

Md. Jubayer Hossain

Department of Public Health, Daffodil International University

Daffodil Smart City, Birulia 1216

+8801622345906 | hossainmj.me@gmail.com | mdjubayerhossain.com

Research Statement

My research focuses on developing reproducible bioinformatics pipelines for large-scale bulk and single-cell RNA-seq data to uncover molecular biomarkers, immunological signatures, and therapeutic targets in cancer and neurological diseases. By integrating multi-omics data analysis, machine learning, and biology, I aim to advance computational methods that bridge biological discovery and clinical translation.

Education and Training

Master of Science in Microbiology

Jagannath University

Dhaka, Bangladesh

2019–2020

- Relevant Coursework: Genomics, Proteomics, Bioinformatics, Research Methodology and Scientific Writing

Bachelor of Science in Microbiology

Jagannath University

Dhaka, Bangladesh

2016–2019

- Relevant Coursework: Biostatistics, Public Health and Hygiene, Bioinformatics I, Bioinformatics II

Interdisciplinary Computational Biology

BRAC University

Dhaka, Bangladesh

January 2026

- Mastered high-resolution protein structure prediction and molecular visualization using AlphaFold2/3 and ChimeraX to investigate complex protein-ligand interactions
- Applied AI/ML frameworks and docking protocols to structural data from PDB and UniProt, culminating in an interdisciplinary project on predictive molecular modeling.

AI in Public Health

Child Health Research Foundation (CHRF)

Dhaka, Bangladesh

Sep 2023

- Developed predictive models and data pipelines for disease surveillance, focusing on the calculation of typhoid incidence and vaccine impact in Bangladesh.
- Built LLM-based code generation and computer vision frameworks for medical image analysis while analyzing the ethical implications of algorithmic bias.

Technical Skills

- Programming Languages:** Python, R, Julia, SQL, Bash, JavaScript
- Bioinformatics:** bulk RNA-Seq, Single-Cell, Spatial Omics, and Multi-omics
- Workflow Management & Reproducibility:** Nextflow, Docker, Git, GitHub
- Computing Environments:** UNIX/Linux, HPC clusters, Cloud computing (AWS, Google Cloud)

Research Experience

Visiting Scholar

Department of Public Health, Daffodil International University

Dhaka, Bangladesh

Sep 2025 – Present

- Investigated computational frameworks within public health genomics to facilitate large-scale biomarker discovery and clinical validation.
- Developed a specialized curriculum for AI-driven disease modeling, integrating neural network architectures into biomedical informatics.
- Synthesized heterogeneous digital health records with genomic datasets to evaluate predictive models for clinical outcomes.
- Led a multidisciplinary research initiative to translate raw sequencing data into evidence-based public health insights.

- Analyzed genomic and transcriptomic datasets using computational tools in R and Python, with emphasis on high-throughput sequencing data.
- Designed and delivered training modules on sequence alignment algorithms, RNA-seq analysis pipelines, and data preprocessing workflows.
- Evaluated and implemented peer-reviewed bioinformatics methods for transcriptomic data, ensuring reproducibility and analytical rigor.
- Facilitated technical discussions on structural biology and next-generation sequencing (NGS) data interpretation.

Data Enumerator

Perth, Australia

School of Allied Health, Curtin University

Oct 2022 – Nov 2022

- Collected and curated quantitative and qualitative data from 40 clinicians for a Delphi study on Patient Outcomes Measurement in Interprofessional Tuberculosis Care.
- Performed rigorous data validation and cleaning, achieving 100% dataset accuracy for downstream statistical analysis.

