XIAORU DONG

228 Buckman Drive, Gainesville, FL 32603

(217)-778-5683 · xdong1@ufl.edu

www.linkedin.com/in/xiaoru-dong

EDUCATION

Ph.D., Biostatistics May 2024

University of Florida

Advisor: Rhonda Bacher, Ph.D.

Dissertation: Improving Analysis and Modelling of Dynamic, High-Dimensional Single-Cell Data

B.S., Statistics December 2018

University of Illinois at 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

- Initiate and 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.

- Independently provide statistical support and conduct single-cell RNA-seq analysis for projects on genetic disorders of the immune system, elucidating underlying pathology and contributing to the development of targeted therapeutic strategies.

- Mentor graduate students by providing guidance on statistical methods and data analysis, helping them develop their skills and achieve their research goals.

Graduate Research Assistant August 2019 - May 2024

University of Florida

- Collaborated with Dr. Todd Brusko and Dr. Philip Efron and their teams at the University of Florida College of Medicine on data analysis projects in immunology, focusing on type 1 diabetes and sepsis.

- Conducted advanced analysis using existing software packages and developed custom scripts for single-cell RNA-seq (e.g., 10X) and bulk RNA-seq data, identifying critical disease-associated factors and biomarkers.

- Applied machine learning methods, such as random forest and lasso, to analyze high-dimensional omics datasets, including genomic, epigenomic, cytokine, and immunophenotyping data.

- Contributed to study design, data analysis, preparation of statistical sections in grant applications, and publication of findings in scientific journals.

- Developed novel computational methods for scRNA-seq data analysis using R, including creating innovative algorithms for data-driven scRNA-seq trajectory evaluation aimed at identifying the most accurate trajectory.

- Provided mentorship and support in R programming to lab members, aiding in data analysis tasks and troubleshooting.

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).

- Provided 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.

- Delivered detailed grading and constructive feedback, contributing to measurable improvements in student performance and academic success.

PUBLICATIONS

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

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. **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

5. **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

6. 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

7. 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

SOFTWARE

GitHub Software Package

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.

PRESENTATIONS

National Talks

1. "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

2. "Enhancing Feature Selection and Estimation in Immunophenotype Data Analysis." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. March 2024.

3. "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.

4. "Linear Regression with Highly Correlated Variables." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. January 2023.

5. "Exploring the CausalCell Approach to Single-Cell Data." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. September 2022.

6. "Human Immune Phenotyping (HIP): Age Prediction Model." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. April 2022.

7. "Trajectory Goodness of Fit in Single-Cell RNA-seq Data." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. December 2021.

8. "Understanding RNA Velocity." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. August 2021.

9. "Gene Selection for Trajectory Inference in Single-Cell RNA-seq Data." Brusko Lab, University of Florida Diabetes Institute, Gainesville, FL, USA. January 2021.

National Posters

10. "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.

11. "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

12. "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.

13. "Analyzing Inclusion Criteria of 7000 Cochrane Systematic Reviews." Undergraduate Research Symposium. Contributed poster. University of Illinois at Urbana-Champaign, Champaign, IL, USA. April 2018.

14. "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

- BMC Bioinformatics

- Scientific Reports

- Bioinformatics

Service in Statistics

- 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)