Title: Is anxiety a pathway to Alcohol Use Disorders? A phenome-wide association study of the GABRA2 coding variant rs279858.

Authors
Alison K. Merikangas PhD MPH1,2,3, Rachel L. Kember PhD4,5, Martin H. Plawecki MD PhD6, Chella Kamarajan PhD7, Grace Chan PhD8, Lance Bauer PhD8, Jacquelyn L. Meyers PhD7, John I. Nurnberger Jr MD PhD8,10, John Kramer PhD11, Bernice Porjesz PhD7, Howard J. Edenberg PhD10,12, Laura Almasy PhD1,2,3

Affiliations
1 Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia
2 Penn-CHOP Lifespan Brain Institute, University of Pennsylvania
3 Department of Genetics, Perelman School of Medicine, University of Pennsylvania
4 Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania
5 Mental Illness Research Education and Clinical Center, Corporal Michael J. Crescenz VA Medical Center
6 Department of Psychiatry, Christian Sarkine Autism Treatment Center, Riley Hospital for Children at Indiana University Health, Indiana University School of Medicine
7 Department of Psychiatry and the Henri Begleiter Neurodynamics Laboratory, State University of New York Downstate Medical Center
8 Department of Psychiatry, University of Connecticut School of Medicine
9 Stark Neurosciences Research Institute, Department of Psychiatry, Indiana University School of Medicine
10 Department of Medical and Molecular Genetics, Indiana University School of Medicine
11 Department of Psychiatry, University of Iowa College of Medicine
12 Department of Biochemistry & Molecular Biology, Indiana University School of Medicine

Correspondence
Alison K. Merikangas, PhD, MPH
Children's Hospital of Philadelphia
Department of Biomedical and Health Informatics
Abramson Building 1016A
3615 Civic Center Blvd
Philadelphia, PA 19104-4319
merikangaa@chop.edu
267-426-5258

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Alcohol use disorders (AUDs) and related electrophysiological endophenotypes have been associated with the GABRA2 gene. However, the causal variants in GABRA2 and their mechanisms of influence on AUD and its correlates have not been established. Here we investigate the phenotypic spectrum of a GABRA2 coding variant (rs279858) through a phenome-wide association study (PheWAS) in two open-source datasets.

We applied the PheWAS approach to identify a broad range of phenotypes associated with rs279858 in the MRC IEU OpenGWAS PheWAS and the Open Targets Genetics Portal. These databases extend the array of phenotypes beyond those available in electronic health records (EHR) to include numerous non-medical phenotypes and traits. We then followed up the results from those exploratory associations by examining the genetic correlations between our “top hits” and alcohol- and smoking-related phenotypes.

In both data sources, rs279858 (C effect allele) was associated with anxiety-related phenotypes, including reduced risk-taking behavior and an increase in nervous feelings, as well as reduced number of lifetime sexual partners. Follow-up analyses revealed that these phenotypes were genetically correlated with each other and with alcohol- and smoking-related phenotypes.

This work illustrates the utility of the PheWAS approach, particularly for phenotypes that extend beyond those that are typically captured in EHR data. In fact, the associations described here are all behavioral rather than clinical phenotypes. We postulate that these traits may be related to anxiety or behavioral inhibition that has been identified as a risk factor for AUD, and may represent pathophysiological intermediaries between GABRA2 and AUD.
Introduction

Alcohol Use Disorder (AUD) is a complex condition defined by a pattern of unhealthy, problematic use of alcohol with significant clinical impairment. There are both genetic and environmental contributions to its development. AUD has considerable public health, social, and economic costs with an estimated 5.9% of people meeting past-year criteria [1] and 13.8% in the United States meeting criteria for lifetime AUD [2]. Despite a reported heritability of nearly 50%, and substantial effort to establish the underlying genetic etiology, few genes aside from the alcohol metabolizing genes \textit{ADH1B} and \textit{ALDH2} have been consistently associated with AUD in population-based genome-wide association studies (GWAS)[3]. However, a recent large meta-analysis including 435,563 individuals of European descent has now identified 29 associated loci for ‘problematic alcohol use’ (PAU)[4].

Both linkage and association studies of samples enriched for AUD have implicated the gene encoding gamma-aminobutyric acid type A receptor subunit alpha2 (\textit{GABRA2}) in the etiology of alcohol use disorder [5-8]. The gamma-aminobutyric acid receptor subunit alpha-2 protein is a subunit of the GABA-A receptor. RNA-seq studies have demonstrated higher \textit{GABRA2} expression in brain, kidney, and small intestine as compared to other body tissues [9,10]. \textit{GABRA2} has also been implicated in EEG changes in those with AUD, especially the beta rhythm of the resting EEG [5,11,12]. \textit{GABRA2} may act via neuromodulation since it has been associated with both AUD and EEG phenotypes [5].

