

SAGNIK BANERJEE

437 Bessey Hall, 2200 Osborn Dr, Ames, IA 50011

(725)-333-0696, sagnik@iastate.edu

<https://github.com/sagnikbanerjee15>

EDUCATION

2016-2021 (Expected)	Ph.D. in Bioinformatics and Computational Biology <i>Iowa State University, Ames, USA</i> Major Advisor: Dr. Carson Andorf Co-Major Advisor: Dr. Karin Dorman
2012-2014	M.E. in Computer Science and Engineering <i>Jadavpur University, Kolkata, India</i> Thesis Advisor: Dr. Mita Nasipuri GPA: 9.22/10
2008-2012	B.Tech. in Computer Science and Engineering <i>Institute of Engineering and Management, Kolkata, India</i> Project Advisor: Dr. Tamal Chakrabarti GPA: 9.02/10

PROFESSIONAL EXPERIENCE

Jan 2018 – Present	Graduate Research Assistant <i>Agricultural Research Service (ARS) Research Participation Program Oak Ridge Institute for Science and Education (ORISE)</i> Currently working on improving gene annotations in cereal genomes using a variety of approaches
Jan 2019 – May 2019	Graduate Teaching Assistant <i>Bioinformatics and Computational Biology Program</i> <i>Iowa State University, Ames, USA</i> Graded homeworks Conducted recitation classes
Aug 2017 – Dec 2017	Graduate Teaching Assistant <i>Department of Genetics, Developmental and Cell Biology</i> <i>Iowa State University, Ames, USA</i> Conducted laboratory classes (255L) instructing Human Anatomy for 45 students in total
May 2017 – Aug 2017	Graduate Research Assistant <i>Department of English</i> <i>Iowa State University, Ames, USA</i> Used machine learning algorithms to predict a writer's native language from essay samples written in English
May 2017 – Aug 2017	Graduate Research Assistant <i>Department of Plant Pathology and Microbiology</i>

	<p><i>Iowa State University, Ames, USA</i> Investigated transcription factors involved in gene regulation for combating Blumeria-induced infection in barley Discover novel barley protein candidate interactors of blumeria effectors in barley-blumeria infectome</p>
Aug 2016 – May 2017	<p>Graduate Research Assistant <i>Bioinformatics and Computational Biology Program</i> <i>Iowa State University, Ames, USA</i> Correlated eight biological replicates of ATAC-seq across two time points Designed a software to remove bias from gene ontology annotations Investigated protein-protein interaction between barley and blumeria proteins from Y2H-Seq data</p>
Jul 2015 – Jul 2016	<p>Assistant Professor <i>Department of Electronics and Communication Engineering</i> <i>Institute of Engineering and Management, Kolkata, India</i> Taught Data Structures and Database Management Systems to undergraduate students and conducted lab sessions for 330 students Handled Neural Networks classes for master's students for 15 students at master's level Investigated Post Translational Modification (Glycosylation) and Protein-Protein Interaction</p>
Jan 2015 – Jul 2015	<p>Assistant Professor <i>Department of Computer Science and Engineering</i> <i>Techno India – Batanagar, Kolkata, India</i> Taught C and Java to undergraduate students and conducted lab work Undertook research work on Protein Domain Prediction</p>
Jul 2014 – Jan 2015	<p>Researcher <i>CMATER Lab, Department of Computer Science and Engineering</i> <i>Jadavpur University, Kolkata, India</i> Completed project titled "Computational Proteomics" Carried out research on Post Translational Modification (Phosphorylation)</p>
Jul 2013 – Jun 2014	<p>Teaching Assistant <i>Department of Computer Science and Engineering</i> <i>Jadavpur University, Kolkata, India</i> Assisted professors in carrying out lab work for the Object-Oriented Programming lab and the Numerical Methods lab Assisted professors in carrying out daily teaching activities</p>

