

VAISHNAV P VARMA

New Delhi, India | +91-9677399947 | vaishnavpvarma@gmail.com
🌐 vaishnavpvarma.github.io | 📄 vaishnav-p-varma | 📁 vaishnavpvarma

EDUCATION

Alagappa University | Karaikudi, India Jul. 2023 – May 2025
M.Sc. Bioinformatics GPA: 7.77/10.00

- Thesis: *Multi-Epitope Vaccine Design Against Hepatitis B Virus Using Reverse Vaccinology*
- Supervisor: Dr. M. Karthikeyan | Dept. of Bioinformatics

Bharathiar University | Coimbatore, India Sep. 2020 – May 2023
B.Sc. Biotechnology GPA: 7.2/10.00

- Thesis: *Evaluation of the Periodicity and Impact of Consumption of Convenience Foods Among Age Groups*
- Supervisor: Dr. S. R. Madhan Shankar | Dept. of Biotechnology

RESEARCH INTERESTS

- Computational method development and machine learning for integrative multi-omics analysis in infectious disease and host–pathogen genomics.
- Predictive multi-omics modelling of host–virus epigenetic interactions driving viral latency and persistence.
- Graph and deep learning approaches for network-based integration of multi-omic and clinical data.

RESEARCH EXPERIENCE AND PROJECTS

CSIR – Institute of Genomics and Integrative Biology | New Delhi, India Aug. 2024 – Present
Project Associate-I | INGEN-HOPE Lab | PI: Dr. Rajesh Pandey

- Executed end-to-end ONT dengue surveillance pipeline (Dorado, minimap2, Clair3, ARTIC) across longitudinal clinical cohorts.
- Developed Maxscriber, a Linux CLI tool automating end-to-end clinical metadata transcription, patient stratification, QC validation, and statistical visualisation for ONT dengue cohorts.
- Performed haplotype-resolved CpG methylation analysis on long-read human WGS data using Modkit; implemented phased differential methylation profiling and gene-level epigenetic annotation.
- Leading ongoing SNP-based genomic characterisation of *ESKAPE* pathogens from ONT data across two active national and international collaborative projects.

Multi-Epitope Vaccine Design Against Hepatitis B Virus Dec. 2024 – May 2025
M.Sc. Dissertation | Alagappa University | Supervisor: Dr. M. Karthikeyan 📄 📁

- Screened 111,000+ HBV protein sequences across 8 genotypes; filtered for conservation, antigenicity, and absence of human homology, yielding 9 validated epitope candidates.
- Assembled and structurally validated a multi-epitope construct achieving >97% global population coverage (MHC-I/II binding).
- Automated 2,000+ epitope immuno-toxicity filtering steps using custom Perl scripts, reducing manual curation time significantly.
- Modelled host immune response via C-ImmSim; codon-optimised final construct for heterologous expression in *E. coli*.

CSIR – Institute of Microbial Technology | Chandigarh, India May 2024 – Jul. 2024
Research Intern | Supervisor: Dr. Srikrishna Subramanian

- Benchmarked five prokaryotic genome annotation pipelines on *Thermodesulfobium narugense*, a low-GC, under-characterised bacterium, revealing systematic CDS prediction discrepancies across tools.
- Identified a Signal Peptidase II gene absent from the NCBI-PGAP annotation; confirmed structural validity using AlphaFold3 fold prediction and DALI structural comparison.
- Supported functional assignment via PSAURON scoring and DIAMOND-Blastp domain mapping, illustrating the inadequacy of single-pipeline annotation for rare microbial genomes.

PATENTS AND PUBLICATIONS

Manuscripts in preparation

- Pandey, R., Ravi, V., **Varma, V. P.**, et al. (In preparation). Longitudinal dengue genomic surveillance using Oxford Nanopore long-read sequencing.
- Pandey, R., Mehta, P., Maurya, R., **Varma, V. P.**, et al. (In preparation). Haplotype-resolved methylation profiling in long-read human whole-genome sequencing.

Book Chapter

- **Varma, V. P.**, Arun, A., Kiran, A., & Senthil, R. D. (2023). Evaluation of the periodicity and impact of consumption of convenience foods among age groups. In *Food and Fermentation – Recent Perspectives* (pp. 32–39). L'Ordine Nuovo Publication. ISBN: 978-93-92995-46-0.

Patent

- **Varma, V. P.**, Shankar, S. R. M., & Iswarya, B. R. (2024). An Apparatus for Collecting Tree Sap. Patent No. 520964, The Patent Office, Government of India. Granted: 06 Mar. 2024.

