Md. Jubayer Hossain

Department of Public Health, Daffodil International University
Daffodil Smart City, Birulia 1216

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

Technical Skills

Programming	Python, R, Julia, SQL, Bash, JavaScript
Data Science & ML	NumPy, Pandas, PyTorch, TensorFlow, scikit-learn, XGBoost, SHAP
Bulk RNA-Seq	FastQC, STAR, HISAT2, Salmon, Kallisto, DESeq2, limma, edgeR
Single-Cell RNA-Seq	Seurat, Scanpy, CellRanger, Harmony, scVI-tools, scVelo, SingleR
Spatial Transcriptomics	Squidpy, SpatialData, Seurat (spatial), Giotto, STUtility
Microbiome	QIIME 2, DADA2, Phyloseq, VSEARCH, Kraken2, MetaPhlAn
Workflow & Reproducibility	Nextflow, nf-core, Snakemake, Docker, Singularity, Git, GitHub Actions
Visualization	ggplot2, Seaborn, Matplotlib, Plotly, ComplexHeatmap, MultiQC, Quarto
Computing	UNIX/Linux, HPC clusters, SLURM, AWS, Google Cloud, Conda, Mamba

Research Experience

Visiting Scholar *Sep 2025 – Present*
Department of Public Health, Daffodil International University *Dhaka, Bangladesh*

- 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.

Research Associate (Bioinformatics) *2017 – 2019*
Bio-Bio-1: Bioinformatics Research Discussion Group *Dhaka, Bangladesh*

- 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.

Data Enumerator *Oct 2022 – Nov 2022*
School of Allied Health, Curtin University *Perth, Australia*

- 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.

Education

Master of Science in Microbiology *2019 – 2020*
Jagannath University *Dhaka, Bangladesh*

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

Bachelor of Science in Microbiology *2016 – 2019*
Jagannath University *Dhaka, Bangladesh*

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

Training & Certifications

Interdisciplinary Computational Biology *January 2026*
BRAC University *Dhaka, Bangladesh*

- 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 *Sep 2023*
Child Health Research Foundation (CHRF) *Dhaka, Bangladesh*

- 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.

Leadership Experience

Founder, Bioinformatics Scientist & CEO *Jan 2025 – Present*
DeepBio Limited *Dhaka, Bangladesh*

- 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 *June 2020 – Present*
CHIRAL Bangladesh (Non-profit Research Institute) *Dhaka, Bangladesh*

- 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.

Teaching Experience

Lead Organizer & Instructor *June 2020 – Present*
DeepBio Academy, Bioinformatics and Computational Biology *Dhaka, Bangladesh*

- 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.

Faculty (Bioinformatics) *Aug 2023 – Present*
University of Dhaka (cBLAST) *Dhaka, Bangladesh*

- 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.

Program Lead *July 2025 – Dec 2025*
GSA Bioinformatics Internship *Dhaka, Bangladesh*

- 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.

Teaching Assistant *Oct 2022 – Dec 2022*
Department of Microbiology, Jagannath University *Dhaka, Bangladesh*

- 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.

Community Engagement

Nextflow Ambassador *2025 – Present*
Seqera *Barcelona, Spain*

- 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.
- Facilitated technical workshops on workflow orchestration, training researchers to manage complex multi-omic datasets with high analytical rigor.

GBD Senior Collaborator *2025 – Present*
Institute for Health Metrics and Evaluation *Washington, USA*

- 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.

Publications

*Note: * indicates corresponding author*

1. 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. https://doi.org/10.1038/s41598-025-13315-6

2. 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

3. Bari, S.M., Fuad, M., **Hossain, M.J.** et al. (2025). A meta-analysis of public RNA-Seq data identifies conserved stress responses in rainbow trout. *BMC Genomics 26*, 999. https://doi.org/10.1186/s12864-025-12127-2

4. Shanta, A. S., Islam, N., Al Asad, M., Akter, K., Habib, M. B., **Hossain, M. J.**, Nahar, S., Godman, B., & Islam, S. (2024). Resistance and co-resistance of metallo-beta-lactamase genes in diarrheal and urinary-tract pathogens in Bangladesh. *Microorganisms, 12*(8), 1589. https://doi.org/10.3390/microorganisms12081589

