

Rodolfo Aramayo, PhD

Computational Biologist & Molecular Geneticist

Large Protein Models | Variant Interpretation | Computational Genomics

College Station, Texas | rodolfo@aramayo.org

[Website](#) | [LinkedIn](#) | [GitHub](#) | [ORCID](#)

PROFESSIONAL SUMMARY

Interdisciplinary computational biologist and molecular geneticist working at the interface of protein-language-model evaluation, variant-interpretation research, and genomics. Develops and rigorously evaluates reproducible systems built on ESM-family protein language models (PLMs) for zero-shot missense scoring, translating DNA, RNA, and protein biology into robust computational representations. Combines deep experimental intuition with benchmark design, scientific software engineering, high-performance computing, and leadership of cross-disciplinary teams.

Professional focus: Representing, comparing, and interrogating biological information across scales to reveal hidden patterns and enable biology-grounded AI.

TECHNICAL EXPERTISE

Protein Language Models & Evaluation: ESM1b, ESM1v, ESM2 through 15B parameters, ESMC, ESM3, zero-shot variant scoring, per-target ROC/AUROC, stratified benchmarking, parameter scaling, accuracy-compute tradeoffs, failure analysis, reproducibility, release gating.

Variant Interpretation & Genomics: AlphaFold pLDDT, ClinVar, dbSNP, Ensembl, HGVS, MANE Select, VCF, research-focused VUS prioritization, genome and transcriptome assembly, NGS, comparative genomics, bulk and single-cell multiomics.

Scientific Computing: Python, R, Bash/Shell, PyTorch, scikit-learn, pandas, NumPy, Bioconductor, Git/GitHub, Linux/Unix, Slurm, HPC, parallel computing, portable CPU/GPU workflows, cloud computing.

Biological Foundation: Molecular genetics, RNA biology, epigenetics, genome engineering, strain construction, DNA/RNA/protein purification and characterization, experimental design, biotechnology consulting.

SELECTED AI & COMPUTATIONAL BIOLOGY WORK

Large Protein Models for Variant Prediction | Current Research

- Develop and evaluate systems using ESM1b, ESM1v, ESM2 models through 15B parameters, ESMC, and ESM3 for zero-shot missense-variant scoring; this work builds on existing protein foundation models rather than training them.
- Design per-target ROC/AUROC, stratified benchmarking, and parameter-scaling analyses to compare predictors, quantify performance, and evaluate accuracy-compute operating points.
- Interpret model scores with AlphaFold pLDDT, structural context, deterministic and chemistry-based signals, curated ClinVar/dbSNP/Ensembl evidence, HGVS nomenclature, and MANE Select transcripts.
- Investigate failure modes associated with training-data composition, low-confidence and disordered protein regions, and edge cases.

MutScan | Current Research - Manuscript in Preparation, 2026

- Developing a reproducibility-aware model-audit and evidence-integration framework across 15 disease-associated human proteins, comparing zero-shot missense scores with deterministic and chemistry-based signals, curated variant evidence, and structural context for research evaluation rather than clinical classification.
- Verified bit-deterministic outputs across 84 replicate runs and generated per-run provenance manifests linking inputs, outputs, and execution environments to support regression testing and release decisions.

Computational Genomics & Industry Applications

- Build reproducible workflows for genome and transcriptome assembly, expression quantification, comparative genomics, and multiomics across Illumina, PacBio, Ion Torrent, bulk, and single-cell assays.
- Completed an industry-related bioinformatics project examining how ribosomal-protein expression affects human protein production in large-scale fermentation.
- Co-authoring a 2026 manuscript in preparation on changes in ribosomal transcription in HEK293 suspension cells infected by recombinant AAV under cGMP manufacturing conditions.

PROFESSIONAL EXPERIENCE

Texas A&M University, Department of Biology | College Station, Texas

Associate Professor | Current faculty appointment

- Lead an interdisciplinary research program spanning molecular genetics, computational genomics, protein and RNA biology, variant interpretation, genome engineering, and reproducible scientific software.
- Translate complex biological questions into experimental designs, computational representations, functional code, benchmark plans, and rigorously tested deliverables.
- Supervise and mentor postdoctoral researchers, graduate students, and undergraduates; developed Computational Genomics, Genomics, Digital Biology, Advanced Eukaryotic Genetics and Epigenetics, and Information in Biology courses.
- Release reusable genomics software and research artifacts through GitHub and Zenodo, including tools for sequence processing, FASTA/GFF3 reconciliation, transcriptional profiling, visualization, and genome assembly.

Selected Biotechnology Consulting & Advisory Engagements | Intermittent

Selected engagements: Base10, Advisor | F2G, Advisory Board | Genencor, Consultant | Schering-Plough, Consultant | Scriptgen, Consultant

- Advised biotechnology and pharmaceutical organizations on genetics and genomics; designed industry-related experiments and evaluated results.

Earlier Computational Biology & Biotechnology Foundations

- Developed nucleic-acid and protein-analysis software, biotechnology and agriculture optimization algorithms, laboratory-automation approaches, and biological database systems at CENARGEN / EMBRAPA; completed advanced genetics research appointments at the University of Wisconsin-Madison and Stanford University.

SELECTED SCIENTIFIC & SOFTWARE IMPACT

- **MutScan** - Reproducibility-aware model-audit and evidence-integration framework across 15 disease-associated proteins; manuscript in preparation, 2026.
- **Proteins as Dynamic Networks** - Sequence, structure, allostery, and molecular motion; manuscript in preparation with a [public research record](#), 2026.
- **Microbial genomics** - Co-authored draft genome analysis of three *Spiroplasma* strains associated with *Drosophila*, *Microbial Genomics*, 2025 ([DOI](#)); co-authored the de novo assembly and annotation of the complete *Myxococcus xanthus* DZ2 genome, 2022 ([DOI](#)).
- **Foundational genetics and genomics** - Contributed to the *Neurospora crassa* genome in *Nature* ([DOI](#)); discovered meiotic transvection in fungi in *Cell* ([DOI](#)).
- **Open scientific software** - Publicly released genomics and visualization tools through [GitHub](#) and [Zenodo](#).

LEADERSHIP & SCIENTIFIC SERVICE

Editor, *PeerJ* (2018-Present) | Section Editor, Genetics and Genomics, *PLOS ONE* (2009-2012) | Extensive postdoctoral, graduate, and undergraduate mentoring

EDUCATION & TRAINING

PhD, Genetics - University of Georgia

MS and BS, Molecular Biology - University of Brasilia

Advanced research training in Genetics - University of Wisconsin-Madison and Stanford University

LANGUAGES

Spanish - Advanced | Portuguese - Advanced