

Supplemental Figures
Hatoum et al., Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders

Supplemental Figure 1. Circos Plots from Hi-C

Supplemental Figure 2. Manhattan plots of substance-specific findings

Supplemental Figure 3. African Ancestry (AA) Multivariate GWAS Manhattan Plots

Supplemental Figure 4. Manhattan Plot of Cross-Ancestry GWAS for pleiotropic variants. Enrichment of Chemical and Perturbagen Gene Sets

Supplemental Figure 5. Polygenic Architecture of Substance Use Disorder Phenotypes.

Supplemental Figure 6. Enrichment of Chemical and Perturbagen Gene sets in FUMA for the transcriptome-wide association analysis (using MultiXcan) of *addiction-rf*.

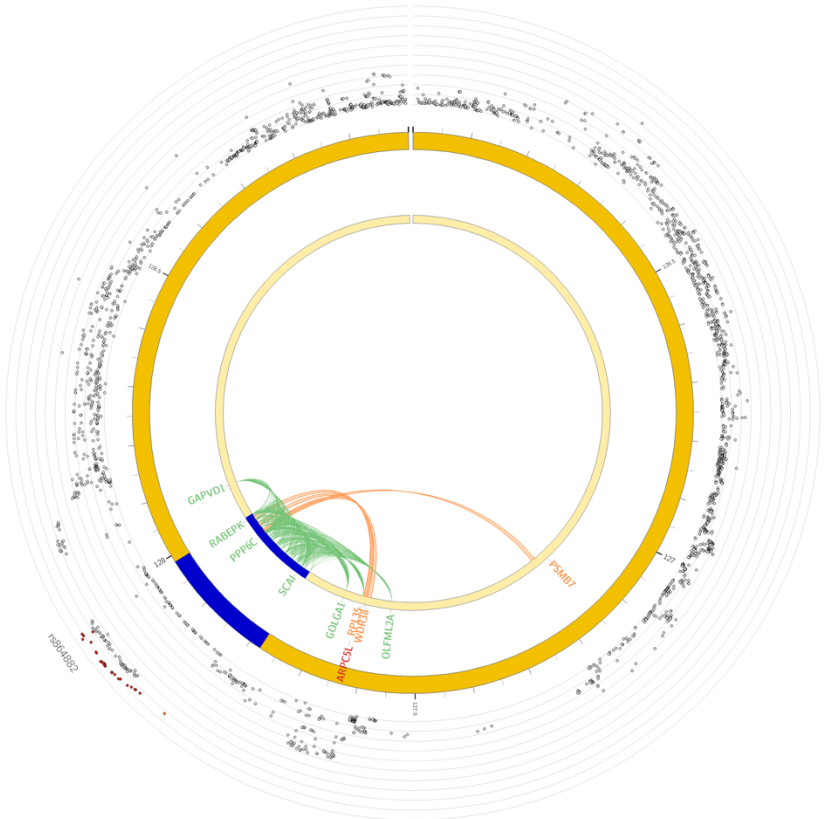
Supplemental Figure 7. PheWAS of *Addiction-rf* accounting for (A) SUD diagnosis and (B) Tobacco Dependence in the BioVU.

Supplemental Figure 8. Factor Structure of 4 SUD GWAS.

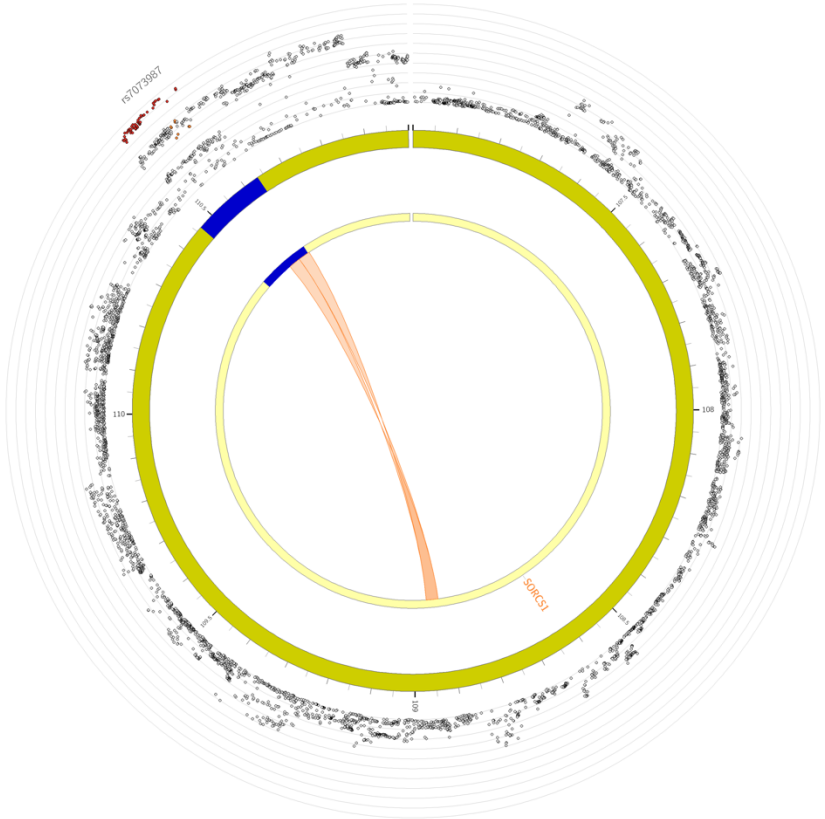
A circular genome browser view of the human genome, likely representing a specific chromosome or a circularized genomic region. The outermost track shows genomic coordinates in megabases (Mb), ranging from approximately 170.5 to 171.5 Mb. The inner tracks display various genomic features and gene annotations. Key genes and features labeled include:

