

Nathan LaPierre

nathanrlapierre.com
NathanL2012@gmail.com

RECENT POSITIONS

Postdoctoral Scholar Aug. 2022 - Present
Department of Human Genetics, University of Chicago
Labs of Xin He and Matthew Stephens

Graduate Student Researcher Sep. 2016 - June 2022
Department of Computer Science, University of California, Los Angeles
Labs of Eleazar Eskin and Wei Wang

EDUCATION

PhD in Computer Science, June 2022
University of California, Los Angeles

M.S. in Computer Science, 3.74 GPA, March 2019
University of California, Los Angeles

B.S. in Applied Computer Science, 3.86 GPA (*magna cum laude*), December 2015
George Mason University

PUBLICATIONS

N. LaPierre*, B. Fu*, S. Turnbull, E. Eskin, and S. Sankararaman, "Leveraging family data to design Mendelian Randomization that is provably robust to population stratification," *bioRxiv*, Jan. 2023.

A. Khan, J. Bagley, **N. LaPierre**, et al, "Genetic pathways regulating the longitudinal acquisition of cocaine self-administration in inbred and recombinant inbred mice," *bioRxiv*, Nov. 2022.

J. Karlin, L. Gai, **N. LaPierre**, et al, "Ensemble neural network model for detecting thyroid eye disease using external photographs," *British Journal of Ophthalmology*, Sep. 2022.

F. Meyer et al (many authors), "Critical Assessment of Metagenome Interpretation-the second round of challenges," *Nature Methods*, vol. 19, pp. 429440, Apr. 2022.

C. Cinelli, **N. LaPierre**, B. L. Hill, S. Sankararaman, and E. Eskin, "Robust Mendelian randomization in the presence of residual population stratification, batch effects and horizontal pleiotropy," *Nature Communications*, vol. 13, no 1., pp. 1093, Mar. 2022.

M. Alser et al (many authors), "Packaging, containerization, and virtualization of computational omics methods: Advances, challenges, and opportunities," *arXiv*, Mar. 2022.

N. LaPierre*, K. Taraszka*, H. Huang, R. He, F. Hormozdiari, and E. Eskin, "Identifying Causal Variants by Fine Mapping Across Multiple Studies," *PLOS Genetics*, vol. 17, no. 9, pp. e1009733, Sept. 2021.

J. Bloom et al (many authors), "Massively scaled-up testing for SARS-CoV-2 RNA via next-generation sequencing of pooled and barcoded nasal and saliva samples,"

Nature Biomedical Engineering, vol. 5, no. 7, pp. 657-665, July 2021.

N. LaPierre, M. Alser, E. Eskin, D. Koslicki*, and S. Mangul*, “Metalign: Efficient alignment-based metagenomic profiling via containment min hash,” *Genome Biology*, vol. 21, pp. e242, Sep. 2020.

N. LaPierre, R. Egan, W. Wang, and Z. Wang, “De novo Nanopore read quality improvement using deep learning,” *BMC Bioinformatics*, vol. 20, no. 1, pp. e552, Dec. 2019.

N. LaPierre, C. Ju, G. Zhou, and W. Wang, “MetaPheno: A Critical Evaluation of Deep Learning and Machine Learning in Metagenome-Based Disease Prediction,” *Methods*, vol. 166, pp. 74-82, Aug. 2019.

N. LaPierre*, S. Mangul*, M. Alser, I. Mandric, N.C. Wu, D. Koslicki, and E. Eskin, “MiCoP: Microbial Community Profiling method capable of detecting low abundance viral and fungal organisms in metagenomic samples,” *BMC Genomics*, vol. 20, no. 5, pp. e423, June 2019.

M. A. Rahman, **N. LaPierre**, H. Rangwala, and D. Barbara, “Metagenome sequence clustering with hash-based canopies,” *Journal of bioinformatics and computational biology*, vol. 15, no. 6, pp.1740006, Oct. 2017.

M. A. Rahman, **N. LaPierre**, and H. Rangwala, “Phenotype Prediction from Metagenomic Data Using Clustering and Assembly with Multiple Instance Learning (CAMIL),” *IEEE/ACM transactions on computational biology and bioinformatics*, Oct. 2017.

N. LaPierre, M. A. Rahman, and H. Rangwala, “CAMIL: Clustering and Assembly with Multiple Instance Learning for Phenotype Prediction,” in *IEEE International Conference on Bioinformatics and Biomedicine*, Shenzhen, China, 2016.

