

# Zachary Rodriguez, PhD

Philadelphia, PA · [linkedin.com/in/zbrodriguez](#) · [github.com/frogriguez](#) · [Google Scholar](#) · [ORCID 0000-0002-4591-803X](#)

Genomics and clinical informatics leader with a PhD in evolutionary biology and postdoctoral training in computational psychiatry. Technical Director of the Penn Medicine Biobank, leading Nextflow pipeline development and genomic data operations for a 300K+ participant biobank. Experienced across GWAS/ExWAS/PheWAS methodology, HPC & cloud infrastructure (AWS, Azure, GCP), and cross-institutional research collaboration.

## EXPERIENCE

Technical Director

2023 – Present

*Penn Medicine Biobank (PMBB), Perelman School of Medicine, University of Pennsylvania*

- Lead development of scalable Nextflow pipelines (GWAS, ExWAS, PheWAS) for 500K+ sample biobank datasets across local HPC and cloud (AWS, Azure, GCP)
- Curate and release the annual PMBB dataset (300K+ participants: genomics, EHR, imaging); project manager for 5+ genomic research teams
- Direct data governance for genomic/EHR data requests, cohort queries, and variant-level extractions across Penn Medicine

Postdoctoral Researcher, Computational Mental Health

2020 – 2023

*Center for Computation & Technology, Louisiana State University*

- Developed ML and NLP pipelines predicting mental health symptoms from archival clinical data; supervised the technical side of 5 research projects
- Ensured HIPAA- and IRB-compliant data integration and management across multi-site collaborations

NSF Graduate Research Fellow

2014 – 2020

*Museum of Natural Science, Louisiana State University*

- Led 3 independent genomics research projects on reptile evolutionary biology from design to publication; secured \$180,000+ in external funding
- Published in *Science Advances*; work covered by NPR, AP, and Science News

## EDUCATION

PhD, Ecology, Evolution, & Genetics

2020

*Louisiana State University*

BS, Biological Sciences

2013

*Clarkson University*

## SKILLS

- Pipelines:** Nextflow, Snakemake, Docker, Singularity, SLURM/ PBS
- Genomics:** GWAS, ExWAS, PheWAS, GSEA, TWAS, PLINK, SAIGE, VEP
- Languages:** Python, R, SQL, Bash
- Cloud:** AWS, Azure, GCP, Databricks
- Clinical:** EHR/HIPAA, IRB compliance, OMOP
- ML/Stats:** NLP, LLMs, Bayesian & mixed models

## SELECTED PUBLICATIONS

- Palmer, Hill, Hodgson, et al., **Rodriguez, Z.B.**, et al., for BRaVa (2026). [The Biobank Rare Variant consortium powers the discovery of rare genetic associations through global collaboration](#). *medRxiv*.
- **Rodriguez, Z.B.**, Guare, L., Cardone, K.M., Carson, C.C., Caruth, L., & Verma, A. PMBB Geno-Pheno Toolkit: A suite of scalable, reproducible pipelines for cross-biobank association analyses. *Bioinformatics* (under review).
- Verma, A., **Rodriguez, Z.B.**, Guare, L., Cardone, K., & Carson, C. (2024). [Command line to pipeLine: Cross-biobank analyses with Nextflow](#). *PSB 2025 Proceedings*, 696–701.
- **Rodriguez, Z.B.**, Perkins, S.L., & Austin, C.C. (2018). [Multiple origins of green blood in New Guinea lizards](#). *Science Advances*, 4(5), eaao5017.
- Cohen, A.S., **Rodriguez, Z.B.**, et al. (2022). [Natural language processing and psychosis: On the need for comprehensive psychometric evaluation](#). *Schizophrenia Bulletin*, 48(5), 939–948.