

Human genetics and epigenetics of alcohol use disorder

Hang Zhou^{1,2,3,4} and Joel Gelernter^{1,2,5,6}

¹Department of Psychiatry, Yale School of Medicine, New Haven, Connecticut, USA. ²Veterans Affairs Connecticut Healthcare System, West Haven, Connecticut, USA. ³Department of Biomedical Informatics and Data Science, ⁴Center for Brain and Mind Health, ⁵Department of Genetics, and ⁶Department of Neuroscience, Yale School of Medicine, New Haven, Connecticut, USA.

Alcohol use disorder (AUD) is a prominent contributor to global morbidity and mortality. Its complex etiology involves genetics, epigenetics, and environmental factors. We review progress in understanding the genetics and epigenetics of AUD, summarizing the key findings. Advancements in technology over the decades have elevated research from early candidate gene studies to present-day genome-wide scans, unveiling numerous genetic and epigenetic risk factors for AUD. The latest GWAS on more than one million participants identified more than 100 genetic variants, and the largest epigenome-wide association studies (EWAS) in blood and brain samples have revealed tissue-specific epigenetic changes. Downstream analyses revealed enriched pathways, genetic correlations with other traits, transcriptome-wide association in brain tissues, and drug-gene interactions for AUD. We also discuss limitations and future directions, including increasing the power of GWAS and EWAS studies as well as expanding the diversity of populations included in these analyses. Larger samples, novel technologies, and analytic approaches are essential; these include whole-genome sequencing, multiomics, single-cell sequencing, spatial transcriptomics, deep-learning prediction of variant function, and integrated methods for disease risk prediction.

Introduction

Alcohol use disorder (AUD) is a chronic relapsing disorder that progresses through a three-stage addiction cycle involving neurocircuitry in the basal ganglia, extended amygdala, and prefrontal cortex (1). Different terminologies and diagnostic approaches have been applied over the decades to AUD and related traits, including alcohol dependence and alcohol abuse based on DSM-IV (2) and prior editions and AUD based on DSM-5 (3). Diagnostic criteria include tolerance to the effects of alcohol; withdrawal in the absence of alcohol use; inability to control or reduce alcohol intake; preoccupation with alcohol to the detriment of work, family, and social priorities; and others (4). In addition, ICD diagnostic codes are widely used in clinical settings and reflected in electronic health records to diagnose AUD and related disorders. If not otherwise specified, henceforth, we use a broad definition of AUD to encompass both alcohol dependence and alcohol use disorder.

AUD and excessive alcohol use contribute greatly to the global disease burden and causing substantial adverse health effects (5). However, only three medications (disulfiram, naltrexone, and acamprosate) are approved by the US FDA for treating AUD (4). Thus, a lack of treatment options persists despite the clinical importance of the problem.

Genetics of AUD: candidate genes and linkage studies

AUD is a complex disorder with significant environmental and genetic components. Genetic influences on AUD have long been established (6–8), and family and twin studies have reported about approximately 0.50 (95% CI, 0.43–0.53) genetic heritability (9–13). Three waves of genetic studies conducted in the past decades have identified susceptibility genes (14–18). The first wave involved candidate gene studies. In this Review, we omit from discussion underpowered studies. Candidate genes related to ethanol metabolism were intensively investigated for associations with AUD (19–25). Alcohol dehydrogenases (ADHs), such as *ADH1B* and *ALDH2*, are enzymes that oxidize ethanol into acetaldehyde, and aldehyde dehydrogenases are enzymes that catalyze aldehydes to their corresponding acids. The importance of functional variants such as rs1229984 (encoding His48Arg) in *ADH1B*, and rs671 (encoding Glu504Lys) in *ALDH2* is well-established. The increased catalytic efficiency of *ADH1B* (conferred by the His48 allele) or lower activity of *ALDH2* (by the Lys504 allele, which is common exclusively in East Asians) leads to accumulation of acetaldehyde and flushing (26, 27), which discourages further alcohol intake, thus protecting against AUD (28–31). Another coding variant, *ADH1B**rs2066702 (Arg369Cys), has been associated with AUD but only in African populations (unless specified otherwise, the African ancestry samples in studies mentioned in this article were African-Americans) as it is nonpolymorphic in other populations (32, 33). Numerous additional candidate genes have failed to survive the GWAS era, and these will not be discussed here.

The second wave of genetic studies involved linkage studies, which utilize family data to identify genomic regions associated with AUD (34, 35). This wave was followed by positional candidate

Conflict of interest: JG holds US patent 10,900,082 titled, “Genotype-guided dosing of opioid agonists,” issued January 26, 2021. JG is paid for his editorial work in the journal *Complex Psychiatry*.

Copyright: © 2024, Zhou et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

Reference information: *J Clin Invest.* 2024;134(16):e172885.

<https://doi.org/10.1172/JCI172885>.

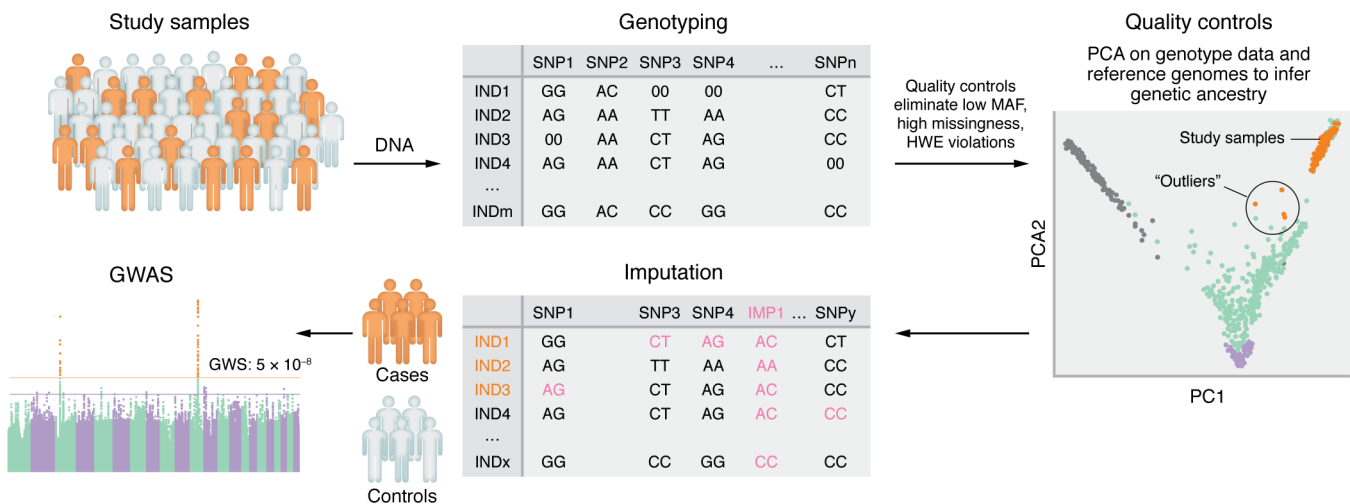


Figure 1. Workflow of GWAS. In a typical GWAS study, participants are recruited and provide written informed consent and blood or saliva samples for DNA extraction and genotyping using microarray (“00” indicates missing genotype call). Basic quality controls are performed to remove SNPs with low minor allele frequencies (MAF), high genotype missingness rate, or violation of Hardy Weinberg Equilibrium expectations (HWE) and remove samples with high genotype missingness. Since genetic factors often differ according to ancestry, principal component analysis (PCA) is performed on the data after quality controls with reference genomes – for example, the 1000 Genomes Project (165) – to infer the genetic ancestries of the study samples and remove genetic outliers (the results from different ancestry groups can then be combined by meta-analysis). Then, the remained samples and the data after quality control are imputed for millions more variants (imputed genotypes and SNPs [IMP], labeled in purple) using reference genomes (165–168). Imputation takes advantage of known patterns of linkage disequilibrium to provide useful data for many more variants than are genotyped directly. A study trait, in the context of either case-control status (for example, AUD) or continuous measurement (for example, AUD criterion counts), is assessed in the cohort. Regression models implemented in computational tools (169–175) are applied to test the association between each variant and the studied trait within the genetically inferred population group, adjusting for covariates including age, sex, and the top principal components of ancestry. Variants with $P < 5 \times 10^{-8}$ are considered genome-wide significant (GWS) after multiple testing corrections for the number of independent genomic regions evaluated (176).

gene studies to target the relevant genes or variants. Several candidate genes have been identified using linkage analyses, including genes that encode GABA receptors (*GABRA1* and *GABRA2*) (36–39), *CHRM2* (cholinergic receptor muscarinic 2) (40, 41), and others. However, these genes were not identified in later GWAS.

GWAS of AUD

Different from the previous candidate-gene studies, GWAS is a hypothesis-free method that scans genome-wide common variants using microarray genotyping or sequencing to identify associations with study traits (Figure 1) (42, 43). Substantial progress has been made in the wave of genetic studies of AUD using GWAS (Figure 2 and Table 1).

In 2009, the first GWAS of AUD was conducted in a German sample comprising 487 cases of AUD and 1,358 population-based controls; no variants reached the genome-wide significant (GWS) threshold (44). In 2011, the same team augmented the sample size by recruiting more participants and identified a variant located between *ADH1B* and *ADH1C*. In this study, the polygenic risk score (PRS, a method that quantifies an individual’s genetic predisposition to a particular trait or disease by summing the effects of multiple genetic variants across the genome) for AUD was investigated for the first time to test the association with AUD in independent samples, including splitting the study samples into two halves randomly, plus two samples from the National Library of Medicine’s Database of Genotypes and Phenotypes (45). In 2010, a study of Dutch and Australian samples was the first AUD GWAS to apply imputation for missing SNPs using the HapMap

reference panel (46). The study reported no GWS results for AUD, but three SNPs were identified for comorbid AUD and nicotine dependence (47). No association was identified in a general community sample in Australia, but this study discussed the polygenic nature of AUD and projected the need for larger sample size (48). Reanalyses of these cohorts were undertaken to enhance statistical power (49–52).