N. LaPierre and H. Rangwala, “Predicting Clinical Phenotype using OTU-based Metagenome Representation,” in *IEEE International Conference on Data Mining Workshop on Biological Data Mining and its Applications in Healthcare*, Atlantic City, New Jersey, 2015, pp. 156-163.

(* Authors contributed equally)

PRESENTATIONS

“Leveraging family data to design Mendelian Randomization that is provably robust to population stratification,” in *International Conference on Research in Computational Molecular Biology (RECOMB)*, Apr. 2023.

“Accounting for Isoform Expression in eQTL Mapping,” in *CSHL Genome Informatics*, virtual, Nov. 2021.

“Metalign: Efficient alignment-based metagenomic profiling via containment min hash,” in *Intelligent Systems for Molecular Biology (ISMB) HitSeq*, virtual, July 2020.

“Identifying Causal Variants by Fine Mapping Across Multiple Studies” in *International Conference on Research in Computational Molecular Biology (RECOMB)*, virtual, June 2020.

“Metalign: Efficient alignment-based metagenomic profiling via containment min hash,” in *RECOMB-Seq*, virtual, June 2020.

“Identifying Causal Variants by Fine Mapping Across Multiple Studies” in American Society for Human Genetics Annual Meeting, October 2019.

“MiniScrub: de novo long read scrubbing using approximate alignment and deep learning” in Amazon Web Services - UCLA Computational Medicine Symposium, February 2019.

“MiniScrub: de novo long read scrubbing using approximate alignment and deep learning” in American Society for Human Genetics Annual Meeting, October 2018.

“MiCoP: Microbial Community Profiling method capable of detecting low abundance viral and fungal organisms in metagenomic samples” in American Society for Human Genetics Annual Meeting, October 2017.

“CAMIL: Clustering and Assembly with Multiple Instance Learning for Phenotype Prediction” in IEEE International Conference on Bioinformatics and Biomedicine, December 2016.

- Won a conference travel grant sponsored by NSF

“Predicting Clinical Phenotype using OTU-based Metagenome Representation” in IEEE International Conference on Data Mining workshop on Biological Data Mining and its Applications in Healthcare, November 14, 2015

- Won a travel grant from the Undergraduate Student Travel Fund of the Office of Student Scholarship, Creative Activities, and Research at GMU

“Developing a Computational Pipeline for Metagenomic State Classification with Feature Engineering” in Volgenau School of Engineering Undergraduate Research Celebration, April 2015

- Won Outstanding Undergraduate Research Project Award for poster

RESEARCH EXPERIENCE

PhD Student

University of California, Los Angeles

September 2016 - June 2022

- Working with Professors Wei Wang, Eleazar Eskin, and Harold Pimentel
- Developed a method for identifying causal variants (statistical fine mapping) by leveraging information from multiple studies using a Bayesian approach
- Developing multiple Mendelian Randomization methods providing sensitivity analysis and robustness to confounding from population structure
- Contributed to the effort to develop & deploy a saliva-based COVID-19 test (“Swab-Seq”) that is now deployed campus-wide at UCLA
- Developing a method for isoform-aware eQTL mapping
- Developed multiple published methods that perform accurate abundance profiling of microbial organisms based on metagenomic reads
- Developed a method that uses deep learning to improve long sequencing read quality, in collaboration with scientists from the Joint Genome Institute
- Two-time teaching assistant (courses: Computational Genetics and Machine Learning in Genetics), given an average rating of 8 out of 9 by my students
- Mentored four undergraduate summer students, two of whom are now PhD students

Graduate Research Assistant (GRA) / Student Researcher

George Mason University

January - August 2016 (June-August as GRA)

- Worked with Professor Huzefa Rangwala
- Developed CAMIL, a pipeline that uses multiple instance learning techniques based on whole metagenome shotgun sequence reads to predict whether or not a patient has a disease.
- CAMIL paper accepted into IEEE BIBM 2016 (19% acceptance rate).
- Second author of paper on using canopy clustering and locality sensitive hashing to reduce clustering time for biological datasets.

