A phenome-wide association and Mendelian randomisation study of alcohol use variants in a diverse cohort comprising over 3 million individuals

Mariela V. Jennings,[a,q] José Jaime Martínez-Magaña,[b,q] Natasia S. Courchesne-Krak,[a,q] Renata B. Cupertino,[a,q] Laura Vilar-Ribó,[c,i,j] Sevim B. Bianchi,[a] Alexander S. Hatoum,[d] Elizabeth G. Atkinson,[e] Paola Giusti-Rodriguez,[g] Janitza L. Montalvo-Ortiz,[b,f] Joel Gelernter,[h,p] María Soler Artigas,[c,i,j,k] 23andMe, Inc. Research Team,[l] Sarah L. Elson,[l] Howard J. Edenberg,[m] Pierre Fontanillas,[l] Abraham A. Palmer,[a,n] and Sandra Sanchez-Roige[a,n,o,*]

[a]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA
[b]Division of Human Genetics, Department of Psychiatry, Yale University School of Medicine, Orange, West Haven, CT, USA
[c]Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain
[d]Department of Psychology & Brain Sciences, Washington University in St. Louis, St Louis, MO, USA
[e]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA
[f]National Center of Posttraumatic Stress Disorder, VA CT Healthcare Center, West Haven, CT, USA
[g]Department of Psychiatry, University of Florida College of Medicine, Gainesville, FL, USA
[h]VA CT Healthcare Center, Department Psychiatry, West Haven, CT, USA
[i]Department of Mental Health, Hospital Universitari Vall d'Hebron, Barcelona, Spain
[j]Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain
[k]Department of Genetics, Microbiology, and Statistics, Faculty of Biology, Universitat de Barcelona, Barcelona, Spain
[l]23andMe, Inc., Sunnyvale, CA, USA
[m]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA
[n]Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA
[o]Division of Genetic Medicine, Department of Medicine, Vanderbilt University, Nashville, TN, USA
[p]Departments Psychiatry, Genetics, and Neuroscience, Yale Univ. School of Medicine, New Haven, CT, USA

eBioMedicine
2024;103: 105086
Published Online 4 April 2024
https://doi.org/10.1016/j.ebiom.2024.105086

Summary

Background Alcohol consumption is associated with numerous negative social and health outcomes. These associations may be direct consequences of drinking, or they may reflect common genetic factors that influence both alcohol consumption and other outcomes.

Methods We performed exploratory phenome-wide association studies (PheWAS) of three of the best studied protective single nucleotide polymorphisms (SNPs) in genes encoding ethanol metabolising enzymes (*ADH1B*: rs1229984-T, rs2066702-A; *ADH1C*: rs698-T) using up to 1109 health outcomes across 28 phenotypic categories (e.g., substance-use, mental health, sleep, immune, cardiovascular, metabolic) from a diverse 23andMe cohort, including European (N ≤ 2,619,939), Latin American (N ≤ 446,646) and African American (N ≤ 146,776) populations to uncover new and perhaps unexpected associations. These SNPs have been consistently implicated by both candidate gene studies and genome-wide association studies of alcohol-related behaviours but have not been investigated in detail for other relevant phenotypes in a hypothesis-free approach in such a large cohort of multiple ancestries. To provide insight into potential causal effects of alcohol consumption on the outcomes significant in the PheWAS, we performed univariable two-sample and one-sample Mendelian randomisation (MR) analyses.

Findings The minor allele rs1229984-T, which is protective against alcohol behaviours, showed the highest number of PheWAS associations across the three cohorts (N = 232, European; N = 29, Latin American; N = 7, African American). rs1229984-T influenced multiple domains of health. We replicated associations with alcohol-related behaviours, mental and sleep conditions, and cardio-metabolic health. We also found associations with understudied traits related to neurological (migraines, epilepsy), immune (allergies), musculoskeletal (fibromyalgia), and reproductive health (preeclampsia). MR analyses identified evidence of causal effects of alcohol consumption on liability for 35 of these outcomes in the European cohort.

*Corresponding author. Department of Psychiatry, University of California San Diego, La Jolla, CA, USA.
E-mail address: sanchezroige@ucsd.edu (S. Sanchez-Roige).
[q]Equal contribution.

Interpretation Our work demonstrates that polymorphisms in genes encoding alcohol metabolising enzymes affect multiple domains of health beyond alcohol-related behaviours. Understanding the underlying mechanisms of these effects could have implications for treatments and preventative medicine.

Funding MVJ, NCK, SBB, SSR and AAP were supported by T32IR5226 and 28IR-0070. SSR was also supported by NIDA DP1DA054394. NCK and RBC were also supported by R25MH081482. ASH was supported by funds from NIAAA K01AA030083. JLMO was supported by VA 1IK2CX002095. JLMO and JJMM were also supported by NIDA R21DA050160. JJMM was also supported by the Kavli Postdoctoral Award for Academic Diversity. EGA was supported by K01MH121659 from the NIMH/NIH, the Caroline Wiess Law Fund for Research in Molecular Medicine and the ARCO Foundation Young Teacher-Investigator Fund at Baylor College of Medicine. MSA was supported by the Instituto de Salud Carlos III and co-funded by the European Union Found: Fondo Social Europeo Plus (FSE+) (P19/01224, PI22/00464 and CP22/00128).

Keywords: Alcohol; PheWAS; Metabolising enzyme genes; *ADH1B*; *ADH1C*; rs1229984

Research in context

Evidence before this study

Alcohol use and the development of alcohol use disorder (**AUD**) are heritable traits. Genetic variants in several key alcohol metabolising enzyme genes are associated with alcohol metabolism, alcohol consumption, AUD and related traits. However, other effects of these genetic variants have not been well explored. We searched PubMed from Jan 1, 1990 to Jan 1, 2022 for reviews, genome-wide association studies (**GWAS**), and Mendelian randomisation (**MR**) analyses of alcohol metabolising enzyme genes. Search terms were "alcohol metabolising enzymes" OR the various genes encoding the enzymes (e.g., *ADH1B*) AND "review" OR "GWAS" OR "genome-wide association study" OR "Mendelian randomisation". We found recent key reviews describing alcohol dehydrogenases, aldehyde dehydrogenases, and AUD, which included thorough discussions of the key variants, as well as GWAS and MR analyses that explored the role of alcohol metabolism on various health-related traits.

Added value of this study

We performed a phenome-wide association study (**PheWAS**) of three variants in alcohol metabolising enzyme genes (*ADH1B*: rs1229984-T, rs2066702-A; *ADH1C*: rs698-T) for which there is strong evidence of an effect on alcohol-related traits and at least modest allele frequency (>1%) in three populations (European N ≤ 2,619,939; Latin American N ≤ 446,646; African American N ≤ 146,776). This constitutes the largest and most diverse dataset ever used to explore these genetic variants. We examined up to 1109 physical and mental health traits available in the 23andMe cohort, most of which have rarely been examined in published GWASs. We identified significant associations with conditions known to be comorbid with alcohol consumption, including mental health, sleep, immune conditions, cardio-metabolic conditions, gastrointestinal conditions, and cancer, as well as a number of associations with understudied traits, including allergies, migraines, epilepsy, myopia, rosacea, preeclampsia, malaria, and fibromyalgia. The MR analyses identified evidence of causal effects of genetic liability for alcohol consumption based on three SNPs in ethanol metabolising enzyme genes on 35 outcomes, which may partly explain common co-occurring traits and disorders of individuals who consume alcohol and may offer opportunities to develop preventative strategies for alcohol use and morbidity. MR results were also inconclusive for other traits identified via PheWAS, suggesting that these SNPs may be having effects on alcohol consumption and these other various outcomes via common biological pathways, and not merely a consequence of alcohol use.

Implications of all the available evidence

Our findings expand the existing knowledge of the relationships between alcohol metabolising enzyme gene variants and reveal new phenotypic associations across multiple ancestries.

Introduction

Heavy alcohol consumption is a heritable trait associated with numerous negative social and health outcomes.[1] Individual differences in how people metabolise alcohol alter the risk for excessive alcohol consumption and development of alcohol use disorders (**AUD**).[2,3] The primary enzymes responsible for alcohol metabolism are alcohol dehydrogenases (**ADHs**) and aldehyde dehydrogenases (**ALDHs**).[2,3] Some individuals have certain single nucleotide polymorphisms (**SNPs**) of the alcohol dehydrogenase 1B (*ADH1B*) and 1C (*ADH1C*) genes that change their amino acid sequence such that they metabolise alcohol to acetaldehyde at a faster rate. Similarly, individuals with coding polymorphisms in

aldehyde dehydrogenase 2 (*ALDH2*) convert acetaldehyde to acetate more slowly. Both of these polymorphisms influence the unpleasant effects of acetaldehyde concentrations[3–5] thereby reducing alcohol consumption and serving as protective factors against the development of AUD and alcohol-related health problems, such as alcohol-related liver disease and cancer.[3–5]

Single nucleotide polymorphisms (**SNPs**) located in genes that encode ethanol metabolising enzymes have been consistently associated with alcohol-related traits both by candidate gene studies and genome-wide association studies (**GWAS**).[3,6–17] As extensively reviewed by others,[3,4] rs1229984-T, a missense variant in *ADH1B*, is associated with reduced alcohol consumption and is protective against AUD and several similar alcohol phenotypes in individuals of European and Asian ancestries.[5,7,8,11,13,14,16,18–30] Additionally, rs2066702-A, another missense variant in *ADH1B*, which is primarily found in individuals of African ancestry, is similar to rs1229984-T.[13,15,20] The functional variant rs671-A in *ALDH2*, which is primarily found in individuals of Asian ancestry,[3,13,21,25,31,32] is associated with the alcohol flush reaction, and is also associated with reduced alcohol consumption and is strongly protective against AUD.[3] Other variants, including rs698-T in *ADH1C*, which is known to contribute to alcohol metabolism, have also been implicated in alcohol use behaviours and risk for AUDs.[3,9,12,14,33,34]

Intriguingly, these variants also have effects on other social and health traits beyond those directly related to alcohol (Supplementary Table S1). These associations could reflect common genetic factors (pleiotropy) that confer liability to both alcohol-related behaviours and AUD comorbidities, or could arise as a consequence of the effects of these variants on alcohol consumption. In this study, we focused on three of the best studied protective variants in the ethanol metabolising enzyme genes (*ADH1B*: rs1229984-T, rs2066702-A; *ADH1C*: rs698-T). Different ancestry groups often have differences in allele frequencies and linkage disequilibrium patterns (**LD**; i.e., frequency with which variants are inherited together). Therefore, we studied the role of these variants on outcomes via phenome-wide association analyses (**PheWAS**) across three diverse populations ($N \leq 2{,}619{,}939$, European; $N \leq 446{,}646$, Latin American; $N \leq 146{,}776$, African American) from the 23andMe research cohort. We included rs2066702-A for its prior association with AUD in individuals of African ancestry,[3,7] despite its very low frequency in other populations. *ALDH2* rs671-A, which is strongly negatively associated with alcohol consumption in Asian populations,[3,35] could not be tested due to extremely low allele frequency in the populations studied, which did not include an East Asian cohort. We employed a hypothesis-free PheWAS approach to examine up to 1109 self-reported traits across 28 phenotypic categories (e.g., substance-use related, mental health, sleep, immune, cardiovascular, metabolic) that had been previously collected by 23andMe, most of which have never been published as GWAS or explored in non-European ancestry populations, meaning that these data are not available from commonly consulted resources such as the EMBL-EBI GWAS catalog.[36] Lastly, we examined causality via two- and one-sample Mendelian randomisation (**MR**) analyses.[37,38] We anticipated replication of previously reported associations with alcohol-related behaviours and discovery of additional associations across numerous categories given our increased sample size and more ancestrally diverse dataset.

