



Published in final edited form as:

Nat Ment Health. 2025 February ; 3(2): 253–265. doi:10.1038/s44220-024-00353-8.

Identification of risk variants and cross-disorder pleiotropy through multi-ancestry genome-wide analysis of alcohol use disorder

Romain Icick, M.D., Ph.D.^{1,2,*c}, Alexey Shadrin, Ph.D.^{1,c}, Børge Holen, M.D.¹, Naz Karadag, M.Sc.¹, Nadine Parker, Ph.D.¹, Kevin S. O’Connell, Ph.D.¹, Oleksandr Frei, Ph.D.^{1,3}, Shahram Bahrami, Ph.D.¹, Margrethe Collier Høegh, Ph.D.¹, Trine Vik Lagerberg, Ph.D.¹, Weiqiu Cheng, Ph.D.¹, Tyler M. Seibert, M.D., Ph.D.^{4,5,6}, Srdjan Djurovic, Ph.D.^{1,8,9}, Anders M. Dale, Ph.D.^{1,7,10,11,12}, Hang Zhou, Ph.D.¹³, Howard J. Edenberg, Ph.D.¹⁴, Joel Gelernter, M.D.¹³, Olav B. Smeland, M.D., Ph.D.¹, Guy Hindley, M.D.¹, Ole A. Andreassen, M.D., Ph.D.^{1,10}

¹NORMENT Centre, Institute of Clinical Medicine, University of Oslo and Division of Mental Health and Addiction, Oslo University Hospital, 0407 Oslo, Norway

²Université Paris Cité, INSERM U1144, F-75006, France

³Center for Bioinformatics, Department of Informatics, University of Oslo, PO box 1080, Blindern, 0316 Oslo, Norway

⁴Department of Radiation Medicine and Applied Sciences, University of California, San Diego, La Jolla, CA, USA

⁵Department of Radiology, University of California, San Diego, La Jolla, CA 92093, USA

⁶Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA

⁷Multimodal Imaging Laboratory, University of California San Diego, La Jolla, CA 92093, USA

⁸Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

⁹NORMENT Centre, Department of Clinical Science, University of Bergen, Bergen, Norway

* **Corresponding author:** romain.icick@aphp.fr. CSAPA « Espace Murger », Département de Psychiatrie et de Médecine Addictologique, Hôpital Fernand Widal, 200 rue du Faubourg Saint-Denis, 75010 Paris, France. Fax: +33140054213.

^cThese authors contributed equally.

Author Contributions Statement

RI, AS, GH and OAA were responsible for the study concept and design. RI, AS and GH contributed to the acquisition of summary data. AS was the main data analyst. BH, NK, OBS, SD, HE, HZ and OAA assisted with data analysis and interpretation of findings. RI, AS, NK and BH prepared figures. BH drafted the manuscript. OAA, OBS, WC, TVL, MCH, KOC, NK, NP, OF, SB, TS, AD and JG provided critical revision of the manuscript for important intellectual content. All authors critically reviewed content and approved final version for publication.

Competing Interests Statement

Dr. Dale is a founder of and holds equity in Cortechs.ai and serves on its scientific advisory board; he is a member of the scientific advisory boards of HealthLytx and the Mohn Medical Imaging and Visualization Center (Bergen, Norway); and he receives funding through a research agreement between General Electric Healthcare and UCSD. Prof. Andreassen has received speaking honoraria from Lundbeck and has served as a consultant for HealthLytx. The other authors report no financial relationships with commercial interests.

Code availability statement

The code for genetic correlation (<https://github.com/brielin/Popcorn>) and MiXeR (<https://github.com/precimed/mixer>) are publicly available.

All of these sources are referred to in the manuscript main text or supplementary methods.

¹⁰KG Jebsen Centre for Neurodevelopmental disorders, University of Oslo, Oslo, Norway

¹¹Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA

¹²Department of Neurosciences, University of California San Diego, La Jolla, CA 92093, USA

¹³Department of Psychiatry, Yale University, New Haven, CT06511, USA. Veterans Affairs Connecticut Healthcare System, West Haven, CT06516, USA

¹⁴Department of Biochemistry and Molecular Biology & Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA.

Abstract

Alcohol use disorder (AUD) is highly heritable and burdensome worldwide. Genome-wide association studies (GWASs) can provide new evidence regarding the aetiology of AUD. We report a multi-ancestry GWAS focusing on a narrow AUD phenotype, using novel statistical tools in a total sample of 1,041,450 individuals [102,079 cases; European, 75,583; African, 20,689 (mostly African-American); Hispanic American, 3,449; East Asian, 2,254; South Asian, 104; descent]. Cross-ancestry functional analyses were performed with European and African samples. Thirty-seven genome-wide significant loci (105 variants) were identified, of which seven were novel for AUD and six for other alcohol phenotypes. Loci were mapped to genes, which show altered expression in brain regions relevant for AUD (striatum, hypothalamus, and prefrontal cortex) and encode potential drug targets (GABAergic, dopaminergic and serotonergic neurons). African-specific analysis yielded a unique pattern of immune-related gene sets. Polygenic overlap and positive genetic correlations showed extensive shared genetic architecture between AUD and both mental and general medical phenotypes, suggesting they are not only complications of alcohol use but also share genetic liability with AUD. Leveraging a cross-ancestry approach allowed identification of novel genetic loci for AUD and underscores the value of multi-ancestry genetic studies. These findings advance our understanding of AUD risk and clinically-relevant comorbidities.

Introduction

Severe alcohol use disorder (AUD) is a chronic and devastating illness characterized by maladaptive patterns of alcohol use. AUD is common, reaching up to 14% lifetime prevalence in the U.S. (1) and up to 30% in Southern and Eastern African sub-Saharan countries (2). Current treatment options for AUD show modest and variable efficacy (3). Further, AUD leads to several complications, including impaired lipid metabolism, liver and cardiovascular dysfunction, psychiatric comorbidity and cognitive impairment, all contributing to the high mortality (4).

AUD is a complex, heritable disorder, with twin heritability estimated at ~50% (5). Genome-wide association studies (GWASs) of AUD have begun to elucidate its genetic underpinnings and characterize its polygenic architecture. Most available GWAS findings consistently suggest distinct patterns of genetic correlations for AUD with other mental phenotypes as compared to phenotypes of drinking patterns (6–10). Thus, one of the challenges for genetic studies of alcohol-related traits is genetic heterogeneity (7,9,11),

which makes it critical to map the genetic architecture of clinically-relevant phenotypes with clear definitions. Several previous studies have combined DSM-based AUD case definition with quantitative and/or screening measures from tools such as the Alcohol Use Disorders Identification Test (AUDIT), which shows an interesting two-factors genetic architecture (8,12,13), from which the factor “problems” can be used in combination AUD cases (9,11) to increase sample sizes and maximise power for genetic discovery. A recent study identified 110 risk loci in a GWAS of a combined phenotyped termed “problematic alcohol use (PAU)”, considered a relevant proxy for the genetic study of AUD due to their strong genetic correlation (~85%) (10). The use of such broad definitions tend to capture relatively heterogenous phenotypes. As a result, substantial gaps remain in the knowledge of the genetic underpinnings of AUD, and how they differ from those of broader alcohol-related phenotypes - even when they build on alcohol-related problems.

To fill this knowledge gap, more studies focusing on DSM- or ICD-based AUD phenotype should be performed, since it represents the most burdensome alcohol-related clinical phenotype for patients and caregivers. AUD is, in part, defined by the distress/burden induced by the disorder (14). Likewise, including samples from multiple ancestries by meta-analysis may improve GWAS ability to detect and infer causality and biological relevance of associated genetic variants (15) – including for AUD (6,7). To make progress regarding the genetic underpinnings of AUD, one has to leverage the latest methods for post-GWAS analyses to translate genome-wide significant loci into biological function. To date, the main genes associated with genome-wide significant (GWS) loci in AUD/PAU implicate alcohol metabolism [three alcohol dehydrogenase genes and an aldehyde dehydrogenase gene (16)], response to stress (corticotropin releasing hormone and fibroblast growth factor genes), opioid signalling (*OPRM1*) and metal transport (*SLC39A8*). These findings encourage further functional annotations for druggable targets [see, e.g., (17) and (10)] to prioritize subsequent preclinical and clinical research to discover new drugs and repurpose existing ones for AUD. *In silico* functional genomic tools that link loci to genes to expression patterns across specific tissues and cell types can improve the discovery of biological pathways involved in AUD with potential for clinical translation. Other methods could be leveraged to explore the genetic architecture of AUD overlapping with mental traits and disorders and with general medical conditions and risk factors. For instance, genome-wide genetic correlation (r_g) has been useful in revealing the relationship between AUD and other phenotypes in samples of European descent (hereafter termed “EUR samples”) [see, e.g. (6,8,12)], and in revealing how much these genetic correlation patterns differed between alcohol consumption and AUD or PAU (6,8,12). Such analyses in samples of African descent (hereafter termed “AFR samples”) have been made possible using recently-developed analytical methods such as POPCORN, but remain scarce [see, e.g. (10)]. However, r_g remains unable to capture scenarios of a mixture of positive and negative correlations. The MiXeR method is able to characterize overlapping genetic architectures beyond r_g in polygenic disorders (18).

We aimed to boost discovery of genetic loci associated with AUD, leveraging novel GWAS data across multiple ancestries and relying on diagnostic criteria for defining AUD cases. We applied novel analytical tools to the multi-ancestry and to ancestry-specific samples to better characterize (i) the genetic architecture of AUD, (ii) its genetic overlap with clinically-

relevant mental and general medical traits, disorders and risk factors - to disentangle genetic risk beyond the direct effect of alcohol consumption and (iii) to investigate molecular pathways of AUD including potential druggable targets.