- Genes and Transcripts:** *ANKRD45*, *ANKRD45-AS1*, *ANKRD45-AS2*, *ANKRD45-AS3*, *ANKRD45-AS4*, *ANKRD45-AS5*, *ANKRD45-AS6*, *ANKRD45-AS7*, *ANKRD45-AS8*, *ANKRD45-AS9*, *ANKRD45-AS10*, *ANKRD45-AS11*, *ANKRD45-AS12*, *ANKRD45-AS13*, *ANKRD45-AS14*, *ANKRD45-AS15*, *ANKRD45-AS16*, *ANKRD45-AS17*, *ANKRD45-AS18*, *ANKRD45-AS19*, *ANKRD45-AS20*, *ANKRD45-AS21*, *ANKRD45-AS22*, *ANKRD45-AS23*, *ANKRD45-AS24*, *ANKRD45-AS25*, *ANKRD45-AS26*, *ANKRD45-AS27*, *ANKRD45-AS28*, *ANKRD45-AS29*, *ANKRD45-AS30*, *ANKRD45-AS31*, *ANKRD45-AS32*, *ANKRD45-AS33*, *ANKRD45-AS34*, *ANKRD45-AS35*, *ANKRD45-AS36*, *ANKRD45-AS37*, *ANKRD45-AS38*, *ANKRD45-AS39*, *ANKRD45-AS40*, *ANKRD45-AS41*, *ANKRD45-AS42*, *ANKRD45-AS43*, *ANKRD45-AS44*, *ANKRD45-AS45*, *ANKRD45-AS46*, *ANKRD45-AS47*, *ANKRD45-AS48*, *ANKRD45-AS49*, *ANKRD45-AS50*, *ANKRD45-AS51*, *ANKRD45-AS52*, *ANKRD45-AS53*, *ANKRD45-AS54*, *ANKRD45-AS55*, *ANKRD45-AS56*, *ANKRD45-AS57*, *ANKRD45-AS58*, *ANKRD45-AS59*, *ANKRD45-AS60*, *ANKRD45-AS61*, *ANKRD45-AS62*, *ANKRD45-AS63*, *ANKRD45-AS64*, *ANKRD45-AS65*, *ANKRD45-AS66*, *ANKRD45-AS67*, *ANKRD45-AS68*, *ANKRD45-AS69*, *ANKRD45-AS70*, *ANKRD45-AS71*, *ANKRD45-AS72*, *ANKRD45-AS73*, *ANKRD45-AS74*, *ANKRD45-AS75*, *ANKRD45-AS76*, *ANKRD45-AS77*, *ANKRD45-AS78*, *ANKRD45-AS79*, *ANKRD45-AS80*, *ANKRD45-AS81*, *ANKRD45-AS82*, *ANKRD45-AS83*, *ANKRD45-AS84*, *ANKRD45-AS85*, *ANKRD45-AS86*, *ANKRD45-AS87*, *ANKRD45-AS88*, *ANKRD45-AS89*, *ANKRD45-AS90*, *ANKRD45-AS91*, *ANKRD45-AS92*, *ANKRD45-AS93*, *ANKRD45-AS94*, *ANKRD45-AS95*, *ANKRD45-AS96*, *ANKRD45-AS97*, *ANKRD45-AS98*, *ANKRD45-AS99*, *ANKRD45-AS100*.
- Other Genes:** *ANKRD45-AS101*, *ANKRD45-AS102*, *ANKRD45-AS103*, *ANKRD45-AS104*, *ANKRD45-AS105*, *ANKRD45-AS106*, *ANKRD45-AS107*, *ANKRD45-AS108*, *ANKRD45-AS109*, *ANKRD45-AS110*, *ANKRD45-AS111*, *ANKRD45-AS112*, *ANKRD45-AS113*, *ANKRD45-AS114*, *ANKRD45-AS115*, *ANKRD45-AS116*, *ANKRD45-AS117*, *ANKRD45-AS118*, *ANKRD45-AS119*, *ANKRD45-AS120*, *ANKRD45-AS121*, *ANKRD45-AS122*, *ANKRD45-AS123*, *ANKRD45-AS124*, *ANKRD45-AS125*, *ANKRD45-AS126*, *ANKRD45-AS127*, *ANKRD45-AS128*, *ANKRD45-AS129*, *ANKRD45-AS130*, *ANKRD45-AS131*, *ANKRD45-AS132*, *ANKRD45-AS133*, *ANKRD45-AS134*, *ANKRD45-AS135*, *ANKRD45-AS136*, *ANKRD45-AS137*, *ANKRD45-AS138*, *ANKRD45-AS139*, *ANKRD45-AS140*, *ANKRD45-AS141*, *ANKRD45-AS142*, *ANKRD45-AS143*, *ANKRD45-AS144*, *ANKRD45-AS145*, *ANKRD45-AS146*, *ANKRD45-AS147*, *ANKRD45-AS148*, *ANKRD45-AS149*, *ANKRD45-AS150*, *ANKRD45-AS151*, *ANKRD45-AS152*, *ANKRD45-AS153*, *ANKRD45-AS154*, *ANKRD45-AS155*, *ANKRD45-AS156*, *ANKRD45-AS157*, *ANKRD45-AS158*, *ANKRD45-AS159*, *ANKRD45-AS160*, *ANKRD45-AS161*, *ANKRD45-AS162*, *ANKRD45-AS163*, *ANKRD45-AS164*, *ANKRD45-AS165*, *ANKRD45-AS166*, *ANKRD45-AS167*, *ANKRD45-AS168*, *ANKRD45-AS169*, *ANKRD45-AS170*, *ANKRD45-AS171*, *ANKRD45-AS172*, *ANKRD45-AS173*, *ANKRD45-AS174*, *ANKRD45-AS175*, *ANKRD45-AS176*, *ANKRD45-AS177*, *ANKRD45-AS178*, *ANKRD45-AS179*, *ANKRD45-AS180*, *ANKRD45-AS181*, *ANKRD45-AS182*, *ANKRD45-AS183*, *ANKRD45-AS184*, *ANKRD45-AS185*, *ANKRD45-AS186*, *ANKRD45-AS187*, *ANKRD45-AS188*, *ANKRD45-AS189*, *ANKRD45-AS190*, *ANKRD45-AS191*, *ANKRD45-AS192*, *ANKRD45-AS193*, *ANKRD45-AS194*, *ANKRD45-AS195*, *ANKRD45-AS196*, *ANKRD45-AS197*, *ANKRD45-AS198*, *ANKRD45-AS199*, *ANKRD45-AS200*.
- Other Genes:** *ANKRD45-AS201*, *ANKRD45-AS202*, *ANKRD45-AS203*, *ANKRD45-AS204*, *ANKRD45-AS205*, *ANKRD45-AS206*, *ANKRD45-AS207*, *ANKRD45-AS208*, *ANKRD45-AS209*, *ANKRD45-AS210*, *ANKRD45-AS211*, *ANKRD45-AS212*, *ANKRD45-AS213*, *ANKRD45-AS214*, *ANKRD45-AS215*, *ANKRD45-AS216*, *ANKRD45-AS217*, *ANKRD45-AS218*, *ANKRD45-AS219*, *ANKRD45-AS220*, *ANKRD45-AS221*, *ANKRD45-AS222*, *ANKRD45-AS223*, *ANKRD45-AS224*, *ANKRD45-AS225*, *ANKRD45-AS226*, *ANKRD45-AS227*, *ANKRD45-AS228*, *ANKRD45-AS229*, *ANKRD45-AS230*, *ANKRD45-AS231*, *ANKRD45-AS232*, *ANKRD45-AS233*, *ANKRD45-AS234*, *ANKRD45-AS235*, *ANKRD45-AS236*, *ANKRD45-AS237*, *ANKRD45-AS238*, *ANKRD45-AS239*, *ANKRD45-AS240*, *ANKRD45-AS241*, *ANKRD45-AS242*, *ANKRD45-AS243*, *ANKRD45-AS244*, *ANKRD45-AS245*, *ANKRD45-AS246*, *ANKRD45-AS247*, *ANKRD45-AS248*, *ANKRD45-AS249*, *ANKRD45-AS250*, *ANKRD45-AS251*, *ANKRD45-AS252*, *ANKRD45-AS253*, *ANKRD45-AS254*, *ANKRD45-AS255*, *ANKRD45-AS256*, *ANKRD45-AS257*, *ANKRD45-AS258*, *ANKRD45-AS259*, *ANKRD45-AS260*, *ANKRD45-AS261*, *ANKRD45-AS262*, *ANKRD45-AS263*, *ANKRD45-AS264*, *ANKRD45-AS265*, *ANKRD45-AS266*, *ANKRD45-AS267*, *ANKRD45-AS268*, *ANKRD45-AS269*, *ANKRD45-AS270*, *ANK*

