

ABHISHEK KHATRI

Email: khatriabhi2319@gmail.com | Phone: +91 6230992504

Portfolio: <https://fgcsl.github.io/portfolio/index.html>

Bioinformatics IT professional with 6+ years of experience in designing, validating, and maintaining bioinformatics pipelines and web-based platforms under the Government of India DBT projects. Strong expertise in **Shell scripting, Bash, Python**, and working in **Linux-based environments**. Proven ability to support researchers in genomic data analysis, tool development, and deploying internal product-like platforms for large-scale scientific use.

EDUCATION

Master of Computer Applications **09/2016 - 07/2018**

Lovely Professional University

Bachelor of Computer Applications **05/2013 - 07/2016**

Himachal Pradesh University

WORK EXPERIENCE

System Analyst **12/2023 - Present**

International Center for Genetic Engineering and Biotechnology (ICGEB), New Delhi

- Developed **ThyPredict**, a machine learning pipeline for classifying thyroid histopathology images.
- Developing a Multi-omics toolbox/Webserver transcriptomics (single and bulk RNA-seq), GWAS, Proteomics.
- Created **Dockerized** and **Conda-based** environments to ensure reproducibility.
- Collected and managed project-related data from multiple DBT-Apex research centers by developing a web-based data entry form with a **MySQL** backend to streamline data submission and curation.
- Designed and implemented a shell-based **File Tagger Tool** to manage large research data files with metadata tagging. Reconfigured and installed it on a separate lab server for cross-project use.
- Built and maintained the **Technology Transfer Portal** using WordPress to facilitate access to research tools and technologies.

Freelance Project

04/2025 - 06/2025

- Collaborated with researchers to convert their Python-based code and machine learning model into a **user-friendly graphical interface**. While the core logic and model were provided by the researchers, I designed and implemented the complete **GUI** to make the tool accessible for non-programmers.

Senior Bioinformatics Programmer

08/2023 - 12/2023

National Institute of Biomedical Genomics (NIBMG), Kalyani

- Developed a command-line tool for **biological file tagging and metadata management** using Shell, R, and MySQL.

Project Associate

12/2022 - 08/2023

University of Hyderabad, Hyderabad

- Designed and implemented the **Indian Metagenomics Database** for DBT-CMI, involving materialized views, dynamic maps, and curated sample metadata.
- Performed exploratory data analysis using a shell script to support needs for collaborative data research projects.
- Contributed to large-scale data upload, processing, and integration into MySQL databases.

Project Coordinator-II / Data Analyst

01/2019 - 11/2022

CSIR-IHBT, Palampur

- Created captivating and interactive pipelines for both routine genomics data analyses and tailored pipelines for collaborative research projects.
- Developed **Indian Himalayan metagenome** database and **AutoQii2**: Automated Pipeline for analysing 16S rRNA amplicon-based datasets
- Collaborated closely with bench scientists to fully understand and actively support their research objectives
- Installed, upgraded, and configured Bioinformatics software & systems and all associated middleware in LINUX environments

PROFESSIONAL SKILLS

- **Languages/Scripting:** Python, Shell, Bash
- **Databases:** MySQL, MariaDB, PostgreSQL (basic)

- **Version Control:** Git, GitHub
- **Environment:** Linux/Unix (CentOS, Ubuntu), Conda, Docker (basic), LAMP stack
- **NGS Platforms:** Illumina, Oxford Nanopore
- **Cloud/Servers:** Web-hosted tools, server-side development, HPC
- **Other:** WordPress, PHP, HTML/CSS (used for bioinformatics portals)

PUBLICATIONS

- A Novel Three-Stage AI-Assisted Approach for Accurate Differential Diagnosis and Classification of NIFTP and Thyroid Neoplasms, *Endocrine Pathology*, 2025 (CIF =11.3)
- IHM-DB: a curated collection of metagenomics data from the Indian Himalayan region, and AutoQii2: Automated pipeline for 16S amplicon-based analysis. *Database (Oxford)*, 2023 (CIF=4.462, # First author)
- Deep learning uncovers distinct behavior of rice network to pathogens response. *iScience*, (Cell Press), 2022 (CIF=5.74)
- Quality assessment, safety evaluation, and microbiome analysis of night soil compost from lahaul valley of northwestern Himalaya. *Waste Management*, 2022 (CIF =7.145)
- Comparative genomics and molecular adaptational analysis of *Arthrobacter* from Sikkim Himalaya provided insights into its survivability under multiple high-altitude stress *Genomics*, 2021 (CIF =5.736, # Joint first author)