Results

GWAS meta-analysis

The meta-analysis of AUD in the multi-ancestry sample identified 105 genome-wide significant (GWS) risk variants from 37 loci (Figure 1). Seven loci were novel for AUD and six for other alcohol-related phenotypes (*e.g.* alcohol consumption, lifetime alcohol use). SNP-based heritability was 0.075 ($se=0.004$, $p<1E-17$) for the EUR GWAS and 0.053 ($se=0.017$, $p=1.4E-3$) for the AFR GWAS. Eight loci were unique to the multi-ancestry sample, seven to the EUR sample. In the latter, there were 88 GWS risk variants from 35 loci. In the AFR sample we identified eight GWS risk variants from one locus represented by the lead variant rs1229987; this variant was in the putative regulatory region of *ADH1B* and in high LD with rs2066702, a known functional locus (7,8,16) (Figure 1). There was no evidence of population stratification according to the by METAL output : MA [$\lambda_{GC} = 1.102$, $\lambda_{1000} = 1.001$], EUR [$\lambda_{GC} = 1.142$, $\lambda_{1000} = 1.001$, LDSC_intercept_with_EUR_LD = 1.0251 (0.0095)], AFR [$\lambda_{GC} = 1.034$, $\lambda_{1000} = 1.001$, LDSC_intercept_with_AFR_LD = 1.0306 (0.0111)].

Supplementary Table 2 shows the location and functional significance of independent GWS variants for AUD for each locus, by ancestry. When genes were mapped using any correspondence with variant position, effect on gene expression or effect on chromatin conformation (see methods section), there were 1366 genes in the multi-ancestry sample, 482 in the EUR sample and 18 in the AFR sample (Supplementary Table 3). These were used by FUMA for further annotation (GENE2FUNCTION step). Supplementary Table 2 only shows genes mapped according to the variant position combined with both effects on gene expression and on chromatin conformation data, which were: 41 for the multi-ancestry sample, 40 for EUR sample and one for AFR sample. The six loci that were novel regarding all alcohol-related phenotypes, *i.e.*, any consumption measure and any problems measure (PAU, dependence, AUD) were mapped to the following genes: *ERI3*, *BARHL2*, *SRFBP1* or *LOX*, *RP11-756H20/ADH1A*, *CNTLN*. Twenty-eight locus boundaries overlapped across the multi-ancestry and the EUR samples. When considering the total number of 73 AUD loci, nine (12%) had significant between-study heterogeneity - a source of potentially low reliability (four in the multi-ancestry sample, four in the EUR sample, and one in the AFR sample; see Methods section).

Among the 27 lead SNPs common to both the EUR and AFR samples, nineteen had a concordant direction of effect in both samples. The sign test confirmed global concordance at $p=0.0261$. Three variants from the highly significant *ADH1B* locus were replicated in the AFR sample after Bonferroni correction. All these replicated variants had concordant directions (decreased AUD risk) of effect compared to the EUR sample. The eight variants with discordant effects in the AFR *vs.* the EUR sample all had p -values in the original GWAS >0.215 . Besides, none of them showed plausible functional impact.

Functional analyses: cells, tissues and gene sets

Among 217 cell types tested, ten were significantly associated with increased/decreased expression of genes to which the lead SNPs mapped in the multi-ancestry sample, five in the EUR and one in the AFR sample. (Figure 2). These cell types mainly included cortical GABAergic neurons, but the multi-ancestry sample also elicited serotonergic and dopaminergic neurons from the adult midbrain and excitatory and inhibitory neurons from the prefrontal cortex. The AFR sample also elicited dopaminergic hippocampus neurons. When conditioning associations across datasets and across similar cell types (step3 of the FUMA pipeline <https://fuma.ctglab.nl/tutorial#workflow>), all associations remained significant in the multi-ancestry sample, vs. two in the EUR (both developmental GABAergic neurons) and none in the AFR samples (see Supplementary Table 5). These findings were remarkably consistent with tissue enrichment analysis (Figure 3), where significantly enriched tissues were mostly brain-related (six in the multi-ancestry, three in EUR sample), including the cortex, hippocampus and nucleus accumbens. No tissue was significantly enriched in the AFR sample.

The EUR+AFR meta-analysis of the phenotype “drinks/week” identified 77 risk variants (40 genomic risk loci) and yielded similar tissue enrichment as AUD (Figure 3D). We mapped 185 genes to these variants, based on their position and either effect on gene expression or on chromatin interaction (credibly mapped genes, see Supplementary Table 5A for the complete gene list mapped by FUMA). Amongst these, 82 (44%) were not common to those mapped to SNP from the drinks/week analysis [MVP (6) +GSCAN Phase 1 (19), N =287]. Only one single cell type was associated at any step of the FUMA analysis (Bonferroni-corrected p -value =0.03 for mediolateral neuroblast type 5, complete list in Supplementary Table 5B).

Also, using MAGMA, there were four significant gene sets in the multi-ancestry analysis, three in EUR, and 14 in AFR (Supplementary Table 6). All samples were enriched for the alcohol dehydrogenase activity geneset. Compared to EUR, the multi-ancestry sample yielded additional ‘response to alkaloids’ and ‘maintenance_of_presynaptic_active_zone_structure’ gene sets. Interestingly, the AFR analysis elicited nine gene sets related to immunity and inflammation, two to cancer risk, and one to cell aging.

Investigating gene enrichment patterns for Drugbank associations (**Supplementary table 7 A-C**) across ancestries, we identified 180 unique DrugBank hits from 41 genes in the multi-ancestry sample, 174 from 41 genes for the EUR sample, and 22 from one gene for the AFR sample. All AFR sample hits were also found in the other samples, and none of the AFR sample hits were related to the nervous system [most were related to alcohol dehydrogenase (*ADH*) gene]. There were 68 hits related to the nervous system ATC category in both the multi-ancestry and the EUR samples, 14% of which were related to interactions between antipsychotics and ADH genes. The remaining were mostly targets of tricyclic and tetracyclic antidepressants. The six different DrugBank hits between the multi-ancestry and EUR samples were related to three genes, *CDK5R1*, *POR* and *MDH2*, and yielded citric acid as the only unique medication with gene targets specifically in the

multi-ancestry sample. We also identified tiapride, a drug often used against agitation in alcohol withdrawal syndrome, and memantine which has been tested and used off-label against AUD (20). Drugbank hits were similar across EUR and AFR ancestries, except for Blood and blood forming organs (Fisher exact test $p=0.037$) (**Supplementary Table 7D**).

Genetic overlap (MiXeR)

Univariate MiXeR showed that AUD was moderately polygenic (7.8–7.9k ‘causal’ variants), and discoverability was 0.0025 (SD=0.0002). Bivariate MiXeR (Figure 4). showed large polygenic overlap between AUD and other substance use phenotypes. Thus, AUD shared 62% of its ‘causal’ variants with CUD, 76% with drinks/week, 95% with OUD and 74% with age at smoking initiation. Overlap was substantial, albeit smaller, between AUD and non-substance related mental disorders. AUD shared more than half its loci with ADHD (52%), bipolar disorder (53%), major depression (60%) and schizophrenia (62%), and mental traits, with a particularly large genetic overlap with neuroticism (96%). There was also a large overlap between AUD and systolic blood pressure (37%).

Genetic correlation (rg)

The genetic correlation of AUD was moderate between EUR and AFR ancestries ($rg=0.65$, $FDR=9.3 \times 10^{-7}$) (Figure 5). Genetic correlation patterns were overall similar in significance and magnitude across mental traits and disorders. Significant, positive correlations ranged from 0.16 (liver age) to 0.85 (opioid use disorder). Significant, negative correlations ranged from -0.4 (education level) to -0.23 (cognitive performance). The strongest correlations were found for ADHD (EUR/AFR $rg=0.47/0.3$), age at smoking initiation (EUR/AFR $rg=0.54/0.37$) and OUD (EUR/AFR $rg=0.85/0.81$). The rg between AUD and CUD, major depression, schizophrenia, bipolar disorder and PTSD were only significant in the EUR sample (highest $p_{corrected}=0.0054$ for PTSD). Consistent patterns were also found for cognitive traits (EUR/AFR $rg=-0.23/-0.4$) and educational attainment (EUR/AFR $rg=-0.32/-0.40$). Significant genetic correlation was also observed between AUD and heart failure ($rg=-0.22$, $p_{corrected}=7 \times 10^{-6}$), liver age ($rg=0.16$, $p_{corrected}=0.011$) and abdominal age ($rg=-0.17$, $p_{corrected}=5.4 \times 10^{-3}$) in the EUR samples. The complete rg results are presented in **Supplementary Table 8**.

Discussion

This multi-ancestry meta-analysis identified 105 genome-wide significant AUD risk variants from 37 independent genomic loci, including seven novel loci for AUD and six novel loci for other alcohol-related traits. Compared to the EUR sample, multi-ancestry meta-analysis resulted in a slight increase in loci discoverability and a stronger increase in biological diversity. AUD loci implicated genes enriched in key brain regions and cells. We confirmed the strong association between *ADH1B* SNPs and AUD in both EUR and AFR samples. We also confirmed extensive shared genetic liability between AUD and other substance use phenotypes, mental traits and disorders (especially ADHD and neuroticism) and other medical conditions, consistent with and extending previous work with AUD and PAU phenotypes (6–9).

To the best of our knowledge, the current study represents the largest GWAS meta-analysis of DSM/ICD-defined AUD, with 1,041,450 participants; identifying ~50% more loci than recently published AUD GWASs (6,9). A multi-ancestry meta-analysis of PAU (AUD + AUDIT-problem subscale; N=1,079,947) identified a total of 110 genetic risk variants across diverse ancestries (10). Comparing multi-ancestry samples findings from both studies (105 in ours, 100 in Zhou et al.'s), we found that 99 (94%) variants overlapped. Overall, our study identified almost as many risk variants as with the PAU phenotype, yet focusing on a narrow AUD phenotype. Although this suggests that six variants may be relatively specific to AUD *vs.* PAU, the only two different mapped genes were pseudogenes. As with the current study, Zhou et al. (10) also provided evidence for the role of the dopamine receptor type 2 gene (*DRD2*), in line with previous studies (6,21). We note that our estimation of the final men/woman sex-ratio in the meta-analysis sample was =2. The age range was comparable to population-based surveys of AUD. Thus, the age and sex ratios in our study are consistent with large-scale, nationally-representative samples of AUD (1). We did not perform sex-specific analysis due to the lack of individual-level data available for the vast majority of the final study sample.