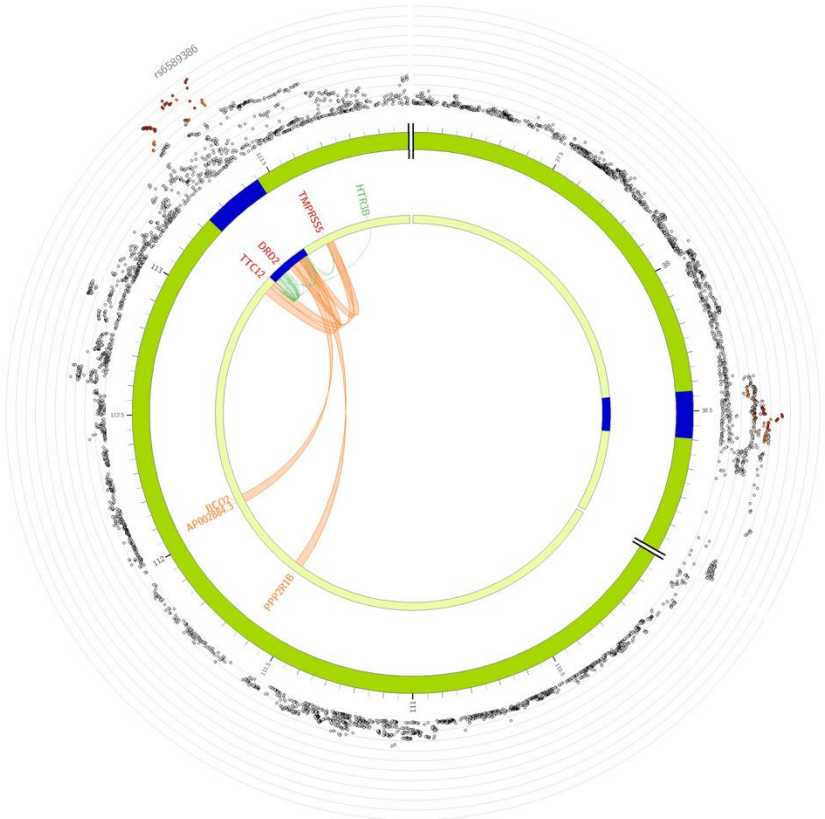
Chromosome 9



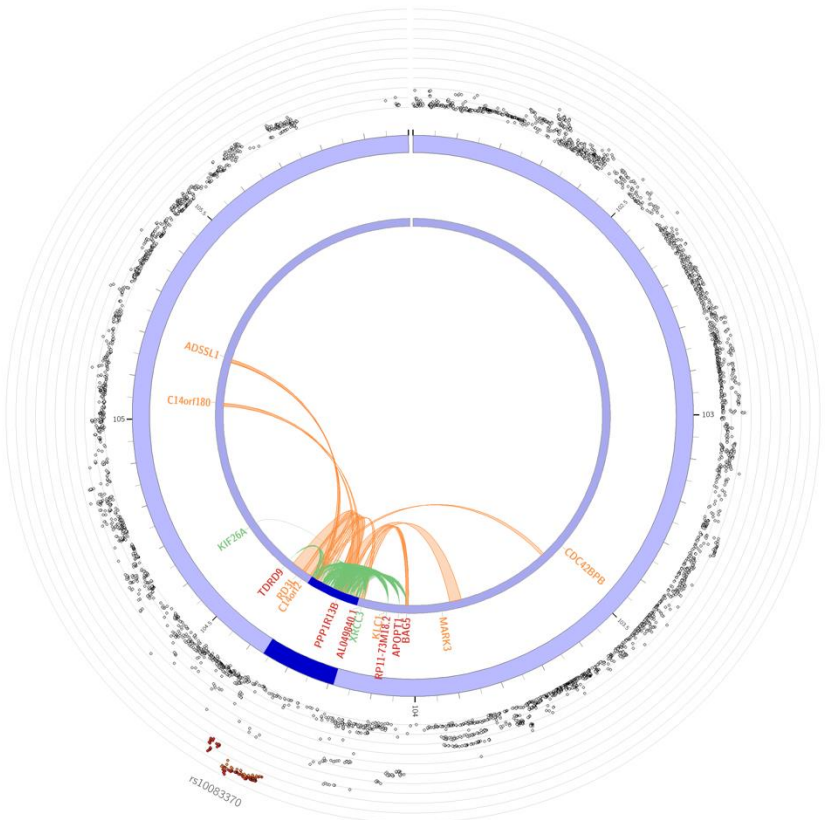
Chromosome 10



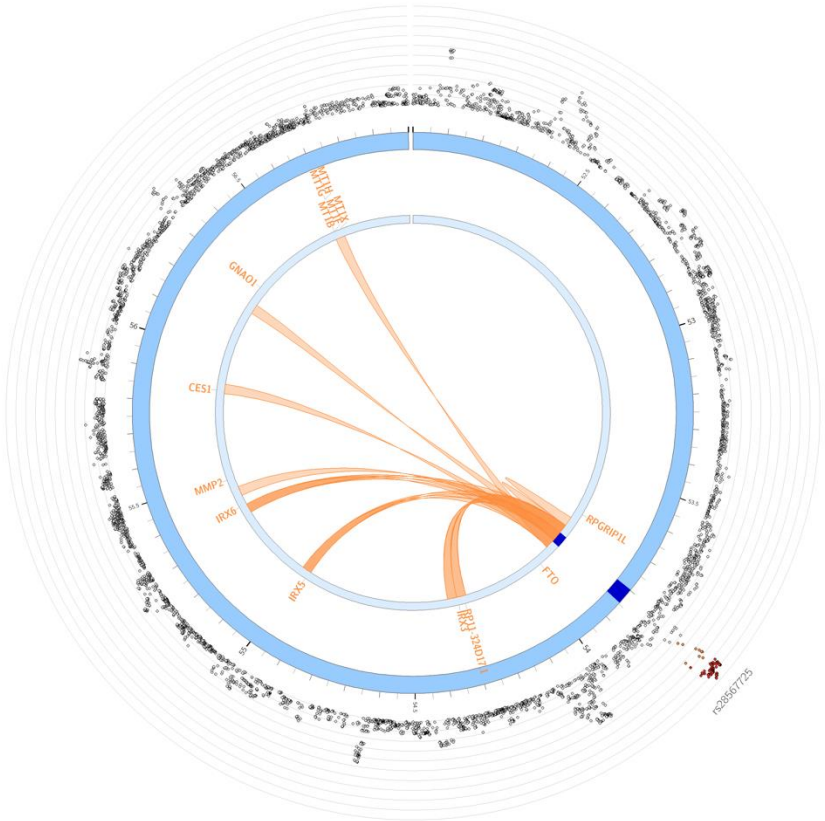