In addition to AUD, associations have been reported between \textit{GABRA2} and several substance misuse phenotypes, including vulnerability to both nicotine dependence (ND) and cannabis dependence (CD)[13] as well as subjective response to alcohol and other alcohol-related reward processes [14,15]. \textit{GABRA2} has also been implicated in the disrupted structural connectome of reward circuits in heroin abusers [16], and in other neuropsychiatric-related phenotypes, including epilepsy [17], age at onset of Alzheimer's disease [18], and risk-taking behavior [19], as well as antipsychotic-associated weight gain [20].

Though multiple single nucleotide polymorphisms (SNPs) within \textit{GABRA2} have been associated with both alcohol use and electrophysiological phenotypes, multiple studies have shown the strongest association signal for rs279858 and SNPs that are in linkage disequilibrium (LD) with this lead variant [5,21]. The rs2 79858 SNP in \textit{GABRA2} is one of the few loci from the candidate gene analysis era that was associated with alcohol dependence in the Study of Alcohol Addiction: Genetics and Addiction (SAGE)[8]. Specifically, G/G carriers appear to be more prone to AUD than A/A or A/G carriers [22,23]. However, in some publications, genotypes were called from the complementary DNA strand with alleles C and T, in which case the C allele is the one associated with higher risk. A functional role for this synonymous change in rs279858 has not been established. Researchers have postulated that its actions involve the modulation of both cognition and reward systems [16]. However, this SNP has not been associated with AUD in the largest case-control GWAS of AUD [4] and PAU [24] to date, but it has been implicated in phenotypic manifestations of alcohol use in some race/ethnicity subgroups [25] and in sex-specific analyses [26].

In parallel to the GWAS approach that attempts to identify markers that may influence the risk of AUD across the genome, phenome-wide association studies (PheWAS) are employed to identify a range of phenotypes that may be associated with a particular genetic variant. PheWAS test the association between a genetic variant (e.g., a SNP, gene, or polygenic score), and an array of phenotypes that range from International Classification of Diseases (ICD) codes in electronic health records (EHR)[27,28] to data from epidemiologic studies and biobanks that have been supplemented with questionnaires that assess...
lifestyle and behavioral traits [29]. PheWAS can identify associations of risk loci with novel traits that have not been previously linked to AUD. This data-driven approach can highlight numerous additional traits or conditions that may be associated with a genetic variant that might not have been investigated in analyses driven by known relationships among phenotypes. The PheWAS approach, e.g., PheWAS of polygenic scores [30,31] and other AUD-associated variants [32] has also been employed to elucidate comorbid conditions that may also contribute to the developmental trajectory of AUD.

Here we apply the PheWAS approach to several open-source tools in order to examine the phenotypic spectrum of an AUD-associated GABRA2 coding variant (rs279858) that may provide indices of potential causal systems and pathways. We employ three approaches to address the exploratory nature of this work including: a) correcting for multiple comparisons; b) conducting a “pseudo” replication by using two curated databases containing multiple large studies; and c) inspecting the genetic correlations within the identified phenotypes and between these phenotypes and alcohol-related phenotypes.

Materials and Methods

PheWAS data sources

The PheWAS approach [27] was used to investigate the range of phenotypes associated with rs279858 using two online tools; the Medical Research Council (MRC) Integrative Epidemiology Unit (IEU) at the University of Bristol OpenGWAS PheWAS tool [33] (https://gwas.mrcieu.ac.uk/phewas/), and the Open Targets Genetics Portal [34] (https://genetics.opentargets.org/). Supplementary Table 3 lists the included data sources for both tools.

The MRC IEU OpenGWAS PheWAS queries a manually curated collection of eighteen batches of GWAS summary datasets encompassing over 14,000 individual GWAS that includes the UK Biobank (UKBB) and other national databases such as the Japanese Biobank [35] (https://biobankjp.org/en/), and Finngen (https://www.finngen.fi/en), as well as the National Human Genome Research Institute (NHGRI)-European Bioinformatics Institute (EBI) GWAS Catalog summary statistics database. We searched rs279858 using their tool (https://gwas.mrcieu.ac.uk/phewas/), which only reports results filtered to p<0.001, and then used the ieugwas R package (R version 4.0.4) to connect to the application programming interface (API) and downloaded the complete results file (version 3.8.5).