Several genes mapped to the novel AUD loci had previously been associated with other alcohol-related phenotypes (*e.g.*, drinking quantity or frequency, lifetime alcohol use). This suggests some molecular mechanisms are shared between such alcohol-related phenotypes and AUD. It is noticeable that about half of the AUD loci were related to regulatory regions and not to altered protein structure or function. This supports the role of regulatory mechanisms in the pathophysiology of AUD, as in other complex phenotypes (22). The single AFR GWS locus was mapped to an intron of *ADH1B*, likely reflecting the strong LD with the functional *ADH1B* variant rs2066702 (0.78 – 0.98 across 1000 genomes AFR populations). When looking at FUMA cell type enrichment from the drinks/week EUR+AFR meta-analysis, only one significant association emerged, in contrast with our findings using the AUD phenotype - despite similar brain tissue enrichment. This highlights the additional insights gained by focusing on a more strictly defined, clinically relevant phenotype, potentially implicating brain circuits in AUD as opposed to the broader phenotype of alcohol consumption. Interestingly, 44% of credibly mapped genes mapped to AUD loci did not overlap with those mapped to drinks/week loci. We believe this further reflects both fundamental differences and similarities in the genetic vulnerability to AUD *vs.* alcohol consumption, which may partly underlie their distinct phenotypic expression.

A key finding in the current study was that ancestral diversity improved functional annotation of GWS risk variants, and allowed the discovery of cell types, tissue types and gene sets with potential relevance to the neurobiology of AUD. This was more than expected from the modest increase in the number of GWS loci in the multi-ancestry *vs.* EUR samples. Since most of the cell types enriched in the multi-ancestry meta-analysis correspond to the top signals of the EUR sample (significant or not), we believe the multi-ancestry meta-analysis showed an actual power increase. First, functional analyses doubled the number and diversity of significantly enriched cell types in multi-ancestry *vs.* EUR samples. These cell types included GABAergic, serotonergic (*Sert⁺*) and dopaminergic (labelled DA1 in EUR) neurons from the midbrain, and prefrontal neurons. This is consistent with previous evidence supporting the involvement of this brain region and

these cell types in AUD, particularly GABAergic neurons in the midbrain in a previous GWAS of maximum alcohol consumption (23) and - more generally - in emotion processing (24) and the action of benzodiazepine medication against alcohol withdrawal syndrome. Second, eight tissues were associated with GWS AUD loci in multi-ancestry meta-analysis, versus 5 in the EUR sample, and mostly included striatal brain regions. Interestingly, the multi-ancestry meta-analysis revealed pathways of relevance to the neuroimaging findings in AUD (25) and to the cell type analysis, including the substantia nigra, frontal cortex and nucleus accumbens. The association with the hypothalamus, a region that regulates liquid intake, could be related to the consumption component of AUD, in line with our previous findings associating hypothalamus with both alcohol consumption and AUD loci (21) and with experimental data regarding altered hypothalamic-pituitary-adrenal axis after chronic alcohol exposure (26). The current findings further illustrate the complementarity of the tissue-level and cell-level bioinformatic approaches to open therapeutic avenues in AUD (27,28). The FUMA cell_type enrichment showed no significant findings for the drinks/week phenotype, in contrast with our findings using the AUD phenotype - despite similar brain tissue enrichment. This highlights the additional insights gained by focusing on a more strictly defined, clinically relevant phenotype. In these analyses, as in state-of-the-art GWAS [e.g. (29)], we applied data from large, coordinated biological experiments implemented in relevant repositories online. Thus, our approach is a cost-effective way to summarize the effects of SNPs on gene expression / cell type representation / drug targets. However, focused experimental biological work is needed to confirm the mechanistic understanding, as in, e.g., (30).

The significant findings implicated alcohol dehydrogenase genes in all samples, but the AFR sample showed additional enrichment in several immunity/inflammation-related pathways (14 sets vs. four in multi-ancestry and two in EUR). A recent meta-analysis reported associations between cytokine levels and AUD (31). Although such a difference in the number of gene sets in AFR compared to EUR samples may represent a degree of noise in the analysis, MAGMA has shown high detection power with little type I error inflation (32). Several DrugBank hits were relevant to AUD and/or comorbid conditions showing significant genetic correlations with AUD (Figure 5). Most of these associations involved drugs routinely used in schizophrenia and bipolar disorders (e.g., first- and second-generation antipsychotics clozapine, aripiprazole, risperidone), in major depressive disorders (tricyclic and tetracyclic antidepressants, e.g., imipramine, mirtazapine), and lamotrigine used to prevent bipolar depression. Interestingly, this analysis also involved dopamine and inamrinone, which are used to treat congestive heart failure, lomitapide, a novel lipid-lowering agent (in relationship with the genetic correlation between AUD and abdominal age) and glutathione, a modulator of the redox function in the liver (thus potentially related to the genetic correlation between AUD and liver age). These findings add indirect support to the extensive shared genetic vulnerability between AUD and other major mental health conditions we identified using multiple methods. Interestingly, the targets of antipsychotics were related to two main pathways. One involved the dopamine receptor type 2 (*DRD2*) and its co-effectors (e.g., *ANKK1*, *CDK5R1*), the other involving genes from the aldehyde dehydrogenase family. The first could relate to conditions showing high comorbidity rates and strong genetic correlation with AUD. The second suggests that there could be

pharmacokinetic interactions between AUD (*i.e.*, heavy alcohol use, which is common in AUD) and the use of antipsychotics. This finding is relevant, since most studies to date have assessed these interactions at the level of CYP450 function (33). Finally, memantine, identified as interacting with the dopaminergic pathway (*DRD2* gene), has previously been investigated in AUD and cocaine use disorder trials (20). There were no ancestry-specific features in this analysis, with little enrichment in multi-ancestry analyses compared with EUR samples. This suggests that statistical power was the most important determinant of our findings regarding the potential druggability of proteins linked to mapped genes. More efforts will thus be required to before implementing treatment alternatives for comorbid AUD and other mental disorders based on GWAS findings in non-EUR samples.

The current sample size was large enough to apply MiXeR to a wider range of traits and disorders than in our previous work on AUD (21). We evidenced that substance use phenotypes are highly polygenic, and estimated the polygenicity of AUD to 7.2k - 8.5k causal variants - roughly similar to what we recently showed for other mental disorders (5.6k for ADHD to 14.5k for major depression) (18). The extensive genetic overlap between AUD and other substance use and mental disorders (notably ADHD, bipolar disorder, major depression and schizophrenia; including global genetic correlation, Figure 5, and polygenic overlap, Figure 4) could partly explain the high comorbidities. MiXeR also revealed a particularly large polygenic overlap between neuroticism and AUD (92%), similar to estimates found between neuroticism and schizophrenia (98%) or neuroticism and bipolar disorder (99%) (18). This suggests that most risk variants associated with neuroticism are also associated with these psychiatric disorders. However, this needs to be interpreted in light of global genetic correlations (r_g) between phenotypes and absolute numbers of risk variants (*i.e.*, polygenicity) for each phenotype. First, r_g between AUD and neuroticism is significant, yet, relatively weak (0.31); revealing a mixture of effect directions for the shared variants. Second, neuroticism is ~58% more polygenic than AUD (12.2k vs. 7.7k risk variants, respectively). This means that, while a large proportion of AUD variants are shared with neuroticism, a large number of neuroticism variants are unique to neuroticism. In comparison, MDD has an equivalent polygenicity (13.9K risk variants) to that of neuroticism. Therefore, while previous analyses from our group (18) have shown that a lower proportion of MDD variants are overlapping with neuroticism than AUD, this comprises ~9.8k risk variants - thus 31% more than the shared genetic component of AUD and neuroticism (7.5k) - and includes variants with mostly convergent direction of effects ($r_g = 0.68$). Taken together, our findings support the hypothesis that the shared genetic component of mental disorders (18,21) also includes AUD and partly relies on a global shared liability to neuroticism, although with a complex pattern of effect distributions. To date we have no evidence as to whether the shared variants between neuroticism and AUD are the same as those shared between neuroticism and other phenotypes. As regards general medical conditions, more GWAS data, especially from non-EUR samples, are needed to reliably estimate polygenic overlap with AUD. Still, we report substantial polygenic overlap between AUD and systolic blood pressure (36%) in the absence of significant genetic correlation.

There were mostly consistent patterns of genetic correlations across ancestries, especially regarding AUD and mental traits and disorders. Amongst them, OUD showed the highest

correlation (0.74) - significantly extending recent work in the AFR population (10). We compare these findings to epidemiological association between AUD and tobacco smoking/nicotine dependence as compared to OUD. In the general population survey NESARC-I (DSM-IV criteria), 46% of the respondents, who fulfilled criteria for past-year AUD also fulfilled the criteria for nicotine dependence (34). This is much higher than the 2.4% prevalence estimated for OUD diagnoses identified in AUD participants (35). This large difference in absolute prevalence, however, is not reflected by adjusted odds ratios in the same sample. While the adjusted risk for nicotine dependence in AUD respondents was 2.7 times higher than for non-AUD respondents (34), this odds ratio raises to 7.7 for OUD (34). Since these odds ratios are adjusted for shared clinical and socio-demographic risk factors, the findings support the hypotheses that there is shared genetic vulnerability between AUD and OUD. Nicotine dependence is more widespread in the general population than either AUD or OUD, and could thus be associated with more diverse risk factors, especially environmental and social risk. Further support for shared genetic vulnerability between AUD and OUD lies in the use of medications approved worldwide for AUD patients (and for OUD patients in countries where opioid agonists remain unavailable), which target the brain opioidergic system by weakly antagonizing OPRM1 receptors (36). There is extensive evidence showing the importance of opioid signalling in the vulnerability to and expression of alcohol self-administration in rodents (37,38) and AUD in humans (39–41). Our findings are also consistent with the hypothesis of a putative addiction (genetic) risk factor, which has emerged through the study of multiple phenotypes using multivariate analysis (10,42,43). Using bivariate MiXeR, we characterized these patterns of shared genetic vulnerability beyond genetic correlation. We showed that the high genetic correlation between AUD and OUD was linked to ~7.4 thousand genetic variants with mostly concordant directions shared between the disorders, while ~6.9 thousand variants remain unique to OUD.. Also, of particular interest were the correlations between AUD and conditions that are usually attributed to the toxic effects of alcohol (*e.g.*, MRI-predicted abdominal age and heart failure). This suggests a more mixed picture of the toxic effects of alcohol and shared molecular underpinnings. We plan to compare these correlation patterns by using individual-level genetic and phenotypic data regarding alcohol consumption *vs.* AUD in the near future.

