**Gerstein Lab Experience with Network Reconstruction**

We were among the earliest to use statistical models to perform systematic network reconstruction of PPI networks by integrating high-throughput data. The task of identifying interactions between biological elements is known as network reconstruction. It requires integrating both large-scale yet noisy data from high-throughput experiments and reliable yet incomplete data from small-scale experiments. We have much previous work in this area. Since then, we have developed many methods for reconstructing a variety of biological networks. We have developed JEME, a method for reconstructing the enhancer-target gene interaction network by integrating epigenetic data from hundreds of human samples, allowing each gene to be regulated by multiple enhancers in a cell- and tissue-specific manner. Using our methods, we have successfully reconstructed different types of networks in human and other model organisms, like TF binding and microRNA (miRNA) binding networks in human cell lines and in C. elegans, respectively based on data from ENCODE and modENCODE. More recently, we merged and uniformly processed data from PsychENCODE and multiple other sources to build a harmonized data resource, which we used to reconstruct cell-type specific gene regulatory networks in the adult brain.

**Gerstein Lab Experience with Variant Prioritization using Networks**

Reconstructed networks enable us to infer the direct and indirect effects of variants, and we have worked closely for ~15 years to use networks as a means of prioritizing and investigating variants. We have leveraged network properties (such as centrality) to evaluate the functional significance of genomic variants, revealing that the correlations between them depend on the network type. We devised methods to identify generalized hierarchies (with loop structures). We have also developed HirNet (a method for determining network hierarchies). We also investigated the roles of network bottlenecks (i.e., nodes through which many "shortest paths'' pass) in the context of protein networks.

Genomic variants can also lead to disruptions of network connections. As a result, some recurrent patterns may be perturbed, such as TFs that frequently co-regulate target genes. We have developed the DiNeR method for identifying such changes and analyzing their consequences to downstream gene expression programs. On a larger scale, some network perturbations may propagate to cause major network rewiring. We have developed the TopicNet method to measure such rewiring in transcriptional regulatory networks. We have also applied this idea to study network rewiring in cancer cells, as part of our efforts in producing a general resource for cancer research based on ENCODE data. Finally, to predict the molecular effects of non-coding variants on genes in a cell-type specific manner. We found that we have developed GRAM, a generalized model the TF features with high importance show high network centrality.

**Gerstein Lab Experience with Analyzing scATAC-seq and scRNA-seq Data**

We have extensive experience applying bulk and single-cell genomics techniques based on our involvement in the original single cell mapping component of the BRAIN initiative and more recently PsychENCODE. We integrated bulk ATAC-seq with sc-RNA using the SCENIC pipeline to create a first-generation map of transcription factor regulatory networks during cortical neurogenesis, the major process we propose to model here. This work, and our recently completed nuc-Seq and ATAC-seq in postmortem human brain in autism and controls (n = 20), strongly supports the feasibility of the proposed studies.