson Laboratory, ME
2025 Keynote: Novartis Advances in Foundation Models Workshop, Cambridge, MA
2025 AstraZeneca AIxBio Frontiers Seminar, Cambridge, MA
2025 Columbia University Biomedical Informatics Seminar, New York, NY
2025 UCLA Bioengineering Seminar, Los Angeles, CA
2025 MIT AI+BIO hosted by MIT CSAIL, MIT Biology and NSF, Cambridge, MA
2025 Genentech Research and Early Development, South San Francisco, CA
2025 Aging Research and Drug Discovery, Copenhagen, Denmark
2025 Stanford AI & Biomedicine Seminar, Stanford, CA
2025 Xaira Therapeutics, South San Francisco, CA
2025 Recursion Pharmaceuticals, Salt Lake City, UT
2025 Guest Lecture, Department of Bioengineering and Therapeutic Sciences, UCSF
2025 Society for Lab Automation and Screening Conference 2025
2024 US Science Day, UCB Pharma, Cambridge, UK
2024 Stanford Graph Learning Workshop 2024, Stanford, CA
2024 AstraZeneca, Cambridge, UK
2023 Arc Institute, Palo Alto, CA
2023 EvolutionaryScale, New York, NY
2023 Stanford Graph Learning Workshop 2023, Stanford, CA
2023 Genentech Research and Early Development, South San Francisco, CA
2022 Altos Labs, Palo Alto, CA
2022 Stanford Network Biology Journal Club, Kundaje/Bassik Lab, Stanford Genetics
2022 Recursion Pharmaceuticals, Salt Lake City, UT
2022 Stanley Qi Lab Group Meeting, Stanford Bioengineering
2020 Society for Lab Automation and Screening Conference 2020
2019 Guest Lecture, School of Engineering, Tufts University
2018 ReWork, Deep Learning for Healthcare, Boston

Student Mentorship

2025 Dhruv Gautam, MS Computer Science, UC Berkeley
 2025 Mingze Dong, PhD Computational Biology, Yale University
 2025 Rohan Shah, BS Computer Science, University of Pennsylvania
 2025 Alishba Imran, BS Computer Science, UC Berkeley
 2024 Sanjay Nagaraj, BS Computer Science, Stanford University
 2024 Andrew Lee, MS Computer Science, Stanford University
 2023 Jaanak Prashar, BS Computer Science, Stanford University
 2023 Ayush Agrawal, BS Computer Science, Stanford University

Teaching

2021 Teaching Assistant, CS-270: Modeling Biomedical Systems, Stanford University
 2022 Teaching Assistant, CS-270: Modeling Biomedical Systems, Stanford University

Reviewing

Nature, Nature Biotechnology, Nature Aging, Nature Methods, NeurIPS, PLOS Biology, PLOS Computational Biology, Nucleic Acids Research, MICCAI

Software

GEARS	300+ GitHub stars
Universal Cell Embeddings (UCE)	200+ GitHub stars
BioDiscoveryAgent	90+ GitHub stars

Awards & Honors

2024 **Best Poster** ICLR 2024 Machine Learning for Genomics Exploration Workshop
 2023 **Best Poster** Top 3 out of 200 at Stanford BioX Interdisciplinary Initiatives (\$500)
 2023 **Innovation Award** Society for Lab Automation and Screening (\$10K)
 2022 **Best Poster** Intelligent Systems for Mol. Biology (ISMB) (Machine Learning Track)
 2022 **Best Poster** Single Cell Genomics meets Data Science, Munich (\$500)
 2019 **Full PhD funding** Awarded by GSK, including tuition + full-time pay (\$1M+)
 2018 **GSK Exceptional Science Award** Deep learning for cellular images (\$17K)
 2010-13 **Merit Certificates, Academic Excellence** (x4), Vellore Inst. of Tech.
