Yutong Wang

ytwang@berkeley.edu https://yu-tong-wang.github.io

RESEARCH INTERESTS

statistical machine learning, high dimensional statistics, probabilistic modeling, statistical genetics
high-throughput sequencing, spatial transcriptomics, multi-omics, CRISPR-screening, functional genomics, immunology

I am passionate about developing and applying rigorous computational methods to elucidate complex biological landscapes, and advance our understanding of pressing biomedical frontiers.

EDUCATION

Ph.D. in Biostatistics *University of California, Berkeley, CA* 2020 - Present
Minor: Computational and Genomic Biology
Advisor: Yun S. Song

M.A. in Biostatistics *University of California, Berkeley, CA* 2018-2020
Thesis: Joint Integration Analysis of Single-cell RNA Sequencing and Imaging Data to Infer Adipogenesis and Lipid Maturation
Committee: Yun S. Song, Aaron Streets, Haiyan Huang

B.S. in Mathematics and Applied Mathematics *Tianjin University, China* 2014-2018

PUBLICATIONS & MANUSCRIPTS

* indicates equal contribution, † indicates co-corresponding authors.

Manuscripts

1. **Yutong Wang***, Fanding Zhou*, Antoine Koehl, Yun S. Song, Decoding cell-cell communication through deep learning reveals novel ligand-receptor pairs from spatial transcriptomics. Machine Learning in Computational Biology (MLCB) 2023, Selected as Oral (top 15% of submissions).
2. **Yutong Wang**, Anushka Gupta, Chen Cao, Rodrigo Cotrim Chaves, Xiannian Zhang, Aaron Streets, Yun S. Song, Paired label-free imaging and RNA-sequencing of mature adipocytes at the single-cell level. In Preparation. (2023+)
3. **Yutong Wang**, Yun S. Song, Spatial Segmentation using Mixture Experts. In Preparation. (2023+)

Refereed Journal Articles

1. Youjin Lee*†, Derek Bogdanoff*, **Yutong Wang***, George Hartoularos, Jonathan M. Woo, Cody T. Mowery, Hunter M. Nisonoff, David S. Lee, Yang Sun, James Lee, Sadaf Mehdizadeh, Joshua Cantlon, Eric Shifrut, David N. Ngyuen, Theodore L. Roth, Yun S. Song, Alexander Marson†, Eric D. Chow†, Chun Jimmie Ye†, XYZeq: Spatially-resolved single-cell RNA-sequencing reveals expression heterogeneity in the tumor microenvironment. *Science Advances* **7**, eabg4755 (2021)

PROFESSIONAL EXPERIENCE

Graduate Student Researcher *University of California, Berkeley, CA* 2019 - Present
Develop and implement novel statistical methods and machine learning algorithms for large-scale high-throughput sequencing data, including spatial transcriptomics, multi-modality data analysis to understand cell-cell communication and characterize heterogeneity in cellular state and functions.

Computational Biologist Summer Intern *Chan Zuckerberg, Biohub, Inc., San Francisco, CA* May - August, 2022
Proposed a novel Bayesian hierarchical partitioning model aimed at characterizing spatial transcriptomic landscapes, specifically addressing the limitations of existing methods in capturing overdispersion and spatial heterogeneity in tissue data.

Research Assistant *University of Pennsylvania, PA* 2017
Designed and implemented computational experiments to compare gene expression recovery methods for single cell RNA sequencing data.

PRESENTATIONS

Contributed Talks

1. Decoding cell-cell communication through deep learning reveals novel ligand-receptor pairs from spatial transcriptomics
   - The 18th Machine Learning in Computational Biology (MLCB2023) Conference, Seattle, WA 2023
   - IGI-EECS Machine Learning Workshop, Innovative Genomics Institute, UC Berkeley, CA 2023
2. Spatial Segmentation using Mixture Experts
   - Center for Targeted Machine Learning and Causal Inference (CTML) Lightning Talks, UC Berkeley, CA 2023
3. Joint integration analysis of paired single-cell imaging and RNA sequencing of mature adipocytes
   - Stanford-Berkeley Women in CS/EE Research Meetup, Stanford, CA 2022
   - Women in Data Science, UC Berkeley, CA 2022
4. Single-cell and spatial transcriptomics data analysis with Seurat in R
   - Doctoral Seminar in Computational Biology, UC Berkeley, CA 2023
   - Computational Biology Skills Seminar, Center for Computational Biology, UC Berkeley, CA 2021
5. XYZeq: Spatially-resolved single-cell RNA-sequencing reveals expression heterogeneity in tumor microenvironment
   - Computational Genomes Meeting, Chan Zuckerberg, Biohub, Inc., San Francisco, CA 2021
   - Computational and Genomic Biology Retreat, Center for Computational Biology, UC Berkeley, CA 2021