Insufficient degree of genetic diversity in study populations has been a persistent challenge in human genetic studies, with the majority of study participants being of European ancestry (53, 54). Including non-European populations in AUD GWAS could help illuminate the shared and specific genetic architectures across populations. Three GWAS of AUD extended the gene discovery effort to more populations (55–57). However, no GWS signals were identified in these studies. Subsequently, several GWAS of AUD were performed of East Asian samples. The first was a study of a Korean sample with 396 unrelated individuals, which identified both the *ADH1B**rs1229984 and *ALDH2**rs671 (58). Thus far, the well-known functional coding variants rs1229984 and East Asian-specific rs671 have been confirmed by the GWAS approach. Other studies also identified the *ALDH2* region to be associated with AUD in East Asian samples (59–62), with no additional risk variants identified beyond these two regions. Meta-analyzing (a method that combine GWAS results from two or more separate cohorts) newly recruited samples with previously published summary data provides an opportunity to uncover additional risk variants. In 2014, a study involving more than 10,000 individuals of African and European ancestries was conducted, combining

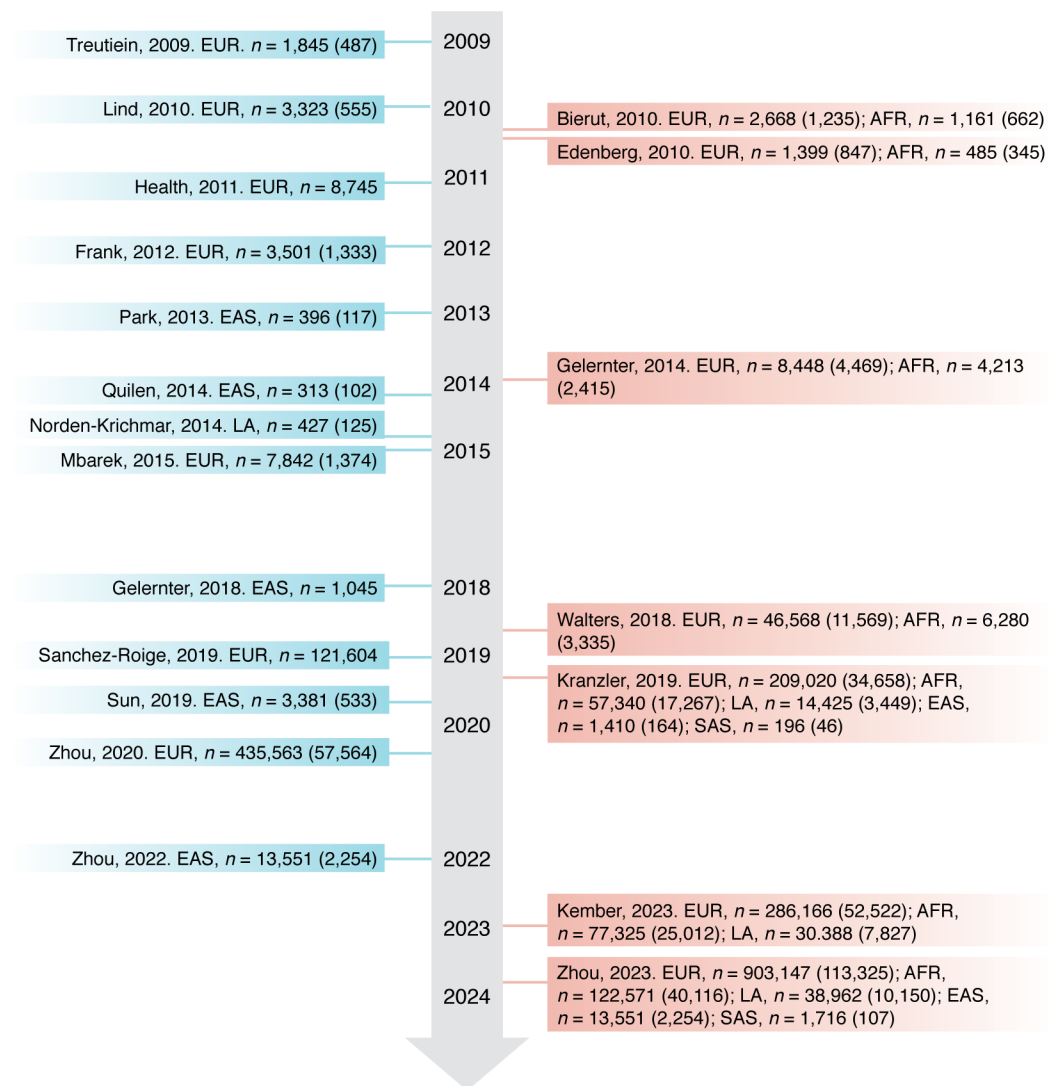


Figure 2. Timeline of the GWAS of AUD. Only studies with new samples are included here. Note that Sanchez-Roige et al. (72) presented here is the GWAS of AUDIT-P in the UK Biobank. Studies with multiple ancestries are listed on the right. Numbers in the brackets are numbers of AUD cases included in the study. Table 1 highlights these studies in more detail. EUR, European; AFR, African; EAS, East Asian; LA, Latin American; SAS, South Asian.

several cohorts. In both case-control analysis and criterion-count analysis, the ADH gene region was confirmed, and a strong association with the coding variant rs2066702 (Arg369Cys) in *ADH1B* was identified in African samples. Four other loci were associated with AUD in the criterion-count analysis: two in European ancestry samples and two in African ancestry samples (63). A large meta-analysis of AUD from the Psychiatric Genomics Consortium combined 28 studies of individuals of both European ($n = 46,568$) and African ($n = 6,280$) ancestries, confirming associations with the ADH gene cluster; however, no additional risk variants were discovered (64). This study also investigated the genetic correlations between AUD and many other traits, observing significant correlations with psychiatric disorders, substance use traits, and socioeconomic status (educational attainment and Townsend deprivation score). PRS derived from the European GWAS showed weaker predictions in independent African sample than the PRS derived from African GWAS, indicating limited portability of PRS across ancestries (65).

Besides DSM or ICD diagnosis, AUD can be assessed using the Alcohol Use Disorders Identification Test (AUDIT), a 10-item questionnaire developed by the WHO to measure hazardous or harmful

drinking in the past year (66). Questions 1–3 are aimed at assessing alcohol consumption levels (AUDIT-C), and questions 4–10 are focused on evaluating problematic alcohol drinking (AUDIT-P). The AUDIT is useful to screen for AUDs (67, 68). Thus, such a questionnaire could be implemented as a cost-effective strategy for phenotyping samples in large-scale cohorts or biobanks.

The first two GWAS of AUDIT scores identified no association (69, 70). A later study of AUDIT in two population-based cohorts, the UK Biobank (71) and 23andMe (69), totaling 141,932 participants, identified 15 independent signals in 11 genomic loci for AUDIT total score, many of them novel (72). Four loci were associated with AUDIT-P subscore, including the ADH region, *KLB* (encoding β -klotho), and *SLC39A8* (solute carrier family 39 [zinc transporter], member 8). Another key finding from this study is that the genetic architecture of AUDIT-P differed from AUDIT-C, and AUDIT-P is genetically correlated with AUD more strongly than AUDIT-C. Transcriptome-wide association study (TWAS) (73) identified 26 genes whose predicted gene expression in brain tissues were associated with AUDIT.

A study in the Million Veteran Program (MVP) (74) investigated both AUD and alcohol consumption (measured by AUDIT-C)

