# Nikolai G. Vetr, PhD

**Contact Information**

**Phone**: (602) 578-9196  
**Email**: nikgvetr@stanford.edu

**LinkedIn**: linkedin.com/in/nikolai-vetr  
**GitHub**: github.com/NikVetr/

## Skills & Interests

- Probability Modeling
- Multiomic Data Analysis
- Time Series Modeling
- Bayesian Methods
- Monte Carlo Methods
- Causal Inference
- Computer Vision
- Artificial Neural Networks
- Machine Learning
- Optimization Methods
- Science Communication
- Nat. Lang. Processing
- Evolutionary Biology
- Exercise Biology
- Population Genetics
- Data Visualization
- High Perf. Computing
- Biomedical Data Science

## Recent Papers

Abell, N., **Vetr, N.***, Montgomery, S., et al. 2024. *A Survey of High Depth Allele-Specific Expression Across Normal Tissues and Ovarian Cancers.* In Prep.

**Vetr, N.**, Gay, N., and Montgomery, S. 2024. *The impact of exercise on gene regulation in association with complex trait genetics.* Nature Communications 15(3346): 1-14. DOI: 10.1038/s41467-024-45966-w.

**MoTrPAC Study Group**[†]. 2024. *Temporal dynamics of the multi-omic response to endurance exercise training across tissues.* Nature 629(8010): 174-183. DOI: 10.1038/s41586-023-06877-w.

Gates, K., Panicker, A., Biendarra-Tiegs, S., **Vetr, N.**, et al. *Shotgun Immunoproteomics for Identification of Nonhuman Leukocyte Antigens Associated With Cellular Dysfunction in Heart Transplant Rejection.* Transplantation 106(7). DOI: 10.1097/TP.0000000000004012

[*]*dual first authorship*, [†]*Author Group: 2 (of 8)*  
*(see second page for a description of work)*

## Leadership

| | | |
|---|---|---|
| **Founder**, Applied Bayesian Statistics Research Cluster, *UC-Davis* | | *2019 - 2020* |
| **President**, Board of Directors, *Wild Animal Initiative* | | *2020 - Present* |
| **President**, Board of Directors, *Rethink Priorities* | | *2023 - Present* |

## Languages

**Programming:** R, Stan, BASH, Python, C++, CSS, HTML, JS  
**Natural:** Russian, English, Spanish

## Education & Training

**Postdoc**, Montgomery Lab, Stanford University — *Current*  
Pathology + Genetics + Biomedical Data Science

**PhD**, University of California, Davis — *2020*  
Dissertation: *Exploring and Extending Multivariate Brownian Diffusion Models of Phenotypic Evolution for Bayesian Phylogenetic Inference*  
Anthropology + Population Biology + Data Science & Informatics

**BA**, Vanderbilt University — *2013*  
Earth & Environmental Sciences + Ecology, Evolution & Organismal Biology  
Departmental Honors, *summa cum laude*

## Teaching

| | | |
|---|---|---|
| **Associate Instructor**, University of California, Davis | | *2015 - 2020* |
| Human Evolution + Primate Evolution + Human Evolutionary Biology | | |
| **Carpentries Instructor**, Data & Software Carpentries | | *2019* |
| **Course Coordinator**, Workshop in Applied Phylogenetics | | *2019* |

**Projects**  Previous **dissertation work** focused on computational methods development in Bayesian phylogenetics. This included developing:

- computational tricks (eg factorizations) to improve scaling of multivariate likelihood calculation
- efficient, tunable proposal distributions and informative (non-$I$ centered) prior distributions for correlation matrices (rank-one updates and downdates w/ Givens rotations to reduce Brownian rate matrix decomposition from $O(n^3)$ to $O(n^2)$)
- extensive simulation experiments to systematically characterize the conditions under which proposed methods succeed and fail
- novel data visualizations (eg *cumulative average resolution* curves) to assist interpretation of the above
- hybrid EM/MCMC approaches to reliably fit otherwise intractable multimodal models (eg phylogenetic diffusion + multivariate ordinal probit)

Current **postdoctoral work** focuses on probability modeling of multiomic dynamics in the biomedical context. Projects here include:

- extensive data analysis and figure generation for preclinical MoTrPAC landscape publication, including integrative proteomic + transcriptomic time series analysis, multiomic & multitissue clustering + network analysis, batch effect adjustment, dimensionality reduction, and uncertainty / error propagation, among other work
- developing novel causal inferential methods, effect standardization workflows, and complex trait enrichment methods, while also leveraging more conventional bioinformatic tools (eg for heritability enrichment)
- performing principled, multilevel modeling of loss-of-heterozygosity and allele-specific expression in a multitissue dataset to identify signatures of ovarian cancer
- developing meta-scientific tools to streamline common computational workflows (eg converting Stan models to interactive R code for predictive simulation, calibration analysis, outlier detection, etc.)

Various **personal projects** leverage scientific and statistical computing workflows to achieve diverse aims. These have entailed writing programs that:

- adaptively compress and vectorize raster images using techniques from computer vision, signal processing, and graph theory
- generate custom crossword puzzles that conform to shapes derived from user-supplied pictures
- glue together API calls to create a fully SMS-based, multimodal digital assistant
- infill and text-wrap a passage of text within a user-specified image
- efficiently predict user ratings and provide media recommendations for popular movies
- interface with external software to assist in metalworking, woodworking, pen plotting, and other physical modification of objects

**Service**  **Journal Review**: *Evolution* (2017), *Science Communications* (2018), *Cell Reports* (2021), *Human Genetics and Genomics Advances* (2022)
**Grant Review**: *WAI Grants* (2021, 2022, 2023)

**Selected Grants & Awards**

| | |
|---|---|
| NIH T15 | *2021* |
| Excellence in Data Science Community Training and Outreach | *2019, 2020* |
| Outstanding Graduate Student Teaching Award Nominee | *2016, 2019, 2020* |
| 1st Place Picnic Day Exhibit Award in "Secrets of Nature" Category | *2017* |
| NSF Graduate Research Fellowship | *2015* |