Our study has limitations. There was insufficient statistical power in several cross-disorder analyses involving AUD, calling for urgent action to gather more genetic data in non-EUR populations. The effect direction across EUR and AFR ancestries was concordant for 19/27 lead SNPs and was confirmed by the sign test. To go further, we compared the proportion of variants with concordant *vs.* discordant directions of effect between our study (19 concordant, eight discordant, 67%) and the most recent PAU GWAS [64 concordant, 12 discordant; 84% (10)]. Fisher's exact test *p* was not significant (*p* =0.092) between the studies - suggesting similar proportions of variants with concordant *vs.* discordant directions of effect. Additionally, the multi-ancestry sample analyses had strong statistical power and few loci showed significant between-study heterogeneity. As in most standard modern GWASs, gene enrichment at the cellular and tissue level in our study was estimated *in silico*. Although this has the advantage of representing the combination of lead SNPs that show individually small effect sizes, it cannot represent direct effect of SNPs toward the facilitation or the resistance to the effects of a given drug. Some *in silico* data used

for functional annotation are sourced from EUR samples only. Additionally, even when the reference SNP and LD maps from the 1000 genomes project are sourced from relatively diverse non-EUR samples, AUD cases of AFR ancestry were recruited in Western countries (US, notably) and no data was available in terms of genetic heterogeneity for these particular samples. Same applies to *FUMA* analyses, for which the detail regarding LD structure is higher for EUR than for other ancestries. Overall, caution is thus advised when interpreting findings for functional or cross-ancestry analyses. Finally, the almost exclusive use of summary statistics, preventing us from performing analyses that could help to identify clinical subgroups and control for potential mediating factors. However, this remains the only way to leverage very large GWAS samples to date, especially for case-control analyses.

The current study leveraged multi-ancestry samples to discover several novel AUD risk loci and improve the biological diversity of associated molecular pathways, cell types and brain regions implicated in AUD.

Methods

This was a multi-ancestry and within-ancestry meta-analysis on AUD based on cross-sectional, case-control genome-wide association studies. The meta-analysis thus represents both a *de novo* analysis and a replication attempt of previous findings obtained with alcohol-related phenotypes.

GWAS samples

Alcohol Use Disorder.—We extracted summary statistics with *p*-values and *Z*-scores from recent GWASs (Table 2) relevant to the AUD phenotype defined according to the DSM-5 or the International Classification of Diseases (ICD) 9/10 (abuse and/or dependence). We included AUD GWASs showing global Linkage Disequilibrium Score Regression (LDSC) genetic correlation (*rg*) >0.8 with each other (44) (Supplementary Figure 1). Based on these criteria, we selected the following GWAS results for inclusion:

- Million Veteran Program (MVP): ICD 9/10 alcohol abuse/dependence (AUD) - and severe acute intoxication (either one inpatient or two outpatient diagnoses, ICD9 codes 303 to 303.03) (9,45). All diagnoses were based on validated electronic health records from clinical encounters at settings affiliated with the U.S. Veterans Affairs system. Participants were genotyped after specific, written informed consent. We analyzed MVP AUD data downloaded from the dbGaP website (accession phs001672.v9.p1). Although we did not plan to include alcohol-related traits other than AUD, the summary statistics from the MVP GWAS were not available without intoxication cases, which only represented 0.3% of the sample (N =226);
- FINNGEN [https://www.finngen.fi/en/access_results, R6 public release (46)]: ICD-9/10 abuse/dependence (AUD) based on validated electronic health records from in- or outpatient care settings in Finland;

- UK Biobank (UKB): ICD 10 abuse/dependence (AUD, see above) based on validated electronic health records from inpatient and primary care settings in United Kingdom (individual level genotypes under accession number 27412);
- Psychiatric Genomics Consortium (PGC) [<https://pgc.unc.edu/for-researchers/download-results/> (7)]: DSM-IV alcohol dependence (considered equivalent to severe DSM-5 AUD) diagnosed by trained clinicians' ratings or semi-structured interviews. These data are not publicly available without the meta-analysis with the FinnGen sample, and require request to the PGC workgroup.

Combining the samples from all ancestries yielded a total multi-ancestry sample of 1,041,450 individuals, including 102,079 AUD cases (Effective sample size - $N_{eff} = 321,343$ – see Supplementary Methods). Ancestry-specific sample sizes enabled ancestry-specific analyses for EUR and AFR samples only, as supported by visual examination of the QQ plots (Supplementary Figure 2A). In all the original samples, participants were genotyped after specific, written informed consent, which also included consent for further analyses and conservation. According to the Norwegian ethics, no review board is required to analyse GWAS summary statistics, which do not contain individual data from the initial study participants.

GWAS samples for comorbid disorders and traits.—To investigate the genetic architecture of AUD overlapping with mental traits and disorders and with general medical conditions and risk factors that represent frequent comorbidities and complications of alcohol use and AUD (47), we performed cross-disorder analyses using GWAS data downloaded between June 15 and July 1st 2022 (Supplementary Table 1 & Supplementary Methods). For each phenotype, we selected the largest sample size GWAS with data available to our group (*i.e.* publicly available or already acquired).

Statistical analysis

The AUD GWASs included in the meta-analysis were adjusted for sex and the first six to 10 principal components of ancestry (depending on the original GWAS). We applied the same procedure to perform our own AUD GWAS in the UK Biobank sample using Regenie 3.6 (48), which allowed us to keep related individuals for this sample, increasing EUR sample size by 1,408 cases and 70,225 controls (~15%). GWAS sample size-weighted meta-analysis was performed with METAL (2020 version) (49), with p -values $< 5 \times 10^{-8}$ considered genome-wide significant. Heterogeneity of effects across studies was measured using the I^2 statistic generated by METAL for each variant. We considered significant heterogeneity in case of Chi^2 p -value < 0.05 and $I^2 > 50\%$, as recommended (https://handbook-5-1.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm). LD score intercept was calculated using linkage disequilibrium score regression (LDSC) (50). To estimate statistical power and population stratification, QQ plots were produced and genetic inflation factors λ_{GC} and λ_{1000} were estimated using custom scripts (Supplementary Figure 2 and <https://github.com/precimed>). We also investigated the concordance of findings across ancestries. First, we extracted the AFR sample p -values of loci, which reached genome-wide significance (GWS) in the EUR sample, in order to examine their level of association with

AUD in the AFR sample. Here we considered nominal significance, applying Bonferroni correction for significance, $p < 0.05/n$, where n is the number of loci actually extracted in the AFR sample. Second, we tested for the concordance of effect direction (“sign”) of these loci across samples. To do so, we used a one-sided exact binomial test to test the hypothesis that sign concordance was randomly distributed (proportion=0.5), given the total number of variants and the number of variants with concordant effects in the EUR and AFR samples. There is statistically significant sign concordance when $p < 0.05$.

Definition of genomic loci

Genomic loci were defined using the standard procedure applied in Functional Mapping and Annotation of Genome-Wide Association Studies, version 1.5.3. The list of GWS loci was compared to GWS loci from previous GWASs on alcohol related phenotypes. We report the results from novelty checking separately regarding the AUD phenotype, then combined regarding AUD + other behavioral phenotypes related to alcohol. First, we searched for GWS loci associated with the keyword “alcohol” in the curated lists *GWAScatalog* (<https://www.ebi.ac.uk/gwas/>, accessed February 15, 2024) and *ieu PheWAS* (<https://www.ebi.ac.uk/gwas/>, version 7.6.6). Second, and importantly for the current study, we retrieved the list of GWS loci from the most recent and largest studies on alcohol-related traits, namely PAU (10), and drinking frequency (51). The AUD phenotypes comprised DSM-IV dependence and DSM-5 AUD, while the other phenotypes comprised consumption in drinks/week, consumption in drinks/occasion, beverage preference, problematic alcohol use, heavy vs. light drinking, lifetime alcohol use, and various subscores of assessment tools (criterion count from the DSM-IV, AUDIT). We considered locus boundaries 1 Mbp downstream and 1Mbp upstream of the lead SNPs coordinates. This makes our novelty checking procedure particularly conservative.

Functional annotation

GWS loci were annotated with FUMA (version 1.5.3) by mapping loci to lead SNPs and lead SNPs to credibly mapped genes, defined by at least two converging signals among positional, gene expression and chromatin interaction mapping (FUMA *SNP2GENE*). These genes were linked to cell types using the *ad hoc* FUMA function (52) from 15 available human tissue types. We included the datasets based on the largest number of cell types available (217 cell types), and report findings from all analyses steps (step 1 shown on Figure 2 with the statistical threshold corresponding to step 2, and step 3 described in the main text - see Supplementary Methods file for further details on this analysis). We also report the functional impact of candidate SNPs on protein structure, chromatin conformation and tissue-specific gene expression using *ad hoc in silico* databases, and identified Drugbank hits (FUMA *GENE2FUNCTION*) from credibly-mapped mapped genes. In order to compare our downstream findings for AUD with other alcohol consumption phenotypes, we used summary statistics for the phenotype “drinks per week” [as in a previous study from our group (21)] and meta-analyzed the EUR and AFR samples to pass it through FUMA *SNP2GENE* and *celltype* functions. See Supplementary Methods and Figure 2B file for further details on this analysis. The complete datasets description is available at <https://fuma.ctglab.nl/tutorial#celltype>.

