
Nikolai G. Vetr

✉ nikgvetr@stanford.edu | ☎ +1 (602) 578-9196 | in linkedin.com/in/nikolai-vetr | 🌐 github.com/NikVetr
🏠 Nikolai Vetr | 📄 Full CV | 📍 San Francisco Bay Area

Summary

Computational biologist and data scientist focused on Bayesian statistics, multi-omic data integration, and machine learning methods development. Extensive experience developing and evaluating novel statistical models and computational tools to address complex biological questions across varied domains.

Education

PhD University of California, Davis

Anthropology + Population Biology + Data Science & Informatics

2020

Dissertation: “Exploring and Extending Multivariate Brownian Diffusion Models of Phenotypic Evolution for Bayesian Phylogenetic Inference”

BA Vanderbilt University

Earth Environmental Sciences + Ecology, Evolution & Organismal Biology

2013

– Departmental Honors, *summa cum laude*

Experience

Stanford University – Postdoctoral Researcher

Stanford, CA | 2020 – Present

Montgomery Lab (Pathology, Genetics, Biomedical Data Science)

- Integrated differential expression results with drug, perturbation target, and biobank databases with approx. 1TB of input data to provide causal inferential insights into the results of a large-scale animal experiment (*Nat. Comms.*, 2024).
- Made key contributions to multi-omic analysis (GWAS, eQTL, transcriptomics, proteomics, ATAC-seq) of exercise response data for the MoTrPAC, identifying molecular transducers of physical activity (*Nature*, 2024).
- Developed novel Bayesian probability models and analysis pipelines for differential allele-specific expression analysis in ovarian cancer using high-depth mmPCR-seq (*ASHG*, 2024).
- Derived and implemented powerful and efficient computational methods for Bayesian data analysis, including novel correlation matrix prior and proposal distributions, efficient model reparameterizations, and code translation utilities. Runtime improved by multiple orders of magnitude vs. conventional approaches on typical datasets.

Rethink Priorities – Board President / Director

Global | 2023 – Present

- Led Board of Directors for research non-profit focused on cause prioritization, overseeing strategic direction, executive performance, and financial governance, leveraging data modeling skills (eg time series of fundraising performance).
- Provided statistical consulting on Bayesian modeling, causal inference, and data analysis for projects spanning GenAI, global health, survey response data, and computational biology.

University of California, Davis – Instructor / Consultant

Davis, CA | 2015 – 2020

- Developed and delivered lectures, labs, and assessments for upper-division courses in Human and Primate Biology.
- Founded and led interdisciplinary *Applied Bayesian Statistics Research Cluster* (70+ members), providing training and consultation on Bayesian and Deep Learning methods to the academic research community.
- Led workshops and served as a certified Carpentries instructor, teaching coding and data science skills to researchers.
- Won multiple awards for biology and data science outreach activities; also nominated several times for Outstanding Instructor award.

Technical Skills

Programming: R, Stan, BASH, Python, C++, JS, Git, AWS, RevBayes

Statistical / ML Methods: Bayesian Inference, Monte Carlo Methods, Multilevel Models, GLMs / GLMMs, Time Series Analysis, Causal Inference, Optimization, Computer Vision, NLP, Measurement Error, Sparse Estimation, Outlier Detection, Signal Processing, Spatial Autocorrelation, Missing Data Imputation, ODEs, D/CTMCs

Bioinformatics: Multi-omic Data Integration (Genomics, Transcriptomics, Proteomics, Epigenomics), Association Mapping, ASE Analysis, Population and Quantitative Genetics, Interactive / Dynamic Data Visualization, HPC

Domain Knowledge: Molecular Biology, Human Genetics, Exercise Physiology, Evolutionary Biology, Human Anatomy

Languages: English (Native), Russian (Native), Spanish (Proficient)