4a(i) Preliminary Results: Identifying regulatory interactions, enhancers, and constructing regulatory networks

We have developed a set of computational tools to identify gene regulatory elements and analyze regulatory networks. PeakSeq \cite{19122651} and MUSIC \cite{25292436} are efficient ChIP-Seq data processing algorithms that can be used to locate transcription factor (TF) binding sites. In order to establish regulatory relationships between TFs and target genes, we introduced a probabilistic model-based method, TIP (Target Identification from Profiles) \cite{22039215}. We have applied machine-learning methods that integrate multiple genomics features to classify human regulatory regions from ENCODE data of more than 100 transcription factor binding sites. A computational pipeline was developed to identify potential enhancers from regions classified as gene-distal regulatory modules. \cite{22950945}. Making use of the potential enhancers, we developed the Function-based Prioritization of Sequence Variants (FunSeq) tool \cite{24092746} for identification of candidate drivers in tumor genomes, and more recently, a more elaborate and flexible framework, FunSeq2, integrating various genomic and cancer resources to prioritize cancer somatic variants, especially regulatory noncoding mutations. \cite{25273974}.