

Md. Jubayer Hossain

Bioinformatics Scientist · Transcriptomics (Bulk & Single-Cell RNA-Seq) and AI in Biology

132/3, Azimpur, Dhaka, Bangladesh

✉ hossainmj.me@gmail.com 🏠 mdjubayerhossain.com ☎ 0000-0001-9323-4997 🐙 GitHub 🌐 LinkedIn

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 neurodegenerative diseases. By integrating multi-omics data analysis, machine learning, and biology, I aim to advance precision medicine through interpretable and scalable computational frameworks that bridge the gap between 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

Specialized Training: AI in Public Health

Child Health Research Foundation (CHRF)

Dhaka, Bangladesh

2023

- Acquired Skills: Epidemiological Modeling, Disease Surveillance using AI, Time Series Analysis (LSTM, ARIMA), Medical Image Analysis (X-ray, MRI)

Research Experience

Visiting Researcher

Department of Public Health, Daffodil International University

Dhaka, Bangladesh

2025– Present (Part-time)

- Collaborated with DIU faculty and students on ongoing research projects, grant proposals, and publications in public health and health informatics, resulting in 2 manuscripts under preparation.
- Delivering workshops, online courses, and research mentorship to strengthen capacity building in public health research.

GBD Collaborator

Institute for Health Metrics and Evaluation

Washington, USA

2025– Present (Part-time,
Voluntary)

- Contributed critical analytical feedback and statistical modeling to 4 Global Burden of Disease (GBD) studies, including one on routine childhood vaccination coverage published in The Lancet.
- Supported manuscript development and provided methodological input for 3 ongoing high-impact GBD papers (breast cancer, headache disorders, and mental disorders), all currently under peer review.
- Enhanced epidemiological insights by integrating and validating datasets from 190+ countries, improving the robustness of global health policy evidence.

Research Intern

International Center for Diarrhoeal Disease Research, Bangladesh

Dhaka, Bangladesh

Feb 2022–Jun 2022

- Designed and implemented a comprehensive search strategy across PubMed, Embase, and PsycINFO, screening 8,000+ articles to support a systematic review on intimate partner violence in South Asia.
- Streamlined the abstract screening process in Covidence, identifying 400+ potentially relevant studies.
- Conducted full-text screening of 120+ articles and extracted key variables for meta-analysis.

- Drafted a systematic review protocol based on PRISMA guidelines, ensuring methodological transparency and reproducibility.

Data Enumerator

Australia

School of Allied Health, Curtin University

Oct 2022–Nov 2022

- Collected both quantitative and qualitative data from 40 Bangladeshi doctors for a Delphi study on Patient Outcomes Measurement in Interprofessional Tuberculosis Care.
- Validated and cleaned all survey responses, achieving 100% accuracy in the final dataset.
- Enhanced the reliability of study outcomes by ensuring high-quality and complete data capture.

Team Member

Australia

Bio-Bio-1: Bioinformatics Research Discussion Group

2017–2019

- Engaged in collaborative research discussions and pursued continuous learning opportunities in bioinformatics.
- Contributed to the design of training modules and preparation of learning content to support group activities.
- Took initiative in team projects and developed leadership skills through active participation.

Leadership Experience

Founder & Chief Executive Officer

Dhaka, Bangladesh

DeepBio Limited (C-202963/2025)

2025 - Present (Self-employed)

- Founded and lead a bioinformatics company specializing in cancer genomics, transcriptomics, and AI-driven healthcare solutions.
- Recruited and manage a multidisciplinary team of 13 scientists (4 CADD Scientists I, 5 Bioinformatics Scientists I, 4 Computational Biologists) to drive computational biology and bioinformatics research.
- Developed and implemented scalable bulk RNA-seq and single-cell analysis pipelines, enabling translational medicine and biomarker discovery for cancer and neurodegenerative diseases.
- Launched a professional training program in bioinformatics and computational biology, training 40 students across Bangladesh, Pakistan, India, and African countries to strengthen the regional talent pipeline.
- Established collaborations with academic and industry partners to expand computational biology research capacity in South Asia.

Founder & Executive Director

Dhaka, Bangladesh

Center for Health Innovation, Research, Action, and Learning - Bangladesh

2020 - Present (Self-employed)

- Founded and expanded a non-profit research institute empowering young scientists in public health, bioinformatics, geospatial health, and AI for health, now structured into 4 research divisions.
- Trained over 3,000 undergraduate students nationwide on the fundamentals of research methods, data science, and scientific writing.
- Published 20+ peer-reviewed articles in international journals (Q1, Q2) and presented 40+ conference papers, advancing knowledge in public health, bioinformatics, and health informatics.
- Guided the preparation of 15+ manuscripts currently under peer review, strengthening the pipeline of early-career researcher contributions.
- Secured collaborative partnerships with Thalassemia Foundation Bangladesh, One Health Lab, Jahangirnagar University and Sher-e-Bangla Agricultural University, and Michigan State University to foster capacity building in biomedical research across Bangladesh.