Methods

23andMe cohort and data collection

Our cohort consisted of research participants from 23andMe, Inc., a direct-to-consumer genetics company. Demographic information about this sample is presented in Supplementary Table S2. Ancestry was determined through an analysis of local ancestry,[39] as described in detail in the Supplementary Material. Only individuals who were categorised as being of European, Latin American, or African American ancestries (which were the largest available ancestral groups at the time of the analysis) based on empirical genotype data[39] were retained. Individuals of African American and Latin American ancestries are admixed with broadly varying contributions from Europe, Africa and the Americas, and sometimes other populations as well. The distributions of the length of segments of European and African American ancestry are very different between individuals of African American and Latin American ancestry, because of distinct admixture timing between the three ancestral populations in the two ethnic groups. Therefore, we trained a logistic classifier that takes one individual's length histogram of segments of African, European, and American ancestries, and predicts whether the individual is likely of African American or Latin American descent. Additional genotyping information is included in the Supplementary Material. Randomisation and blinding did not apply to the current study.

An overview of the data collection process has been previously described.[40] All participants provided answers to online surveys available on the 23andMe platform. The traits were self-reported and pertained to 28 phenotypic categories (e.g., substance-use related, mental health, sleep, immune, cardiovascular, metabolic) that we identified.

Ethics

All 23andMe research participants included in the analyses provided informed consent and volunteered to participate in the research online, under a protocol approved by the external AAHRPP-accredited IRB,

Ethical & Independent (E&I) Review Services. As of 2022, E&I Review Services is part of Salus IRB (https://www.versiticlinicaltrials.org/salusirb).

Variant selection

We selected three protective SNPs for alcohol consumption (*ADH1B*: rs1229984-T; rs2066702-A; *ADH1C*: rs698-T; Supplementary Figures S1–S3) based on their roles in alcohol metabolism and results from diverse large-scale GWAS ([7–17]; see[3,4] for comprehensive reviews), that were available in the 23andMe cohort and passed the quality control criteria (Table 1, Supplementary Tables S3–S7). The functional SNP *ALDH2* (rs671-A, Supplementary Figure S4), which is heavily studied in individuals of East Asian ancestry, could not be tested in this study due to near zero (<0.002) allele frequency in the populations studied, which did not include East Asians. Imputed variant statistics are shown in Supplementary Table S4.

Phenome-wide association analyses

We employed single-SNP PheWASs to examine associations between the three selected SNPs and 1316 traits in the European cohort, 962 traits in the Latin American cohort and 785 traits in the African American cohort from the 23andMe database (Fig. 1). We excluded traits with <1,000 responses (15.73%, 8.84% and 11.46% from the European, Latin American and African American cohorts respectively), based on a prior simulation study for power analysis of PheWAS.[41] The final number of traits tested was 1109 in the European cohort, 883 in the Latin American cohort and 695 in the African American cohort, with the sample size ranging from 1,012 to 2,619,939 in the European cohort, 1,006–446,646 in the Latin American cohort, and 1,000–146,776 in the African American cohort (Fig. 1 and Supplementary Table S8). Most traits had sample sizes >10,000, which provided power (80%) to detect modest associations (e.g., beta = 0.02; Supplementary Figure S5). For case–control comparisons, we computed association test results by logistic regression. For quantitative traits, association tests were performed by linear regression. We assumed additive allelic effects and included covariates for age at the time of the analysis (as determined by participant date of birth), sex, and the top five ancestry-specific principal components, to account for residual population structure within each group. Within each cohort, we used a 5% FDR Benjamini & Hochberg correction for multiple testing.[42]

Mendelian randomisation

In the PheWAS, we examined the association between three SNPs in ethanol metabolising enzyme genes, well-studied for their protective effects on alcohol consumption,[3,4] and various outcomes. Some of the associations could be the result of horizontal pleiotropy (i.e., a SNP is independently associated with both alcohol consumption and these other traits[43]) or vertical pleiotropy, also known as causality (i.e., a SNP is associated with these other traits as a consequence of alcohol consumption[38,44]). To explore whether the associations between alcohol consumption and these outcomes are causal, we conducted univariable two-sample MR analysis[37,38,45–47] (for a flowchart see Supplementary Figure S6), following the STROBE guidelines (Supplementary Material). We extracted the summary statistics (beta and SE) for the same SNPs used in the PheWAS from a previously published GWAS of alcohol consumption measured in an independent cohort (UK Biobank, N = 160,824[9];). Alcohol consumption was measured by the Alcohol Use Disorder Identification Test (**AUDIT**) Consumption scores (items 1–3[48]), which we treated as our exposure. Summary statistics (beta and SE) for the PheWAS significant outcomes were extracted from the European cohort results. Exposure and outcomes were thus derived from similar underlying populations (i.e., European ancestry, middle-aged

rsID	chr:pos:alleles	Protective allele	Gene	Effect predicted by VEP	eQTL in GTEx	eQTL tissues	eQTL *p* values	Amino acid	freq_EUR	freq_LA	freq_AA
rs2066702	4:100229017:G:A	A	*ADH1B*	missense variant	Y (*ADH1A*)	Adipose-Subcutaneous; Breast—Mammary Tissues; Skin—Sun Exposed (Lower Leg)	8.00E-07; 5.70E-06; 8.60E-05	R (Arg) > C (Cys)	0.002	0.017	0.189
rs1229984	4:100239319:T:C	T	*ADH1B*	missense variant	N	–	–	H (His) > R (Arg), H (His) > P (Pro)	0.035	0.068	0.012
rs698	4:100260789:T:C	T	*ADH1C*	intron variant (*ADH1B*), missense variant, downstream gene variant (*ADH1C*)	Y (*ADH1C*)			I (Ile) > F (Phe), I (Ile) > V (Val)	0.588	0.707	0.856
rs3114045	4:100252560:T:C	T	*ADH1B*	intron variant (*ADH1B*)	Y (RP11-571L19.8; *METAP1*)	Adipose—Subcutaneous; Skin–Sun Exposed (Lower leg); Nerve—Tibial; Thyroid	3.00E-07; 2.70E-06; 1.80E-04; 2.70E-04		0.135	0.299	0.071

Shown are the alcohol-protective allele frequencies. *VEP*, variant effect predictor. Additional eQTL information is presented in Supplementary Tables S3 and S9.

Table 1: **Alcohol variants included in the PheWAS and Mendelian randomisation analyses.**

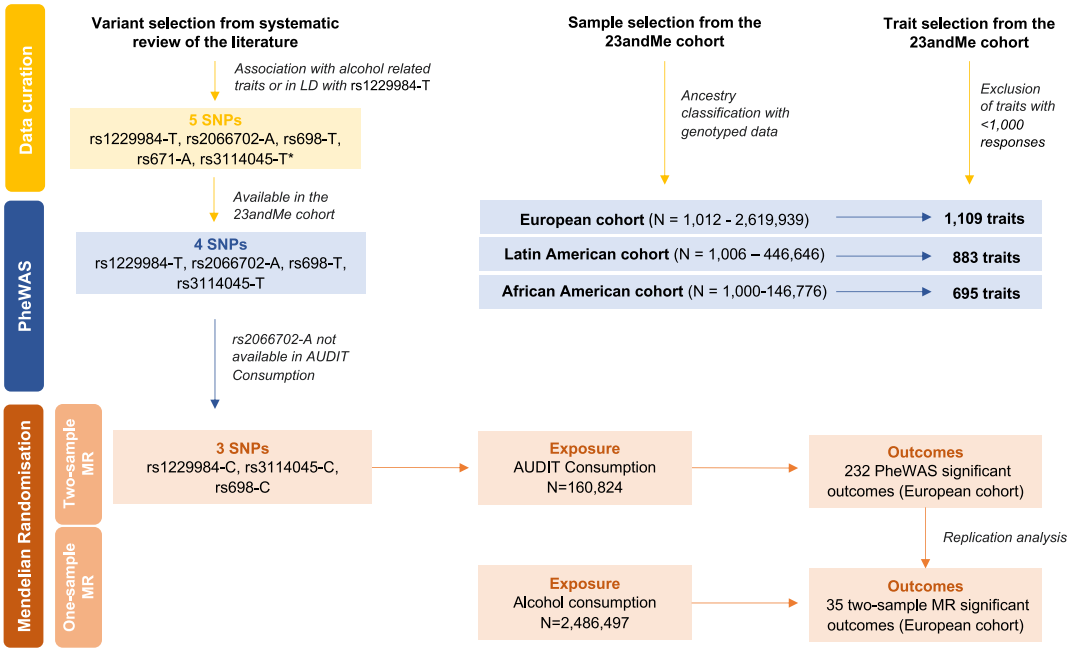

***Fig. 1:* Overview of the study**. Four variants (rs1229984-T, rs2066702-A, rs698-T, rs671-A) were selected based on a systematic literature review of genes (*ADH1B, ADH1C, ALDH2*) implicated in alcohol use disorder and other alcohol-related traits. rs671-A could not be tested due to extremely low allele frequency in the populations studied. We tested the association of the three remaining SNPs and 1109 traits available in 23andMe via a phenome-wide association study (**PheWAS**). We included traits with >1000 responses per cohort (EUR N = 1012–2,619,939; LA N = 1006–446,646; AA N = 1000–146,776). To assess evidence of a causal effect of more alcohol consumption by ethanol metabolising enzyme SNPs and PheWAS significant traits, a two-sample MR analysis and a one-sample replication Mendelian randomisation (**MR**) analysis were conducted. SNP effects associated with AUDIT Consumption and number of drinks consumed over the past 2 weeks were treated as exposures for the two- and one-sample MR analyses respectively. *In order to have a minimum of three SNPs for MR analyses, an additional *ADH1B* SNP (rs3114045) was added to replace replaced rs2066702, as this SNP was not available in the AUDIT Consumption GWAS.

adults). We restricted the MR analysis to the European cohort because there are currently no well-powered GWAS of alcohol consumption in individuals of Latin American or African American populations that we could use for the analysis.

