Tara Friedrich

+1 (???) ???-???? · tfriedrich.solutions@gmail.com
Open to opportunities in the SF Bay Area or remote

Programming Languages: Python, R, SQL
Data Science: Pandas, Numpy, Scikit-learn, PyTorch, Matplotlib, Seaborn, Ggplot2, Dplyer, Bioconductor, Keras, RShiny
Computational Resources: High Performance Cluster, AWS (S3, EC2, Batch, Lambda), Google Cloud Compute, Kubernetes, Docker, Argo, WDL, Nextflow, Prefect, DNAnexus, Git, JIRA, SQL (MySQL, Spark), noSQL (MongoDB, DynamoDB), Neo4j
Statistics and Machine Learning: Hypothesis Testing, Linear and Generalized Linear Modeling, Multivariate Modeling, Clustering, Dimensionality Reduction, Supervised and Unsupervised Classification, Random Forests, HMM, Kaplan-Meier
Bioinformatics Tools: Dragen, BWA, VEP, FastQC, MultiQC, Samtools, Bedtools, Seurat, Scanpy, Kalisto, DESeq2, EdgeR

Summary

Scientist with 10+ years of experience integrating multimodal clinical and genomic data where statistics and machine learning are fundamental to building reliable products and solving real problems.

Experience

Bioinformatics Engineer Consultant - New Frontiers — April 2025–Present

· Scaled and refactored antibody screening and maturation pipelines using Nextflow to improve reproducibility, performance, and cross-campaign reuse.

· Validated LIMS outputs by writing targeted biological and statistical queries to detect inconsistencies, confirm expected assay behavior, and ensure data integrity for downstream pipelines.

· Built a scalable pipeline to iteratively cluster 2 billion antibody sequences by sequence similarity.

· Contributed to a graph database to model target candidates and their relationships, enabling scalable comparison of sequence similarity, functional annotations, and experimental metrics across discovery campaigns.

· Delivered monthly presentations on project progress and communicated broader technical and strategic insights in company-wide presentations.

Senior Bioinformatics Scientist - Natera — Feb 2022–August 2024

· Partnered with lab scientists to enhance RNA-seq assays, focusing on detecting oncogenic fusions in noisy FFPE tumor samples, leading to improved assay sensitivity and specificity.

· Designed novel quality metrics and visualization tools for germline and somatic pipelines, enabling biomarker-driven patient stratification and deeper characterization of assay performance.

· Built scalable, cloud-native pipelines that reduced turnaround time for WES and RNA-seq QC analyses by 40%.

· Developed an integrated classifier to identify pathogenic genetic variants by combining population genetics, functional annotations, machine learning-based predictors (e.g., AlphaFold), and curated clinical metadata.

Bioinformatics Scientist - Fabric Genomics — Nov 2019–Jan 2022

· Maintained and enhanced the ACE algorithm, a production-grade tool for prioritizing clinically pathogenic variants, by implementing comprehensive regression and unit testing to ensure reliability and compliance in a regulated environment.

· Enhanced variant prediction classifiers to increase precision in identifying likely pathogenic variants and flagging borderline cases for use by CAP/CLIA regulated labs using statistics and machine learning.

· Utilized clinical ontologies (HPO, SNOMED, ICD-9/10) to map patient symptoms to structured phenotypes and prioritize candidate variants for rare disease diagnosis and clinical interpretation.

· Leveraged distributed computing frameworks like Apache Spark to scale data classification workflows and optimize ETL processes, streamlining data analysis pipelines for large-scale datasets.

· Actively participated in code reviews, planning sessions, and best practices discussions, fostering a collaborative and high-quality development environment.

Education

Ph.D. Bioinformatics - University of California, San Francisco

· Statistics and machine learning to identify enhancers and genes predictive of tissue developmental phenotypes.

B.S. Biochemistry - University of California, Los Angeles

Publications

Cordero, G.A., Holloway, A.K., Friedrich, T., Eme, J., Eckalbar, W., Kusumi, K., Janzen, F.J., Hicks, J.W., Conlon, F.L., Bruneau, B.G., Pollard, K.S. **Comparative transcriptomics of the heart illustrates challenges to extending classical theory to gene expression trends in animal development.** Genes, Development and Evolution (Manuscript in preparation).

Friedrich, T.*, Betty, B.*, Mason, M., VanderMeer, J., Zhao, J., Eckalbar, W., Logan, M., Pollard, K. S., Illini, N., Ahituv, N. **Bat accelerated regions identify a bat forelimb specific enhancer in the HoxD locus**. PLoS Genetics 12(3):e1005738 (2016). * Co-first authors.

Eckalbar, W., Schlebusch, S., Mason, M., Gill, Z., Booker, B., Nishizaki, S., Nday, C., Terhune, E., Nevonen, K., Makki, N., Friedrich, T., VanderMeer, J., Pollard, K.S., Carbone, L., Wall, J., Illing, N., Parker, A., Ahituv, N. **Genomic characterization of the developing bat wing.** Nature Genetics 48, 528-536 (2016).

Myers, S.A., Petted, S., Chatterjee, N., Friedrich, T., Tomoda, K., Krings, G., Thomas, S., Broeker, M., Maynard, J., Thomson, M., Pollard, K., Yamanaka, S., Burlingame, A. L., Panning, B. **SOX2 O-GlcNAcylation alters its protein-protein interactions and genomic occupancy to modulate gene expression in pluripotent cells.** eLife 5:e10647 (2016).

Kostka, D., Friedrich, T., Holloway, A. K., Pollard, K.S. **motifDiverge: a model for assessing the statistical significance of gene regulatory motif divergence between two DNA sequences.** Statistics and Its interface 8(4), 463–476 (2015).

Fogel, B. L., Wexler, E., Wahnich A., Friedrich T., Vijayendran C., Gao F., Parikshak N., Konopka G., Geschwind D.H. **RBFOX1 Regulates Both Splicing and Transcriptional Networks in Human Neuronal Development.** Hum. Mol. Genet. 1–16 (2012).

Konopka, G., Friedrich, T., Davis-Turak, J., Winden, K., Oldham M.C., Gao, F., Chen, L, Wang, G., Luo, R., Preuss, T.M., Geschwind, D.H. **Human-Specific Transcriptional Networks in the Brain.** Neuron 75, 601–617 (2012).

Schein, S., Sands-Kidner, M., Friedrich, T. **The physical basis for the head-to-tail rule that excludes most fullerene cages from self-assembly.** Biophys. J. 94, 938–57 (2008).

Schein, S., Friedrich, T. **A geometric constraint, the head-to-tail exclusion rule, may be the basis for the isolated-pentagon rule in fullerenes with more than 60 vertices.** Proc. Natl. Acad. Sci. U. S. A. 105, 19142–7 (2008).

Nagarajan, S., Friedrich, T., Garcia, M., Kambham, N., Sarwal, M.M. **Gastrointestinal leukocytoclastic vasculitis: an adverse effect of sirolimus.** Pediatr. Transplant. 9, 97–100 (2005).