ZHONGYI (JAMES) GUO

Email: zhongyig123@gmail.com LinkedIn: zhongyi-guo Phone: (+1) 607-262-3415

EDUCATION

MS, Epidemiology and Clinical Research, Stanford University (**GPA: 3.86/4.00**) 06/2025
Core Coursework: Translational Bioinformatics, Causal Inference, Design of Clinical Trials, Meta-analysis, Biostatistics, Data Analysis in SAS, Deep Learning in Genomics, Data Visualization

BS, Biometry & Statistics *and* **Biological Sciences,** Cornell University (**GPA: 3.57/4.30**) 05/2023
Honors: *cum laude*, Dean's List
Core Coursework: Machine Learning, Data Science, Quantitative Genomics, Statistical Computing, Theory of Statistics, Probability Inference, Linear Algebra, OOP & Data Structure, Biostatistics, pre-med

Biochemistry major, Case Western Reserve University (**GPA: 4.0/4.0**; transferred to Cornell) 05/2020
Core Coursework: Calculus I & II & III, Micro & Macroeconomics, Intro. Biology, Intro. Chemistry

SELECTED WORK

Research Associate, the Corces Lab, Gladstone Institutes 07/2025 – Present

- Pioneered **an R package** implementing a **permutation-test** pipeline to reduce false positives in scRNA-seq differential gene expression analysis of Alzheimer's disease to **identify therapeutic targets,** integrating eight public Synapse datasets to demonstrate real-world use.
- Applied **AlphaGenome** to study how genetic variants of the *MAPT* gene **alter splicing patterns**.
- **Removed doublets,** performed dimensionality reduction & clustering of scATAC-seq data in **ArchR**.
- Communicated statistical findings and effective visualizations clearly to **diverse audiences**.

SELECTED RESEARCH

Graduate Student Researcher, the Kundaje Lab, Stanford University 04/2024 – 06/2025

- Preprocessed and integrated high-throughput **scRNA-seq and scATAC-seq** data of psychiatric disorders, applying summary statistics, quality control, feature engineering, and dimension reduction.
- Performed differential gene expression and peak accessibility analyses at **pseudobulk** (using Wald test and LRT in *DESeq2*) and **single-cell levels** (using Wilcoxon rank-sum test and MAST in **Seurat**).
- Trained and interpreted cell type-specific ChromBPNET, a CNN-based DL model, to assess effects of potential genetic disease variants on transcription factor binding and chromatin accessibility.
- Delivered clear updates to professors and colleagues regularly, **simplifying complex statistics**.

Graduate Student Researcher, the Graff Lab, UCSF (remote) 10/2023 – 06/2025

- Applied **t-test, PCA, PLS-DA, random forest, logistic & linear regression, chemical similarity enrichment analysis, pathway analysis, and WGCNA** to LC-MS metabolomics data, characterizing metabolites/biomarkers associated with Black-White **prostate cancer disparities.**
- Interpreted statistical findings and **communicated them through clear scientific writing** and **effective data visualizations**; defended before Stanford Epidemiology faculty and students.

TECHNICAL SUMMARY

Languages: R, Python, SAS (Base certified), SQL, LaTeX, HTML, CSS, Javascript (D3.js)
Tools: Git, GitHub, Conda, RStudio, Jupyter Notebook, High Performance Computing, MS Office
Core Skills: statistical computing, quantitative modeling, data-driven & cross-functional collaboration
Aesthetics: Adobe (Illustrator, Photoshop, After Effect), Procreate

Last Updated on 02/172026

LEADERSHIP

Epidemiology M.S. Student Ambassador, Stanford University 05/2024 – 06/2025

- Responded to inquiries from applicants, explaining Master's curriculum and research resources.

Social Media Content Editor, Chinese Rainbow Network (CRN) 03/2024 – 01/2025

- Edited newsletter and visual content on mental health resources for the Chinese LGBTQ+ community.

Scientific Review Editor (Education Branch), Community HealthEd 03/2022 – 05/2023

- Validated accuracy of cited information and converted scientific jargon to plain language to ensure accessibility in weekly newsletters, focusing on neurological, psychiatric, and prenatal health.

Orientation Leader, Cornell University 08/2021 and 01/2022

- Designed welcoming icebreakers and guided mini campus tours to welcome new students.

TEACHING

Cornell University

- **Beta Tester** and **Teaching Assistant,** Introduction to Data Science 01/2023 – 05/2023
- **Grader,** Probability Models & Inference 09/2022 – 12/2022
- **Teaching Assistant,** Laboratory in Genetics and Genomics 01/2021 – 05/2021

JNC Study Abroad Platform

- **Teaching Assistant,** Introductory Biology *and* Physics I 07/2022 – 08/2022

PUBLICATION

[1] (Submitted to Nature Neuroscience, 2025) Shin, J.‡, Brady, E.‡, Chen, C., Lauderdale, K., Agrawal, A., Zhang, Y., Jiang, X., Nambiar, P., Herbert, J., Mallen, D., Ly, K., **Guo, Z.**, Sant, C., Thomas, R., Miller, S., Cobos, I., Palop, J.. APOE4 and Aβ synergize to drive neuronal network dysfunction and lysosomal-ER proteostasis dysregulation in the preclinical stages of Alzheimer's disease.

Manuscripts in preparation:

[1] (In Progress; Aiming for BMC Medicine) **Guo, Z.‡**, Chen, D.‡, Stopsack, K. H., Soule, P., Ajit, D., Ramamoorthy, P., Hoffmann, T. J., Chan, J. M., Mucci L. A., Graff, R. E.. Metabolomic Disparities Between Black and White Men with Metastatic Hormone-Sensitive Prostate Cancer: A Pilot Study.

[2] (In Progress; Aiming for Cell) Qu, P., Wang, T., Jessa, S., **Guo, Z.**, Guo, H., Purmann, C., Monte, E., Jiang, L., Yang, X., Zhou, B., Kundu, S., Kundaje, A., Wong, W., Hallmayer, J. F., Urban, A. E., Snyder, M. P.. Multi-modal functional genomics analysis of bipolar disorder and schizophrenia. (Title is tentative.)

[3] (In Progress; Aiming for Nature Neuroscience) Sant, C., **Guo, Z.**, Corces, M. R.. Preventing false discoveries in Alzheimer's disease single-cell sequencing data using permutation testing. (Title is tentative.)

‡ indicates co-first authorship.