The SNPs extracted were each independently (r^2 < 0.27) associated with alcohol consumption at a genome-wide significant level (p < 5.00E-08). rs2066702-A was not available in the alcohol consumption GWAS.[9] Because some MR methods require a minimum of three SNPs,[49] we expanded our PheWAS results to include another *ADH1B* variant (rs3114045; Supplementary Table S9, Supplementary Figure S6), which was well genotyped and imputed in the 23andMe cohort, and was significant in the alcohol consumption GWAS.[9] We evaluated instrument strength by assessing the proportion of variance in the phenotype explained by each instrument and the F-statistics, where a value under 10 may indicate a weak instrument.[50]

Next, we harmonised the exposure-outcome datasets (i.e., consistent direction of effect between exposure and outcome associations, where the effect alleles for all SNPs were associated with more alcohol consumption), and applied an inverse variance weighted (**IVW**) analysis to give an overall estimate of causal effect.[51] The causal estimate of the IVW analysis indicates the causal increase in the outcome (or log odds of the outcome for a binary outcome) per unit change in the exposure. Because IVW assumes that all variants are valid instrumental variables in the absence of horizontal pleiotropy (Supplementary Figure S10), any deviation from this assumption can introduce bias in MR analysis estimates if the SNPs affect the outcome via a different pathway. To evaluate the robustness of our interpretations, we compared the IVW results with two other methods known to be more robust to horizontal pleiotropy. Specifically, we used weighted median MR, which allows for 50% of the instrumental variables to be invalid, and MR-Egger regression, which relaxes the assumption that the effects of the variants on the outcome are entirely mediated via the exposure.[52–54] This is achieved by not constraining the intercept, where a non-zero intercept indicates directional horizontal pleiotropy. To test the suitability of the MR-Egger method, we calculated the I^2_{GX} statistic to quantify the degree of regression dilution bias due to measurement error of

SNP-exposure effects.[50] Where there was evidence of violation of negligible measurement error (**NOME**, I^2 0.6–0.9), MR-Egger was performed with simulation extrapolation (**SIMEX**) correction.[50,55] SIMEX relies on simulation to estimate or reduce bias due to measurement error, considering additional data sets with increasing measurement error variance. Lastly, we tested for heterogeneity between the individual SNPs included in the genetic instrument using Cochran's Q test of heterogeneity. To assess whether any single SNP was driving effect estimates, we conducted leave one SNP out MR analyses. All analyses were conducted using the *TwoSampleMR* package (v0.5.6)[56] in *R* (v 4.2.1).

Triangulating results across multiple MR methods provides the strongest evidence of causal inference.[57] In our case, the IVW method was the main analysis and the other methods served as sensitivity analyses. We thus present the estimates from the IVW method in the manuscript; results for all three methods are shown in Supplementary Table S14. Significance was determined after multiple testing correction based on the number of outcomes tested (Supplementary Figure S6).

Finally, we performed one-sample MR as a form of replication using the *MendelianRandomization* package (v. 0.6.0)[49] in *R* (v. 4.2.1). We included the summary statistics for alcohol consumption (i.e., number of drinks consumed over the past 2-weeks) as the exposure and any significant outcomes obtained via two-sample MR. Our analyses were not pre-registered, and therefore our results should be considered exploratory.[58]

Role of funders

The funders, as listed in the Acknowledgements section, did not have any role in study design, data collection, data analyses, interpretation, or writing of the report.

Results

The European cohort included up to 2,619,939 individuals (56.20% female; mean age = 51.77 [SD = 18.71]); the Latin American cohort included 446,646 individuals (57.30% female; mean age = 42.89 [SD = 17.74]); and the African American cohort included 146,776 individuals (58.30% female; mean age = 44.27 [SD = 17.93]). Alcohol consumption (≥1 drink) in the past two weeks was reported by 65.29% (European cohort), 61.25% (Latin American cohort), and 57.16% (African American cohort). Additional sociodemographic and health-related outcome results can be found in Supplementary Table S2.

The list of PheWAS association results after 5% FDR correction is available in Supplementary Table S11 (European), Supplementary Table S12 (Latin American), and Supplementary Table S13 (African American). The standardised effect sizes were generally small (Supplementary Figures S7–S9; European, β = –6.25 to 3.65; Latin American, β = –2.70 to 2.32; African American, β = –1.26 to 1.85), as is expected for a single SNP's effect on a complex trait. Overall, all findings that were observed in two or more cohort groups showed the same direction of effect and similar effect sizes (Supplementary Figures S11–S14).

ADH1B rs1229984

rs1229984-T (protective allele (T) frequency: 0.06 EUR, 0.08 LA, 0.04 AA) showed the highest number of FDR-significant PheWAS associations (Fig. 2, Supplementary Table S11) in the populations studied, despite its low frequency in those populations.

In the European cohort, rs1229984-T was associated with 232 traits from each of the 28 categories we considered; we describe some of the most prominent findings below:

Alcohol-related behaviours. The alcohol-protective rs1229984-T allele was most strongly associated, as expected, with alcohol-related traits [e.g., less alcohol consumption (β = –0.25, p < 1.00E-216), less preference for whisky (β = –0.05, p = 2.39E-13), more alcohol flush (β = 0.12, p = 2.63E-34), more wine headaches (β = 0.28, p = 1.53E-117), less heavy drinking (>3 daily servings; β = –0.60, p = 2.97E-44), less concern regarding a need to stop drinking alcohol (β = –0.45, p = 1.54E-144)]. We also found associations with conditions that are often comorbid with alcohol consumption, such as gastrointestinal disorders [e.g., less cirrhosis (β = –0.25, p = 1.19E-09); less gastroesophageal reflux disease (**GERD**; β = –0.03, p = 1.47E-08)], renal conditions [i.e., more kidney stones (β = 0.07, p = 1.21E-22)], as well as respiratory conditions that tend to be associated with smoking, which commonly co-occurs with alcohol consumption [e.g., less chronic bronchitis (β = –0.11, p = 3.61E-05)].

Central nervous system (***CNS***). We found associations between the protective allele (rs1229984-T) and mental health traits [e.g., lower rates of self-reported post-traumatic stress disorder (β = –0.04, p = 2.21E-05), bipolar disorder (β = –0.03, p = 4.12E-03), and depression (β = –0.01, p = 5.43E-03)], and neurological traits [e.g., fewer reports of memory loss (β = –0.10, p = 1.51E-05), lower rates of epilepsy (β = –0.04, p = 9.79E-03), and less restless leg syndrome (β = –0.05, p = 2.30E-07)]. In addition, we observed associations with sleep health [e.g., less insomnia (β = –0.03, p = 6.05E-07), less sleep paralysis (β = –0.08, p = 3.15E-16), less REM sleep behaviour disorder (β = –0.10, p = 1.42E-09), less sleep apnea (β = 0.03, p = 1.13E-04), higher sleep quality (β = 0.01, p = 1.01E-09)].

Immune, cardiovascular, endocrine/metabolic. Many of the associations we observed were related to immune, cardio-metabolic, and endocrine/metabolic traits (Fig. 3). For example, we found associations between rs1229984-T and multiple immune traits [e.g., fewer allergies (β = –0.07, p = 4.80E-11), less Crohn's or ulcerative colitis (β = –0.11, p = 7.45–16)]; cardiovascular

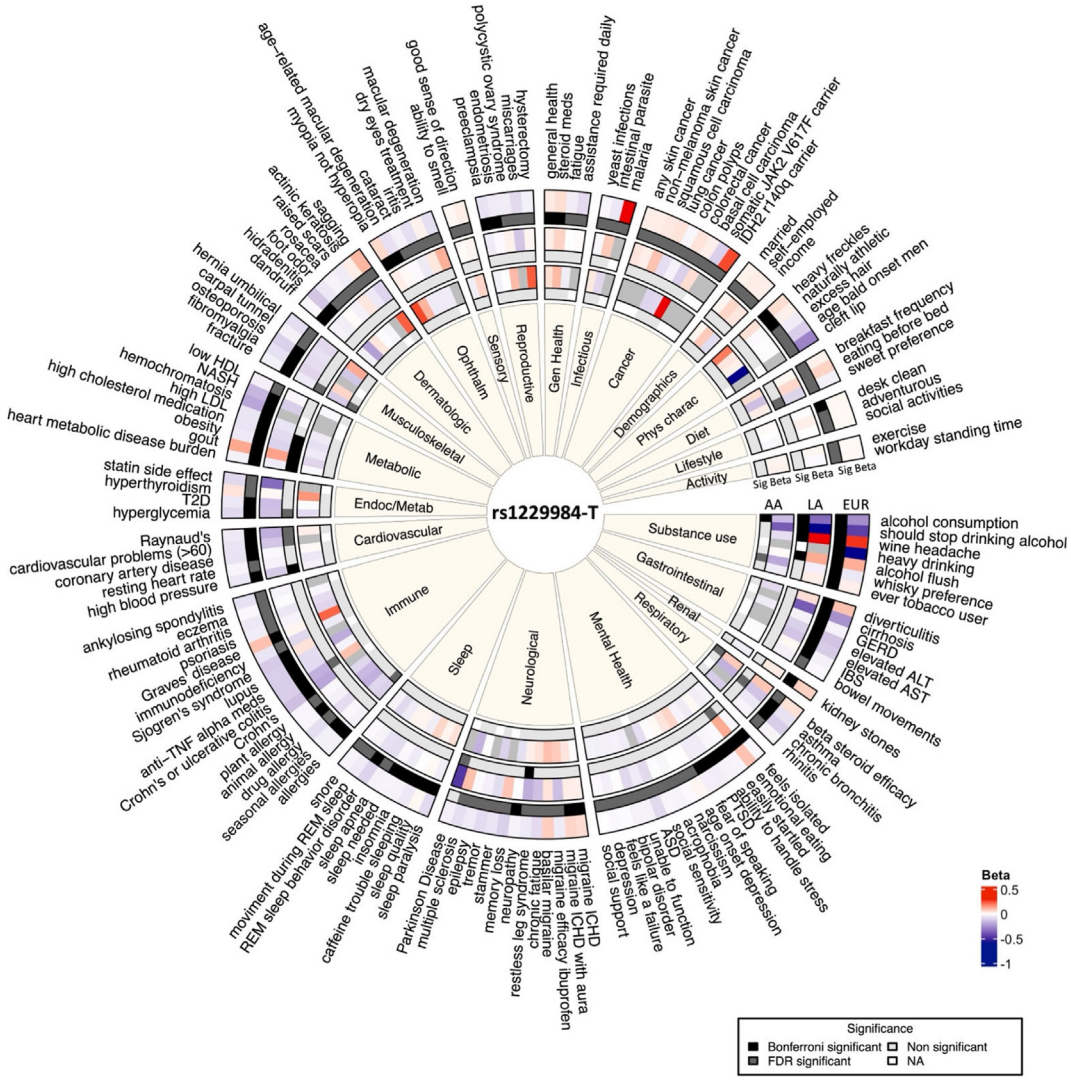

Fig. 2: FDR (5%)-significant PheWAS associations between rs1229984-T and a selection of traits from multiple categories in the European (EUR), Latin American (LA), and African American (AA) cohorts. The full list of FDR-significant associations is shown in Supplementary Tables S11–S13.

traits [e.g., fewer self-reports of high blood pressure (β = –0.05, p = 4.93E-23), less coronary artery disease (β = –0.06, p = 4.60E-07)]; and metabolic/endocrine traits [e.g., less obesity (β = –0.05, p = 3.11E-25), less hyperglycaemia (β = –0.08, p = 7.14E-13), less type II diabetes (β = –0.05, p = 7.68E-09), fewer self-reports of medications to reduce cholesterol (β = –0.06, p = 1.57E-36), and less heart metabolic disease burden (β = –0.03, p = 2.69E-42)].