5. 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. *The Lancet*. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(25)01037-2/

6. GBD 2023 Headache Collaborators (**Hossain, M. J.**, 2025). Global, regional, and national burden of headache disorders, 1990–2023. *Lancet Neurology*. https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(25)00402-8/fulltext

7. **Hossain, M. J.***, Das, M., Shahjahan, M., Islam, M. W., & Towhid, S. T. (2025). Clinical and hematological manifestation of dengue patients in 2022 outbreak. *Health Science Reports, 8*, e70356. https://doi.org/10.1002/hsr2.70356

8. Das, M., & **Hossain, M. J.*** (2025). Young stroke in Bangladesh: Addressing rare cases with diagnostic challenges. *Stroke and Vascular Neurology*. https://svn.bmj.com/content/early/2025/04/07/svn-2025-004178

9. **Hossain, M. J.***, Sony, S. A., Fariha, F. T. J., & Hossen, S. (2025). Preventing the silent threat: A perspective on preparing Bangladesh for Human Metapneumovirus (HMPV). *Health Science Reports, 8*, e71101. https://doi.org/10.1002/hsr2.71101

10. **Hossain, M. J.***, Das, M., & Munni, U. R. (2024). Urgent call for compulsory premarital screening: A crucial step towards thalassemia prevention in Bangladesh. *Orphanet Journal of Rare Diseases, 19*, 326. https://doi.org/10.1186/s13023-024-03344-1

11. Islam, M. W., Shahjahan, M., Azad, A. K., & **Hossain, M. J.*** (2024). Factors contributing to antibiotic misuse among parents of school-going children in Dhaka City, Bangladesh. *Scientific Reports, 14*, 2318. https://doi.org/10.1038/s41598-024-52313-y

12. **Hossain, M. J.***, Das, M., Islam, M. W., Shahjahan, M., & Ferdous, J. (2024). Community engagement and social participation in dengue prevention. *Health Science Reports, 7*, e2022. https://doi.org/10.1002/hsr2.2022

13. **Hossain, M. J.***, Azad, A. K., Shahid, M. S. B., Shahjahan, M., & Ferdous, J. (2024). Prevalence, antibiotic resistance pattern for bacteriuria from patients with urinary tract infections. *Health Science Reports, 7*, e2039. https://doi.org/10.1002/hsr2.2039

14. Akter, M. M., & **Hossain, M. J.*** (2024). Food consumption patterns and sedentary behaviors among university students. *Health Science Reports, 7*, e70259. https://doi.org/10.1002/hsr2.70259

15. **Hossain, M. J.***, Islam, M. W., Munni, U. R., et al. (2023). Health-related quality of life among thalassemia patients in Bangladesh using the SF-36 questionnaire. *Scientific Reports, 13*(1). https://doi.org/10.1038/s41598-023-34205-9

16. Towhid, S.T., **Hossain, M. J.**, Sammo, M.A.S., & Akter, S. (2022). Perception of Students on Antibiotic Resistance and Prevention. *European Journal of Biology and Biotechnology*. https://doi.org/10.24018/ejbio.2022.3.3.341

17. **Hossain, M. J.**, Towhid, S.T., Sultana, S., Mukta, S.A., Gulshan, R., Miah, M.S. Knowledge and Attitudes towards Thalassemia among Public University Students in Bangladesh. *Microbial Bioactives, 5*(2). https://doi.org/10.25163/microbbioacts.526325

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. [Submitted to *Health Science Reports*].

2. Islam, M. W., Fariha, F. T. J., Ahmed, M. Z., Ferdous, J., et al. & **Hossain, M. J.*** (2025). Bioinformatics-driven multi-omics profiling of CDK1 and CDK6 identifies prognostic and therapeutic roles in breast cancer. [Under Review in *Health Science Reports*].

3. Billah, M. M., Mabsurah, K., Ahammad, K., et al. & **Hossain, M. J.*** (2025). Multi-omics pan-cancer analysis reveals an immunological role and prognostic potential of WDR76. [Under Review in *Discover Oncology*].