PROFESSIONAL MEMBERSHIPS

- Bioinformatics and Drug Discovery Society (BIDDS) | Member, ID: SMBIDDS17-572 | Feb. 2025 – Present

FELLOWSHIPS AND CERTIFICATIONS

- CSIR-IMTECH Post-Graduate Training Fellowship | Awarded for competitive selection to the Bioinformatics Centre | May – Jul. 2024
- NPTEL: Big Data Computing | Indian Institute of Technology | Aug. – Oct. 2023
- National Training on Genetic Engineering Using CRISPR Technology | IIT Kharagpur | Nov. 2022

TECHNICAL SKILLS

Programming & Infrastructure:

- **Programming & Software Development:** Python, R, Bash/Shell, Perl, C++, Go; algorithm design, data structures, statistical modelling, scientific software engineering, task automation
- **Development Environment & Reproducibility:** Linux/Unix (Ubuntu, CentOS, Fedora), Git/GitHub, conda/mamba, SLURM, tmux/screen; version-controlled, reproducible research workflows
- **Workflow & Pipeline Engineering:** Nextflow (nf-core), Snakemake; scalable pipeline design, parallel execution, HPC job scheduling
- **AI-Augmented Development:** LLM-assisted coding workflows (Claude, OpenCode); local model deployment (Ollama); API integration via OpenRouter; prompt engineering for scientific automation

Data Science & ML:

- **Data Science & Machine Learning:** pandas, NumPy, SciPy, scikit-learn, Biopython; hypothesis testing, classification, clustering, feature engineering, biological data integration
- **Data Visualisation:** matplotlib, seaborn, ggplot2, plotly, UpSetR; publication-quality figure design

Long-Read Sequencing & Genomics:

- **Sequencing Platforms:** Oxford Nanopore Technologies: MinION (Mk1B/Mk1C), GridION, PromethION - P2Solo and P24; real-time data acquisition, post-sequencing QC, and downstream computational analysis
- **Bioinformatics Pipelines (End-to-End):**
 - Basecalling: Dorado, Guppy (signal-to-sequence modelling)
 - Alignment: minimap2, BWA-MEM, Bowtie2 (k-mer indexing, read mapping)
 - Assembly & Polishing: Flye (assembly graphs), Medaka
 - Variant Calling: Clair3 (haplotype-aware, long-read); QC & Post-processing: samtools, bcftools, bedtools, VCFtools, tabix; Variant Annotation: SnpEff
 - Haplotype-resolved methylation: Modkit, Nanopolish; CpG profiling, differential methylation analysis

Genome Annotation & Comparative Genomics:

- Annotation: PROKKA, BAKTA, RAST, DFAST, DRAM, BEACON, PSAURON; functional annotation, metabolic reconstruction, structure-assisted gene validation (AlphaFold3)

- Comparative genomics & phylogenomics: DIAMOND, InterProScan, KAAS, IGV; MAFFT (multiple sequence alignment), Augur (phylodynamics)
- Metagenomics & taxonomic profiling: Kraken2, eggNOG-mapper; taxonomic classification, functional profiling

Structural Bioinformatics & Molecular Modelling:

- Structure prediction & homology modelling: AlphaFold, SWISS-MODEL, MODELLER, PSIPRED; DALI (structural alignment); PyMOL, UCSF Chimera, Discovery Studio
- Molecular docking & dynamics: AutoDock Vina, PyRx, HADDOCK; GROMACS (MD simulation: system setup, equilibration, production runs, trajectory analysis — RMSD, RMSF, binding free energy)

Immunoinformatics, Databases & Custom Tools:

- **Immunoinformatics & Vaccine Design:** NetMHCpan, NetMHCIIpan, ABCPred, VaxiJen, AllerTOP, IFNepitope, ToxinPred, IEDB tools, C-ImmSim (immune simulation), JCAT, CD-HIT; epitope prediction, reverse vaccinology, multi-epitope vaccine design
- **Biological Databases & Resources:** NCBI (SRA, GenBank, RefSeq), UniProt, KEGG, IEDB, HBVdb, GISAID, MySQL; programmatic data retrieval via REST APIs
- **Custom Tools:** Maxscriber (Linux CLI tool for clinical metadata transcription, patient stratification, QC validation, and statistical visualisation for ONT dengue cohorts)

CO-CURRICULUM

- **Bioinformatics Trainer** February 2025
@ WHO SEARO Regional Training Workshop on Dengue Whole Genome Sequencing
Using Oxford Nanopore Technology, CSIR-IGIB, New Delhi
- **Developer of Maxscriber** 
Linux CLI automation tool for end-to-end clinical metadata management in ONT genomic surveillance projects
- **Open-source contributor** 
Custom Perl scripts for high-throughput immunoinformatics data parsing

ADDITIONAL INFORMATION

- *Languages:* English (Fluent), Hindi (Fluent), Tamil (Conversational), Malayalam (Native)
- *Scientific Interests:* Long-read sequencing algorithms, epigenomics, pangenomics, machine learning for multi-omics, metagenomics