Other. We also found associations between rs1229984-T and several other traits, such as musculoskeletal [e.g., fewer fractures (β = –0.04, p = 1.54E-04), less fibromyalgia (β = –0.07, p = 5.84E-09)] and dermatologic traits [e.g., less rosacea (β = –0.03, p = 1.57E-03)], eye health [e.g., less age-related macular degeneration (β = –0.07, p = 1.66E-05)], and reproductive health [e.g., less preeclampsia or high blood pressure during pregnancy (β = –0.06, p = 4.72E-07) and less endometriosis liability (β = –0.05, p = 1.34E-06)].

Overall, the *alcohol-protective* rs1229984-T allele was associated with *better* health outcomes, including general health (β = 0.03, p = 9.90E-55), less chronic fatigue (β = –0.08, p = 2.270E-11), and less daily assistance required (β = –0.08, p = 2.53E-03). However, we also found that 32 out of the 232 associations (13.79%) were with *worse* health outcomes, including more self-reports of lifetime tobacco use (β = 0.03, p = 1.25E-09), a mental health trait [i.e., more emotional eating (β = 0.09,

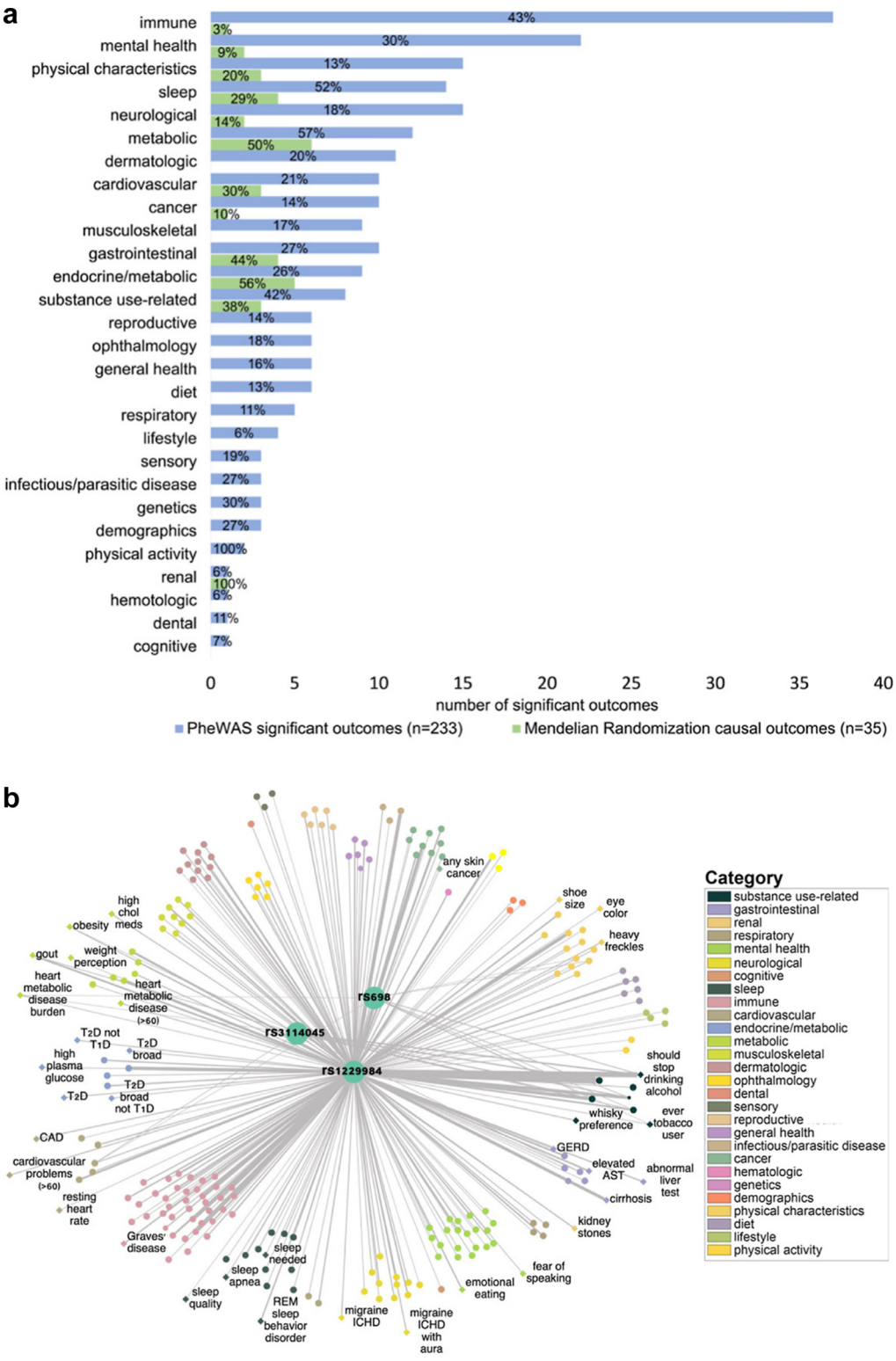

Fig. 3: **a**) Integrated bar chart featuring significant traits identified in both the PheWAS and MR analyses, grouped by categories. Each bar represents the number of outcomes for each category. The percentage represents the proportion of significant outcomes from the total possible number of outcomes within that specific category. **b**) Integrated network analysis of PheWAS and MR results in the European cohort. Each node in the network denotes one of the 233 PheWAS or 35 MR significant traits (annotated), with colour indicating the category of the association, the width of the line denoting the magnitude (beta) of the effect size, and the size of the shape reflecting the association strength (log10 *p* values).

p = 2.13E-06)], neurological conditions [e.g., more migraine with aura (β = 0.06, p = 4.37E-11)], immune and endocrine/metabolic conditions [e.g., more Graves' disease (β = 0.09, p = 1.09–05), more gout (β = 0.13, p = 4.74E-37), more hyperthyroidism (β = 0.14, p = 8.46E-03)], infectious/parasitic diseases [e.g., more malaria (β = 0.30, p = 3.93E-03)], conditions related to eye health [i.e., more myopia (β = 0.06, p = 9.74E-09)], and several cancers [e.g., more skin cancer (β = 0.03, p = 1.41E-04), more lung cancer (β = 0.04, p = 1.65E-03)].

In the Latin American cohort, we identified similar findings: rs1229984-T was associated with 29 traits across 13 categories (Supplementary Table S12):

Alcohol-related behaviours. Again, the strongest associations of the protective allele were observed with alcohol-related traits [e.g., less alcohol consumption (β = –0.24, p = 7.09E-216), more wine headaches (β = 0.29, p = 1.68E-22), fewer concerns regarding a need to stop drinking alcohol (β = –0.42, p = 3.44E-24)], and conditions related to alcohol consumption and smoking, such as respiratory disorders [i.e., less asthma (β = –0.10, p = 1.37E-03)].

CNS. We found associations with neurological traits [i.e., less restless leg syndrome (β = –0.13 p = 5.04E-07) and Parkinson's disease (β = –0.34 p = 1.41E-04)].

Immune, cardiovascular, endocrine/metabolic. Many of the associations were again related to immune, cardiometabolic, and endocrine/metabolic traits. For example, we found associations with immune traits [e.g., fewer allergies (β = –0.05, p = 2.72E-07) and less Crohn's disease (β = –0.12, p = 1.02E-03)]; cardiovascular traits [e.g., lower rates of high blood pressure (β = –0.08, p = 4.94E-10) and fewer cardiovascular problems at an older age (β = –0.09, p = 3.91E-04)]; endocrine/metabolic traits [e.g., less obesity (β = –0.07, p = 1.35E-12), less hyperglycaemia (β = –0.10, p = 1.30E-04), and less heart metabolic disease burden (β = –0.03, p = 6.26E-16)].

Other. Again, we found associations with understudied traits such as musculoskeletal conditions [e.g., fewer fractures (β = –0.03, p = 9.84E-04)]. We also noted that 3 out of the 20 significant associations (10.34%) were unexpected, including a mental health trait [i.e., more emotional eating (β = 0.12, p = 8.57E-04)], a dermatologic trait [i.e., more keloids (β = 0.06, p = 2.23E-04)], and a metabolic trait [i.e., more gout (β = 0.14, p = 6.95E-06)].

In the African American cohort, which was the smallest cohort, rs1229984-T was associated with 7 traits (Supplementary Table S13). The strongest associations were with alcohol-related traits [e.g., less alcohol consumption (β = –0.15, p = 1.12E-12) and less preference for whisky (β = –0.19, p = 1.37E-04)]. In addition, we found associations with similar immune [e.g., fewer allergies (β = –0.14, p = 4.35E-08)] and musculoskeletal traits [i.e., fewer fractures (β = –0.11, p = 1.43E-04)] observed in the other cohorts. In this cohort only, rs1229984-T was associated with lower height (β = –0.05 p = 1.87E-05).

ADH1B rs2066702

rs2066702-A (protective allele (A) frequency: 0.002 EUR, 0.017 LA, 0.189 AA) was associated with less alcohol consumption in the European (β = –0.09, p = 3.21E-07), Latin American (β = –0.09, p = 1.82E-08), and African American (β = –0.08, p = 3.78E-19) cohorts, but no other alcohol-related behaviours were found. In the Latin American cohort, rs2066702-A was associated with three additional traits, including a mental health trait [i.e., lower rates of self-reported depression (β = –0.08, p = 7.04E-05)] and two metabolic traits [i.e., lower BMI (β = –0.05, p = 2.50E-10) and lower weight (β = –0.04, p = 1.54E-07)]. No other associations were found in the European and African American cohorts.

ADH1C rs698

rs698-T (protective allele (T) frequency: 0.61 EUR, 0.67 LA, 0.82 AA), was associated with less alcohol consumption in the European (β = –0.04, p = 6.44E-136), Latin American (β = –0.05, p = 7.39E-33) and African American (β = –0.04, p = 1.94E-05) cohorts.

