XIAORU DONG

228 Buckman Drive, Gainesville, FL 32603
(217)-778-5683 · xdong1@ufl.edu
https://xiaorudong.github.io

EDUCATION

Ph.D., Biostatistics May 2024
University of Florida
Advisor: Rhonda Bacher, Ph.D.
Dissertation: Improving Analysis and Modeling of Dynamic, High-Dimensional Single-Cell Data

B.S., Statistics December 2018
University of Illinois Urbana-Champaign

RESEARCH INTERESTS

My research interests center on the intersection of biology and statistics, focusing on developing statistical and computational tools for genomics and biomedical research. I specialize in analyzing complex datasets, especially human single-cell data from next-generation sequencing technologies using advanced statistical techniques. My aim is to drive novel biological discoveries and enhance our understanding of genomic processes through innovative and data-driven approaches.

EXPERIENCE

Postdoctoral Associate June 2024 - Present
University of Florida

- Lead the development of advanced statistical methods using machine learning techniques to analyze complex human biomedical datasets, such as flow cytometry data, to gain insights into disease-related changes in the immune system.
- Create and validate a novel statistical framework and a pipline to identify condition-specific effects and sgRNA interactions in CRISPR screening data, enhancing the precision and interpretability of gene function and interaction analyses across diverse experimental conditions.
- Design and implement an innovative statistical approach for integrating and interpreting multi-omics data, including functional genomics, single-cell epigenomics, and single-cell transcriptomics data, to enhance the understanding of complex biological processes and interactions.
- Conduct independent, in-depth single-cell RNA-seq analyses to investigate genetic disorders of the immune system, driving the development of targeted therapeutic strategies by uncovering disease-specific molecular pathways.
- Mentor graduate students in statistical methods and data analysis techniques, fostering their skills and supporting their progress toward achieving research objectives.

Graduate Research Assistant August 2019 - May 2024
University of Florida

- Collaborated with clinical and translational investigators (e.g., Dr. Todd Brusko, Dr. Philip Efron, and Dr. Laura Jacobsen) on biomedical studies involving patients with type 1 diabetes and sepsis, providing statistical expertise in the analysis of high-dimensional datasets with direct relevance to clinical applications.

- Performed advanced analyses of single-cell (e.g., 10X Genomics) and bulk RNA-seq data using established software and custom-developed pipelines in both local and high-performance computing environments, identifying critical disease-associated factors and biomarkers.
- Applied robust statistical methods to analyze high-dimensional omics datasets, including genomic, epigenomic, cytokine, and immunophenotyping data, uncovering novel insights into underlying disease mechanisms.
- Developed computational methods for scRNA-seq data analysis in R, including novel algorithms for data-driven trajectory evaluation, enhancing the accuracy of trajectory inference and quantifying trajectory properties affected by analytical decisions.
- Contributed to statistical sections for grant applications and manuscripts by preparing statistical analysis plans, generating publication-quality figures, and interpreting results from complex statistical models.
- Provided training and consultation in experimental design and statistical software (R, SAS) to multidisciplinary teams, promoting rigorous analytical practices and reproducible workflows in collaborative biomedical research.

Graduate Teaching Assistant August 2019 - May 2023
University of Florida

- Mentored undergraduate and graduate students across various departments in five classes including Public Health Computing, Biostatistical Methods, and Statistical Analysis of Genetic Data, enhancing engagement and performance in both online and in-person classes (20+ students per class).
- Offered weekly office hours for coding guidance in R and SAS, clarifying complex concepts and contributing to the students' practical understanding of statistical analysis.
- Led interactive classroom discussions and facilitated group activities to encourage critical thinking and foster a collaborative learning environment.
- Created and implemented innovative teaching materials, including assignments and digital learning platforms (Canvas), leading to improved student learning outcomes and course effectiveness.
- Provided detailed grading and constructive feedback, directly contributing to measurable enhancements in student academic achievement.

