

# GRAHAM LARUE

📞 (831) 420-7720    ✉ graham@glarue.net    🔗 LinkedIn    🐙 GitHub

## SUMMARY

---

I'm a scientist with broad interests and a focus on bioinformatics, data engineering, and building tools that enable empirically grounded insight and discovery. In both academic and industry settings, I've emphasized rigor, reproducibility, clarity, and collaboration in tackling complex biological questions. Alongside my work in biotech, I bring years of experience in the service industry, which fostered a people-centered approach rooted in respectful, clear, and responsive communication. With over a decade of experience in Python and bioinformatics, I value the open-source ethos in software and science, and aim to develop maintainable, robust tools and workflows that support diverse research goals.

## EDUCATION

---

<b>Ph.D.</b> , Quantitative and Systems Biology <i>University of California Merced</i>	2016 – 2022
<b>M.S.</b> , Cell and Molecular Biology <i>San Francisco State University</i>	2013 – 2015
<b>B.S.</b> , Biology <i>University of California Santa Cruz</i>	2002 – 2008

## WORK EXPERIENCE

---

<b>Bioinformatics Data Scientist</b> <i>Mammoth Biosciences</i>	Nov. 2022 – May 2025 <i>Brisbane, CA</i>
--	---

- Large-scale metagenomic protein discovery and analysis for early R&D; sequence- and domain-based homolog search, clustering, filtering and sub-selection.
- Protein and DNA sequence generation and analysis using variational autoencoder and generative AI approaches.
- Custom bioinformatics workflow development and deployment on AWS Batch infrastructure.
- Collaborative software development, CI/CD and code review using GitHub and Jupyter notebooks.
- SQL database design, population, and maintenance for biological sequence data storage, retrieval, and analysis.

<b>Graduate Student Researcher</b> <i>University of California Merced</i>	Aug. 2016 – Aug. 2022 <i>Merced, CA</i>
--	--

Bioinformatics and molecular evolution doctoral research in Dr. Scott Roy's group, focused on the evolutionary dynamics of minor spliceosomal introns across diverse eukaryotic genomes. I developed software to classify genomic sequences using machine learning techniques, identify orthologous gene clusters using graph-based algorithms, and extract and catalog annotated genomic feature sequences from thousands of genome and annotation files. Additional experience with RNA-seq-based gene annotation pipelines and Linux system administration.

<b>Summer Graduate Intern</b> <i>DOE Joint Genome Institute</i>	Jun. 2019 – Sep. 2019 <i>Walnut Creek, CA</i>
--	--

Big data comparative genomic analysis of intron lengths in thousands of fungal genomes.

**Staff Research Associate***University of California Berkeley*

Jul. 2011 – Jul. 2013

*Berkeley, CA*

Electrophysiological research on mammalian retinal architecture in the labs of Professor Frank Werblin (Neurobiology) and Dr. Long-Sheng Fan (EE).

**PROGRAMMING**

---

**Advanced:** Python, Bash**Intermediate:** Pandas, SQL, Git**Basic:** R, Julia, Perl, C++**TECHNICAL EXPERIENCE**

---

- **Bioinformatics:** Designed and built: novel SVM-based intron sequence classification ML algorithm (intronIC), custom ortholog and intron homology identification pipelines, command-line JGI data retrieval utility (jgi-query). Experienced in: de novo transcriptome assembly using RNA-seq, metagenomic genome assembly, annotation and protein discovery, RNA-seq-based eukaryotic gene annotation, phylogenetic tree construction, sequence clustering, machine learning and generative AI-based sequence generation.
- **Computing:** CI/CD, Linux system administration, Git, GNU coreutils, AWS, SQL, Jupyter, Slurm.
- **Data visualization:** Matplotlib, Seaborn, Plotly, Altair.

**SELECTED PUBLICATIONS**

---

- **Larue, G.E.**, Roy, S.W., 2023. “Where the minor things are: a pan-eukaryotic survey suggests neutral processes may explain much of minor intron evolution.” *Nucleic Acids Res.*, Volume 51, Issue 20, 10 November 2023, Pages 10884–10908.  
<https://doi.org/10.1093/nar/gkad797>
- **Larue, G.E.**, Eliáš, M., Roy, S.W., 2021. “Expansion and transformation of the minor spliceosomal system in the slime mold *Physarum polycephalum*.” *Current Biology*, 31:3125–3131.e4.  
<https://doi.org/10.1016/j.cub.2021.04.050>
- Moyer, D.C.\*; **Larue, G.E.\***; Hershberger, C.E.; Roy, S.W.; Padgett, R.A., 2020. “Comprehensive database and evolutionary dynamics of U12-type introns.” *Nucleic Acids Res.*, Volume 48, Issue 13, 27 July 2020, Pages 7066–7078.  
<https://doi.org/10.1093/nar/gkaa464> (\*co-first author)
- Hudson, A.J.; McWatters, D.C.; Bowser, B.A.; Moore, A.N.; **Larue, G.E.**; Roy, S.W.; Russell, A.G., 2019. “Patterns of conservation of spliceosomal intron structures and spliceosome divergence in representatives of the diplomonad and parabasalid lineages.” *BMC Evolutionary Biology*, 19:162.  
<https://doi.org/10.1186/s12862-019-1488-y>
- Slabodnick, M.M.; Ruby, J.G.; Reiff, S.B.; Swart, E.C.; Gosai, S.; Prabakaran, S.; Witkowska, E.; **Larue, G.E.**; Fisher, S.; Freeman, R.M.; Gunawardena, J.; Chu, W.; Stover, N.A.; Gregory, B.D.; Nowacki, M.; Derisi, J.; Roy, S.W.; Marshall, W.F.; Sood, P., 2017. “The Macronuclear Genome of *Stentor coeruleus* Reveals Tiny Introns in a Giant Cell.” *Current Biology*, 27:569–575.  
<https://doi.org/10.1016/j.cub.2016.12.057>