Table 1. GWAS of AUD

Study ^A	Details	Findings ^B	Note
Treutlein, 2009 (44)	EUR, <i>n</i> = 1,845; Ncase = 487	No GWS	All males
Lind, 2010 (47)	EUR, <i>n</i> = 3,323; Ncase = 555	No GWS	First imputed
Edenberg, 2010 (56)	EUR, <i>n</i> = 1,399; Ncase = 847 AFR, <i>n</i> = 485; Ncase = 345	No GWS	
Bierut, 2010 (55)	EUR, <i>n</i> = 2,668; Ncase = 1,235 AFR, <i>n</i> = 1,161; Ncase = 662	No GWS	
Health, 2011 (48)	EUR, <i>n</i> = 8,754	No GWS	AUD factor score
Frank, 2012 (45)	EUR, <i>n</i> = 3,501; Ncase = 1,333	rs1789891	All males
Park, 2013 (58)	EAS, <i>n</i> = 396; Ncase = 117	<i>ADH1B</i> *rs1229984 and <i>ALDH2</i> *rs671	
Gelernter, 2014 (63)	EUR, <i>n</i> = 5,131 AFR, <i>n</i> = 4,629 EUR, <i>n</i> = 8,448; Ncase = 4,469 AFR, <i>n</i> = 4,213; Ncase = 2,415	<i>ADH1B</i> *rs1229984; <i>ADH1B</i> *rs2066702 in AFR and 4 other loci <i>ADH1B</i> *rs1229984; <i>ADH1B</i> *rs2066702 in AFR	Ordinal trait Case-control
Quillen, 2014 (59)	EAS, <i>n</i> = 313; Ncase = 102	<i>ALDH2</i> *rs671	All males
Norden-Krichmar, 2014 (57)	LA, <i>n</i> = 427; Ncase = 125	NO GWS	
Mbarek, 2015 (70)	EUR, <i>n</i> = 7,842; Ncase = 1,374	NO GWS	AUDIT case/control
Gelernter, 2018 (60)	EAS, <i>n</i> = 1,045	<i>ALDH2</i> locus	Ordinal trait
Walters, 2018 (64)	EUR, <i>n</i> = 46,568; Ncase = 11,569 AFR, <i>n</i> = 6,280; Ncase = 3,335	<i>ADH1B</i> *rs1229984; <i>ADH1B</i> *rs2066702 in AFR	
Sanchez-Roige, 2019 (69)	EUR, <i>n</i> = 20,328	NO GWS	AUDIT
Sanchez-Roige, 2019 (72)	EUR, <i>n</i> = 141,932	15 independent variants in 11 ^C loci for AUDIT; <i>KLB</i> , <i>ADH</i> region, <i>SLC39A8</i> for AUDIT-P	AUDIT, AUDIT-P
Kranzler, 2019 (75)	EUR, <i>n</i> = 209,020; Ncase = 34,658 AFR, <i>n</i> = 57,340; Ncase = 17,267 LA, <i>n</i> = 14,425; Ncase = 3,449 EAS, <i>n</i> = 1,410; Ncase = 164 SAS, <i>n</i> = 196; Ncase = 46	10 loci in EUR; 2 in AFR; 2 in LA; 15 independent variants in all analyses	
Sun, 2019 (61)	EAS, <i>n</i> = 3,381; Ncase = 533	<i>ADH</i> region and <i>ALDH2</i>	All males
Zhou, 2020 (78)	EUR, <i>n</i> = 435,563; Ncase = 57,564	24 independent variants with AUD; 29 with PAU	PAU
Zhou, 2022 (62)	EAS, <i>n</i> = 13,551; Ncase = 2,254	<i>ADH</i> region and <i>ALDH2</i>	
Kember, 2023 (77)	EUR, <i>n</i> = 286,166; Ncase = 52,522 AFR, <i>n</i> = 77,325; Ncase = 25,012 LA, <i>n</i> = 30,388; Ncase = 7,827	19 ^D independent variants in EUR; 4 in AFR; 1 in LA; 32 ^E in all analyses	Less stringent AUD definition was also tested
Zhou, 2023 (80)	EUR, <i>n</i> = 903,147; Ncase = 113,325 AFR, <i>n</i> = 122,571; Ncase = 40,116 LA, <i>n</i> = 38,962; Ncase = 10,150 EAS, <i>n</i> = 13,551; Ncase = 2,254 SAS, <i>n</i> = 1,716; Ncase = 107	85 independent variants in EUR; <i>ADH1B</i> *rs1229984 and rs2066702 in AFR; <i>ADH1B</i> *rs1229984 in LA; <i>ADH</i> region and <i>ALDH2</i> in EAS; 110 in all analyses	PAU

Studies are sorted by publication date. EUR, European; AFR, African; EAS, East Asian; LA, Latin American; SAS, South Asian; GWS, genome-wide significance; ADH, alcohol dehydrogenase; AUDIT, Alcohol Use Disorders Identification Test; PAU, problematic alcohol use; Ncase, number of cases. ^AOnly the studies with new samples are included here. ^BThe genome-wide significant results from discovery samples are presented in this table, not including some results combined with replication samples. ^CAfter conditional analyses. ^DConditional analyses were not performed in each ancestry; this number is a proxy number. ^EIndependent signals in the cross-ancestry meta-analysis plus the ancestry-specific signals.

in five population groups, including European, African, East Asian, Latin American, and South Asian populations (75). This study included 274,391 participants, with 55,584 diagnosed with AUD based on the ICD codes. Fifteen independent variants (after conditional analyses) in 10 loci were identified in multiple ancestries, including 10 in European, 2 in African, and 2 in Latin American ancestries. Partitioning heritability analysis to investigate how the cell type-specific functional categories of the genome contrib-

ute to the heritability of a complex disease (76) indicated that the CNS was the most significantly enriched cell type group for AUD, confirming with genetic evidence that AUD is a brain-related disorder. This study also delivered a key finding that the genetic architecture of alcohol consumption (measured by AUDIT-C) differs from that of AUD (a similar pattern was observed between AUDIT-C and AUDIT-P, ref. 72), stressing that analyzing AUD or AUDIT-P separately from alcohol consumption traits would

reduce heterogeneity. Prior to these two key papers, it was not recognized that quantity/frequency versus dependence measures differed genetically and therefore biologically. Another cross-ancestry study used longitudinal data from MVP and confirmed this difference between AUD and AUDIT-C and identified novel loci with both traits. Specifically, this study identified a set of variants with effects on AUD that are not mediated through alcohol consumption (i.e., AUDIT-C) (77).

A subsequent study of problematic alcohol use (PAU), a proxy phenotype of AUD, combined AUD from the MVP and Psychiatric Genomics Consortium and AUDIT-P from UK Biobank and identified 29 independent risk variants in 435,563 EUR participants (78). In this study, the genetic correlation between AUDIT-P and AUD was estimated to be 0.71 (standard error = 0.05), justifying the proxy-phenotype meta-analysis of PAU across these data sets. This study noted the heterogeneity among these phenotypes and discussed that associations specific to each definition could have been attenuated. A total of 327 known drug-gene interactions were found for 16 associated genes, with *DRD2* having the most drug interactions ($n = 177$) followed by *BDNF* ($n = 68$) and *PDE4B* ($n = 36$). Phenome-wide PRS analysis in the independent biobank BioVU confirmed the genetic correlations between PAU and substance use and psychiatric disorders. Pathways including reactome ethanol oxidation and ethanol and alcohol metabolism were the most significantly enriched for AUD. TWAS showed significant enrichments in several brain tissues, including the cerebellum and cortex, further illustrating the tissue-specific mechanisms of this brain-related disease. Mendelian randomization analysis (79), a set of methods that uses genetic variants as instrumental variables to estimate the causal relationship between exposure and outcome, suggested liability to substance use, psychiatric status, risk-taking behavior, and cognitive performance having causal effects on the liability to PAU.

Findings of the 2023 multiancestry GWAS of PAU

Thus far, studies have identified risk genes associated with AUD in multiple ancestries and have repeatedly confirmed the associations of several genes, mostly in populations with European ancestry. In 2023, a multiancestry study of PAU with more than 1 million participants revealed numerous novel findings (80). 85 independent risk variants were identified in participants with European ancestry, and 110 risk variants in total were identified in either within-ancestry or cross-ancestry meta-analysis. Cross-ancestry fine-mapping identified credible sets in 13 loci (a set of plausible causal variants within each locus — these sets of putative causal variants are called “credible sets”) containing a single variant. There were 34 additional credible sets containing 2–5 variants. Taken together, these results provided a list of target variants for future experimental functional studies. Leveraging information from multiple ancestries, the cross-ancestry PRS association findings were greater than those using single-ancestry PRS (81).

This study examined overlapped genes by both gene-based association analysis and TWAS (82) in brain tissues and/or chromatin interaction analysis (83) using Hi-C brain annotations. Many genes showed convergent evidence linking association to PAU with brain biology through gene expression (TWAS) and

chromatin interaction (Hi-C) analyses. Translating genetic results into clinical applications is an important goal of human genetic studies, and previous studies have demonstrated the possibilities (84–86). Through two types of drug-repurposing analyses, this study identified existing medications as potential treatments for AUD. The first analysis searched the independent genetic signals in Open Targets (87) for druggability and medication target status. Many genes were druggable, including *DRD2*, *CACNA1C*, *DPYD*, *PDE4B*, *KLB*, *BRD3*, *NCAM1*, *FTO*, *MAPT*, *OPRM1*, and *GABRA4*. The second drug repurposing analysis, using TWAS results, found that 287 compounds were significantly correlated with the transcriptional pattern associated with risk for AUD. These compounds include trichostatin-a, melperone, triflupromazine, spironolactone, amlodipine, and clomethiazole. Trichostatin-a has effects on preventing the development of alcohol withdrawal-related anxiety in rats (88), clomethiazole is used to treat alcohol withdrawal syndrome (89), and spironolactone reduces alcohol use in both rats and humans with convergent evidence (90). This study provided a list of potential medications and targets for future pharmacological studies for AUD.

Limitations of AUD GWAS and future directions

While the field of AUD genetics has made considerable progress, substantial gaps persist (similar to other psychiatric disorders, ref. 91). Here, we highlight some limitations of the current AUD studies (Table 2) with the hope that gaps may be filled with new data sets, technologies, analytic methods, and research directions in the future.

(a) Different definitions of AUD and proxy phenotypes (e.g., AUDIT-P) have shared genetic architecture, resulting in improved power in gene discovery when they are combined from different cohorts (78, 80). However, they are not identical traits. Deep phenotyping (either using same definition or focusing on subphenotypes) in larger cohorts could reduce the phenotypic heterogeneity and increase the possibility of identifying trait-specific associations and pathways (92).

(b) Some studies have endeavored to include samples in multiple ancestries (55, 56, 63, 64, 75, 80), but the sample sizes in the non-European ancestries are smaller than sample sizes in the European ancestries — a common issue in human genetic studies (53, 54). Recruitment of individuals of diverse genetic ancestries is a critical next step in this field. With more multiethnic biobanks becoming available, including MVP, the Global Biobank Meta-analysis Initiative (93), and the All of Us Research Program (94), we anticipate that the gap in diversity will diminish. Funding agencies should also direct attention to studies that propose recruitment focused on non-European ancestry participants.

