Submitter Name: Dongbing Lai Submitted Email: dlai@iu.edu

Genome-wide meta-analyses of cross substance use disorders in European, African, and Latino ancestry populations

Dongbing Lai[1], Michael Zhang[1], Nick Green[1], Marco Abreu[1], Tae-Hwi Schwantes-An[1], Clarissa C. Parker[2], Shanshan Zhang[3,4], Fulai Jin[3], Anna Sun[5], Pengyue Zhang[5], Howard J. Edenberg[1,6], Yunlong Liu[1], Tatiana Foroud[1]

[1]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; [2]Department of Psychology and Program in Neuroscience, Middlebury College, Middlebury, VT; [3]Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH; [4]Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, MA; [5]Department of Biostatistics and Heath Data Science, Indiana University School of Medicine, Indianapolis, IN; [6]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN

Many individuals suffer from multiple substance use disorders (SUDs) simultaneously. They typically have worse consequences and require complicated and costly treatment. SUDs share common features and previous studies have demonstrated the existence of SUD-shared genes. Searching for SUD-shared genes will help us understand the etiologies of SUDs and facilitate the development of efficient and effective prevention and treatment strategies. In this study, we performed the largest multivariate analyses to date to search for SUD-shared genes using samples of European (EA, effective sample size (EN)=1,467,929), African (AA, EN=159,000), and Latino (LA, EN=45,727) ancestries. We also performed cross-ancestry analyses (total EN=1,672,656). As we were searching for SUD-shared genes, we innovatively focused on variants having cross-SUD and cross-population concordant effects and identified 45 loci. Through gene-based analyses, gene mapping, and gene prioritization, we identified 250 SUD-shared genes. Using the single-cell RNA sequencing data generated by the BRAIN initiative cell census network, we found that SUD-shared genes were highly expressed in amygdala, cortex, hippocampus, hypothalamus, and thalamus, primarily in neuronal cells. Cross-SUD concordant variants explained ~50% of the heritability of each SUD in EA. Using samples from the All of Us research program and the Indiana Biobank, we found that those 5% individuals having the highest polygenic scores were approximately twice as likely to have SUDs as others in EA and LA. Using real-world data of medical and pharmacy claims from a large national insurer cohort (Clinformatics®), we identified five drugs targeting identified SUD-shared genes that may be repurposed to treat SUDs.