Quantification of polygenic overlap

We applied univariate MiXeR, version 1.3 (53) to estimate SNP-based heritability, discoverability (*i.e.* the average magnitude of additive genetic effects among trait-influencing variants), and polygenicity (*i.e.* the number of trait-influencing variants expected to explain 90% of heritability) of AUD. We used bivariate MiXeR to quantify total polygenic overlap between AUD in EUR and other phenotypes of interest at the genome-wide level (54). Briefly, MiXeR infers characteristics of the genetic architecture of complex traits based on GWAS summary data. First, MiXeR assumes that, for each trait, common genetic variants are a mixture of “causal” variants and noncausal variants, each with its own normal distribution. Based on this assumption, univariate MiXeR estimates the polygenicity, which is the fraction of causal variants, and the discoverability, which is the variance of effect size per causal variant using maximum likelihood estimation. Once SNP-based heritability (h^2_{SNP}) has been derived from these estimates, polygenicity can be restricted to the number of causal variants with strongest effects that explain 90% h^2_{SNP} . This threshold prevents extrapolating model parameters into variants with infinitesimally small effects. Bivariate MiXeR then assumes that common genetic variants can be described as a mixture of four components for a pair of traits. These components are shared “causal” variants, unique “causal” variants for trait one, unique “causal” variants for trait two, and non-causal variants. Based on univariate MiXeR output from each trait, bivariate MiXeR can estimate the shared component’s polygenicity and genetic correlation within the shared component (r_{gs}). Importantly the shared component’s polygenicity does not depend on effect directions. The genome-wide genetic correlation (r_{g}) is then derived from these estimates. r_{gs} and r_{g} depend on effect directions, conversely to the shared component’s polygenicity. Statistical details and validation procedures are available in the MiXeR method paper (54). To evaluate MiXeR reliability, we reported analyses with Akaike Information Criterion (AIC) differences > 0 , as previously reported (18). Due to model fit and use of a European reference genome, MiXeR could only be applied to the EUR sample.

Cross-disorder genetic correlations across European and African ancestries

Cross-ancestry genetic correlations (r_{g}) were estimated between AUD and relevant mental traits and disorders and general medical risk factors and disorders (described above) using Popcorn 3 (55).

Correction for multiple testing

We applied Bonferonni correction for multiple testing to GWS loci, genetic correlation and tissue enrichment. Cell type specificity analyses applied both Bonferonni (step one) and FDR (steps two and three).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments and funding sources

This work was partly performed on the TSD (Tjeneste for Sensitive Data) facilities, owned by the University of Oslo and operated and developed by the TSD service group at the University of Oslo, IT Department (USIT).

Computations were also performed on resources provided by UNINETT Sigma2, the National Infrastructure for High-Performance Computing and Data Storage in Norway. The authors gratefully acknowledge support from NIH (NS057198, EB000790, 1R01MH124839), the Research Council of Norway (229129, 213837, 223273, 324252, 248980, 334920), the South-East Norway Regional Health Authority (2017–112, 2019–108), and KG Jebsen Stiftelsen (SKGJ-MED-021). This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement nos. 847776, 964874, and 801133 (Marie Skłodowska-Curie grant agreement). Romain Icick thanks the INTPART program (PI: Dr TV Lagerberg) for supporting his fellowship at NORMENT. *Core funding of third party data acquisition*: Million Veteran Program (MVP), United States Department of Veterans Affairs; FINNGEN, Business Finland and the pharmaceutical industry partners (see <https://www.finngen.fi/en/funding>); UK Biobank (UKB), Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and the Northwest Regional Development Agency (see <https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/our-funding>); Psychiatric Genomics Consortium (PGC), US National Institute of Mental Health and the US National Institute of Drug Abuse (see <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5756100/>).

Data availability statement

The summary statistics used for the current study were obtained from third party, and are thus not fully available to all authors. Two of them are publicly available [FINNGEN, https://www.finngen.fi/en/access_results, R6 public release; Psychiatric Genomics Consortium (PGC), <https://pgc.unc.edu/for-researchers/download-results/>]. We accessed UK biobank individual-level genotype to extend publicly available summary statistics for alcohol-related traits GWAS that are otherwise publicly available. Likewise, MVP data can be accessed through the dbGaP website, study phs001672.v9.p1.

The summary statistics and data underlying the figures are thus only available to people with registered access to MVP data, upon request to Romain Icick (romain.icick@aphp.fr) or Alexey Shadrin (alexey.shadrin@medisin.uio.no).

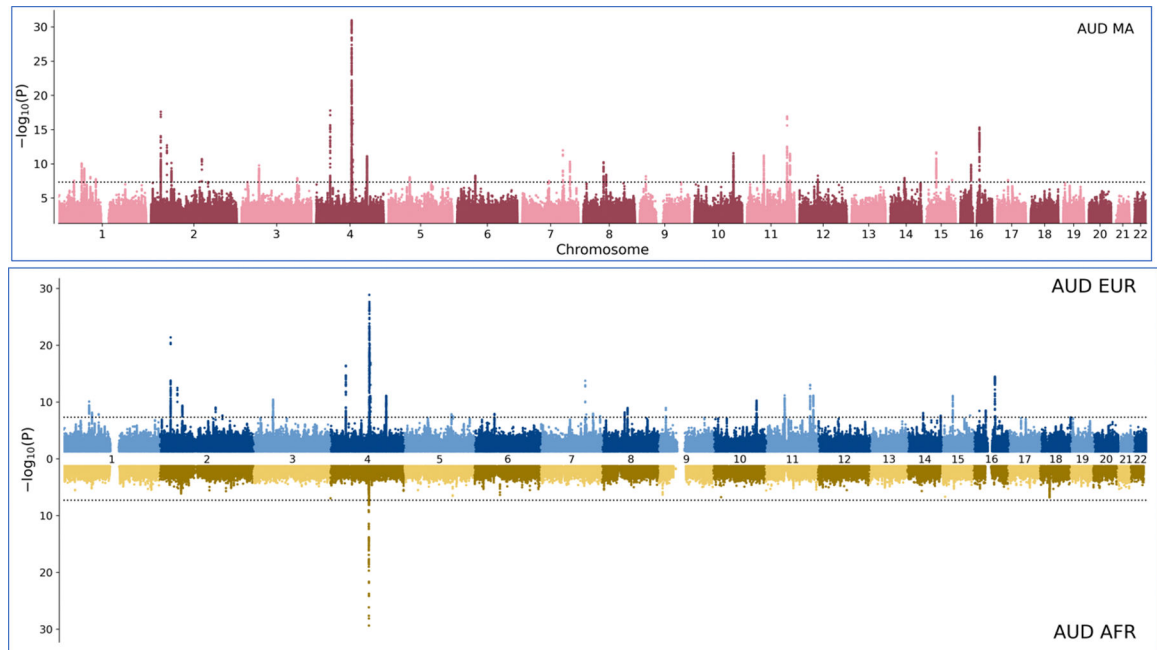
References

1. Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, et al. Epidemiology of DSM-5 Alcohol Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry*. 2015 Aug;72(8):757–66. [PubMed: 26039070]
2. Milanzi EB, Ndasauka Y. Prevalence of Alcohol Addiction in Africa. In: Ndasauka Y, Kayange GM, editors. *Addiction in South and East Africa : Interdisciplinary Approaches* [Internet]. Cham: Springer International Publishing; 2019 [cited 2022 Sep 6]. p. 215–28. Available from: 10.1007/978-3-030-13593-5_13
3. Carvalho AF, Heilig M, Perez A, Probst C, Rehm J. Alcohol use disorders. *The Lancet*. 2019 Aug 31;394(10200):781–92.
4. Rehm J, Shield KD. Global Burden of Alcohol Use Disorders and Alcohol Liver Disease. *Biomedicine*. 2019 Dec;7(4):99. [PubMed: 31847084]
5. Verhulst B, Neale MC, Kendler KS. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med*. 2015 Apr;45(5):1061–72. [PubMed: 25171596]
6. Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat Commun*. 2019 Apr 2;10(1):1499. [PubMed: 30940813]
7. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018 Dec;21(12):1656–69. [PubMed: 30482948]
8. Sanchez-Roige S, Palmer AA, Fontanillas P, Elson SL, Adams MJ, Howard DM, et al. Genome-wide association study meta-analysis of the Alcohol Use Disorder Identification Test (AUDIT) in two population-based cohorts. *Am J Psychiatry*. 2019 Feb 1;176(2):107–18. [PubMed: 30336701]