(c) AUD is a highly polygenic disorder, with hundreds of variants at least contributing to the risk (80, 95). The “brute force” GWAS approach requires a larger sample size to identify more risk variants. Unlike other traits or behaviors that can be measured directly and assessed in large populations or biobanks — for example, GWAS of height (96), educational attainment (97), and alcohol consumption (98) have been conducted in 3–5 million participants — clinical diagnosis of AUD in large cohorts is still lagging. Similar to point (a), increasing sample size and incorporating multiple ancestries could improve the power and resolution of causal vari-

Table 2. Challenges of genetic studies of AUD

Challenges	Explanation	Solutions
1. Deep phenotyping	Heterogeneity due to different definitions	Use same definition to reduce heterogeneity, and focus on subphenotypes to identify trait-specific variants
2. Genetic diversity	Most studied samples are of European ancestry	Collect individuals of diverse genetic ancestries, e.g., through global biobanks or targeted recruitment of non-European populations
3. Polygenicity	Hundreds of thousands of variants contributed to the risk with small effects	Increase sample size; integrate data from biobanks
4. Incomplete genome	Lack of information on rare variants and structural variants	Whole-genome sequencing could extend the gene discovery to the entire genome
5. Functions	Most identified variants are in noncoding regions with unknown functions	Deep-learning methods to predict variant functions; validate the functions using experimental essays
6. Pathways	Biological pathways are not known	Utilize novel data sets (epigenetics, multi-omics) and technologies (single-cell sequencing, spatial transcriptomics) to investigate the mechanisms
7. Confounding	Indirect genetic effects and confounding effects may bias the results	Dissect the indirect genetic effects from direct genetic effects and take confounding factors into account
8. Prediction	Low performance of risk prediction based on PRS	Leverage big data and AI technology, perhaps in the near future

PRS, polygenic risk score; AI, artificial intelligence.

ant fine-mapping (80). Besides the well-known functional coding variants in the alcohol metabolic genes, most variants identified through large GWAS have small to very small effects on the risk of AUD, reducing the yield of the extensive effort of following functional studies on individual variants. This is a common issue in the genetic study of complex traits.

(d) Current GWAS studies have mostly used SNP arrays and post hoc imputation to fill in common variants, which does not allow analysis of the full genome because some parts of the genome are not fully “covered” — i.e., there are unassessed variants in some genomic regions that cannot be tested for association, for technical reasons. A typical SNP array can capture from 600,000 (for example, Illumina PsychArray) to 1.8 million (for example, Illumina Multi-Ethnic Genotyping Array) variants. After imputation and application of standard quality controls for the variants, typical analyzable numbers of high-quality variants vary from 5 to 15 million, depending on the original array SNP density, sample size, and genetic ancestry (from a population genetics point of view, African populations have more common variants than other populations due to their evolutionary history). Given the inherent missing information from different steps, GWAS meta-analyses can only cover a subset of variants of the whole genome, indicating that much of the genome is missing in the current genetic studies of AUD. Whole-genome sequencing (WGS), which can detect essentially all variants (including rare variants and structural variants) without ascertainment bias, could provide better opportunities to investigate the full genetic architecture of the trait.

Several whole-exome sequencing (WES) studies and one WGS study of AUD have been conducted recently (99–102). A phenotype-wide WES study of 170,979 individuals (6,320 cases) from the UK Biobank identified two common variants in the *ADH1C* gene associated with AUD ($P < 2 \times 10^{-9}$), using either an additive

or a dominant model (102). A WES study combining 469,835 individuals from the UK Biobank data (13,121 cases) and 3,789 individuals from the Yale-Penn cohort (2,562 cases) with multiple ancestries identified the well-known functional variant *ADH1B**rs1229984 and several common variants in *ADH1C*. Gene-based tests accounting for the burden from loss-of-function, missense, and synonymous variants identified novel genes *CNST* and *IFIT5* (101). A low-coverage WGS study of AUD-related life events and two affective symptoms in 742 American Indians and 1,711 European Americans identified both common and rare novel variants (103).

(e) Most variants identified by GWAS are in noncoding regions with unknown functions (104). The top associated variants in each risk locus are not necessarily the causal variants for AUD. Although post-GWAS fine-mapping analysis could identify a

credible set of potential causal variants (105–107), further efforts are needed to interpret and validate the variants’ functions. In recent years, novel analytic approaches like deep learning (a subset of machine learning) have been successfully implemented in biomedical research. For example, deep-learning methods contribute to prediction of protein structure (108, 109), pathogenic missense variants (110, 111), and regulatory functions of genome variations (112–115). Combining novel computational tools and cutting-edge functional essays like genome editing (116–118) could help assess the variants’ effects at scale.

(f) Although hundreds of risk variants have been identified and many have been repeatedly replicated in GWAS, indirect genetic effects (also called “genetic nurture”), which are effects of alleles in parents on offspring through the environment (119), have not been distinguished from direct genetic effects on AUD. Methods have been developed to impute parental genotypes using family data (120), which could be used to improve estimates of direct genetic effects for AUD. Confounding effects, including socioeconomic status, may also bias the results. For example, educational attainment influences many psychiatric and nonpsychiatric traits (97) and has a genetic correlation $r_g = -0.21$ with AUD, which needs to be considered in future studies.

(g) Another profound gap is that the current predictive performance of PRS for AUD based on GWAS common variants — i.e., using genetic variation to predict risk in genotyped individuals — is strongly statistically significant but numerically still weak and has not yet entered the range of clinical utility. Despite the increase in sample size, the SNP-based heritability (h^2) by GWAS is low (h^2 ranges from 5.6% to 12.7% with liability-scale h^2 ranging from 8.9% to 16.2%, refs. 64, 72, 75, 78, 80) compared with the total heritability but comparable to what is observed for many other genetically complex traits. PRS presently has limited power for AUD predic-

tion (explained variance measured by pseudo R^2) in independent cohorts; thus, the clinical use of the current PRS of AUD is not imminent. Possibly, the success of artificial intelligence in other areas could extend into predicting AUD risk, with more genomic and large-scale electronic health records data available by integrating improving genomic data with other trait predictors.

(h) Finally, genetic studies have confirmed that AUD is partly a brain-related disorder (75). Genes with expression perturbation in specific brain tissues have been prioritized (72, 80), but the biological pathways from genetics to the etiology of AUD are largely unclear. There are major exceptions though: the mechanism of the effect on risk of alcohol-metabolizing enzyme variation is well understood. Many biological processes play roles in the pathways, such as gene expression, functional regulation, protein perturbation, metabolites, and other mediating traits. To understand the pathway mechanisms, studies beyond genetics are warranted, including, but not limited to epigenetics (discussed in *Epigenetics of AUD*), multiomics, single-cell sequencing, and the latest spatial transcriptomics.

Epigenetics of AUD

Epigenetic studies of AUD have emerged as an important avenue for understanding the complex interplay among genetics, environment, and gene regulation in the development and progression of AUD. Epigenetic factors include transcription factors, noncoding RNAs, DNA modifications, or histone modifications that alter the gene expression and consequently affect phenotypes, without changing the DNA sequence (121, 122). While epigenetic status is highly heritable and affected by environmental factors, including alcohol exposures, certain epigenetic changes in specific brain regions have been implicated in the etiology of AUD (123).

Although most epigenetic studies in humans have focused on alcohol consumption (which is not the main focus of this Review; we focus here on use *disorder* rather than use), some studies have explored DNA methylation patterns in individuals with AUD and identified differential methylation in specific genomic regions (reviewed in refs. 124–128). These changes are often observed in genes related to neurobiological processes, neurotransmitter systems, and immune responses. For example, significantly greater DNA methylation in the *HERP* promoter was reported in patients with AUD than controls (129), while a higher level of DNA methylation in the promoter region of the *OPRM1* gene was observed in AUD (130). A study of postmortem human brains found an overall decrease in methylation in the long-terminal repeat retrotransposons in the frontal cortex (131). Notably, these brain samples, along with those used in several follow-up epigenetic studies of AUD, were mostly from the New South Wales Tissue Resource Centre (NSW TRC) at the University of Sydney (132). However, no global methylation differences were observed between AUD cases and controls in the frontal cortex (133). DNA hypermethylation was also reported in other genes, including *SNCA* (134), *MAOA* (135), *DAT* (136), *NGF* (137), *AVP* (138), *PDNY* (139), and *GABRD* (140). In a study of 285 African Americans and 249 European Americans using a custom-designed methylation array of 384 CpGs in 82 candidate genes, a significant CpG site was identified in the *HTR3A* (5-hydroxytryptamine receptor 3A) promoter region in European Americans. Several other suggestive CpGs were

also reported in either African Americans (in genes *GABRB3* and *POMC*) or European Americans (in genes *NCAMI*, *DRD4*, *MBD3*, *HTR2B*, and *GRINI*) (141). As for studies of genetic variation, it is unclear whether “candidate gene” results will be proven stable over time in epigenetic studies.

Epigenome-wide association studies of AUD

Most DNA methylation studies in AUD have focused on individual gene regions and did not produce replicable results; epigenome-wide scans are needed to identify AUD-related epigenetic changes at scale. Similar to the waves of technologies in genomic studies, microarrays and next-generation sequencing techniques have been applied to epigenetic studies of AUD (Table 3).

