BRHANU FENTAW ZNABU

25674436@nebraska.edu | brhanufenbme@gmail.com | 531-254-0990 | Lincoln, NE

linkedin.com/in/brhanu-fentaw-znabu-23115619a | github.com/brhanufen

EDUCATION

Ph.D. Biomedical Engineering | *University of Nebraska–Lincoln*   Aug 2024 – Present
- Advisor: Nicole R. Sexton, PhD | Nebraska Center for Virology
- Focus: Machine learning and genomic foundation models for viral host-range prediction and epidemic emergence

M.Sc. Biomedical Science and Engineering | *GIST, South Korea*   2021 – 2024
- GPA: 4.0/4.5 | Advisor: Euiheon Chung, PhD
- Thesis: Deep behavioral phenotyping of diabetic neuropathy mouse model using open-field test

B.Sc. Biomedical Engineering | *Jimma University, Ethiopia*   2014 – 2018
- GPA: 3.64/4.0 | Graduated with Distinction (Ranked 4th of 180)

RESEARCH EXPERIENCE

Graduate Research Assistant | *Sexton Lab, University of Nebraska–Lincoln*   Aug 2024 – Present
Nebraska Center for Virology | School of Biological Sciences
- Developed ArboFM, a domain-adapted genomic foundation model (continued MLM pre-training of DNABERT-2 on 120K+ arbovirus genome windows) for predicting epidemic emergence across five RNA virus families (9,299 genomes, 362 species). Achieved AP = 0.978 on Flaviviridae under species-grouped evaluation. [In preparation for Bioinformatics, Oxford]
- Designed a leakage-aware evaluation framework using DNABERT-2 embeddings on 3,031 flavivirus genomes with windowed embedding strategy (1,000 nt windows, 500 nt stride). Quantified a 15-percentage-point performance inflation from phylogenetic leakage via species-grouped vs. accession-grouped cross-validation. Identified UpA-containing motifs linked to the OAS/RNaseL innate immunity pathway through genome-localized attribution and 6-mer enrichment analysis. [Submitted to PLOS Computational Biology]
- Extended the leakage-aware framework to a cross-family comparison across four arbovirus families (Flaviviridae, Togaviridae, Peribunyaviridae, Rhabdoviridae; 6,052 genomes, 124,883 windows). Built a 4×4 cross-family transfer learning matrix demonstrating that host-switching grammar is predominantly family-specific (Jaccard ≈ 0), supporting convergent evolution of immune evasion strategies. [In preparation for Virus Evolution, Oxford]
- Built species-stratified ML classifiers (logistic regression, random forest) on 1,285 curated flavivirus genomes with 97 compositional features (dinucleotide O/E ratios, RSCU codon usage). Achieved PR-AUC = 1.000 for ISFV vs. dual-host classification. Identified CpG suppression as the primary discriminative feature. [In preparation for Journal of Virology]
- Performed retrospective temporal validation of ArboFM showing it could detect Zika, chikungunya, and West Nile virus epidemic lineages before their documented emergence events, including within-species lineage discrimination (WNV Lineage 1a vs. Lineage 2, CLES = 0.898).
- Conducted all computational analyses on NVIDIA V100 GPUs (Holland Computing Center HPC), including domain-adaptive pre-training, embedding generation (120K+ windows), gradient × input attribution, and large-scale cross-validation pipelines.

Research Assistant | *Neurophotonics Lab, GIST*   Mar 2021 – Jun 2024
- Developed unsupervised autoregressive and Hidden Markov models to identify temporal dependence and transitions between behavioral states in a diabetic neuropathy mouse model.
- Benchmarked deep learning-based 3D pose estimation (DeepLabCut) against traditional 2D behavioral analysis using logistic regression.

PUBLICATIONS

1. Znabu BF, Sexton NR. **A genomic foundation model for predicting arbovirus epidemic emergence across RNA virus families.** *Bioinformatics* (In preparation)
2. Znabu BF, Sexton NR. **Deciphering the host-switching grammar of flaviviruses using foundation model embeddings: a leakage-aware evaluation framework.** *PLOS Computational Biology* (Submitted)
3. Znabu BF, Sexton NR. **Cross-family comparison of host-switching grammar in arthropod-borne viruses reveals convergent evolution of family-specific immune evasion strategies.** *Virus Evolution* (In preparation)
4. Znabu BF, Sexton NR. **Genome-scale classification of flavivirus host range using composition and codon-usage signatures.** *Journal of Virology* (In preparation)
5. Ashiquzzaman A*, Lee E*, Znabu BF*, Sakib AN, Chung G, Kim SS, Kim YR, Kwon H-S, Chung E. **MoSeq based 3D behavioral profiling uncovers neuropathic behavior changes in diabetic mouse model.** *Scientific Reports 15, 15114* (Published 2025)

CONFERENCE PRESENTATIONS

1. Znabu BF, Sexton NR. **Deciphering the host-switching grammar of flaviviruses using foundation model embeddings: a leakage-aware evaluation framework.** *Annual Engineering Symposium, University of Nebraska–Lincoln* (Poster, Feb 2026)
2. Znabu BF, Sexton NR. **Cracking the viral code: codon usage and CpG patterns predict host specificity.** *Nebraska Center for Virology Annual Symposium* (Poster, Oct 2025)

TECHNICAL SKILLS

Programming: Python (primary), R, C++, MATLAB

ML / Deep Learning: PyTorch, scikit-learn, Hugging Face Transformers, DNABERT-2, Nucleotide Transformer

Foundation Models: Domain-adaptive pre-training (masked language modeling), embedding extraction, transfer learning, gradient-based attribution and interpretability

ML Methods: Logistic regression, random forest, SVM, MLP classifiers, PCA, UMAP, t-SNE, cross-validation design (species-stratified, temporal splits), ablation studies

Biological Data: Viral genomics (NCBI GenBank/Virus), TCGA RNA-seq, Biopython, codon usage analysis, dinucleotide composition, sequence curation and taxonomic harmonization

Infrastructure: GPU computing (NVIDIA V100), HPC clusters (Slurm), CD-HIT, Git/GitHub

Data Science: NumPy, pandas, SciPy, matplotlib, seaborn, survival analysis (Kaplan–Meier, Cox PH)

SELECTED COURSE PROJECTS

Computational Methods in Bioinformatics (CSCE 871) | *UNL*   Fall 2025
- Validated a single-cell-derived hypoxia/glycolysis gene signature in 576 TCGA-LUAD bulk RNA-seq samples; demonstrated significant survival association (HR = 1.55, p = 0.002) and enrichment in Squamoid expression subtype (ANOVA p = 2.3×10^{-6}).
- Built SVM classifiers (linear and RBF kernels) with grid search and feature selection (SelectKBest) on the Duke breast cancer gene expression dataset (7,129 features); implemented two-stage normalization and PCA-based dimensionality reduction.

TEACHING EXPERIENCE

Lecturer | *Hawassa University, Ethiopia*   Aug 2019 – Jan 2021
- Taught Biomechanics, Biomaterials, and Biophysics to undergraduate biomedical engineering students.

AWARDS & GRANTS

- Korean Government Scholarship for Master's Study, GIST (2020)
- Research Grant for Colostomy Device Development, Hawassa University (2020)
- Best B.Sc. Thesis Award, Ethiopian Science, Technology, and Innovation (2018)
- Graduated with Distinction, Ranked 4th of 180, Jimma University (2018)