SHAHROZ RAHMAN

shahrozrahman.me@gmail.com +92-344-6497495 Linkedin

Scholar Github Portfolio

PERSONAL PROFILE

Bioinformatics graduate with training in computational analysis of biological data and experience working with genome-scale datasets and omics-based studies. My academic background combines biological sciences with programming and data analysis, including the use of Python, R, and Linux-based bioinformatics tools. I have worked on genome-wide analyses, sequence analysis, and the interpretation of high-throughput sequencing data, and have contributed to several research publications during my undergraduate training. I am seeking to pursue a master's degree to further develop my expertise in computational genomics, bioinformatics workflows, and large-scale biological data analysis while gaining deeper research experience in interdisciplinary biological research.

EDUCATION

Government College University Faisalabad **Sep 2019 - Aug 2023**

BS Bioinformatics

- Graduated with **3.72/4.00** CGPA

Relevant Modules:

- Cell and Molecular Biology, Molecular Genetics, Introduction to Computer Programming, Data Structure and Algorithms, Calculus, Biological Data Retrieval, Object-Oriented Programming, Database Management Systems, Bioinformatics Software Development, Biological Data Analysis, Genomics, Proteomics, Bioinformatics Scripting, Systems Biology

Thesis:

- Identification of Novel Covid-19 Variants in Pakistan

RESEARCH EXPERIENCES

Determination of Fatty Acid Profile Research Intern, AYUB Research **Oct 2022 - Nov 2022**

Oil Technology Laboratory of Oilseeds Research Institute

- Gained expertise in Solvent Extraction Technique for extracting oil from major oilseed crops, modern analytical techniques for determination of Fatty acid profile of oilseed crops especially through Near Infrared Reflectance Spectroscopy (NIR) and report writing.

Computational Data Analyst Research Assistant **Sep 2023 - Jul 2025**

Functional Genomics & Tissue Culture Lab

- Developed proficiency in both wet-lab and computational techniques, including real-time PCR, gel electrophoresis, genome-wide gene family identification, NGS data analysis, and variant calling. Contributed to the development of a biosystem for capturing phenotypic variation and implemented a machine-learning pipeline for automated image analysis, improving accuracy by about 30%. Built strong bioinformatics and programming skills in Python, R, and Bash, focusing on biological data analysis and computational workflows. Conducted genomic studies on multiple gene families associated with stress responses and contributed to several peer-reviewed research publications.

Bioinformatics Data Analyst (Freelancer) **Aug 2025 - Present**

Research Specialist, Biological Data Analyst

- I am currently providing bioinformatics and computational analysis support for research projects involving genomic and high-throughput biological datasets. I analyze sequence and omics data using Python, R, and Linux-based tools, develop reproducible analysis pipelines, and assist researchers with data interpretation, visualization, and preparation of results for scientific publications.

PUBLICATIONS

1. **Rahman S**, Alshehri D, Siham A.B, Ikram A, Rizwan M, Rafique M, Siraj H, Azeem F, "Genome-wide Identification of bZIP transcription factors and their role toward drought stress and fruit development in *Mangifera indica* (Mango)". Crop and Pasture Science. **2025**, https://doi.org/10.1071/CP25008
2. **Rahman S**, Ikram AR, AlHusnain L, Fiaz S, Rafique MU, Ali MA, AlKahtani MD, Attia KA, Azeem F. "Genome-wide profiling of bZIP transcription factors in *Camelina sativa*: implications for development and stress response". BMC Genomic Data. **2024**. https://doi.org/10.1186/s12863-024-01270-6
3. **Rahman S**, Rehman A, Waqas M, Mubarik MS, Alwutayd K, AbdElgawad H, Jalal A, Azeem F, Rizwan M. "Genome-wide exploration of bZIP transcription factors and their contribution to alkali stress response in *Helianthus annuus*". Plant Stress. **2023**, https://doi.org/10.1016/j.stress.2023.100204
4. **Rahman S**, Ikram AR, Azeem F, Tahir ul Qamar M, Shaheen T. "Precision Genome Editing with CRISPR-Cas9". In Plant Functional Genomics: Methods and Protocols, Volume 2 **2024** Apr 25 (pp. 355-372). New York, NY: Springer US, https://doi.org/10.1007/978-1-0716-3782-1_21
5. Rehman A, Alwutayd KM, Alshehri D, Alsudays IM, Azeem F, **Rahman S**, Abid M, Shah AA. "Regulatory role of AGC genes in heat stress adaptation in maize (*Zea mays*)". Functional Plant Biology. **2024**, https://doi.org/10.1071/FP23282
6. Rahman A, Batool S, Shahid M, **Rahman S**, Waqas M, Azeem F. "Computational Approaches for Retrotransposon Investigations". In Functional Genomics: Methods and Protocols, Volume 3 **2025**. Taylor & Francis, 10.1201/9781032663494-5
7. Haris M, Rasheed S, **Rahman S**, Fatima I, Mazhar A, Azeem F. "Tissue Culture Optimization: Maize Callus Induction and Factors Affecting the Callus Efficiency". Integr. Plant Biotechnol. **2025**, https://doi.org/10.55627/pbiotech.003.02.1288
8. Haris M, **Rahman S**, Qamar MTU, Azeem F. "HCMdrugs: A Web-Based Database Cataloging the Drugs Available for the Treatment of Hypertrophic Cardiomyopathy". Journal of Computational Biophysics and Chemistry, **2026.** https://doi.org/10.1142/S2737416525500954
9. Mudassar M, Shahid MT, Ahmed J, Naeem MZ, Shahid MB, Azeem F, **Rahman S**, Jahangeer M, Kausar A, Baki GB, Khaliq M. "Effects of Exposure to Environmental Pollutants on Male Fertility". Annual Methodological Archive Research Review, **2025**. https://doi.org/10.63075/5kxcr887

