

ANKUR SHARMA, PhD

Senior ML & Agentic AI Engineer | Bioinformatics | Cloud Architecture

ankurs103@gmail.com · +65 8402 6093 · Singapore (PR) · [LinkedIn](#) · [GitHub](#) · [Portfolio](#)

PROFESSIONAL SUMMARY

Senior ML and Agentic AI Engineer with 8+ years shipping production systems across clinical genomics, cloud infrastructure, and machine learning. Built multi-agent systems that cut mean time to resolution by 67%, cloud pipelines processing 6,000+ genomic samples at 40% lower compute cost, and open-source agentic AI tooling for variant interpretation. PhD in Biological Sciences (NTU Singapore). Seeking Senior/Staff ML Engineer, Agentic AI Engineer, or Research Engineer roles.

TECHNICAL SKILLS

Agentic AI & LLMs: LangGraph, LangChain, Claude/Anthropic API, OpenAI GPT-4, multi-agent orchestration, RAG, tool-calling/function agents, prompt engineering, LLM routing, Pydantic v2

Machine Learning & Data Science: Python, scikit-learn, PyTorch, Pandas, NumPy, SciPy, statistical modelling, survival analysis (Cox PH), feature engineering, MLflow, Seaborn/Matplotlib

Cloud & Infrastructure: AWS (EC2, S3, Lambda, Batch, Step Functions, DynamoDB, RDS), Docker, Nextflow (DSL2), CI/CD (GitHub Actions), IaC (CloudFormation/CDK), Kubernetes

Bioinformatics & Genomics: NGS (WGS, WES, RNA-seq, ChIP-seq, WGMS, single-cell, spatial), DRAGEN, GATK, samtools, bcftools, pysam, variant calling & interpretation, ACMG/AMP, R/Bioconductor, Snakemake

Regulatory & Clinical: Design Verification Testing (DVT), clinical validation (sensitivity, specificity, LOD), FDA compliance, cGMP, SOPs, HIPAA

OPEN-SOURCE PROJECTS

agentic-genomics — GenomicsCopilot

github.com/ankurgenomics/agentic-genomics

LangGraph agent for explainable variant interpretation. 7-node pipeline: VCF ingest → MyVariant.info annotation (gnomAD, ClinVar, SpliceAI) → frequency filter → Phrank-style HPO semantic similarity scoring → ACMG-lite classification (PVS1 with LoF-intolerance gating, Richards 2015 Table 5 combining rules) → LLM synthesiser → LLM critic that fact-checks claims against evidence JSON. Every run emits a machine-readable reasoning trace. 51 tests, 81% coverage, CI/CD via GitHub Actions.

Stack: LangGraph, Claude API, Pydantic v2, pysam, Streamlit, Typer CLI, pytest, ruff

genomics-skills — Agent-Callable Genomics Skill Library

github.com/ankurgenomics/genomics-skills

Modular library of 8 pure-Python skills for downstream genomics analysis: pan-cancer expression profiling (TCGA, 9,479 samples), Kaplan-Meier survival analysis (Cox PH regression), GO/KEGG pathway enrichment, PubMed literature search with NLP digest, protein variant mapping (AlphaFold/PDB), 3D structure viewing, and publication-quality volcano plots. LLM-powered skill routing via Claude Haiku. Parquet caching for instant repeat queries.

Stack: Python, Claude API (LLM routing), cBioPortal/MyVariant/NCBI/PDB REST APIs, Pandas, Matplotlib, pytest (14 tests)

PROFESSIONAL EXPERIENCE

Senior Scientist, Bioinformatics

Jul 2025 – Present

Illumina · Singapore

- Lead **Design Verification Testing (DVT)** and regulatory validation for TSO500, NIPT16, and VeriSeq diagnostic products — ensuring FDA compliance and clinical-deployment quality.
- Designed clinical concordance studies measuring sensitivity, specificity, accuracy, LOD, and reproducibility for regulatory submissions.
- Optimised TruSight Oncology 500 assay workflows for comprehensive genomic profiling in oncology; contributed to NIPT16 prenatal testing validation.

Senior Scientist — Bioinformatics & Cloud

Jan 2022 – Jul 2025

Mirxes · Singapore

- Built production-grade cloud data-analysis pipelines: **40% ↓ compute cost, 50% ↓ storage footprint, 30% ↓ turnaround time.**

- Processed **6,000+ genomic samples** with standardised Nextflow workflows and rigorous QC on AWS.
- Architected AWS infrastructure integrating **15 services** (Batch, Lambda, Step Functions, S3, DynamoDB) to manage **400 TB** of genomic data.
- Led team of 5 scientists delivering clinical diagnostic assay workflows across cancer types — 33% ↓ analysis time.
- Technical lead for the **Singapore National Precision Medicine** project; identified SGD 0.5M in business opportunities.

Scientist, Assay Development

Jul 2021 – Dec 2021

Vela Diagnostics · Singapore

- Implemented automated verification procedures and QC protocols; maintained regulatory documentation (NCR, CAPA, DR).
- Developed testing framework that cut non-conformance reports by 30% and lifted product quality by 15%.

PhD Research Fellow, Biological Sciences

2016 – 2021

Nanyang Technological University · Singapore

- Multi-omic study (RNA-seq, ChIP-seq, Hi-C) of age-dependent transcriptional and epigenetic alterations in mouse hepatocytes.
- Built reproducible NGS analysis pipelines for transcriptome, histone modifications, and 3D chromatin data.
- Head of NTU 3-Minute Thesis team (23 members, Nanyang Awards); led operations for TEDxNTU (78-person team, 1,500+ audience).

Technical Supervisor

2011 – 2013

Zydus Cadila Healthcare Ltd. · Ahmedabad, India

- Developed process-validation and QC pipelines for 13 pharmaceutical products under cGMP guidelines — 30% ↑ efficiency, 10% ↓ processing time.

EDUCATION

PhD in Biological Sciences — Nanyang Technological University, Singapore (2016–2021)

Dissertation: Age-dependent transcriptional and epigenetic alterations in mouse hepatocytes · DOI: 10.32657/10356/155390

M.Tech in Biotechnology — Anna University, India (2013–2015)

B.Tech in Biotechnology — Uttar Pradesh Technical University, India (2007–2011)

PUBLICATIONS

Sharma, A. (2021). *Age-dependent transcriptional and epigenetic alterations in mouse hepatocytes*. PhD Dissertation, NTU Singapore. DOI

Sharma, A. et al. (2019). Significance of hepatocyte polyploidization in liver physiology and pathology. *Cell Symposia, Chicago*.

Kumari, M., Taritla, S., **Sharma, A.**, & Jayabaskaran, C. (2018). Antiproliferative and antioxidative bioactive compounds in marine-derived endophytic fungus. *Frontiers in Microbiology*.