

JONATHAN LIU

jhkliu42@gmail.com ♦ 510-366-8951 ♦ www.jonathan-liu.com ♦ San Francisco, CA

Expert in single cell data science, high throughput phenotypic screening, and multimodal analysis.

Passionate about empowering cross-functional collaboration between experimentation and computation.

EDUCATION

University of California, Berkeley
PhD in Physics

Aug 2016 - Aug 2021

California Institute of Technology
BS in Applied Physics with Honors

Sept 2011 - June 2015

TECHNICAL SKILLS

Programming:	Python, R, MATLAB, Java, Mathematica
Data Science:	bash, Nextflow, AWS, git, Docker, Pandas, scikit-learn, PyTorch
Statistics:	Hypothesis testing, Bayesian inference, stochastic modeling, applied ML
Computer Vision:	Classical and ML-based segmentation, spot detection, temporal analysis
Bio-imaging:	live imaging, RNA-FISH, <i>in situ</i> sequencing, IHC, histopathology
Bioinformatics:	RNA-seq (bulk, single cell/nucleus), perturb-seq, spatial transcriptomics
Mathematics:	Calculus, linear algebra, differential equations, probability, statistics

SCIENTIFIC EXPERIENCE

insitro

Senior Data Scientist

Sep 2024 - Present

Lead Data Scientist

Jun 2023 - Aug 2024

- Founding member of computational biology team, assisted with hiring to grow team to 6+ in a year
- Devised strategy and led analysis for high throughput single cell CRISPR-based screening of iPSC-derived cell lines in diverse modalities (pooled optical screening, perturb-seq) for 3 internal pipelines
- Integrated results from *in vitro* experiments with human clinical biomarker data and genetic evidence to enable multimodal investigation of disease biology
- Collaborated with wet lab and ML teams to develop self-supervised applied ML methodologies for drug target discovery – highlights include successfully reaching collaboration milestone with external partner for \$10M cash payment and several nominated targets for preclinical drug discovery
- Tech lead for neuroscience assay development (fluorescence-based calcium imaging, multi-electrode arrays), focusing on scaling up proof-of-concept experiments to productionized internal platforms
- Led statistical meta-analysis of high throughput screening strategies to optimize size, spend, and choice of platform for upcoming large-scale screens
- Mentored two data scientists; managed one intern

Chan Zuckerberg Biohub San Francisco

Data Scientist II

Jan 2023 - Jun 2023

Data Scientist I

Aug 2021 - Dec 2022

Contract Data Scientist

Jan 2021 - Jun 2021

- Head of spatial transcriptomics working group; achievements include securing \$300k purchase of Vizgen MERSCOPE instrument, winning \$10k grant awarded from 10x Genomics for Visium research, and publishing first-author manuscript of subsequent technical benchmarking analysis
- Managed numerous active projects involving both internal researchers and external scientific collaborators; used spatial transcriptomics to study systems such as developing zebrafish embryos, pathogen-infected mouse lungs, and memory formation in mouse brains. Highlights include publication in *Nature* investigating long-term memory in neuron-astrocyte synergy.

- Built and maintained internal image analysis and single-cell bioinformatics pipelines in Python for Vizgen MERSCOPE data processing and downstream analysis (TB of data per experiment)
- Managed internal and external seminar series; duties included fostering internal dissemination of on-going research and inviting external speakers from academia and industry
- Mentored three research associates, one scientist, and two junior data scientists; managed one intern
- Co-led hiring search for one research associate and one scientist, resulting in two successful hires

UC Berkeley

NDSEG Graduate Research Fellow

Aug 2017 - Aug 2021

Graduate Student Researcher

Jun Aug 2016 - Jul 2017

- Investigated biophysical models of gene regulation through live imaging fluorescence microscopy of nascent RNA transcription in eukaryotes, with 2 first-author publications and 1 third-author publication
- Used image analysis tools such as machine learning segmentation to convert single-cell microscopy data into large time series (several TB, hundreds of cells over minutes with time resolutions of seconds)
- Applied Markov Chain Monte Carlo statistical inference techniques to generate single-cell datasets of kinetic transcriptional parameters
- Developed and investigated theoretical models of gene regulation (e.g. deterministic ODEs, stochastic simulations)
- 4 trainees supervised (2 graduate, 2 undergraduate)

