Refining the scope of genetic influences on alcohol misuse through environmental stratification and gene-environment interaction

Jeanne E. Savage, PhD¹, Christiaan A. de Leeuw, PhD¹, Josefin Werne, MSc¹, Spit for Science Working Group, Danielle M. Dick, PhD⁷, Danielle Posthuma, PhD¹,³, Sophie van der Sluis, PhD³

¹. Department of Complex Trait Genetics, Centre for Neurogenomics and Cognitive Research, VU University, Amsterdam, The Netherlands

². Department of Psychiatry, Robert Wood Johnson Medical School, Rutgers Addiction Research Center, Rutgers University, Piscataway, NJ, USA

³. Department of Child and Adolescent Psychology and Psychiatry, section Complex Trait Genetics, Amsterdam Neuroscience, VU University Medical Center, Amsterdam, The Netherlands


*Correspondence concerning this article should be addressed to Jeanne E. Savage de Boelelaan 1085 1081HV Amsterdam Netherlands Phone: +31 20 598 2832 Fax: +31 20 598 6926 Email: j.e.savage@vu.nl

Acknowledgements: This research was funded by a grant to J.E.S. (VI.VENI.201G-064) from The Netherlands Organization for Scientific Research (NWO) and has been conducted using the UK Biobank Resource (application no. 16406). We would like to thank the many UK Biobank participants and staff. Analyses were carried out on the Genetic Cluster Computer hosted by the Dutch National computing and Networking Services SURFsara. Spit for Science has been supported by Virginia Commonwealth University, P20AA017828, R37AA011408, K02AA018755, P50AA022537, and K01AA024152 from the National Institute on Alcohol Abuse and Alcoholism, and UL1RR031990 from the National Center for...
Research Resources and National Institutes of Health Roadmap for Medical Research. This research was also supported by the Center for the Study of Tobacco Products at Virginia Commonwealth University. The content is solely the responsibility of the authors and does not necessarily represent the views of the funding agencies. The funding agencies had no role in the study design, data analysis, manuscript preparation, or decision to submit for publication. We would like to thank the Spit for Science participants for making this study a success, as well as the many University faculty, students, and staff who contributed to the design and implementation of the project.

**Competing interests:** The authors report no competing interests.

**Data availability:** Genome-wide summary statistics will be made publicly available via [https://ctg.cnclr.nl/software/summary_statistics/](https://ctg.cnclr.nl/software/summary_statistics/) upon publication. Raw data from this study are available to qualified researchers via UK Biobank ([https://www.ukbiobank.ac.uk/](https://www.ukbiobank.ac.uk/)) and dbGaP (phs001754.v4.p2).
Abstract

**Background**: Gene-environment interaction (GxE) is likely an important influence shaping individual differences in alcohol misuse (AM), yet it has not been extensively studied in molecular genetic research. In this study, we utilize a series of genome-wide gene-environment interaction (GWEIS) and *in silico* annotation methods with the aim of improving gene identification and biological understanding of AM.

**Methods**: We carried out GWEIS for four AM phenotypes in the large UK Biobank sample (*N* = 360,314), with trauma exposure and socioeconomic status (SES) as moderators of the genetic effects. Exploratory analyses compared stratified GWAS and GWEIS modelling approaches. We applied functional annotation, gene- and gene-set enrichment, and polygenic score analyses to interpret the GWEIS results.

**Results**: GWEIS models showed few genetic variants with significant interaction effects across all gene-environment pairs. Enrichment analyses identified moderation by SES of the genes *NOXA1*, *DLGAP1*, and *UBE2L3*, on drinking quantity and the gene *IFIT1B* on drinking frequency. Except for *DLGAP1*, these genes have not previously been linked to AM. The most robust results (GWEIS interaction *p* = 4.59e-09) were seen for SES moderating the effects of variants linked to immune-related genes on a pattern of drinking with versus without meals.

**Conclusions**: Even in large samples, GxE effects are difficult to detect at the molecular level. Our results highlight several genes and a potential mechanism of immune system functioning behind the moderating effect of SES on the genetic influences on AM. While GWEIS seems to be a preferred approach over stratified GWAS, modelling molecular GxE effects remains a challenge that will require larger consortia and more in-depth phenotypic measurement.

**Keywords**: genome-wide environment interaction, stratified GWAS, alcohol, trauma, SES
Introduction

Alcohol misuse (AM) involves multiple dimensions of heavy or problematic alcohol use behaviors that often result in clinical consequences, leading to a major global impact on human health and wellbeing (World Health Organization, 2018). AM has many varied manifestations, but all are strongly influenced by genetics (Dick et al., 2011, Verhulst et al., 2015). However, these genetic influences are subject to moderation by numerous cultural, environmental, and developmental circumstances (Savage et al., 2018b, Young-Wolff et al., 2011, Dick and Kendler, 2012) in a phenomenon known as gene-environment interaction (G×E).

Recent progress in the field through large-scale data collection has begun to reveal the specific genes linked to AM (Kranzler et al., 2019, Liu et al., 2019, Zhou et al., 2020). Such studies typically use a genome-wide association (GWAS) approach to agnostically scan the genome for a statistical association between millions of genetic variants (single nucleotide polymorphisms or SNPs) and a measured phenotype such as quantity of alcohol consumption or alcohol use disorder (AUD) diagnoses. However, much of the ~50% heritability of AM estimated by twin studies (Verhulst et al., 2015) has not yet been accounted for. Explanations for this “missing” heritability include hypotheses that the SNP effect sizes are too small to detect with the current levels of statistical power, that the heritability resides in rare variants that are not easily genotyped with current microarray-based study designs, or that SNP effects are not well characterized because they differ across groups of individuals or environments (i.e., G×E). The field has thus far primarily focused on addressing the first two hypotheses, by increasing the number of participants in GWAS efforts – now frequently including over a million individuals – and expanding whole genome sequencing technologies. However, there have been relatively few attempts to consider other factors like G×E, even though this can readily be investigated with existing data and methodology.
A long history of twin research has demonstrated that genes generally have a stronger influence on AM in contexts in which there are fewer restrictions on alcohol consumption, particularly social restrictions (whether explicit or implicit). For example, the impact of genes is enhanced in countries with low taxes on alcohol sales and in relationships with heavy drinking partners, and decreased in more restrictive contexts such as social groups with a high proportion of drinking abstention and among youth whose access to alcohol is controlled (Dick and Kendler, 2012, Savage et al., 2018b). These moderation effects also appear to differ across the specific dimensions of AM, with, for example, stronger genetic influences on alcohol consumption but weaker genetic influences on AUDs under conditions of high socioeconomic status (SES) (Davis and Slutske, 2018, Barr et al., 2018, Pasman et al., 2020).