9. Zhou H, Sealock JM, Sanchez-Roige S, Clarke TK, Levey DF, Cheng Z, 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 Jul;23(7):809–18. [PubMed: 32451486]
10. Zhou H, Kember RL, Deak JD, Xu H, Toikumo S, Yuan K, et al. Multi-ancestry study of the genetics of problematic alcohol use in over 1 million individuals. *Nat Med*. 2023 Dec 7;1–9. [PubMed: 36694061]
11. Zhou H, Kember RL, Deak JD, Xu H, Toikumo S, Yuan K, et al. Multi-ancestry study of the genetics of problematic alcohol use in >1 million individuals [Internet]. medRxiv; 2023 [cited 2023 Feb 7]. p. 2023.01.24.23284960. Available from: <https://www.medrxiv.org/content/10.1101/2023.01.24.23284960v2>
12. Mallard TT, Savage JE, Johnson EC, Huang Y, Edwards AC, Hottenga JJ, et al. Item-Level Genome-Wide Association Study of the Alcohol Use Disorders Identification Test in Three Population-Based Cohorts. *Am J Psychiatry*. 2022 Jan;179(1):58–70. [PubMed: 33985350]
13. Deak JD, Levey DF, Wendt FR, Zhou H, Galimberti M, Kranzler HR, et al. Genome-Wide Investigation of Maximum Habitual Alcohol Intake in US Veterans in Relation to Alcohol Consumption Traits and Alcohol Use Disorder. *JAMA Netw Open*. 2022 Oct 27;5(10):e2238880. [PubMed: 36301540]
14. Heilig M, MacKillop J, Martinez D, Rehm J, Leggio L, Vanderschuren LJMJ. Addiction as a brain disease revised: why it still matters, and the need for consilience. *Neuropsychopharmacology*. 2021 Sep;46(10):1715–23. [PubMed: 33619327]
15. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, et al. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell*. 2019 Oct 17;179(3):589–603. [PubMed: 31607513]
16. Edenberg HJ, McClintick JN. Alcohol Dehydrogenases, Aldehyde Dehydrogenases, and Alcohol Use Disorders: A Critical Review. *Alcohol Clin Exp Res*. 2018 Dec;42(12):2281–97. [PubMed: 30320893]
17. Lago SG, Bahn S. The druggable schizophrenia genome: from repurposing opportunities to unexplored drug targets. *Npj Genomic Med*. 2022 Mar 25;7(1):1–13.
18. Hindley G, Frei O, Shadrin AA, Cheng W, O'Connell KS, Icick R, et al. Charting the Landscape of Genetic Overlap Between Mental Disorders and Related Traits Beyond Genetic Correlation. *Am J Psychiatry*. 2022 Sep 7;appiajp21101051.
19. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, 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 Feb;51(2):237–44. [PubMed: 30643251]
20. Montemitto C, Angebrandt A, Wang TY, Pettorusso M, Abulseoud OA. Mechanistic insights into the efficacy of memantine in treating certain drug addictions. *Prog Neuropsychopharmacol Biol Psychiatry*. 2021 Dec 20;111:110409. [PubMed: 34324921]
21. Icick R, Shadrin A, Hølen B, Karadag N, Lin A, Hindley G, et al. Genetic overlap between mood instability and alcohol-related phenotypes suggests shared biological underpinnings. *Neuropsychopharmacology*. 2022 Oct;47(11):1883–91. [PubMed: 35953530]
22. Sun BB, Kurki MI, Foley CN, Mechakra A, Chen CY, Marshall E, et al. Genetic associations of protein-coding variants in human disease. *Nature*. 2022 Mar 3;603(7899):95–102. [PubMed: 35197637]
23. Gelernter J, Sun N, Polimanti R, Pietrzak RH, Levey DF, Lu Q, et al. Genome-wide Association Study of Maximum Habitual Alcohol Intake in >140,000 U.S. European and African American Veterans Yields Novel Risk Loci. *Biol Psychiatry*. 2019 Sep 1;86(5):365–76. [PubMed: 31151762]
24. Müller-Oehring EM, Jung YC, Sullivan EV, Hawkes WC, Pfefferbaum A, Schulte T. Midbrain-Driven Emotion and Reward Processing in Alcoholism. *Neuropsychopharmacology*. 2013 Sep;38(10):1844–53. [PubMed: 23615665]
25. Bogdan R, Hatoum AS, Johnson EC, Agrawal A. The Genetically Informed Neurobiology of Addiction (GINA) model. *Nat Rev Neurosci*. 2023 Jan;24(1):40–57. [PubMed: 36446900]
26. Bøhle K, Otterholt E, Bjørkly S. Is There an Association Between Salivary Cortisol and Dropping Out of Inpatient Substance Addiction Treatments? A Prospective Repeated Measures Study. *Subst Abuse Res Treat*. 2022;16:11782218221106797.

27. Fish KN, Joffe ME. Targeting prefrontal cortex GABAergic microcircuits for the treatment of alcohol use disorder. *Front Synaptic Neurosci* [Internet]. 2022 [cited 2023 Apr 13];14. Available from: <https://www.frontiersin.org/articles/10.3389/fnsyn.2022.936911>
28. Trojak B, Soudry-Faure A, Abello N, Carpentier M, Jonval L, Allard C, et al. Efficacy of transcranial direct current stimulation (tDCS) in reducing consumption in patients with alcohol use disorders: study protocol for a randomized controlled trial. *Trials*. 2016 May 17;17(1):250. [PubMed: 27188795]
29. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*. 2022 Apr;604(7906):502–8. [PubMed: 35396580]
30. Cooper YA, Teyssier N, Dräger NM, Guo Q, Davis JE, Sattler SM, et al. Functional regulatory variants implicate distinct transcriptional networks in dementia. *Science*. 2022 Aug 19;377(6608):eabi8654. [PubMed: 35981026]
31. Adams C, Conigrave JH, Lewohl J, Haber P, Morley KC. Alcohol use disorder and circulating cytokines: A systematic review and meta-analysis. *Brain Behav Immun*. 2020 Oct;89:501–12. [PubMed: 32805393]
32. Leeuw CA de, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput Biol*. 2015 Apr 17;11(4):e1004219. [PubMed: 25885710]
33. Cacabelos R, Hashimoto R, Takeda M. Pharmacogenomics of antipsychotics efficacy for schizophrenia. *Psychiatry Clin Neurosci*. 2011;65(1):3–19. [PubMed: 21265934]
34. Falk DE, Yi H ye, Hiller-Sturmhöfel S. An epidemiologic analysis of co-occurring alcohol and tobacco use and disorders: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Alcohol Res Health J Natl Inst Alcohol Abuse Alcohol*. 2006;29(3):162–71.
35. Stinson FS, Grant BF, Dawson DA, Ruan WJ, Huang B, Saha T. Comorbidity between DSM-IV alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug Alcohol Depend*. 2005 Oct 1;80(1):105–16. [PubMed: 16157233]
36. McPheeters M, O'Connor EA, Riley S, Kennedy SM, Voisin C, Kuznacik K, et al. Pharmacotherapy for Alcohol Use Disorder: A Systematic Review and Meta-Analysis. *JAMA*. 2023 Nov 7;330(17):1653–65. [PubMed: 37934220]
37. Lepreux G, Shinn GE, Wei G, Suko A, Concepcion G, Sirohi S, et al. Recapitulating phenotypes of alcohol dependence via overexpression of Oprk1 in the ventral tegmental area of non-dependent TH::Cre rats. *Neuropharmacology*. 2023 May 1;228:109457. [PubMed: 36764577]
38. Borruto AM, Fotio Y, Stopponi S, Petrella M, De Carlo S, Domi A, et al. NOP receptor antagonism attenuates reinstatement of alcohol-seeking through modulation of the mesolimbic circuitry in male and female alcohol-preferring rats. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol*. 2021 Nov;46(12):2121–31.
39. Nutt DJ. The role of the opioid system in alcohol dependence. *J Psychopharmacol Oxf Engl*. 2014 Jan;28(1):8–22.
40. Hansson AC, Gründer G, Hirth N, Noori HR, Spanagel R, Sommer WH. Dopamine and opioid systems adaptation in alcoholism revisited: Convergent evidence from positron emission tomography and postmortem studies. *Neurosci Biobehav Rev*. 2019 Nov;106:141–64. [PubMed: 30243576]
41. Inder WJ, Joyce PR, Ellis MJ, Evans MJ, Livesey JH, Donald RA. The effects of alcoholism on the hypothalamic-pituitary-adrenal axis: interaction with endogenous opioid peptides. *Clin Endocrinol (Oxf)*. 1995 Sep;43(3):283–90. [PubMed: 7586596]
42. Hatoum AS, Colbert SMC, Johnson EC, Huggett SB, Deak JD, Pathak GA, et al. Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders. *Nat Ment Health*. 2023 Mar;1(3):210–23. [PubMed: 37250466]
43. Hatoum AS, Johnson EC, Colbert SMC, Polimanti R, Zhou H, Walters RK, et al. The addiction risk factor: A unitary genetic vulnerability characterizes substance use disorders and their associations with common correlates. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol*. 2022 Sep;47(10):1739–45.

44. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 2015 Nov;47(11):1236–41. [PubMed: 26414676]
45. Zhou H, Kalayasiri R, Sun Y, Nuñez YZ, Deng HW, Chen XD, et al. Genome-wide meta-analysis of alcohol use disorder in East Asians. *Neuropsychopharmacology.* 2022 Sep;47(10):1791–7. [PubMed: 35094024]
46. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature.* 2023 Jan;613(7944):508–18. [PubMed: 36653562]
47. Im PK, Wright N, Yang L, Chan KH, Chen Y, Guo Y, et al. Alcohol consumption and risks of more than 200 diseases in Chinese men. *Nat Med.* 2023 Jun;29(6):1476–86. [PubMed: 37291211]
48. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, et al. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet.* 2021 Jul;53(7):1097–103. [PubMed: 34017140]
49. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010 Sep 1;26(17):2190–1. [PubMed: 20616382]
50. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015 Mar;47(3):291–5. [PubMed: 25642630]
51. Saunders GRB, Wang X, Chen F, Jang SK, Liu M, Wang C, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature.* 2022 Dec;612(7941):720–4. [PubMed: 36477530]
52. Watanabe K, Umi evi Mirkov M, de Leeuw CA, van den Heuvel MP, Posthuma D. Genetic mapping of cell type specificity for complex traits. *Nat Commun.* 2019 Jul 19;10(1):3222. [PubMed: 31324783]
53. Holland D, Frei O, Desikan R, Fan CC, Shadrin AA, Smeland OB, et al. Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLOS Genet.* 2020 mai;16(5):e1008612. [PubMed: 32427991]
54. Frei O, Holland D, Smeland OB, Shadrin AA, Fan CC, Maeland S, et al. Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat Commun.* 2019;10(1):1–11. [PubMed: 30602773]
55. Brown BC, Ye CJ, Price AL, Zaitlen N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *Am J Hum Genet.* 2016 Jul 7;99(1):76–88. [PubMed: 27321947]



1

Figure 1: Ancestry Specific Genetic Architecture of AUD in the multi-ancestry analysis (top, red, AUD multi-ancestry) and for the European (top, blue, AUD EUR) and African (bottom, yellow, AUD AFR) samples.

Meta-analysis (METAL) was performed from 2-sided multiple regressions adjusted for sex and the ten first ancestry components for each SNP included in the original GWASs, based on the recessive model. $-\log_{10}$ (raw p-values) are shown on the y-axis. The dotted line represents the threshold for statistical significance: 5×10^{-8} considering Bonferroni correction. The x-axis represents increasing chromosome numbers from 1 to 22 and positions. Y-axis is truncated to $-\log_{10}(P) = 32$.

