jacob.krol@colorado.edu
github.com/jakekrol

Jacob D Krol

linkedin.com/in/jacob-krol-b3b784156

Education

Ph.D. Computer Science | University of Colorado Boulder | Boulder, CO **Aug 2024 — present**

- Advisor: Ryan Layer

Graduate-level Non-degree | University of Colorado Anschutz Medical Campus | Aurora, CO **Sep 2023 — Dec 2023**

- GPA: 4.00/4.00
- BIOS 7747: Machine Learning for Biomedical Applications, graduate course offered by the Colorado School of Public Health

Bachelor of Science, Neuroscience | Michigan State University | East Lansing, MI **Sep 2020 — Aug 2022**

- GPA: 3.89/4.00
- Graduation awards: *Cum Laude*
- Semester awards: Dean's List

Math and Science Transfer Program | Washtenaw Community College | Ann Arbor, MI **Sep 2017 — Apr 2020**

- GPA: 3.52/4.00
- Semester awards: Honor Roll

Research Interests

Computational biology, Statistics, Algorithms, Graph theory, Genomics, Natural language processing, History, Evolution

Research Experience

Graduate Research Assistant **Aug 2024 — Present**

Layer lab | Department of Computer Science | PI: Ryan Layer
University of Colorado Boulder | Boulder, CO

NIH/NLM Data Science and Informatics Scholars Summer Intern **Jun 2024 — Aug 2024**

Protein and Genome Evolution Research Group | PI: Aravind Iyer
National Institute of Health (NIH/NLM/NCBI) | North Bethesda, MD

- Developed methods for discovering genomic conflict systems through natural language processing (NLP) and network analysis.
- Trained Word2Vec model to learn context-aware gene embeddings and extracted genomic conflict systems from cosine similarity networks of the gene embedding space.
- Identified co-localized genes/proteins N-grams (N>=2) associated with biological conflict systems across the evolutionary tree using entropy and frequency analysis.
- Presented findings at two poster events on NIH campus.
- Member of Transformers for AI journal club.

Information Sciences Professional **Nov 2022 — May 2024**

JRavi Lab | Department of Biomedical Informatics | PI: Janani Ravi
University of Colorado School Anschutz School of Medicine | Aurora, CO

- Developed an internal R package for machine learning analysis of antimicrobial resistant (AMR) bacterial pathogens.
- Full stack web development for *MolEvolvR: a web app for characterizing proteins using molecular evolution and phylogeny.*, Krol, et al., 2023; bioRxiv, DOI: doi.org/10.1101/2022.02.18.461833 (jravilab.org/molevolvr).
- Submitted graduate-level fellowships to national agencies: 1) NSF GRFP and 2) DOE CSGF.
- Presented research talks and software development workshops at international conferences: 1) R Bioconductor 2023 and 2) Great Lakes Bioinformatics 2023.
- Led workshops on shell programming, Git version control, remote linux computing, and Docker.
- Managed the lab's GitHub organization: code review plus managing issues and pull requests.
- Peer-mentored Ph.D. (3) and undergraduate students (3) on R/Python/Shell programming, version control, data wrangling, developing ML models, hypothesis testing, web development, and presenting scientific research.
- Assisted in public deployment of R Shiny dashboard for the publication *The Phage-shock-protein (PSP) Envelope Stress Response: Discovery of Novel Partners and Evolutionary History. DOI: doi.org/10.1101/2020.09.24.301986*
- Performed system administrator tasks (e.g., server onboarding, dependency/user/data/resource management) for lab's server.

Student Research Assistant **May 2022 — Nov 2022**

Krishnan Lab & Malmstrom Lab, Department of Computational Mathematics Science and Engineering (CMSE)
Michigan State University, East Lansing, MI

- Developed machine learning classifiers trained on viral omics data to predict plant virus host and virus taxonomy.
- Analyzed the association of protein domains with viral host phenotype through Fisher's Exact test hypothesis testing
- Extracted latent viral genomic features with principal component analysis (PCA).
- Presented the approach and results in project meetings.

Publications

1. **Krol**, **J. D.** *et al.* MolEvolvR: A web-app for characterizing proteins using molecular evolution and phylogeny. *bioRxiv. DOI: doi.org/10.1101/2022.02.18.461833* (2023).

Awards and Funding

- **Awarded** National Institute of Health (NIH) Intramural Research Training Award (IRTA) Summer Student Traineeship. **2024**
- **Awarded** NSF ACCESS explore project allocation for high performance computing: **2024**
  "Processing sequences into feature datasets for microbial genotype to phenotype machine learning."
- Submitted Winter 2023 Department of Energy Computational Science Graduate Fellowship (DOE CSGF): **2023**
  A multi-modal deep learning technique predict antibiotic resistance via computational chemistry and bacterial genomics. (Not awarded).
- Submitted Fall 2023 NSF Graduate Research Fellowship Program (GRFP): **2023**
  "*Microbial phenotype prediction with graph machine learning methods". (Not awarded)*.

Presentations

Research and technical talks

- R/Bioconductor: Cancer and Evolution track. *MolevolvR a web-app for protein characterization.* **Jul 2023**
  Boston University, Boston, MA.
- Great Lakes Bioinformatics Conference. *MolEvolvR a web-app for protein characterization.* **May 2023**
  McGill University, Montreal, CA.
- Great Lakes Bioinformatics Conference. *How and when to build a web-app or R 2023 package?* **May 2023**
  McGill University, Montreal, CA.

Posters

- National Institute of Health Summer Poster Day *The language of genomic conflict systems* **Aug 2024**
  National Institute of Health, Bethesda, MD.
- American Society for Microbiology Rocky Mountain Branch *Classifying antimicrobial resistance in high-impact pathogens* **Apr 2024**
  University of Colorado Boulder, Boulder, CO.
- CU Department of Biomedical Informatics Annual retreat. *Classifying antimicrobial resistance in high-impact pathogens* **Sep 2023**
  University of Colorado Anschutz, Aurora, CO.

Peer Mentees

PhD Students

- Charmie Vang, Biomedical Sciences program, CU Anschutz **2023**
- Keenan Manpearl, Computational Bioscience program, CU Anschutz **2023**
- Jill Bilodeaux, Microbiology program, CU Anschutz **2023**

Undergraduates

- Skylar Stefonowicz, B.S. Biology, Metropolitan State University of Denver **2024**
- Ethan Wolfe, B.S. Biochemistry & Molecular Biology with CMSE and additional minors, Michigan State University **2022 — 2024**

References

- **Aravind Iyer, Ph.D.**; aravind@ncbi.nlm.nih.gov
  — Senior Investigator of Protein and Genome Evolution Research Group, NLM/NCBI
- **Janani Ravi, Ph.D.** ; janani.ravi@cuanschutz.edu
  — Assistant Professor and Principal Investigator of the JRavi Lab, Department of Biomedical Informatics, CU Anschutz
- **Arjun Krishnan, Ph.D.** ; arjun.krishnan@cuanschutz.edu
  — Associate Professor and Principal Investigator of the Krishnan Lab, Department of Biomedical Informatics, CU Anschutz