In the European cohort, rs698-T was associated with 9 additional traits:

Alcohol-related behaviours. The protective allele rs698-T was associated with two additional alcohol-related traits [i.e., more wine headaches (β = 0.02, p = 9.48E-05), and fewer concerns regarding a need to stop drinking alcohol (β = –0.05, p = 5.55E-12)].

CNS. We found an association with a sleep trait [i.e., less movement during REM sleep (β = –0.02, p = 1.76E-04)].

Cardiovascular, endocrine/metabolic. We observed associations with the same cardiovascular [i.e., lower rates of high blood pressure (β = –0.01, p = 1.86E-05)] and metabolic traits [i.e., less heart metabolic disease burden (β = –0.01, p = 1.55E-04)] that were found to be associated with rs1229984-T.

Other. rs698-T was associated with physical characteristic traits [i.e., height (β = –0.003, p = 1.55E-04) and body type (β = –0.02, p = 5.83E-05)]. Only one trait was associated with the protective rs698-T allele in the unexpected direction [i.e., more reports of lifetime tobacco use (β = 0.01, p = 1.11E-04)].

In the Latin American cohort, rs698-T was associated with 2 additional traits, which included an alcohol-related trait [i.e., fewer concerns regarding a need to stop drinking alcohol (β = –0.08, p = 4.66E-05)], and one metabolic trait [i.e., reduced heart metabolic disease burden (β = –0.01, p = 1.27E-04)].

We did not observe any significant associations in the African American cohort.

Mendelian randomisation

Using three SNPs in ethanol metabolising enzyme genes that were robustly and independently associated with alcohol consumption measured via AUDIT scores, two-sample MR in the European cohort provided

evidence of a causal effect of genetically determined alcohol consumption on 35 outcomes (15.09% of the 232 PheWAS significant traits) after accounting for multiple testing correction and sensitivity analyses (Supplementary Figures S10, S15–S18, Supplementary Table S14). The proportion of the phenotypic variance of the exposure explained by each instrument (rs1229984-C, R^2 = 0.04; rs3114045-C, R^2 = 0.02; rs698-C, R^2 = 0.01) and the F-statistics were adequate (F > 32.58), indicative of valid instrument strength.

Alcohol-related behaviours. MR results showed strong evidence of a causal effect of increased alcohol consumption on higher preference for whisky (IVW OR = 1.34, 95% CI = 1.25–1.44, p = 5.39E-17), and increased concern about the need to stop drinking alcohol (IVW OR = 12.09, 95% CI = 10.11–14.46, p = 1.00E-164). Gastrointestinal conditions that are often associated with alcohol consumption, such as cirrhosis (IVW OR = 3.90, 95% CI = 2.58–5.88, p = 9.18E-11), elevated aspartate aminotransferase (**AST**), which is a clinical biomarker of liver damage (IVW OR = 1.90, 95% CI = 1.50–2.42, p = 1.64E-07) and GERD (IVW OR = 1.17, 95% CI = 1.11–1.22, p = 6.07E-10), showed evidence of being causally associated with alcohol consumption.

CNS. We found evidence of a causal effect of increased alcohol consumption on one mental health trait [i.e., fear of speaking (IVW OR = 1.21, 95% CI = 1.11–1.30, p = 2.88E-06)] and sleep alterations [e.g., REM sleep behaviour disorder (IVW OR = 1.76, 95% CI = 1.50–2.06, p = 4.53E-12), poor sleep quality (IVW OR = 1.07, 95% CI = 1.05–1.10, p = 7.14E-13)].

Cardiovascular, endocrine/metabolic traits. We found evidence of a causal effect of increased alcohol consumption on cardiovascular conditions [e.g., coronary artery disease (IVW OR = 1.43, 95% CI = 1.26–1.63, p = 5.86E-08)], metabolic/endocrine conditions [e.g., obesity (IVW OR = 1.30, 95% CI = 1.24–1.36, p = 3.32E-28), type 2 diabetes (**T2D**, IVW OR = 1.35, 95% CI = 1.23–1.47, p = 2.59E-11), high plasma glucose (IVW OR = 1.20, 95% CI = 1.11–1.29, p = 6.17E-07), and heart metabolic disease burden (IVW OR = 1.17, 95% CI = 1.15–1.20, p = 2.08E-51)].

Other. While the majority of the associations were positive (i.e., higher alcohol consumption increasing the odds for these outcomes), 28.57% of the associations were negative [e.g., lower self-reports of: ever using tobacco (IVW OR = 0.85, 95% CI = 0.79–0.91, p = 1.49E-05), emotional eating (IVW OR = 0.61, 95% CI = 0.52–0.72, p = 5.58E-09), migraines (IVW OR = 0.75, 95% CI = 0.69–0.81, p = 2.84E-12), kidney stones (IVW OR = 0.67, 95% CI = 0.60–0.74, p = 1.13E-13), gout (IVW OR = 0.48, 95% CI = 0.43–0.54, p = 1.26E-09), and Grave's disease (IVW OR = 0.57, 95% CI = 0.46–0.70, p = 2.47E-07)]. We identified evidence of causal effect of increased alcohol consumption on other domains, such as physical characteristics [e.g., heavy freckles (IVW OR = 0.74, 95% CI = 0.67–0.82, p = 8.58E-09)], that may be indicative of residual biases not fully accounted for by sensitivity analyses.

For all outcomes, the direction of effect was consistent across the three MR methods tested (IVW, weighted median, MR-Egger), and the high I^2_{GX} values of the SNP-exposure associations indicated low bias from regression dilution (i.e., attenuation towards the null of an association as a result of measurement error). There was some heterogeneity in the SNP effects, funnel plots identifying an influential potential outlier (rs1229984-C; Supplementary Figure S16); nonetheless, all results were still significant following leave-one-out analyses (Supplementary Table S14).

All of these associations were replicated in the one-sample MR (Supplementary Table S15).

Discussion

The societal and economic consequences of alcohol consumption are extensive,[59,60] and the importance of the role of ethanol metabolising enzyme variants on alcohol-related behaviours is undeniable.[3,4] This is the largest study investigating the role of three of these variants in a diverse population comprising over 3 million individuals. Through a hypothesis-free approach aimed at investigating phenome-wide associations across numerous traits and categories, our study reveals three key insights. First, we provide further evidence that the alleles that encode the more active alcohol metabolising enzymes[3] of these functionally important variants offer protection against excessive alcohol consumption across three ancestries.[61] We replicated associations between three variants in alcohol metabolising enzyme genes and alcohol-related phenotypes–from alcohol consumption to wine headache to heavy drinking. Second, our PheWAS provides evidence of even further-reaching effects of these variants on human health, including replicating associations with mental, sleep, and cardio-metabolic health from other large datasets like the UK Biobank, and revealing associations of these SNPs with numerous understudied traits with no prior GWAS data. Third, MR analyses suggest that some of these associations may be mediated, at least in part, by genetically determined alcohol consumption, while others may reflect common genetic factors across the traits studied. Identifying shared pathways of risk between alcohol consumption and these outcomes could have high translational potential, allowing tailoring of treatments for multiple medical conditions.

It is well-established that alcohol consumption is associated with poor mental health.[62] PheWAS associations with mental health traits (e.g., PTSD, depression and earlier onset of depression, bipolar disorder) were primarily identified with rs1229984-T in the European cohort, which is largely due to greater sample size and therefore discovery power compared to the other

ancestries. This aligns with GWAS in individuals of European ancestry showing strong genetic correlations between alcohol-related behaviours and mental health traits[7–9,13,14]; our PheWAS results show that this individual SNP contributes to those genetic correlations. MR, which could only be performed in the European ancestry cohort due to the lack of well-powered alcohol consumption GWAS in other populations, provided limited evidence that alcohol consumption contributes to mental disorders, consistent with a comprehensive systematic review (including 63 studies) identifying causal links between mental disorders and increased alcohol consumption without evidence for the reverse.[63]

Another cluster of associations related to two of the alcohol-protective alleles we studied (rs1229984-T and rs698-T) were with a variety of sleep conditions. Prior large-scale GWAS of alcohol consumption behaviours in individuals of European ancestry revealed significant genetic correlations with sleep alterations,[7–9,13,14] implying that there are common underlying biological pathways for alcohol consumption and sleep health. The MR analyses in our study support these associations by identifying a potential effect of more alcohol consumption on 29% of the PheWAS significant sleep traits, consistent with prior work,[63–67] and extending to other sleep behaviours, such as REM sleep disorder. Our MR results (based only on the SNPs we studied) are somewhat contradictory with a prior MR study (based on polygenic scores) suggesting a causal effect of another sleep behaviour (insomnia) on alcohol consumption and dependence but not vice versa[64]; however, we did not test for bi-directional associations in the current study. While collectively this line of work suggests the potential for simultaneously treating both excessive alcohol consumption and sleep disturbances,[68] future research should explore bi-directional associations in larger and more diverse samples and consider potential confounding factors such as psychiatric and metabolic disorders not accounted for in this or prior studies.[69,70]

Beyond mental health, the breadth of information available in our 23andMe cohort revealed PheWAS associations with other neurological conditions that often co-occur with alcohol use, such as migraines. The association between alcohol-protective alleles and increased risk for migraines was specifically for rs122998-T. Our results add to previous observational studies reporting that migraines can be triggered by alcohol consumption,[71,72] and prior MR studies linking genetically determined alcohol consumption with lower migraine risk,[73] as well as providing evidence for the reverse.[74]

Alcohol consumption also tends to be associated with a range of other medical conditions, such as gastrointestinal or respiratory health. In our study, nearly half of the gastrointestinal conditions that were found to be significant in the European PheWAS, such as cirrhosis and GERD, were suggested as a direct result of alcohol consumption in our MR analysis. This is consistent with a recent MR study based on polygenic scores showing a causal effect of alcohol consumption on gastrointestinal diseases, such as duodenal ulcer, cirrhosis, and chronic pancreatitis,[75] and supporting observational studies showing that chronic alcohol consumption, especially in excess, can lead to a range of gastrointestinal symptoms (e.g.,[76–78]). We also observed associations with respiratory conditions, such as self-reported diagnosis of asthma and, for the first time, the efficacy of beta steroids in the treatment of asthma. Although prior prospective studies have found a U-shaped relationship between alcohol consumption and incidents of asthma,[79] our MR analyses were discrepant, either because we were unable to stratify by alcohol consumption levels, or because comorbid smoking consumption, which is heavily associated with alcohol consumption and is a strong contributor to respiratory conditions, was not included in this analysis.