The first epigenome-wide association study (EWAS) on AUD involved peripheral blood samples from 10 AUD cases and 10 controls of East Asian ancestry. In total, 865 hypomethylated and 716 hypermethylated CpG sites were identified (defined as expression difference score ≥ 20) (142). The second study involved a cohort of 128 East Asian males, with 63 individuals diagnosed with AUD. In this study, significantly lower levels of methylation were observed in cases compared with controls, with 1,702 hypomethylated and 8 hypermethylated sites reaching $FDR P < 0.005$ (143). A study of 33 patients from alcohol treatment centers and 33 individuals acting as healthy controls (abstinent from alcohol for six months) — mostly European males — identified 56 differentially methylated CpG sites. None of these sites remained significant after the 30-day inpatient treatment program (144). A longitudinal study of East Asian samples identified 149 hypermethylated and 51 hypomethylated genes ($P < 0.01$) between healthy (1990–1992) and dependent phases (2003–2009) in 10 individuals (145). Given the changes in methylation status with the change of alcohol use status in the cases, further studies of AUD with a more robust experiment design that eliminates the ongoing effects of alcohol use are warranted. Similar patterns were observed in a larger longitudinal study of 99 in-patient AUD cases and 95 matched individuals acting as controls. Blood samples were collected in two phases from the affected participants, one during acute alcohol withdrawal and the other after two weeks of recovery in the treatment centers. Compared with the controls, 9,845 CpGs were identified during alcohol withdrawal and 6,094 after two weeks. Comparing the two phases within cases revealed 2,876 differentially methylated CpG sites, suggesting reversibility of alcohol- and withdrawal-related methylation (146). EWAS in 18 discordant monozygotic twin pairs, i.e., one affected and one not in each pair, identified 77 differentially methylated regions at $FDR < 0.05$ (147). A larger EWAS of AUD in 539 blood samples reported 5,101 significant differentially methylated CpG sites after FDR correction. Of these, 96 CpG sites were replicated in a second cohort of 43 AUD cases and 43 controls (148).

The studies mentioned above were conducted primarily on blood samples, with later investigations focusing on postmortem brain samples from NSW TRC to identify DNA methylation changes associated with AUD. A study specifically analyzed prefrontal cortex samples from 23 AUD cases and 23 controls, with a particular emphasis on sex-stratified analyses due to a previous observation of a sex-biased methylome (149). Among 32 males (16 AUD cases), 1,201 hypermethylated and 611 hypomethylated CpG sites

Table 3. Epigenome-wide association studies of AUD

Study	Tissue	N	Array	Finding	Note
Zhao, 2013 (142)	Blood	EAS, <i>n</i> = 20; Ncase = 10	450,000	865 hypomethylated; 716 hypermethylated CpG sites	Significant sites were defined as expression difference scores ≥ 20
Zhang, 2013 (141)	Blood	EAS, <i>n</i> = 128; Ncase = 63	27,000	1,702 hypomethylated and 8 hypermethylated	FDR <i>P</i> < 0.005
Philibert, 2014 (144)	Blood	EUR ^a , <i>n</i> = 66; Ncase = 33	450,000	56 differentially methylated CpG sites	All 56 sites are not significant with alcohol abstinence
Weng, 2015 (145)	Blood	EAS, <i>n</i> = 10; Ncase = 10 ^b	27,000	Not clear	All males
Ruggeri, 2015 (147)	Blood	EUR, <i>n</i> = 36; Ncase = 18	385,000	77 CpGs, 68% hypermethylated	Monozygotic twin pairs
Wang, 2016 (150)	PFC (BA9)	EUR, <i>n</i> = 46; Ncase = 23	450,000	In males, 1,201 hypermethylated and 611 hypomethylated	Only perform sex-stratified analyses
Hagerty, 2016 (152)	Precuneus	EUR, <i>n</i> = 96; Ncase = 49	450,000	561 hypomethylated; 485 hypermethylated	
Lohoff, 2018 (151)	PFC (BA9)	EUR ^c , <i>n</i> = 46; Ncase = 23	450,000	1 hypermethylated and 2 hypomethylated	
Gatta, 2021 (153)	PFC (BA10)	EUR, <i>n</i> = 46; Ncase = 23	850,000	Not clear	
Witt, 2020 (146)	Blood	EUR, <i>n</i> = 194; Ncase = 99	850,000	Alcohol withdrawal: 9,845 CpGs; recovery: 6,094 CpGs	FDR <i>P</i> < 0.05; All are males
Lohoff, 2021 (148)	Blood	Mixed ^d , <i>n</i> = 539; Ncase = 336	850,000	5,101 CpG sites	FDR <i>P</i> < 0.05
Meng, 2021 (154)	PFC (BA9), NAc	EUR, <i>n</i> = 86; Ncase = 39	450,000	Two differentially methylated regions; no single CpG sites	All are males; permutation <i>P</i> < 0.05
Clark, 2022 (157)	4 cell types in blood; PFC (BA10)	EUR, <i>n</i> = 1,132; Ncase = 323 EUR, <i>n</i> = 50; Ncase = 25	MBD-Seq MBD-seq	3 CpGs T cells; 1,397 in monocytes No single CpG sites	FDR <i>P</i> < 0.1
Zillich, 2022 (155)	ACC, PFC (BA9), CN, VS, PUT	EUR, <i>n</i> = 111; Ncase = 53	850,000	2 CpGs in CN; 18 in VS	FDR <i>P</i> < 0.05
White, 2024 (156)	PFC (BA9), NAc	EUR, <i>n</i> = 119; Ncase	850,000	53 CpG in NAc; 31 in DLPFC; 105 in meta-analysis	FDR <i>P</i> < 0.05

PFC, prefrontal cortex; NAc, nucleus accumbens; ACC, anterior cingulate cortex; CN, caudate nucleus; VS, ventral striatum; PUT, putamen; MBD-Seq, methyl-CG binding domain sequencing. ^aMost are of European ancestry, and four are of other ancestries. ^bOnly compared different phases within cases. ^cSame samples were used in Wang et al., 2016. ^d289 European Americans and 249 African Americans.

were discovered at an FDR level of 0.05; however, no significant results were observed in females (150). Reanalysis of the same data, combining both sexes, identified three CpG sites after FDR-based correction (151). EWAS of 49 AUD cases and 47 controls using the 450,000 methylation array identified 561 hypomethylated CpGs and 485 hypermethylated CpGs were reported with $P < 1 \times 10^{-7}$ (152). A small EWAS on brain samples from 23 AUD cases and 23 controls, using a higher density methylation array (850,000 CpG sites) found no GWS loci (see Figure 1C in ref. 153), although there were 1,218 CpGs with $P_{\text{nominal}} \leq 0.001$ (153).

Several studies investigated multiple brain regions simultaneously. For instance, a study of two brain regions, including the prefrontal cortex and nucleus accumbens in 86 individuals, did not identify single CpG sites but identified two differentially methylated regions (permutation $P < 0.05$) mapping to the upstream regions of *ZFP57* and *DLGAP2* genes (154). An EWAS on five brain regions from 111 individuals identified two differentially methylated CpG sites in the caudate nucleus region and 18 in the ventral striatum, with no significant findings in the other three regions (155). Despite utilizing brain samples from the same repository (NSW TRC), the reported findings across these studies were largely inconsistent.

In an additional study on AUD, brain samples from 119 individuals from the Lieber Institute for Brain Development Human Brain Repository were analyzed. This investigation identified 53 CpGs associated with AUD in the nucleus accumbens and 31

in the dorsolateral prefrontal cortex (BA9) at FDR $P < 0.05$, with no overlap across the two regions. The investigators conducted a meta-analysis across the two regions, revealing an additional 21 CpGs, bringing the total to 105 unique AUD-associated CpGs in 120 genes (156). Comparing the results of this study with previous research on brain regions (152, 155, 157), only one intergenic CpG, cg00402668 in the BA9, reached a “look-up” level of significance (FDR $P < 0.05$) in BA10 from Clark et al. (157), but not in other brain regions. On the gene level, three annotated genes in this study overlapped with genes from Hagerty et al. (152). When testing for overlap of the top 1% of CpG sites across studies, a significant enrichment for the nucleus accumbens results in this study was observed in the putamen and ventral striatum results from Zillich et al. (155).

Cross-tissue studies have advanced the understanding of overlapping epigenetic mechanisms. In a study of 1,132 blood samples consisting of four cell types and 50 brain samples, over 21 million CpG sites were assessed using methyl-CG binding domain sequencing. No significant associations were observed for the whole blood or brain. However, for T cells and monocytes, 3 CpGs and 1,397 CpGs were identified at FDR $P < 0.1$, respectively. One CpG site in the *DLGAP1* gene was significantly replicated, and an additional 34 sites were nominally replicated in an independent sample of 73 AUD cases and 339 controls. Beyond DNA methylation, this study also assessed hydroxymethylation for over

26 million CpG sites in the brain samples. While no individual site reached methylome-wide significance, the study observed significant overlap between the top sites in blood cell type-specific EWAS and both methylation and hydroxymethylation EWAS in the brain. This represents the first exploration of cell type-specific methylation for AUD in blood and considers the role of brain hydroxymethylation in AUD (157).

Other epigenetic studies of AUD

Besides studies of DNA methylation for AUD, limited studies of noncoding RNAs and histone modifications in the human brain have been conducted. microRNA (miRNAs) are small noncoding RNAs that regulate target mRNA expression and/or translation, with important roles in a variety of biological processes (158). A study of miRNAs in the prefrontal cortex of 27 individuals identified 12 upregulated miRNAs (FDR $P < 0.05$) in AUD cases ($n = 14$) compared with controls ($n = 13$), suggesting a regulatory role of miRNAs in AUD gene expression (159). A candidate gene study of long noncoding RNA *BDNF-AS* in the human amygdala implicated a regulatory effect on *BDNF* expression in early-onset (before age 21) AUD cases ($n = 11$) compared with controls ($n = 22$) but not late-onset AUD (160).

A study of the transcriptome using RNA-Seq and histone H3 lysine 4 trimethylation (H3K4me3) using ChIP-Seq in postmortem brain hippocampus samples from the University of Miami Brain Bank identified 11 differentially expressed genes with FDR $P < 0.05$ in AUD cases ($n = 8$) compared with controls ($n = 8$). However, no H3K4me3 changes reached FDR $P < 0.05$ or overlapped with expression changes (161). Reanalyzing the data using network approaches identified 7 coexpression modules enriched for H3K4me3-associated changes in AUD cases compared with controls, suggesting relationships between this epigenetic mark and gene expression (162). The small sample size and low power indicate that these results should be taken with caution.