The Open Targets Genetics Portal includes data from Finngen (https://www.finngen.fi/en), the National Human Genome Research Institute (NHGRI)-European Bioinformatics Institute (EBI) GWAS Catalog summary statistics database [36] (GWAS Catalog, https://www.ebi.ac.uk/gwas/), and UKBB (http://www.nealelab.is/uk-biobank). We searched rs279858 on the Open Targets Genetics site (https://genetics.opentargets.org/, version 22.02) which maps this SNP to its location GRCh38: 4:46,312,576 (GRCh37: 4:46,314,593, https://genetics.opentargets.org/var/4_46312576_T_C). Only traits associated at p<0.005 are returned and used for analysis.

PheWAS statistical significance

Some traits that have been examined in independent GWAS analyses of the biobanks have also been deposited in the GWAS Catalog. Therefore, traits may appear in both the OpenGWAS and Open Targets results, and on occasion multiple times in each. For such situations, we opted to retain the results from the instance with the smallest p-value. We adjusted the p-value to correct for the 16482 traits that appeared in the downloadable OpenGWAS dataset, which resulted in a Bonferroni-corrected significance threshold for this study of p<4.00x10^{-6}. The Open Targets tool provides a Bonferroni-
corrected phenome-wide significance threshold of $p<9.77 \times 10^{-5}$. The overlap in contributing data sources in shown in Supplementary Table 3.

SNP-based heritability

In order to determine the likelihood of detecting genetic signals for alcohol-related phenotypes in the UKBB, we examined the SNP-based heritability ($h^2$) in the UKBB generated by the Neale Lab [37]. They provide an automated calculation of SNP-heritability via partitioned linkage disequilibrium score regression (LDSC) [38], with $h^2$ values using a liability scale for binary phenotypes. We used this information to gauge the overall genetic signal for UKBB alcohol-related phenotypes queried in our PheWAS analysis, by examining the heritability and statistical significance of component items of the Alcohol Use Disorders Identification Test (AUDIT) [39] present in the UKBB.

Genetic correlation

We used results generated by the Neale Lab on the genetic correlation ($r_g$) between phenotypes in the UKBB via LDSC (https://ukbb-rg.hail.is/) to explore unexpected results from the PheWAS. Statistically significant phenotypes from the PheWAS analyses were compared to one another, and to the available alcohol, tobacco, and cannabis use phenotypes. This information can be used to evaluate whether the identified phenotype may be indexing another trait that does not demonstrate statistical significance due to sample size. A Bonferroni corrected p-value of $2.92 \times 10^{-4}$ was employed to correct for the number of correlations that we examined across the selected traits (171 pairs).

Results

The PheWAS results at $p<0.005$ from both the MRC IEU OpenGWAS and the Open Targets PheWAS are shown in Supplementary Table 1 and 2, respectively. Supplementary Table 3 lists the data sources and highlights their overlapping and independent contributions to the results presented here.

MRC IEU OpenGWAS

The strongest associations for rs279858 (C effect allele) were with reduced risk-taking behavior, an increase in nervous feelings, reduced number of lifetime sexual partners, and reduced length of mobile phone use. Table 1 shows the top ten associations of rs279858 in the IEU analysis. For traits that appeared more than once, the instance with the smallest p-value was retained. Figure 1a and Supplementary Table 1 show the complete results at $p<0.005$. 
Table 1: Top ten associations of rs279858 in the OpenGWAS Analysis.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Data Source</th>
<th>Total N</th>
<th>Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk taking</td>
<td>IEU UKBB</td>
<td>446,279</td>
<td>-0.0052</td>
<td>2.10×10⁻⁸</td>
</tr>
<tr>
<td>Nervous feelings</td>
<td>IEU UKBB</td>
<td>450,700</td>
<td>0.0045</td>
<td>4.60×10⁻⁷</td>
</tr>
<tr>
<td>Lifetime number of sexual partners</td>
<td>IEU UKBB</td>
<td>378,882</td>
<td>-0.0090</td>
<td>1.10×10⁻⁶</td>
</tr>
<tr>
<td>Length of mobile phone use</td>
<td>IEU UKBB</td>
<td>456,972</td>
<td>-0.0134</td>
<td>2.80×10⁻⁶</td>
</tr>
<tr>
<td>Natural Killer T Absolute Count</td>
<td>EBI Database</td>
<td>3,653</td>
<td>0.1123</td>
<td>1.02×10⁻⁷</td>
</tr>
<tr>
<td>HLA DR on HLA DR+ T cell</td>
<td>EBI Database</td>
<td>3,060</td>
<td>-0.1071</td>
<td>4.10×10⁻⁷</td>
</tr>
<tr>
<td>Natural Killer T %T cell</td>
<td>EBI Database</td>
<td>3,669</td>
<td>0.1046</td>
<td>4.34×10⁻⁷</td>
</tr>
<tr>
<td>Natural Killer T %lymphocyte</td>
<td>EBI Database</td>
<td>3,669</td>
<td>0.1020</td>
<td>7.47×10⁻⁷</td>
</tr>
<tr>
<td>Milk type used: Never/rarely have milk</td>
<td>Neale UKBB</td>
<td>360,806</td>
<td>-0.0017</td>
<td>8.08×10⁻⁵</td>
</tr>
<tr>
<td>MAP kinase-activated protein kinase 2</td>
<td>Sun et al. 2018 [40]</td>
<td>3,301</td>
<td>0.0941</td>
<td>1.35×10⁻⁴</td>
</tr>
</tbody>
</table>