4. Sony, S. A., Kundu, L. R., Limon, M. H., et al. & **Hossain, M. J.*** (2025). Predicting early antenatal care initiation at the first trimester among reproductive women in Bangladesh using machine learning. [Under Review in *BMC Pregnancy and Childbirth*].

5. **Hossain, M. J.***, Shahariar, M., Barsha, L.H.J., et al. (2024). Lack of knowledge and training about antibiotic resistance among community pharmacists in Bangladesh. [Under Review in *Health Science Reports*].

Working Papers

1. **Fezf2-Mediated Cortical Development: Multi-Omics Analysis**. Integrative analysis of bulk and single-cell RNA-Seq data to elucidate Fezf2's role in cortical neurogenesis. *Tech:* Python, Seurat, Scanpy, DESeq2. GitHub

2. **Single-Cell Meta-Analysis of Microglial Activation in Alzheimer's Disease**. Meta-analysis of single-cell RNA-Seq datasets to identify microglial states and activation patterns. *Tech:* Python, Scanpy, R/Seurat, Harmony, scVI-tools. GitHub

3. **Integrative Spatial-scRNA-seq Atlas of Immunotherapy Resistance Mechanisms Across Cancer Types**. Comprehensive spatial transcriptomics and single-cell analysis to map tumor microenvironment and resistance mechanisms. *Tech:* Python, Scanpy, Seurat, Squidpy, SpatialData, scVI-tools. GitHub

Conference Presentations

Poster Presentations

1. Hossain, M. J., Das, M. (2024). Diagnostic and Electrophysiological Features of Hirayama Disease in Young Adult Male: A Case Report. *International Conference of Public Health, IEDCR, Dhaka, Bangladesh*.

2. Das, M., Kajol, M. M., Mim, N. A., et al. & **Hossain, M. J.** (2024). Level of knowledge regarding breast cancer and breast self-examination among school and college-going girls in Bangladesh. *International Conference of Public Health, IEDCR, Dhaka, Bangladesh*.

3. Shahariar, M., Joarder Barsha, L. H., Shahjahan, M., et al. & **Hossain, M. J.** (2024). Lack of knowledge and training about antibiotic resistance among community pharmacists in Bangladesh. *International Conference of Public Health, IEDCR, Dhaka, Bangladesh*.

4. Afrin, N., Howlader, G., Bhattacharjee, A., et al. & **Hossain, M. J.** (2024). Influence of Meteorological Factors on Dengue Incidence: A 23-Year Retrospective Analysis in Bangladesh. *International Conference of Public Health, IEDCR, Dhaka, Bangladesh*.

5. Akter, N., Mim, N. A., & **Hossain, M. J.** (2024). Particulate Matter PM2.5 Pollution and Air Quality Index Trends Evaluation with Meteorological Factors in Dhaka City. *International Conference of Public Health, IEDCR, Dhaka, Bangladesh*.

Oral Presentations

1. Tuhin, M. A. A., Mim, N. A., Akter, N., et al. &**Hossain, M. J.** (2024). Exploring the Factors Influencing Heat Stress Risks Using the Discomfort Index Method in Bangladesh. *ICESRM 2024, Mawlana Bhashani Science and Technology University, Bangladesh*.

2. Mim, N. A., Akter, Tuhin, M. A. A., et al. &**Hossain, M. J.** (2024). Temporal Trends and Factors Influencing Diarrhea Prevalence Among Children in the Rangamati Hill Tract Region. *ICESRM 2024, Mawlana Bhashani Science and Technology University, Bangladesh*.

3. Nayeem, M. U., Mrittika, M. A., **Hossain, M. J.**, et al. (2023). Quantitative Microbial Risk Assessment from Vancomycin-Resistant Enterococcus. *36th Bangladesh Society of Microbiologists Annual Conference, SUST, Sylhet, Bangladesh*.

4. **Hossain, M. J.**, Habiba, U., Rozario, C., et al. (2023). Factors influencing heat stress risk in Bangladesh. *9th International Public Health Conference, BMRC, Bangladesh*.

