

Mahima Mahabaleshwar Siddheshwar

(317) 506-5008 | mahima.siddheshwar@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

Bioinformatics Scientist with 3+ years of experience designing, validating, and implementing NGS analysis pipelines for R&D and diagnostic projects. Skilled in developing reproducible omics workflows using Python, R, and Bash in Unix-based environments, with proficiency in statistical analyses for RNA-seq, WES, and variant calling. Hands-on collaborator with wet-lab teams on RNA extraction and NGS library preparation, ensuring compliance under GxP standards.

TECHNICAL SKILLS

- **Genomic Analysis & Pipelines:** scRNA-seq, Seurat, Scanpy, Bulk RNA-Seq, DESeq2, edgeR, limma, Multi-omics Integration, WGCNA, Variant Calling, GATK, Pathway Enrichment, GSEA, KEGG, Reactome, Spatial Transcriptomics, Omics Data, NGS Analysis Pipelines
- **Bioinformatics Tools & Resources:** STAR, BWA, SAMtools, BEDTools, ANNOVAR, Ensembl, GEO, TCGA, UCSC Genome Browser, Cytoscape, IGV, Jira
- **Programming & Data Science:** Python, R, Bash/Shell Scripting, SQL, Git Version Control, Scientific Programming, Workflow Development Languages
- **Machine Learning & Statistics:** PCA, t-SNE, UMAP, K-means, Hierarchical Clustering, LASSO Regression, SVM, Model Evaluation
- **Cloud, HPC & Infrastructure:** HPC Cluster Computing, SLURM, AWS, Google Cloud Platform, GCP, Unix/Linux Workflow Automation
- **Data Visualization & Reporting:** Volcano Plots, Heatmaps, Interactive Dashboards, Plotly, Power BI, Tableau
- **Wet Lab & Quality Control:** Cell Culture, DNA/RNA extraction, PCR, SDS-PAGE, FACS-based workflows, NGS library preparation
- **Clinical, Quality & Regulatory:** HIPAA, GxP/GLP, ALCOA+, FAIR Principles, LIMS, FDA Documentation & Audit Readiness

PROFESSIONAL EXPERIENCE

Elucidata Corporation

Jul 2025 - Oct 2025

Bioinformatics Analyst Intern – R&D

USA

- Optimized AI-assisted NGS analysis pipelines by ~30% via prompt refinement and LLM evaluation, improving biological reasoning for DEG and pathway analysis.
- Evaluated genomic workflow frameworks across 100+ curated research scenarios, ensuring biologically consistent and reproducible outputs before release.
- Performed LLM evaluation and prompt engineering using Python scientific programming, strengthening biological reasoning for KG, bulk/scRNA-seq, Synthea, and DEG tasks by 25–40%.
- Verified Synthea (synthetic health data) pipelines on cloud compute by confirming disease-specific cohort accuracy across 50+ repeatable test runs.
- Strengthened end-user research usability by stress-testing AI/ML agents across 100+ biological queries, improving genomics output quality & KG-driven insights.

IU BioHealth Informatics Research Center

Sep 2023 - May 2025

Research Assistant – Global Health Data Analytics

USA

- Integrated team efforts to optimize RNA-seq preprocessing on HPC SLURM, cutting runtime by 40% using a Snakemake/Nextflow workflow development languages framework.
- Engineered reproducible NGS analysis pipelines using STAR, SAMtools, and BEDTools, generating high-quality, fully traceable gene-level feature matrices.
- Led QC & preprocessing automation using Bash scripting and SLURM batch jobs, applying scientific programming best practices to improve pipeline stability by 25% & reduce manual team effort.
- Created pathway and network visualizations in Cytoscape, translating enrichment results into interpretable, biologically meaningful findings.

Maxgen Technologies (Client: Hetero Pharma)

Jul 2021 - Aug 2023

Bioinformatics Specialist

India

- Supported scRNA-seq clustering, annotation & biomarker analysis, improving insights by 25% and contributing to assay and preclinical decisions.
- Executed SOP approved RNA-seq & variant workflows under GxP/GLP/ALCOA+, ensuring 100% reproducibility & reducing gaps by 40%.
- Collaborated with R&D, engineering, and IT data teams to resolve pipeline issues, boosting workflow scalability by 20% and accelerating review cycles.
- Collaborated hands-on with wet-lab scientists on PCR-based RNA extraction, NGS library preparation, and assay interpretation, delivering analysis reports that aligned with experimental goals and facilitated on-time project delivery

EDUCATION

Luddy School of Informatics, Indiana University Indianapolis

Aug 2023 - May 2025

Master of Science, Bioinformatics

Indiana, USA

Sapthagiri College of Engineering (Affiliated to VTU)

Aug 2016 - Aug 2020

Bachelor of Technology, Biotechnology

Bangalore, India

ACADEMIC PROJECTS

Machine Learning Pipeline for Breast Cancer Biomarker Discovery

- Built a scalable ML framework to harmonize 7 GEO cohorts (~814 samples) using SVA/Limma, reducing batch-driven technical noise.
- Benchmarked five ML models (SVM, RF, ANN, LR, GBM) with stratified 80/20 validation, achieving 98.9% accuracy and AUROC >0.99.
- Validated a 28-gene signature on external datasets, identifying high-confidence biomarkers (POSTN, LSR) with $p < 1e-31$ for qPCR validation.

WGCNA on COVID-19 and RSV Transcriptomes

- Implemented an end-to-end RNA-seq pipeline (FastQC, HISAT2, featureCounts) on HPC, achieving >89% alignment across 68 samples.
- Applied WGCNA to identify key gene modules, showing strong correlation between MEyellow and COVID-19 severity ($r = 0.89$).
- Identified three shared antiviral hub genes (OASL, TXN, RBCK1) across COVID-19 and RSV datasets for downstream functional validation.

Streamlining RNA-Seq Data Analysis Framework

- Architected an automated RNA-seq processing & visualization framework on HPC, ensuring data integrity across 68 samples with >89% alignment rate.
- Profiled four pipelines for efficiency and quality, identifying t-SNE + Hierarchical Clustering as optimal (ARI = 0.60) with biological label agreement.
- Reduced computational runtime by 35% and applied the optimized framework to DEA, identifying 14,667 significant COVID-19 genes ($p < 0.05$).

CERTIFICATIONS

- **Biosciences Program Certification - Biocon Academy × KGI (2021):** Industry training in biologics R&D, manufacturing, QC/QA, RA with hands-on lab, fermentation, mammalian cell culture, and drug-development projects.