

Novel insights into the common heritable liability to addiction: a multivariate genome-wide association study

Tabea Schoeler^{1,2*}, PhD
Jessie Baldwin^{2,3}, PhD
Andrea Allegrini², PhD
Wikus Barkhuizen², PhD
Andrew McQuillin⁴, PhD
Nicola Pirastu⁵, PhD
Zoltán Kutalik^{1,6,7}, PhD
Jean-Baptiste Pingault^{1,3}, PhD

¹ Department of Computational Biology, University of Lausanne, Switzerland

² Department of Clinical, Educational and Health Psychology, Division of Psychology and Language Sciences, University College London, London, UK

³ Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

⁴ Division of Psychiatry, University College London, London, UK

⁵ Centre for Global Health Research, Usher Institute, University of Edinburgh, UK

⁶ University Center for Primary Care and Public Health, University of Lausanne, Switzerland

⁷ Swiss Institute of Bioinformatics, Lausanne, Switzerland

*Correspondence: t.schoeler@ucl.ac.uk

Abstract

Addiction to nicotine, alcohol and cannabis commonly co-occurs, which is thought to partly stem from a common heritable liability. To elucidate its genetic architecture, we modelled the common liability to addiction, inferred from genetic correlations among six measures of dependence and frequency of use of nicotine, alcohol and cannabis. Forty-two genetic variants were identified in the multivariate genome-wide association study on the common liability to addiction, of which 67% were novel and not associated with the six phenotypes. Mapped genes highlighted the role of dopamine (e.g., dopamine D2 gene), and showed enrichment for several components of the central nervous systems (e.g., mesocorticolimbic brain regions) and molecular pathways (dopaminergic, glutamatergic, GABAergic) that are thought to modulate drug reinforcement. Genetic correlations with other traits were most prominent for reward-related behaviours (e.g., risk-taking, cocaine and heroin use) and mood (e.g., depression, insomnia). These genome-wide results triangulate and expand previous preclinical and human studies focusing on the neurobiological substrates of addiction, and help to elucidate the common genetic architecture underlying addiction.

Introduction

Addiction is one of the leading causes of preventable harms¹, and considerable research efforts are made to better understand its aetiology. Addiction is typically not restricted to just one class of substance, as dependence of accessible psychoactive substances such as nicotine, alcohol and cannabis often co-occurs². This co-occurring pattern of use has been shown to be particularly detrimental to the individual and to society as a whole³.

Aetiological models posit that addiction to multiple substances stems from a common liability to addiction^{4,5} – a latent continuous trait accounting for the shared risk of developing addiction to different substances. Based on findings from genomic^{6–8} and behavioural genetic studies^{9,10}, it is assumed that this common liability includes a genetic component. Indeed, genetic correlations between use of different classes of psychoactive substances are substantial, as estimated in twin (up to $r_g \sim 0.89$ ^{11–13}) and genome-wide association (GWA) studies (up to $r_g \sim 0.70$ ^{7,14–16}). The underlying molecular mechanisms of this common heritable liability to addiction are, however, not fully understood. While it has been shown that genotypic variations contribute to the common heritable liability to addiction¹⁷, an investigation into the specific genome-wide effects has yet to be conducted. Furthermore, although increasingly large GWA studies have identified growing numbers of genetic risk variants associated with individual substance use phenotypes⁷, it remains unclear as to whether associated risk variants reflect shared (common) or non-shared (substance specific) risk across different addiction phenotypes. Some identified genetic variants likely operate through substance-specific pharmacological pathways, as is the case for variants affecting nicotinic receptors (e.g. genes coding for nicotinic acetylcholine receptors, such as *CHRNA3-CHRNA5*¹⁸) or alcohol metabolism (e.g. variants in the alcohol dehydrogenases gene family, such as *ADH1B*^{7,15}, *ADH1C*^{14,15}). Other variants may affect common pathways, such as variants associating with two or more classes of psychoactive substances (e.g. *BDNF*^{7,19,20}, *PDE4B*^{7,20,21} or *DRD2*^{7,15,18,20,22}) or behavioural phenotypes (e.g., the top variant identified for cannabis use disorder, which also associates with ADHD and risk-taking²³).

The aforementioned genetic overlap complicates research on causes and consequences of addiction, and distilling shared (common) from non-shared (substance specific) genetic risk is pivotal to the interpretation of genome-wide discoveries. One way of scrutinizing putative pleiotropic variants is to explicitly model the genetic overlap among different phenotypes indexing addiction, using multivariate methods such as genomic structural equation modelling (genomic SEM²⁴). Applying this method has already been helpful in characterising shared genetic influences across dimensions of psychopathology^{24–31} and cognition^{32–35}. In addition to assessing shared effects of suspected pleiotropic variants, genomic SEM also has the potential to identify novel genetic variants not previously identified in univariate GWA studies on individual phenotypes²⁹. This is expected, since shared risk is thought to be expressed indirectly via the common liability, resulting in inherently small effects, which hampers detection of pleiotropic variants in univariate GWA analyses. A multivariate GWA can therefore boost discovery of shared variants directly associated with a common heritable liability. While genetically informed methods using polygenic scores have already explored risk factors involved in the common liability to addiction^{36,37}, a multivariate GWA of the shared and non-shared genetic architecture can further deepen our understanding of biological pathways underlying addiction to multiple substances. Indeed, leveraging genetically informed methods would allow us to revisit long-theorized biological pathways underlying addiction, and to triangulate evidence from behavioural genetic¹¹, brain

imaging³⁸, candidate gene³⁹ and preclinical studies⁴⁰ focusing on the role of genetics and neural substrates of addiction. Together, such triangulated findings would help researchers and clinicians to better understand biological and developmental pathways involved in risk of developing addiction to commonly used and abused psychoactive substances.

To unravel the genetic architecture underlying addiction to nicotine, alcohol and cannabis, here we conduct a multivariate GWA analysis on their common heritable liability. To model the common liability to addiction to these substances, we include phenotypes indexing clinical (diagnosis of dependence) as well as quantitative (frequency of use) measures of addiction to nicotine, alcohol and cannabis. More specifically, we conduct a multivariate GWA of the common heritable liability, with the aim to

- a. identify putative genetic variants associated with the common liability (i.e., shared/pleiotropic variants) and variants specific to the use of different classes of substances (i.e., non-shared)
- b. characterize the functional features of genetic variants associated with the common liability
- c. assess the genetic correlations between the common liability with other complex traits
- d. evaluate the validity of the causal claims imposed by a common liability model of addiction

Materials and methods

GWA summary datasets

We screened GWA summary statistics of addiction-related phenotypes for the most commonly used and misused psychoactive substances, namely nicotine, alcohol and cannabis. For each substance class, we included one clinical (diagnosis of dependence) and one quantitative (frequency of use) measure of addiction. The following summary statistics were included, derived from samples of individuals of European ancestry: Data on alcohol use disorder ($n=28,757$)¹⁴ and cannabis use disorder ($n=358,534$)²³ was obtained from the Substance Use Disorders working group of the Psychiatric Genomics Consortium (PGC-SUD). Nicotine dependence ($n=244,890$) was taken from GWAS ATLAS⁴¹ (cf. [URLs](#)). Frequency of cigarette ($n=245,876$) and alcohol use ($n=513,208$), obtained from the GWA meta-analysis by Liu et al.⁷. The phenotypes were measured as the number of cigarettes/alcoholic drinks consumed during periods of consumption (assessed through the questions “How many cigarettes do/did you smoke per day?” and “How many drinks do/did you have each week/month?”). Frequency of cannabis use ($n=24,798$, from the question “Considering when you were taking cannabis most regularly, how often did you take it?”) was obtained from the Neale Lab UKBB summary statistics (cf. [URLs](#)). Additional details of each of the included summary statistic files can be found in **sTable 1** (Supplement). As the included phenotypes are considered to reflect different dimensions of addiction, we coined their shared genetic architecture the ‘common heritable liability to addiction’.

Genomic model of the common heritable liability to addiction

We first estimated the genetic correlations (r_g) among the individual phenotypes using genomic structural equation modelling (genomic SEM²⁴) version 0.0.3. The method uses an extension of LD-score regression⁴² and accounts for sample overlap across studies through the LD-score intercept. In confirmatory factor analysis, we fitted a structural equation model with a single latent factor, representing a common liability to addiction to nicotine, alcohol and cannabis, onto which the six indicators loaded (**Figure 1b**). Equality constraints were imposed on paths belonging to the same pattern of substance use, i.e., equal weights across measures of dependence, and equal weights across measures of frequency of use. Correlated residuals were included to allow for within-substance class associations, as depicted in **Figure 1b**. The Diagonally Weighted Least Squares (DWLS) estimator was used and model fit was assessed based on the Comparative Fit Index (CFI) and the standardized root mean square residual (SRMR).

Multivariate genome-wide association analysis

For the multivariate GWA on the common heritable liability to addiction, the summary statistics for the individual addiction phenotypes were restricted to single nucleotide polymorphisms (SNPs) contained in the 1000 genomes phase 3 reference sample (with a minor allele frequency (MAF) > 1%) and SNPs that were present in all GWA summary datasets included in the analysis. Genomic control was applied to all summary statistics showing evidence of uncontrolled confounding (LD score intercept > 1), by multiplying standard errors by the LD score intercept. To identify lead SNPs after conducting the GWA on the common liability, we selected LD-independent SNPs ($r^2 < 0.1$ within 250 kb) based on genome-wide significance ($p < 5 \times 10^{-8}$).

