GUAN'AO YAN

(+1) 213-469-7747 ◇ yanguana@msu.edu ◇ https://www.guanaoy.com
428 S Shaw Ln East Lansing, MI 48824

RESEARCH AREA

My research interests lie in developing new statistical methods for understanding real-world data. Specific research topics include

- Statistical Bioinformatics
  - Statistical methods for analyzing high-dimensional single-cell and spatial omics data
  - Using synthetic data to enhance the statistical rigor in single-cell and spatial omics data analysis
- General Statistical Methodologies: High-dimensional model inference and variable selection
- Statistics in Education: Statistical methods for promoting education equity (reported by Forbes)

POSITIONS

Assistant Professor, Michigan State University 08/2025 -
Department of Computational Mathematics, Science & Engineering

EDUCATION

University of California, Los Angeles 09/2020 - 06/2025
Ph.D. in Statistics
Thesis: *Advancing Statistical Rigor in Single-Cell and Spatial Omics Analysis Through In Silico Control Data*
Advisor: Jingyi Jessica Li

Zhejiang University 09/2017 - 03/2020
M.Sc in Probability and Mathematical Statistics
Advisor: Yi Zhang

Shandong University 09/2013 - 06/2017
B.Sc in Mathematics and Applied Mathematics
B.Ec in Economics

GRANTS & AWARDS

National Science Foundation (NSF) Travel Award, STATGEN 2025 2025
Outstanding PhD Student Award, University of California, Los Angeles 2025
Dissertation Year Fellowship ($38,000), University of California, Los Angeles 2024
JXTX+CSHL Genome Informatics 2023 Scholarship ($2,000), Cold Spring Harbor Laboratory 2023
Don Ylvisaker Award for the Best Practice of Statistics, University of California, Los Angeles 2023
Interdisciplinary Opportunity Award ($10,000), NSF-Simons CMCF 2022
Most Promising Statistician Award, University of California, Los Angeles 2022
Summer Mentored Research Fellowship ($6,000), University of California, Los Angeles 2021
China National Scholarship 2018
Merit Graduate Student Award, Zhejiang University 2018
Outstanding Student Award, Shandong University 2014 - 2016
Hua Loo-Keng Talent Scholarship, Chinese Academy of Sciences 2015

PUBLICATIONS & MANUSCRIPTS

†*Indicating co-first author*

E. Huang, T. Fu, L. Zhang, **G. Yan**, R. Yamamoto, S. Terrazas, T.L. Nguyen, C. Gonzalez-Figueroa, A. Khanbabaei, J.H. Bahn, R. Varada, K. Amoah, J. Hervoso, M.T. Paulsen, B. Magnuson, M. Ljungman, J.J. Li, and X. Xiao (2025). Genetic variants affecting RNA stability influence complex traits and disease risk. *Nature Genetics*, 1-11.

Z. Li, Z. M. Patel, D. Song, S. N. Yasa, R. Cannoodt, **G. Yan**, J. J. Li and L. Pinello (2025). Systematic benchmarking of computational methods to identify spatially variable genes. *Genome Biology*, 26, 285.

G. Yan, S. Hua and J.J. Li (2025). Categorization of 34 computational methods to detect spatially variable genes from spatially resolved transcriptomics data. *Nature Communications*, 16, 1141.

G. Yan, J.J. Li and M. Biggin (2024). Question-Score Identity Detection (Q-SID): A statistical algorithm to detect collusion groups with error quantification from exam question scores. *arXiv*, 2407.07420. (Under review at *Journal of the American Statistical Association*) [Website] [Forbes article] [Podcast]

J. Zhao, F. Lao, **G. Yan** and Y. Zhang (2024). How data heterogeneity affects innovating knowledge and information in gene identification: A statistical learning perspective. *Journal of Innovation & Knowledge*, 9-3.

G. Yan, D. Song and J.J. Li (2023). scReadSim: a single-cell RNA-seq and ATAC-seq read simulator. *Nature Communications*, 14(1), 7428. [Software] [Website]

D. Song, Q. Wang, **G. Yan**, T. Liu and J.J. Li (2023). scDesign3 generates realistic in silico data for multimodal single-cell and spatial omics. *Nature Biotechnology*, 1-6. [Software]

S. Tang, H. Wang, **G. Yan**, L. Zhang (2022). Empirical likelihood based tests for detecting the presence of significant predictors in marginal quantile regression. *Metrika*, 1-31.

S. Chen†, **G. Yan**†, W. Zhang, J. Li, R. Jiang and Z. Lin (2021). RA3 is a reference-guided approach for epigenetic characterization of single cells. *Nature Communications*, 12(1), 1-13. [Software]

J. Zhao, **G. Yan** and Y. Zhang (2021). Robust estimation and shrinkage in ultrahigh dimensional expectile regression with heavy tails and variance heterogeneity. *Statistical Papers*, 1-28.