In addition, we identified that carriers of the alcohol-protective alleles we studied had a more favourable cardiovascular (e.g., high blood pressure, coronary artery disease) and endocrine-metabolic (e.g., obesity, hyperglycaemia, hyperthyroidism, T2D, cholesterol medication) profile, consistent with early candidate gene studies[80,81] and prospective cohort studies.[82] Although these loci are known for their contribution to ethanol metabolism in hepatic tissues, transcripts of *ADH1B* are also highly expressed in multiple other tissues, including cardiac and adipose tissues.[3,4,83,84] Therefore, it is not surprising that variants in *ADH1B* have been implicated in lipid metabolism in humans[84–91] and in cholesterol and triglyceride absorption in animal models, regardless of alcohol consumption.[92] While our PheWAS results suggest that these associations may partially arise from common genetic underpinnings, our MR results contribute to the growing body of literature showing that the association with cardiometabolic outcomes may also be, at least in part, a direct consequence of alcohol drinking,[46,93–95] and suggest that decreasing alcohol use could have widespread effects on cardiometabolic health. This is in contrast to prior large-scale GWAS (incorporating genome-wide variants) reporting negative genetic associations between alcohol consumption and *better* cardiometabolic health,[8,9,11,33] which may be partially attributed to biases, such as differences in socioeconomic status, measurement errors, or changes in alcohol consumption over time in the different cohorts.[9,19,96] Multivariable MR including additional variables, such as BMI and high blood pressure, or experimental analyses, such as those using animal models, may assist in further disentangling the role of alcohol consumption in cardiometabolic health.[81,84,93]

Our study also revealed broad implications of variations in ethanol metabolising enzyme genes for domains of health that are relatively unexplored. For

example, across all three ancestries, a large number of the PheWAS associations were with immune traits, particularly between rs1229984-T and various types of allergies. The nature of these associations is still unclear. While some epidemiological studies have found positive associations between alcohol consumption and allergies,[97–99] others have failed to identify significant associations,[99,100] or have only found associations with lifetime alcohol consumption, but not the quantity consumed,[97] or with levels of Immunoglobulin E (a type of antibody that plays a crucial role in the immune system), but not with allergic disease. Consistent with these findings, our MR analysis did not establish these associations as causal, indicating other common mechanisms whereby these variants influence the development and course of allergies. Beyond allergies, our PheWAS results revealed an association between rs1229984-T (protective allele) and other immune conditions, including lower incidence of rheumatoid arthritis, lupus, and many others. MR did not identify a causal relationship between alcohol consumption and most of these outcomes, consistent with prior studies presenting conflicting findings and challenging prior observational studies that may be more highly subjected to residual confounders, such as socioeconomic and lifestyle factors. Furthermore, our PheWAS identified associations with eye health (e.g., myopia, cataracts, age related macular degeneration), which is consistent with previous prospective observations[101–103] and MR studies identifying an association between genetically predicted alcohol consumption and risk of developing age-related macular degeneration. However, these results are discrepant with our and other MR studies showing inconsistent results between alcohol consumption and eye health,[104,105] emphasising the need for more in-depth studies exploring the effects of alcohol consumption on the visual system. Finally, we also identified associations between rs1229984-T and fractures in all three ancestries. This is perhaps unsurprising, as persistent alcohol consumption (along with many other risk factors such as smoking, low BMI) has been linked to higher odds of developing osteoporosis and bone fractures,[106–108] and to increased injuries as a result of heavy drinking.[109] Although a previous MR study identified an association between alcohol dependence (based on a single SNP, rs1229984) and fractures, our MR results were primarily null, suggesting that additional studies, particularly those using a broader range of SNPs and examining more severe alcohol-related phenotypes, are needed to draw more robust conclusions.

Variants in genes encoding ethanol metabolising enzymes are also known to affect the risk of developing cancer.[3,4,6,110,111–113] We identified *unexpected* positive associations between the protective alleles of ethanol metabolising enzyme variants that encode more active enzymes and higher rates for several types of cancer (e.g., skin and lung cancer) in the European cohort. It is not clear whether this association is mediated by alcohol use, by alcohol metabolism in non-hepatic tissues, by both mechanisms, or due to other issues. Our MR results identified one association between increased alcohol consumption and one type of cancer (skin cancer) but not others. These results suggest an indirect pathway explaining the effect of rs1229984 on cancer other than drinking, and replicate the mixed findings observed in prior MR studies for other cancer types.[110,114–117] Future studies exploring additional confounders (smoking, poor diet), increasing granularity in alcohol-behaviours (e.g., only including individuals drinking amounts higher than the recommended thresholds), and using higher numbers of cancer cases, can contribute to clarifying the pathogenic pathways by which alcohol contributes to carcinogenesis.

Lastly, our study discovered several ancestry-specific associations (e.g., endocrine/metabolic conditions only being reported in individuals of European and Latin American ancestry; mental health conditions primarily reported in individuals of European ancestry), many of which may simply reflect the difference in power. Another possible explanation for these differences could be attributed to environmental differences among participants from the European, Latin American, and African American cohorts,[118–120] such as education and socioeconomic factors (Supplementary Table S2) that may affect both drinking patterns and disease ascertainment.

Several limitations are also worth mentioning. First, there are hundreds to thousands of additional SNPs and other classes of genetic variation beyond the three SNPs that we studied–however, the three SNPs we tested have by far the highest effect sizes on alcohol consumption in the cohorts that we studied. Although we included diverse cohorts, our study lacked many major ancestry groups such as East Asians and South Asians, which were not available at the time of the analysis. Including those of East Asian ancestry would have allowed testing of rs671, a SNP in *ALDH2* that has the most profound effect of any studied SNP in alcohol metabolising enzymes.[3,21,31,32,35] Another important caveat is that our results may be biased by ascertainment and related characteristics of the sample. 23andMe participants are volunteers, and they tend to be older than the general population, with higher education and socioeconomic status, and low levels of alcohol consumption (Supplementary Table S2). Indeed, approximately 30% of individuals in this study reported no current alcohol drinking. Thus, this cohort includes both lifetime abstainers and former drinkers who may have reduced or quit drinking because of other detrimental outcomes, which are known to introduce biases in genetic studies.[19] This cohort composition may explain some of our paradoxical *inverse* relationships between alcohol-protective variants and increased liability for some

conditions (e.g., more lifetime tobacco use, cancer, Graves' disease). The mediating effects of age and alcohol exposure on the reported associations could be assessed in future studies, and therefore replication using younger populations,[121] longitudinal data[122] and cohorts with lower socioeconomic status and higher levels of alcohol consumption is warranted. Moreover, our PheWAS used minimal (self-reported) phenotyping to allow a large number of diverse phenotypes to be surveyed within a single study; in particular, alcohol phenotypes pertain to recent use at the time of the analyses and not lifetime use, which would be preferable. Although we have studied the largest diverse cohort to date, some traits were still underpowered for some SNPs (Supplementary Figure S5). Similarly, while we examined them separately, most conditions we studied are correlated. For example, the comorbidity between alcohol consumption and high blood pressure is necessary to consider when assessing the elevated likelihood of cardio-metabolic disease. Thus, despite our attempt to disentangle the mechanisms underlying pairs of comorbidities with problematic alcohol consumption, the risk may be better represented by a matrix of comorbidities with many shared genetic and non-genetic pathways. In this study, the genetic liability for alcohol consumption was defined by using three specific SNPs in the ADH region. MR is less prone to the bias of observational studies, particularly in relation to reverse causation and confounding, providing a more robust estimate of causal relationships between exposure and outcomes. Incorporating established instruments, particularly rs1229984,[3,4] provides an additional layer of support for the validity of our MR analysis; however, relying solely on three variants with modest effects on the magnitude of alcohol consumption can diminish predictive accuracy. Furthermore, although our MR results were robust to different assumptions and biases and to sensitivity analyses accounting for horizontal pleiotropy,[46] some confounding may remain (e.g., associations with physical characteristics). Because we are only testing one pathway related to ethanol metabolism, it is possible that other molecular causal pathways could play a significant role in these associations. Similarly, there may be other unknown or unmeasured confounding factors we have not accounted for in MR analyses, such as dynastic effects (e.g., parents with high alcohol consumption genetic propensity conferring some unrelated advantage/disadvantage on these outcomes) or assortative mating (genetic resemblance between spouses due to selecting each other based on genetically influenced traits).[123] Our results should therefore be replicated with even larger sample sizes (or by using multiple SNPs in combination to increase the variance explained in alcohol consumption) and employing additional study designs, such as within families, to counteract those potential confounders.

Conclusion

This study performs an in-depth exploration of the relationships between variants in alcohol metabolising enzyme genes and thousands of other traits across three large ancestral groups. We provide evidence that these loci contribute directly or indirectly to several leading causes of disability. Interventions that decrease alcohol consumption may be a promising therapeutic target for a constellation of other disorders. Future research using larger sample sizes and/or observational longitudinal studies in multiple ancestries may allow for further investigation into these complex associations.

Contributors

SSR conceived of the idea for the paper, with input from AAP. PF, MVJ, JJMM. and the 23andMe team contributed analyses and data. PF and SLE accessed and verified the original data. All contributing authors wrote and edited the paper, and approved the final version.

Data sharing statement

Data collection was not performed as part of this study. Summary statistics from the PheWAS and MR analyses are presented in Supplementary Tables S11–S15.

Declaration of interests

PF, SLE and members of 23andMe Research Team are employees of 23andMe, Inc., and hold stock or stock options in 23andMe. ASH has filed for a provisional Patent entitled "Methods of Treatments for Multi-Drug or Broad Addiction Liability", University patent application T-020489. ASH reports consulting fees from Psychiatric Genomics Consortium Suicide working group, speaker fees from Jackson Laboratories and a grant from National Institute on Alcoholism and Alcohol Abuse (NIAAA K01 AA030083-01). CK received a new investigator award from TRDRP and consulting fees and stock options from CARI Health, Inc. JG is paid for editorial work for the journal 'Complex Psychiatry' and is named as an inventor on PCT patent application #15/878,640 entitled: "Genotype-guided dosing of opioid agonists" filed January 24, 2018 and issued on January 26, 2021 as U.S. Patent No. 10,900,082. AP received consulting fees from universities and companies and holds a biomedical patent (related to inhibitions of the gene *Glo1*), both unrelated to this work. SSR received consulting fees from the Externalizing Consortium as well as a Royal Society Alcoholism Early Career Award honoraria. EA reports consulting fees from PGC-PTSD and LATINO consortium. All other authors declare no conflicts of interest.

Acknowledgements

We would like to thank the research participants and employees of 23andMe for making this work possible. This research was conducted as part of the 23andMe Research Innovation Collaboration program (research.23andme.com/collaborate/#research).

We would like to thank Brittany Leger for her assistance creating figures for this paper.