To determine whether the effects of the identified lead SNPs are likely to act through the common liability, we applied the heterogeneity test as implemented in Genomic SEM. The

resulting Q-statistic (Q_{SNP}) is a χ^2 -distributed test statistic, with significant Q_{SNP} estimates ($p < 5 \times 10^{-8}$) indicating that the SNP effect does not act entirely through the common liability. Based on Q_{SNP} , we selected only SNPs that did not show evidence of heterogeneity ($Q_{\text{SNP}} p \geq 5 \times 10^{-8}$) before conducting functional follow-up analyses of SNPs associated with the common liability.

Two complementary strategies were used to map SNPs to genes: (i) positional mapping of lead SNPs and (ii) expression quantitative trait loci (eQTL) mapping of lead SNPs. For positional mapping (i), g:Profiler^{43,44} and Phenoscanner⁴⁵ were used, mapping SNPs to genes based on being physically located inside a gene. SNPs identified in GWA studies often reside in non-coding regions of the genome (intergenic and intronic), but can nevertheless influence gene expression of nearby genes. We therefore applied (ii) eQTL mapping on the lead SNPs (*cis*-eQTL, variants in a +/- 1 megabase window around the transcription start site of a given gene) using Qtlizer⁴⁶. Qtlizer integrates a number of eQTL databases (e.g. GTEx Portal⁴⁷, Haploreg⁴⁸, GRASP⁴⁹, GEUVADIS⁵⁰, SCAN⁵¹, seeQTL⁵², Blood eQTL Browser⁵³, pGWAS⁵⁴, ExSNP⁵⁵ and BRAINEAC⁵⁶), to assign noncoding SNPs to their cognate genes.

Finally, to explore previously identified associations of lead SNPs with other phenotypes, we searched the PhenoScanner⁴⁵, a database of genotype-phenotype associations from existing GWA studies. For comparability, the same set of analyses was employed to characterise SNPs associated with the individual addiction phenotypes, including only lead SNPs that did not solely operate through the common liability ($Q_{\text{SNP}} p < 5 \times 10^{-8}$). That way, the results obtained from gene annotation and functional analyses are interpretable with respect to shared (i.e., effects operating through the common liability) and non-shared (i.e., effects specific to the individual addiction phenotypes) risks.

Pathway enrichment analysis of genetic variants associated with the common liability

To identify the most likely biological pathways underlying the common heritable liability to addiction, we used Data-driven Expression-Prioritized Integration for Complex Traits (DEPICT⁵⁷) and Pathway SCoring ALgorithm (PASCAL)⁵⁸. DEPICT was used to test for tissue/cell type enrichment of a set of LD-independent SNPs ($r^2 < 0.05$ within 500 kb) outside genome-wide significance ($p < 5 \times 10^{-5}$). PASCAL was used to test for enrichment of all SNPs, using three gene sets (BIOCARTA, KEGG, REACTOME) curated by the Molecular Signatures Database (MSigDB⁵⁹) and gene sets defined by DEPICT. Prior to running the analyses, the GWA on the common liability was filtered according to the Q_{SNP} statistic, retaining only SNPs operating through the common liability ($Q_{\text{SNP}} p \geq 5 \times 10^{-8}$). Results were corrected for multiple testing using false discovery rate (FDR) correction (controlled at 5%). Further details regarding the application of the two methods can be found in the Supplement (sMethods).

Genetic correlations between the common liability and other complex traits

Bivariate LD score regression analyses were performed in Genomic SEM, to estimate the genetic correlations between SNPs operating through the common liability (i.e., SNPs with $Q_{\text{SNP}} p \geq 5 \times 10^{-8}$) and 35 other traits related to physical features (e.g., height, body mass index), personality (e.g., risk-taking, neuroticism), social variables (e.g., socioeconomic status, education) and mental health (e.g., schizophrenia, depression). A complete list of the included GWA summary statistics can be found in **sTable 1** (Supplement). FDR correction (controlled at 5%) was used to adjust for multiple testing.

Evaluation of the causal claims implied by the common liability theory

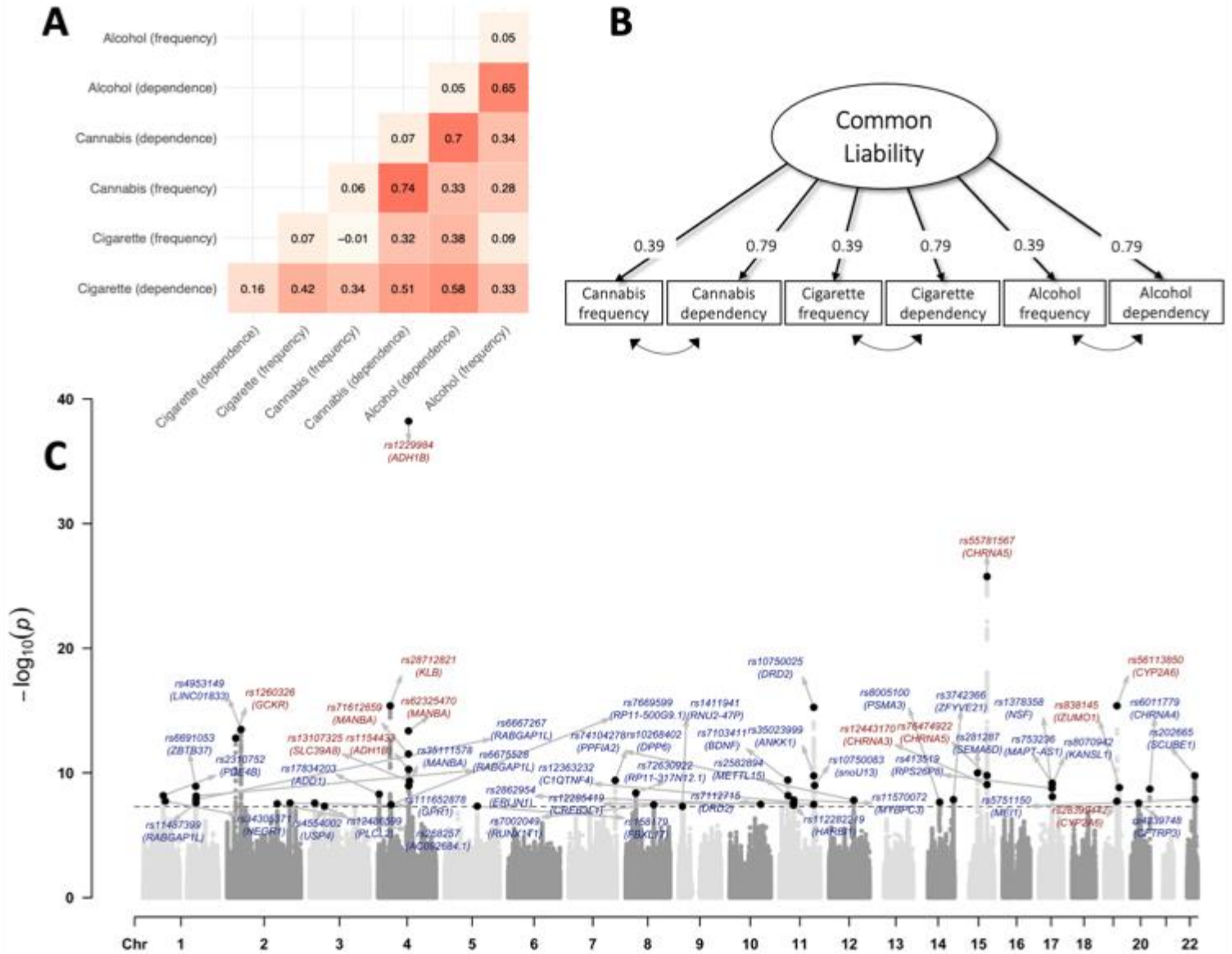
It is important to note that the common liability to addiction, constructed from the observed covariance structure of phenotypes reflecting addictive behaviours, constitutes only a statistical model. As such, the model's fit statistics do not examine the validity of the causal claims imposed by the model. For example, due to statistical equivalence, models with distinct interpretations (e.g. network models versus latent factor models) can have equivalent fit when constructed from the same data⁶⁰. Therefore, model fit reveals little about the underlying causal relationships between the observed and latent variables. To assess the factor model-implied causal pathways, where the common liability factor is thought to causally influence all its indicators, Mendelian Randomization (MR) analysis was used to assess the effects between the common liability and the individual substance use phenotypes. Inverse variance weighted (IVW) MR implemented in TwoSampleMR⁶¹ package was applied to all analyses. The genetic markers instrumenting the common liability were selected based on genome-wide significance ($p < 5 \times 10^{-8}$) and Q_{SNP} , retaining only SNPs that operated through the common liability ($Q_{\text{SNP}} p \geq 5 \times 10^{-8}$). To facilitate comparability of the MR estimates, the beta estimates for the included SNPs were standardized by dividing the z-scores by the square root of the sample size before conducting MR (cf. Supplement).

