

Shriram Bhat

sb65 [at] illinois [dot] edu — github.com/artorias111 — ttaggc.vercel.app

EDUCATION

University of Illinois Urbana-Champaign · Urbana, IL

2022 – Present

Ph.D. Candidate, Department of Evolution, Ecology and Behavior · Advisor: Chris Cheng DeVries

Expected July 2027

Birla Institute of Technology and Science (BITS) Pilani · Hyderabad, India

2017 – 2022

B.Eng. Mechanical Engineering and M.Sc. Biological Sciences (Dual Degree)

RESEARCH EXPERIENCE

Graduate Researcher, Cheng Lab, University of Illinois Urbana-Champaign

2022 – Present

Comparative genomics of Arctic gadids: chromosome-scale assembly and population genomics of species boundaries

- Built chromosome-scale reference genomes for multiple Arctic gadid species from PacBio HiFi and Hi-C data, resolving highly heterozygous and repeat-dense loci in genomes with no close reference.
- Paired those assemblies with population-genomic analysis of low-coverage whole-genome data to test species boundaries within *Gadus*, where morphology and existing taxonomy disagree.
- Wrote the lab's assembly and annotation infrastructure as reusable Nextflow pipelines (assembly, haplotig purging, contamination screening, transposable-element and gene annotation, mitochondrial genome assembly) orchestrating 10+ standard tools; applied to 19 genome assemblies to date, 9 of them scaffolded to chromosome level. Now the default workflow for 4 researchers across two labs, and released open-source.
- Reduced assembly turnaround from ~1 week of hand-run scripts to ~2 days, and annotation from up to 2 weeks of per-tool debugging to a predictable ~4 days.
- Used k -mer spectra to characterize heterozygosity and repeat content up front, and to choose assembly strategies for organisms where standard parameter defaults fail.

Evolutionary loss of the antifreeze glycoprotein (AFGP) trait in notothenioid fishes

- Traced the genomic loss of the AFGP locus across 6 notothenioid species (5 assembled and annotated in house, 1 from NCBI), resolving the structural mechanism underlying loss of the trait: large-scale deletion with associated transposable-element activity.
- Annotated and compared assemblies of varying quality to distinguish true structural variation from assembly artifact, a recurring failure mode in non-model comparative genomics.

Bioinformatics Collaborator & Undergraduate Mentor, University of Illinois Urbana-Champaign

2024 – Present

- Facilitated a collaborating research group with bioinformatics analyses, including genome assembly, annotation, and computational workflow development.
- Mentored undergraduate researchers from biochemistry and mathematics backgrounds, teaching Linux/HPC workflows, computational genomics, and reproducible data analysis.

Undergraduate Researcher, Andersen Lab, Northwestern University

Aug 2021 – May 2022

*Natural variation in telomere length across *Caenorhabditis* wild strains*

- Identified orthologs of telomere-complex proteins across *C. elegans*, *C. briggsae*, and *C. tropicalis* to test whether telomere-length variation tracks sequence divergence in the maintenance machinery.
- Processed population-scale genomic datasets on the Quest HPC cluster using R (Tidyverse) and shell pipelines.

PUBLICATIONS & MANUSCRIPTS

- York J. M., **Bhat S.**, Kim J., Cardenas L., Cheng C.-H. C. Chromosomal genome assemblies of Antarctic *Notothenia coriiceps* and temperate relative *Paranotothenia angustata*. *bioRxiv* (2025). doi:10.1101/2025.10.11.681823. *Under review*.
- **Bhat S.**, et al. Structural genomic differences and mitonuclear discordance at a contested species boundary: Greenland cod and Pacific cod. *Manuscript in preparation*.

• **Bhat S.**, et al. Chromosome-scale genome assemblies of the Antarctic toothfish *Dissostichus mawsoni* and the Patagonian toothfish *D. eleginoides*. *Manuscript in preparation*.

SOFTWARE & OPEN SOURCE

Repositories at github.com/artorias111

- **nf-core/scrnaseq**: contributor to the community single-cell pipeline. Also active in nf-core/rare-disease.
- **FeatureFlow**: Nextflow pipeline for genome feature annotation, integrating EarlGrey (transposable elements), BRAKER (gene models) and ncRNA detection methods into one reproducible run.
- **simple-asm**: Nextflow pipeline for HiFi+Hi-C genome assembly with QC and contamination screening built in.
- **pg4Findr**: Rust tool for G-quadruplex motif detection in sequencing reads and genome assemblies.

TALKS

Comparative genomics to decipher evolutionary loss of the AFGP genotype in notothenioid fishes at the EEB Colloquium, UIUC	2025
Comparative genomics to decipher the evolution of the antifreeze glycoprotein genotype in notothenioids. International Congress on the Biology of Fish	2024
Genomic basis of antifreeze trait divergence. Midwest Ecology and Evolution Conference	2023

AWARDS & FELLOWSHIPS

University Block Grant Fellowship, University of Illinois Urbana-Champaign	2022
SIB Summer Research Award: resolving gadid genome assemblies with DNA language models (<i>awarded; not pursued</i>)	2025
Odum-Kendeigh Summer Award, Dept. of Evolution, Ecology and Behavior, UIUC	2024
Thomas Frazzetta Award for Outstanding Teacher in EEB, UIUC	2025
List of Teachers Ranked as Excellent by Their Students, UIUC	2023, 2024, 2026
EEB Summer Stipend Award, Dept. of Evolution, Ecology and Behavior, UIUC	2023

TEACHING & SERVICE

Course Assistant , Computational Genomics for Clinicians (Mayo Clinic / Illinois Alliance)	Summer 2025
• Supported a week-long computational genomics course for clinicians and wet-lab scientists by moderating hands-on labs, troubleshooting analyses, and answering questions; set up the Jupyter infrastructure and authored the course manual.	
Graduate Teaching Assistant , University of Illinois Urbana-Champaign	2023 – Present
• Programming for Genomics, Organismal Biology, Genetics, Evolution, and Physiology; taught for 6 semesters.	
Peer Reviewer	2024
• Pitman et al., “The genome sequence of the European chub, <i>Squalius cephalus</i> .” <i>Wellcome Open Research</i> .	

TECHNICAL SKILLS

Languages	Python, Rust, R (tidyverse), Bash/Zsh
Genomics	Genome assembly (PacBio HiFi, Hi-C scaffolding), genome annotation, variant calling, bulk RNA-seq (differential expression), population genetics (selection scans, PCA), transposable element analysis, <i>k</i> -mer analysis, Phylogenetic analysis
ML & Data	PyTorch, scikit-learn, statsmodels, pandas, introductory work with genomic language-model embeddings
Tools	Nextflow, samtools, bcftools, minimap2, bwa-mem2, seqkit, Git, Docker, Conda/Pixi, SLURM, coding agents, L ^A T _E X