YIREN SHAO

http://www.yirenshao.com
(984) 261-5682 yiren.shao@outlook.com

EDUCATION

DUKE UNIVERSITY, School of Medicine, Durham, NC
Master of Biostatistics, May 2023, GPA: 3.85/4.0 **Program Scholarship**
Major coursework: Data Structure and Algorithm, Statistical Inference, High Dimensional Statistics, Computational Sequence Biology, Statistical Computing, Applied Biostatistical Method I&II, Generalized Linear Model

TONGJI UNIVERSITY, School of Mathematical Science, Shanghai, China
Bachelor of Science, **Applied Mathematics**, July 2021, GPA: 87.51/100. **Scholarship for Excellent Academic Performance**
Major coursework: Differential Equations, Numerical Analysis, Mathematical Analysis, Complex Analysis, Computing Method, Functional Analysis, Advanced Algebra, Topology, Deep Learning, Database Management

PUBLICATIONS

- ***Shao Y***, Gao Q, Wang L, Nixon A, Chan C, Li Q, Xie J; **B-Lightning: using bait genes for marker gene hunting in single-cell data with complex heterogeneity.** Briefings in Bioinformatics, 2025 Feb 10. doi: 10.1093/bib/bbaf033
- Cieri N, Hookeri N, Stromhaug K, Li L, Keating J, Díaz-Fernández P, Gómez-García de Soria V, Stevens J, Kfuri-Rubens R, ***Shao Y***, Kooshesh KA, Powell K, Ji H, Hernandez GM, Abelin J, Klaeger S, Forman C, Clauser KR, Sarkizova S, Braun DA, Penter L, Kim HT, Lane WJ, Oliveira G, Kean LS, Li S, Livak KJ, Carr SA, Keskin DB, Muñoz-Calleja C, Ho VT, Ritz J, Soiffer RJ, Neuberg D, Stewart C, Getz G, Wu CJ. **Systematic identification of minor histocompatibility antigens predicts outcomes of allogeneic hematopoietic cell transplantation.** Nat Biotechnol. 2024 Aug 21. doi: 10.1038/s41587-024-02348-3. PMID: 39169264.

ABSTRACTS

- Cieri N, ***Shao Y***, Powell K, Smith C, Ho VT, Kerrigan M, Stevens J, Lane WJ, Kim HT, Neuberg D, Stewart C, Getz G, Ritz J, Soiffer RJ, Wu CJ. **Minor histocompatibility antigens cross-reactive against gut-tropic viral epitopes facilitate acute GvHD of the gut.** Transplantation and Cellular Therapy, Official Publication of the American Society for Transplantation and Cellular Therapy, 2025 Feb, doi: 10.1016/j.jtct.2025.01.057.
- Cieri N, ***Shao Y***, Powell K, Ho VT, Stevens J, Lane WJ, Kim HT, Neuberg D, Stewart C, Getz G, Ritz J, Soiffer RJ, Wu CJ. **Acute GvHD of the gut is associated with minor histocompatibility antigens cross-reactive against gut-tropic viral epitopes**. Blood, 2024 Nov 5, doi: 10.1182/blood-2024-198976.

PRESENTATIONS

- "**Interpretable dimension reduction for pseudo-time trajectories**," The 18th Annual CFAR Fall Scientific Retreat, Duke Center for AIDS Research, Duke University, Durham, NC, September 22, 2022
- "**An iterative gene fishing method to detect senescence gene markers**," CHSI Meta-lab Seminar, Duke Center for Human Systems Immunology, Duke University, Durham, NC, September 27, 2022

RECENT ACADEMIC EXPERIENCE

DANA-FARBER CANCER INSTITUTE, Boston, MA 2023-2025
***Computational Biologist*, Supervisors: Professor Catherine Wu, Dr. Nicoletta Cieri**
Project: Systematic identification of minor histocompatibility antigens predicts outcomes of allogeneic hematopoietic cell transplantation