Results

Genomic model of the common heritable liability to addiction

The correlations among the individual cigarette, alcohol and cannabis use phenotypes are presented in the heatmap in **Figure 1a** (cf. Supplement, **sTable 2** for estimates). Genetic correlations varied widely between the individual substance phenotypes, ranging from $r_g = -0.01$ to $r_g = 0.74$ (mean $r_g = 0.40$; $SD = 0.21$). All standardized factor loadings are presented in **Figure 1b** (cf. also **sTable 3**, Supplement), showing that the constrained loadings were estimated to be 0.39 for frequency measures of substance use and 0.79 for substance dependence measures. On average, the common factor accounted for 38.81% (range 15.21%-62.41%) of the genetic variance in the six substance use phenotypes. The genomic common liability model showed evidence of a good model fit ($CFI = 0.97$, $SRMR = 0.07$, cf. **sTable 4**).

Figure 1. Multivariate genome-wide association study of the shared genetic architecture of cigarette, alcohol and cannabis use



Panel A. Heat map displaying the genetic correlations among the six substance use phenotypes. Shown are the genetic correlations between each of the cigarette, alcohol and cannabis phenotypes, with SNP-heritability estimates displayed down the diagonal. The mean genetic correlation is $r_g=0.4$ [sd=0.21, median=0.34 and range (-0.01-0.74)]. **Panel B.** Genomic structural equation model fitted on the genetic covariance matrices of the individual cigarette, alcohol and cannabis use phenotypes. Squares represent observed variables (the measured cigarette, alcohol and cannabis use phenotypes). The circle represents the latent variable, i.e., the common heritable liability to addiction, derived through factor analysis of the genetic correlations between the individual substance use phenotypes. Single-headed arrows are regression paths constrained to be equal across measures of frequency of use and dependence. **Panel C.** Manhattan plot of the SNP effects obtained from the multivariate genome-wide association analysis on the common liability. Labels are provided for the LD-independent genome-wide significant SNPs (i.e., SNPs above the horizontal line, with $p < 5 \times 10^{-8}$) and gene names obtained through positional mapping. The x-axis refers to chromosomal position, the y-axis refers to the p -value on a $-\log_{10}$ scale. Genetic variants coloured in red index variants that showed heterogeneous effects across the individual cigarette, alcohol and cannabis use phenotypes ($Q_{SNP} < 5 \times 10^{-8}$), indicating that their effects operate not entirely through the common liability. Genetic variants coloured in blue index genetic variants that did not show heterogeneous effects across the individual cigarette, alcohol and cannabis use phenotypes ($Q_{SNP} \geq 5 \times 10^{-8}$), indicating that their effects are likely to operate through the common liability.

Genetic variants associated with the common heritable liability

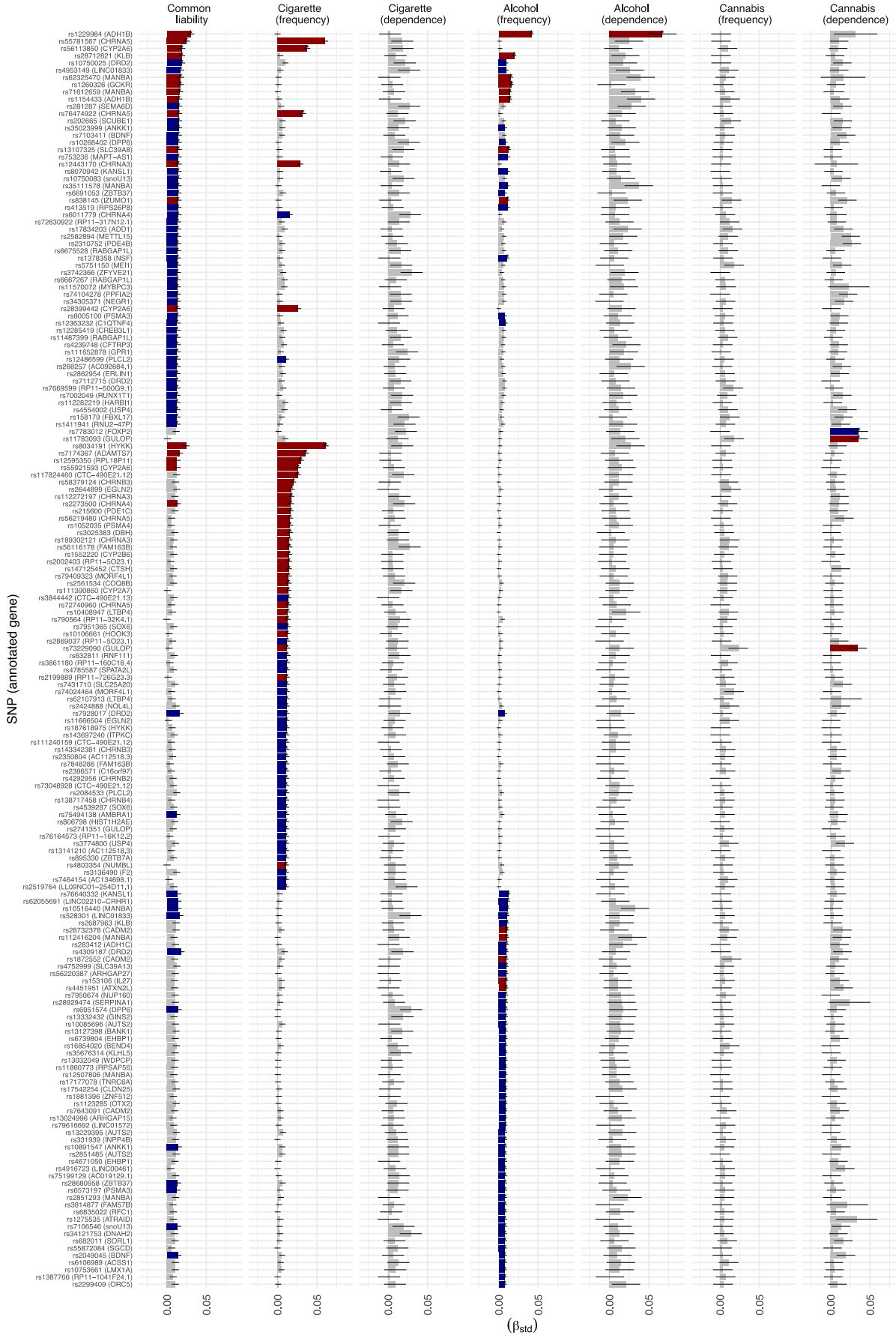
6,500,152 SNPs were included in the GWA of the common liability. Since our SNP estimates were derived from overlapping samples, we used the formula developed by Mallard et al.⁶² to derive the effective samples size (cf. sMethods, Supplement), which was estimated to be $N=187,062$. The Manhattan and Q-Q (quantile-quantile) plot of the common liability GWA are shown in **Figure 1c** and **sFigure 1** (Supplement), respectively. The main results for all GWA analyses, including the multivariate analysis on the common liability and the six univariate analyses on the individual substance use phenotypes are summarised in **sTable 5-7** and **sFigure 1** (Supplement). In brief, the GWA on the common liability identified 3,509 genome-wide ($p < 5 \times 10^{-8}$) SNPs, tagging 55 LD-independent SNPs. After removing SNPs showing significant heterogeneity ($Q_{\text{SNP}} p < 5 \times 10^{-8}$), 42 SNPs operating through the common liability remained (cf. SNPs highlighted in blue in the Manhattan plot, **Figure 1c**). Of the 42 SNPs, 28 (66.67%) were novel, i.e., have not been associated with any of the individual substance use phenotypes. Positional mapping showed that the top five SNPs (rs10750025, rs4953149, rs281287, rs202665, rs35023999) operating through the common liability lay mostly outside coding regions, located close to *DRD2*, *LINC01833*, *SEMA6D*, *SCUBE1* and *ANKK1*, respectively. Further inspection through eQTL mapping indicated that the aforementioned SNPs acted as eQTLs for positionally mapped genes, highlighting their putative role in the common liability via gene expression (cf. **sTable 8**, Supplement). A search in the PhenoScanner database⁴⁵ indicated that the five lead SNPs operating through the common liability have previously been linked to a number of behavioural phenotypes, such as neuroticism, irritability, smoking status or time spent in front of the computer (cf. **sTable 9**, Supplement). Of note, 13 of the 55 SNPs associated with the common liability still showed heterogeneous effects across the individual substance use phenotypes ($Q_{\text{SNP}} p < 5 \times 10^{-8}$, highlighted in red in the Manhattan plot, **Figure 1c**). Those SNPs can be considered as false discoveries, which may result from a single or a subset of SNPs with large effects on the individual substance use phenotypes³⁵. Among all 6,500,152 SNPs included in the common liability GWA, 2,356 (0.04%) showed heterogeneous effects.