Chromosome 11



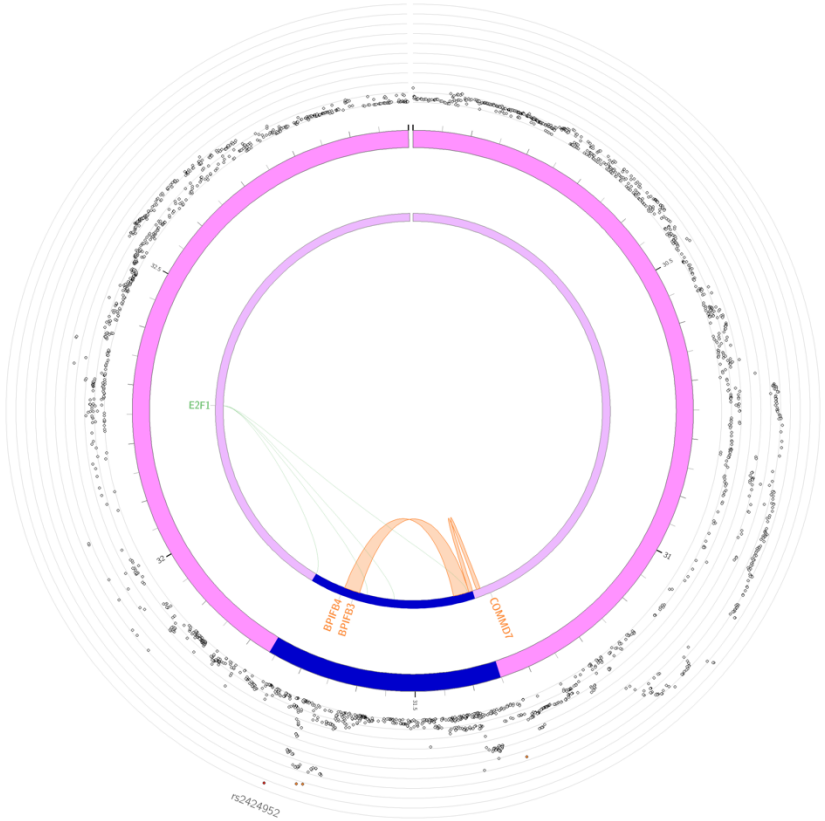
Chromosome 14



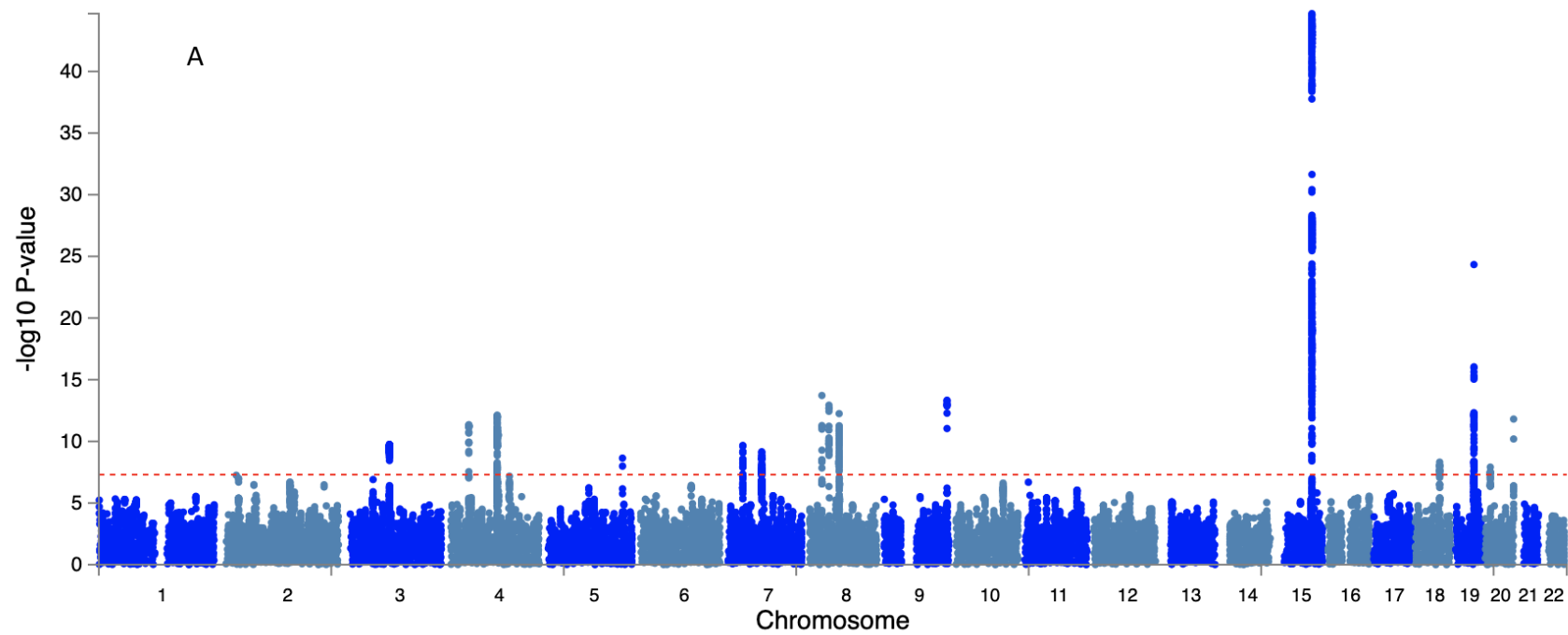
