Pleiotropy Detection Between All Gwas Catalog Traits and Alcohol Use Disorder Symptomalogy and Comorbidity in Admixed Population Via Machine Learning Method

Qian Peng[1], Kirk C. Wilhelmsen[2], Cindy L. Ehlers[1]

[1] Department of Neuroscience, The Scripps Research Institute, La Jolla, USA;
[2] Department of Genetics and Neurology, University of North Carolina, Chapel Hill, USA

Alcohol use disorders (AUD) are complex diseases that are postulated to have polygenic inheritance and are often comorbid with many other disorders. The comorbidities may be partially attributed by pleiotropy between AUD and other disorders. AUD rates vary across ethnic groups and populations, and are particularly high in some American Indian populations. Identifications of its specific genetic variants have been challenging.

Many methods exist for pleiotropy detection across the genomes. They either require all individuals in the sample collection having the phenotype data of all traits being studied, such as multivariate methods or PheWAS; or require GWAS summary statistics from populations of similar ancestries in cross-sample analysis, such as polygenic risk scores or cross-trait LD score regression methods. We have developed a method that can take a database of variant-trait association and determine if a subset of the diseases/traits are potentially pleiotropic with the disorder under study in a special admixed population.

To apply the method, we focused on the entire set of variants in the NHGRI-EBI catalog of published GWAS that had significant associations with any common diseases or traits. We investigated two independent populations: Amerindians (AI) from extended pedigrees, and Euro-Americans (EA) from UCSF Family Alcohol Study. Participants in both samples have whole genome sequencing data and were assessed using the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA). The AI is an admixed population, and both cohorts have relatedness. Linear mixed model (LMM) was thus employed to control for the population structures and familial relatedness. Regularized regressions including Lasso and elastic net were used in conjunction with LMM to select the best sets of SNPs from the GWAS catalog to be associated with alcohol-related life events that reflect the severity of AUD in both AI and EA, and similar severity measures for comorbid substance use disorders including marijuana, cocaine and stimulant, and multi-substance (MSUD) in AI. Once the variants were selected for each cohort, permutation tests were used to determine disease enrichments in the selected variants.

220--250 SNPs were selected for each trait and cohort. Inflammatory measurements were the most enriched category for alcohol-related life events and MSUD in AI, followed by other traits and neurological disorders for alcohol. In contrast, liver enzyme measurements were the most enriched for alcohol-related life events in EA. Neurological disorders were the most enriched for marijuana, while immune system disorders in addiction to neurological disorders were the most enriched for cocaine and stimulant use in AI. We further used functional and network analyses to investigate the underlying biological functions of the selected potentially pleiotropic variants.

Our results suggest that there is evidence for pleiotropy between alcohol-related symptomatology and other complex disorders, and the underlying genetic factors are likely polygenic and population-dependent.

(Supported by: K25AA025095, AA010201, DA030976)

Keywords: Pleiotropy; Polygenic; Complex Traits; Variant Selection; Alcohol Use Disorders