For comparison, we also evaluated the GWA results of the individual substance use phenotypes, focusing on significant variants ($p < 5 \times 10^{-8}$) from the original GWA studies that showed heterogeneous effects ($Q_{\text{SNP}} p < 5 \times 10^{-8}$). Here, a number of variants appeared to be specific with respect to the class of substance. For alcohol use, the most prominent SNP was rs1229984, a variant on the alcohol dehydrogenase 1B gene (*ADH1B*). As shown in **Figure 2**, this variant was associated with all alcohol use phenotypes, but none of the cigarette or cannabis use phenotypes. Other identified SNPs related to the alcohol dehydrogenase group included rs1154433 and rs283412 (intron variants located on *ADH1B* and *ADH1C*, respectively) and rs1154433 acting as a *cis*-eQTL for *ADH1A* and *ADH4*. For cigarette use, a number of SNPs related to nicotinic pathways showed genome-wide effects, mapping onto nicotinic receptor genes (e.g., rs76474922 and rs58379124, intron variants located on *CHRNA5* and *CHRN3*, respectively) and variants affecting gene expression of *CHRNA2*, *CHRNA3*, *CHRNA4*, *CHRNA5*, *CHRN4* and *CHRN5* (cf. *cis*-eQTLs rs8034191, rs7174367, rs76474922, rs12595350, rs2273500, rs1052035, rs72740960, rs73229090). Only two SNPs were associated with cannabis use phenotypes, of which one variant (rs7783012, *FOXP2*) appeared to operate via the common liability ($Q_{\text{SNP}} p \geq 5 \times 10^{-8}$). This is supported by studies implicating rs7783012 in a number of phenotypes related to externalising

behaviours (e.g., ADHD, risk-taking and risky sexual behaviour⁴⁵), rather than pathways pointing to the endocannabinoid system specifically.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Figure 2. Associations of genetic variants with the common liability (blue) and the individual substance use phenotypes (red)



Plotted are the standardized beta coefficients (β_{std}) and their confidence intervals (cf. Supplement for details and corresponding formula) obtained from the multivariate genome-wide association (GWA) analysis on the common liability (column 1) and the univariate GWA analyses on the individual substance use phenotypes (columns 2-7). Displayed are genetic variants associated ($p < 5 \times 10^{-8}$) with at least one of the individual substance use phenotypes and/or the common liability. Bars coloured in grey index genetic variants that are not significantly associated ($p \geq 5 \times 10^{-8}$) with their respective phenotype. Bars coloured in red index genetic variants that showed heterogeneous effects across the individual cigarette, alcohol and cannabis use phenotypes ($Q_{SNP} p \geq 5 \times 10^{-8}$), indicating that their effects are unlikely to entirely operate through the common liability. The complete set of estimates can be found in **Table 7**. The asterisks (*) highlight genetic variants that were identified as LD-independent SNPs following clumping

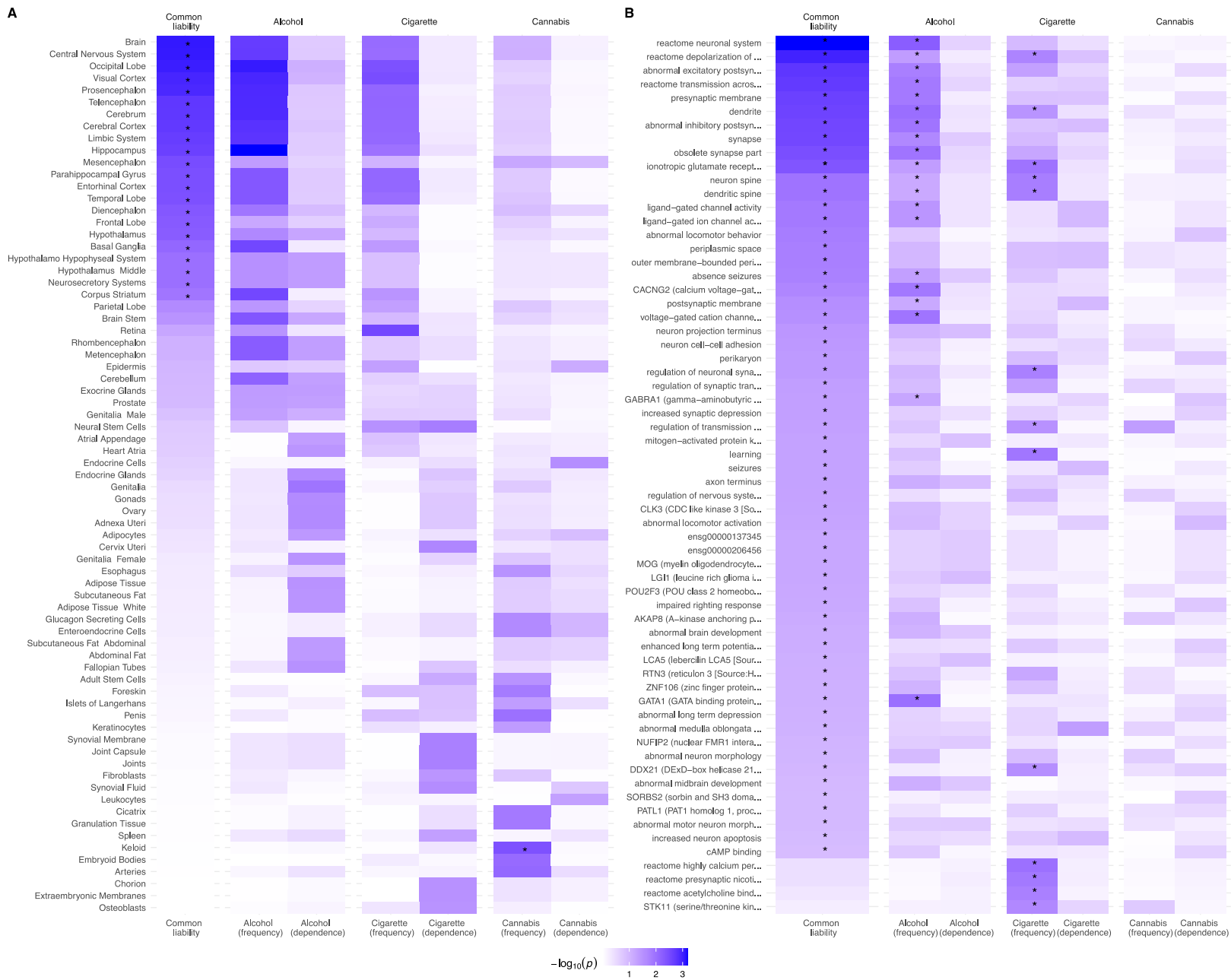
Pathway enrichment analyses of genes associated with the common heritable liability

Testing for tissue and cell type enrichment in DEPICT revealed 22 pathways associated (with FDR controlled at 5%) with the common liability (**Figure 3A**), which were all part of central nervous system tissues. A pattern of regional enrichment highlighted the role of a number of brain structures involved in the mesocorticolimbic brain circuits theorized to underlie addictive behaviours; notably, enrichment was present for regions involved in reward and emotion processing (limbic structures), motivation (basal ganglia), memory (hippocampus, parahippocampal gyrus, entorhinal cortex) and cognitive control (frontal lobe areas). In addition, the findings suggested that other brain structures may play a role in risk of addiction, such as areas involved in stress response (hypothalamus), or those relevant for visual processing (e.g., parahippocampal cortex, visual cortex, occipital lobe). Given this widespread network of brain areas, our results feed into theories suggesting that genetic risk to addiction is not solely the manifestation of altered limbic reward processes⁶³. Since the highlighted pathways were most prominently enriched for the common liability, and less so for individual substance use phenotypes, the aforementioned brain circuits may tap into somewhat distinct features characterizing the common liability.

Findings from PASCAL (**Figure 3B**) complement such conclusions, highlighting the role of neuronal signalling pathways in the common liability. Overall, 481 pathways were significantly (FDR controlled at 5%) enriched for the common liability, of which the top pathways related to broader categories of neurotransmitter functioning (e.g., neural system, abnormal excitatory postsynaptic currents, abnormal excitatory postsynaptic currents, transmission across chemical synapses). While the GWA analysis on the common liability identified the dopamine receptor D2 gene as the lead gene, pathway enrichment analyses implicated not only dopamine (e.g., dopamine neurotransmitter release cycle, dopamine binding), but also a number of other neurotransmitter systems, notably glutamate (e.g., ionotropic glutamate receptor complex, glutamate receptor activity), GABA (GABA-A receptor activity) or serotonin (serotonin neurotransmitter release cycle). Other identified pathways echoed the function of brain regions described above, such as those involved in learning and cognition (e.g., visual and associative learning, memory, cognition) or stress response (e.g., cAMP binding). Finally, a number of pathways were uniquely associated with the individual substance use phenotypes, mainly including those related to nicotinic pathways for cigarette use (e.g., presynaptic nicotinic acetylcholine receptors, postsynaptic nicotinic acetylcholine receptors, acetylcholine binding and downstream events). For alcohol, the ethanol oxidation pathway did not remain significant after correction for multiple testing, and there was no significant enrichment for cannabis use phenotypes.

All estimates obtained from DEPCIT and PASCAL are included in **sTable 11-12** (Supplement).