Statistically significant findings are highlighted in bold font. IEU = Integrative Epidemiology Unit; EBI = European Bioinformatics Institute; Sun et al. 2018 [40] is an independent study included in the OpenGWAS database.

Open Targets Genetics
In the Open Targets analysis, rs279858 (C effect allele) was associated with a reduced level of self-reported number of lifetime sexual partners, reduced risk-taking behavior, an increase in nervous feelings, a lower Townsend deprivation index at recruitment (a measure of socioeconomic status within a population encompassing employment, car and home ownership, and household overcrowding by postcode)[41], getting up earlier in the morning, never/rarely having milk, and a reduced length of mobile phone use. Table 2 shows the top ten associations for rs279858 in the Open Targets analysis. Figure 1b and Supplementary Table 2 show the complete results at p<0.005. No alcohol, tobacco, or other substance use phenotypes were among the top rs279858 associations in either PheWAS.
Table 2: Top ten associations of rs279858 in the Open Targets Analysis.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Data Source</th>
<th>Total N</th>
<th>Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifetime number of sexual partners</td>
<td>Neale UKBB</td>
<td>296609</td>
<td>-0.0098</td>
<td>$1.39 \times 10^{-06}$</td>
</tr>
<tr>
<td>Risk taking</td>
<td>Neale UKBB</td>
<td>348549</td>
<td>-0.0250</td>
<td>$3.80 \times 10^{-06}$</td>
</tr>
<tr>
<td>Nervous feelings</td>
<td>Nagel et al. 2018 [42]</td>
<td>373121</td>
<td>0.0101</td>
<td>$1.58 \times 10^{-05}$</td>
</tr>
<tr>
<td>Townsend deprivation index at recruitment</td>
<td>Neale UKBB</td>
<td>360763</td>
<td>-0.0283</td>
<td>$4.87 \times 10^{-05}$</td>
</tr>
<tr>
<td>Getting up in morning (Morningness)</td>
<td>Neale UKBB</td>
<td>360231</td>
<td>0.0072</td>
<td>$6.04 \times 10^{-05}$</td>
</tr>
<tr>
<td>Never/rarely have milk</td>
<td>milk type used</td>
<td>Neale UKBB</td>
<td>360806</td>
<td>-0.0531</td>
</tr>
<tr>
<td>Length of mobile phone use</td>
<td>Neale UKBB</td>
<td>356618</td>
<td>-0.0125</td>
<td>$8.11 \times 10^{-05}$</td>
</tr>
<tr>
<td>Renal/kidney failure</td>
<td>non-cancer, self-reported</td>
<td>Neale UKBB</td>
<td>361141</td>
<td>0.4043</td>
</tr>
<tr>
<td>Both eyes</td>
<td>which eye(s) affected by astigmatism</td>
<td>Neale UKBB</td>
<td>9749</td>
<td>0.1130</td>
</tr>
<tr>
<td>Oat cereal</td>
<td>cereal type</td>
<td>Neale UKBB</td>
<td>299898</td>
<td>0.0214</td>
</tr>
</tbody>
</table>

Statistically significant findings are highlighted in bold font. Nagel et al. 2018 [42] is an independent study included in the Open Targets database.