Another study investigated gene coexpression and its relationship with multiple epigenetic modifications for AUD in brain tissues from the NSW TRC. Central and basolateral nucleus of amygdala and superior frontal cortex from 17 AUD cases and 15 controls were assessed for gene expression, cortex from 6 AUD cases and 6 controls were assessed for global H3K4 methylation and DNA methylation of long-terminal repeat retrotransposons, and cortex from 5 AUD cases and 5 controls were assessed for histone H3K4me3 (ChIP-Seq). This study identified critical cellular components and previously unrecognized epigenetic determinants of gene coexpression relationships and discovered novel markers of chromatin modifications in the human brain (131).

In conclusion, epigenetic studies have provided limited insights into the molecular mechanisms underlying AUD. Considering the known genetic and etiologic complexity of AUD risk, and the contributions of both genes and environment, larger samples will be required to draw durable conclusions about AUD epigenetics. The integration of DNA methylation, histone modifications, and noncoding RNAs into our understanding of AUD pathogenesis holds promise for identifying novel therapeutic targets and developing personalized interventions. As technology advances and research methodologies are refined, the field of epigenetics is expected to profoundly contribute to unraveling the

complexities of AUD. However, the field is presently limited greatly by a lack of well-powered investigations.

Limitations of epigenetic studies in AUD

To date, there has been very limited progress in unraveling the epigenetic landscape of AUD. The heterogeneity of AUD, coupled with varying degrees of alcohol exposure in different groups and stages of the disorder, introduce complexities in interpreting findings, especially in small samples. Additionally, the dynamic nature of epigenetic modifications requires sophisticated study designs to capture temporal changes throughout AUD development. Comparatively, genetic studies of AUD, mainly through GWAS, have had much larger sample sizes (many orders of magnitude), while current epigenetic studies, particularly in the human brain, are small, with the largest study to date involving 119 individuals (156). Finally, the overlapping findings across tissues and studies are minimal (126, 156). While this could be attributed to tissue/context-specific epigenetic changes, it also raises the possibility of false-positive results.

As for GWAS of AUD, larger-scale epigenetic studies will be required to generate replicable findings and prioritize robust genomic variations for future pharmacological studies. Other future directions that have the potential to improve our understanding of the epigenetic mechanisms of AUD include the following. First, longitudinal studies tracking the epigenetic changes over time (163) could help us understand the temporal relationship between alcohol consumption and epigenetic modifications and work out the cause and effect on AUD. Second, integrating epigenetic data with other omics data (e.g., transcriptomics) should help us gain a better understanding of relevant regulatory mechanisms. Third, it is important to investigate how environmental factors (e.g., stress and diet) interact with epigenetic factors to influence the risk of AUD. Fourth, the current EWASs are focused on tissue-level changes; performing cell type-specific epigenetic analyses (e.g., single-cell epigenome data, ref. 164) can provide deeper insight into the molecular mechanisms.

Conclusions

Genetic studies of AUD have greatly advanced our understanding of its complex etiology, while epigenetic studies have made limited progress. Although these studies have provided valuable insights, challenges and gaps in our comprehension persist, emphasizing the need for continued research and exploration incorporating larger samples with deeper phenotyping in more diverse populations. Integration of comprehensive insights from both genetic and epigenetic studies holds promise for the development of targeted and personalized therapeutic strategies, representing a crucial step forward in addressing the multifaceted nature of AUD. Future research should aim to increase statistical power, expand the study populations to encompass diverse groups, and thus refine our understanding of the mechanisms involved. Overcoming limitations and translating research findings into effective clinical interventions for AUD should be at the forefront of ongoing efforts.

Acknowledgments

The authors are supported by grants from the NIH (R01AA026364, R01DA037974, P50AA012870, R21CA252916,

U54AA027989, and RM1HG011558) and the Department of Veterans Affairs (1I01CX001849). HZ was also supported by a NARSAD Young Investigator grant (no. 27835) from the Brain & Behavior Research Foundation.

Address correspondence to: Hang Zhou, Department of Psychiatry, Yale School of Medicine; Veterans Affairs Connecticut Healthcare System Annex Center, Room 1325, 200 Edison Rd, Orange, Connecticut 06477, USA. Email: hang.zhou@yale.edu.

- Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology*. 2010;35(1):217–238.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. American Psychiatric Association; 1994.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. American Psychiatric Association; 2013.
- Kranzler HR, Soyka M. Diagnosis and pharmacotherapy of alcohol use disorder: A Review. *JAMA*. 2018;320(8):815–824.
- GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2018;392(10152):1015–1035.
- Smith M, et al. Developmental changes and polymorphism in human alcohol dehydrogenase. *Ann Hum Genet*. 1971;34(3):251–271.
- Bosron WF, Li TK. Genetic determinants of alcohol and aldehyde dehydrogenases and alcohol metabolism. *Semin Liver Dis*. 1981;1(3):179–188.
- Pickens RW, et al. Heterogeneity in the inheritance of alcoholism: A study of male and female twins. *Arch Gen Psychiatry*. 1991;48(1):19–28.
- Agarwal DP, et al. Racial differences in biological sensitivity to ethanol: the role of alcohol dehydrogenase and aldehyde dehydrogenase isozymes. *Alcohol Clin Exp Res*. 1981;5(1):12–16.
- Kendler KS, et al. A twin-family study of alcoholism in women. *Am J Psychiatry*. 1994;151(5):707–715.
- Reed T, et al. Genetic predisposition to organ-specific endpoints of alcoholism. *Alcohol Clin Exp Res*. 1996;20(9):1528–1533.
- Heath AC, et al. Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychol Med*. 1997;27(6):1381–1396.
- Verhulst B, et al. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med*. 2015;45(5):1061–1072.
- Dick DM, Bierut LJ. The genetics of alcohol dependence. *Curr Psychiatry Rep*. 2006;8(2):151–157.
- Gelernter J, Kranzler HR. Genetics of alcohol dependence. *Hum Genet*. 2009;126(1):91–99.
- Edenberg HJ, Foroud T. Genetics and alcoholism. *Nat Rev Gastroenterol Hepatol*. 2013;10(8):487–494.
- Tawa EA, et al. Overview of the genetics of alcohol use disorder. *Alcohol Alcohol*. 2016;51(5):507–514.
- Deak JD, et al. Genetics of alcohol use disorder: a review. *Curr Opin Psychol*. 2019;27:56–61.
- Thomasson HR, et al. Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. *Am J Hum Genet*. 1991;48(4):677–681.
- Borras E, et al. Genetic polymorphism of alcohol dehydrogenase in Europeans: the ADH2*2 allele decreases the risk for alcoholism and is associated with ADH3*1. *Hepatology*. 2000;31(4):984–989.
- Chambers GK, et al. The genetics of alcoholism in Polynesians: alcohol and aldehyde dehydrogenase genotypes in young men. *Alcohol Clin Exp Res*. 2002;26(7):949–955.
- Whitfield JB. Alcohol dehydrogenase and alcohol dependence: variation in genotype-associated risk between populations. *Am J Hum Genet*. 2002;71(5):1247–1250.
- Edenberg HJ. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res Health*. 2007;30(1):5–13.
- Peng Q, et al. Association and ancestry analysis of sequence variants in ADH and ALDH using alcohol-related phenotypes in a Native American community sample. *Am J Med Genet B Neuropsychiatr Genet*. 2014;165B(8):673–683.
- Peng Q, et al. Associations between genomic variants in alcohol dehydrogenase genes and alcohol symptomatology in American Indians and European Americans: distinctions and convergence. *Alcohol Clin Exp Res*. 2017;41(10):1695–1704.
- Mizoi Y, et al. Relationship between facial flushing and blood acetaldehyde levels after alcohol intake. *Pharmacol Biochem Behav*. 1979;10(2):303–311.
- Crabb DW, et al. Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant. *J Clin Invest*. 1989;83(1):314–316.
- Harada S, et al. Possible protective role against alcoholism for aldehyde dehydrogenase isozyme deficiency in Japan. *Lancet*. 1982;2(8302):827.
- Goedde HW, et al. Population genetic studies on aldehyde dehydrogenase isozyme deficiency and alcohol sensitivity. *Am J Hum Genet*. 1983;35(4):769–772.
- Li D, et al. Strong protective effect of the aldehyde dehydrogenase gene (ALDH2) 504lys (*2) allele against alcoholism and alcohol-induced medical diseases in Asians. *Hum Genet*. 2012;131(5):725–737.
- Bierut LJ, et al. ADH1B is associated with alcohol dependence and alcohol consumption in populations of European and African ancestry. *Mol Psychiatry*. 2012;17(4):445–450.
- Edenberg HJ, et al. Association of alcohol dehydrogenase genes with alcohol dependence: a comprehensive analysis. *Hum Mol Genet*. 2006;15(9):1539–1549.
- Ehlers CL, et al. Association of the ADH2*3 allele with a negative family history of alcoholism in African American young adults. *Alcohol Clin Exp Res*. 2001;25(12):1773–1777.
- Long JC, et al. Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from an autosome-wide scan in an American Indian population. *Am J Med Genet*. 1998;81(3):216–221.
- Reich T, et al. Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet*. 1998;81(3):207–215.
- Covault J, et al. Allelic and haplotypic association of GABRA2 with alcohol dependence. *Am J Med Genet B Neuropsychiatr Genet*. 2004;129B(1):104–109.
- Edenberg HJ, et al. Variations in GABRA2, encoding the alpha 2 subunit of the GABA(A) receptor, are associated with alcohol dependence and with brain oscillations. *Am J Hum Genet*. 2004;74(4):705–714.
- Enoch MA, et al. GABRG1 and GABRA2 as independent predictors for alcoholism in two populations. *Neuropsychopharmacology*. 2009;34(5):1245–1254.
- Ittiwut C, et al. GABRG1 and GABRA2 variation associated with alcohol dependence in African Americans. *Alcohol Clin Exp Res*. 2012;36(4):588–593.
- Wang JC, et al. Evidence of common and specific genetic effects: association of the muscarinic acetylcholine receptor M2 (CHRM2) gene with alcohol dependence and major depressive syndrome. *Hum Mol Genet*. 2004;13(17):1903–1911.
- Luo X, et al. CHRM2 gene predisposes to alcohol dependence, drug dependence and affective disorders: results from an extended case-control structured association study. *Hum Mol Genet*. 2005;14(16):2421–2434.
- Klein RJ, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308(5720):385–389.
- McCarthy MI, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9(5):356–369.
- Treutlein J, et al. Genome-wide association study of alcohol dependence. *Arch Gen Psychiatry*. 2009;66(7):773–784.
- Frank J, et al. Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. *Addict Biol*. 2012;17(1):171–180.
- International HapMap C. The international HapMap Project. *Nature*. 2003;426(6968):789–796.
- Lind PA, et al. A genomewide association study of nicotine and alcohol dependence in Australian and Dutch populations. *Twin Res Hum Genet*. 2010;13(1):10–29.
- Heath AC, et al. A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry*. 2011;70(6):513–518.
- Wang KS, et al. A meta-analysis of two genome-wide association studies identifies 3 new loci for alcohol dependence. *J Psychiatr Res*. 2011;45(11):1419–1425.
- Zuo L, et al. Genome-wide association study of alcohol dependence implicates KIAA0040 on chromosome 1q. *Neuropsychopharmacology*. 2012;37(2):557–566.
- Wang JC, et al. A genome-wide association study of alcohol-dependence symptom counts in extended pedigrees identifies C15orf53. *Mol*