Predicting Clinical Phenotype using OTU-based Metagenome Representation

George Mason University

January 2015 - November 2015

- Worked with Professor Huzefa Rangwala
- Developed a computational pipeline that uses clustering and classification methods to quickly and accurately predict whether a patient has a disease based on a case/control metagenomic dataset
- Paper accepted for publication, and poster presentation won Outstanding Undergraduate Project Award (see publications and presentations sections)

INTERNSHIPS

Graduate Student Researcher

Summer 2017

DOE Joint Genome Institute / Lawrence Berkeley National Lab

- Developed a method that uses a Convolutional Neural Network (Deep learning method) to improve Oxford Nanopore long read quality *de novo*
- Method improves both read accuracy and downstream *de novo* assembly
- Helped pioneer use of deep learning within the lab

Security Engineering Intern

Summers 2013-2015

Sony Corporation of America

- Wrote secure and scalable software and worked with Big Data in order to help analyze, detect, and prevent attacks on Sony's networks
- Used Python, Javascript (Node, Express, Meteor), MongoDB, X/HTML, CSS
- Developed a workplace communications system using Meteor.js; solo project
- Assisted in the development of a web application based on a searchable database system using Node.js, Express.js, and MongoDB
- Assisted in the development of a network forensics system; wrote backend python scripts and XML web layouts

TEACHING EXPERIENCE

Graduate Teaching Assistant

Spring 2020

CS/BIOINFO 122/222 - Algorithms in Bioinformatics, Computer Science Department, UCLA

Graduate Teaching Assistant

Winter 2020

CS/BIOINFO 124/224 - Machine Learning in Genomics, Computer Science Department, UCLA

Undergraduate Teaching Assistant

Fall 2014 - Spring 2015

CS 306 - Computer Law and Ethics, Computer Science Department, George Mason University

- Received Outstanding Undergraduate Teaching Assistant Award
- Assisted students with legal research, writing, and oral communication
- Responsible for grading student assignments worth 25% of their grade; one of the only Undergraduate Teaching Assistants entrusted with this responsibility

AWARDS, FELLOWSHIPS, SCHOLARSHIPS

Selected Honors

- *Honorable Mention in NSF Graduate Research Fellowship*, 2015-16
- *Outstanding Undergraduate Student Award*, given to overall best undergraduate student in Computer Science at George Mason University, May 2016
- *Outstanding Academic Achievement Award*, given for outstanding performance in Computer Science, May 2016
- *Outstanding Undergraduate Research Project* for presentation of my research project at Volgenau School of Engineering Undergraduate Research Symposium, April 2015
- *Outstanding Undergraduate Teaching Assistant* for two semesters of excellence as a teaching assistant, April 2015
- *Dean's List* every semester

Fellowships and Merit-Based Scholarships

- *ModEling and uNdersTanding human behaviOR (MENTOR) NSF Training Grant*, 2018-19 academic year at UCLA
- *NIH T32 Doctoral Training Fellowship*, 2016-18 academic years at UCLA
- *Bersoff Endowed Scholarship*, Outstanding Academic Achievement, Awarded twice in 2015 and 2014
- *SWIFT Scholarship*, Outstanding Academic Achievement, Awarded in 2015

Honors Societies and Organizations

- *Honors College* at George Mason University
- *Alpha Lambda Delta Honor Society*
- *Golden Key International Honour Society*

Technical Competitions

- *Top 5 at VTHacks*, a software development competition at Virginia Tech with over 45 teams, April 2014
- *2nd Place in the Technical Innovation Challenge*, a week-long competition at George Mason University to design a viable software product, jointly refereed by Computer Science and Business Departments, October 2014

VOLUNTEER EXPERIENCE

- Executive Curriculum Planner and Mentor, Community Programming Initiative
George Mason University Honors College and SRCT 2014 - 2015
- Volunteer effort to teach basic programming to local 5th-8th grade students
 - Planned and developed parts of the curriculum for the sessions, such as designing games and hints to help the students create them
 - Mentored the elementary and middle school students during the sessions

COMPUTATIONAL SKILLS

Languages / Scripting: Python, R, C, Matlab/Octave, bash

Deep Learning Frameworks: Keras, TensorFlow, PyTorch

Other Technologies: Git, Docker, LaTeX

Operating Systems: Linux, Windows, macOS

Experience with:

- Machine Learning Algorithms: Deep Learning, Linear Regression, Logistic Regression, SVMs, Random Forests, etc
- Machine Learning Applications: Clustering, Classification, Bioinformatics
- Bioinformatics methods experience: statistical fine mapping, alignment, assembly, read error correction, metagenomics, abundance profiling
- SGE and Slurm grid/cluster computing systems

MEMBERSHIPS *Student-Run Computing and Technology (SRCT)* 2013 - 2016
AND LEADERSHIP

- Student organization at George Mason University that works on software projects and competitions to benefit the university and broader local community
- Secretary and Member of Executive Board, Fall 2015 Semester