Sumaira Zaman

5A Poplar Lane, Storrs, CT 06268 • Phone: 203-560-6723 • E-Mail: sumaira.zaman@uconn.edu • Website: https://sumairazaman.weebly.com/

Education

PhD in Computer Science & Engineering, GPA 3.7/4.0, Expected 2022, University of Connecticut, Storrs, CT

M.S. in Ecology & Evolutionary Biology, GPA 3.7/4.0, August 2018, University of Connecticut, Storrs, CT

B.S. in Biomedical Engineering, Minor in Computer Science, GPA 3.8/4.0, May 2017, University of Connecticut, Storrs, CT

Experience

University of Connecticut - Computational Biology Lab

- Studying different parameters of protein domain evolution and its impact on the inference of gene family evolution.
- Investigating statistical methods e.g. Bayesian machine learning to mitigate the effects of protein domain evolution on the inference of gene family phylogeny.
- Evaluating protein domain architectures in 12 fly species across 7,000 gene families, and studying their impact on multiple sequence alignment and reconstruction of gene tree phylogeny.

University of Connecticut – Plant Computational Genomics Lab March 2015 – Present

- Evaluating and analyzing appropriate bioinformatics methods for annotating genomes of non-model species (specifically conifers) in order to generate accurate gene models for comparative genomics studies.
- Investigate and implement a deep learning approach towards genome annotation for non-models, inclusive of more features such as protein domain, free energy, gene structure, and other critical components.
- Developing OrthoQuery, a Tripal module using a Drupal content management system for TreeGenes database, allowing user to analyze, investigate and visualize orthologous relationships between selected plant species.

Jackson Laboratory – Computational Methods in Genomic Medicine Lab

- Built a pipeline to analyze raw sequencing data from ATAC-seq and identify peaks in the genome using shell scripts, pre-exiting bioinformatics software, and high-performance computing.
- Determined which ATAC-seq features (sequence length, peak height) and genetic features (transcription factor binding sites, GC content) are relevant for enhancer prediction through principal component analysis done in **R**.
- Implemented Decision Trees, Support Vector Machines, and Naïve Bayes in MATLAB to predict enhancer regions within the human genome given specific ATAC-seq and genetic features.

University of Connecticut - Molecular Genetics Lab

- Identifying and analyzing genes under positive selection in wallaby and closely related species with RNA-seq data and maximum likelihood methods (CodeML).
- Wet lab experience including gel electrophoresis, polymerase chain reaction, RNA extraction, and prepping samples for sequencing.
- Cell culture maintenance including sub-culturing and cryopreservation.

Conferences

- A. Uyar, A. Kaygun, **S. Zaman**, E. Marquez, J. George, D. Ucar , "Identification of Enhancer Sequences by ATAC-seq Open Chromatic Profiling", in *Human Genome Conference*, Houston, Texas, Feb. 28 March 2, 2016.
- S. Zaman, C. Fisher, J. Whalen, S. Jogdeo, R. Cronn, J. Wegrzyn "Bioinformatic Evaluation of Transcriptomic Frame Selection Method in Non-Model Species". Poster presentation at Plant Animal Genome Conference, San Diego, California, Jan. 14-18, 2017.
- **S. Zaman**, M. Caballero, J. Wegrzyn, "Comparative Evaluation of Gene Annotation Methods in Conifer Megagenomes". Oral presentation at Plant Animal Genome Conference, San Diego, California, Jan. 14-18, 2018.
- **S. Zaman**, E. Grau, S. Buehler, S. Ficklin, Jill Wegrzyn, "Orthoquery A Tripal Module for Supporting Identification and Visualization of Orthogroups". Lightening Talk at Galaxy Community Conference. Portland, Oregon. June 2018.

June 2015 – Aug 2015

August 2018 – Present

June 2014 - Aug 2014

- S. Zaman, A. Ferreira, M. Caballero, J. Wegrzyn, "Refining Annotation Methodology for Comparative Genomics in Conifers." Plant Animal Genome Conference, San Diego, California, Jan. 14-18, 2019.
- **S. Zaman**, E. Grau, S. Buehler, S. Ficklin, Jill Wegrzyn, "Orthoquery: A Tripal Database Module to Assess and Visualize Gene Family Evolution". Plant Animal Genome Conference, San Diego, California. Jan. 14-18, 2019.
- S. Zaman, A. Ferreira, M. Caballero, R. Whetten, J. Wegrzyn, "Improving Annotation Methodology in Conifers." IUFRO Tree Biotechnology Conference. Raleigh, North Carolina. June 24-27, 2019.

Publications

- A. Trouern-Trend, T. Falk, **S. Zaman**, M. Caballero, D.B. Neale, C.H. Langley, A. Dandekar, K.A. Stevens, J.L. Wegrzyn, "Comparative Genomics of Six Juglans Species Reveals Patterns of Disease-associated Gene Family Contractions." 2019. (Accepted in *Plant Journal*).
- T. Falk, N. Herndon, E. Grau, S. Buehler, P. Richter, S. Zaman, E.M. Baker, R. Ramnath, S. Ficklin, M. Staton, F.A Feltus, S. Jung, D. Main, J.L Wegrzyn, "Growing and cultivating the forest genomics database, TreeGenes." *Database*, Volume 2018.
- D. Neale, P. McGuire, [and 20 others including **S. Zaman**], "A Reference Genome Sequence for Douglas-fir", *Genes Genomes Genetics* (Accepted Aug. 2017)

Teaching

- Undergraduate & Graduate Teaching Assistant for Introductory Programming Supervised laboratory programming assignments, taught students the basic concepts of programming, introduced microcontrollers, and graded assignments.
- Graduate Teaching Assistant for Practical Genomics Responsibilities include supervising bioinformatics labs and grading. Lectured on Github and Python and designed coding challenges for students.
- Graduate Teaching Assistant for Data Structure Responsibilities include reinforcing the concepts of data structure through programming assignment and assisting students during office hours and through online platforms (Piazza).

Awards/Leadership

- Placed 1_{st} amongst 20 teams in Junior Design and top 10 amongst 120 teams in Innovation Quest by developing and marketing a respiration monitoring system for infants as a viable consumer product for detecting SIDS (Sudden Infant Death Syndrome).
- Recipient of OUR (Office of Undergraduate Research) Travel Award to present "Bioinformatic Evaluation of Transcriptomic Frame Selection Method in Non-Model Species" at Plant and Animal Genome 2017 conference in San Diego.
- Recipient of the Leadership Scholarship, awarded to students who achieve a competitive grade point average in high school and demonstrate commitment towards multicultural diversity programs and initiatives through their leadership.
- Recipient of the Walter M. Rose Scholarship, awarded to students in the School of Engineering who demonstrate academic excellence and involvement in the community.
- Galaxy Community Conference 2018 Travel Fellowship to present "OrthoQuery: A Tripal Module to Assess Gene Family Evolution in Sequenced Gymnosperms".
- IUFRO Tree Biotechnology 2019 Travel Award to present "Improving Annotation Methodology in Conifers".

Technical Skills & Relevant Coursework

- Programming Languages: Python, Matlab, Java, JavaScript, R, C, Assembly
- Libraries:
- 5+ year experience in high performance computing, Linux OS, & bash scripting
- Version control through Github & Gitlab
- Relevant Coursework: Algorithms, Computational Biology, Computational Genomics, Machine Learning, Bayesian Machine
 Learning, Dynamic Modeling of Biological Networks, Computational Problem in Evolutionary Genomics, Computer Methods in
 Molecular Evolution, Phylogenetics, Genetics, Theory of Computation, Discrete Optimization