AWARDS & ACHIEVEMENTS

Spring 2020	<p>Travel award <i>Bioinformatics & Computational Biology program</i> <i>Iowa State University, USA</i> Awarded \$150 travel fellowship for coming in 3rd at a peer-evaluated Student Seminar course</p>
Spring 2018	<p>Travel award <i>Bioinformatics & Computational Biology program</i> <i>Iowa State University, USA</i> Awarded \$200 travel fellowship for coming in 2nd at a peer-evaluated Student Seminar course</p>

Mar 2018	Student Poster Award <i>Bioinformatics & Computational Biology program</i> <i>Iowa State University, USA</i> 3 rd prize winner for Student Poster Presentation at the 4 th Annual Bioinformatics and Computational Biology Symposium
Aug 2017	Recognized Reviewer <i>Applied Soft Computing Journal</i> <i>Elsevier</i> Awarded for completing 70 manuscript reviews between 2015-2017
Dec 2016	Outstanding Reviewer <i>Applied Soft Computing Journal</i> <i>Elsevier</i> Awarded for being within the top 10 th percentile of reviewers for the Journal
2008	Certificate of Merit <i>Central Board of Secondary Education, India</i> Awarded Certificate of Merit in Chemistry for being among the top 0.1% of successful candidates, India
2008	Central Sector Scholarship <i>Central Board of Secondary Education, India</i> Awarded for exceptional performance in class XII exams, India

PUBLICATIONS

Velásquez-Zapata, V., Elmore, J. M., **Banerjee, S.**, Dorman, K. S., Wise R. P. (2020). Y2H-SCORES: A statistical framework to infer protein-protein interactions from next-generation yeast-two-hybrid sequence data. (*Under review in Plos Computational Biology*)

Banerjee, S., Velásquez-Zapata, V., Fuerst, G., Elmore, J. M., Wise R. P. (2020). NGPINT: A Next-generation protein-protein interaction software. Briefings in Bioinformatics (*Accepted*)

Elmore, M., **Banerjee, S.**, Pedley, K., Ruck, A., Whitham, S. (2020). De novo transcriptome of *Phakopsora pachyrhizi* uncovers putative effector repertoire during infection. *Physiological and Molecular Plant Pathology*, 101464.

Hunt, M., **Banerjee, S.**, Surana, P., Liu, M., Fuerst, G., Mathioni, S., Nettleton D., & Wise, R. P. (2019). Small RNA discovery in the interaction between barley and the powdery mildew pathogen. *BMC genomics*, 20(1), 610.

Vajjala, S., & **Banerjee, S.** (2017). A study of N-gram and Embedding Representations for Native Language Identification. *Bronze Sponsors*, 240.

Banerjee, S., Ghosh, D., Basu, S., & Nasipuri, M. (2016). JUPred_SVM: Prediction of Phosphorylation Sites Using a Consensus of SVM Classifiers. In *Proceedings of Fifth International Conference on Soft Computing for Problem Solving* (pp. 553–560). Springer.

Banerjee, S., Ghosh, D., Basu, S., & Nasipuri, M. (2016). JUPred_MLP: Prediction of Phosphorylation Sites Using a Consensus of MLP Classifiers. In *Proceedings of the 4th International Conference on Frontiers in Intelligent Computing: Theory and Applications (FICTA)*. Springer.

Banerjee, S., Guha, S., Dutta, A., & Dutta, S. (2015). Improvement of protein disorder prediction by brainstorming consensus. In *Computing and Communication (IEMCON), 2015 International Conference and Workshop on* (pp. 1–7). IEEE.

Banerjee, S., Mitra, B., Chatterjee, A., Santra, A., & Chatterjee, B. (2015). Identification of relevant physico chemical properties of amino acids with respect to protein glycosylation prediction. In *Computing and Communication (IEMCON), 2015 International Conference and Workshop on* (pp. 1–7). IEEE.