Chromosome 16

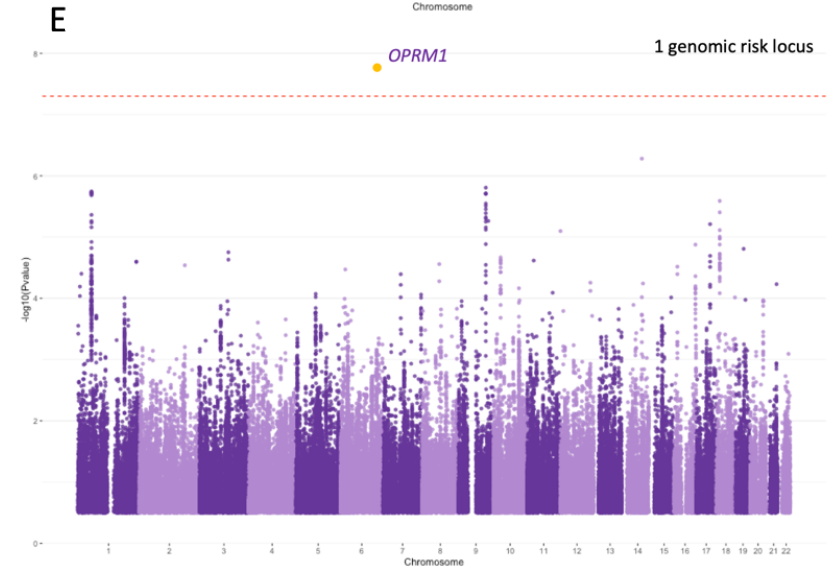
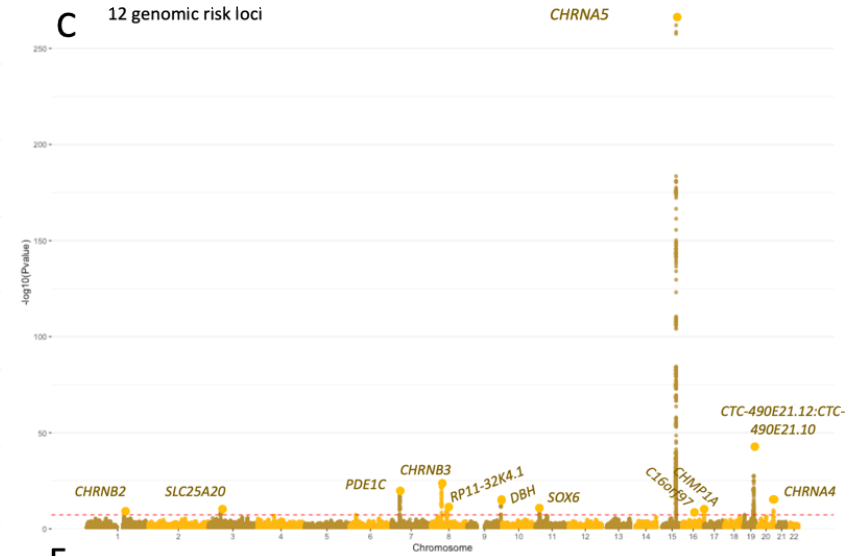
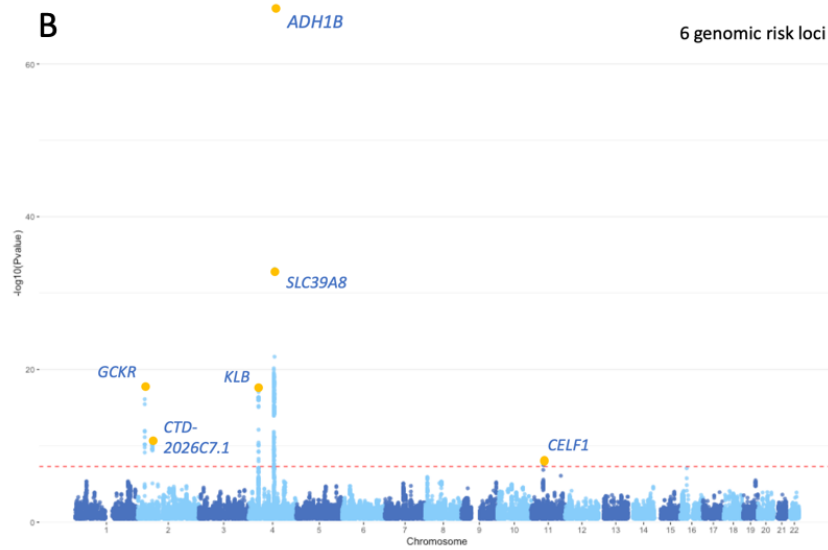


