NITIN SHUKLA

nitinshukla1912@gmail.com · (+91)-7984243353

Gandhinagar, Gujarat 382010

SUMMARY

A research fellow working in the field of genomics and bioinformatics with 3+ years of work experience in handling diverse amounts of biological data. Competent in working with HPC computing environment. Proficient in understanding the depth of NGS data and its applications to understand complex metagenomic and transcriptomic landscapes. Experienced in disease biology, genetics, and molecular biology, with a keen enthusiasm for continual learning and collaborative engagement within the scientific community. Adept in staying abreast of advancements in bioinformatics research.

TECHNICAL PROFICIENCIES

Languages: R, Bash

Version Control: Git, GitHub

High-throughput Data & Platform: WGS/WES, RNA-Seq, Microarray, ILLUMINA (MiSeq, NovaSeq), DRAGEN, BaseSpace, ThermoFisher (S5plus, S5)

Tools & Software: GATK, bcftools, samtools, STAR, TRINITY, FeatureCounts, mirdeep2, Cytoscape, Mafft, SPADES, Megahit, MetageneMark2, conda/mamba, renv, Axiom Analysis Suite, plink1.9

Methods & Concepts: Over representation analysis, Gene set enrichment (GSEA), Gene Ontology enrichment (DAVID, Enrichr, ShinyGO), Differential Gene expression (DESeq2, EdgeR), Dimensionality reduction (t-SNE, PCA, MDS hierarchical clustering), Sequence Assembly & Alignment (BWA/STAR/BowTie), Gene abundance estimation (HTSeq-counts), Network enrichment (WGCNA, STRING, CleuGo, Cytohubba), Trio-analsyis, Mutation and genetic aberration analysis (VariantAnnotator, gnomAD, GeneCards), Statistical Analysis (Pearson/Spearman correlation, Student t-test, Wilcox rank sum test, ANOVA, Kruskal-Wallis, Regression Model, AUC, and ROC analysis), TWIST comprehensive research panel of Virus and virome database.

PROFESSIONAL EXPERIENCE

Gujarat Biotechnology Research Centre, Gandhinagar   **Jun 2023 – Present**

Senior Research Fellow

- Experience in analyzing multi-omics (WGS, WES, RNA-Seq, Metagenomic, Microarray) data and perform large scale genomic analysis including quality control, data preprocessing, running pipelines, downstream analysis, and result visualization (ggplot2/gplots).
- Analyzed multidimensional high content data sets, integration of these data sets with clinical metadata, and assessment of correlations across datasets to identify potential therapeutic targets in patients suffering from coronary artery disease.
- Proficient coding skills in R and bash to wrangle, analyze and visualize data.
- Sound knowledge of statistical concepts and implementation of statistical tests for hypothesis driven research.
- Parallelization of jobs over HPC to run routine pipelines for large number of samples in addition to data storage and management (NAS).
- Cloud based genomic data analysis 16S Microbiome and whole shotgun metagenome using DIAMOND + MEGAN workflow.
- Deign Affymetrix SNP Chip on *Panthera leo leo* using 23 WGS, including trio-analysis, to discern rare genetic variants of Asiatic Lion.
- Significant gene expression profiles in *Bos indicus* organs affected by virus LSDV. Proficiently utilized R for NGS data visualization.
- Concordance analysis genotype-phenotype correlations in AMR, Heatmaps, Phylodynamic analysis, and diverse graphical representation in the form of geo location MAPs.
- Ability to effectively communicate results in the form of Manuscripts, Articles and Reports. Largely independent collaborations with diverse research groups to provide computational support as well as contributions to grants.

Project 1: Development of Camelid Single Domain Antibody (SdAb)

- Bio panning data from BP0 to BP5, sequence obtained from MiSeq (250x2) chemistry.
- Aligned the data with reference IMGT reference library of Alpaca downloaded from https://www.imgt.org
- MIXCR was utilized to align, assemble and export clonotypes obtained from the library.
- Statistical analysis and for data visualization used (VDJtools/Immunarch)

Project 2: Dysregulated Non-coding RNA Identification in CAD Patients

- Using GRCH38.p14 reference genome, mapped Raw sequencing data obtained from NextSeq 2000.
- Implemented miRDeep2 workflow for identification of novel miRNAs along with prediction of target genes with (miRbase, miRDB, TargetScan).
- Validation using HMDD + qRT-PCR. Gene ontology and pathway enchainment analysis with DAVID.

Project 3: TATVAM (Translational Applications for Therapeutic from Veterinary and Allied Microbials)

- Rumen gut microbiome profiles in Sheep and Goat under different dietary conditions (17% Moringa leaves).
- Processed (.bcl) files from NovaSeq 6000 for Diamond BLAST + MEGAN analysis using AnnoTree database.
- Visualized phylogenetic trees in MEGAN6 for subsequent statistical analysis on MOCHI.

