

Cory Gardner

PHD CANDIDATE • DEPARTMENT OF COMPUTER SCIENCE

Saint Louis University, St. Louis, MO 63103

(314) 717-7638 | cory.gardner@slu.edu

Education

PhD Candidate in Computer Science, Saint Louis University (St. Louis, MO)

2022 – Present

- Advisor: Dr. Tae-Hyuk Ahn
- Dissertation Focus: Multimodal Deep Learning and HPC for Scalable Metagenomics

Open Science Grid (OSG) School, University of Wisconsin (Madison, WI)

2024

MS in Bioinformatics & Computational Biology, Saint Louis University (St. Louis, MO)

2020 – 2021

BS in Sociology, Colorado State University (Pueblo, CO)

2014 – 2017

Research Interests

Computational Biology & Genomics in general, and in particular:

- Advanced sequence representation methods for genomic data analysis
- Multimodal deep learning approaches integrating genomic, transcriptomic, and phenotypic data
- Applications of deep learning to biological sequence understanding
- Development of efficient computational methods for metagenomic data processing, including taxonomic profiling and classification

High-Performance Computing & Parallel Processing in general, and in particular:

- GPU-accelerated image analysis for medical and biological applications
- Parallel algorithm optimization for processing large-scale genomic datasets
- Distributed computing strategies for computationally intensive bioinformatics tasks
- Memory-efficient implementations for handling terabyte-scale biological data

Machine Learning Privacy & Security in general, and in particular:

- Membership and property inference attack detection in biomedical ML
- Federated learning approaches for healthcare data privacy
- Privacy-preserving data fusion for multimodal medical data
- Adversarial model testing and vulnerability discovery frameworks

Academic/Professional Experience

Graduate Research Assistant – Ahn Lab

Saint Louis University, 2021 – Present

- CT-image analysis (GPU-based CUDA acceleration of CT images for quality control)
- Metagenomic Analysis (Critical Assessment of Massive Data Analysis)
- RNA-seq (Novel Candidate Genes Differentially Expressed in Glyphosate-Treated Horseweed)
- Long read *de-novo* assembly and annotation (First haplotype-resolved assembly of *Saccharomyces bayanus*)
- Computational Privacy (Privacy and Inference Attacks on Multimodal Machine Learning Models)
- T-Cell analysis (sequence diversity for pathogen exposure prediction)
- Anomaly detection (Novel methods for anomaly detection in images for improved efficiency and accuracy)
- ChipSeq (CTCF regulates the long gene expression program in neurons)
- Federated learning for improved privacy in medical applications

- Data Structures
- Distributed Computing
- Concurrent and Parallel Programming
- Introduction to Bioinformatics

Supplemental Instructor

Saint Louis University, 2024

- Introduction to Bioinformatics II

Publications and Presentations

Sirasani JP, **Gardner C**, Jung G, Lee H, and Ahn TH. "Bioinformatics approaches of blood and tissue microbiome analyses: challenges and perspectives." *Briefings in Bioinformatics* (2025). DOI: 10.1093/bib/bbaf176

Gardner C, et al. "Chromosome-level subgenome-aware de novo assembly of *Saccharomyces bayanus* provides insight into genome divergence after hybridization." *Genome Research* (2024). DOI: 10.1101/gr.279364.124

Gardner C, "Long-read Genome Assembly," *Lightning Talk, Donald Danforth Plant Science Center*, St. Louis, MO (2023).

Yang Y, **Gardner C**, Gupta P, Peng Y, Piasecki C, Millwood RJ, Ahn TH, Stewart CN Jr. "Novel Candidate Genes Differentially Expressed in Glyphosate-Treated Horseweed," *Genes (Basel)* (2021). DOI: 10.3390/genes12101616

Gardner C, Ahmad S, Ahn TH. "Metagenomic Data Analysis with Probability-Based Reduced Dataset Representation," *Presentation at CAMDA Conference, ISMB* (2020).

Gardner C, Min B, Ahn TH. "PaDiM-Wavelet: Memory-Efficient Anomaly Detection for Heterogeneous Computing Systems," *Poster Presentation, GCASR 2025*, Loyola University. Chicago, IL (Upcoming, 2025).

Research Projects**Multimodal Deep Learning and HPC for Scalable Metagenomics — 2024 – 2026**

This project develops a scalable multimodal deep learning framework integrating diverse metagenomic data, including sequencing reads, taxonomic profiles, and environmental metadata, to enhance accuracy and interpretability in microbiome analyses. This research addresses memory-intensive assembly and CPU-heavy profiling bottlenecks in terabyte-scale datasets, enabling high-throughput, accurate microbial analyses with significant impacts on diagnostics, therapeutics, and environmental microbiology.

- **My Role:** Lead researcher; responsible for pipeline design, multimodal representation learning, GPU optimization, HPC task distribution, and benchmarking.
- **Required Skills:** Multimodal deep learning, high-performance computing (MPI, CUDA, distributed computing), metagenomic analysis (OMEGA, Kraken2, HUMAnN3), large-scale data handling and optimization.