Chromosome 20

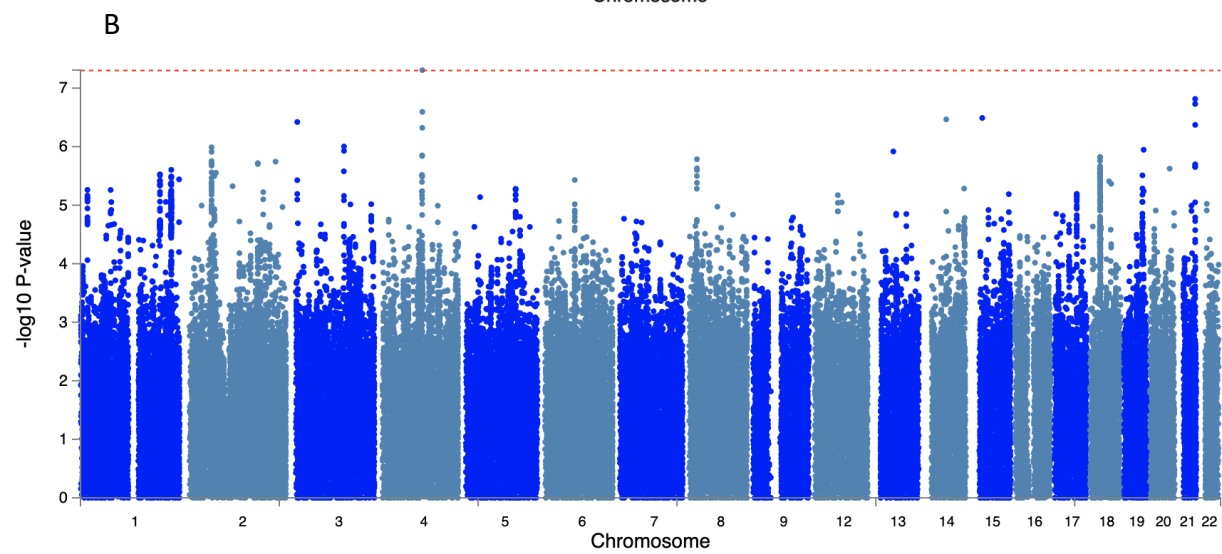
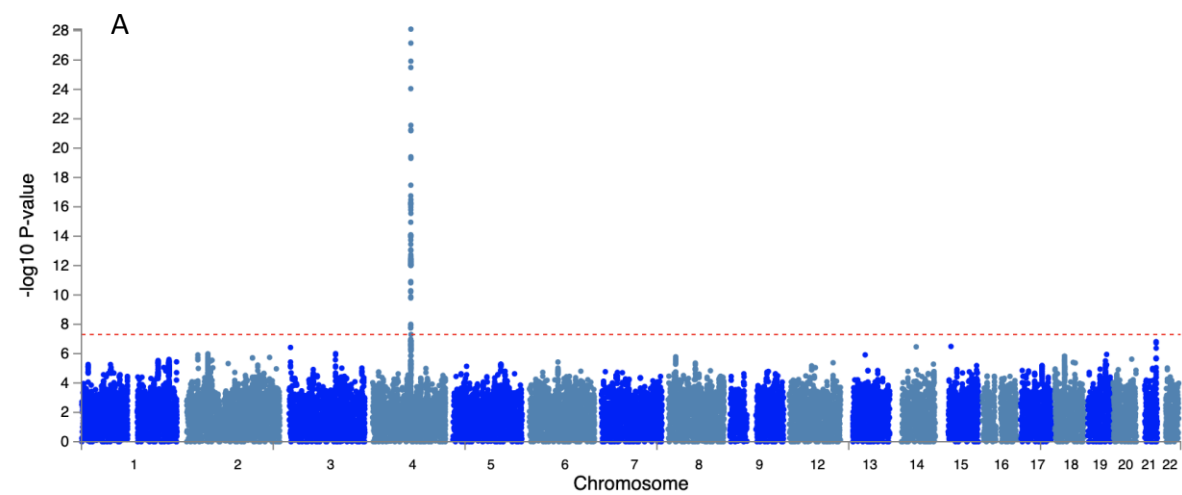


Supplemental Figure 1. Circos Plots for chromosomes where top SNPs demonstrated chromatin contact. Chromosome is a circle with SNPs plotted by their negative log 10 p-value (lead SNPs annotated). Orange lines represent chromatin contact with a gene. Green lines represent cis-eQTLs with a gene.

