Novel pathway transcriptomics method greatly increases detection of molecular pathways associated with the drugs traits

C. Chatzinakos [1]; N. Gillespie; PGC-SUD consortium; K. Kendler; S.-A. Bacanu [1]

1) Virginia Commonwealth University (VCU), Richmond, VA.

Genetic signal detection in genome-wide association studies (GWAS) is improved by pooling information from multiple single nucleotide polymorphism (SNP). Because many genes influence trait via gene expression, it is of interest to combine information from Quantitative Trait Loci (eQTLs) in a gene or genes in the same pathway. Our group has developed a transcriptomic method that uses eQTLs to infer associations between traits and predicted expression of a gene under study i.e. Gene-level Joint Analysis of functional SNPs in Cosmopolitan Cohorts (JEPEGMIX). However, due to the O(n2) computational burden for computing linkage disequilibrium (LD) between numerous (n) SNPs, such methods are not yet applicable to large sets. To overcome this obstacle, we propose JEPEGMIX2, a novel O(n) transcriptomics method, which 1) computes LD for gene statistics and 2) uses LD and GWAS summary statistics to rapidly test for the association between a trait and the expression of genes even in very large pathways. It first computes chromosome arm pathway $\chi 2$ tests as Mahalanobis statistics of Z-scores for genes in the pathway and chromosome arm. Subsequently, due to variants on different chromosome arms being quasi-independent, JEPEGMIX2 pathway $\chi 2$ statistics and degrees of freedom are estimated based on the sum of their chromosome arm counterparts. JEPEGMIX2 controls type I error at or below nominal rates for the entire set of genes/pathways and each gene/pathway separately. To underscore JEPEGMIX2's potential for greatly increasing the power to uncover genetic signals over existing (non-transcriptomics) pathway methods, we applied JEPEGMIX2 to recent meta- and mega-analyses of licit and illicit substance use and misuse based on nicotine, alcohol and cannabis phenotypes. When compared to non-transcriptomic pathways, JEPEGMIX2 application results in a larger number of pathway signals, some of them quite unexpected.