Selected Publications

1. Ahmed, M. Z., Billah, M. M., Ferdous, J., & **Hossain, M. J.*** (2025). Pan-cancer analysis reveals immunological and prognostic significance of CCT5 in human tumors. *Scientific Reports*, 15, 14405. <https://doi.org/10.1038/s41598-025-88339-z>
2. Fariha, F. T. J., Fuad, M., Saha, C. S., Hossen, S., & **Hossain, M. J.*** (2025). Comprehensive bioinformatics analysis reveals prognostic significance and immunological roles of WNT gene family in breast cancer. *Sci Rep* 15, 34490 (2025). <https://doi.org/10.1038/s41598-025-13315-6>
3. Bari, S.M., Fuad, M., **Hossain, M.J.** et al. A meta-analysis of public RNA-Seq data identifies conserved stress responses in rainbow trout. *BMC Genomics* 26, 999 (2025). <https://doi.org/10.1186/s12864-025-12127-2>
4. GBD 2023 Vaccine Coverage Collaborators (**Hossain, M. J.**, 2025). Global, regional, and national trends in routine childhood vaccination coverage from 1980 to 2023 with forecasts to 2030: A systematic analysis for the Global Burden of Disease Study 2023. *The Lancet*. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(25\)01037-2/](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(25)01037-2/)
5. GBD 2023 Headache Collaborators (**Hossain, M. J.**, 2025). Global, regional, and national burden of headache disorders, 1990–2023: a systematic analysis for the Global Burden of Disease Study 2023. *Lancet Neurology*. [https://www.thelancet.com/journals/laneur/article/PIIS1474-4422\(25\)00402-8/fulltext](https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(25)00402-8/fulltext)

Papers Under Review

1. Islam, M. W., Rahman, M. M., Naznin, H., Hossain, M. S., Akter, T., Shatabde, Z. A., & **Hossain, M. J.*** (2025). “Integrative bioinformatics analysis reveals COL13A1 and COL23A1 as potential diagnostic and prognostic biomarkers in thyroid cancer”. *Under Review*.
2. Islam, M. W., Fariha, F. T. J., Ahmed, M. Z., Ferdous, J., Lota, H. B., Hossain, M. S., Kundu, P., Shahjahan, M., & **Hossain, M. J.*** (2025). “Bioinformatics-driven multi-omics profiling of CDK1 and CDK6 identifies prognostic and therapeutic roles in breast cancer”. *Under Review*.
3. Mabsurah, K., Yeana, I. J., Sumaiya, M., Islam, T., Ahammad, K., Bhattacharjee, A., Ferdous, J., Billah, M. M., & **Hossain, M. J.*** (2025). “Pan-cancer analysis identifies potential prognostic and immunological role of RAS-degrading WDR76 in human tumors”. *Under Review*.

Working Papers

1. **Hossain, M. J.**, Shahjahan, M., & Fuad, M. (2025). Fezf2-mediated cortical development: Multi-omics analysis. *Working Paper/GitHub*. <https://github.com/hossainlab/fezf2-multiomics>

2. **Hossain, M. J.**, Shahjahan, M., & Fuad, M. (2025). Single-cell meta-analysis of microglial activation in Alzheimer's disease. *Working Paper/GitHub*. <https://github.com/hossainlab/alzheimers-microglia-meta-analysis>
3. **Hossain, M. J.**, Shahjahan, M., & Fuad, M. (2025). Integrative spatial-scRNA-seq atlas of immunotherapy resistance mechanisms across cancer types. *Working Paper/GitHub*. <https://github.com/hossainlab/spatial-scRNA-immune-resistance>