Strong links between trauma exposure and alcohol in the twin literature (Magnusson et al., 2012) prompted early molecular G×E research which showed that trauma moderates the effects of candidate genes involved in stress response (Baranger et al., 2016, Lieberman et al., 2016) and alcohol metabolism (Sartor et al., 2014). More recently, Polimanti et al. (2018) demonstrated that such moderation effects are detectable in a genome-wide gene-environment interaction (GWEIS) design. A replication study showed the gene-by-trauma effects to be specifically associated with AUD symptoms and not with alcohol consumption (Hawn et al., 2018).

Interaction effects are challenging to detect in molecular genetic studies because their effect sizes are typically an order of magnitude smaller than the already small main effects of SNPs (McClelland and Judd, 1993). However, a GWEIS approach can be fruitful with a sufficiently well-powered sample (Werme et al., 2021, Wendt et al., 2021). Even when individual SNP effect sizes are too small to detect, GWEIS can allow for insight into the underlying biology of a trait through bioinformatics annotation and follow-up methods such as gene-based and gene-set analysis. Approaches using polygenic scores (PGS), which...
aggregate genetic “risk” across many SNPs, have also demonstrated that molecular G×E functions largely the same way as implied mathematically by twin studies. That is, PGS indexing higher risk for AM tend to show stronger associations in more permissive contexts, such as belonging to deviant peer groups (presenting issues with conduct and antisocial behavior), and weaker effects in more restrictive contexts, such as under high parental involvement (Salvatore et al., 2014, Barr et al., 2017). Like twin-studies, PGS do not reveal the specific genes involved or mechanisms of genetic effect, but the results from such studies echo the notion that individual SNP G×E effects are relevant and waiting to be uncovered.

In an alternative approach to GWEIS, recent studies have shown that stratified GWAS can also improve gene discovery and relevance of genetic associations by splitting the sample into subgroups based on known environmental exposures. For example, a study stratifying groups based on experience of childhood sexual abuse, a major risk factor for depression, uncovered a significant genetic association for depression that was specific to the no-abuse group (Peterson et al., 2017). More generally, studies that subset cases based on known or suspected etiological differences (e.g., early- vs. late-onset Alzheimer’s disease (Kamboh, 2022); typical vs. atypical symptom profiles of depression (Milaneschi et al., 2017); familial vs. complex forms of disorders (Hinney et al., 2010)) can generate meaningful disease insights. This approach takes into account the genetic heterogeneity that is widespread in complex phenotypes, including AM, in which different sets of genetic and/or environmental risk and protective factors are relevant for different individuals. Although statistical power is reduced when stratifying the sample due to the smaller numbers in each subgroup, this is expected to be offset by increased homogeneity, and therefore stronger genetic effects.

To our knowledge, only a few GWEIS and no stratified GWAS have thus far been conducted with AM phenotypes, despite the extensively observed phenomenon of G×E in the twin literature. A stratified GWAS approach fundamentally examines the same question as a
modeled G×E approach, but without a formal statistical test comparing whether the genetic
effect sizes differ between groups. As both study designs are currently being used in the field,
it is important to know which approach is best suited to the aims of gene identification and
interpretation for AM. The primary goal of the current study was thus to apply molecular
approaches to examine G×E for multiple understudied AM phenotypes in order to uncover
genetic variants, genes, and biological pathways that have differential effects on AM given
environmental influences. As a secondary exploratory aim, we compared multiple current
approaches for implementing G×E on a genome-wide level in order to inform our choice of
analysis for the main research question. By incorporating known environmental moderators
of the genetic effects on AM, we hope to improve gene discovery and increase the clinical
relevance of the identified genetic associations.

Materials and Methods

Measures

Data was obtained from the UK Biobank (Bycroft et al., 2018), a population-based
sample of adults in the UK with self-report surveys, linked electronic health records, and
genotypic data from imputed genome-wide microarrays. The National Research Ethics
Service Committee North West–Haydock ethically approved this initiative (reference
11/NW/0382) and participants provided informed written consent. Data were accessed under
application #16406.

Genetic analyses in the current study included up to 360,314 unrelated participants of
European ancestry with non-missing data. We selected four AM phenotypes for analysis
(Table 1): a broad indicator of lifetime AUD (BroadAUD), frequency and quantity of alcohol
consumption (FreqAlc and QuantAlc), and an indicator of whether alcohol is usually
consumed with meals (DrinkMeals). Recent GWAS of these phenotypes (Savage et al., 2023)
demonstrated that they represent distinct, clinically relevant dimensions of AM and have
robust SNP-based heritability, which is needed to ensure sufficient statistical power for stratified GWAS and G×E analyses. We also selected two robust environmental factors known to moderate the genetic effects on alcohol consumption and problems, respectively: SES, indexed for all participants with the Townsend Deprivation Index (Townsend et al., 1988), and lifetime exposure to traumatic events, measured by self-report in an online survey administered to approximately 157,000 participants (Table 1).

Data Analysis

Based on the twin and epidemiological literature, we chose to investigate five gene-environment pairs: BroadAUD with trauma, QuantAlc with SES, FreqAlc with SES, and DrinkMeals with both trauma and SES. This is because G×E effects have been shown between alcohol problems and trauma and between consumption and SES, as described above, while DrinkMeals is a measure recently shown to be part of a genetically distinct dimension of AM (Savage et al., 2023), but which has not been studied in a G×E context.

Table 2 provides an overview of the analyses performed and the measures involved. The purpose of this pre-selection of environments and phenotypes, as opposed to analyzing all possible combinations, was to reduce multiple testing and avoid data mining.

G×E model selection.

Prior to the primary analyses, we explored multiple ways of examining gene-environment interaction that are often employed in this type of research. Each method seeks to answer the same question (i.e., “are genetic effects equal across levels of an environmental moderator?”) but with different strengths and weaknesses.

(1) Stratified GWAS: splits the sample into two groups based on a dichotomous environmental exposure and runs a separate GWAS in each group. The results are easily interpretable when there is a significant association in one group but not the other and it is a very simple analysis to carry out. However, there is no formal test of
equality across groups to demonstrate whether an effect that is significantly different
from zero in one group is also significantly different from the effect in the other
group. As such, this is not a true test of G×E and may reflect other influences such as
power differences between groups.