Memory-Efficient Anomaly Detection for Heterogeneous Systems — 2024 – Present

This project focuses on developing novel methods for image anomaly detection tailored for memory efficiency and deployment on heterogeneous computing systems (e.g., CPU/GPU). The research explores techniques such as wavelet-enhanced Patch Distribution Modeling (PaDiM-Wavelet) to enhance accuracy and performance, particularly under resource constraints inherent in complex computing environments.

- **My Role:** Developing and implementing novel anomaly detection algorithms (e.g., PaDiM-Wavelet); optimizing implementations for memory efficiency and parallel execution on heterogeneous platforms; evaluating model performance and accuracy; preparing results for publication/presentation (GCASR 2025 poster).
- **Required Skills:** Anomaly detection algorithms, image processing, wavelet transforms, performance optimization, algorithm design and evaluation.

De Novo Haplotype-Level Assembly of Hybrid Yeast — 2021 – 2024

The accurate assembly of hybrid genomes presents unique computational challenges due to the presence of multiple subgenomes with varying levels of divergence. This project focused on generating a chromosome-level subgenome-aware *de novo* assembly of the *Saccharomyces bayanus* yeast genome to study post-hybridization genome evolution. The resulting assembly revealed novel insights into subgenome divergence following hybridization events.

- **My Role:** Designed and executed a comprehensive assembly pipeline including read quality control, *de novo* subgenome-aware assembly, and sequence polishing; developed methods for subgenome discrimination and parental haplotype attribution; validated assembly results via phylogenetic analysis; published findings in *Genome Research*.
- **Required Skills:** Oxford Nanopore long-read sequencing data analysis, genome assembly tools, high-performance computing, bioinformatics workflow design and implementation.

Privacy-Preserving Machine Learning for Biomedical Data — 2024 – Present

Medical data present challenges for machine learning applications due to strict privacy requirements. This project addresses membership and property inference attacks (MIA/PIA) against machine learning models trained on biomedical data, where adversaries may extract private information about training data or individuals. The work focuses on developing privacy-preserving techniques that maintain accuracy while protecting sensitive information.

- **My Role:** Investigating differential susceptibility to privacy attacks between multimodal models and the unimodal models on which they are based.
- **Required Skills:** Privacy-preserving machine learning techniques, deep learning for medical imaging, adversarial machine learning and attack simulation.

GPU-Accelerated CT Image Analysis — June 2020 – August 2021

Quality control of imaging data, particularly CT scans, is computationally demanding and often traditionally performed on CPU architectures with significant processing time requirements. This research collaboration with Xavis focused on developing optimized CUDA implementations for image preprocessing, image quality assessment, and CT rotation to achieve substantial performance improvements while maintaining accuracy.

- **My Role:** Implemented parallel algorithms for CT image analysis tasks; designed CUDA kernels for noise level estimation, artifact detection, and contrast evaluation; integrated accelerated components into existing workflows.
- **Required Skills:** CUDA programming and GPU architecture knowledge, medical image processing techniques, parallel algorithm design and optimization, performance benchmarking and validation.

RNA-seq Analysis of Glyphosate-Treated Horseweed — 2020 – 2021

Identified novel candidate genes involved in glyphosate resistance via RNA-seq differential expression analysis, offering insights into resistance mechanisms in agricultural weeds.

- **My Role:** Computational analyses including RNA-seq preprocessing, differential expression analysis, and candidate gene identification.
- **Required Skills:** RNA-seq pipeline (edgeR, DESeq2), statistical analysis, computational biology methods.

Metagenomic Analysis & Sample Origin Prediction — September 2019 – August 2020

Environmental metagenomic samples contain complex, interacting microbial communities that can serve as biological signatures of their origin location. This project focused on developing computational methods for analyzing low-biomass microbial samples and predicting their geographical origins using machine learning techniques applied to taxonomic profiles generated through next-generation sequencing.

- **My Role:** Implemented taxonomic classification pipeline; designed and trained ensemble machine learning models for geolocation prediction; presented the work at CAMDA 2020, receiving the Best Presentation Award.
- **Required Skills:** Metagenomic data analysis and taxonomic classification, machine learning for biological data, high-performance computing for large-scale datasets, scientific communication and visualization.

Awards & Honors

BITWISE Scholarship

2020 – 2021

National Science Foundation — Merit-based scholarship for research in bioinformatics and high-performance computing

Conference Fellowship

2020

International Society of Computational Biology

Best Presentation Award

2020

CAMDA International Conference — Recognized for outstanding research presentation on metagenomic sample classification

Skills

- **Programming Languages:** C++, Python, R, CUDA
- **HPC & Parallel Computing:** MPI, CUDA, Slurm, HTC and Open Science Grid, Distributed Computing, GPU Optimization
- **Frameworks & Technologies:** PyTorch, TensorFlow, Bioinformatics Tools (OMEGA, Kraken2, HUMAnN3)
- **Analysis Skills:** Statistical Modeling, Data Visualization, Multimodal Deep Learning
- **Soft Skills:** Technical Writing, Research Presentation, Mentoring and Teaching, Cross-disciplinary Collaboration