Department of Biological Sciences

PhD Open Seminar

Speaker: **Ankit Gupta**   Advisor: **Dr. Vineet K. Sharma**
Date & Time: **Tuesday, 27th Feb., 2018 at 10:00 a.m.**   Venue: **L3, LHC**

Functional Insights into Genomes and Metagenomes through Integrative Omics and Machine Learning Approaches

Recent advancements in high-throughput sequencing technologies have enabled us to access the genetic and functional information inherent in individual genomes and metagenomes [1]. Several large-scale global efforts are underway for the characterization of genomes, and complex microbiomes such as the human gut microbiome, which plays a key role in the well-being of the host. The understanding of the human gut microbiome holds unprecedented potential in diagnostics, therapeutics, and in identifying the microbial markers associated with gut diseases [2]. Thus, we examined the taxonomic, functional and metabolic dysbiosis in colorectal carcinoma in the Indian population using multi-omics approaches [3]. We identified potential microbial markers such as microbial genes and metabolites, which can be used for the early-stage detection of colorectal cancer. The study also revealed a unique cancer-associated gut microbiome of the Indian population compared to other western countries, which holds significance due to the low incidence of colorectal cancer in the Indian population. However, the major bottleneck for any such large scale metagenomic analysis is the unavailability of efficient computational methods for functional analysis and also for the comparative analysis of healthy and diseased microbiomes. For this task, I have developed a unique tool 'MP3' using machine learning to predict and compare complete and partial pathogenic proteins from genomic and metagenomic datasets [4, 5]. Similarly, 'MicroTAXI' was developed for the ab-initio taxonomic classification using the proteome information instead of the 16S rRNA gene sequence [6]. Another challenge was to recover assembled genomes from metagenomic data, for which I developed a unique 'Binning-Assembly' approach for the reconstruction of bacterial and viral genomes from a fragmented pool of genome 'metagenome' [7]. We also applied the similar approaches for the genomic analysis of higher eukaryotes to decipher the inherent genomic clues, constructed complete genome sequences, identified genes and their role in early developmental pathways, and innate and adaptive immunity. The results from above work will be presented.

References:

1) Hardwick, S.A.et al., (2017). Mercer, *Reference standards for next-generation sequencing.* Nat Rev Genet, 2017. 18(8): p. 473-484.

2) DeSantis, T.Z., et al., (2017). *Microbial markers in the diagnosis of colorectal cancer: the promise, reality and challenge.* Future Microbiol, 2017. 12: p. 1341-1344.

3) **Gupta, A.,** et al. Assessing the role of location and diet on the gut microbiome of patients suffering from colorectal carcinoma using multi-omics approaches. *(manuscript under preparation)*

4) **Gupta, A.**, et al. (2014). MP3: a software tool for the prediction of pathogenic proteins in genomic and metagenomic data. PloS One, 9(4), p.e93907.

5) Sharma, A.K.*, **Gupta, A.***, et al., 2015. Woods: A fast and accurate functional annotator and classifier of genomic and metagenomic sequences. **Genomics**, 106(1), pp.1-6. **(* equal contrubution)**

6) **Gupta, A.** and Sharma, V.K. (2015). Using the taxon-specific genes for the taxonomic classification of bacterial genomes. BMC Genomics, 16(1), p.396.

7) **Gupta, A.,** et al., (2016). Reconstruction of bacterial and viral genomes from multiple metagenomes. Frontiers in Microbiology, 7.