J. Zhao†, **G. Yan**† and Y. Zhang (2019). Semiparametric expectile regression for high-dimensional heavy-tailed and heterogeneous data. *arXiv*, 1908.06431. (In press at *Applied Mathematics-A Journal of Chinese Universities*)

PATENTS

M. Biggin, J.J. Li, **G. Yan**. Systems and methods for detecting collusion in student testing using graded scores or answers for individual questions (Serial No. 17/450,984; US Patent 11,915,615 B2)

SOFTWARES

Q-SID An online anti-collusion proctoring system [Website]

scReadSim Python package of synthetic read simulator designed for the single-cell multiomics data [Software]

RA3 R package of "RA3 is a reference-guided approach for epigenetic characterization of single cells" [Software]

PRESENTATIONS

Oral Presentations

· STATGEN 2025, Minneapolis, USA 5/2025
· Jonsson Comprehensive Cancer Center Gene Regulation Seminar, Los Angeles, USA 5/2025
· NHGRI Genome Tech Dev Working Group, Jackson Laboratory, USA 1/2025
· Department of Mathematical Sciences, New Jersey Institute of Technology, USA 12/2024
· Department of Statistics and Data Sciences, University of California, Los Angeles, USA 11/2024
· Joint Statistical Meetings, Portland, USA 08/2024
· Institute for Computational and Experimental Research in Mathematics, Providence, USA 12/2023
· Jonsson Comprehensive Cancer Center Gene Regulation Seminar, Los Angeles, USA 11/2023
· Institute for Quantitative and Computational Biosciences Research Seminar, Los Angeles, USA 12/2022
· NSF-Simons Center for Multiscale Cell Fate 5th Annual Symposium, Irvine, USA 10/2022
· The 7th International Conference on Statistics and Probability, IMS-China, Dalian, China 07/2019

Poster Presentations

· Cold Spring Harbor Laboratory Genome Informatics Conference, New York, USA 12/2023
· RECOMB/ISCB Conference on Regulatory & Systems Genomics, Los Angeles, USA 11/2023
· Chan Zuckerberg Initiative Single-Cell Biology 2023 Annual Meeting, Carlsbad, USA 11/2023
· ISMB/ECCB, Lyon, France 07/2023
· Los Angeles Bioscience Ecosystem Summit, Los Angeles, USA 05/2023
· Jonsson Comprehensive Cancer Center Retreat Poster Session, Los Angeles, USA 05/2023
· Institute for Quantitative and Computational Biosciences Poster Session, Los Angeles, USA 09/2022

TEACHING & MENTORING

Teaching Assistant

· STATS 205, Hierarchical Linear Models, UCLA Spring 2024
· STATS 203, Large Sample Theory, UCLA Winter 2024
· Statistical Science with Applications to Epidemiology, ElevatePro Summer 2021
· MATH 1001, Advanced Mathematics, Zhejiang University Fall 2019
· MATH 1001, Advanced Mathematics, Zhejiang University Fall 2018

Workshop Instructor

· Presenter, Jonsson Comprehensive Cancer Center Workshop, UCLA Dec 2024
"Categorization of 34 Computational Methods to Detect Spatially Variable Genes from Spatially Resolved Transcriptomics Data"
· Coordinator & Presenter, QCBio Workshop, UCLA May 2022
"Statistical Methods for Enhancing the Rigor in Single-cell RNA-seq Data Analysis"

Guest Lecturer

· STATS 205, Hierarchical Linear Models, UCLA Spring 2024
· BIOINFO 229, Current Topics in Bioinformatics, UCLA Winter 2024

Undergraduate Student Mentor

· Weijian Wang, Zhejiang University 12/2022 - Present
· Zhiyin Liu, Hong Kong University of Science and Technology 12/2022 - Present
· Shuo Hua, Tsinghua University 06/2022 - 12/2022

PROFESSIONAL EXPERIENCE

Graduate Student Reseracher 09/2021 – Present
Department of Statistics & Data Science
University of California, Los Angeles
Advisor: Jingyi Jessica Li

Research Scientist Intern 06/2024 – 09/2024
Data and Statistical Sciences
Product Development Department
Genentech, Inc.
"Enhancing Statistical Practice in Recurrent Event Endpoint Data Analysis"

Research Scientist Intern 06/2023 – 09/2023
Data and Statistical Sciences
Product Development Department
Genentech, Inc.
"Dynamic Monitoring of Ongoing Clinical Trials for Early Decision Making"

Research Assistant 06/2019 – 04/2020
Department of Statistics
The Chinese University of Hong Kong
Advisor: Zhixiang Lin

Undergraduate Researcher 06/2016 – 06/2017
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Advisor: Jianming Xia

PROFESSIONAL SERVICES

Conference and Session Organization

· Session Organizer, STATGEN 2025 Conference May 2025
"Spatially Variable Gene Detection: Advances and Applications in Spatial Transcriptomics data"

Statistical Consulting Services

· Jonsson Comprehensive Cancer Center 2022 - 2024

ASSISTANCE WITH GRANT PROPOSALS

Single-Cell Biology Data Insights Grant 07/01/2022 – 12/31/2023
Chan-Zuckerberg Initiative $200,000
"Enhancing Rigor and Reliability of Single-Cell Data Science" (PI: Jingyi Jessica Li)

R35 GM140888 06/01/2021 – 05/31/2026
NIH / NIGMS MIRA for established investigators $1,848,665
"Statistical Methods for Elucidating Regulatory Mechanisms and Functional Impacts of Transcriptome Variation at Population and Single-cell Scales" (PI: Jingyi Jessica Li)

REVIEWING ACTIVITIES

Reviewer for Scientific Journals
Nature Communications, Bioinformatics, PLOS Computational Biology, Bioinformatics Advances, BMC Bioinformatics, Biometrics

Reviewer for Scientific Conferences
ACM BCB 2025 Conference, 2025 NCME Annual Meeting

PROFESSIONAL AFFILIATIONS

Institute of Mathematical Statistics	2023 - Present
American Statistical Association	2022 - Present
American Society of Human Genetics	2022 - 2023
UCLA Jonsson Comprehensive Cancer Centers	2022 - Present
International Indian Statistical Association	2022 - 2023