Figure 3. Pathway enrichment analyses of genes associated with the common heritable liability

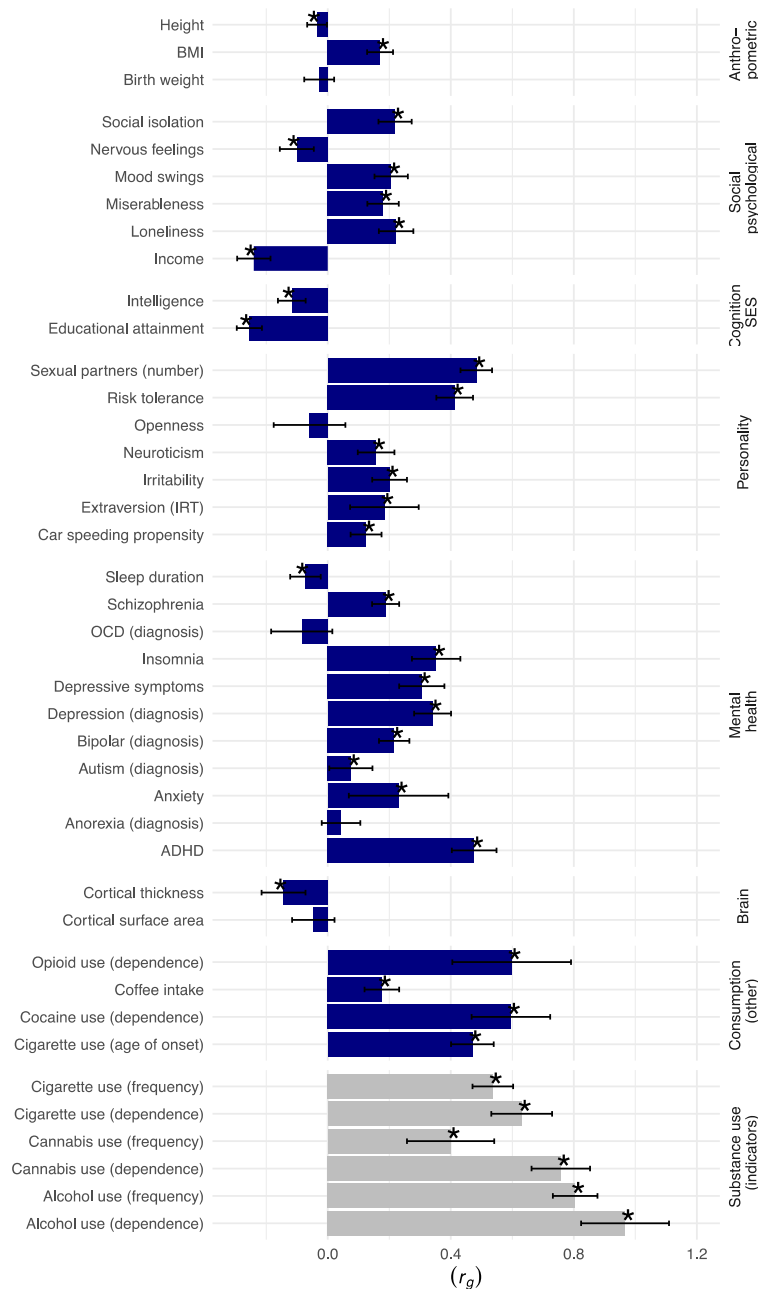


Shown are the results obtained from pathway enrichment analysis conducted in DEPICT and PASCAL. The common liability GWA results (filtered according to $Q_{SNP} p < 5 \times 10^{-8}$) and the individual substance use GWA summary statistics were used as the input. The violet shading indexes the significance level corresponding to each tested pathway. The asterisk marks pathways that remained significant after correction for multiple testing (False Discovery Rate controlled at 5%). **Panel A** highlights results obtained from the tissue/cell type enrichment analysis done in DEPICT. Displayed are the $-\log_{10}(p)$ -value for all pathways that were significant ($p < 0.05$) in at least one of the included GWA studies. **Panel B** depicts results obtained from pathway analysis done in PASCAL, using gene-sets curated by the Molecular Signatures Database ($n=1077$ sets) and DEPICT ($n=14462$ sets). For the common liability, $n=478$ pathways were significant after FDR correction for multiple testing. Displayed in the figure are the 15 most significant pathways per GWA study. The full set of results is listed in **sTable 10-11** (Supplement).

Genetic correlations between the common liability and other complex traits

Using the input from the Q_{SNP} -filtered GWA of the common liability and the GWA summary statistics for 41 traits (cf. **sTable 1** in Supplement for details), we found significant correlations with 36 complex traits after correction for multiple testing (**Figure 4**). As expected, the largest positive correlations were present between the common liability and its “constituents”, i.e., the cigarette, alcohol and cannabis use phenotypes used to derive the common liability (mean $r_g=0.68$). Among the other traits, the largest genetic correlations were present for cocaine and opioid dependence (both $r_g=0.60$), number of sexual partners ($r_g=0.49$), ADHD ($r_g=0.48$) and risk tolerance ($r_g=0.41$). Moderate genetic correlations were also present for a number of traits relating to mood, including insomnia ($r_g=0.35$) and depression ($r_g=0.34$). No significant (FDR controlled at 5%) associations were found with birth weight, openness, obsessive-compulsive disorder, anorexia and cortical surface area.

Figure 4. Genetic correlations between the common liability and other traits

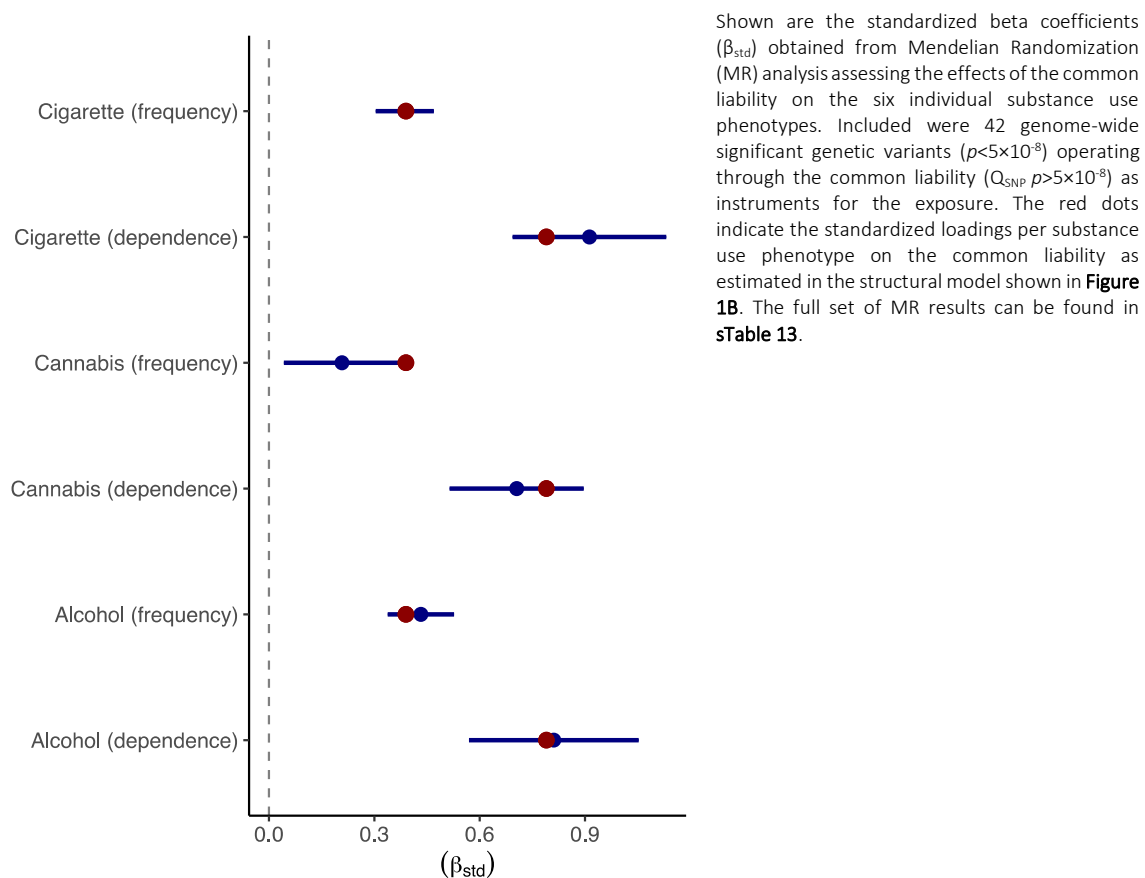


Shown are the genetic correlations (r_g) between the common liability GWA (filtered according to $Q_{SNP} p < 5 \times 10^{-8}$) and 41 other phenotypes, including 35 other traits (highlighted in blue) and the six individual substance use phenotypes used to derive the common liability (highlighted in grey). The asterisk indexes significant genetic correlations after correction for multiple testing (false discovery rate controlled at 5%, corrected for 41 tests). The full set of results is reported in **sTable 12** (Supplement).

Evaluation of the causal relationships implied by a common liability model

Figure 5 displays the results from Mendelian Randomization (MR) analyses, assessing paths running from the common liability to the individual substance use phenotypes. Using 42 Q_{SNP} -filtered LD-independent SNPs from the common liability GWA, the MR findings provide support for a causal interpretation of the initial descriptive common liability model (cf. **Figure 1b**) – that is, the common liability increases the risk of addiction to nicotine, alcohol and cannabis. More specifically, the loadings obtained from the genomic factor model of the common heritable were recovered using genetic markers instrumenting the common liability. As shown, the standardized causal effects obtained in MR were comparable to the factor loadings of the indicators (highlighted in red in **Figure 5**), as evident for measures of dependence [mean MR estimate: 0.79 (0.10 SD)] and measures of frequency of substance use [mean MR estimate: 0.33 (0.12 SD)]. The hypothesized structural model also asserts absence of causal effects between indicators belonging to a different class of substance (e.g., cannabis dependence \rightarrow alcohol dependence), which was in line with our MR results. Finally, reverse causation (effects of the specific substance use indicators on the common liability to addiction) was indicated for three of the indicators. Further discussion on the interpretation of reverse causation in this context is included in the Supplement, together with the full set of MR results (cf. **sTable 13**).

