**Betweenness Centrality in Weighted PPI Network Identifies Critical Genes and Potential Drugs for Major Depressive Disorder**

**Noah Black1, Christine Cai2, Sophie Cai3, Alana Schroden4**

*112th, 29th, 39th, 411th, H. H. Dow, Midland, MI*

**Introduction:** Major Depressive Disorder (MDD) is a psychological disease affecting millions of people. Characterized by a lack of neurotransmitters, MDD has a myriad of symptoms including persistent sadness which may render patients unable to perform basic activities. The genetic basis of MDD remains elusive. We sought to change this by researching betweenness centrality in protein-protein interaction (PPI) networks. Betweenness centrality is the amount of information communicated by a particular node in a network. PPIs display predicted and verified protein interactions in an undirected, unweighted network. By combining these constructs, we elucidate genes that contribute most to the dysregulation of MDD, and possible treatments.

**Methods:** We utilized the NCBI's GEO resource to find dataset GSE98793, which compares the mRNA expression of peripheral blood in MDD to control patients. We employed GEOquery to download the normalized expression values of each sample. Additionally, we used GEO2R to identify the top 300 genes by greatest positive logFC value, which we used to create a PPI via STRING-db. The PPI was reconstructed inside of Python with the Networkx library. Using sample expression data, we assigned each edge the weight of the corresponding Spearman’s correlation coefficient squared (ρ2), which represents the monotonic variance explained by each variable on the other. We then calculated the betweenness centrality of each node and extracted the top ten genes contributing to information flow in the positive dysregulation network. To identify potential treatments, we used DGIdb to research the known interactions between drugs and each gene of interest. We used this information to identify drugs that modulate gene expression of the MDD group towards the control group, or that interact with two or more genes of interest.

**Results:** Our Networkx graph contained 153 nodes. Once the most central genes were identified, we referenced the genes with MDD literature. Eight of the top ten genes have been associated with MDD, which serves to verify our results. Genes that have not been associated with MDD include ITGB5 and CLU. Through our DGIdb analysis, we identified eighteen potential drugs to be used for MDD treatment. Sixteen drugs inhibit one of our genes of focus, which are all upregulated in MDD patients. Celecoxib has unknown interactions with two genes of focus. And Bevacizumab inhibits VEGFA and has an unknown interaction with MMP9.

**Conclusion:** Betweenness centrality in weighted PPIs could facilitate studies into the construction of disease dysregulation. Our experiment describes a pipeline for discovering critical genes and possible drugs to treat diseases. Additional research could investigate the role of MMP9, EGF, and SERPINB2 in MDD, as those genes have the greatest betweenness centrality in our network. We suggest drug studies focus on Bevacizumab, as that drug has the most absolute and beneficial interactions of all investigated compounds. Of final note, we used peripheral blood as a biomarker for brain transcriptomic dysregulation because that technique has been used for the study of other neurological diseases, however more research is needed to apply this technique to MDD with certainty.

**Keywords:** Betweenness Centrality, Network, PPI