Selected Conference Presentations

1. Ahmad, F., Fuad, M., Shahjahan, M., & **Hossain, M. J.** (2025). Identification of Subtype-Specific Prognostic Biomarkers in Esophageal Carcinoma through Integrated Bioinformatics Analysis. 9th BSM Annual Conference, University of Dhaka, Bangladesh. *GitHub*. https://github.com/gsabioinfinternship/ESCA_Meta
2. Aditee, L. M., Fuad, M., Shahjahan, M., & **Hossain, M. J.** (2025). Network-Driven Bioinformatics Approach Uncovers Prognostic Signatures and Therapeutic Targets in Pancreatic Ductal Adenocarcinoma. 9th BSM Annual Conference, University of Dhaka, Bangladesh. *GitHub*. https://github.com/gsabioinfinternship/PAAD_Meta
3. Ahamed, M. S., Fuad, M., Shahjahan, M., & **Hossain, M. J.** (2025). Bioinformatic Analysis of Transcriptomic Data Reveals Molecular Signatures Driving Colorectal Cancer. 9th BSM Annual Conference, University of Dhaka, Bangladesh. *GitHub*. https://github.com/gsabioinfinternship/COADREAD_Meta
4. Tabassum, R., Fuad, M., Shahjahan, M., & **Hossain, M. J.** (2025). Comparative Transcriptomic Meta-Analysis Reveals Shared and Distinct Transcriptional Networks and Pathway Crosstalk in Hepatocellular Carcinoma for Biomarker and Therapeutic Target Identification. 9th BSM Annual Conference, University of Dhaka, Bangladesh. *GitHub*. https://github.com/gsabioinfinternship/LIHC_Meta

Invited Talks

1. **Computational Biology and Bioinformatics Research in Resource-Limited Settings: Strategies, Tools and Opportunities** (2025). Speaker, Jagannath University Higher Study and Research Society. In this talk, I delivered the strategies and techniques essential for mastering computational biology and bioinformatics research.
2. **Mastering Biomedical Data Management** (2024). Speaker, IFMSA Bangladesh, Dhaka, Bangladesh. In the era of big data and advanced technologies, the management of biomedical data presents both immense opportunities and significant challenges. In this talk, I explored the strategies and techniques essential for mastering biomedical data management.
3. **Undergraduate Research - Importance, Benefits, and Challenges** (2022). Speaker, CHIRAL Bangladesh, Dhaka, Bangladesh. This talk was designed to enhance the undergraduate research experience by focusing on critical communication skills for success in research and broad transferable professional skills.
4. **State the Art of Microbial Genome Analysis** (2022). Speaker, Jagannath University, Dhaka, Bangladesh. This talk was about the art of genomic data analysis. I explored genomic data analysis steps typically include data collection, quality check and cleaning, processing, modeling, visualization and reporting.

Teaching Experience

Lead Organizer & Instructor

DeepBio Academy, Bioinformatics and Computational Biology

Dhaka, Bangladesh

June 2020 – Present

- Designed comprehensive bioinformatics curriculum specializing in Single-Cell RNA-seq (scRNA-seq) workflows and multi-omic data integration.
- Benchmarked computational tools for transcriptomics and genomics to ensure pipeline reproducibility and analytical rigor.
- Supervised a technical research team in processing high-throughput omics datasets to identify transcriptional drivers of disease.

- Mentored undergraduate students in the development of computational research portfolios, emphasizing open-science standards and version control.

Faculty (Bioinformatics)

University of Dhaka (cBLAST)

Dhaka, Bangladesh

Aug 2023 – Present

- Instructed specialized modules in Biomedical Machine Learning, focusing on the application of Python-based neural networks to sequence analysis.
- Directed student research projects applying deep learning frameworks to identify regulatory elements in complex biological datasets.
- Facilitated technical practicums for 100+ trainees on scalable data science methodologies for high-throughput genomic data.
- Evaluated computational frameworks for performance, establishing rigorous criteria for the optimization of biological models.
- Bridge theoretical computational principles with experimental biology to foster collaborative research outputs.

Program Lead

GSA Bioinformatics Internship

Dhaka, Bangladesh

July 2025 – Dec 2025

- Conducted a pan-cancer transcriptomics meta-analysis across four organ systems to identify conserved regulatory networks in tumorigenesis.
- Optimized high-throughput pipelines for Bulk and Single-Cell RNA-seq, implementing advanced batch-effect correction and normalization protocols.
- Mentored research interns in the utilization of High-Performance Computing (HPC) environments and R/Bioconductor analytical frameworks.
- Analyzed complex transcriptional networks using Gene Set Enrichment Analysis (GSEA) to uncover novel oncogenic drivers.
- Managed collaborative research workflows across three institutions, resulting in the generation of peer-reviewed conference abstracts.