PROJECTS

CRC-DataLab: Tool for Colorectal Cancer Genomic Research **Sep 2025**

- Tool for analyzing gene expression and genomic variants in colorectal cancer. It integrates pre-analyzed data, statistical analyses, machine learning, and text mining to provide actionable insights.

Single-Cell RNA-Seq Explorer: Interactive Web Application **Mar 2025**

- Streamlit application designed to help users explore and analyze single-cell transcriptomic data with interactive visualizations and clustering tools. This app provides features for dimensionality reduction, marker gene identification, and customizable data input formats.

Bio-Wrangler: Bioinformatics Data Wrangling Package **Dec 2024**

- Bio-Wrangler is a Python package developed for seamless manipulation of bioinformatics formats such as FASTA, FASTQ, VCF, and GFF. It provides tools to load, filter, summarize, and merge datasets into DataFrames, allowing efficient data analysis.

SKILLS

- **Computational Coding:** Proficient in Python (Biopython, Streamlit, ML libraries, Tkinter), R, and Bash scripting for bioinformatics workflows and data analysis. Experienced in pipeline development, automation, and data visualization. Familiar with Linux command-line environments, Jupyter Notebook, RStudio, and Galaxy for NGS analysis, with working knowledge of HTML, CSS, and JavaScript.
- **Bioinformatics & Data Analysis:** Experienced in NGS and omics data analysis, including DNA/RNA sequencing, differential gene expression (DEG) analysis, and single-cell sequencing workflows. Background in genomics, transcriptomics, proteomics, and multi-omics integration. Familiar with variant calling, motif discovery, ChIP-Seq analysis, molecular docking, protein structure prediction, and metagenomics pipelines.
- **Research & Scientific Writing:** Skilled in biological data mining, statistical analysis, and computational analysis of large biological datasets. Experienced in scientific writing, manuscript preparation, and bioinformatics documentation for research publications.
- **Wet Lab Techniques:** Basic experience in molecular biology techniques including PCR, RT-PCR, SDS-PAGE, ELISA, and related laboratory procedures.

ADDITIONAL CERTIFICATIONS AND CONFERENCES

Bioinformatics Command Line Training Workshop, Bioinformatics.pk **Sep 2023**

- Organized a workshop on bioinformatics data analysis through the online platform bioinformatics.pk. Conducted a comprehensive 2-hour session to train students in utilizing command-line tools for efficient bioinformatics data analysis.

NGS Data Analysis: Basics to Advances, GCUF **Nov 2023**

- Participated as an organizer and speaker at a nationally recognized three-day workshop hosted by Government College University Faisalabad, Pakistan. The workshop, titled NGS Data Analysis: Basics to Advances was a resounding success, providing training to over a thousand students.

Hands-on Bioinformatics Starter Training Program **Aug 2022**

- I have participated as a speaker and organizer in a three-day "Hands-on bioinformatics starter training program"workshop, which was held online and was organized by bioinformatics.pk.

A guide to multiomics pathway analysis **Jul 2021**

- Participated in a webinar hosted by EMBL-EBI, In the webinar, the presenter discussed the methods for performing comparative multiomics pathway analyses using the ReactomeGSA pathway analysis system.

REFERENCES

- **Dr. Farrukh Azeem,** Department of Bioinformatics, GCUF Pakistan

Email: farrukh@gcuf.edu.pk

- **Dr. Muhammad Zubair,** Department of Bioinformatics, GCUF Pakistan

Email: drmzubair@gcuf.edu.pk

- **Dr. Muhammad Tahir ul Qamar,** Department of Bioinformatics, GCUF Pakistan

Email: tahirulqamar@gcuf.edu.pk