- Psychiatry*. 2013;18(11):1218–1224.
52. Wetherill L, et al. Family-based association analysis of alcohol dependence criteria and severity. *Alcohol Clin Exp Res*. 2014;38(2):354–366.
 53. Peterson RE, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell*. 2019;179(3):589–603.
 54. Sirugo G, et al. The missing diversity in human genetic studies. *Cell*. 2019;177(1):26–31.
 55. Bierut LJ, et al. A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci U S A*. 2010;107(11):5082–5087.
 56. Edenberg HJ, et al. Genome-wide association study of alcohol dependence implicates a region on chromosome 11. *Alcohol Clin Exp Res*. 2010;34(5):840–852.
 57. Norden-Krichmar TM, et al. Protective variant associated with alcohol dependence in a Mexican American cohort. *BMC Med Genet*. 2014;15:136.
 58. Park BL, et al. Extended genetic effects of ADH cluster genes on the risk of alcohol dependence: from GWAS to replication. *Hum Genet*. 2013;132(6):657–668.
 59. Quillen EE, et al. ALDH2 is associated to alcohol dependence and is the major genetic determinant of “daily maximum drinks” in a GWAS study of an isolated rural Chinese sample. *Am J Med Genet B Neuropsychiatr Genet*. 2014;165B(2):103–110.
 60. Gelernter J, et al. Genomewide association study of alcohol dependence and related traits in a Thai population. *Alcohol Clin Exp Res*. 2018;42(5):861–868.
 61. Sun Y, et al. Genome-wide association study of alcohol dependence in male Han Chinese and cross-ethnic polygenic risk score comparison. *Transl Psychiatry*. 2019;9(1):249.
 62. Zhou H, et al. Genome-wide meta-analysis of alcohol use disorder in East Asians. *Neuropsychopharmacology*. 2022;47(10):1791–1797.
 63. Gelernter J, et al. Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry*. 2014;19(1):41–49.
 64. Walters RK, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018;21(12):1656–1669.
 65. Martin AR, et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet*. 2017;100(4):635–649.
 66. Saunders JB, et al. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption—II. *Addiction*. 1993;88(6):791–804.
 67. Allen JP, et al. A review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol Clin Exp Res*. 1997;21(4):613–619.
 68. Boschloo L, et al. The performance of the Alcohol Use Disorder Identification Test (AUDIT) in detecting alcohol abuse and dependence in a population of depressed or anxious persons. *J Affect Disord*. 2010;126(3):441–446.
 69. Sanchez-Roige S, et al. Genome-wide association study of alcohol use disorder identification test (AUDIT) scores in 20 328 research participants of European ancestry. *Addict Biol*. 2019;24(1):121–131.
 70. Mbarek H, et al. The genetics of alcohol dependence: Twin and SNP-based heritability, and genome-wide association study based on AUDIT scores. *Am J Med Genet B Neuropsychiatr Genet*. 2015;168(8):739–748.
 71. Bycroft C, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–209.
 72. Sanchez-Roige S, et al. Genome-wide association study meta-analysis of the Alcohol Use Disorders Identification Test (AUDIT) in two population-based cohorts. *Am J Psychiatry*. 2019;176(2):107–118.
 73. Barbeira AN, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun*. 2018;9(1):1825.
 74. Gaziano JM, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol*. 2016;70:214–223.
 75. Kranzler HR, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat Commun*. 2019;10(1):1499.
 76. Finucane HK, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet*. 2015;47(11):1228–1235.
 77. Kember RL, et al. Genetic underpinnings of the transition from alcohol consumption to alcohol use disorder: shared and unique genetic architectures in a cross-ancestry sample. *Am J Psychiatry*. 2023;180(8):584–593.
 78. Zhou H, 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.
 79. Sanderson E, et al. Mendelian randomization. *Nat Rev Methods Primers*. 2022;2:6.
 80. Zhou H, et al. Multi-ancestry study of the genetics of problematic alcohol use in over 1 million individuals. *Nat Med*. 2023;29(12):3184–3192.
 81. Ruan Y, et al. Improving polygenic prediction in ancestrally diverse populations. *Nat Genet*. 2022;54(5):573–580.
 82. Barbeira AN, et al. Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet*. 2019;15(1):e1007889.
 83. Sey NYA, et al. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat Neurosci*. 2020;23(4):583–593.
 84. Nelson MR, et al. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47(8):856–860.
 85. Backman JD, et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature*. 2021;599(7886):628–634.
 86. King EA, et al. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet*. 2019;15(12):e1008489.
 87. Ochoa D, et al. The next-generation Open Targets Platform: reimaged, redesigned, rebuilt. *Nucleic Acids Res*. 2023;51(d1):D1353–D1359.
 88. Pandey SC, et al. Brain chromatin remodeling: a novel mechanism of alcoholism. *J Neurosci*. 2008;28(14):3729–3737.
 89. Sychla H, et al. Comparison of clomethiazole and diazepam in the treatment of alcohol withdrawal syndrome in clinical practice. *Eur Addict Res*. 2017;23(4):211–218.
 90. Farokhnia M, et al. Spironolactone as a potential new pharmacotherapy for alcohol use disorder: convergent evidence from rodent and human studies. *Mol Psychiatry*. 2022;27(11):4642–4652.
 91. Derks EM, et al. Ten challenges for clinical translation in psychiatric genetics. *Nat Genet*. 2022;54(10):1457–1465.
 92. Cai N, et al. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat Genet*. 2020;52(4):437–447.
 93. Zhou W, et al. Global Biobank Meta-analysis Initiative: Powering genetic discovery across human disease. *Cell Genom*. 2022;2(10):100192.
 94. All of Us Research Program Investigators, et al. The “all of us” research program. *N Engl J Med*. 2019;381(7):668–676.
 95. Gelernter J, Polimanti R. Genetics of substance use disorders in the era of big data. *Nat Rev Genet*. 2021;22(11):712–729.
 96. Yengo L, et al. A saturated map of common genetic variants associated with human height. *Nature*. 2022;610(7933):704–712.
 97. Okbay A, et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet*. 2022;54(4):437–449.
 98. Saunders GRB, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature*. 2022;612(7941):720–724.
 99. Gentry AE, et al. Case-only exome variation analysis of severe alcohol dependence using a multivariate hierarchical gene clustering approach. *PLoS One*. 2023;18(4):e0283985.
 100. Hill SY, Hostyk J. A whole exome sequencing study to identify rare variants in multiplex families with alcohol use disorder. *Front Psychiatry*. 2023;14:1216493.
 101. Wang L, et al. Multi-ancestry whole-exome sequencing study of alcohol use disorder in two cohorts [preprint]. <https://doi.org/10.1101/2024.04.05.24305412>. Posted on medRxiv April 8, 2024.
 102. Wang Q, et al. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature*. 2021;597(7877):527–532.
 103. Peng Q, et al. Genetic loci for alcohol-related life events and substance-induced affective symptoms: indexing the “dark side” of addiction. *Transl Psychiatry*. 2019;9(1):71.
 104. Edwards SL, et al. Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet*. 2013;93(5):779–797.
 105. Huang H, et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature*. 2017;547(7662):173–178.
 106. Yuan K, et al. Fine-mapping across diverse ancestries drives the discovery of putative causal variants underlying human complex traits and diseases [preprint]. <https://doi.org/10.1101/2023.01.07.23284293>. Posted on medRxiv July 9, 2023.

