TITLE:

A COMPREHENSIVE REVIEW AND EVALUATION OF PROTEIN SEQUENCE AND PROTEIN TO PROTEIN INTERACTION NETWORK:

The concept of bioinformatics is merged with a holistic approach to data mining techniques. the solutions for the specific bioinformatics applications such as protein complex detection, proteins are fundamental to the cell structural components proteins are dynamically generated by a cell that executes the genetic program. some proteins have a unique function, the majority of them interact with other proteins forming protein-protein interaction networks to carry out processes in the cell. such processes include cell cycle control, differentiation, signaling, translation, transcription, protein folding, post transactional modification, and transportation, protein complexes are used to help us to predict the functions of proteins. There is evidence that many disease mechanisms involve protein complexes. For Example, The complex "RNA polymerase II" can produce mRNAs, snRNAs, many diseases are usually caused by an erroneous production of some protein complex various experiment technologies such as yeast two-hybrid purification generated a large amount of protein-to-protein interaction data. detecting overlapping protein complexes is a non-trivial task for two reasons, firstly the number of protein complexes is unknown for a given PPI network, secondly, the protein complexes may have many reliable interactions with other complexes. In this research, the fuzzy detection algorithm is applied with prorank to detect the overlapping protein complexes in the PPI network. to evaluate prorank based fuzzy quantitatively, the algorithm is applied to detect the overlapping protein complexes. in yeast datasets, PPI-D1, PPI-D2, collins, and krogan-core, DIP(Database of Interacting Proteins) Phenotypes such as yeast heterogeneous and yeast homogeneous and proposed algorithm with popular clustering methods namely clusterONE and ProRank, ProRank for same dataset the predicted complexes are compared with two references namely cmplex-D1 and complex-D2 from MIPS(Munich Information Center for Protein sequence finally the results are compared using some evaluation criteria to measure the proRank based Fuzzy algorithm. The main objective is to predict the protein function of yeast from different perspective such as Protein structural class prediction, protein complex detection, alignment and pattern Identification. Specific objectives are as follows:

To align multiple protein sequences

- To detect the overlapping macromolecule protein complexes with smart performance and robustness

- Improved algorithm for predicting the secondary structure class of a particular protein sequence

- A hybrid pattern matching algorithm to perform protein secondary structure prediction