Banerjee, S., Basu, S., Ghosh, D., & Nasipuri, M. (2015). PhospredRF: Prediction of protein phosphorylation sites using a consensus of random forest classifiers. In *Computing and Communication (IEMCON), 2015 International Conference and Workshop on* (pp. 1–7). IEEE.

Banerjee, S., Nag, S., Tapadar, S., Ghosh, S., Guha, S., & Bakshi, S. (2015). Improving protein protein interaction prediction by choosing appropriate physiochemical properties of amino acids. In *Computing and Communication (IEMCON), 2015 International Conference and Workshop on* (pp. 1–8). IEEE.

Banerjee, S., Basu, S., & Nasipuri, M. (2015). Big Data Analytics and Its Prospects in Computational Proteomics. In *Information Systems Design and Intelligent Applications* (pp. 591–598). Springer.

Banerjee, S., Chakrabarti, T., & Sinha, D. (2013). A Faster Fitness Calculation Method for Genetic Algorithm Based Multiple Protein Sequence Alignment, *4*(7), 768–772.

Banerjee, S., Chakrabarti, T., & Sinha, D. (2013). Finding all Occurrences of a Pattern by a Genetic Algorithm based Divide-and-Conquer Method. *International Journal of Computer Applications*, *64*(18).

Banerjee, S., Chakrabarti, T., & Sinha, D. (2012). A Genetic Algorithm Based Pattern Matcher.

ACADEMIC PROJECTS

Jan 2019 - Present

Discovery and characterization of orphan genes in Arabidopsis

Iowa State University, USA

Using RNA-Seq data and microarray data to find Orphan genes

May 2018 - Present

De-novo Assembly of the Transcriptome of the fungus Soybean Rust using error corrected long reads

Iowa State University, USA

Investigating a plethora of methods to assemble the transcriptome of soybean rust *de-novo*

Jan 2018 - Present	Improvement of Gene Annotation in Newly Sequenced Genomes <i>Iowa State University, USA</i> Compared various gene annotation strategies Currently developing a novel pipeline to generate best annotations using both short and long reads
Mar 2017 - Present	Identification and Characterization of Host Targets of Powdery Mildew Effector Proteins <i>Iowa State University, USA</i> <ul style="list-style-type: none"> □ Engineered a novel pipeline which is capable of, but not limited to, the following: <ul style="list-style-type: none"> ○ Cleaning up of reads to remove adapter sequences ○ Removing of short reads which maps to spurious places ○ Optimizing mapping of reads to remove vector sequences ○ Detecting differentially enriched genes which are expressed both in-frame and out-of-frame □ Currently developing a statistical model to optimize recognition of known flanked sequences □ Developed software to identify protein interactors
May 2017 – May 2018	Detection of phasiRNAs in Barley-Blumeria infectome <i>Iowa State University, USA</i> <ul style="list-style-type: none"> □ Developed a software to detect phasiRNAs in barley and blumeria using a hyper-geometric distribution
May 2017 – Aug 2017	Native Language Identification shared task project <i>Iowa State University, USA</i> <ul style="list-style-type: none"> □ Improved the performance of native language identification using an ensemble of machine learning classifiers
Jan 2013 - May 2014	A consensus-based approach towards Protein Secondary Structure Prediction <i>Jadavpur University, India</i> <ul style="list-style-type: none"> □ Designed an ensemble classifier comprising of multiple Support Vector Machines and augmented it with other existing protein secondary structure predictors to further improve prediction
Jul 2011 - May 2012	An optimized pattern matching algorithm <i>Institute of Engineering & Management, India</i> <ul style="list-style-type: none"> □ Designed an algorithm which could perform faster string search on compressed data □ Compressed both the text and the pattern using Huffman's algorithm, where the dictionary used for compressing pattern was same as that of the text □ Rewrote Knuth-Morris-Pratt algorithm (KMP) and Boyre Moore (BM) algorithm such that they could match text and patterns represented in the form of bits
Jul 2011 – May 2012	Learner's Management System <i>Institute of Engineering & Management, India</i> <ul style="list-style-type: none"> □ Designed a website which would cater to the daily needs of academia □ Designed an online whiteboard, using HTML5, which could be used for communication between teachers and students □ Added a module which could check a program against a set of previously provided input and output thereby eliminating the need for manual inspection