Publications

Note: * indicates corresponding author

Bioinformatics & Computational Biology

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 (2025). <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. 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>
4. Bari, S. M., Fuad, M., & **Hossain, M. J.*** (2025). Transcriptomic insights into heat and hypoxia stress in rainbow trout: A meta-analysis of public RNA-Seq data [Preprint]. *Research Square*. (Accepted by *BMC Genomics*) <https://doi.org/10.21203/rs.3.rs-6710656/v1>

Global Health

1. GBD 2023 Vaccine Coverage Collaborators. (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/)
2. **Hossain, M. J.***, Das, M., Shahjahan, M., Islam, M. W., & Towhid, S. T. (2025). Clinical and hematological manifestation of dengue patients in 2022 outbreak: A tertiary care hospital-based cross-sectional study. *Health Science Reports*, 8, e70356. <https://doi.org/10.1002/hsr2.70356>
3. Das, M., & **Hossain, M. J.*** (2025). Young stroke in Bangladesh: Addressing rare cases with diagnostic challenges and much-needed solutions. *Stroke and Vascular Neurology*. <https://svn.bmj.com/content/early/2025/04/07/svn-2025-004178>
4. **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>
5. **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>
6. 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>
7. **Hossain, M. J.***, Das, M., Islam, M. W., Shahjahan, M., & Ferdous, J. (2024). Community engagement and social participation in dengue prevention: A cross-sectional study in Dhaka City. *Health Science Reports*, 7, e2022. <https://doi.org/10.1002/hsr2.2022>
8. **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>
9. Akter, M. M., & **Hossain, M. J.*** (2024). Food consumption patterns and sedentary behaviors among the university students: A cross-sectional study. *Health Science Reports*, 7, e70259. <https://doi.org/10.1002/hsr2.70259>
10. **Hossain, M. J.***, Islam, M. W., Munni, U. R., Gulshan, R., Mukta, S. A., Miah, M. S., Sultana, S., Karmakar, M., Ferdous, J., & Islam, M. A. (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>
11. Towhid, S.T., **Hossain, M. J.**, Sammo, M.A.S., & Akter, S. Perception of Students on Antibiotic Resistance and Prevention: An Online, Community-Based Case Study from Dhaka, Bangladesh. *European Journal of Biology and Biotechnology*, <https://doi.org/10.24018/ejbio.2022.3.3.341>
12. **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>

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., 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 in Health Science Reports].
3. Billah, M. M., Mabsurah, K., Ahammad, K., Yeana, I. J., Sumaiya, M., Islam, T., Bhattacharjee, A., Ferdous, J., & **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., Chowdhury, T. B. K., & **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. Shahjahan, M., Towhid, ST., Sheikh, MK., Hasan, MM., Rahman, MH., Sazid, MS. (2024). Lack of knowledge and training about antibiotic resistance among community pharmacists in Bangladesh: a cross-sectional study. [Under Review in Health Science Reports].

Conferences

*Note: *Indicates a presenter, #indicates a co-author, and * indicates a corresponding author*

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, Institute of Epidemiology, Disease Control and Research (IEDCR), Dhaka, Bangladesh
2. *Das, M., Kajol, M. M., Mim, N. A., Maruf, M. F. I., Hirko, K., ***Hossain, M. J.** (2024). Level of knowledge regarding breast cancer and breast self-examination among school and college-going girls in Bangladesh: A cross-sectional study. International Conference of Public Health, Institute of Epidemiology, Disease Control and Research (IEDCR), Dhaka, Bangladesh
3. Shahariar, M., Joarder Barsha, L. H., *Shahjahan, M., Towhid, S. T., Sheikh, M. K., Hasan, M. M., Rahaman, M. H., Sazid, M. S., ***Hossain, M. J.** (2024). Lack of knowledge and training about antibiotic resistance among community pharmacists in Bangladesh: A cross-sectional study. International Conference of Public Health, Institute of Epidemiology, Disease Control and Research (IEDCR), Dhaka, Bangladesh
4. *Afrin, N., Howlader, G., Bhattacharjee, A., Akter, N., Mesu, M. H., Mazumder, E., ***Hossain, M. J.** (2024). Influence of Meteorological Factors on Dengue Incidence: A 23-Year Retrospective Analysis in Bangladesh. International Conference of Public Health, Institute of Epidemiology, Disease Control and Research (IEDCR), Dhaka, Bangladesh
5. *Akter, N., Mim, N. A., ***Hossain, M. J.** (2024). Particulate Matter PM2.5 Pollution and Air Quality Index (AQI) Trends Evaluation with Meteorological Factors in Dhaka City. International Conference of Public Health, Institute of Epidemiology, Disease Control and Research (IEDCR), Dhaka, Bangladesh