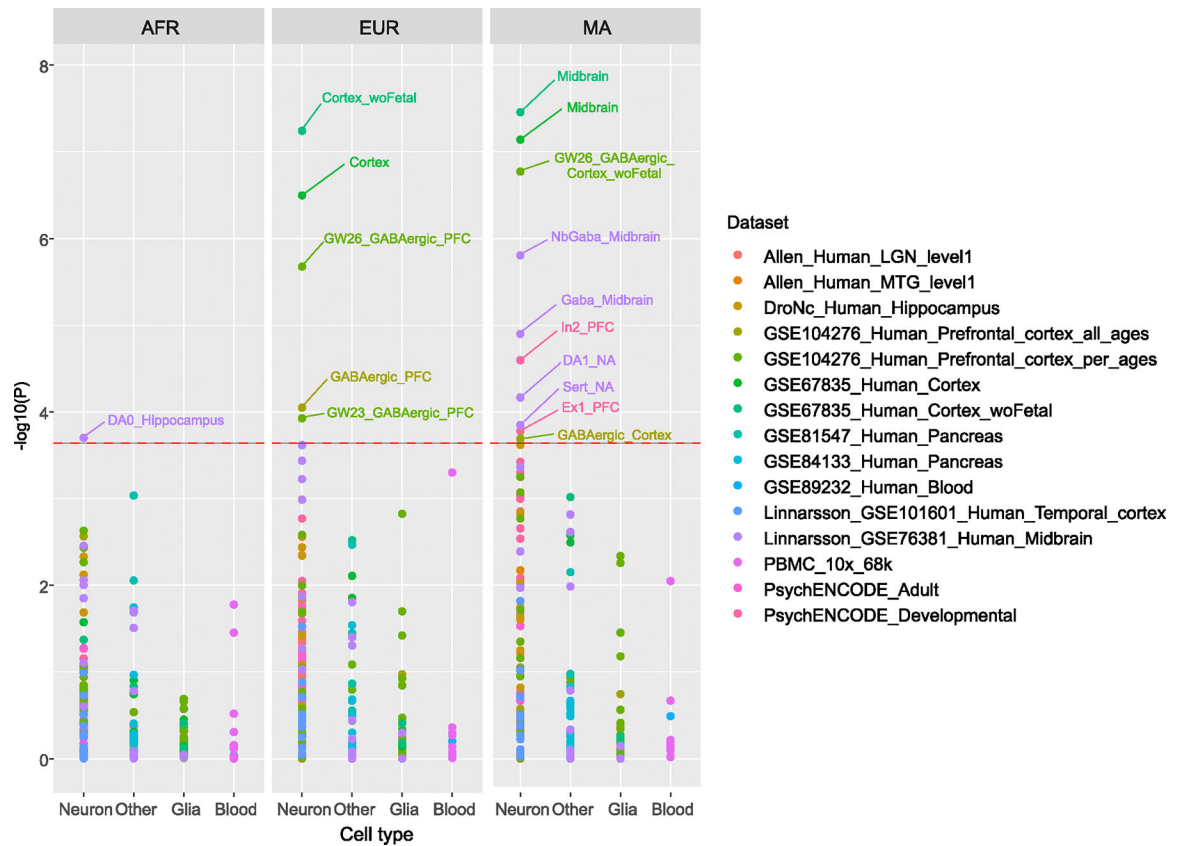
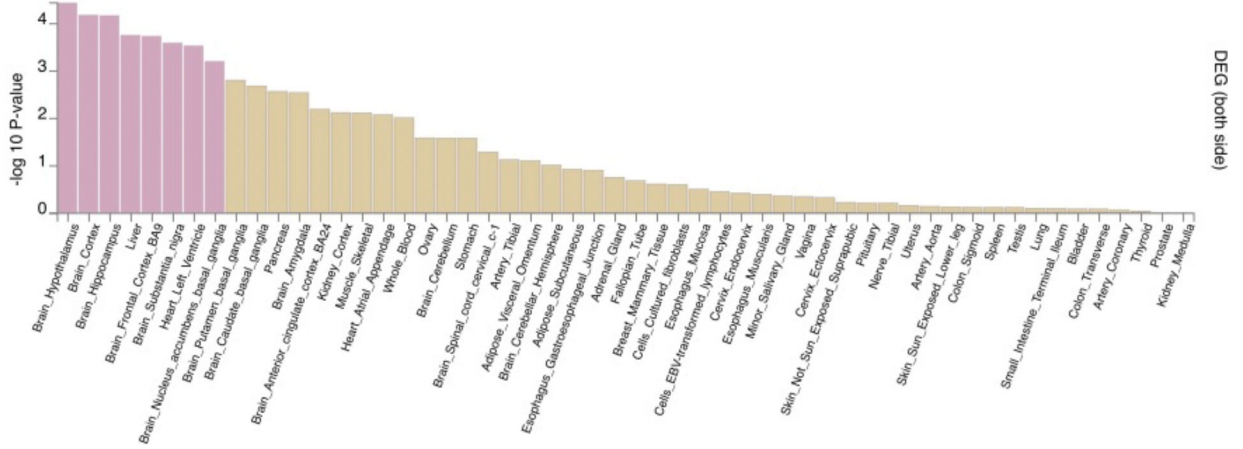


Figure 2. Independent cell types associated with the GWAS meta-analysis results in the African (AFR), European (EUR) and multi-ancestry samples.

Results from FUMA step three analysis obtained with 217 Human cell types. Celltypes are first associated with genes mapped to our study loci using multiple regression analysis. In step two, a systematical step-wise conditional analysis per dataset is performed, by setting thresholds for proportional significance (PS) of the conditional P-value of a cell type relative to the marginal P-value as described in the table. Finally, step three tests, for each pair of cell types from different datasets, three 2-sided regression models, which incorporate the effect of the average expression from the other dataset conditions. P-values are finally Bonferroni-corrected for the number of celltypes considered for analysis in step three (here, threshold at $p = 0.00023$). woFetal, dataset considered without developing cells; GW, gestation week; PFC, prefrontal cortex; exCA1, hippocampal cornu ammonis excitatory neurons.

A) Multi-ancestry



B) AFR

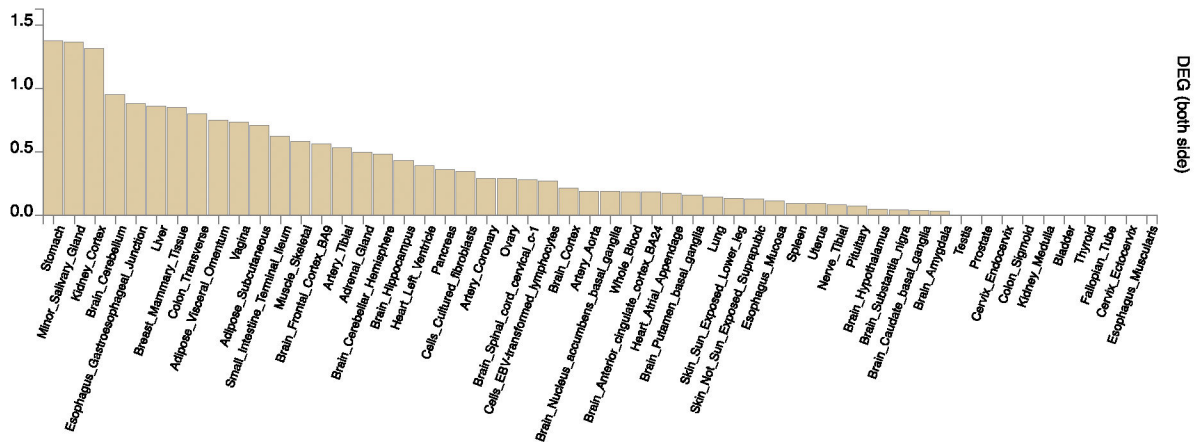
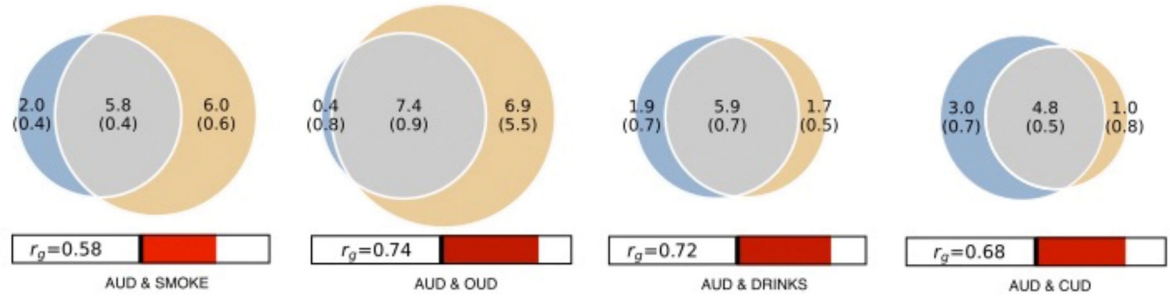


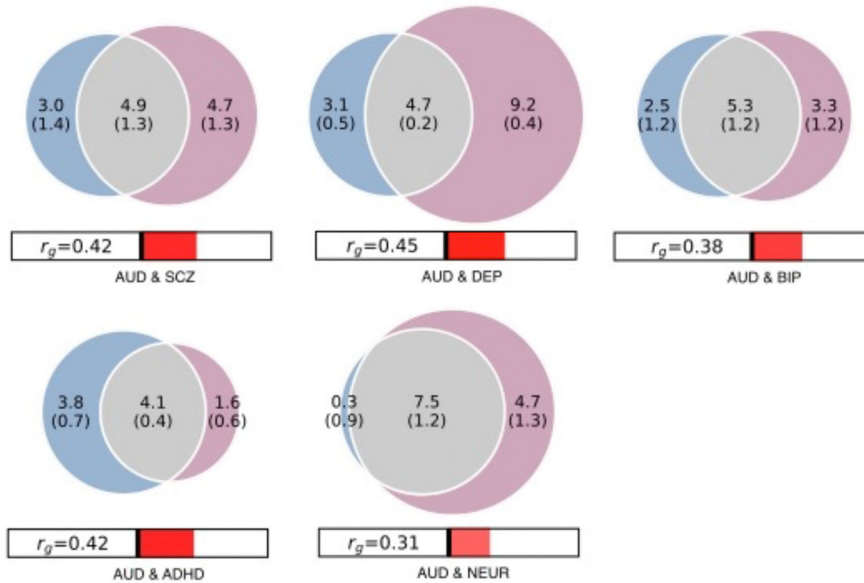
Figure 3: Tissue-specific gene expression enrichment from the multi-ancestry (top panel), African (AFR, middle panel) and European (EUR, bottom panel) analyses.

two-sided t-test for any one of labels against all others. For this, expression values were normalized (zero-mean) following to a log 2 transformation of expression value (EPKM or TPM). Significant enrichment is considered for genes with with P-value < 0.00093 after Bonferroni correction for 54 tissue types and absolute log fold change > 0.58 ; represented in pink. DEG, differentially-enriched genes.

