Search Engine to Map FDA Approved Drugs to Diseases Based on Microarray Data Mined from GEO

Feldmann, Axel

Every drug that is introduced into an organism has an effect in up or down regulating the genes within that organism. A better understanding of the differentially expressed genes within that organism could lead to advances towards personalized medicine and treatments for many as of now untreatable conditions. This project seeks to find differentially expressed genes in existing drug-perturbation experiments found in the Gene Expression Omnibus (GEO). Accessing online, parsing, and then completing statistical tests upon the microarray data found in GEO, it was possible to find significant gene expression changes across over one thousand drug perturbation experiments in GEO. The data was tested with standard ANOVA methods as well as a novel method called Characteristic Direction in an effort to verify its effectiveness, and indeed Characteristic Direction proved more effective than ANOVA methods. Further data analysis permitted the coupling of the final data to liver cancer expression profiles and finding combination drug therapies that work to reverse the disease state. The entire project was accomplished in-silico and could possibly create a widespread new method for mining existing data on GEO. The data found can be analyzed with a myriad of different methods in order to find clusters of drugs, networks, and upstream regulators. While the drug-perturbation data found by the experiment is useful by itself, the methods devised to extract the data are made readily available and can be used to find any other set of data on GEO with considerable ease.

Awards Won:
Third Award of $1,000