RESEARCH PROJECTS

A). Prediction of biomolecular function of proteins

Our group over the past decade has come up with many novel approaches for computational inference of biomolecular function of proteins.[1-6] However, there is still a long way to go.[7] We still cannot computationally predict what function(s) a protein may switch during post-translation modification, oligomerization or even in context of multi-domain proteins. Our lab has for the first time shown that it is possible to use molecular dynamics information as a fingerprint like evidence for biomolecular function. We would like to further advance the work leveraging a combination of static and dynamic information on proteins to computationally infer their function in various biological states. Student interested in this project needs to be inclined in learning about protein structure and have a good understanding of physics and maths. He/she will also need to do computer programming.

References

1. Inference of protein function from protein structure. Debnath Pal and David Eisenberg. Structure **13**, 121-130 (2005)
2. Functionally important segments in proteins dissected using Gene Ontology and geometric clustering of peptide fragments. K. Manikandan, Debnath Pal, S. Ramakumar, Nathan E Brener, Seetharama Iyenger, Guna Seetharaman. Genome Biology , 9:R52(2008)
3. Inferring molecular function: contributions from functional linkages. Arturo Medrano-Soto, Debnath Pal and David Eisenberg. Trends in Genetics **24**, 587-590 (2008)
4. *De novo* inference of protein function from coarse-grained dynamics. Pratiti Bhadra and Debnath Pal. Proteins: Structure, Function and Bioinformatics, **82**, 2443-2454 (2014)
5. Identifying functionally important cis-peptide containing segments in proteins and their utility in molecular function annotation. Sreetama Das, Suryanarayanrao Ramakumar and Debnath Pal. FEBS Journal, **281**, 5602-5621 (2014)
6. Molecular dynamics information improves *cis*-peptide based function annotation of proteins. **Under Submission.**
7. On Gene Ontology and Function Annotation. Debnath Pal. Bioinformation **1**, 97-98 (2006).

B). Analysis of large-scale genomic datasets

The amount of publicly available genomic datasets are increasing rapidly. However, much of the analysis ignores presence of repeat sequences, problems with using the same reference genome and inadequate coverage with many multi-reads not mapped owing to constraints of efficiency. Our group in collaboration has developed accurate short read mapping technology[1-4] , where the present deficiencies in biological analysis can be addressed. In light of these, we would like to develop our own Genome Analysis Pipeline through which we are able to analyse single nucleotide polymorphism, copy number variation, and other chromosomal aberrations more accurately. Student interested in this project need to be inclined to learn about genome, transcriptome and chromosomal organization and develop a good understanding of cellular processes of specific interest. He/she will need to have good understanding of statistics and do computer programming.

References

1. Hardware accelerator for alignment of short reads in sequencing platforms. Santhi Natarajan, Debnath Pal and S. K. Nandy. (Patent filed, 2015)
2. Mapping of short reads in sequencing platforms. Santhi Natarajan, Debnath Pal and S. K. Nandy. (Patent filed, 2015)
3. Data streaming in hardware accelerator for alignment of short reads. Santhi Natarajan, Debnath Pal and S. K. Nandy. (Patent filed, 2015)
4. AccuRA: Accurate Alignment of Short Reads on Scalable Reconfigurable Accelerators. Santhi Natarajan, Krishna Kumar N., Debnath Pal and S. K. Nandy. International Conference on Embedded Computer Systems: Architechtures. MOdelling and Simulation (SAMOS) XVI, **Accepted** , (2016)

C). Structure based analysis of protein interaction networks in biology

Our group has been working in the area of protein-protein interaction and quaternary structure assessment over the past several years. In the initial years we have focussed primarily in attending to the deficiencies in the existing approaches of protein-protein docking. However, to make further improvement it is necessary to perform detailed structure based analysis of known protein-protein interaction networks. We currently have highly accurate methods for quaternary structure assessment and would like to leverage this information further to come up with insights into interaction networks in cells. The results obtained could also be very useful for helping crystallographers to co-crystallize protein complexes for structure determination. The knowledge will be pooled to improve our in-house docking pipeline.

1. Accurate Detection of Cyclic and Dihedral Point Group Symmetry of Proteins in Lattice Structures. **Under Submission.**
2. PRUNE and PROBE - two modular web services for protein-protein docking. Pralay Mitra and Debnath Pal. Nucleic Acids Research, **39**, W229-W234 (Webserver Issue, 2011)
3. Combining Bayes classification and point group symmetry under Boolean framework for enhanced protein quaternary structure inference. Pralay Mitra and Debnath Pal. Structure **19**, 304-312 (2011)
4. Using correlated parameters for improved ranking of protein-protein docking decoys. Pralay Mitra and Debnath Pal. Journal of Computational Chemistry **32**, 787-796 (2011)
5. dockYard - a repository to assist modeling of protein-protein docking. Pralay Mitra and Debnath Pal. Journal of Molecular Modeling **17**, 599-606 (2011)