Gujarat Biotechnology Research Centre, Gandhinagar   **Mar 2021 – Jun 2023**

   Junior Research Fellow

Project 1: Aquaculture (Genomic Selection of Elite Fish Varieties for Seed Improvement)

- Analyzed Kidney, Brain, Gills tissue data for salinity treated and normal Labeo rohita freshwater fish Bioproject accession number: (PRJNA853878).
- Differential expression analysis, PCA, Heatmap for the visualization of Data.
- Identification of ncRNA (miRNA and lncRNA) and Target gene prediction using miRbase, miRTarBase, and RNAcentral.
- Correlating them with mRNA expressions to understand ce-RNA hypothesis in R using cor() function.
- Network analysis in Cytoscape, using various packages (Cytohubba, ClueGo) gene enrichment, ontology, and Kegg pathway analysis.
- Alternate Splicing event analysis in the transcriptome data to understand different spliced sites Exon Skipping, A3SS, A5SS, mutually exclusive exons and intron retention.
- Codon usage analysis and variant calling to understand SNP effect in adaptation of organism towards higher salinities.

Project 2: Detection of SAR CoV2 from Human and Wastewater Samples

- Performed analysis of the SARS CoV-2 sequence data using reference-based mapping against Wuhan-Hu-1 with bwa, bcftools/vcftools for variant calling and generating consensus using bedtools. The lineages were assessed with pangolin and nextclade.
- Phylogenetic based analysis, protein modelling of viruses using alphafold, Swiss-model and i-TASSER.
- As author submitted 746 Curated covid genomes on GISAID and NCBI SRA under (Bioproject accession number: (PRJNA625669).
- Built Nextstrain (Real-time tracking of pathogen evolution), for around 10,000 SARS CoV-2 genomes for visualization and tracking mutations with the help of Augur (phylogenetic analysis) and Auspice (interactive visualisations of phylogenomic data).
- Prepared automated shell script for understanding the lineage abundances from mixed SARS CoV-2 wastewater samples using Freyja tool (andersen-lab), for both single-end and paired-end sequencing data.

EDUCATION

Parul University, Faculty of Applied Sciences, Vadodara   **2018 - 2020**
Msc. Biotecnology – CGPA (8.92)

Parul University, Faculty of Applied Sciences, Vadodara   **2015 - 2018**
Bsc. Biotecnology – CGPA (8.31)

PUBLICATIONS

2025

- Comprehensive Analysis of Extensive Drug-Resistant Salmonella Typhi in Gujarat Region, India: Genomic Findings and Prospective Alternative Therapy (Accepted) **Microbiology Spectrum**
- Ruminal microbial responses to Moringa oleifera feed in lactating goats (Capra hircus): A metagenomic exploration (https://doi.org/10.1016/j.nbt.2025.01.006) **New Biotechnology**

2024

- Moringa leaf meal exerts growth benefits in small ruminants through modulating the gastrointestinal microbiome (https://doi.org/10.1007/s00253-024-13265-5) **AMAB**
- lncRNA-miRNA-mRNA network in kidney transcriptome of Labeo rohita under hypersaline environment (https://doi.org/10.1038/s41597-024-03056-y) **Nature Scientific Data**

2023

- Resistant cumin cultivar, GC-4 counters Fusarium oxysporum f. sp. cumini infection through up-regulation of steroid biosynthesis, limonene and pinene degradation and butanoate metabolism pathways (https://doi.org/10.3389/fpls.2023.1204828) **Frontiers in Plant Science**
- Integrative miRNA-mRNA network analysis to identify crucial pathways of salinity adaptation in brain transcriptome of Labeo rohita (https://doi.org/10.3389/fgene.2023.1209843) **Frontiers in Genetics**
- Interplay of gene expression and regulators under salinity stress in gill of Labeo rohita (https://doi.org/10.1186/s12864-023-09426-x) **BMC Genomics**

2022

- Kidney transcriptome response to salinity adaptation in rohu fish (Labeo rohita) (https://doi.org/10.3389/fphys.2022.991366) **Frontiers in Physiology**
- Resistance evolution against host-directed antiviral agents: Buffalopox virus switches to use p38-Y under long-term selective pressure of an inhibitor targeting p38-α (https://doi.org/10.1093/molbev/msac177) **Molecular Biology and Evolution**

2021

- A simple and quick PCR based method for detection of Omicron variant of SARS-CoV-2 (https://doi.org/10.1101/2021.12.20.21268053) **medRxiv**

POSTERS & CONFERENCES

- <u>Conference:</u> 48th Annual meeting and Internation conference of the Indian Society of Human Genetics (ISHG) 2024, was part of organizing committee of the workshop and seminars organized by Institute of Human Genetics and GBRC, Ahmedabad Jan 2024.
- <u>Workshop:</u> Part of GATK 3 days' workshop by BROAD institute at IIT-Gandhinagar, Jan 2024
- <u>Poster:</u> Interplay of gene expression and regulators under salinity stress in gill of Labeo rohita, National Conference part of Vibrant Gujarat - Bhartiya Vigyan Sammelan, Science City Ahmedabad December 2023.
- <u>Talk:</u> Pre-activated antiviral innate immunity in the upper airways controls early SARS-CoV-2 infection in children Author: J.Loske.