Figure 5. Mendelian Randomization analysis assessing causality between the common liability and the individual substance use phenotypes



Discussion

To dissect shared from non-shared genetic liability to addiction to nicotine, alcohol and cannabis, we conducted a multivariate genome-wide association (GWA) study of a common heritable liability to addiction. The top genetic variant operating through the common liability (rs10750025, located on the dopamine receptor D2 [*DRD2*] gene) provides support for the role dopamine in risk of addiction. Functional follow-up of common liability-associated genes further highlighted the role of widespread neuronal signalling pathways and neurotransmitter functioning beyond dopamine, such as GABAergic and glutamatergic pathways. Brain areas implicated in the common liability to addiction spanned limbic and cortical areas involved in reward, motivation, memory and cognitive control. The genetic overlap between the common liability and other complex traits was most prominent for other measures of addiction (e.g., cocaine and heroin use), as well as impulsive behaviours (e.g., risk-taking, ADHD) and mood (e.g., depression, insomnia). For cigarette and alcohol use, risk genes not operating via the common liability translated into specific pharmacogenomic pathways, such as nicotinic acetylcholine receptor functioning. Distinct pathways for cannabis use were, however, not identified.

Shared and non-shared genetic risks involved in risk of addiction to nicotine, alcohol and cannabis

In line with existing evidence, we found substantial genetic correlations between measures of cigarette, alcohol and cannabis use. This allowed us to model the common heritable liability to addiction, which explained substantial variance (average=39%) in genetic liabilities to individual addiction phenotypes. *DRD2* was identified as the lead gene operating via the common liability – a pathway believed to be a common mechanism by which addictive substances exert their acute pleasurable effects. *DRD2* in particular is a frequently studied gene implicated in addictive behaviours, given its central role in modulating the dopamine reward system that mediates the reinforcing effects of addictive substances. Indeed, *DRD2* has been identified in numerous genome-wide studies on cigarette^{7,8}, alcohol^{7,8,20} and cannabis use⁸. Other notable genes linked to the common liability included brain-derived neurotrophic factor (*BDNF*), corticotropin-releasing hormone receptor 1 (*CRHR1*), G protein-coupled receptor 1 (*GPCR1*), Ankyrin Repeat and Kinase Domain Containing 1 (*ANKK1*) or cAMP responsive element binding protein 3 like 1 (*CREB3L*). While these genes have already been scrutinized in studies using animal models and candidate gene approaches^{64,65} aiming to understanding the molecular basis of addiction, this is the first study confirming their involvement using a ‘hypothesis-free’ (i.e., genome-wide) approach.

Our results also highlight the role of neural signalling pathways involved in the common heritable liability to addiction, particularly synaptic functioning and a range of neurotransmitter systems beyond dopamine (GABA, glutamate, serotonin). Indeed, while dopaminergic mechanisms have been the traditional focus in addiction research, a growing body of research is now assessing the role of wider-ranging and interconnected neurotransmission systems in addiction vulnerability, involving GABAergic, glutamatergic and serotonergic projections that contribute to modulating reward reinforcement and drug-seeking behaviour^{66–68}. In line with this, enrichment analysis implicated the central nervous system and a network of brain areas in the common liability, including circuits involved in the processing of information related to reward (limbic structures), motivation (basal ganglia), memory (hippocampus) and cognitive control (frontal lobe areas). Since the discussed pathways were most prominently related to the common

liability, rather than the individual substance use liabilities, the identified pathways may reflect common neural substrates characterizing addiction vulnerability. Indeed, genetic variants associated with the individual cigarette and alcohol use phenotypes showed a different molecular footprint when compared to variants associated with the common liability. Substance-specific genetic risk, defined as genetic risk not operating via the common liability, was most notably expressed in pharmacological pathways relevant to nicotine and alcohol. For cigarette use, six of the 16 central nervous system-expressed nicotinic acetylcholine receptor genes were associated with at least one of the cigarette use phenotypes. For alcohol use, four genes belonging to the alcohol dehydrogenase family associated with at least one of the alcohol use phenotypes. The existence of these substance-specific genetic risks may help explain why some individuals become addicted to either nicotine or alcohol. For cannabis use, there was less clear evidence for distinct genetic risk, as only one cannabis-associated gene (*GULOP*) did not operate through the common liability.

Finally, it is assumed that the brain reward pathways partly link to addictive behaviours via some intermediate complex behaviours, such as risk-taking, sensation seeking or impulsivity. While this remains to be formally tested, this idea corroborates with our findings of genetic correlations between the common liability and maladaptive behaviours, including ADHD, risk-taking and cocaine and heroin dependence. We also found large genetic correlations between the common liability with internalising symptoms, including depression and insomnia.

Implications for the aetiology of addiction

Before jumping to aetiological conclusions regarding the shared and non-shared risks involved in addiction, it is important to assess if the core causal claims imposed by the common liability model hold true. Bi-directional MR was used to evaluate key assumptions, namely (1) the common liability to addiction has direct effects on all its indicators (i.e., the individual substance use phenotypes) and (2) there is no reverse causation. A third claim typically made by strict latent factor models asserts that there are no mutual direct effects between the individual indicators. In our study, this assumption was relaxed as we allowed for residual correlations between phenotypes belonging to the same class of substance (e.g., between frequency of cigarette use and cigarette dependence). Rather, our structural model posits that (3) all observed correlations between phenotypes belonging to a *different* class of substance are explained by the common liability. Overall, MR findings provided support for assumption (1), as causality ran from the common liability to all of the individual substance use phenotypes. While effects in the reverse direction were also present for three indicators – at odds with assumption (2) – this may reflect unaccounted pleiotropy (cf. Supplement for further discussion on this point). In line with assumption (3), there were no causal relationships between phenotypes belonging to a different class of substance. Together, since core assumptions are met (common liability → individual phenotypes), the results suggest that a common liability to addiction may usefully capture most relationships between substances. Embracing such conceptualization would have important implications for intervention. First, targeting modifiable features of the common liability should reduce risk of addiction to nicotine, alcohol and cannabis. For example, pharmacological treatments targeting dopamine, glutamate and GABA function may reduce craving and the euphoric/rewarding responses to cigarettes, alcohol and cannabis^{69–74}. Second, interventional targeting of only one specific class of substance (e.g., nicotine) unlikely leads to reductions in use

of another class (e.g., alcohol). This conclusion is somewhat inconsistent with previous evidence in rodents showing reductions in alcohol use following the administration of nicotinic treatments (e.g. varenicline^{75,76}), although evidence from RCTs in humans is mixed^{77,78} and efficacy may not translate into the long-term⁷⁹.

Future directions

An important goal for future work would be to extend this multivariate analysis to a broader array of substance classes (e.g., cocaine, opiate) and addiction phenotypes (e.g., tolerance, craving, withdrawal, relapse), once the required GWA data becomes available. Such efforts will allow exploration of more fine-grained structural models. Furthermore, multivariate approaches as employed here are not just important in terms of GWA discovery, but also essential to reducing biases; in our study, a substantial proportion of GWA-significant SNPs associated with the individual substance use phenotypes appeared to be mediated by the common heritable liability. As such, future GWA studies powerful enough to detect small genetic effects will likely tag an increasing number of SNPs with horizontal pleiotropic effects when examining addiction phenotypes (i.e., direct effects on several phenotypes). As such, modelling heritable latent factors as done in this study, and/or accounting for its contribution as recently proposed^{80,81} is therefore paramount when using genetically informed causal inference methods that are sensitive to the presence of heritable confounding.

Taken together, our results confirm that a common heritable liability partially explains the high co-occurrence of addiction to nicotine, alcohol, and cannabis. Functions of the implicated genes converged on broad central nervous system pathways beyond the dopaminergic pathways long-hypothesised in risk of addiction.

Data access

Summary statistics of the common liability GWA analysis will be available upon publication of this work. References to all publicly available summary statistic files included in this work are listed in **sTable 1**.

Code availability

The code used to conduct the analyses presented in this work is available on GitHub (https://github.com/TabeaSchoeler/TS2021_CommonLiabAddiction)

URLs

MsigDB data (<https://www.gsea-msigdb.org/gsea/index.jsp>)

PASCAL (<https://www2.unil.ch/cbg/index.php?title=Pascal>)

G:Profiler (<https://biit.cs.ut.ee/gprofiler/page/r/>)

PhenoScanner (<https://github.com/phenoscanner/phenoscanner>)

TwoSampleMR (<https://mrcieu.github.io/TwoSampleMR/>)

GenomicSEM (<https://github.com/MichelNivard/GenomicSEM>)

DEPICT (<https://data.broadinstitute.org/mpg/depict/>)

Neale Lab UKBB summary statistics (<http://www.nealelab.is/uk-biobank/>)

GWAS ATLAS (<https://atlas.ctglab.nl/>)

Acknowledgements

This study would not have been possible without the use of publicly available genome-wide summary data and software tools. The authors gratefully acknowledge these resources, and thank the research participants, the research teams and institutions that have contributed to this research. We are grateful to Dr. Judit Cabana-Domínguez and Prof. Bru Cormand for providing the results from the cocaine dependence genome-wide association study.

