Md. Masud Rana

Address: House #691, Natore road, Ghoramara-6100, Rajshahi, Bangladesh
Telephone (Mobile): +880 1713846449 Email: ringku_740@yahoo.com

Education:

PhD in Bioinformatics	National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences (CAS) & China National Center for Bioinformation (CNCB), Beijing, China	2017-2020 2023-2024
MSc in Statistics	Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh (3.750 in a scale of 4)	2013-2014
BSc in Statistics	Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh (3.401 in a scale of 4)	2008-2013

Research Experience:

National Genomics Data Center, Beijing Institute of Genomics, CAS & CNCB, Beijing, China **PhD Fellow (CAS-TWAS President's Fellowship)**	**Supervisor: Prof. LIU Fan & Prof. ZHANG Zhang**	2017-2020 2023-2024

Novel methods for identifying biomarkers of drug-induced toxicity in complex toxicogenomics datasets

- Developed novel statistical models, algorithms, and the software tool 'ToxAssay', which automates data management and post-marker discovery analyses to efficiently identify biomarkers and uncover molecular mechanisms of drug-induced toxicity. This method identified core genes and adverse outcome pathways, revealing key mechanisms underlying glutathione depletion.
- Developed the algorithm 'ALVA', integrating statistical and machine learning techniques to prioritize compound targets based on perturbed gene expression. ALVA identified targets related to carcinogenesis, enhancing safety assessments and contributing to the development of targeted therapies for drug-induced hepatotoxicity and cancer.
- Developed a robust stacking ensemble learning (RSEL) algorithm to classify imbalanced data for predicting chemically induced liver toxicity endpoints with high accuracy.

DNA based facial recognition

- We successfully integrated DNA phenotyping with computer vision biometric processes to evaluate the feasibility of DNA-based facial recognition.
- Developed a novel metric learning method, 'WSML,' to match observed phenotypes with predicted phenotypes from DNA, enabling the identification of unknown individuals. This method was evaluated on both simulated and real datasets, outperforming existing distance metric learning approaches.

HEQEP sub-project (CP-3603, W2 and R3), Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh **Research Assistant**	**Supervisor: Prof. Md. Nurul Haque Mollah**	2015-2017

Statistical modeling for omics data analysis in bioinformatics

- We developed a novel robust copula-based statistical model for multi-phenotype GWAS analysis, accounting for phenotype dependence as a function of genetic variation. The model, evaluated through simulation studies, outperformed single-trait GWAS and other multi-trait models.
- Contributed to a team analyzing microbiome data, including processing datasets into phyloseq objects, modeling with over-dispersed and zero-inflated models (Poisson, NB, ZIP, ZINB), and conducting diversity analysis using Vegan. Additionally, evaluated microbial community composition, differential abundance, and associations between clinical metadata and microbial features using MaAsLin2.
- Contributed to a team developing the logistic probabilistic hidden variable model (LPHVM) to co-cluster compound doses and potential target genes. We demonstrated that LPHVM outperformed hierarchical clustering-based co-clustering methods in identifying more accurate gene-compound dose clusters.
- Personally extended the hierarchical naïve Bayes model through mathematical derivation for three-level data classification.

Bioinformatics Lab (Dry), Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh	**Supervisor: Prof. Md. Nurul Haque Mollah**	2013-2015

MSc thesis student

Discovery of toxicogenomic biomarkers using the multistage hierarchical nested ANOVA approach

In my thesis, I proposed a multistage hierarchical nested ANOVA model to analyze toxicogenomics data and identify biomarker genes. I developed a ranking method to pinpoint the toxic dose of compounds and successfully identified genes involved in the glutathione metabolism pathway responsible for oxidative stress. These findings are part of a patent application for novel therapeutic targets in drug-induced toxicity.

Publications:

Rana MM, Mollah MNH, Albujja MH, Hadi SS and Liu F (2024) ToxAssay: A hierarchical model-driven tool for advanced toxicogenomics biomarker discovery. *Bioinformatics* (under review)

Rana MM, Hasan MN, Hossen MA, Albujja MH, Hadi SS, Mollah MNH and Liu F (2024) Prioritizing compound targets by perturbed gene expression using ALVA refines biomarkers in carcinogenesis (Submitted to *BMC Bioinformatics*)

Rana MM et al. A robust copula-based statistical model for multi-trait genome-wide association studies (Manuscript prepared and under review by co-authors, expected submission December 2024.)

