

Predictive Models for Toxicity using rat in-vivo Gene Expression Data (ToxExpress)

Microarray technology is a powerful tool that allows simultaneous investigation of gene expression changes in thousands of genes to various stimuli. The application of microarray technology to toxicology has led to emergence of the discipline of toxicogenomics. Several previous studies have shown the utility in using gene expression profiles for prediction of drug-induced toxicity. In this talk, I will discuss the work done here at Pfizer in building predictive models for toxicity using rat *in-vivo* gene expression data (obtained from GeneLogic's ToxExpress database). Performance of several model-building methods (such as PLS, Random Forest, SVM) using expression profiles from different time points and doses was compared. We found that random forest models at high dose had better performance over the others. We also tested the utility of using protein-protein interaction networks to identify "important" features and perform network-based classification that for some types of toxicities improved performance of the classifier as compared to the gene-based approach.