Funding

T.S. is funded by a Wellcome Trust Sir Henry Wellcome fellowship (grant 218641/Z/19/Z). J.R.B is funded by a Wellcome Trust Sir Henry Wellcome fellowship (grant 215917/Z/19/Z). A.M. is supported by the UCLH NIHR BRC. Z.K. is funded by the Swiss National Science Foundation (# 310030-189147). J.B.P. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 863981) and is supported by the Medical Research Foundation 2018 Emerging Leaders 1st Prize in Adolescent Mental Health (MRF-160-0002-ELP-PINGA). The funding organisations listed were not involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Competing interests

No competing interests to declare.

References

1. Degenhardt L, Charlson F, Ferrari A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry*. 2018;5(12):987-1012. doi:10.1016/S2215-0366(18)30337-7
2. Palmer RHC, Button TM, Rhee SH, et al. Genetic etiology of the common liability to drug dependence: Evidence of common and specific mechanisms for DSM-IV dependence symptoms. *Drug Alcohol Depend*. 2012;123:S24-S32. doi:10.1016/j.drugalcdep.2011.12.015
3. Connor JP, Gullo MJ, White A, Kelly AB. Polysubstance use - diagnostic challenges, patterns of use and health. *Curr Opin Psychiatry*. 2014;27(4):269-275. doi:10.1097/YCO.0000000000000069
4. Vanyukov MM, Tarter RE, Kirillova GP, et al. Common liability to addiction and “gateway hypothesis”: Theoretical, empirical and evolutionary perspective. *Drug Alcohol Depend*. 2012;123:3-17. doi:10.1016/j.drugalcdep.2011.12.018
5. Vanyukov MM, Tarter RE, Kirisci L, Kirillova GP, Maher BS, Clark DB. Liability to substance use disorders: 1. Common mechanisms and manifestations. *Neurosci Biobehav Rev*. 2003;27(6):507-515. doi:10.1016/j.neubiorev.2003.08.002
6. Nivard M, Verweij K, Minică C, et al. Connecting the dots, genome-wide association studies in substance use. *Mol Psychiatry*. 2016;21:733-735. doi:10.1038/mp.2016.14
7. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*.

- 2019;51(2):237-244. doi:10.1038/s41588-018-0307-5
8. Jang S-K, Saunders G, Liu M, Jiang Y, Liu DJ, Vrieze S. Genetic correlation, pleiotropy, and causal associations between substance use and psychiatric disorder. *Psychol Med*. Published online August 7, 2020:1-11. doi:10.1017/S003329172000272X
 9. Kendler KS, Schmitt E, Aggen SH, Prescott CA. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Arch Gen Psychiatry*. 2008;65(6):674-682.
 10. Ystrom E, Reichborn-Kjennerud T, Neale MC, Kendler KS. Genetic and environmental risk factors for illicit substance use and use disorders: joint analysis of self and co-twin ratings. *Behav Genet*. 2014;44(1):1-13.
 11. Swan GE, Carmelli D, Cardon LR. The consumption of tobacco, alcohol, and coffee in caucasian male twins: A multivariate genetic analysis. *J Subst Abuse*. 1996;8(1):19-31. doi:10.1016/S0899-3289(96)90055-3
 12. Agrawal A, Grant JD, Lynskey MT, et al. The genetic relationship between cannabis and tobacco cigarette use in European- and African-American female twins and siblings. *Drug Alcohol Depend*. 2016;163:165-171. doi:10.1016/j.drugalcdep.2016.04.011
 13. True WR, Xian H, Scherrer JF, et al. Common genetic vulnerability for nicotine and alcohol dependence in men. *Arch Gen Psychiatry*. 1999;56(7):655. doi:10.1001/archpsyc.56.7.655
 14. Walters RK, Polimanti R, Johnson EC, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018;21(12):1656-1669. doi:10.1038/s41593-018-0275-1
 15. Zhou H, Sealock JM, Sanchez-Roige S, et al. Genome-wide meta-analysis of problematic

- alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nat Neurosci*. 2020;23(7):809-818. doi:10.1038/s41593-020-0643-5
16. Quach BC, Bray MJ, Gaddis NC, et al. Expanding the genetic architecture of nicotine dependence and its shared genetics with multiple traits. *Nat Commun*. 2020;11(1):5562. doi:10.1038/s41467-020-19265-z
 17. Hatoum AS, Johnson EC, Polimanti R, et al. The Addiction Genetic Factor (a): A unitary genetic vulnerability characterizes substance use disorders and their associations with common correlates. *medRxiv*. Published online January 1, 2021:2021.01.26.21250498. doi:10.1101/2021.01.26.21250498
 18. Quach BC, Bray MJ, Gaddis NC, et al. Expanding the Genetic Architecture of Nicotine Dependence and its Shared Genetics with Multiple Traits: Findings from the Nicotine Dependence GenOmics (iNDiGO) Consortium. *bioRxiv*. Published online January 1, 2020:2020.01.15.898858. doi:10.1101/2020.01.15.898858
 19. Warnault V, Darcq E, Morisot N, et al. The BDNF Valine 68 to Methionine Polymorphism increases compulsive alcohol drinking in mice that is reversed by tropomyosin receptor kinase B activation. *Biol Psychiatry*. 2016;79(6):463-473. doi:10.1016/j.biopsych.2015.06.007
 20. Evangelou E, Gao H, Chu C, et al. New alcohol-related genes suggest shared genetic mechanisms with neuropsychiatric disorders. *Nat Hum Behav*. Published online 2019. doi:10.1038/s41562-019-0653-z
 21. Brazel DM, Jiang Y, Hughey JM, et al. Exome Chip meta-analysis fine maps causal variants and elucidates the genetic architecture of rare coding variants in smoking and alcohol use.

- Biol Psychiatry*. 2019;85(11):946-955. doi:10.1016/j.biopsych.2018.11.024
22. Henderson JA, Buchwald DS, Howard B V, et al. Genetics of smoking behavior in American Indians. *Cancer Epidemiol Biomarkers Prev*. Published online August 27, 2020:cebp.0026.2020. doi:10.1158/1055-9965.EPI-20-0026
 23. Johnson EC, Demontis D, Thorgeirsson TE, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. *The Lancet Psychiatry*. 2020;7(12):1032-1045. doi:10.1016/S2215-0366(20)30339-4
 24. Grotzinger AD, Rhemtulla M, de Vlaming R, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav*. 2019;3(5):513-525. doi:10.1038/s41562-019-0566-x
 25. Luningham JM, Poore HE, Yang J, Waldman ID. Testing structural models of psychopathology at the genomic level. *bioRxiv*. Published online 2018:502039. doi:10.1101/502039
 26. Lee PH, Anttila V, Won H, et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell*. 2019;179(7):1469-1482.e11. doi:10.1016/j.cell.2019.11.020
 27. Mallard TT, Linnér RK, Grotzinger AD, et al. Multivariate GWAS of psychiatric disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities. *bioRxiv*. Published online January 1, 2020:603134. doi:10.1101/603134
 28. Thorp JG, Campos AI, Grotzinger AD, et al. Symptom-level genetic modelling identifies novel risk loci and unravels the shared genetic architecture of anxiety and depression. *medRxiv*. Published online January 1, 2020:2020.04.08.20057653.

doi:10.1101/2020.04.08.20057653

29. Grotzinger AD, Mallard TT, Akingbuwa WA, et al. Genetic Architecture of 11 Major Psychiatric Disorders at Biobehavioral, Functional Genomic, and Molecular Genetic Levels of Analysis. *medRxiv*. Published online January 1, 2020:2020.09.22.20196089.
doi:10.1101/2020.09.22.20196089
30. Karlsson Linner R, Mallard TT, Barr PB, et al. Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. *bioRxiv*.
Published online January 1, 2020:2020.10.16.342501. doi:10.1101/2020.10.16.342501
31. Peyre H, Schoeler T, Liu C, et al. Combining multivariate genomic approaches to elucidate the comorbidity between autism spectrum disorder and attention deficit hyperactivity disorder. *J Child Psychol Psychiatry*. Published online July 7, 2021:jcpp.13479.
doi:10.1111/jcpp.13479
32. Davies G, De la Fuente, Javier, Grotzinger AD, Tucker-Drob EM, Deary IJ. Genetic “General Intelligence,” Objectively Determined and Measured. *bioRxiv*.
33. Demange PA, Malanchini M, Mallard TT, et al. Investigating the Genetic Architecture of Non-Cognitive Skills Using GWAS-by-Subtraction. *bioRxiv*. Published online January 1, 2020:2020.01.14.905794. doi:10.1101/2020.01.14.905794
34. Hatoum AS, Morrison CL, Mitchell EC, et al. GWAS of Over 427,000 Individuals Establishes GABAergic and Synaptic Molecular Pathways as Key for Cognitive Executive Functions. *bioRxiv*. Published online January 1, 2020:674515. doi:10.1101/674515
35. de la Fuente J, Davies G, Grotzinger AD, Tucker-Drob EM, Deary IJ. A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. *Nat Hum*

