Pathway-based Diagnostic Biomarker Identification for Post-Traumatic Stress Disorder

Ruoting Yang1, Bernie J. Daigle Jr1, Linda R. Petzold1,2,3, Francis J. Doyle III1,4

1Institute for Collaborative Biotechnologies, 2Department of Mechanical Engineering, 3Department of Computer Science, 4Department of Chemical Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106-9560

Tel: 805-893-2692 E-mail: ruoting, bdaigle@gmail.com, petzold, doyle@engineering.ucsb.edu

Summary: The aim of this research is to identify diagnostic biomarkers for post-traumatic stress disorder (PTSD) based on mouse transcriptomic data. We integrate biological pathways to filter redundant and noisy information from the genetic expression profiles. Both unsupervised and supervised classification methods were applied to search for robust pathway biomarkers.

Post-Traumatic Stress Disorder (PTSD) is an anxiety disorder that can develop after exposure to trauma. Current techniques to diagnose PTSD depend on psychological evaluation, which has large false positive rates. For precise diagnosis, it is compelling to identify biological parameters that can be easily measured in laboratory assays. Our goal is to find accurate and robust genetic indicators, or biomarkers, for PTSD. Both brain and blood samples of a mouse model were harvested and analyzed using DNA microarrays. Traditional biomarker models for classification select individual genes with largest discriminative power. However, because of heterogeneity, the gene-based biomarkers have poor reproducibility, poor interpretability, and large redundancy. Here, we propose a pathway-based method that integrates known biological pathways into biomarker models. Only the most discriminative genes in each pathway are selected and condensed into a measure of pathway activity. Then we applied both unsupervised hierarchical clustering and two state-of-the-art supervised classification methods, Support Vector Machine (SVM) and Nearest Shrunken Centroid (NSC), to search for robust biomarkers based on pathway activities.

![Fig. 1. Comparison of gene biomarkers and pathway biomarkers a) using unsupervised hierarchical clustering method. The mouse samples are listed in columns, while gene/pathway markers are in rows. The red/green color stands for upregulated/downregulated genes, and black means no difference. b) Comparison using supervised SVM classifier. The classifier performance was measured by AUC (Area under ROC curve), which plots true positive rate versus false positive rate for a binary classifier as its discrimination threshold is varied. Standard 100 times 5-fold cross-validation was conducted to report the mean AUC with its confidence interval.](image)

In conclusion, by incorporating prior biological pathway knowledge, the resulting classifiers enhance the robustness to noise and reduce potential redundancy of genes. In so doing, the pathway biomarkers achieve better predictive power and more clearly indicate the underlying mechanisms for PTSD.

Acknowledgement: We would like to thank our collaborators Dr. Marti Jett and Rasha Hammamieh for providing the experimental data. Project supported by the Institute for Collaborative Biotechnologies through Grant W911NF-09-D-0001 from the U. S. Army Research Office.