MVJ, NCK, SBB, SSR and AAP were supported by T32IR5226 and 28IR-0070. JG was supported by NIH grant (5P50AA012870-23) and USVA grant (5IO1CX001849-04). SSR was also supported by NIDA DP1DA054394. NCK and RBC were also supported by R25MH081482. ASH was supported by funds from NIAAA K01AA030083. JLMO was supported by VA 1IK2CX002095. JLMO and JJMM were also supported by NIDA R21DA050160. JJMM was also supported by the Kavli Postdoctoral Award for Academic Diversity. EGA was supported by K01MH121659 from the NIMH/NIH, the Caroline Wiess Law Fund for Research in Molecular Medicine and the ARCO Foundation Young Teacher-Investigator Fund at Baylor College of Medicine. MSA was supported by the Instituto de Salud Carlos III and co-funded by the European Union Found: Fondo Social Europeo Plus (FSE+) (P19/01224, PI22/00464 and CP22/00128). The content is solely the responsibility of the authors and

does not necessarily represent the official views of the National Institutes of Health.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2024.105086.

References

1 World Health Organization. *Global status report on alcohol and health 2018*. World Health Organization; 2018.
2 Zakhari S. Overview: how is alcohol metabolized by the body?. 2006;29:10.
3 Edenberg HJ, McClintick JN. Alcohol dehydrogenases, aldehyde dehydrogenases, and alcohol use disorders: a critical review. *Alcohol Clin Exp Res*. 2018;42:2281–2297.
4 Polimanti R, Gelernter J. *ADH1B*: from alcoholism, natural selection, and cancer to the human phenome. *Am J Med Genet B Neuropsychiatr Genet*. 2018;177:113–125.
5 Thomasson HR, Edenberg HJ, Crabb DW, et al. Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. *Am J Hum Genet*. 1991;48:677–681.
6 Hart AB, Kranzler HR. Alcohol dependence genetics: lessons learned from genome-wide association studies (GWAS) and post-GWAS analyses. *Alcohol Clin Exp Res*. 2015;39:1312–1327.
7 Walters RK, Polimanti R, Johnson EC, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018;21:1656–1669.
8 Sanchez-Roige S, Palmer AA, Fontanillas P, 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;176:107–118.
9 Mallard TT, Savage JE, Johnson EC, et al. Item-level genome-wide association study of the alcohol use disorders identification test in three population-based cohorts. *Am J Psychiatry*. 2021;179:58.
10 Gelernter J, Polimanti R. Genetics of substance use disorders in the era of big data. *Nat Rev Genet*. 2021;22:712–729. https://doi.org/10.1038/s41576-021-00377-1.
11 Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*. 2019;51:237–244.
12 Saunders GRB, Wang X, Chen F, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature*. 2022;612:720–724.
13 Kranzler HR, Zhou H, Kember RL, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat Commun*. 2019;10:1499.
14 Zhou H, Sealock JM, Sanchez-Roige S, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nat Neurosci*. 2020;23:809–818.
15 Gelernter J, Sun N, Polimanti R, 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;86:365–376.
16 Deak JD, Levey DF, Wendt FR, 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;5:e2238880.
17 Kember RL, Vickers-Smith R, Zhou H, 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;180:584–593.
18 Brazel DM, Jiang Y, Hughey JM, et al. Exome chip meta-analysis fine maps causal variants and elucidates the genetic architecture of rare coding variants in smoking and alcohol use. *Biol Psychiatry*. 2019;85:946–955.
19 Dao C, Zhou H, Small A, et al. The impact of removing former drinkers from genome-wide association studies of AUDIT-C. *Addiction*. 2021;116:3044–3054.
20 Gelernter J, Kranzler HR, Sherva R, et al. Genome-wide association study of alcohol dependence:significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry*. 2014;19:41–49.
21 Jorgenson E, Thai KK, Hoffmann TJ, et al. Genetic contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic genome-wide association study. *Mol Psychiatry*. 2017;22:1359–1367.
22 Kapoor M, Wang J-C, Wetherill L, et al. A meta-analysis of two genome-wide association studies to identify novel loci for maximum number of alcoholic drinks. *Hum Genet*. 2013;132:1141–1151.
23 Karlsson Linnér R, Biroli P, Kong E, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat Genet*. 2019;51:245–257.
24 Lai D, Wetherill L, Bertelsen S, et al. Genome-wide association studies of alcohol dependence, DSM-IV criterion count and individual criteria. *Genes Brain Behav*. 2019;18.
25 Matoba N, Akiyama M, Ishigaki K, et al. GWAS of 165,084 Japanese individuals identified nine loci associated with dietary habits. *Nat Hum Behav*. 2020;4:308–316.
26 Park BL, Kim JW, Cheong HS, et al. Extended genetic effects of ADH cluster genes on the risk of alcohol dependence: from GWAS to replication. *Hum Genet*. 2013;132:657–668.
27 Thompson A, Cook J, Choquet H, et al. Functional validity, role, and implications of heavy alcohol consumption genetic loci. *Sci Adv*. 2020;6:eaay5034.
28 Watanabe K, Stringer S, Frei O, et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet*. 2019;51:1339–1348.
29 Zhong VW, Kuang A, Danning RD, et al. A genome-wide association study of bitter and sweet beverage consumption. *Hum Mol Genet*. 2019;28:2449–2457.
30 Polimanti R, Kranzler HR, Gelernter J. Phenome-wide association study for alcohol and nicotine risk alleles in 26394 women. *Neuropsychopharmacology*. 2016;41:2688–2696.
31 Quillen EE, Chen X-D, Almasy L, 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;165:103–110.
32 Zhou H, Kalayasiri R, Sun Y, et al. Genome-wide meta-analysis of alcohol use disorder in East Asians. *Neuropsychopharmacology*. 2022;47:1791–1797.
33 Schumann G, Liu C, O'Reilly P, et al. *KLB* is associated with alcohol drinking, and its gene product β-Klotho is necessary for FGF21 regulation of alcohol preference. *Proc Natl Acad Sci*. 2016;113:14372–14377.
34 Sanchez-Roige S, Fontanillas P, Elson SL, et al. Genome-wide association study of alcohol use disorder identification test (AUDIT) scores in 20 328 research participants of European ancestry: GWAS of AUDIT. *Addict Biol*. 2019;24:121–131.
35 Yoshida A, Hsu LC, Yasunami M. Genetics of human alcohol-metabolizing enzymes. Cohn WE, Moldave K *Prog Nucleic Acid Res Mol Biol*. 1991;40:255–287.
36 Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47:D1005–D1012.
37 Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *Int J Epidemiol*. 2003;32:1–22.
38 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23:R89–R98.
39 Durand EY, Do CB, Mountain JL, Macpherson JM. Ancestry composition: a novel, efficient pipeline for ancestry deconvolution. *bioRxiv*. 2014. https://doi.org/10.1101/010512.
40 Ehm MG, Aponte JL, Chiano MN, et al. Phenome-wide association study using research participants' self-reported data provides insight into the Th17 and IL-17 pathway. *PLoS One*. 2017;12:e0186405.
41 Verma A, Bradford Y, Dudek S, et al. A simulation study investigating power estimates in phenome-wide association studies. *BMC Bioinf*. 2018;19:120.
42 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57:289–300.
43 Millard LAC, Munafò MR, Tilling K, Wootton RE, Davey Smith G. MR-pheWAS with stratification and interaction: searching for the causal effects of smoking heaviness identified an effect on facial aging. *PLoS Genet*. 2019;15:e1008353.
44 Shikov AE, Skitchenko RK, Predeus AV, Barbitoff YA. Phenome-wide functional dissection of pleiotropic effects highlights key molecular pathways for human complex traits. *Sci Rep*. 2020;10:1037.