(2) GWEIS

a. **GWEIS-individual test**: Regression models with an interaction provide a
formal test of G×E by means of the regression coefficient and associated p-
value for the both the main effect term and interaction term, individually.
These models are easy to apply in current genetic analysis software but can be
difficult to interpret, especially for continuous moderators. There may also be
reduced statistical power due to the many covariate-by-SNP and covariate-by-
environment interaction terms that need to be included in the model to avoid
spurious results (Dick et al., 2015).

b. **GWEIS-joint test**: Regression models with a joint test of significance of the
SNP main and interaction effects (rather than only the interaction term) have
most of the same strengths and weaknesses as (2a). However, by using a 2
degree of freedom F test with the (usually larger) main effects, they have more
statistical power to identify a relevant SNP that may be involved in an
interaction, but cannot determine whether identified associations are due to G
or G×E specifically (Winham and Biernacka, 2013).

We examined each of these 3 models using BroadAUD and trauma as a test case. For
the stratified GWAS (1), the sample was split into subgroups with (n=78,208) or without
(n=48,383) a lifetime experience of any type of trauma, and a GWAS of BroadAUD was run
within each subgroup. For the GWEIS models (2a & 2b), the trauma measure was treated as a
continuous burden score (log+1 transformed to reduce skew) summed across the individual
traumatic events. As the aim of this exploratory analysis was to maximize gene discovery, we compared the number of genome-wide significant \((p < 5e-8)\) and suggestive \((p < 5e-6)\) loci detected between analyses and the strength of associations present. To evaluate the overall concordance between the methods, we estimated genetic correlations using the results from each approach using linkage disequilibrium score regression (LDSC) (Bulik-Sullivan et al., 2015). Genetic correlations were calculated using default software settings and European ancestry LD scores provided with the software. As the GWEIS-joint F-test does not have an associated signed effect direction, we examined Spearman’s correlations between the rank of SNP \(p\)-values from the GWEIS-joint and each other method. We also carried out gene-set enrichment and polygenic score analyses, described below, to interpret and compare the results from the different models. Here, we aim to see whether adding environmental information can tell us more about the underlying biology of AM and whether convergent results are seen between different models. As there is currently no gold standard among the multiple approaches being used to investigate \(G \times E\), we qualitatively compare the gene discovery success and interpretability of results to select the model used for the main analysis of the other AM phenotypes.

**Genetic association analyses**

All stratified GWAS and GWEIS analyses were carried out using the generalized linear model \(\text{\texttt{--glm}}\) in PLINK2 (Chang et al., 2015) and controlled for sex, genotyping array, and within-ancestry principal components as covariates. Continuous measures (QuantAlc, FreqAlc, SES, and trauma count) were standardized prior to interaction analyses. All interaction models included full SNP\(\times\)covariate and environment\(\times\)covariate terms to avoid potential confounding effects (Dick et al., 2015). Full details of the genotyping, quality control, ancestry assignment, and general analysis pipeline have been described previously (Savage et al., 2018a). As \(G \times E\) analyses can be prone to inflated type I error rates, we applied
a sandwich estimator to obtain robust standard errors and adjusted \( p \)-values for all genome-wide significant loci identified in the main analyses. These analyses were carried out in R version 4.2.2 (R Core Team, 2015) using the method described by Werme et al. (2021) and the \textit{sandwich} package (Zeileis, 2006).

GWAS summary statistics were uploaded to the online platform FUMA (Watanabe et al., 2017) for annotation, gene mapping, and enrichment testing of the associated loci. We used default settings to define associated loci and candidate SNPs based on linkage disequilibrium (LD; \( r^2 > .6 \)) with an independently significant (uncorrected \( p < 5e-8 \)) lead SNP, and then annotated the function of those candidate SNPs from bioinformatics databases.

Genes were mapped from candidate SNPs based on their position within the gene, statistical links to expression levels of the gene (eQTLs), or 3-dimensional chromatin interactions between the SNP and the promoter region of the gene. We annotated genes in all significant loci (uncorrected \( p < 5e-8 \)), but consider individual SNPs to be credible only if the robust \( p \)-value was significant after Bonferroni correction for 5 GWASs (robust \( p < 1e-8 \)).

FUMA was also used to implement gene-based and gene-set testing with MAGMA (de Leeuw et al., 2015). This tests for the enrichment of genome-wide SNP association signal, in aggregate, within individual genes and sets of genes belonging to known biological pathways in the Molecular Signatures Database (curated and gene ontology sets; Liberzon et al., 2011) or with tissue-specific gene expression patterns (GTEx Consortium, 2015). When significant loci were identified, we also used the GENE2FUNC analysis in FUMA to test whether the genes mapped from these loci were similarly enriched in biological pathways and tissues. Per analysis, Bonferroni correction for the number of genes (19,628), gene-sets (10,678), or tissues (54) was applied.

\textit{Polygenic risk score (PRS) prediction}
To test replicability of the GWEIS results, we used the summary statistics from these analyses to calculate interaction-based polygenic scores (iPRSG\textsuperscript{GE}, as described in Werme et al. (2021)) in an independent sample. For any given SNP and environment, the number of effect alleles each individual carried was multiplied by their observed level of the environmental moderator and by the regression coefficient for the GWEIS SNP\times environment interaction. These values were summed across all SNPs to create a single individual-level score, which was then used to predict individual differences in AM measures.

iPRSG\textsuperscript{GE} were calculated in “Spit for Science” (S4S) (Dick et al., 2014), a longitudinal study of college students at a large, urban, public university. The study enrolled multiple cohorts of incoming students who completed yearly self-report surveys and provided a DNA sample. All participants provided informed consent and the S4S study was approved by the university Institutional Review Board. Data was collected and managed by the secure, web-based REDCap system of electronic data capture tools (Harris et al., 2009). The current study used data from $N = 3,524$ S4S participants for whom genotyping and four years of data collection was complete. Samples were restricted to those of European ancestry passing genetic quality control thresholds (Webb et al., 2017, Dick et al., 2014, Peterson et al., 2017).