OTHER EXPERIENCE

Beyond Academia - Co-director, team member

Aug 2019 - Jun 2021

- Co-director of volunteer organization that hosts an annual two-day conference for 300+ current PhDs and postdocs and features over 100 speakers
- Managed transition to virtual conference due to COVID-19 pandemic and oversaw 10x growth (3000+ registrations across 6 continents with >50% attendance) with widespread attendee satisfaction
- Fundraised and wrote grant proposals to secure \$60k yearly budget
- Project management experience organizing ~10 member team in tasks involving event logistics, conference speaker recruitment, finance, and media outreach

Mentorship and Communication

- Wrote 4 layperson-targeted articles for *Berkeley Science Review*, QB3-Berkeley, and *Physics Today*
- Invited speaker at various science events, highlights include:
 - Frontline Genomics Spatial Biology (2022) - subject matter expert on spatial biology webinar targeting broad biologist audience
 - WonderFest (2021) - long-standing Bay Area science outreach program
 - STIx on the Hill (2019) - sole graduate student speaker at DoD science policy event
- Advisor for the Berkeley Summer Undergraduate Research Fellowship program, providing mentorship and support for several hundred undergraduate researchers (*Summer 2019*)

SELECTED PUBLICATIONS

For a full publication list see my Google Scholar profile.

1. Wenfei Sun, Zhihui Liu, Xian Jiang, Michelle B Chen, Hua Dong, **Jonathan Liu**, *et. al.* "Spatial transcriptomics reveal neuron-astrocyte synergy in long-term memory." *Nature*, Mar 2024.
2. Rahul Atmaramani, *et. al.* "Deep Learning Analysis on Images of iPSC-derived Motor Neurons Carrying fALS-genetics Reveals Disease-Relevant Phenotypes." *bioRxiv*, Jan 2024.
3. **The COVID Tissue Atlas Consortium.** "Single-nuclei characterization of pervasive transcriptional signatures across organs in response to COVID-19." *eLife*, Oct 2023.

4. **Jonathan Liu***, Vanessa Tran*, *et. al.* “Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing.” *Life Science Alliance*, Dec 2022. (*equal authorship)
5. **Jonathan Liu**, *et. al.* “Real-time single-cell characterization of the eukaryotic transcription cycle reveals correlations between RNA initiation, elongation, and cleavage.” *PLoS Computational Biology*, May 2021.
6. Elizabeth Eck*, **Jonathan Liu***, *et. al.* “Quantitative dissection of transcription in development yields evidence for transcription factor-driven chromatin accessibility.” *eLife*, Oct 2020. (*equal)
7. Matthias Morasch, **Jonathan Liu**, *et. al.* “Heated gas bubbles enrich, crystallize, dry, phosphorylate, and encapsulate prebiotic molecules.” *Nature Chemistry*, Jul 2019 (cover article).

HONORS AND AWARDS

2022 CytoData Hackathon Prize: Member of winning team in CytoData Symposium 2022 hackathon focusing on image analysis and computational biology

2017 NDSEG Graduate Fellowship: 1 of 195 graduate students selected for four-year fully-funded graduate research fellowship, supported by the Department of Defense. (\$200k total)

2015 U.S. Fulbright Student Fellowship: 1 of about 100 students selected for yearlong visiting research appointment in Germany, funded by the U.S. and German Fulbright programs. (\$15k total)

OTHER WRITING

I enjoy writing as a medium for expressing my thoughts on topics such as graduate education in STEM, the intersection of science and community, and science communication. I've published articles in venues such as Physics Today and QB3 Berkeley. See <https://www.jonathan-liu.com/#writing> for full list of publications.