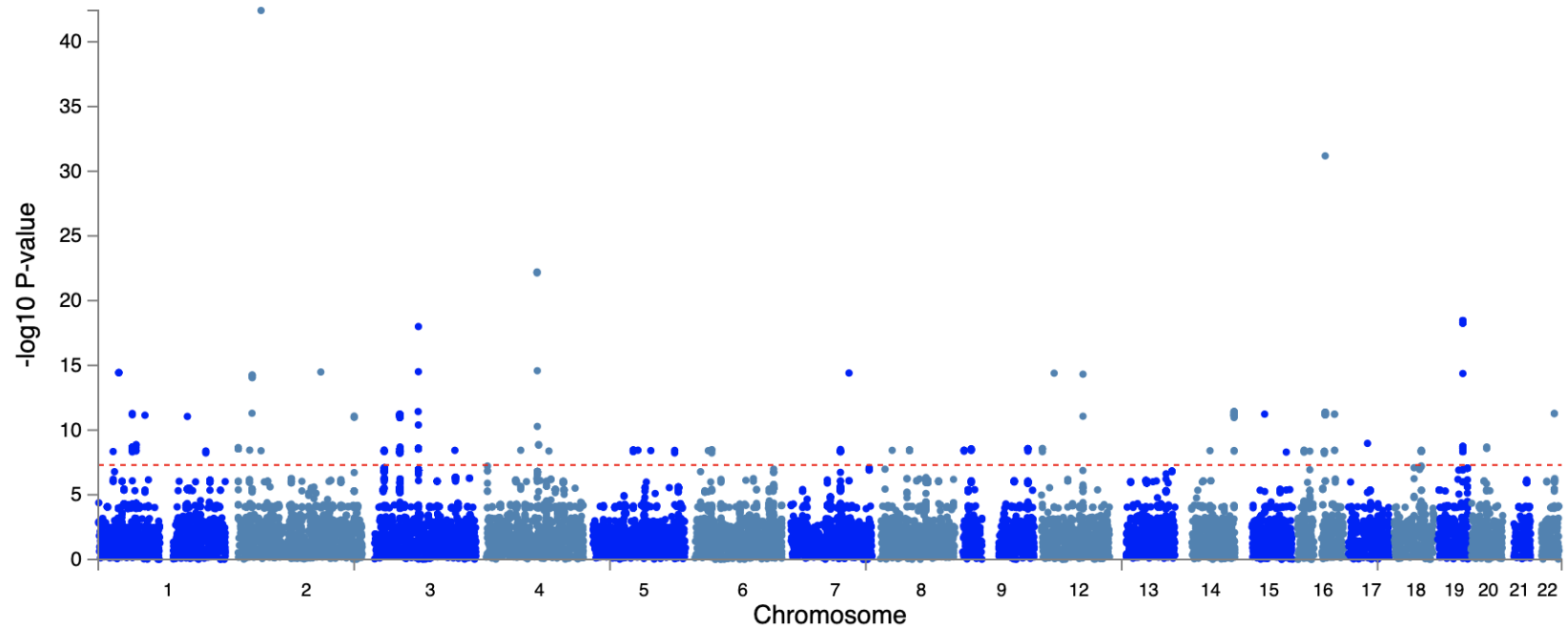




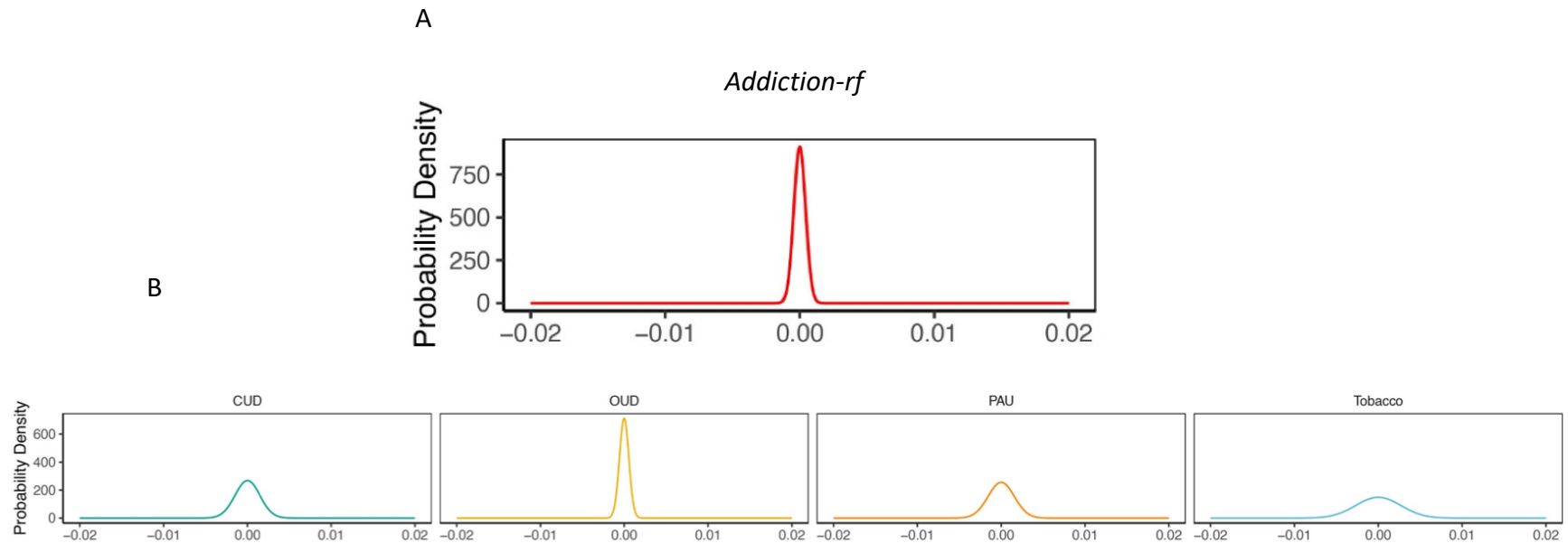
Supplemental Figure 2. Manhattan plots of substance-specific findings. (A) Initial Q-SNP analysis p-value. Panels B-E are based on ASSET analyses. (B) SNPs associated with PAU only (C) SNPs associated with PTU only (D) SNPs associated with CUD only and (E) SNPs associated with OUD only



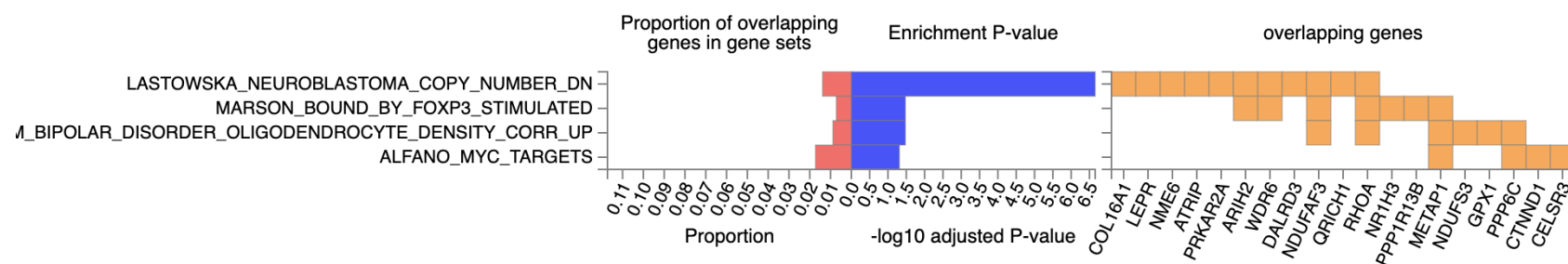
Supplemental Figure 3. African Ancestry (AA) Multivariate GWAS Manhattan Plots. **(A)** GWAS results from ASSET including all SNPs, note the significant peak on chromosome 4 that maps to ADH1B and was alcohol specific. **(B)** Loci showing evidence of pleiotropy between at least two different use disorder variables.