Alcohol use measures assessed in the yearly surveys included questions about typical drinking frequency (number of drinking days per month, averaged across time), typical drinking quantity (gm ethanol per month, following the calculations of (Salvatore et al., 2016) and averaged across time), and DSM-5 AUD from the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA; Bucholz et al., 1994). A dichotomous BroadAUD diagnosis was derived based on meeting criteria for any level of AUD (2+ clustered symptoms in the past-year) at any assessment wave. A lifetime count of exposure to traumatic events such as an assault or natural disaster was measured by the Life Events Checklist (Gray...
et al., 2004). There was no direct measure of SES available, but we used a proxy measure based on the sum of mother’s and father’s years of education.

We derived $iPRSG \times E$ in S4S for each of the gene-environment pairs for which GWEIS was applied in UKB, as well as standard PRS from the SNP main effects in the same model. GWEIS summary statistics were weighted using PRS-CS “auto” version and the UKB LD reference panel provided with the software (Ge et al., 2019). Scores were derived in PLINK2 using the –score method. SNPs were first filtered on imputation INFO score > .8, MAF > .05, and missingness < .025. All continuous measures were standardized prior to score creation and analysis. Individual $iPRSG \times E$ were used in a linear or logistic regression model in R to predict their corresponding alcohol measures (FreqAlc, QuantAlc, BroadAUD). There was no corresponding measure of DrinkMeals in S4S, so $iPRSG \times E$ from the two GWEIS of DrinkMeals were used to predict the other three alcohol measures. All regression models included 10 within-ancestry PCs, sex, and age as covariates, as well as the main effects of the environmental moderator and main effects PRS. Explained variance was calculated as the change in (Nagelkerke’s) $R^2$ between a covariate-only and a covariate + $iPRSG \times E$ model.

To facilitate model comparison, we additionally created standard PRS from the stratified trauma/no trauma GWAS of BroadAUD. The procedure described above was used to create and test these PRS except that no environmental measure was included in the score calculation. Stratified PRS were used to predict BroadAUD in subgroups of S4S participants with and without lifetime trauma exposure, respectively. As there were no signed effect sizes produced by the GWEIS-joint model, it was not possible to derive comparative PRS with this model. Bonferroni correction was applied for the 11 models tested.

**Results**

*GxE model selection*
Exploratory analysis of BroadAUD with trauma indicated that the difference between different model implementations was small, but more significant loci \( (p < 5\times10^{-8}) \) were identified in GWEIS than in stratified GWAS (Figure 1; 1 vs. 0 loci). The same single locus (chr7: 117544507-117595829, near the gene \textit{CTTNBP2}) was significant in both the GWEIS-individual and GWEIS-joint approaches (Figure 1c; 1e). However, the fact that this locus was significant in the GWEIS-individual main effect, while showing no signal in the GWEIS-individual interaction effect (min. \( p = 0.23 \)) suggests that the joint test prioritizes genomic regions that are not necessarily enriched for environmental moderation at even a sub-threshold level. Additionally, several suggestive loci \( (p < 5\times10^{-6}) \) with associations specific to the GWEIS-individual interaction were seen (Figure 1d), which were not apparent in the GWEIS-joint model.

The mean \( \chi^2 \) statistic, indicative of the strength of association signal identified, was similar across methods (stratified trauma group: 1.043, stratified no trauma group: 1.028, GWEIS-individual main effect: 1.016, GWEIS-individual interaction effect: 1.039, GWEIS-joint: 1.038). Genetic correlations between the results of the different modelling approaches are shown in Table 3. Genetic effects were largely similar between the stratified groups \( (r_g = .83) \) and the stratified GWAS effects were mostly captured by the GWEIS-individual SNP main effects \( (r_g = .46 - .90) \). The GWEIS-joint signal largely overlapped with the GWEIS-individual main and interaction effects \( (r_g = .64 - .65) \). Gene-set enrichment analysis revealed no significant genes, gene-sets, or tissues for any of the models after Bonferroni correction for multiple testing. Polygenic score analyses showed that, after correcting for 11 models tested \( (\alpha = .05/11 = .0045) \), neither iPRS\(^{G\times E}\) from the GWEIS-individual model nor PRS from the stratified GWAS predicted BroadAUD in an independent sample (Table 4).

Overall, comparison of the different approaches did not point to any one model providing substantially greater biological insight than another, but modeled \( G\times E \) had an edge.
in gene identification over stratified GWAS. As the GWEIS-individual approach allows for both a statistical test of the interaction and discrimination between the SNP main and interaction effects, we selected this model for all further analyses and refer to it simply as “GWEIS” henceforth.

Genetic association analyses

For BroadAUD (Figure 1c-d), the GWEIS identified no significant loci for the SNP×trauma interaction effect, although one intergenic locus on chromosome 12 (chr12:92736886-92775269) nearly reached genome-wide significance with a p-value of 8.02e-08 (robust p-value = 4.15e-08). Using positional, eQTL, and chromatin interaction mapping, this locus was linked to 10 genes of which CLLU1 and CLLU1OS were the most proximal (Supplementary Table 1). There were no significant gene-based or gene-set-based associations.

For QuantAlc, significant SNP×SES interaction effects (Figure 2a) were found in loci on chromosome 4 (chr4:24517180-24646772, p = 2.96e-08; robust p = 2.89e-07), in the gene DHX15, and in an intergenic region on chromosome 10 (chr10:44172209-44755269, p = 3.28e-08; robust p = 6.98e-07). Additionally, two lone significant SNPs were found at chr4:100067606 (p = 4.28e-08; robust p = 2.31e-07), in a non-coding RNA linked to expression of the alcohol metabolism gene ADH4, and on chr5:139695229 (p = 4.38e-08; robust p = 8.90e-08), in an intergenic region. These four loci were mapped to 17 genes (Supplementary Table 1), including ADH1A, ADH4 and ADH5. MAGMA indicated 3 significant gene-based associations (NOXA1, DLGAP1, UBE2L3) after Bonferroni correction (Table 5), and one significant gene-set association within the GO Biological Pathways set of genes involved in regulation of developmental growth (p = 2.31e-06).

For FreqAlc (Figure 2b), GWEIS identified one significant locus for the SNP×SES interaction (chr16:5143432-5176998, p = 1.91e-08; robust p = 5.73e-08). The genes ALGI
and \textit{FAM86A} were implicated by positional mapping, including a candidate exonic
nonsynonymous variant in \textit{FAM86A}, and the genes \textit{MGRN1} and \textit{SEC14L5} were linked via
eQTL SNPs in the locus which affect expression levels of these genes in the brain
\textbf{(Supplementary Table 1). The gene IFIT1B} was identified in gene-based analysis ($p = 4.50e-08$). There were no gene-sets or tissues enriched for association.