Jun 2011 – Jul 2011

Creation of a calculator using an 8051 microcontroller

ICEE Design Technologies, Kolkata, India

- ☐ Created a basic calculator using an 8051 microcontroller
- ☐ Improved the calculator by programming it to solve quadratic equations

MENTORING EXPERIENCE

- Served as a **Mentor** for Larry Davis, a first year PhD student in Bioinformatics and Computational Biology Program (BCB), Iowa State University
- Provided **Letter of Recommendation (LOR)** for Sourav Ghosh, a prospective MS student in the Department of Computer Science and Engineering, Institute of Engineering and Management, Kolkata, India, currently pursuing MS at TU Dresden, Germany
- Mentored 2 graduate students (MS Level) to support their advancement in research activity. These include, but are not limited to:
 - Designing a research problem
 - Searching the correct sets of input data
 - Developing a pipeline to integrate several different algorithms
 - Understanding and interpreting the results

POSTER PRESENTATIONS

- Presented a poster titled “FINDER: A fully automated pipeline to FIND accurate gene structures from proteins and RNA-Seq expression data” in the 11th ACM Conference on Bioinformatics, Computational Biology and Health Informatics (**September 2020**)
- Presented a poster titled “A computational pipeline to identify host targets of fungal virulence proteins using next generation yeast two-hybrid” in the 26th Conference on Intelligent Systems for Molecular Biology (**July 2018**)
- Presented a poster titled “A computational pipeline to identify host targets of fungal virulence proteins using next generation yeast two-hybrid” in the Loomis Symposium (**May 2018**)
- Presented a poster titled “A computational pipeline to identify host targets of fungal virulence proteins using next generation yeast two-hybrid” in the 1st Midwest Statistical Machine Learning Colloquium (**May 2018**)
- Presented a poster titled “A computational pipeline to identify host targets of fungal virulence proteins using next generation yeast two-hybrid” in the 5th Graduate and Professional Student Research Conference (**April 2018**)
- Presented poster titled “A computational pipeline to identify host targets of fungal virulence proteins using next generation yeast two-hybrid” in the 4th Annual Bioinformatics and Computational Biology Symposium (**March 2018**)

TALKS

- Presented at the 6th Annual Bioinformatics & Computational Biology Symposium titled “A computational pipeline to identify host targets of fungal virulence proteins using next generation yeast two-hybrid screening” (**May 2020**)

MEMBERSHIPS IN PROFESSIONAL SOCIETIES

2018-Present International Society for Computational Biology (ISCB)

2017-Present American Association for the Advancement of Science (AAAS)

SERVICE

Institutional

2018-Present	Graduate Professional Student Senate (GPSS) Senator for Department of Plant Pathology and Microbiology
October 2018	Presented a section in the event “The Road Less Travelled” – an outreach event to encourage girls at high-school level to pursue a career in science
April 2018	Volunteered to be a facilitator for the 5th Graduate and Professional Student Research Conference
2017-2018	Vice President of Bioinformatics and Computational Biology Graduate Student Organization (BCBGSO)

Professional

- Reviewed 77 Papers in Applied Soft Computing Journal, Elsevier
- Reviewed 47 Research Papers in BMC Genomics
- Reviewed 9 Research Papers in BMC Bioinformatics
- Program Chair for IEMCON-2015, Vancouver, Canada, October 2015

Community

- Served as the lead judge for State Science and Technology Fair of Iowa (SSTFI) (March 2019)
- Served as a judge for State Science and Technology Fair of Iowa held virtually (SSTFI) (March 2020)