45 Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37:658–665.
46 Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Methods Primers*. 2022;2:6.
47 Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018:k601.
48 Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*. 1993;88:791–804.
49 Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol*. 2017;46:1734–1739.
50 Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int J Epidemiol*. 2016;45:1961–1974.
51 Burgess S, Scott R, Timpson N, Davey Smith G, Thompson S. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. 2015;30.
52 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512–525.
53 Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40:304–314.
54 Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32:377–389.
55 Soler Artigas M, Sánchez-Mora C, Rovira P, Vilar-Ribó L, Ramos-Quiroga JA, Ribasés M. Mendelian randomization analysis for attention deficit/hyperactivity disorder: studying a broad range of exposures and outcomes. *Int J Epidemiol*. 2023;52:386–402.
56 Rasooly D, Patel CJ. Conducting a reproducible Mendelian randomization analysis using the R analytic statistical environment. *Curr Protoc Hum Genet*. 2019;101:e82.
57 Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016;45:1866–1886.
58 Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using Mendelian randomization: the STROBE-MR statement. *JAMA*. 2021;326:1614–1621.
59 Manthey J, Hassan SA, Carr S, Kilian C, Kuitunen-Paul S, Rehm J. What are the economic costs to society attributable to alcohol use? A systematic review and modelling study. *Pharmacoeconomics*. 2021;39:809–822.
60 Shield KD, Rehm J. Societal development and the alcohol-attributable burden of disease. *Addiction*. 2021;116:2326–2338.
61 Macgregor S, Lind PA, Bucholz KK, et al. Associations of ADH and ALDH2 gene variation with self report alcohol reactions, consumption and dependence: an integrated analysis. *Hum Mol Genet*. 2009;18:580–593.
62 Puddephatt J-A, Irizar P, Jones A, Gage SH, Goodwin L. Associations of common mental disorder with alcohol use in the adult general population: a systematic review and meta-analysis. *Addict Abingdon Engl*. 2022;117:1543–1572.
63 Treur JL, Munafò MR, Logtenberg E, Wiers RW, Verweij KJH. Using Mendelian randomization analysis to better understand the relationship between mental health and substance use: a systematic review. *Psychol Med*. 2021;51:1593–1624.
64 Pasman JA, Smit DJA, Kingma L, Vink JM, Treur JL, Verweij KJH. Causal relationships between substance use and insomnia. *Drug Alcohol Depend*. 2020;214:108151.
65 Castillo-Carniglia A, Keyes KM, Hasin DS, Cerdá M. Psychiatric comorbidities in alcohol use disorder. *Lancet Psychiatr*. 2019;6:1068–1080.
66 Koob GF, Colrain IM. Alcohol use disorder and sleep disturbances: a feed-forward allostatic framework. *Neuropsychopharmacology*. 2020;45:141–165.
67 Russo D, Purohit V, Foudin L, Salin M. Workshop on alcohol use and health disparities 2002: a call to arms. *Alcohol*. 2004;32:37–43.
68 Kranzler HR, Hartwell EE. Medications for treating alcohol use disorder: a narrative review. *Alcohol Clin Exp Res*. 2023;47:1224–1237.
69 Chaudhary NS, Wong MM, Kolla BP, Kampman KM, Chakravorty S. The relationship between insomnia and the intensity of drinking in treatment-seeking individuals with alcohol dependence. *Drug Alcohol Depend*. 2020;215:108189.
70 Kolla BP, Mansukhani MP, Biernacka J, Chakravorty S, Karpyak VM. Sleep disturbances in early alcohol recovery: prevalence and associations with clinical characteristics and severity of alcohol consumption. *Drug Alcohol Depend*. 2020;206:107655.
71 Marmura MJ. Triggers, protectors, and predictors in episodic migraine. *Curr Pain Headache Rep*. 2018;22:81.
72 Panconesi A, Bartolozzi ML, Guidi L. Alcohol and migraine: what should we tell patients? *Curr Pain Headache Rep*. 2011;15:177–184.
73 Garcia-Martin E, Martinez C, Serrador M, et al. Alcohol dehydrogenase 2 genotype and risk for migraine. *Headache J Head Face Pain*. 2010;50:85–91.
74 Yuan S, Daghlas I, Larsson SC. Alcohol, coffee consumption, and smoking in relation to migraine: a bidirectional Mendelian randomization study. *Pain*. 2022;163:e342–e348.
75 Yuan S, Chen J, Ruan X, et al. Smoking, alcohol consumption, and 24 gastrointestinal diseases: Mendelian randomization analysis. *Elife*. 2023;12:e84051.
76 Yu C, Tang H, Guo Y, et al. Hot tea consumption and its interactions with alcohol and tobacco use on the risk for esophageal cancer: a population-based cohort study. *Ann Intern Med*. 2018;168:489–497.
77 Samokhvalov AV, Rehm J, Roerecke M. Alcohol consumption as a risk factor for acute and chronic pancreatitis: a systematic review and a series of meta-analyses. *eBioMedicine*. 2015;2:1996–2002.
78 Younossi ZM, Noureddin M, Bernstein D, et al. Role of noninvasive tests in clinical gastroenterology practices to identify patients with nonalcoholic steatohepatitis at high risk of adverse outcomes: expert panel recommendations. *Am J Gastroenterol*. 2021;116:254–262.
79 Lieberoth S, Backer V, Kyvik KO, et al. Intake of alcohol and risk of adult-onset asthma. *Respir Med*. 2012;106:184–188.
80 Choudhary NS, Duseja A. Genetic and epigenetic disease modifiers: non-alcoholic fatty liver disease (NAFLD) and alcoholic liver disease (ALD). *Transl Gastroenterol Hepatol*. 2021;6:2.
81 Holmes MV, Dale CE, Zuccolo L, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. *BMJ*. 2014;349:g4164.
82 Ding C, O'Neill D, Britton A. Trajectories of alcohol consumption in relation to all-cause mortality in patients with cardiovascular disease: a 35-year prospective cohort study. *Addiction*. 2022;117:1920–1930.
83 Tews D, Schwar V, Scheithauer M, et al. Comparative gene array analysis of progenitor cells from human paired deep neck and subcutaneous adipose tissue. *Mol Cell Endocrinol*. 2014;395:41–50.
84 Winnier DA, Fourcaudot M, Norton L, et al. Transcriptomic identification of ADH1B as a novel candidate gene for obesity and insulin resistance in human adipose tissue in Mexican Americans from the veterans administration genetic epidemiology study (VAGES). *PLoS One*. 2015;10:e0119941.
85 Di L, Balesano A, Jordan S, Shi SM. The role of alcohol dehydrogenase in drug metabolism: beyond ethanol oxidation. *AAPS J*. 2021;23:20. https://doi.org/10.1208/s12248-020-00536-y.
86 Beulens JWJ, Rimm EB, Hendriks HFJ, et al. Alcohol consumption and type 2 diabetes. *Diabetes*. 2007;56:2388–2394.
87 Hotamisligil GS, Bernlohr DA. Metabolic functions of FABPs—mechanisms and therapeutic implications. *Nat Rev Endocrinol*. 2015;11:592–605.
88 Hyun J, Han J, Lee C, Yoon M, Jung Y. Pathophysiological aspects of alcohol metabolism in the liver. *Int J Mol Sci*. 2021;22:5717.
89 Kema VH, Mojerla NR, Khan I, Mandal P. Effect of alcohol on adipose tissue: a review on ethanol mediated adipose tissue injury. *Adipocyte*. 2015;4:225–231.
90 Lawlor DA, Benn M, Zuccolo L, et al. ADH1B and ADH1C genotype, alcohol consumption and biomarkers of liver function: findings from a Mendelian randomization study in 58,313 European Origin Danes. *PLoS One*. 2014;9:e114294.
91 Thumser AE, Moore JB, Plant NJ. Fatty acid binding proteins: tissue-specific functions in health and disease. *Curr Opin Clin Nutr Metab Care*. 2014;17:124–129.
92 Iqbal J, Parks JS, Hussain MM. Lipid absorption defects in intestine-specific microsomal triglyceride transfer protein and ATP-

binding cassette transporter A1-deficient mice. *J Biol Chem*. 2013;288:30432–30444.
93 Lankester J, Zanetti D, Ingelsson E, Assimes TL. Alcohol use and cardiometabolic risk in the UK Biobank: a Mendelian randomization study. *PLoS One*. 2021;16:e0255801.
94 Celorrio D, Bujanda L, Chbel F, Sánchez D, Martinez-Jarreta B, de Pancorbo MM. Alcohol-metabolizing enzyme gene polymorphisms in the Basque country, Morocco, and Ecuador: adh in Basque country, Morocco & Ecuador. *Alcohol Clin Exp Res*. 2011;35:879–884.
95 Traversy G, Chaput J-P. Alcohol consumption and obesity: an update. *Curr Obes Rep*. 2015;4:122–130.
96 Xue A, Jiang L, Zhu Z, et al. Genome-wide analyses of behavioural traits are subject to bias by misreports and longitudinal changes. *Nat Commun*. 2021;12:20211.
97 Skaaby T, Kilpeläinen TO, Taylor AE, et al. Association of alcohol consumption with allergic disease and asthma: a multi-centre Mendelian randomization analysis: alcohol consumption, hay fever and asthma. *Addiction*. 2019;114:216–225.
98 Vidal C, Armisen M, Dominguez-Santalla MJ, Gude F, Lojo S, Gonzalez-Quintela A. Influence of alcohol consumption on serum immunoglobulin E levels in atopic and nonatopic adults. *Alcohol Clin Exp Res*. 2002;26:59–64.
99 Linneberg A, Friedrich N, Husemoen LLN, et al. Incidence and remission of specific IgE aeroallergen sensitization from age of 40 to 60 Years, and association with alcohol consumption. *Int Arch Allergy Immunol*. 2010;151:142–148.
100 Assing K, Bodtger U, Linneberg A, Malling HJ, Poulsen LK. Association between alcohol consumption and skin prick test reactivity to aeroallergens. *Ann Allergy Asthma Immunol*. 2007;98:70–74.
101 Adams MKM, Chong EW, Williamson E, et al. 20/20–Alcohol and age-related macular degeneration: the Melbourne Collaborative Cohort Study. *Am J Epidemiol*. 2012;176:289–298.
102 Chong EW-T, Kreis AJ, Wong TY, Simpson JA, Guymer RH. Alcohol consumption and the risk of age-related macular degeneration: a systematic review and meta-analysis. *Am J Ophthalmol*. 2008;145:707–715.
103 Dinu M, Pagliai G, Casini A, Sofi F. Food groups and risk of age-related macular degeneration: a systematic review with meta-analysis. *Eur J Nutr*. 2019;58:2123–2143.
104 Jiang N, Lee YO, Ling PM. Association between tobacco and alcohol use among young adult bar patrons: a cross-sectional study in three cities. *BMC Publ Health*. 2014;14:500.
105 Kanclerz P, Hecht I, Tuuminen R. Is occasional alcohol drinking and smoking related to the development of age-related cataract? *Invest Ophthalmol Vis Sci*. 2023;64:8.
106 Xu X-H, Dong S-S, Guo Y, et al. Molecular genetic studies of gene identification for osteoporosis: the 2009 update. *Endocr Rev*. 2010;31:447–505.
107 De Laet C, Kanis JA, Odén A, et al. Body mass index as a predictor of fracture risk: a meta-analysis. *Osteoporos Int*. 2005;16:1330–1338.
108 Kanis JA, Johansson H, Johnell O, et al. Alcohol intake as a risk factor for fracture. *Osteoporos Int*. 2005;16:737–742.
109 Chikritzhs T, Livingston M. Alcohol and the risk of injury. *Nutrients*. 2021;13:2777.
110 Zhang B, Peng Y-H, Luo Y, et al. Relationship between esophageal squamous cell carcinoma risk and alcohol-related ALDH2 and ADH1B polymorphisms: evidence from a meta-analysis and Mendelian randomization analysis. *Cancer Med*. 2023;12:20437–20449.
111 Hashibe M, McKay JD, Curado MP, et al. Multiple ADH genes are associated with upper aerodigestive cancers. *Nat Genet*. 2008;40:707–709.
112 McKay JD, Truong T, Gaborieau V, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE Consortium. *PLoS Genet*. 2011;7:e1001333.
113 Wu C, Kraft P, Zhai K, et al. Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. *Nat Genet*. 2012;44:1090–1097.
114 Chien H-T, Tsai C-L, Young C-K, et al. Single-nucleotide polymorphism at alcohol dehydrogenase 1B: a susceptible gene marker in oro-/hypopharyngeal cancers from genome-wide association study. *Cancer Med*. 2023;12:19174–19187.
115 Im PK, Yang L, Kartsonaki C, et al. Alcohol metabolism genes and risks of site-specific cancers in Chinese adults: an 11-year prospective study. *Int J Cancer*. 2022;150:1627–1639.
116 Chida K, Oshi M, Roy AM, Sato T, Endo I, Takabe K. Pancreatic ductal adenocarcinoma with a high expression of alcohol dehydrogenase 1B is associated with less aggressive features and a favorable prognosis. *Am J Cancer Res*. 2023;13:3638–3649.
117 Sugimoto Y, Koyanagi YN, Kawakatsu Y, et al. Mediation analysis unveils a carcinogenic effect of ADH1B rs1229984 through mechanisms other than change in drinking intensity: oesophageal cancer case-control study. *Jpn J Clin Oncol*. 2023;53:581–588.
118 Subramanian S. The effects of sample size on population genomic analyses – implications for the tests of neutrality. *BMC Genom*. 2016;17:123.
119 Harpak A, Przeworski M. The evolution of group differences in changing environments. *PLoS Biol*. 2021;19:e3001072.
120 Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, Przeworski M. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife*. 2020;9:e48376.
121 Karcher NR, Barch DM. The ABCD study: understanding the development of risk for mental and physical health outcomes. *Neuropsychopharmacology*. 2021;46:131–142.
122 Saunders GRB, McGue M, Iacono WG, Vrieze S. Longitudinal effects and environmental moderation of ALDH2 and ADH1B gene variants on substance use from age 14 to 40. *Dev Psychopathol*. 2022:1–9.
123 Howe LJ, Lawson DJ, Davies NM, et al. Genetic evidence for assortative mating on alcohol consumption in the UK Biobank. *Nat Commun*. 2019;10:5039.