

TIFFANY TU

PhD Candidate in Computational Biology and Bioinformatics

@ tiffany.tu@duke.edu [in linkedin.com/in/tiffany-tu/](https://www.linkedin.com/in/tiffany-tu/) [📍 Durham, NC 27713](#)

EDUCATION

- 2025 **PhD - Computational Biology and Bioinformatics**, Duke University School of Medicine. **Advisor** : Alejandro Ochoa
Dissertation : *Bias corrections for genetic association studies with applications in pediatric Nephrotic Syndrome*
- 2020 **MS - Biostatistics - Statistical Genetics Track**, Columbia University Mailman School of Public Health
- 2017 **BS - Computational Mathematics**, George Washington University

PROFESSIONAL EXPERIENCE

- Current**
Aug 2021 **Graduate Research Assistant/PhD Student in Ochoa Lab, DUKE UNIVERSITY, Durham, NC**
 - Lead biweekly meetings with collaborative partners from Duke Molecular Physiology Institute and National Institutes of Health (NIH) in discussing biological insights from genetic association (GWAS) pipelines involving phenotypic data curation, sample/variant QC, statistical analysis design, and population structure/imputation methods evaluation
 - Provide single-cell analysis results and interpretation in characterizing Nephrotic Syndrome immune profiles longitudinally to investigate HLA risk haplotype and other risk variants.
 - Optimize algorithm runtime and memory for novel R package LIGERA, which performs fast genetic association tests while fully modelling relatedness via population kinship matrices using a reverse regression model




Rcpp R:provis Plink Admixture SAIGE LocusZoom Web Scraping Bash/Shell Git/BitBucket
- Aug 2023**
Jun 2023 **Early Discovery Genetics Intern, MERCK, Cambridge, MA**
 - Built an end-to-end pipeline for applying Mendelian Randomization to disease phenotype and protein level (pQTL) data to identify unconfounded causal contributions for potential new targets and biomarkers.
 - Developed proficient skillset in genetic data harmonization, post-GWAS algorithm and pipeline implementation, and demonstrated effective communication within the Data, AI, and Genome Sciences (DAGS) department.

TwoSampleMR LDscore LDClump R Unix/Linux AWS
- Aug 2020**
Jan 2020 **Real World Data Science Intern, ROCHE, Little Falls, NJ**
pREDi Data Science (Pharma Research and Early Development Informatics)
 - Implemented a new preference score function in a Roche R-package that enables selection of an appropriate comparator arm in observational studies, as a step in comparing treatment combination pairs to predict patient response
 - Developed an end-to-end analytical pipeline for overall survival of specific mutation combinations by integrating clinicogenomic data to understand molecular mechanism behind drug resistance for non-small cell lung cancer
 - Contributed to RNA-seq differential gene expression analysis across and within cohorts with specific checkpoint inhibitors of interest to identify potential cancer immunotherapy resistant genes

R R:DESeq Bioconductor MultiAssayExperiment Package Unix/Linux Git/Bitbucket roxygen2
- Dec 2019**
Jun 2019 **Bioinformatics Co-op Student, REGENERON PHARMACEUTICALS, Tarrytown, NY**
Bioinformatics Core Services
 - Developed a web application for searching, analyzing, and displaying of LacZ-reporter gene expression image data to improve efficiency of pre-therapeutic target discovery research using RShiny
 - Refined an existing statistical consulting application that recommends tests based on user input and allows for user data upload and analysis by adding in 13 additional statistical tests with detailed descriptions and use cases
 - Maintained experimental image data intake, quality control, and analysis workflow with AWS cloud computing
 - Strengthened skills in data integrity, infrastructure design, and appropriate use of revision control

R R-Shiny Python AWS Unix/Linux HTML CSS Git/Bitbucket
- Aug 2018**
Oct 2017 **Research Associate Level II, ICAHN SCHOOL OF MEDICINE AT MOUNT SINAI, New York, NY**
Precision Immunology Institute
 - Performed high dimensional data reduction analyses using machine learning algorithms, such as Phenograph and t-SNE, for predictive modeling and visualization of immune cell clusters with single cell RNA-Seq data
 - Established analytical pipelines and experimental database build up on viral infections, lung cancer, bladder cancer, and breast cancer for studying how gene diversity shapes human immune responses
 - Performed wet bench experiments such as PBMC blood processing and immunohistochemistry.
 - Undertook administrative tasks by maintaining lab supplies, equipment operation, and lab protocols.

R t-SNE Phenograph Python

Aug 2017	Graduate Research Assistant, UNIVERSITY OF CALIFORNIA, LOS ANGELES, Los Angeles, CA
Jun 2017	<ul style="list-style-type: none"> ➤ Recipient of NIH funded UCLA Bruins in Genomics summer research program and attended two-week structured boot camp covering topics on NGS Analysis, Hi-C, Galaxy, and RNA-Seq held by the QCB Collaboratory. ➤ Strengthened skills in sequencing alignment, data cleaning, and data scraping using RNA-Seq data from public sources such as ENCODE (Encyclopedia of DNA Elements). ➤ Worked on developing an algorithm to infer unknown genetic functions between two species with tight spectral clustering and bipartite node covariates; particularly looking at human and mouse homologs.
	
