Name: Ulvi Bajarani

Student ID: 20539914

The name of presenter: Dr. Marzieh Ayati

The topic of presentation: Bioinformatics: Co-Phosphorylation Networks

The date of presentation: Nov-6-2019

Questions:

1. In paper 2, there is a quote "In this paper, we propose a scoring scheme that (i) integrates disease association and network connectivity in a parameter-free fashion and (ii) incorporates an approximation of the statistical significance of this integrated score. The key idea of the proposed method is to assess the disease association of each interaction in the network and account for the background disease association as an approximation to statistical significance. In this respect, the proposed approach may be thought of a generalization of Newman’s [18] measure of modularity, which was developed for community detection in networks."

How is the background disease association calculated? How might be known that this background disease association gives reliable results?

1. In paper 2, there is a quote "To avoid making assumptions on the distribution of disease association scores, we compute these background scores empirically for each protein pair. For this purpose, we randomize the original GWAS data by permuting the labels of the samples to break the relationship between the genotype and phenotype, while preserving the distribution of genotypes for each locus and also preserving the relationship between loci"

Why are the distribution of genotypes for each locus and relationship between loci preserved?