For DrinkMeals (\textbf{Figure 2c}), there were no significant SNP interaction effects in the
model with trauma count as a moderator. With SES as a moderator (\textbf{Figure 2d}), there were
two loci with significant SNP×SES interactions: chr2:135771974-136787402 ($p = 5.20e-10$; robust $p = 3.35e-03$) and chr11:111541124-111989063 ($p = 1.23e-09$; robust $p = 4.59e-09$; considered credible after Bonferroni correction). Gene-based tests implicated 19 genes (\textbf{Table 4}), several of which contained exonic nonsynonymous variants among the candidate SNPs:
\textit{LCT} (chr2), \textit{FDXACB1}, \textit{ALG9}, \textit{DLAT}, \textit{PIH1D2}, and \textit{SDHD} (chr 11). The 94 genes mapped
by these loci (\textbf{Supplementary Table 1}) were enriched for expression in breast, spleen,
visceral adipose, and esophagus tissues ($p < 4.36e-04$). They were also enriched in multiple
gene-sets involved in immune system functioning, particularly in the immunological
signatures of plasmacytoid dendritic cells (\textbf{Table 6}).

\textbf{PRS prediction}

Across all alcohol phenotypes, PRS and iPRS\textsuperscript{G×E} based on the GWEIS results did not
significantly predict corresponding alcohol measures in an independent sample (\textbf{Table 4}).

\textbf{Discussion}

\textbf{Summary}

In this study, we conducted a large-scale GWEIS on multiple dimensions of alcohol
misuse in a population-based sample. By concentrating on phenotypes with previously
demonstrated SNP heritability and environmental factors (trauma and SES) with known
moderating effects on AM, we maximized our \textit{a priori} likelihood of identifying AM-relevant
gene-environment interactions. Despite this, we found few genomic regions with strong
evidence for G×E and poor out-of-sample prediction from interaction-based polygenic risk
scores. There was only one locus with convincing evidence for G×E at an individual SNP
level (i.e., robust p-value meeting the Bonferroni-corrected genome-wide significance
threshold of 1e-8). This was found for SES moderating the effect of SNPs mapped to
immune-related genes on a pattern of drinking with vs. without meals. Our aggregate gene-
based results, however, nominate several candidate genes which seem to interact with SES to
influence alcohol consumption and drinking patterns.

Interpretation of G×E effects

Trauma exposure has long been viewed as a catalyst for the manifestation of genetic
risk for AM (Magnusson et al., 2012). Yet, in this study, we found little evidence that the
effects of genetic variants are moderated by trauma or that they behave differently in
individuals with vs. without trauma exposure. This is in contrast to previous GWEIS in
smaller samples (N=20,000-50,000 versus the current N=120,000-360,000) which
demonstrated several significant SNP×trauma interactions for AM (Cheng et al., 2021,
Polimanti et al., 2018). This lack of replication may be due to the more specific focus of these
other studies on childhood trauma exposure (Cheng et al., 2021) or in a mostly military
sample (Polimanti et al., 2018). Genetic heterogeneity is a major challenge in gene
identification for AM (Wong and Schumann, 2008), and variability in the type, timing, or
context of trauma exposure may add an additional layer of complexity obscuring the
underlying genetic influences. However, it may also be that earlier results were false positives
which fail to replicate in larger studies, as has been the case for much G×E research.

For interactions involving SES, we are cautious of interpreting individual SNP effects
as most (robust) p-values did not survive multiple testing corrections. However, the gene-
based tests, which aggregate individual SNP G×E effects, pointed to several promising
candidate genes. For drinking quantity, these included the NADPH oxidase activator *NOXA1*, DLG associated protein *DLGAP1*, and ubiquitin conjugating enzyme *UBE2L3*, and, for frequency, the interferon induced protein with tetratricopeptide repeats *IFIT1B*. Except for *DLGAP1*, these genes have not previously been linked to AM. The GWASatlas (Watanabe et al., 2019) reports that *NOXA1* is associated with metabolic phenotypes like cholesterol and body fat, while *UBE2L3* has strong reported associations with numerous immunological traits/diseases and *IFIT1B* has been linked to smoking. *DLGAP1* has associations across a broad range of domains of health and behavior including an early GWAS of alcohol dependence (Wang et al., 2013). Genes whose effects on drinking quantity were moderated by SES were also enriched in regulation of developmental growth, suggesting that growth and/or metabolic processes related to drinking are broadly impacted by SES. Most interesting were the results for DrinkMeals, an AM phenotype which has recently been shown to represent an important distinct dimension of AM (Savage et al., 2023). G×E enrichment in multiple genes related to immune function suggests that genetic predispositions for immune-related conditions might differentially impact patterns of alcohol use at high versus low SES. For example, this could indicate differential patterns - influenced by income or other SES-related factors - of avoiding alcohol versus self-medicating in response to health problems.

To our knowledge, there are no other GWEIS examining SES with which we can compare these results, so the mechanism of action behind these interactions remains to be investigated. PRS prediction into a sample of college students did not show evidence of replication. It is uncertain whether this is due to the results being unreliable, the effect sizes too small, or the AM phenotypes and environmental moderators being too different across the discovery and validation samples (genetic heterogeneity). As substantial sub-threshold association signal was evident in the GWEIS, it may be that even larger samples and/or more refined measures are needed to obtain sufficient statistical power.
Strategies for modelling G×E

Through comparison of multiple G×E study designs, we also suggest that modelling main and interaction effects in a GWEIS individually is a more informative approach for characterizing molecular G×E, than stratified GWAS or a GWEIS with a joint test of the SNP main and interaction effects. We did not find substantial differences between these models in terms of their gene identification success or predictive power, but we found the GWEIS individual test model to be the most interpretable while formally comparing the difference in SNP effect sizes as a function of the level of the environmental moderator. However, our results may be limited by the power of the phenotype these models were applied to. A simulation study that systematically compares these models under different conditions could help to guide future G×E research.