Hossen MA, Reza MS, **Rana MM**, Hossen MB, Shoaib M, Mollah MNH, Han C (2024) Identification of most representative hub-genes for diagnosis, prognosis, and therapies of hepatocellular carcinoma. *Chin Clin Oncol* 13(3):32. doi: 10.21037/cco-23-151

Hasan MN, **Rana MM**, Begum AA, Rahman M and Mollah MNH (2018) Robust co-clustering to discover toxicogenomic biomarkers and their regulatory doses of chemical compounds using logistic probabilistic hidden variable model. *Front. Genet.* 9:516. doi: 10.3389/fgene. 2018.00516

Rana MM, Hasan MN, Ahmed MS and Mollah MNH (2018) A novel computational approach for toxicogenomics biomarker discovery in drug development pipeline. *J. Bio-Sci.* 25: 57-66. doi: 10.3329/jbs.v25i0.37499

Ahmed MS, Kamruzzaman M, **Rana MM**, Akond Z and Mollah MNH (2018) In Silico Analyses of Human Collagen Protein Function Prediction. *J. Bio-Sci.* 24: 55-65. doi: 10.3329/jbs.v24i0.37487

Ahmed MS, Shahjaman M, **Rana MM** and Mollah MNH (2017) Robustification of naïve bayes classifier and its application for microarray gene expression data analysis. *Biomed Res. Int.*, 2017: 3020627. doi: 10.1155/2017/3020627

Conference Proceedings:

Rana MM, Hasan MN, Ahmed MS and Mollah MNH (2017) Toxicogenomics biomarker detection for liver toxicity using multilevel hierarchical modeling. *International Conference on Bioinformatics and Biostatistics for Agriculture, Health and Environment*, Department of Statistic, University of Rajshahi & BBCBA, Bangladesh. pp:77-82.

Rana MM, Singha AC, Ahmed MS and Mollah MNH (2015) Identification of the toxicogenomic biomarker using the multistage hierarchical ANOVA approach, *J Drug Metab Toxicol*, 6, 71.

Ahmed MS, Singha AC, **Rana MM**, Mollah NMH (2015) Bayesian Approach for Prediction of Interface and Non-interface Residues from Protein Sequence. *International Conference on Materials, Electronics &Information Engineering, ICMEIE-2015*, ISBN 978-984-33-894—4.

Presentation:

Invited

- *Training workshop on R statistical programing language, Chittagong Veterinary and Animal Sciences University, Bangladesh.* “Introduction to R, the powerful tool for statistical computing, data analysis and graphic”, February, 2017.
- *2nd National Workshop on Bioinformatics and Computational Biology, Chittagong Veterinary and Animal Sciences University & University of Science and Technology, Chittagong, Bangladesh.* “Introduction to computational toxicogenomics and hands-on training”, December, 2016.

Oral

- *International Conference on Bioinformatics and Biostatistics for Agriculture, Health and Environment, Department of Statistic, University of Rajshahi & BBCBA, Bangladesh*. Toxicogenomics biomarker detection for liver toxicity using multilevel hierarchical modeling. **Rana MM**, Hasan MN, Ahmed MS and Mollah MNH, January, 2017.
- *International Conference on Analysis of Repeated Measures Data, Department of Applied Statistics, East West University, Bangladesh*. Identification of chemically-induced liver toxicity genomic biomarkers using factorial ANOVA approach. **Rana MM**, Hasan MN, Ahmed MS and Mollah MNH, November, 2016.
- *The 2nd international conference on theory and applications of statistics, Dhaka University, Bangladesh*. Nested ANOVA-Simultaneous component analysis (NASCA): An extended tool of ANOVA-SCA for toxicogenomic biomarker discovery. **Rana MM**, Hasan MN, Singha AC, Ahmed MS, and Mollah MNH, December, 2015.
- *International Conference on Applied Statistics, ISRT, University of Dhaka, Bangladesh.* Discovery of the toxicogenomic biomarker using the multistage experimental nested design. **Rana MM**, Singha AC, Ahmed MS, and Mollah MNH, December, 2014.

