Cole Williams

Curriculum Vitae

Center for Computational Molecular Biology
Brown University

✉ cole.w@me.com   📞 (530) 304-7735   🌐 williamscole.github.io

Education

Ph.D. in Computational Biology   Expected 2025
Brown University

- Advisor: Dr. Sohini Ramachandran

- Committee: Dr. Emilia Huerta-Sanchez, Dr. Dan Weinreich, Dr. Amy Williams

M.S. in Computational Biology   2024
Brown University

- Earned during Ph.D. studies

- GPA: 4.00/4.00

B.S. in Genetics & Genomics   2019
University of California, Davis

- GPA: 3.99/4.00

- Honors Thesis: "Inferring genetic relatedness in an endogamous African population"

Research Experience

Graduate Research Fellow   2020–Present
Ramachandran Lab, Brown University

- Developed HAPTIC, a novel inter-chromosomal phasing algorithm using spectral clustering of identity-by-descent (IBD) segments, successfully scaling to analyze 500,000+ UK Biobank samples and validated on 23andMe's dataset.

- Engineered and optimized high-performance computing code for large-scale genetic relatedness analysis, enabling efficient processing of genome-wide IBD segments across millions of sample pairs.

- Developed PONDEROSA, a machine learning algorithm for predicting genealogical relationships from IBD segments, with ongoing development of CNN-based approaches for improved accuracy.

- Led research on genetic privacy implications, resulting in forthcoming publication on policy recommendations and best practices.

Professional Research Assistant   2019–2020
Gignoux Lab, University of Colorado, Anschutz Medical Campus

- Developed clustering methods to identify genetic relatives in the San Antonio Mexican American Families Study (SAMAFS).

- Assisted with biobank team to implement relatedness detection algorithms for the Colorado Center for Personalized Medicine biobank.

Undergraduate Researcher   2018–2019
Henn Lab, University of California, Davis

- Investigated genetic relatedness in the Himba population.

- Began development of PONDEROSA, for inferring genealogical relationships using IBD data.

Publications

Journal Articles

1. Moody, N.M., **Williams, C.M.**, Ramachandran, S., & Fuxjager, M.J. Social mates dynamically coordinate aggressive behavior to produce flexible territorial strategies. *In revision at PLOS Computational Biology.*

2. **Williams, C.M.**, O'Connell, J., Freyman, W.A., 23andMe Research Team, Gignoux, C.R., Ramachandran, S., & Williams, A.L. Phasing millions of samples achieves near perfect accuracy, enabling parent-of-origin classification of variants. *bioRxiv*, May 2024.

3. **Williams, C.M.**, Seelza, B.A., Slack, S.D., Mathias, R.A., Watson, H., Barnes, K.C., Lange, E., Johnson, R.K., Gignoux, C.R., Ramachandran, S., & Henn, B.M. A rapid, accurate approach to inferring pedigrees in endogamous populations. *bioRxiv*, July 2024.

4. Wiley, L.K., Shortt, J.A., Roberts, E.R., Lowery, J., Kudron, E., Lin, M., Mayer, D., Wilson, M., Brunetti, T.M., Chavan, S., Phang, T.L., Pozdeyev, N., Lesny, J., Wicks, S.J., Moore, E.T., Morgenstern, J., Roth, A.N., Shapowitz, E.L., Stewart, A., **Williams, C.M.**, [...], Barnes, K.C., & Gignoux, C.R. Building a vertically integrated genomic learning health system: The biobank at the Colorado Center for Personalized Medicine. *The American Journal of Human Genetics*, 111(1):11–23, January 2024.

5. Swinford, N.A., Prall, S.P., Gopalan, S., **Williams, C.M.**, Sheehama, J., Seelza, B.A., & Henn, B.M. Increased homozygosity due to endogamy results in fitness consequences in a human population. *Proceedings of the National Academy of Sciences*, 120(43):e2309552120, October 2023.

6. Gopalan, S., Berl, R.E.W., Myrick, J.W., Garfield, Z.H., Reynolds, A.W., Bafens, B.K., Belbin, G., Mastoras, M., **Williams, C.**, Daya, M., Negash, A.N., Feldman, M.W., Hewlett, B.S., & Henn, B.M. Hunter-gatherer genomes reveal diverse demographic trajectories during the rise of farming in Eastern Africa. *Current Biology*, 32(8):1852–1860.e5, April 2022.

Conference Abstracts & Presentations

1. **Williams, C.M.**, O'Connell, J., 23andMe Research Team, Gignoux, C.R., Ramachandran, S., & Williams, A.L. Phasing of millions of samples achieves near perfect accuracy, enabling parent-of-origin classification of variants. Poster presentation at the American Society for Human Genetics (ASHG) annual meeting. Washington, D.C. November 2023.

2. **Williams, C.M.**, 23andMe Research Team, Gignoux, C.R., Ramachandran, S., & Williams, A.L. Parent of origin classification of identity by descent segments. Poster presentation at the Probabilistic Modeling in Genomics meeting. Cold Spring Harbor, NY. March 2023.

3. **Williams, C.M.**, Seelza, B.A., Shortt, J.A., Lange, E., Rafaels, N., Crooks, K., Barnes, K.C., Gignoux, C.R., & Henn, B.M. A rapid, accurate approach to inferring pedigrees in endogamous populations. Poster presentation at the American Society for Human Genetics (ASHG) annual meeting. Houston, TX. October 2019.

4. **Williams, C.M.**, Seelza, B.A., & Henn, B.M. A rapid, accurate approach to inferring pedigrees in endogamous populations. Talk presented at the Bay Area Population Genomics meeting. Palo Alto, CA. June 2019.

Technical Skills

Programming

- Advanced: Python, Bash

- Intermediate: R

- Basic: C++

Tools & Technologies

- Version Control: Git/GitHub

- Development Tools: Apptainer/Singularity, VS Code, Jupyter Notebooks

- Workflow management: Nextflow, Snakemake

Data Analysis

- Dimensionality reduction: PCA, LDA

- Machine Learning: scikit-learn, TensorFlow

- Visualization: seaborn, matplotlib

Leadership & Service

Academic Service

- Graduate Teaching Assistant (BIOL495: Statistical Analysis of Biological Data)

- Mentorship of Data Science Master's student

Professional Service

- Co-founder: Science on the Hill outreach program