Limitations and conclusion

Despite substantial increases in sample size and phenotypic breadth over previous GWEIS of AM, the results indicate that we still have limited power to detect G×E at the individual SNP level, although the twin literature indicates that they must exist. Until this issue is resolved, the biological insight provided by such findings is necessarily low. Our study is limited by the relatively shallow measures collected and the known biases in drinking behavior (Xue et al., 2021) due to the “healthy volunteer” bias in UKB (Fry et al., 2017). Nonetheless, we have highlighted several genes and a potential mechanism of immune system functioning behind the moderating effect of SES on the genetic influences on AM. We also emphasize the importance of considering specific dimensions of AM, as there was little overlap in the results between AUD, consumption, and drinking patterns. We conclude that GWEIS may be a preferred approach over stratified GWAS but modelling molecular G×E effects remains a challenge. Larger consortia, more in-depth measures, and longitudinal data collection are needed to resolve the complex gene-environment interplay underlying AM.
References


GENE-ENVIRONMENT INTERACTION IN ALCOHOL MISUSE


GENE-ENVIRONMENT INTERACTION IN ALCOHOL MISUSE

1 GABRIELEN, M. E., JOHNSEN, M. B., SKROVE, M., INDREDAVIK, M. S.,
2 DRANGE, O. K., BJORKESET, O., BØRTE, S., STENSLAND, S. Ø., CHOQUET,
3 H., SOCHERTY, A. R., FAUL, J. D., FOERSTER, J. R., FRITSCHE, L. G.,
4 GABRIELEN, M. E., GORDON, S. D., HAESSLER, J., HOTTENGA, J.-J.,
5 HUANG, H., JANG, S.-K., JANSEN, P. R., LING, Y., MÄGI, R., MATOBA, N.,
6 MCMAHON, G., MULAS, A., ORRÜ, P., PAL VIAINEN, T., PANDIT, A.,
7 REGINSSON, G. W., SKOGHOLT, A. H., SMITH, J. A., TAYLOR, A. E.,
8 TURMAN, C., WILLEMSEN, G., YOUNG, H., YOUNG, K. A., ZAJAC, G. J. M.,
9 ZHAO, W., ZHOU, W., BJÖNNSDOTTIR, G., BOARDMAN, J. D., BOEHNEKE, M.,
10 BOOMSMA, D. I., CHEN, C., CUCCA, F., DA VIES, G. E., EATON, C. B.,
11 EHRINGER, M. A., ESKO, T., FIORILLO, E., GILLESPIE, N. A.,
12 GUDBJARTSSON, D. F., HALLER, T., et al. 2019. Association studies of up to 1.2
13 million individuals yield new insights into the genetic etiology of tobacco and alcohol
14 use. Nature Genetics, 51, 237-244.
15 MAGNUSSON, Å., LUNDHOLM, C., GÖRANSSON, M., COPELAND, W., HEILIG, M.
18 MCCLELLAND, G. H. & JUDD, C. M. 1993. Statistical difficulties of detecting interactions
19 and moderator effects. Psychol Bull, 114, 376-90.
20 MILANESCHI, Y., LAMERS, F., PEYROT, W. J., BAUNE, B. T., BRENN, G., DEGHAN,
21 A., FORSTNER, A. J., GRABE, H. J., HOMUTH, G., KAN, C., LEWIS, C.,
22 MULLINS, N., NAUCK, M., PISTIS, G., PREISIG, M., RIVERA, M., RIETSCHEL,
23 M., STREIT, F., STROHMAIER, J., TEUMER, A., VAN DER AUWERA, S., WRAY,
25 Depression With Atypical Features and Obesity-Related Immunometabolic
26 Dysregulations. JAMA Psychiatry, 74, 1214-1225.
27 PASMAN, J. A., VERWEIJ, K. J. H., ABDELLAOUI, A., HOTTENGA, J. J., FEDKO, I. O.,
29 between polygenic risk and neighborhood environment. Drug Alcohol Depend, 209,
30 107948.
31 PETERSON, R. E., EDWARDS, A. C., BACANU, S. A., DICK, D. M., KENDLER, K. S. &
32 WEBB, B. T. 2017. The utility of empirically assigning ancestry groups in cross-
34 POLIMANTI, R., KAUFMAN, J., ZHAO, H., KRANZLER, H. R., URSANO, R. J.,
35 KESSLER, R. C., GELERNTER, J. & STEIN, M. B. 2018. A genome-wide gene-by-
36 trauma interaction study of alcohol misuse in two independent cohorts identifies
38 R CORE TEAM 2015. R: A language and environment for statistical computing. Vienna,
40 SALVATORE, J. E., ALIEV, F., EDWARDS, A. C., EVANS, D. M., MACLEOD, J.,
41 HICKMAN, M., LEWIS, G., KENDLER, K. S., LOUKOLA, A., KORHONEN, T.,
43 Predict Alcohol Problems in an Independent Sample and Show Moderation by the
45 SALVATORE, J. E., THOMAS, N. S., CHO, S. B., ADKINS, A., KENDLER, K. S. & DICK,
46 D. M. 2016. The role of romantic relationship status in pathways of risk for emerging
47 adult alcohol use. Psychol Addict Behav, 30, 335-44.
48 SARTOR, C. E., WANG, Z., XU, K., KRANZLER, H. R. & GELERNTER, J. 2014. The
49 joint effects of ADH1B variants and childhood adversity on alcohol related


Table 1. Descriptions of the alcohol phenotypes and environmental measures included in gene-environment interaction analyses.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Item name</th>
<th>Description</th>
<th>Field ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>BroadAUD</td>
<td>broad_aud_ph</td>
<td>Aggregate of lifetime diagnoses of AUD or alcohol-related diseases (alcoholic cirrhosis; Korsakov's syndrome; Wernicke's syndrome; alcoholic ataxia) from medical records, a prescription of an alcohol treatment drug (acamprosate, disulfiram, nalmefene), clinical codes for &quot;personal history of alcoholism&quot;, or self-report of an alcohol addiction</td>
<td>20404,20406</td>
</tr>
<tr>
<td>FreqAlc</td>
<td>drinkfreq</td>
<td>Self-report measure of typical drinking frequency</td>
<td>1558</td>
</tr>
<tr>
<td>QuantAlc</td>
<td>quant_ts_gp</td>
<td>Self-report or clinician-reported measures of typical drinking quantity (gm ethanol) per month, log+1 transformed</td>
<td>1568,1578,1588,1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>598,1608,4407,44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18,4429,4440,445</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1,4462,5364</td>
</tr>
<tr>
<td>DrinkMeals</td>
<td>drink_w_meals</td>
<td>Self-report measure of whether alcohol is typically consumed with or without meals, indexing a broader pattern of drinking habits</td>
<td>1618</td>
</tr>
<tr>
<td>Trauma</td>
<td>-</td>
<td>Lifetime exposure to 10 potentially traumatic events such as assault, abuse, combat experience, a life-threatening accident/illness, or witnessing a violent death</td>
<td>20488, 20490,20523,20524,20526-20531</td>
</tr>
<tr>
<td>SES</td>
<td>-</td>
<td>Socioeconomic status indexed at the time of study enrollment with the Townsend Deprivation Index, a continuous scale score derived from national census measures of material deprivation metrics (e.g., unemployment, home ownership) within a participant’s home postcode</td>
<td>22189</td>
</tr>
</tbody>
</table>