Contributed Posters

1. Decoding cell-cell communication through deep learning reveals novel ligand-receptor pairs from spatial transcriptomics
   - Computational and Genomic Biology Retreat, Center for Computational Biology, UC Berkeley, CA 2023

TEACHING EXPERIENCE

Doctoral Seminar in Computational Biology *COMPBIO 293* Fall 2023
Graduate Student Instructor Center for Computational Biology, University of California, Berkeley, CA
Doctoral level course covering a broad spectrum of advanced topics including dimensionality reduction, clustering, RNA-seq analysis, multi-omic and spatial transcriptomics, CRISPR technologies, deep mutation scanning, and predictive models for variant effects and polygenic risk scores.

Python Bootcamp 2022, 2023
Instructor Center for Computational Biology, University of California, Berkeley, CA
8-hour lectures to 122 participants including graduate students, postdoc and faculties about data structures, text manipulation, and file input/output.

Biostatistics Seminar *PBHLTH 292* 2022, 2023
Invited Lecturer Graduate Group in Biostatistics, University of California, Berkeley, CA
Curriculum development and delivery of 6-hour lectures to graduate students on Eugenics in Statistics, Ethics and Algorithmic Fairness in Health Care.

Algorithms for Single-Cell Genomics *COMPBIO 290* Fall 2021
Graduate Student Instructor Center for Computational Biology, University of California, Berkeley, CA
Doctoral level course on algorithms and statistical methods in single-cell genomics. Three main themes of the course are spatial transcriptomics, multi-omics integration, and immune receptor-antigen interactions.

Concepts of Statistics *STAT 135* Spring 2020
Graduate Student Instructor Department of Statistics, University of California, Berkeley, CA
Core upper-division course on statistical theory including parametric estimation, hypothesis testing, statistical tests, and linear regression.

Introduction to Statistics *STAT 2* 2018, 2019
Graduate Student Instructor Department of Statistics, University of California, Berkeley, CA
Responsible for two sections with 50 students, office hours, and grading

MENTORING

Research Mentoring *Song Lab* University of California, Berkeley, CA
- Fanding Zhou: PhD student in Biostatistics 2022 - Present

- Shiyi Yang: Rotating PhD student in Computational Biology 2022

Academic Mentoring *Peer Mentorship Program* University of California, Berkeley, CA

- Andy Kim: Graduate student in Biostatistics 2022
- Joy Nakato: Graduate student in Biostatistics 2022

OUTREACH & SERVICE

Inaugural Student Fellow in Diversity, Equity, Inclusion (DEI) *Graduate Group in Biostatistics, UC Berkeley, CA* 2021 - 2023

1. Led the DEI Committee of $10+$ members, organizing various initiatives and conducting bi-weekly strategy meetings.
2. Orchestrated a two-day Data Science Workshop aimed at providing underrepresented minority high school students equitable exposure to programming and data analysis.
3. Organized and hosted three Graduate Diversity Admissions Fairs.
4. Served as the Student Representative in the Graduate Student Admission Committee.
5. Facilitated town halls, administered climate surveys, and enhanced website accessibility to foster open communication and evaluate the department's culture.
6. Initiated a peer mentorship program and organized alumni panels to bolster community connections.

AWARDS & FELLOWSHIPS

Best Poster Award
Computational and Genomic Biology Retreat, Center for Computational Biology, UC Berkeley, CA 2023

Lorraine McCormick Memorial Scholarship
Association for Women In Science (AWIS) East Bay, San Mateo, CA 2023

STEM*FYI & PPG Foundation Professional Development Grant
Graduate Division, UC Berkeley, CA 2022

Biostatistics Diversity Fellowship
Graduate Group in Biostatistics, UC Berkeley, CA 2021-2023

Biostatistics Block Grant and Non-Resident Student Tuition Award
Graduate Group in Biostatistics, UC Berkeley, CA 2018-2021

Berkeley Wellness Letter Fellowship
School of Public Health, UC Berkeley, CA 2020-2021

Public Health Alumni Association Award
School of Public Health, UC Berkeley, CA 2019

COMPUTING

Languages R, Bash, LaTeX, Python, Tensorflow, Pytorch, Matlab, Mathematica
Tools Git, GitHub, Emacs, Vim

LANGUAGES

Mandarin *Native*
English *Proficient*
German *Intermediate*

Last Updated on November 9, 2023