PUBLICATIONS

*Key: † indicates co-first authors; * indicates corresponding authors.*

Peer-Reviewed Journal Articles

1. **Dong, X.**, Leary, J. R., Yang, C., Brusko, M. A., Brusko, T. M., & Bacher, R.* (2024). Data-driven selection of analysis decisions in single-cell RNA-seq trajectory inference. *Briefings in Bioinformatics, 25*(3), bbae216. https://doi.org/10.1093/bib/bbae216
2. Shapiro, M. R.†, **Dong, X.**†, Perry, D. J.†, McNichols, J. M., Thirawatananond, P., Posgai, A. L., Peters, L. D., Motwani, K., Musca, R. S., Muir, A., Concannon, P., Jacobsen, L. M., Mathews, C. E., Wasserfall, C. H., Haller, M. J., Schatz, D. A., Atkinson, M. A., Brusko, M. A., Bacher, R.*, & Brusko, T. M.* (2023). Human immune phenotyping reveals accelerated aging in type 1 diabetes. *JCI Insight.* https://doi.org/10.1172/jci.insight.170767
3. Jacobsen, L. M.†, Diggins, K.†, Blanchfield, L.†, McNichols, J. A., Perry, D. J., Brant, J., **Dong, X.**, Bacher, R., Gersuk, V. H., Schatz, D. A., Atkinson, M. A., Mathews, C. E., Haller, M. J., Long, S. A.*, Linsley, P. S.*, & Brusko, T. M.* (2023). Responders to low-dose ATG induce CD4 T cell exhaustion in type 1 diabetes. *JCI Insight.* https://doi.org/10.1172/jci.insight.161812

4. Darden, D. B., **Dong, X.**, Brusko, M. A., Kelly, L., Fenner, B., Rincon, J. C., Dirain, M. L., Ungaro, R., Nacionales, D. C., Gauthier, M., Kladde, M., Brusko, T. M., Bihorac, A., Moore, F. A., Loftus, T., Bacher, R., Moldawer, L. L., Mohr, A. M., & Efron, P. A.* (2021). A Novel Single Cell RNA-seq Analysis of Non-Myeloid Circulating Cells in Late Sepsis. *Frontiers in Immunology, 12*, 696536. https://doi.org/10.3389/fimmu.2021.696536

5. Ross, J. J., Wasserfall, C. H., Bacher, R., Perry, D. J., McGrail, K., Posgai, A. L., **Dong, X.**, Muir, A., Li, X., Campbell-Thompson, M., Brusko, T. M., Schatz, D. A., Haller, M. J., & Atkinson, M. A.* (2021). Exocrine Pancreatic Enzymes Are a Serological Biomarker for Type 1 Diabetes Staging and Pancreas Size. *Diabetes, 70*(4), 944–954. https://doi.org/10.2337/db20-0995

Book Chapters

6. **Dong, X.**, & Bacher, R. (2023). Analysis of Single-Cell RNA-seq Data. In B. Fridley & X. Wang (Eds.), *Statistical Genomics* (Vol. 2629, pp. 95–114). Springer US. https://doi.org/10.1007/978-1-0716-2986-4_6

Other Publications

7. **Dong, X.**, & Bacher, R. (2022). Data-driven assessment of dimension reduction quality for single-cell omics data. *Patterns, 3*(3), 100465. https://doi.org/10.1016/j.patter.2022.100465 [Invited Preview]

Pre-prints

8. **Dong, X.**, Goyal, A., Liang, M., Brusko, M. A., Brusko, T. M., & Bacher, R.* (2025). Penalized Linear Models for Highly Correlated High-Dimensional Immunophenotyping Data. *arXiv*. https://arxiv.org/abs/2504.07771

9. Leary, J. R., **Dong, X.**, & Bacher, R.* (2025). Interpretable Trajectory Inference with Single Cell Linear Adaptive Negative-binomial Expression (scLANE) Testing. *bioRxiv*. https://doi.org/10.1101/2023.12.19.572477

SOFTWARE

GitHub Software Packages

1. Escort: Methodology and an R package implementing Escort. This method evaluates a single-cell RNA-seq dataset's suitability for trajectory inference and for quantifying trajectory properties influenced by analysis decisions. Available at: GitHub.

2. berm: Bootstrap-Enhanced Regularization Method. This package implements the Bootstrap-Enhanced Regularization Method (BERM), a statistical approach aiming to enhance the robustness and accuracy of variable selection and coefficient estimation in immunophenotyping datasets. Available at: GitHub.