Note: 1Item names for alcohol measures relate to the descriptions in Savage et al., 2023 (Supplementary Table 1) which provide greater detail on the derivation of these measures. 2Field IDs as listed in the UK Biobank (https://biobank.ctsu.ox.ac.uk/showcase/index.cgi).
Table 2. Overview of the alcohol phenotypes and environmental measures included in gene-environment interaction analyses.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Outcome phenotype levels</th>
<th>N (N cases / N controls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. BroadAUD, stratified GWAS, trauma-exposed subset</td>
<td>Case/control</td>
<td>12,098 / 66,110</td>
</tr>
<tr>
<td>2. BroadAUD, stratified GWAS, non-trauma-exposed subset</td>
<td>Case/control</td>
<td>4,458 / 43,925</td>
</tr>
<tr>
<td>3. BroadAUD - Trauma, GWEIS-individual</td>
<td>Case/control</td>
<td>16,534 / 109,887</td>
</tr>
<tr>
<td>4. BroadAUD - Trauma – GWEIS-joint</td>
<td>Case/control</td>
<td>16,534 / 109,887</td>
</tr>
<tr>
<td>5. FreqAlc – SES, GWEIS-individual</td>
<td>Days per month</td>
<td>360,314</td>
</tr>
<tr>
<td>6. QuantAlc – SES, GWEIS-individual</td>
<td>Quantity per month</td>
<td>332,668</td>
</tr>
<tr>
<td>7. DrinkMeals – Trauma, GWEIS-individual</td>
<td>Drink with/not with meals</td>
<td>51,591/ 16,215</td>
</tr>
<tr>
<td>8. DrinkMeals - SES, GWEIS-individual</td>
<td>Drink with/not with meals</td>
<td>135,288 / 62,246</td>
</tr>
</tbody>
</table>
Table 3. Genetic correlations between genome-wide association results using multiple modelling approaches for characterizing gene-environment interaction for BroadAUD.

<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Stratified GWAS - trauma</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Stratified GWAS – no trauma</td>
<td>0.83</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. GWEIS–individual, main effect</td>
<td>0.90</td>
<td>0.46</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>4. GWEIS–individual, interaction</td>
<td>-0.07</td>
<td>-0.59</td>
<td>-0.38</td>
<td>--</td>
</tr>
<tr>
<td>5. GWEIS–joint 2df test\textsuperscript{a}</td>
<td>0.12</td>
<td>0.27</td>
<td>0.65</td>
<td>0.65</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Spearman’s rank-order correlations calculated between SNP p-values as the joint test has no signed effect.
Table 4. Polygenic score (PGS) and interaction-based polygenic score-by-environment (iPGS<sub>GxE</sub>) prediction of alcohol phenotypes in an independent validation sample.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Outcome</th>
<th>Main effects PGS</th>
<th>Interaction iPGS&lt;sub&gt;GxE&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Beta/OR  P</td>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>BroadAUD stratified (trauma group)</td>
<td>BroadAUD</td>
<td>0.946 .138 .002</td>
<td>- - -</td>
</tr>
<tr>
<td>BroadAUD stratified (no trauma group)</td>
<td>BroadAUD</td>
<td>0.895 .331 .037</td>
<td>- - -</td>
</tr>
<tr>
<td>BroadAUD - Trauma</td>
<td>BroadAUD</td>
<td>0.929 .048 .001</td>
<td>1.039 .390 .000</td>
</tr>
<tr>
<td>DrinkMeals - SES</td>
<td>BroadAUD</td>
<td>0.982 .655 .000</td>
<td>1.279 .285 .000</td>
</tr>
<tr>
<td>DrinkMeals - Trauma</td>
<td>BroadAUD</td>
<td>0.962 .291 .000</td>
<td>1.066 .399 .000</td>
</tr>
<tr>
<td>QuantAlc - SES</td>
<td>QuantAlc</td>
<td>0.027 .151 .001</td>
<td>0.016 .557 .000</td>
</tr>
<tr>
<td>DrinkMeals - SES</td>
<td>QuantAlc</td>
<td>0.003 .864 .001</td>
<td>0.185 .087 .001</td>
</tr>
<tr>
<td>DrinkMeals - Trauma</td>
<td>QuantAlc</td>
<td>0.000 .998 .000</td>
<td>0.049 .162 .001</td>
</tr>
<tr>
<td>FreqAlc - SES</td>
<td>FreqAlc</td>
<td>0.021 .229 .000</td>
<td>0.027 .747 .000</td>
</tr>
<tr>
<td>DrinkMeals - SES</td>
<td>FreqAlc</td>
<td>-0.001 .952 .000</td>
<td>0.058 .597 .000</td>
</tr>
<tr>
<td>DrinkMeals - Trauma</td>
<td>FreqAlc</td>
<td>-0.006 .728 .000</td>
<td>0.069 .052 .001</td>
</tr>
</tbody>
</table>
Table 5. Genes with significant enrichment in SNPxSES interaction effects from the GWEIS.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Gene</th>
<th>CHR</th>
<th>Start BP</th>
<th>End BP</th>
<th># SNPs</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>QuantAlc-SES</td>
<td>NOXA1</td>
<td>9</td>
<td>140317802</td>
<td>140328858</td>
<td>56</td>
<td>4.5643</td>
<td>2.51E-06</td>
</tr>
<tr>
<td>QuantAlc-SES</td>
<td>DLGAP1</td>
<td>18</td>
<td>3496030</td>
<td>4455335</td>
<td>3480</td>
<td>4.583</td>
<td>2.29E-06</td>
</tr>
<tr>
<td>QuantAlc-SES</td>
<td>UBE2L3</td>
<td>22</td>
<td>21903736</td>
<td>21978323</td>
<td>131</td>
<td>4.7598</td>
<td>9.69E-07</td>
</tr>
<tr>
<td>FreqAlc-SES</td>
<td>IFIT1B</td>
<td>10</td>
<td>91137813</td>
<td>91144962</td>
<td>22</td>
<td>5.3841</td>
<td>3.64E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>MAP3K19</td>
<td>2</td>
<td>135722061</td>
<td>135805038</td>
<td>272</td>
<td>4.7591</td>
<td>9.72E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>R3HDM1</td>
<td>2</td>
<td>136289025</td>
<td>136482840</td>
<td>336</td>
<td>5.0336</td>
<td>2.41E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>UBXN4</td>
<td>2</td>
<td>136499189</td>
<td>136542625</td>
<td>107</td>
<td>5.069</td>
<td>2.00E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>LCT</td>
<td>2</td>
<td>136545410</td>
<td>136594750</td>
<td>141</td>
<td>5.2046</td>
<td>9.72E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>MCM6</td>
<td>2</td>
<td>136597196</td>
<td>136633996</td>
<td>88</td>
<td>5.2887</td>
<td>6.16E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>DARS</td>
<td>2</td>
<td>136664247</td>
<td>136743670</td>
<td>152</td>
<td>4.5928</td>
<td>2.19E-06</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>SIK2</td>
<td>11</td>
<td>111473115</td>
<td>111601577</td>
<td>339</td>
<td>4.9598</td>
<td>3.53E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>PPP2R1B</td>
<td>11</td>
<td>111597632</td>
<td>111637151</td>
<td>99</td>
<td>5.0392</td>
<td>2.34E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>ALG9</td>
<td>11</td>
<td>111657010</td>
<td>111750149</td>
<td>224</td>
<td>5.5454</td>
<td>1.47E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>FDXACB1</td>
<td>11</td>
<td>111744780</td>
<td>111751967</td>
<td>21</td>
<td>5.3291</td>
<td>4.94E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>CRYAB</td>
<td>11</td>
<td>111779289</td>
<td>111794446</td>
<td>33</td>
<td>4.9116</td>
<td>4.52E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>HSPB2-C11orf52</td>
<td>11</td>
<td>111783460</td>
<td>111797595</td>
<td>30</td>
<td>4.8833</td>
<td>5.22E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>C11orf52</td>
<td>11</td>
<td>111788756</td>
<td>111797596</td>
<td>18</td>
<td>5.0699</td>
<td>1.99E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>DIXDC1</td>
<td>11</td>
<td>111797868</td>
<td>111893308</td>
<td>195</td>
<td>5.37</td>
<td>3.94E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>DLAT</td>
<td>11</td>
<td>111895538</td>
<td>111935114</td>
<td>82</td>
<td>5.2919</td>
<td>6.05E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>PIH1D2</td>
<td>11</td>
<td>111934734</td>
<td>111944998</td>
<td>22</td>
<td>5.524</td>
<td>1.66E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>C11orf57</td>
<td>11</td>
<td>111944810</td>
<td>111955874</td>
<td>16</td>
<td>5.3897</td>
<td>3.53E-08</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>TIMM8B</td>
<td>11</td>
<td>111955524</td>
<td>111957522</td>
<td>2</td>
<td>5.0433</td>
<td>2.29E-07</td>
</tr>
<tr>
<td>DrinkMeals-SES</td>
<td>SDHD</td>
<td>11</td>
<td>111957497</td>
<td>111990353</td>
<td>66</td>
<td>5.1472</td>
<td>1.32E-07</td>
</tr>
</tbody>
</table>

