

Mid-term project

- **Develop a script which takes in a human protein-protein interaction network from biogrid (<http://www.thebiogrid.org>) database (attached here) and identifies the function of a protein (given as input at command line as a HGNC/Official symbol) based on its interaction partners. The script should be able to accomplish the following**
 - If the number of physical interactions are too low or if the annotation can't be predicted, it should provide information about the same.
 - Should be able to use the gene ontology annotations of proteins available for humans from <http://www.geneontology.org/GO.downloads.annotations.shtml>
 - Tip: you can use any simple guilt by association approach to transfer the annotation i.e, if the majority of the proteins are annotated with a function (based on gene ontology terms) then the predicted function of the query protein would be the same as that of the majority.
 - Note that proteins can have more than one function based on this notion.

Mid-term project is due October 30th