A. Mental disorders and traits (substance)



B. Mental disorders and traits



C. General medical condition

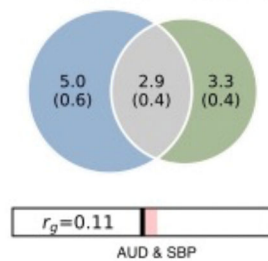


Figure 4: Polygenic overlap between AUD (blue) and clinically relevant phenotypes, after filtering based on estimation of MiXeR stability using the Akaike Informant Criterion.
 A) CUD, cannabis use disorder; OUD, opioid use disorder; DRINKS, drinks / week; SMOKE, age at smoking initiation (yellow); B) ADHD, attention deficit/hyperactivity disorder; BIP, bipolar disorder; DEP, major depression; SCZ, schizophrenia; NEUR, neuroticism (pink); C) SBP, systolic blood pressure (green).

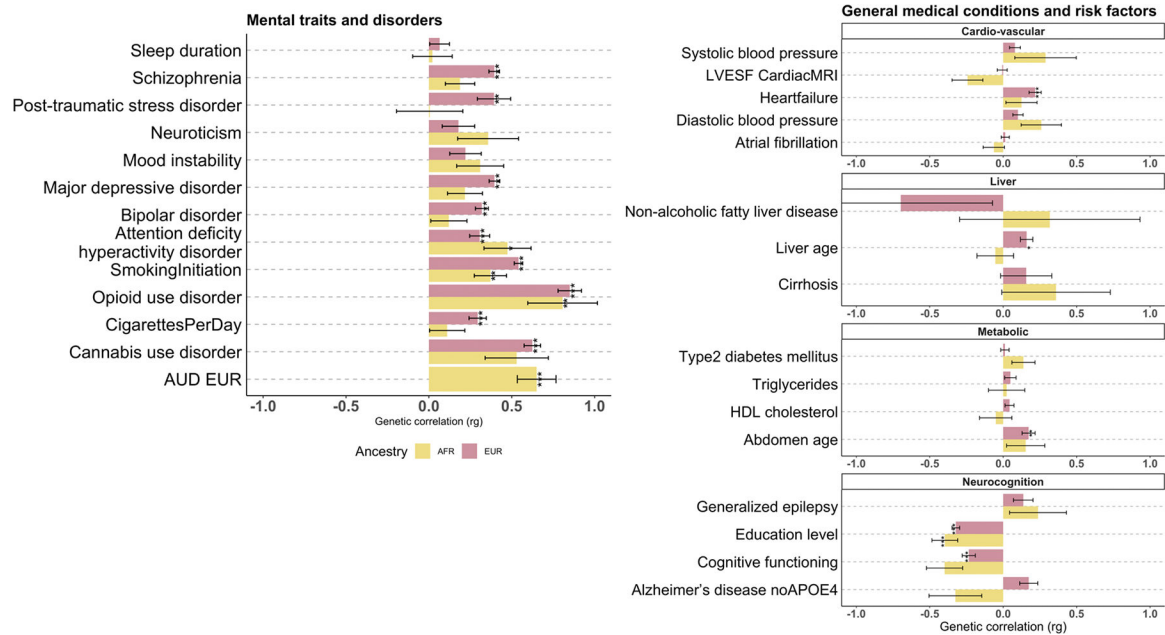


Figure 5: Genetic correlation of AUD with mental traits and disorders (top and middle) and with general medical conditions and risk factors (bottom), including neuropsychiatric diseases, for the AFR and the EUR samples, separately.

The secondary traits are listed. Error bars represent standard error for the Spearman correlation coefficient (two-sided test). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (Bonferroni-corrected). Exact p-values for each correlation is listed in **Supplementary Table 8**. LVESF, left ventricular ejection systolic fraction; MRI, magnetic resonance imaging; HDL, high density lipoproteins; APOE4, Apolipoprotein ϵ 4 locus. AUD EUR, Alcohol Use disorder. EUR, European. Cardiovascular traits included GWAS for abdominal age ($n_{\text{eff}}=36,784$), diastolic ($n_{\text{eff}}=1,073,636$) and systolic blood pressure ($n_{\text{eff}}=1,054,597$), left ventricular ejection systolic fraction (cardiac MRI, $n=36,784$), type 2 diabetes mellitus ($n_{\text{eff}}=497,442$), atrial fibrillation ($n_{\text{eff}}=197,847.1$), and heart failure ($n_{\text{eff}}=180,075.7$). Liver-related traits included GWAS for non-alcoholic fatty liver disease ($n_{\text{eff}}=37557.2$), liver cirrhosis ($n_{\text{eff}}=19140.9$), and liver age ($n_{\text{eff}}=45,552$). Neurocognition traits included Alzheimer’s disease (after exclusion of the apolipoprotein ϵ 4 locus) ($n_{\text{eff}}=306,866.2$), generalized epilepsy ($n_{\text{eff}}=40,227.8$), cognitive functioning ($n_{\text{eff}}=269,867$), and education level ($n_{\text{eff}}=765,283$). Psychiatric disorders and substance use traits included cigarettes per day ($n_{\text{eff}}=337,334$), smoking initiation ($n_{\text{eff}}=1,232,091$), cannabis use disorder ($n_{\text{eff}}=79,107.3$), opioid use disorder ($n_{\text{eff}}=74,635.9$), attention deficit hyperactivity disorder ($n_{\text{eff}}=128,213.8$), bipolar disorder ($n_{\text{eff}}=147,290.9$), major depressive disorder ($n_{\text{eff}}=449,855.9$), post-traumatic stress disorder ($n_{\text{eff}}=69,897.3$), schizophrenia ($n_{\text{eff}}=126,282$). Psychiatric traits included neuroticism ($n_{\text{eff}}=390,278$), mood instability ($n_{\text{eff}}=157,039$) and sleep duration ($n_{\text{eff}}=384,225$). Serum lipids included GWAS for HDL cholesterol and triglycerides ($n_{\text{eff}}=1,320,016$). All samples were of European ancestry, and GWAS data was downloaded from the repositories indicated in the publications between June 15 and July 1st 2022.

Table 2:

GWAS data, including ancestry breakdown.

GWAS	Phenotype	Sample size (cases)	Sample size (cases) by ancestry (n)	Age (years), sex (%)	Maximum number of SNPs (n, millions)
<i>MVP</i>	AUD (+ severe acute intoxication) (ICD-9/10)	305,511 (68,913)	EUR: 221,137 (45,943) AFR: 56,648 (17,267) HA: 14,175 (3,449) EAS: 13,551 (2,254)	30–75 years 8% women	6.8
<i>UK Biobank</i>	Alcohol abuse / dependence (ICD 10)	425,224 (8,201)	EUR: 409,558 (7,910) AFR: 7,045 (87) SAS: 8,621 (104)	30–69 years 56% women	9.9
<i>FinnGen</i>	AUD (ICD-9/10)	260,405 (10,688)	EUR: 260,405 (10,688)	Median 63 years * 56% women (total sample)	20.2
<i>PGC without FinnGen</i>	Alcohol dependence (DSM-IV)	50,310 (14,377)	EUR: 44,030 (11,042) AFR: 6,280 (3,335)	39% women > 18 years Pooled mean 34.7 from 30% of the sample **	10.9
<i>Total (multi ancestry)</i>	AUD (+ severe acute intoxication)	1,041,450 (102,179) <i>Neff</i> =321,343	<ul style="list-style-type: none"> • EUR: 935,130 (75,583) <i>Total Neff</i>=249,626 (FinnGen, 40,997; PGC without FinnGen, 25,267; MVP, 152,333; UKB, 31,029) • AFR: 70,060 (20,689) <i>Total Neff</i>=53,350 (MVP, 48,015; PGC without FinnGen, 4,991; UKB, 343,755) • HA: 14,175 (3,449) <i>Total Neff</i>=10,439 (MVP) • EAS+SAS: 21,972 (2,358) <i>Neff</i>=EAS 7,516 (MVP), SAS 411 (UKB) 	Pooled: Mean age: 51.8 Women: 32%	Multi-ancestry: 24.8 EUR: 19.1 AFR: 9.9

EUR, European; AFR, African; HA, Hispanic American; EAS, East Asian; SAS, South Asian; AUD, alcohol Use Disorder; ICD, International Classification of Diseases; DSM, Diagnostic and Statistical Manual of Mental Disorders; GWAS, Genome-wide association studies; MVP, Million Veteran Program; PGC, Psychiatric Genomic Consortium; SNP, Single nucleotide polymorphism. Neff, effective sample size.

* median age is less relevant for the FINNGEN cohort compared to the others since all individuals are followed-up from birth.

** Pooled mean age was obtained using the cohorts mean ages, except for the PGC sample, where it was extrapolated from the mean age of the Pale-Yenn and SAGE cohorts, which represent ~30% of the total PGC sample - yielding pooled mean age =34.7 based on PMID 27028160.