107. Wellcome Trust Case Control Consortium, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat Genet.* 2012;44(12):1294–1301.
108. Jumper J, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–589.
109. Lin Z, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science.* 2023;379(6637):1123–1130.
110. Cheng J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science.* 2023;381(6664):eadg7492.
111. Gao H, et al. The landscape of tolerated genetic variation in humans and primates. *Science.* 2023;380(6648):eabn8153.
112. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat Methods.* 2015;12(10):931–934.
113. Zhou J, et al. Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. *Nat Genet.* 2018;50(8):1171–1179.
114. Jaganathan K, et al. Predicting splicing from primary sequence with deep learning. *Cell.* 2019;176(3):535–548.e24.
115. Chen KM, et al. A sequence-based global map of regulatory activity for deciphering human genetics. *Nat Genet.* 2022;54(7):940–949.
116. Morris JA, et al. Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science.* 2023;380(6646):eadh7699.
117. Gasperini M, et al. A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell.* 2019;176(1-2):377–390.
118. Wunnemann F, et al. Multimodal CRISPR perturbations of GWAS loci associated with coronary artery disease in vascular endothelial cells. *PLoS Genet.* 2023;19(3):e1010680.
119. Trejo S, Domingue BW. Genetic nature or genetic nurture? Introducing social genetic parameters to quantify bias in polygenic score analyses. *Bio-demography Soc Biol.* 2018;64(3-4):187–215.
120. Young AI, et al. Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nat Genet.* 2022;54(6):897–905.
121. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet.* 2012;13(2):97–109.
122. Holliday R. The inheritance of epigenetic defects. *Science.* 1987;238(4824):163–170.
123. Robison AJ, Nestler EJ. Transcriptional and epigenetic mechanisms of addiction. *Nat Rev Neurosci.* 2011;12(11):623–637.
124. Krishnan HR, et al. The epigenetic landscape of alcoholism. *Int Rev Neurobiol.* 2014;115:75–116.
125. Zhang H, Gelernter J. Review: DNA methylation and alcohol use disorders: Progress and challenges. *Am J Addict.* 2017;26(5):502–515.
126. Longley MJ, et al. Epigenetics of alcohol use disorder—A review of recent advances in DNA methylation profiling. *Addict Biol.* 2021;26(6):e13006.
127. Nieratschker V, et al. Genetics and epigenetics of alcohol dependence. *J Mol Psychiatry.* 2013;1(1):11.
128. Berkel TD, Pandey SC. Emerging role of epigenetic mechanisms in alcohol addiction. *Alcohol Clin Exp Res.* 2017;41(4):666–680.
129. Bleich S, et al. Epigenetic DNA hypermethylation of the *HERP* gene promoter induces down-regulation of its mRNA expression in patients with alcohol dependence. *Alcohol Clin Exp Res.* 2006;30(4):587–591.
130. Zhang H, et al. Hypermethylation of *OPRM1* promoter region in European Americans with alcohol dependence. *J Hum Genet.* 2012;57(10):670–675.
131. Ponomarev I, et al. Gene coexpression networks in human brain identify epigenetic modifications in alcohol dependence. *J Neurosci.* 2012;32(5):1884–1897.
132. Harper C, et al. Neuropathological alterations in alcoholic brains. Studies arising from the New South Wales Tissue Resource Centre. *Prog Neuropsychopharmacol Biol Psychiatry.* 2003;27(6):951–961.
133. Manzardo AM, et al. Global DNA promoter methylation in frontal cortex of alcoholics and controls. *Gene.* 2012;498(1):5–12.
134. Bonsch D, et al. DNA hypermethylation of the alpha synuclein promoter in patients with alcoholism. *Neuroreport.* 2005;16(2):167–170.
135. Philibert RA, et al. MAOA methylation is associated with nicotine and alcohol dependence in women. *Am J Med Genet B Neuropsychiatr Genet.* 2008;147B(5):565–570.
136. Hillemacher T, et al. Promoter specific methylation of the dopamine transporter gene is altered in alcohol dependence and associated with craving. *J Psychiatr Res.* 2009;43(4):388–392.
137. Heberlein A, et al. Epigenetic down regulation of nerve growth factor during alcohol withdrawal. *Addict Biol.* 2013;18(3):508–510.
138. Hillemacher T, et al. Epigenetic regulation and gene expression of vasopressin and atrial natriuretic peptide in alcohol withdrawal. *Psychoneuroendocrinology.* 2009;34(4):555–560.
139. Taqi MM, et al. Prodynorphin CpG-SNPs associated with alcohol dependence: elevated methylation in the brain of human alcoholics. *Addict Biol.* 2011;16(3):499–509.
140. Gatta E, et al. Emerging role of one-carbon metabolism and DNA methylation enrichment on δ -containing GABAA receptor expression in the cerebellum of subjects with alcohol use disorders (AUD). *Int J Neuropsychopharmacol.* 2017;20(12):1013–1026.
141. Zhang H, et al. Array-based profiling of DNA methylation changes associated with alcohol dependence. *Alcohol Clin Exp Res.* 2013;37(suppl 1):E108–E115.
142. Zhao R, et al. Genome-wide DNA methylation patterns in discordant sib pairs with alcohol dependence. *Asia Pac Psychiatry.* 2013;5(1):39–50.
143. Zhang R, et al. Genome-wide DNA methylation analysis in alcohol dependence. *Addict Biol.* 2013;18(2):392–403.
144. Philibert RA, et al. A pilot examination of the genome-wide DNA methylation signatures of subjects entering and exiting short-term alcohol dependence treatment programs. *Epigenetics.* 2014;9(9):1212–1219.
145. Weng JT, et al. Integrative epigenetic profiling analysis identifies DNA methylation changes associated with chronic alcohol consumption. *Comput Biol Med.* 2015;64:299–306.
146. Witt SH, et al. Acute alcohol withdrawal and recovery in men lead to profound changes in DNA methylation profiles: a longitudinal clinical study. *Addiction.* 2020;115(11):2034–2044.
147. Ruggeri B, et al. Association of protein phosphatase PPM1G with alcohol use disorder and brain activity during behavioral control in a genome-wide methylation analysis. *Am J Psychiatry.* 2015;172(6):543–552.
148. Lohoff FW, et al. Epigenome-wide association study and multi-tissue replication of individuals with alcohol use disorder: evidence for abnormal glucocorticoid signaling pathway gene regulation. *Mol Psychiatry.* 2021;26(6):2224–2237.
149. Xu H, et al. Sex-biased methylome and transcriptome in human prefrontal cortex. *Hum Mol Genet.* 2014;23(5):1260–1270.
150. Wang F, et al. DNA co-methylation modules in postmortem prefrontal cortex tissues of European Australians with alcohol use disorders. *Sci Rep.* 2016;6:19430.
151. Lohoff FW, et al. Methylomic profiling and replication implicates deregulation of PCSK9 in alcohol use disorder. *Mol Psychiatry.* 2018;23(9):1900–1910.
152. Hagerty SL, et al. An exploratory association study of alcohol use disorder and DNA methylation. *Alcohol Clin Exp Res.* 2016;40(8):1633–1640.
153. Gatta E, et al. Genome-wide methylation in alcohol use disorder subjects: implications for an epigenetic regulation of the cortico-limbic glucocorticoid receptors (NR3C1). *Mol Psychiatry.* 2021;26(3):1029–1041.
154. Meng W, et al. Genotype-dependent epigenetic regulation of DLGAP2 in alcohol use and dependence. *Mol Psychiatry.* 2021;26(8):4367–4382.
155. Zillich L, et al. Epigenome-wide association study of alcohol use disorder in five brain regions. *Neuropsychopharmacology.* 2022;47(4):832–839.
156. White JD, et al. Alcohol use disorder-associated DNA methylation in the nucleus accumbens and dorsolateral prefrontal cortex [preprint]. <https://doi.org/10.1101/2024.01.17.23300238>. Posted on medRxiv January 17, 2024.
157. Clark SL, et al. Dual methylation and hydroxymethylation study of alcohol use disorder. *Addict Biol.* 2022;27(2):e13114.
158. Filipowicz W, et al. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet.* 2008;9(2):102–114.
159. Lewohl JM, et al. Up-regulation of microRNAs in brain of human alcoholics. *Alcohol Clin Exp Res.* 2011;35(11):1928–1937.
160. Bohnsack JP, et al. The lncRNA BDNF-AS is an epigenetic regulator in the human amygdala in early onset alcohol use disorders. *Transl Psychiatry.* 2019;9(1):34.
161. Zhou Z, et al. Substance-specific and shared transcription and epigenetic changes in the human hippocampus chronically exposed to cocaine and alcohol. *Proc Natl Acad Sci U S A.* 2011;108(16):6626–6631.
162. Farris SP, et al. Epigenetic modulation of brain gene networks for cocaine and alcohol abuse. *Front Neurosci.* 2015;9:176.
163. Ng JW, et al. The role of longitudinal cohort studies in epigenetic epidemiology: challenges and opportunities. *Genome Biol.* 2012;13(6):246.

164. Guo H, et al. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* 2013;23(12):2126–2135.
165. 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68–74.
166. Das S, et al. Next-generation genotype imputation service and methods. *Nat Genet.* 2016;48(10):1284–1287.
167. Browning BL, et al. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet.* 2018;103(3):338–348.
168. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet.* 2010;11(7):499–511.
169. Chang CC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7.
170. Chen MH, Yang Q. GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics.* 2010;26(4):580–581.
171. Loh PR, et al. Mixed-model association for biobank-scale datasets. *Nat Genet.* 2018;50(7):906–908.
172. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42(1):121–130.
173. Zhou W, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet.* 2018;50(9):1335–1341.
174. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat Methods.* 2014;11(4):407–409.
175. Mbatchou J, et al. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet.* 2021;53(7):1097–1103.
176. Pe'er I, et al. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol.* 2008;32(4):381–385.