Curriculum Vitae

Name: Taiba Shamim
Email: taibashamim04@gmail.com
LinkedIn: https://www.linkedin.com/in/taiba-shamim-31b64922b/
GitHub: https://github.com/taiba-shamim

Career Summary
Second-year M.Sc. Bioinformatics student, with sharp commitment and adaptability, hands-on experience in NGS data Analysis**,** GRN reconstruction, **RNA-seq analysis** (bulk-cell seq), **Biomarker Discovery**, protein structure prediction and Structural analysis of Proteins, Skilled in **Python, R, NGS data analysis (Variant Calling),** also familiar to **(galaxy platform)** and practical exposure to **homology modelling & molecular docking.**

Education	
M.Sc. (Bioinformatics) *Jamia Millia Islamia, Delhi* *(Dept. Of Computer Science)* CGPA: 9.34/10 (Ist YEAR)	| 2024 – 2026
Bachelor of Science (Honors) Biotechnology *Mahatma Jyotiba Phule Rohilkhand University, Bareilly* Percentage: 83.4%	| 2021 - 2024
Senior Secondary (10+2) *St. Mary's Convent Senior Secondary School, Amroha (CBSE)* Percentage: 86.6%	| 2019 - 2020
Secondary (10th) *St. Mary's Convent Senior Secondary School, Amroha (CBSE)* Percentage: 87%	| 2017 - 2018

Skills-

- **Programming:** Python, R, Java Script, PHP

- **Linux**

- **NGS Data Analysis –** Trimmomatic, BWA, Bowtie2, Samtools, GATK, multiQC, fastQC, Galaxy Platform, Cufflinks, CNVkit, Deep Variant..

- **Bioinformatics Tools & Databases:** BLAST, NCBI GEO, STRING, SRA, TCGA Cytoscape, SWISS-Model, PyMOL, Cell Designer, MCODE, UCSF Chimera, AutoDock, IGV, Galaxy Platform, DAVID .

- **Protein Structure Prediction:** Homology modelling of **POTE paralogs** using **Swiss-Model**

- **Protein structure visualization** using PyMOL (3TZM, POTEA)

- **Molecular Docking:** using **AutoDock Vina, ChimeraX, PyMOL**

- **Sequence alignment:** BLAST, CLUSTAL-W
- **Pathway Modelling-** Cell designer
- **Laboratory Techniques:** Gel Electrophoresis, PCR, SDS-PAGE, Centrifugation, Broth culture
- **Other Skills:** MS Word, MS PowerPoint, MS Excel, My SQL, HTML.

Projects (self & academic)

1. Gene Regulatory Network Construction using Mutual Information and ARACNE based method.

- Constructed a **Gene Regulatory Network (GRN)** from publicly available **GEO datasets** using **ARACNE** as the key inference method.
- Applied Python-based libraries to process gene expression data and infer regulatory relationships.
- Performed Topological Analysis of the GRN, using Python.
- Utilized visualization tools-**Cytoscape** to represent the network structure.

2. Identification of Differentially Expressed Genes in Inflammatory Bowel Disease

- Analyzed the dataset **GDS1615** related to **Crohn's disease and ulcerative colitis** to identify differentially expressed genes.

3. **Protein Structure Prediction** of POTE paralogs

4. **Phylogenetic Tree construction-**using MEGA 12 of POTE paralogs

5. **Structural Analysis of Proteins –** allosteric sites analysis (using python)

6. **NGS Data Analysis of triple negative breast cancer and ovarian cancer-** quality check , read mapping and assessment, Visualization.

7. **Variant Calling, Variant Filtering, Variant Annotation (SNPs, InDels, CNVs)**

8. **RNA-Seq Analysis** – discovery of novel potential gene biomarkers for the diagnosis, prognosis, and development of drug targets of Breast Cancer.

9. **Functional Annotation, KEGG Pathway Enrichment –** of DEGs Identified in Breast Cancer Dataset using R.

10. **Gene Ranking Using Machine Learning (R-based)**
    Implemented ML models in R to rank top 10 DEGs using FPKM values, identifying biologically significant genes for expression analysis.

Publications-

Ali, A. M., Shah, S. N., Shamim, T., Khan, A., Rizwee, F., Khan, E. A. R., Neeraj, & Qazi, S. (2025). **Patient-centric care and digital health tools.** *In Advances in Protein Chemistry and Structural Biology.*

Elsevier. https://www.sciencedirect.com/science/article/abs/pii/S0079612325001207

Internship

➢ **Jawaharlal Nehru University (School of Computational and Integrative Sciences)-**

Summer Intern (May-July)

- Network Pharmacology, Data Mining, Web Scrapping, Machine Learning.
- Worked on cervical cancer dataset
- **Certificate of Internship -**
  https://drive.google.com/file/d/1F4CcnDI6ybL8oe6ifUpqDnr7uXcC6L13/view?usp=drivesdk

---

Ongoing Project (academic)

- *"Developing a Machine Learning Model for Prognosis of Hepatocellular Carcinoma using RNA-Seq and Clinical Data"*

Awards- Academic Excellence Award- B.Sc Biotechnology, MJPRU Bareilly.

Extracurricular & Leadership

- **Vice President - BioNexus Club Jamia Millia Islamia** – Organized Linux/mini conda setup hands on session.

Languages

- **English:** Fluent
- **Hindi:** Fluent
- **Urdu:** Native