

# Nishat Anjum Bristy

Phone: +14123203654

Email: [nbristy@andrew.cmu.edu](mailto:nbristy@andrew.cmu.edu)

## Research Interest

Computational Biology, Tumor Biology, Phylogenetics, Algorithms, Machine learning

## Education

- 2022 - ongoing **PhD. in Computational Biology (Ongoing)**  
Advisor: Russell Schwartz  
Carnegie Mellon University (CMU-Pitt joint Computational Biology Program)
- 2016-2020 **Bachelors in Computer Science and Engineering**  
Bangladesh University of Engineering and Technology  
CGPA : 3.88/4.00, Position: 7/143. (**Major GPA: 3.94/4.00**)

## Publications

### PREPRINTS

- 2023 **Nishat Anjum Bristy**, Xuecong Fu, and Russell Schwartz. "Sc-TUSV-ext: Single-cell clonal lineage inference from single nucleotide variants (SNV), copy number alterations (CNA) and structural variants (SV)." bioRxiv: the preprint server for biology (2023). [<https://pubmed.ncbi.nlm.nih.gov/38106049/>]

### JOURNAL ARTICLES

- 2019 Nazifa Ahmed Moumi, Badhan Das, Zarin Tasnim Promi, **Nishat Anjum Bristy**, and Md Shamsuzzoha Bayzid. "Quartet-based inference of cell differentiation trees from chip-seq histone modification data." Plos one 14, no. 9 (2019): e0221270. [<https://doi.org/10.1371/journal.pone.0221270>]
- 2022 Metin Balaban\*, **Nishat Anjum Bristy\***, Ahnaf Faisal, Md. Shamsuzzoha Bayzid, and Siavash Mirarab. "Genome-wide alignment-free phylogenetic distance estimation under a no strand-bias model." Bioinformatics advances 2, no. 1 (2022): vbac055. [<https://doi.org/10.1093/bioadv/vbac055>]

## Accepted Posters and Talks

### POSTERS

- AACR 2024 Integrating single nucleotide variants (SNVs), copy number alterations (CNAs), and structural variants (SVs) into single-cell clonal lineage inference.

### TALKS

- RECOMB-CCB 2024 Sc-TUSV-ext: Single-cell clonal lineage inference from single nucleotide variants (SNV), copy number alterations (CNA) and structural variants (SV)
- GLBio 2024 Reconstructing tumor evolutionary histories from single-cell sequencing data incorporating single nucleotide variants (SNV), copy number aberrations (CNA) and structural variants (SV)

## Academic services

- Sub-reviewer RECOMB (2024), RECOMB-CCB (2024)

## Research Experience

- January 2023 - ongoing **Working with Dr. Russell Schwartz on Sc-TUSV-ext: Single-cell clonal lineage inference from single nucleotide variants (SNV), copy number alterations (CNA) and structural variants (SV).**  
(Status: **Manuscript Prepared**, Keywords: Single-Cell Sequencing, Phylogenetics, Cancer, Somatic Variation, Integer Linear Programming)

- Clonal lineage inference is a crucial tool for making sense of cancer development and progression.
- We develop a integer linear programming framework to infer tumor evolutionary trees for single-cell DNA sequencing data with the help of single nucleotide variants, copy number alterations and structural variations.
- In both synthetic and real datasets, we show that incorporating three types of variants help understand the clonal relationships better.

2017 - 2019

**Worked with Dr. Shamsuzzoha Bayzid on**

**Quartet-based inference of cell differentiation trees from ChIP sequence histone modification data**

(Status: **Published (PLOS ONE)**, Keywords: Computational biology, Phylogenetics)

- We leveraged the quartet-based phylogenetic tree estimation techniques to construct cell-type trees.
- We included both normal and cancerous cell types to expand the scope of our work.
- We emphasized on the strength of the standard phylogenetic tree estimation approaches in replacing the traditional intricate *in vitro* methods which explains the relationships among different cell-types.

2019 - 2020

**Worked with Dr. Shamsuzzoha Bayzid on**

**Species tree estimation from Triplet and Quartet trees included in gene trees**

(Status: **Finished**, Keywords: Computational biology, Phylogenetics)

- We developed a highly accurate statistically consistent coalescent based method for estimating species trees from a collection of gene trees based on the number of triplets and quartets.
- We modified the codebase of ASTRAL to support rooted triplets as well as unrooted quartets.