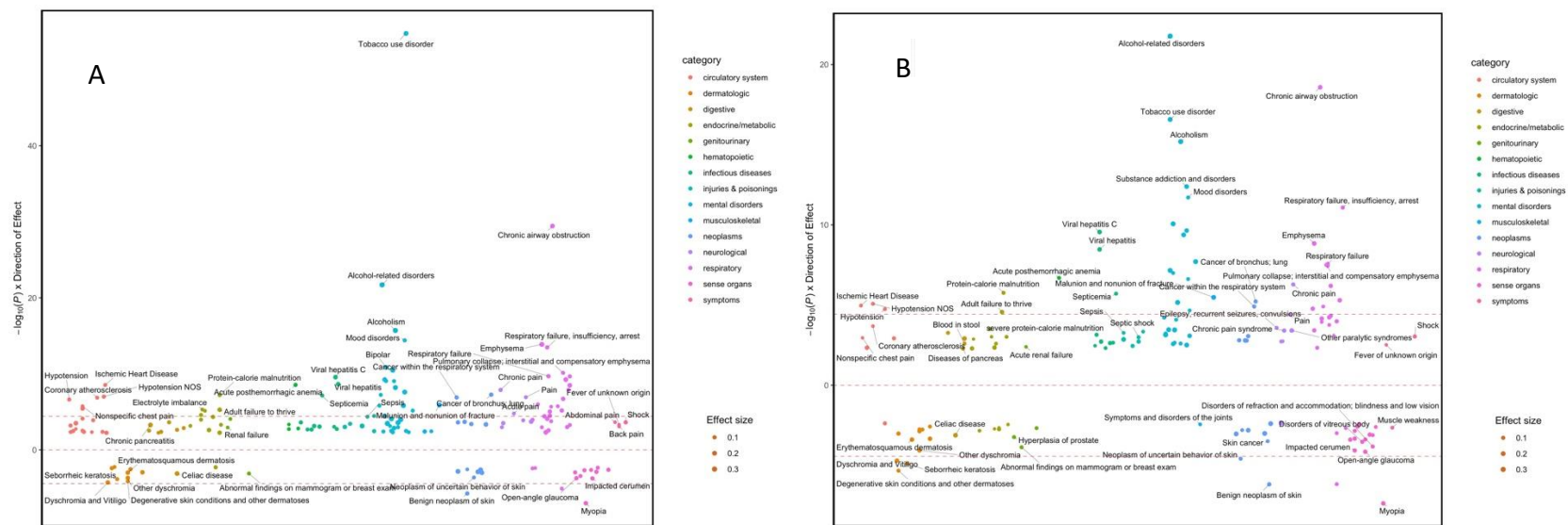
Supplemental Figure 4. Manhattan Plot of Cross-Ancestry GWAS for pleiotropic variants. The plot is sparse and only includes 317,447 loci that existed in both ancestral groups and showed some evidence of pleiotropy within each ancestral group.



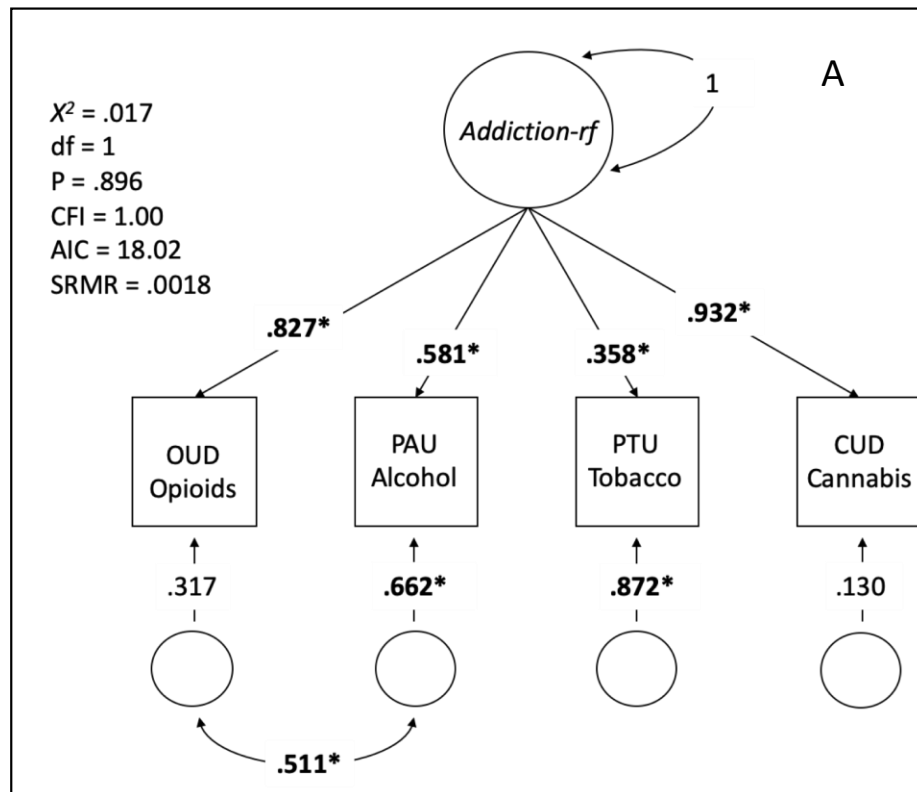
Supplemental Figure 5. Polygenic Architecture of Substance Use Disorder Phenotypes. Distribution of SNP effect sizes for the *addiction-rf* and substance-specific components. Taller distributions mean generally smaller effect sizes. (A) Effect size distribution for *addiction-rf*. (B) Distributions for individual substance use disorder/problematic substance use GWAS that comprised *addiction-rf*. Broadly, the *addiction-rf* had higher polygenicity. PAU = problematic alcohol use, PTU = problematic tobacco use, CUD = cannabis use disorder, OUD = opioid use disorder



Supplemental Figure 6. Enrichment of Chemical and Perturbagen Gene sets in FUMA for the transcriptome-wide association analysis (using MultiXcan) of *addiction-rf*. We plotted the proportion of the genes that overlap with the total set (red), genes individually overlapping (gold), and the Bonferroni corrected $-\log_{10}$ enrichment p-value for each gene set.



Supplemental Figure 7. PheWAS of Addiction-*rf* accounting for (A) SUD diagnosis and (B) Tobacco Dependence in the BioVU electronic health record database (N=66,915). Results are shown by level of statistical significance ($-\log_{10}(p)$ of association on y-axis) and in broad categories. Analyses were conducted using logistic regression, controlling for median age on record, reported gender, and first 10 genetic ancestry PCs. The dashed lines represent Bonferroni correction for 1,335 phecodes ($p_{\text{bon}}=3.745\text{e-}05$). The top associations are annotated. Symptoms represents phenotypes like back pain, loss of height, muscle weakness etc that are not indicative of a disease alone.



Supplemental Figure 8. Factor Structure of 4 SUD GWAS. Factor loadings and model fit were recreated from Hatoum et al. 2021.

We allowed all 4 SUD categories to load on a latent factor with a residual correlation between PAU and OUD was added to account

for their assessment using electronic health records in the MVP cohort. *Addiction-rf=The Addiction Risk-Factor*. **Bold*** represents significance at $p < .05$.