Aug 2016	Mathematical Ecology Research Student, UNIVERSITY OF WISCONSIN, LA CROSSE, La Crosse, WI
Jun 2016	<ul style="list-style-type: none"> ➤ Generated bird and bat population dynamic model subject to wind energy development with branching processes, population inertia, and sensitivity and elasticity simulation. ➤ Collaborated with USGS to generate significant results for future Upper Midwest environmental management decisions on the endangered Indiana Bat species.
	 

PUBLICATIONS

- Cason RK, Chambers E, **Tu T**, Chryst-Stangl M, Huggins K, Lane BM, Ochoa A, Jackson AM, Gbadegesin RA. Genetic risk variants for childhood nephrotic syndrome and corticosteroid response. *Front Pediatr*. 2023 Oct 6;11:1248733.
- Haider H, Oldfield S, **Tu T**, Moreno R, Diendorfer J, Eager E, Erickson R. Incorporating Allee effects into the potential biological removal level. *Natural Resources Modeling*. 2017; 30:e12133

TALKS AND POSTERS

- [Poster] **Tu T**, Ochoa A. Genetic association meta-analysis is susceptible to confounding by between-study cryptic relatedness, **American Society of Human Genetics Annual Meeting (ASHG)**, Denver, CO, November, 2024
- [Poster] **Tu T**, Ochoa A, Sood A, Chan C, Chamber E, Jackson AM, Adeyemo A, Gbadegesin R. Genome-wide association studies of pediatric nephrotic syndrome identify variants associated with corticosteroid response, **American Society of Nephrology (ASN) Kidney Week**, San Diego, CA, October, 2024
- [Talk + Poster] **Tu T**, Ochoa A. The presence of cryptic relatedness creates inflation in meta-analyses of genome-wide association studies, **RECOMB - Genetics Satellite**, Cambridge, MA, April, 2024
- [Poster] **Tu T**, Ochoa A. The presence of cryptic relatedness creates inflation in meta-analyses of genome-wide association studies, **Annual International Conference on Research in Computational Molecular Biology (RECOMB)**, Cambridge, MA, April, 2024
- [Poster] **Tu T**, Gbadegesin RA, Adeyemo A, Ochoa A. A test of genotyping platform bias for multiethnic case/control association studies merging external controls, **American Society of Human Genetics Annual Meeting (ASHG)**, Washington, D.C., November, 2023
- [Talk] Chung HA, Fralish ZD, **Tu T**, Reker D. Profiling Human Microbiome Metabolites via Computational and Experimental Approaches, **Biomedical Engineering Society Annual Meeting (BMES)**, Seattle, WA, October, 2023
- [Poster] Oldfield S, **Tu T**, Eager E, Erickson. Demographic Modeling of Indiana Bat Population Subject to Stress by Wind Energy, **Mathematical Association of America Joint Math Meeting (JMM)**, Atlanta, GA, January, 2016

TEACHING

- **Teaching Assistant for Biostat915**: Conducted weekly office hour discussions and designed practice assignments involving the genetic analysis of global goat populations. Project encompassed addressing genotype missingness, analyzing ancestry distribution/population structure, and simulating phenotypes for a simple GWAS example. (Fall, 2022)
- **Seminar Assistant for Software Carpentry** at Regeneron: Assisted in two-day R/Python training sessions for researchers with minimal programming experience. (2019)

MENTORSHIP

- **Internship panelist** for career dev course BIOTRAIN 898: Gateway to Internship and Experiential Learning (Nov, 2024)
- **Rotation student mentor in Ochoa Lab** (Fall, 2024)
- **Internship panelist** for Duke Computational Biology and Bioinformatics annual program retreat (Nov, 2023)
- Regeneron **Campus Ambassador** for Columbia University (2019-2020)

SCHOLARSHIP

2024	Evolutionary Medicine Summer Institute hosted by The Triangle Center for Evolutionary Medicine (TriCEM), NC State, NC
2023	Advanced Gene Mapping Course; Laboratory of Statistical Genetics, The Rockefeller University, NY
2018	Canadian Bioinformatics 1-Week Workshop for Cancer Genomics; Cold Spring Harbor Laboratory, NY
2016	SAMSI 3-day Undergraduate Workshops for Astronomy; Research Triangle Park, NC
2015	Hong Kong University of Science and Technology Summer Abroad full scholarship; Hong Kong
2013-17	School of Engineering and Applied Sciences Merit-Based Scholarship; George Washington University; Washington, D.C.