Teaching Assistant

Department of Microbiology, Jagannath University

Dhaka, Bangladesh

Oct 2022 – Dec 2022

- Assisted in the delivery of research methodology training, focusing on experimental design and statistical power in public health microbiology.
- Processed public health surveillance data to support departmental research on emerging epidemiological trends.
- Critiqued and refined student research proposals to ensure compliance with modern computational and microbiological standards.
- Communicated complex biological theories through technical presentations to diverse academic audiences.
- Implemented digital data-tracking systems to monitor research milestones and student performance metrics.

Leadership Experience

Founder, Bioinformatics Scientist & CEO

DeepBio Limited

Dhaka, Bangladesh

Jan 2025 – Present

- Lead a team of 13 researchers executing scRNA-seq and bulk RNA-seq analysis pipelines for cancer genomics and transcriptomics projects.
- Develop and benchmark scalable bioinformatics workflows for biomarker discovery in cancer and neurodegenerative diseases.
- Build and maintain AI-integrated computational infrastructure for multi-omics data processing and interpretation.
- Develop and deliver a national training program in computational biology, upskilling undergraduate researchers across Bangladesh.

Founder & Capacity Building Director

CHIRAL Bangladesh (Non-profit Research Institute)

Dhaka, Bangladesh

June 2020 – Present

- Founded and lead a research institute with dedicated divisions in Bioinformatics, Public Health, Geospatial Health, and AI-Healthcare.
- Published 20+ peer-reviewed articles in Q1/Q2 journals and 40+ conference papers in health informatics and computational biology.
- Supervised 15+ manuscripts through analytical development and peer-review preparation in global health and bioinformatics.

- Collaborated with Michigan State University on biomedical research capacity-building initiatives in South Asia.
- Trained 3,000+ early-career researchers in data science, research methodology, and computational biology workflows.

Community Engagement

Nextflow Ambassador

Barcelona, Spain

Seqera

2025 – Present

- Advocated for the adoption of Nextflow-based DSL2 pipelines to ensure reproducibility and scalability in high-throughput single-cell and bulk RNA-seq workflows.
- Engineered technical tutorials and documentation for the deployment of portable scientific containers (Docker/Singularity) in cloud and HPC environments.
- Optimized community-driven bioinformatic workflows, focusing on the integration of standardized pipelines for large-scale transcriptomic data processing.
- Facilitated technical workshops on workflow orchestration, training researchers to manage complex multi-omic datasets with high analytical rigor.
- Collaborated with global health data scientists to implement best practices for version-controlled, open-source computational research.

GBD Senior Collaborator

Washington, USA

Institute for Health Metrics and Evaluation

2025 – Present

- Contributed advanced statistical modeling and analytical validation to Global Burden of Disease (GBD) studies published in high-impact journals like The Lancet.
- Synthesized and audited complex epidemiological datasets from 190+ countries to strengthen the robustness of global health genomic evidence.
- Formulated predictive models for disease burden, utilizing large-scale data integration techniques to identify health trends across diverse populations.
- Evaluated the quality and consistency of multi-national health data, ensuring high-fidelity inputs for global health policy frameworks.
- Collaborated with international research cohorts to refine statistical methodologies for assessing neurological and metabolic disease prevalence.

References

Dr. Syeda Tasneem Towhid

Associate Professor

Department of Microbiology

Jagannath University

Email: towhidst@mib.jnu.ac.bd

Dr. Md. Salequl Islam

Professor

Department of Microbiology

Jahangirnagar University

Email: salequl@juniv.edu

Dr. Zeba Islam Seraj

UGC Professor

Department of Molecular Biology

and Biochemistry

University of Dhaka

Director, cBLAST

Email: zebai@du.ac.bd