- Refined an analytic framework to predict graft-versus-host-disease (GvHD) outcomes and identify potential immunotherapy targets by predicting minor histocompatibility antigens (mHAgs) from single-nucleotide polymorphisms (SNPs) in whole-exome sequencing (WES) data of allogeneic hematopoietic cell transplant (allo-HCT) donor-receipient (D-R) pairs. Extended the evaluation of mHAgs to include HLA class II prediction.
- Processed, integrated and analyzed single-cell RNA sequencing data to build a gene expression atlas of inflammatory GI tissues in allo-HCT patients with GvHD
- Genotyped and analyzed paired single-cell RNA-seq and TCR data of allo-HCT D-R pairs' peripheral blood mononuclear cells three days post transplantation to identify donor-derived proliferating T cells and validate predicted mHAgs
- **Our work established a foundation for personalized post allo-HSCT mHAg repertoire prediction and demonstrated that total and organ-specific mHAg load could independently predict the occurrence of acute GvHD and chronic pulmonary GvHD. This work was published in Nature Biotechnology (doi: 10.1038/s41587-024-02348-3)**

Project: Quantification of viral cross-reactivity across patient-specific minor histocompatibility antigen landscape

- Generated *in silico* human virus epitope library and predicted their HLA class I binding affinity by neural network based algorithms
- Designed and implemented a string metric to quantify homology between viral epitopes and predicted HLA class I mHAgs inferred from allo-HCT D-R WES data to identify cross-reactive mHAgs
- **Revealed that cross-reactivity against CMV contributed to the pathophysiology of GI acute GvHD. This work was published in Blood (doi: 10.1182/blood-2024-198976) and Transplantation and Cellular Therapy (doi: 10.1016/j.jtct.2025.01.057)**

Project: Identification of major HLA mismatch derived epitopes for prognostication in partially-matched transplants

- Established a BLAST and neural network based framework to predict alloantigens derived from major HLA mismatch
- Ran pipeline on 1317 single HLA class I-mismatched unrelated D-R pairs to understand the association between HLA mismatch-derived epitopes and GvHD
- **Revealed that Grade II-IV acute GvHD was more frequent in patients with high predicted HLA mismatch-derived epitope loads.**

Project: Deciphering the molecular basis of cutaneous chronic GvHD following Blaschko Lines

- Processed and analyzed WES data of affected and unaffected dermis and epidermis (AD, AE, UD, UE) biopsies from a post-alloHCT patient with skin chronic GvHD to identify somatic mutations and predict neoantigens
- Quantified mitochondrial DNA heteroplasmy in AD, AE and UE bulk RNA-seq data based on mitochondrial reads to access the extent of immune infiltration across samples
- Processed, analyzed bulk RNA-seq data of AD, AE and UE to investigate the presence of concomitant virome and identify transposable elements (TEs)
- Integrated our bulk RNA-seq data of AD, AE and UE with external bulk and single-cell RNA-seq datasets of autoimmune-disease skin to build a gene expression atlas for skin GvHD

DUKE UNIVERSITY, School of Medicine, Durham, NC 2022-2024

***Research Assistant,* Supervisor: Professor Jichun Xie**

Project: B-Lightning: Using bait genes for marker gene hunting in single-cell data with complex heterogeneity

- Developed an iterative statistical gene-network based algorithm called B-Lightning as a gene marker fishing method using R and C++. **This work was published in Briefings in Bioinformatics (doi: 10.1093/bib/bbaf033)**
- Implemented the algorithm on the simulated and experimental single-cell RNA-seq data, including lung ciliated cells dataset, T cells dataset, breast cancer dataset and Alzheimer's Disease dataset
- Benchmarked B-Lightning against traditional differential expressed genes detecting methods and demonstrated that B-Lightning outperforms existing methods with significant improvement in sensitivity, specificity and robustness

SKILLS AND TECHNIQUES

Technical Skills:	(Proficient) R, PYTHON, WDL, DOCKER, HPC, GIT, JAVA, MATLAB, SPSS, C++, LaTeX (Familiar) C#, JULIA, SQL, Monolix2021
Bioinformatics:	NGS data analysis (including transcriptomic, whole-exome sequencing, TCR sequencing) germline and somatic mutation analysis transposable elements analysis

ADDITIONAL INFORMATION

Language:	Mandarin (Native Speaker) English (Fluent) TOEFL Score: 110 (Reading – 30; Listening – 30 Speaking – 25; Writing – 25)
Interests:	Piano, Golf, Cooking
Professional Associations:	American Statistical Association