2019-2021

**Working with Dr. Siavash Mirarab and Dr. Shamsuzzoha Bayzid on**

**Alignment-free phylogenetic distance estimation under a no strand-bias model model**

(Status: **Published (Bioinformatics advances)**, **Co-first author**, Keywords: Computational biology, Phylogenetics)

- We worked on estimating phylogenetic distances from unaligned sequences, under the 4 parameter time-reversible TK4 model, using a k-mer based approach.
- We have developed a novel technique which is able to deal with mixed-strand data.
- Our model is expected to overcome the limitations of the Jukes-Cantor model for unaligned sequences and show promising accuracy for larger distances.
- This work is a collaboration with University of California, San Diego.

November,

2020 - 2022

**Worked with Dr. Atif Hasan Rahman on**

**An RNN based method for the classification of HIV-1 genomes**

(Status: **Finished**, Keywords: Computational biology, Deep learning)

We proposed an alignment-free subtyping method using RNNs that operates on both time series analysis of the individual character and k-mer frequencies in HIV-1 sequences.

May, 2021

2022

**Worked with Dr. Atif Hasan Rahman on**

**A deep learning based tool for sequence contamination detection**

(Status: **Finished**, Keywords: Computational biology, Deep learning)

We are leveraging the effectiveness of deep learning using a novel metagenomic binning approach, to model the co-occurrence of contigs belonging to the same genome.

May, 2020

2022

**Worked with Dr. Tanzima Hashem on 2 projects - Detecting malnutrition and dehydration from facial images, and A survey on different smartphone image analysis techniques for disease detection**

(Status: **Finished**, Keywords: Machine Learning, HCI)

We worked on detecting malnutrition and dehydration from images of different body parts (face, eyes, lips, etc.) that show significant variation in color and texture due to the disease. Our goal was to build an easily accessible machine learning based smartphone application, which will be able to detect dehydration and malnutrition levels. This work is a collaboration with [International Centre for Diarrhoeal Disease Research, Bangladesh](#).

## Work Experience

### Lecturer

BRAC University (July, 2021 - July, 2022)

### Adjunct Lecturer

Bangladesh University of Engineering and Technology  
(February 2022 - May, 2022)

**Research Assistant** (February 2021 - December, 2021)

**Research lab of Dr. Shamsuzzoha Bayzid**

Department of CSE, BUET

### Volunteer

**Academic Co-ordinator** (2020 - 2021)

**Academic team member** (2015 - 2020)

Bangladesh Mathematical Olympiad (BdMO)

**Accademic Coordinator**

December,

2018

## Grants, honors & awards

2024	<b>CMLH Fellowship</b> in Digital Health Innovation
2024	GLBIO 2024 Travel Fellowship
2020	Winner of the first round of the <b>Sanger Institute Prize Competition, 2020</b> , arranged by Wellcome Sanger Institute
2019, 2020	<b>University merit scholarship</b> for academic excellence, BUET
2019, 2020	<b>Deans List Award</b> , BUET
2011, 2013-14	National round champion at <b>Bangladesh Mathematical Olympiad</b> and <b>Bangladesh Physics Olympiad</b>
2013	Divisional champion at <b>National Creative Talent Hunt</b> , Bangladesh
2010	Received an award from the honorable president of Bangladesh, Zillur Rahman, for a <b>nation-wide essay competition</b>

## Projects

2017	Website of a Bank management system with JSP, HTML, CSS (As part of the Database systems course)
2018-2019	Machine learning based Bank Credit Risk Prediction (A collaboration with <a href="#">Bank Asia</a> )
2019	TCP Reset Attack on video streaming (Computer security)

## Selected Computer Literacy

Languages	Python, C/C++, Java, R, Bash, Matlab, PHP, JavaScript, HTML/CSS, SQL
Libraries	Keras, Tensorflow, NLTK, Pandas, NumPy
Frameworks	Django

## Leisure

Hobbies	Reading, Swimming, Sketching, Learning languages
Languages	Bangla, English, French, Arabic