

Eric M. Kernfeld

Collaborative statistician and computational biologist seeking a Boston-area biotech role with opportunities to grow towards new technologies and biological guestions.

Programming

*** **Python** Machine learning, pipelining, data visualization, and benchmarking across multiple roles. PyTorch, numpy, scipy, Seaborn, scikit-learn, pandas, scanpy

Professional Experience

Diabetes Center of Excellence University of Massachusetts Medical School Bioinformatician September 2016 – August 2020

Derived insights from scRNA-seq, ChIP-seq, ATAC-seq, and Hi-C to advance research on stepwise differentiation from endoderm (paper) to foregut (paper) to thymus (paper). Responsible for close collaboration with biologists and driving projects from base-calling all the way through published papers and external talks.

*** Experimental design, data visualization, hypothesis testing, survival analysis, and machine learning across many roles. ggplot2, tidyverse, linear models, seurat, BioConductor

Academic Experience

Department of Biomedical Engineering **Johns Hopkins University** Ph.D. Research Assistant August 2020 - March 2025

Evaluated diverse algorithms for statistical error control in gene regulatory networks (paper) and for prediction of perturbation transcriptomics, including foundation models and graph neural networks (preprint).

SOL **★★★☆☆** Efficient data retrieval via manually written or automatically constructed queries.

Division of Medical Genetics Research Assistant

University of Washington January – June 2016

Julia *** Fast biochemical simulations and related inverse problems.

Analyzed genetic association data seeking loci that contribute to Alzheimer's Disease (paper). Performed data visualization and survival analysis using R and PLINK.

Data Types

ChIP-sea

Docker

scRNA-sea **** Used extensively in most of my research since 2016.

Department of Statistics Teaching Assistant

University of Washington September 2014 – January 2016

Lectured, prepared materials, graded, and held office hours for statistics courses.

ATAC-sea **★★★☆** Experience with end-to-end analysis of multiple types of scATAC data.

 $\star\star\star\star$

★★★☆

Industry Internships

Data Science Team Data Science Ph.D. Intern **Day Zero Diagnostics** July – September 2022

Used collections of data to screen Compared genome assemblies and antibiotic resistance predictions using short-read for technical issues and test and long-read data. Discovered key contamination events. Built a well-documented software pipeline spanning from raw reads to biologically interpretable data displays.

hypotheses from other sources. **★★★☆☆** Brief experience with alignment,

Data Science Team Statistics Intern

MedGenome, Inc. June – August 2015

Using Python, processed exome sequencing data to reduce PCR bias when identifying copy number variations. Using R, planned experimental design for TCR sequencing.

Software Tools

QC, and background models.

Snakemake $\star\star\star\star$ Used at Day Zero and at UMass for NGS quantification pipelines.

Education

Johns Hopkins University Ph.D. Biomedical Engineering

Baltimore, MD August 2020 - March 2025

AWS ★★★☆☆ Used S3 and EC2 throughout Ph.D. work.

Used throughout Ph.D. work.

University of Washington M.S. Statistics

Seattle, WA September 2014 – June 2016

Tufts University B.S. Mathematics, Summa Cum Laude

Medford, MA September 2010 – May 2014