Poster

- *International Conference on Bioinformatics and Biostatistics for Agriculture, Health and Environment, Department of Statistic, University of Rajshahi & BBCBA*. Robust prediction of human genes expression profile using rat model in japanese toxicogenomics project. Akefa S., **Rana MM**, Hasan MN, Hosen A and Mollah MNH, January, 2017.

- *International Conference on Biological Sciences: Food, Health and Environmental Perspectives, IBSc, University of Rajshahi, Bangladesh*. Identification of genomic biomarkers of detoxification organ toxicity using hierarchical ANOVA approach. **Rana MM**, Singha AC, Ahmed MS and Mollah MNH, August, 2015.

Software Packages:

- ToxAssay (https://github.com/ringku09/toxassay)
- ALVA (https://github.com/ringku09/alva)
- WSML (https://github.com/ringku09/wsml)
- HNB (https://github.com/ringku09/nbm)
- RSEL (https://github.com/ringku09/rsel)

Honors and Awards:

- **CAS-TWAS President's Fellowship Programme for Doctoral Candidates**, 2017.
- **Third Place** in Poster Presentation Competition at the *International Conference on Biological Sciences: Food, Health, and Environmental Perspectives*, August 8, 2015, Institute of Biological Sciences, University of Rajshahi, Bangladesh.
- **First Place** in the *Statistical Computing Competition on C++*, February 13, 2013, Department of Statistics, University of Rajshahi, Bangladesh.

Activities and Membership:

- **Life Member**, Bangladesh Bioinformatics and Computational Biology Association (BBCBA).
- **Life Member**, Bioinformatics Research Group of Rajshahi University (BioRGRU), University of Rajshahi, Bangladesh.
- **Member**, Statistical Association, Department of Statistics, University of Rajshahi, Bangladesh.
- **Sports Secretary** (2012-2013), Statistical Association, Department of Statistics, University of Rajshahi, Bangladesh.

Professional Skills Training:

- Awarded a certificate for completing training in "**2019 Winter School on Frontier and Interdisciplinary Sciences**", International College, University of Chinese Academy of Sciences, Pingtang, Qiannan, Guizhou, China (January 10-15, 2019).
- Awarded a certificate for completing training in "**R-package/software development**", Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh (June 2, 4-5, 2016).
- Awarded a certificate for completing hands-on training in "**Python for Bioinformatics**", Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh (March 26-27, 2016).
- Awarded a certificate for completing training in "**Perl Programming and Online Software Development for Bioinformatics**", University of Rajshahi, Bangladesh (July 31- Feb 1, 2015).
- Certificate in Computer Applications: MS Word, MS Excel, MS PowerPoint, MS-DOS, and LaTeX, SPSS, MAPLE, STATA.

Additional Employment:

- **Data Scientist (Freelance)**, Freelancer, Sydney, Australia (2021-2022).

- **Additional Class Teacher (Mathematics)**, Secondary Education Quality and Access Enhancement Project, Bangladesh (Supported by the World Bank) (2015-2016).

Personal Memoranda:

- Date of Birth: November 16, 1989.
- Nationality: Bangladesh (by birth).
- Religion: Islam.
- Marital Status: Married.

References:

- **Dr. Fan Liu**, Full Professor at Naif Arab University for Security Sciences, Riyadh, Saudi Arabia and Visiting professor at Erasmus University Medical Center, Netherland, E-mail: fliu@nauss.edu.sa, Tel: +966 545114868.
- **Dr. Zhang Zhang**, Professor & Associate Director of National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences & China National Center for Bioinformation, Beijing, China. E-mail: zhangzhang@big.ac.cn, Tel: +86 (10) 8409-7261.
- **Dr. Md. Nurul Haque Mollah**, Professor, Bioinformatics Lab (Dry), Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh. E-mail: mollah.stat.bio@ru.ac.bd, Tel: +880-721-711275.