Oral Presentations

1. *Tuhin, M. A. A., Mim, N. A., Akter, N., Jamal, F., Sakib, M. N., Mehjabin, K., ***Hossain, M. J.** (2024). Exploring the Factors Influencing Heat Stress Risks Using the Discomfort Index Method in Bangladesh. International Conference on Environmental Science and Resource Management 2024 (ICESRM 2024), Department of Environmental Science and Resource Management, Mawlana Bhashani Science and Technology University, Bangladesh.
2. Mim, N. A., Akter, Tuhin, M. A. A., *Hossain, S., Alam, K. E., Mouli, N. J., ***Hossain, M. J.** (2024). Temporal Trends and Factors Influencing Diarrhea Prevalence Among Children in the Rangamati Hill Tract Region: A Time Series Analysis. International Conference on Environmental Science and Resource Management 2024

(ICESRM 2024), Department of Environmental Science and Resource Management, Mawlana Bhashani Science and Technology University, Bangladesh.

3. Nayeem, M. U., Mrittika, M. A., ***Hossain, M. J.**, Azad, A. K., Ferdous, J., Ahmed, S., Sanyal, S. K., Towhid, S. T. (2023). Quantitative Microbial Risk Assessment from Vancomycin-Resistant *Enterococcus faecalis* and *Enterococcus faecium* from a Specific Neighborhood in Dhaka City, Bangladesh. 36th Bangladesh Society of Microbiologists Annual Conference, Shahjalal University of Science and Technology (SUST), Sylhet, Bangladesh, January 2023.
4. ***Hossain, M. J.**, *Habiba, U., Rozario, C., Tuhin, M. L. A., Mabsurah, K., Uddin, N. F. (2023). Factors influencing heat stress risk in Bangladesh. 9th International Public Health Conference by Public Health Foundation at Bangladesh Medical Research Council (BMRC), Bangladesh, November 2023.
5. ***Hossain, M. J.**, *Nowshin, N., Momtaj, Yeana, I. Y., Nohor, N. (2023). Monitoring Water-Borne Disease (*Vibrio cholerae*) using NASA Earth observation data. 9th International Public Health Conference by Public Health Foundation at Bangladesh Medical Research Council (BMRC), Bangladesh, November 2023.

Teaching

Instructor

Dhaka, Bangladesh

Micro-Credentials Academy, Daffodil International University

2025 - Present (Part-time)

- Designed and delivered modules on AI for disease modeling, digital health, and biomedical informatics
- Coordinated training programs that reached diverse student groups, bridging computing and health sciences

Lead Organizer & Instructor

Dhaka, Bangladesh

Professional Training in Bioinformatics and Computational Biology

2025 - Present (Part-time)

- Developed a year-long, university-style program covering genomics, transcriptomics, single-cell RNA-seq, and AI in health, with 40 students enrolled from Bangladesh, Pakistan, India, and African countries.
- Led a team of 4 instructors and 8 teaching assistants and created real-world case study curricula, preparing students for professional research careers in bioinformatics and computational biology.

Program Lead

Dhaka, Bangladesh

GSA Bioinformatics Internship

2025 - Present (Part-time)

- Designed and taught a hands-on bioinformatics curriculum (Bash, R, RNA-seq, scRNA-seq) for interns annually.
- Mentored 5 research interns on an original project: Comparative pan-cancer transcriptomics meta-analysis, which uncovered shared and distinct transcriptional networks with pathway crosstalk in liver, kidney, pancreas, and gastrointestinal cancers for biomarker and therapeutic target discovery.
- Guided interns in RNA-seq meta-analysis, coding, and scientific communication, leading to conference abstracts and manuscript preparation.

Faculty

Dhaka, Bangladesh

Center for Bioinformatics Learning Advancement and Systematics Training, University of Dhaka

2023 - Present (Part-time)

- Designed and delivered specialized courses in Biomedical Machine Learning with Python (20 trainees), Data Science and Machine Learning for Biologists (100+ enrolled), and School of Bioinformatics (15 trainees).
- Supervised student research projects applying ML frameworks to biomedical datasets, guiding participants toward publication-ready outputs and conference presentations.