Note: genes shown are significant after Bonferroni correction for 19,628 protein-coding genes (p < .05/19628=2.55e-06)
Table 6. Gene-sets with significant enrichment in genes mapped by the GWEIS of SNPxSES interaction effects predicting DrinkMeals.

<table>
<thead>
<tr>
<th>GeneSet</th>
<th>N genes</th>
<th>n overlap</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE29618_MONOCYTE_VS_PDC_DN</td>
<td>187</td>
<td>27</td>
<td>1.20e-32</td>
</tr>
<tr>
<td>GSE29618_PDC_VS_MDC_UP</td>
<td>189</td>
<td>27</td>
<td>1.62e-32</td>
</tr>
<tr>
<td>GSE29618_PDC_VS_MDC_DAY7_FLU_VACCINE_UP</td>
<td>190</td>
<td>26</td>
<td>8.62e-31</td>
</tr>
<tr>
<td>GSE29618_BCELL_VS_PDC_DN</td>
<td>195</td>
<td>24</td>
<td>2.89e-27</td>
</tr>
<tr>
<td>GSE29618_BCELL_VS_PDC_DAY7_FLU_VACCINE_DN</td>
<td>195</td>
<td>24</td>
<td>2.89e-27</td>
</tr>
<tr>
<td>GSE29618_MONOCYTE_VS_PDC_DAY7_FLU_VACCINE_DN</td>
<td>192</td>
<td>23</td>
<td>7.44e-26</td>
</tr>
<tr>
<td>GSE17301_ACD3_ACD28_VS_ACD3_ACD28_AND_IFNA5_STIM_CD8_TCELL_DN</td>
<td>193</td>
<td>11</td>
<td>2.74e-09</td>
</tr>
<tr>
<td>GSE16385_ROSIGLITAZONE_IL4_VS_IL4_ALONE_STIM_MACROPHAGE_12H_UP</td>
<td>193</td>
<td>9</td>
<td>4.54e-07</td>
</tr>
<tr>
<td>GSE29618_PRE_VS_DAY7_FLU_VACCINE_MDC_UP</td>
<td>188</td>
<td>8</td>
<td>4.04e-06</td>
</tr>
</tbody>
</table>

Note: gene-sets shown are significant after Bonferroni correction for 10,678 pathways in the Molecular Signatures Database curated and gene ontology sets (p < .05/10678=4.68e-06)
Figures

Figure 1. Manhattan plots of the genomic associations for BroadAUD within the a) trauma-exposed and b) non-trauma-exposed subgroups of a stratified GWAS, and for the c) individual SNP main effects and d) individual SNP $\times$ trauma interaction effects, and e) joint test of SNP main and interaction effects in a genome-wide gene-environment interaction study (GWEIS).
Figure 2. Manhattan plots of the genomic associations for interaction effects from a GWEIS of (a) socioeconomic status (SES) with QuantAlc, (b) SES with FreqAlc, (c) trauma exposure with DrinkMeals, and (d) SES with DrinkMeals.