5. **Hossain, M. J.**, Nowshin, N., Momtaj, et al. (2023). Monitoring Water-Borne Disease (Vibrio cholerae) using NASA Earth observation data. *9th International Public Health Conference, BMRC, Bangladesh*.

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.
2. **Mastering Biomedical Data Management** (2024). Speaker, IFMSA Bangladesh, Dhaka, Bangladesh.
3. **Undergraduate Research - Importance, Benefits, and Challenges** (2022). Speaker, CHIRAL Bangladesh, Dhaka, Bangladesh.
4. **State the Art of Microbial Genome Analysis** (2022). Speaker, Jagannath University, Dhaka, Bangladesh.

Research Mentoring

Bulk RNA-Seq Meta-Analysis of Liver, Kidney, Pancreatic, and Gastrointestinal Cancers for Biomarker and Immune Signature Discovery (July 2025 – present)

- **Rahnuma Tabassum** (Biochemistry & Molecular Biology, Jagannath University) – Liver Cancer
- **Fayez Ahmad** (Notre Dame College) – Esophageal Cancer
- **Sabbir Khan** (Doctor of Veterinary Medicine, Gazipur Agricultural University) – Kidney Cancer
- **Md Shakil Ahamed** (Biotechnology & Genetic Engineering, Mawlana Bhashani Science & Technology University) – Colon Cancer
- **Lamisa Manha Aditee** (Biotechnology, BRAC University) – Pancreatic Cancer

Bioinformatics Workflows

1. **Bulk RNA-Seq Meta-Analysis Pipeline**. Nextflow-based pipeline for meta-analysis of public RNA-Seq data with automated QC and differential expression. *Tech:* Nextflow, nf-core, Python, R/Bioconductor, Docker. GitHub

2. **Single-Cell Meta-Analysis Pipeline**. Pipeline for single-cell RNA-Seq integration, cell type annotation, and batch correction. *Tech:* Python, Scanpy, R/Seurat, Harmony, scVI-tools. GitHub

3. **Salmon RNA-Seq Quantification Pipeline**. End-to-end pipeline for ultra-fast transcript quantification and quality assessment. *Tech:* Salmon, FastQC, MultiQC, R/Bioconductor. GitHub

4. **Biomarker Identification using Machine Learning**. ML framework for biomarker discovery with feature selection and model interpretability. *Tech:* Python, Scikit-learn, XGBoost, SHAP. GitHub

5. **GEO Differential Expression & ML Dataset Prep**. Automated GEO data retrieval, differential expression analysis, and ML dataset generation. *Tech:* R/Bioconductor, Python, GEOquery, limma, DESeq2. GitHub

6. **16S rRNA Microbiome Analysis Pipeline**. Comprehensive microbiome pipeline combining QIIME 2 and DADA2 for taxonomic profiling. *Tech:* QIIME 2, R/DADA2, Phyloseq, VSEARCH. GitHub

7. **16S-OHJU Pipeline**. Bioinformatics pipeline for processing and analyzing 16S rRNA gene sequencing data from microbial community studies. *Tech:* QIIME 2, Python, R, FastQC, MultiQC, Cutadapt. GitHub

8. **Shotgun Metagenomics Pipeline**. Comprehensive automated pipeline for shotgun metagenomic analysis from raw reads through metagenome-assembled genomes (MAGs). *Tech:* Kraken2, MetaPhlAn4, HUMAnN3, MEGAHIT, metaSPAdes, MetaBAT2, CheckM. GitHub

GenAI Projects

1. **DeepTrust AI (Prototype)**. AI-powered tool to detect and counter health misinformation with NLP and credibility scoring features. *Tech:* Python, NLP, Transformers, FastAPI, Streamlit. Demo

2. **GenMed AI (Prototype)**. AI-driven platform accelerating drug discovery through disease data analysis, protein target identification, and therapeutic compound generation. *Tech:* Python, BioPython, RDKit, DeepChem, Transformers, PyTorch. Demo

Media Coverage

1. **Parental lack of antibiotic knowledge imperils child health in Bangladesh: Study**. The Business Standard (TBS), 2024.
2. **What the rise in self-medication tells us about the country"s healthcare system**. The Business Standard (TBS), 2024.

References

References available upon request.