Teaching Assistant

Dhaka, Bangladesh

Department of Microbiology, Jagannath University

Oct 2022 - Dec 2022 (Part-time)

- Assisted in preparing and delivering lectures, course materials, and assessments for 40 undergraduate students in Public Health & Hygiene.
- Mentored students on a public health project, providing guidance on research design, data analysis, and presentation skills.
- Supported faculty in student evaluation and academic advising, fostering stronger student engagement with departmental activities.

Mentoring

1. **Rahnuma Tabassum** (July 2025 - present), Biochemistry and Molecular Biology, Jagannath University: “Bulk RNA-Seq Meta-Analysis of Liver, Kidney, Pancreatic, and Gastrointestinal Cancers for Biomarker and Immune Signature Discovery” (Liver)
2. **Fayez Ahmad** (July 2025 - present), Notre Dame College: “Bulk RNA-Seq Meta-Analysis of Liver, Kidney, Pancreatic, and Gastrointestinal Cancers for Biomarker and Immune Signature Discovery” (Esophageal)
3. **Sabbir Khan** (July 2025 - present), Doctor of Veterinary Medicine, Gazipur Agricultural University: “Bulk RNA-Seq Meta-Analysis of Liver, Kidney, Pancreatic, and Gastrointestinal Cancers for Biomarker and Immune Signature Discovery” (Kidney)
4. **Md Shakil Ahamed** (July 2025 - present), Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University: “Bulk RNA-Seq Meta-Analysis of Liver, Kidney, Pancreatic, and Gastrointestinal Cancers for Biomarker and Immune Signature Discovery” (Colon)
5. **Lamisa Manha Aditee** (July 2025 - present), Biotechnology, BRAC University: “Bulk RNA-Seq Meta-Analysis of Liver, Kidney, Pancreatic, and Gastrointestinal Cancers for Biomarker and Immune Signature Discovery” (Pancreatic)

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.

Projects

DeepTrust AI (Prototype)

<https://deeptrustai.deepbioltd.com/>

- Led the development of DeepTrust AI, an AI-powered tool designed to detect and counter health misinformation, addressing a critical public health challenge in Bangladesh.
- Built and showcased a functional prototype with natural language processing and credibility scoring features.

GenMed AI (Prototype)

<https://genmedai.deepbioltd.com/>

- Directed the design of GenMed AI, an AI-driven platform to accelerate drug discovery by analyzing disease data, identifying protein targets, and generating therapeutic compound designs within minutes instead of years.
- Developed prototype models integrating bioinformatics pipelines and generative AI frameworks.

Differential Expression Analysis and ML Dataset Preparation

https://github.com/bigbiolab/geo_ml

- Download GEO expression datasets programmatically
- Perform differential expression analysis with DESeq2

- Export machine learning-ready dataset for classification or feature selection

Biomarker Identification using Machine Learning

https://github.com/hossainlab/biomarker_identification

- Analyze high-dimensional TCGA genomic data from cancer samples
- Apply machine learning algorithms for feature selection and biomarker discovery

Differential Expression Analysis and ML Dataset Preparation

https://github.com/bigbiolab/geo_ml

- Download GEO expression datasets programmatically
- Perform differential expression analysis with DESeq2
- Export machine learning-ready dataset for classification or feature selection

End-to-end Bulk RNA-Seq Quantification Pipeline using Salmon

<https://github.com/hossainlab/salmon-rnaseq>

- Automatically download and quantify GEO expression datasets
- Perform differential expression analysis using tximport and DESeq2

Skills

Programming & Data Science

Python, R, Julia, SQL; PyTorch, TensorFlow, scikit-learn, PyCaret

Bioinformatics

Bulk RNA-seq (FastQC, STAR, HISAT2, Salmon, Kallisto, DESeq2); Single-cell RNA-seq (Seurat, Scanpy, CellRanger); Nextflow, Docker

Multi-omics Integration

Harmony (integration)

Other Tools

UNIX/Linux, Git

References

Dr. Syeda Tasneem Towhid

Associate Professor
Department of Microbiology
Jagannath University
Email: towhidst@mib.jnu.ac.bd

Dr. A B M Alauddin Chowdhury

Professor & Head
Department of Public Health
Daffodil International University
Email:
headph@daffodilvarsity.edu.bd
dralauddin@daffodilvarsity.edu.bd

Dr. Zeba Islam Seraj

UGC Professor
Department of Molecular Biology
and Biochemistry
University of Dhaka
Director, cBLAST
Email: ziseraj@gmail.com