Interactive Tools and Applications

3. shinyEscort: An interactive Shiny application for guiding trajectory construction in single-cell RNA-seq data. Available at: shinyEscort website.

PRESENTATIONS

National Talks

1. "Realistic Simulation for Benchmarking and Designing CRISPR-Pooled Screen." American Statistical Association (ASA) Section on Statistics in Genomics and Genetics (SSGG). Contributed talk. Minneapolis, "MN", USA. May 2025.

2. "Data-driven Selection of Trajectory in Single-cell RNA-seq Data." International Biometric Society Eastern North American Region (ENAR) Annual Meeting. Contributed talk. Baltimore, MD, USA. March 2024.

Local Talks

3. "Penalized Linear Models for Highly Correlated Immunophenotyping Data." 2025 Annual Meeting of the Florida Chapter of the American Statistical Association (ASA). Contributed presentation. University of Central Florida, Orlando, FL, USA. February 2025.
4. "Enhancing Feature Selection and Estimation in Immunophenotype Data Analysis." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. March 2024.
5. "Data-driven Trajectory Construction in Single-Cell RNA-seq Data." Mark C. K. Yang Award Ceremony. Contributed talk. University of Florida, Gainesville, FL, USA. October 2023.
6. "Trajectory Goodness of Fit in Single-Cell RNA-seq Data." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. December 2021.

National Posters

7. "Data-driven Selection of Trajectory in Single-cell RNA-seq Data." American Statistical Association (ASA) Section on Statistics in Genomics and Genetics (SSGG). Contributed poster. Pittsburgh, PA, USA. May 2024.
8. "Data-driven Evaluation of Trajectories in Single-Cell RNA-seq Data." International Biometric Society Eastern North American Region (ENAR) Annual Meeting. Contributed poster. Nashville, TN, USA. March 2023.

Local Posters

9. "Data-driven Evaluation of Trajectories in Single-Cell RNA-Seq Data." 2023 Annual Meeting of the Florida Chapter of the American Statistical Association (ASA). Contributed poster. University of Florida, Gainesville, FL, USA. March 2023.
10. "Immune Phenotypic Profiling across the Human Lifespan Demonstrates Accelerated Immune Aging in Subjects with Type 1 Diabetes." 2022 College of Medicine Research Day. Contributed poster. University of Florida, Gainesville, FL, USA. April 2022.
11. "Analyzing Inclusion Criteria of 7000 Cochrane Systematic Reviews." Undergraduate Research Symposium. Contributed poster. University of Illinois at Urbana-Champaign, Champaign, IL, USA. April 2018.
12. "Distortions in Scientific Literature - A Replication Analysis of Greenberg's Citation Network of 302 Alzheimer's Science Research Papers." Undergraduate Research Symposium. Contributed poster. University of Illinois at Urbana-Champaign, Champaign, IL, USA. April 2018.

TEACHING

Teaching Assistant
University of Florida

- PHC4094 - Introduction to Biostatistics for Health Science and Public Health. Fall 2021.
- PHC6052 - Introduction to Biostatistical Methods. Fall 2019 & Fall 2020.
- PHC6059 - Introduction to Applied Survival Analysis. Fall 2022.
- PHC6088 - Statistical Analysis of Genetic Data. Spring 2021, Spring 2022 & Spring 2023.
- PHC6089 - Public Health Computing. Spring 2020.

ACADEMIC SERVICE

Journal Reviewer

- Bioinformatics

- Bioinformatics Advances
- BMC Bioinformatics
- Communications Biology
- Scientific Reports

Service in Statistics

- Poster judge. 2024 Florida Genetics Symposium. November 2024.
- Session moderator. Annual Meeting of the Florida Chapter of the American Statistical Association (ASA). March 2023.
- Session moderator. Annual Applied Statistics Symposium of the International Chinese Statistical Association (ICSA). June 2022.

HONORS & AWARDS

- Mark C. K. Yang Student Presentation Award. October 2023
- Student Competition Award, American Statistical Association Florida Chapter. March 2023
- Certificate in Data Science. November 2018
- Successful Participant Award, Mathematical Contest in Modeling. January 2016

PROFESSIONAL SOCIETY MEMBERSHIPS

- International Biometric Society (ENAR)
- American Statistical Association (ASA)