Aymen Maqsood Mulbagal

Boston, MA | (857) 296-2150 | aymenmaqsood.2000@gmail.com | LinkedIn | github

---

EDUCATION

Master of Science in Bioinformatics   **December 2024**

Northeastern University, Boston, MA

Coursework: Bioinformatics Computational Methods, Statistics for Bioinformatics, Machine Learning and Data Mining, Bioinformatics Programming, Genomics in Bioinformatics, Biotech Enterprise and Visual Analytics

Leadership: Head of Bioinformatics, Graduate Biotech and Bioinformatics Association

Bachelor of Engineering in Biotechnology

Ramaiah Institute of Technology, Bangalore, India

Coursework: Cell and Molecular Biology, Immunology, Structural Biology, Genomics, Proteomics, Bioinformatics, Biostatistics, Bio Modeling, Biochemistry, Medical Biotechnology and Bioprocess Engineering

---

WORK EXPERIENCE

Quantitative Biology Co-op   **July 2024 – December 2024**

Vor Bio, Boston, MA

- Designed and deployed Nextflow pipelines for CRISPR-edited NGS data , integrating AWS EC2 to improve scalability and reduce analysis time
- Enhanced in-house tools for off-target analysis and performed comparative evaluations against open-source platforms, delivering structured reports and insights for decision-making
- Executed single-cell genotyping and immunophenotyping on CD33-edited HSPCs using Mission Bio Tapestri data, analyzing thousands of cells to quantify editing efficiency and CD33 loss
- Generated detailed experiment summaries, visualizations, and statistical reports while managing code versioning with Bitbucket and tracking tasks through Jira to ensure efficient and data-driven project execution
- Created interactive R Shiny applications to streamline data preprocessing tasks for bench scientists, enabling efficient handling of experimental data

Research Associate   **January 2024 – June 2024**

Institute for Experiential AI, Northeastern University

- Built a preprocessing pipeline for eCLIP data incorporating demultiplexing, adapter trimming, sequence alignment, and PCR duplicate removal for RNA-Binding Proteins (RBP) binding site analysis utilizing SLURM in HPC
- Detected and analyzed RNA-binding protein peaks using a python-based tool within a Snakemake pipeline

Research Intern   **February 2022 – June 2022**

Cancer Systems Biology LAB, BSSE, Indian Institute of Science Bangalore, India

- Performed RNA-Seq analysis using Bioconductor in R on a Linux HPC cluster, processing over 150 cancer cell line samples to explore EMT dynamics
- Identified key gene targets through GSEA and constructed a regulatory network involving miR-200, SNAIL, ZEB-1, and SLUG, advancing insights into EMT-related gene regulation in oncology

---

PROJECTS

Single-Cell RNA-Seq Analysis of COVID-19 Immune Response

- Investigated immune cell composition and gene expression using scRNA-seq of PBMCs from patients and healthy controls
- Integrated datasets with Scanorama for batch correction and performed gseapy pathway analysis after downstream analysis
- Uncovered dysregulated immune pathways such as TNF and IL-6 signaling pathways

Single-Cell ATAC-Seq Analysis of Hematopoietic Stem Cells

- Utilized Signac and Seurat in R for downstream analysis of single-cell ATAC-seq data from young and aged HSCs
- Study Revealed differential chromatin accessibility peaks associated with aging, revealing key epigenetic changes in HSCs
- Merged and examined multi-sample datasets to contrast chromatin landscapes in young versus aged HSCs, exposing age-associated regulatory shifts

GlycoEnzyme Expression Atlas (Bio-IT Hackathon)

- Designed and implemented an end-to-end pipeline (Python/R) to curate 906 glycoenzyme genes and perform differential expression (DESeq2) on two muscular dystrophy RNA-seq datasets, focusing on glycosylation-related dysregulation.
- Performed KEGG enrichment (clusterProfiler) and Cytoscape network analyses to identify key glycoenzyme-driven pathway disruptions—highlighting GAG biosynthesis and degradation as critical in muscular dystrophy.
- Developed an interactive R Shiny atlas that enables researchers to query glycoenzyme expression, visualize DE results (volcano plots/heatmaps), and explore STRING-based interaction networks; all code and documentation are open-source on GitHub.

TECHNICAL SKILLS

Languages and Frameworks: Python (NumPy, Pandas, PyTorch, Scanpy, Seaborn), R (Tidyverse, ggplot2, Bioconductor, Plotly)
Bioinformatics: NGS analysis (WGS, RNA-seq, Variant Calling with GATK), Single-Cell Analysis (Scanpy, Seurat), Genomics, Version Control (Git), Bioinformatics databases (TCGA, UNIPROT, ENCODE, gTEx)
Bioinformatics Tools: BLAST, ISAAC4, STAR, Trimmomatic, FastQC, Cutadapt, UMI-tools, Bowtie, BWA, Picard, SAMtools, Bedtools, DESeq2, Trinity, IGV, UCSC Genome Browser
Tools and Technologies: HPC (SLURM), Workflows (Snakemake, Nextflow), Cloud (AWS EC2, S3, Docker)

Publications

Qiu, H., Wang, R., Wang, S., Mulbagal, A. M., Anderson, E., & Ge, H. G. (2025). *SANTON: Sequencing Analysis Toolkits for Off-target Nomination*. bioRxiv. https://doi.org/10.1101/2025.05.09.653082