- Behav.* Published online September 7, 2020. doi:10.1038/s41562-020-00936-2
36. Iob E, Schoeler T, Cecil CM, Walton E, McQuillin A, Pingault J-B. Identifying risk factors involved in the common versus specific liabilities to substance use: A genetically informed approach. *Addict Biol.* Published online July 23, 2020:e12944. doi:10.1111/adb.12944
 37. Carey CE, Agrawal A, Bucholz KK, et al. Associations between Polygenic Risk for Psychiatric Disorders and Substance Involvement. *Front Genet.* 2016;7. doi:10.3389/fgene.2016.00149
 38. Fowler J, Volkow N, Kassed C, Chang L. Imaging the addicted human brain. *Sci Pract Perspect.* 2007;3(2):4-16. doi:10.1151/spp07324
 39. Bühler K-M, Giné E, Echeverry-Alzate V, Calleja-Conde J, de Fonseca FR, López-Moreno JA. Common single nucleotide variants underlying drug addiction: more than a decade of research. *Addict Biol.* 2015;20(5):845-871. doi:10.1111/adb.12204
 40. Mayfield J, Arends MA, Harris RA, Blednov YA. Genes and alcohol consumption. In: ; 2016:293-355. doi:10.1016/bs.irn.2016.02.014
 41. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet.* 2018;50(11):1593-1599. doi:10.1038/s41588-018-0248-z
 42. Bulik-Sullivan BK, Loh P-R, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015;47(3):291-295. doi:10.1038/ng.3211
 43. Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research.* 2020;9:709. doi:10.12688/f1000research.24956.1

44. Reimand J, Isserlin R, Voisin V, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc*. 2019;14(2):482-517. doi:10.1038/s41596-018-0103-9
45. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. Kelso J, ed. *Bioinformatics*. 2019;35(22):4851-4853. doi:10.1093/bioinformatics/btz469
46. Munz M, Wohlers I, Simon E, et al. Qtlizer: comprehensive QTL annotation of GWAS results. *bioRxiv*. Published online January 1, 2020:495903. doi:10.1101/495903
47. Ardlie KG, Deluca DS, Segre A V., et al. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science (80-)*. 2015;348(6235):648-660. doi:10.1126/science.1262110
48. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res*. 2016;44(D1):D877-D881. doi:10.1093/nar/gkv1340
49. Eicher JD, Landowski C, Stackhouse B, et al. GRASP v2.0: an update on the Genome-Wide Repository of Associations between SNPs and phenotypes. *Nucleic Acids Res*. 2015;43(D1):D799-D804. doi:10.1093/nar/gku1202
50. Lappalainen T, Sammeth M, Friedländer MR, et al. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature*. 2013;501(7468):506-511. doi:10.1038/nature12531
51. Gamazon ER, Zhang W, Konkashbaev A, et al. SCAN: SNP and copy number annotation. *Bioinformatics*. 2010;26(2):259-262. doi:10.1093/bioinformatics/btp644

52. Xia K, Shabalina AA, Huang S, et al. seeQTL: a searchable database for human eQTLs. *Bioinformatics*. 2012;28(3):451-452. doi:10.1093/bioinformatics/btr678
53. Westra H-J, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013;45(10):1238-1243. doi:10.1038/ng.2756
54. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun*. 2017;8(1):14357. doi:10.1038/ncomms14357
55. Yu C-H, Pal LR, Moulton J. Consensus genome-wide expression quantitative trait loci and their relationship with human complex trait disease. *Omi A J Integr Biol*. 2016;20(7):400-414. doi:10.1089/omi.2016.0063
56. Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci*. 2014;17(10):1418-1428. doi:10.1038/nn.3801
57. Pers TH, Karjalainen JM, Chan Y, et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun*. 2015;6(1):5890. doi:10.1038/ncomms6890
58. Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. Listgarten J, ed. *PLoS Comput Biol*. 2016;12(1):e1004714. doi:10.1371/journal.pcbi.1004714
59. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*.

- 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102
60. Fried EI. Lack of theory building and testing impedes progress in the factor and network literature. *psyarxiv*. Published online 2020. doi:<https://doi.org/10.31234/osf.io/zg84s>
61. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7. doi:10.7554/eLife.34408
62. Mallard TT, Linnér RK, Okbay A, et al. Not just one p: Multivariate GWAS of psychiatric disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities. *bioRxiv*. Published online January 1, 2019:603134. doi:10.1101/603134
63. Goldstein RZ, Volkow ND. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *Am J Psychiatry*. 2002;159(10):1642-1652. doi:10.1176/appi.ajp.159.10.1642
64. Gelernter J, Yu Y, Weiss R, et al. Haplotype spanning TTC12 and ANKK1, flanked by the DRD2 and NCAM1 loci, is strongly associated to nicotine dependence in two distinct American populations. *Hum Mol Genet*. 2006;15(24):3498-3507. doi:10.1093/hmg/ddl426
65. Park J, Sung J-Y, Kim D-K, Kong ID, Hughes TL, Kim N. Genetic association of human Corticotropin-Releasing Hormone Receptor 1 (CRHR1) with Internet gaming addiction in Korean male adolescents. *BMC Psychiatry*. 2018;18(1):396. doi:10.1186/s12888-018-1974-6
66. Al-Hasani R, Gowrishankar R, Schmitz GP, et al. Ventral tegmental area GABAergic inhibition of cholinergic interneurons in the ventral nucleus accumbens shell promotes reward reinforcement. *Nat Neurosci*. Published online August 12, 2021. doi:10.1038/s41593-021-00898-2

67. Fischer AG, Ullsperger M. An update on the role of serotonin and its interplay with dopamine for Reward. *Front Hum Neurosci*. 2017;11. doi:10.3389/fnhum.2017.00484
68. Kalivas PW. The glutamate homeostasis hypothesis of addiction. *Nat Rev Neurosci*. 2009;10(8):561-572. doi:10.1038/nrn2515
69. Liu Y, Sun H, Bao Y, et al. Subjective, cognitive/psychomotor, and physiological effects of aripiprazole in Chinese light and heavy smokers. *Drug Alcohol Depend*. 2009;101(1-2):42-52. doi:10.1016/j.drugalcdep.2008.10.024
70. Kranzler HR, Covault J, Pierucci-Lagha A, et al. Effects of aripiprazole on subjective and physiological responses to alcohol. *Alcohol Clin Exp Res*. 2008;32(4):573-579. doi:10.1111/j.1530-0277.2007.00608.x
71. Rassnick S, Pulvirenti L, Koob GF. Oral ethanol self-administration in rats is reduced by the administration of dopamine and glutamate receptor antagonists into the nucleus accumbens. *Psychopharmacology (Berl)*. 1992;109(1-2):92-98. doi:10.1007/BF02245485
72. Kenny PJ, Chartoff E, Roberto M, Carlezon WA, Markou A. NMDA receptors regulate nicotine-enhanced brain reward function and intravenous nicotine self-administration: Role of the Ventral Tegmental Area and Central Nucleus of the Amygdala. *Neuropsychopharmacology*. 2009;34(2):266-281. doi:10.1038/npp.2008.58
73. Gray KM, Watson NL, Carpenter MJ, LaRowe SD. N-Acetylcysteine (NAC) in young marijuana users: An Open-Label Pilot Study. *Am J Addict*. 2010;19(2):187-189. doi:10.1111/j.1521-0391.2009.00027.x
74. Haney M, Hart CL, Vosburg SK, et al. Effects of baclofen and mirtazapine on a laboratory model of marijuana withdrawal and relapse. *Psychopharmacology (Berl)*. 2010;211(2):233-

244. doi:10.1007/s00213-010-1888-6
75. Feduccia AA, Simms JA, Mill D, Yi HY, Bartlett SE. Varenicline decreases ethanol intake and increases dopamine release via neuronal nicotinic acetylcholine receptors in the nucleus accumbens. *Br J Pharmacol*. 2014;171(14):3420-3431. doi:10.1111/bph.12690
76. Wouda JA, Riga D, De Vries W, et al. Varenicline attenuates cue-induced relapse to alcohol, but not nicotine seeking, while reducing inhibitory response control. *Psychopharmacology (Berl)*. 2011;216(2):267-277. doi:10.1007/s00213-011-2213-8
77. Litten RZ, Ryan ML, Fertig JB, et al. A double-blind, placebo-controlled trial assessing the efficacy of varenicline tartrate for alcohol dependence. *J Addict Med*. 2013;7(4):277-286. doi:10.1097/ADM.0b013e31829623f4
78. O'Malley SS, Zweben A, Fucito LM, et al. Effect of varenicline combined with medical management on alcohol use disorder with comorbid cigarette smoking. *JAMA Psychiatry*. 2018;75(2):129. doi:10.1001/jamapsychiatry.2017.3544
79. Bold KW, Zweben A, Fucito LM, et al. Longitudinal findings from a Randomized Clinical Trial of Varenicline for alcohol use disorder with comorbid cigarette smoking. *Alcohol Clin Exp Res*. 2019;43(5):937-944. doi:10.1111/acer.13994
80. Morrison J, Knoblauch N, Marcus JH, Stephens M, He X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat Genet*. 2020;52(7):740-747. doi:10.1038/s41588-020-0631-4
81. Darrous L, Mounier N, Kutalik Z. Simultaneous estimation of bi-directional causal effects and heritable confounding from GWAS summary statistics. *medRxiv*. Published online January 2020:2020.01.27.20018929. doi:10.1101/2020.01.27.20018929