SNP-based heritability of AUDIT phenotypes in UKBB

Table 3 shows the AUDIT items in the UKBB, their SNP-based heritability, associated p-value, and the sample size for each item. The final column shows the p-value that was generated for that item in our MRC IEU OpenGWAS PheWAS analysis. The AUDIT items that showed statistically significant SNP-based heritability included: "Frequency of drinking alcohol"; "Amount of alcohol drunk on a typical drinking day"; and "Frequency of consuming six or more units of alcohol". Notably, not all of the AUDIT items were included in the PheWAS, and the ones that were included were not associated with rs279858.
Table 3: Heritability of alcohol phenotypes in the UKBB from the MRC IEW OpenGWAS PheWAS

<table>
<thead>
<tr>
<th>AUDIT item description</th>
<th>$h^2$</th>
<th>p-value</th>
<th>N</th>
<th>OpenGWAS PheWAS p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of drinking alcohol</td>
<td>0.073</td>
<td>$4.22 \times 10^{-24}$</td>
<td>117914</td>
<td>0.528</td>
</tr>
<tr>
<td>Amount of alcohol drunk on a typical drinking day</td>
<td>0.052</td>
<td>$1.45 \times 10^{-14}$</td>
<td>108256</td>
<td>--</td>
</tr>
<tr>
<td>Frequency of consuming six or more units of alcohol</td>
<td>0.061</td>
<td>$1.88 \times 10^{-19}$</td>
<td>108485</td>
<td>--</td>
</tr>
<tr>
<td>Frequency of inability to cease drinking in last year</td>
<td>0.033</td>
<td>$1.96 \times 10^{-18}$</td>
<td>64973</td>
<td>--</td>
</tr>
<tr>
<td>Frequency of failure to fulfil normal expectations due to drinking alcohol in last year</td>
<td>0.024</td>
<td>$1.53 \times 10^{-15}$</td>
<td>65054</td>
<td>--</td>
</tr>
<tr>
<td>Frequency of needing morning drink of alcohol after heavy drinking session in last year</td>
<td>0.017</td>
<td>$5.63 \times 10^{-13}$</td>
<td>65099</td>
<td>--</td>
</tr>
<tr>
<td>Frequency of feeling guilt or remorse after drinking alcohol in last year</td>
<td>0.047</td>
<td>$8.79 \times 10^{-13}$</td>
<td>65099</td>
<td>--</td>
</tr>
<tr>
<td>Frequency of memory loss due to drinking alcohol in last year</td>
<td>0.023</td>
<td>$2.30 \times 10^{-10}$</td>
<td>65029</td>
<td>--</td>
</tr>
<tr>
<td>Ever been injured or injured someone else through drinking alcohol: No</td>
<td>0.040</td>
<td>$1.03 \times 10^{-10}$</td>
<td>118002</td>
<td>0.454</td>
</tr>
<tr>
<td>Ever been injured or injured someone else through drinking alcohol: Yes, but not in the last year</td>
<td>0.019</td>
<td>$2.87 \times 10^{-11}$</td>
<td>118002</td>
<td>0.429</td>
</tr>
</tbody>
</table>

Statistically significant findings are highlighted in bold font. AUDIT = Alcohol Use Disorders Identification Test; $h^2$ = SNP-based heritability.

Genetic correlation

The genetic correlations between the statistically significant UKBB phenotypes from the PheWAS, and UKBB alcohol, tobacco, and substance use phenotypes are shown in Figure 2 (also see the full results in Supplementary Table 4). After a conservative Bonferroni correction ($p<2.92 \times 10^{-4}$), statistically significant genetic correlations were observed among the majority of the PheWAS phenotypes with the exception of the dietary phenotypes [i.e., “Milk type used: Never/rarely have milk” and “Cereal type: Oat cereal (e.g., Ready Brek, porridge)"], which were nominally correlated with the other PheWAS traits ($p<0.05$) but did not survive correction for multiple testing. The top hits from the PheWAS were also genetically correlated with most of the smoking and alcohol phenotypes at a standard $p<0.05$. However, “Getting up in morning”, “Nervous feelings”, and the dietary phenotypes were not genetically correlated with smoking after employing correction for multiple comparisons. Similarly, fewer of the top hits were genetically correlated with alcohol phenotypes after control for multiple comparisons. Notably, “nervous feelings” was not genetically correlated with alcohol use. In addition, risk-taking and lifetime number of sexual partners were genetically correlated with being a previous drinker.

Discussion

In both open-source PheWAS analyses, the C allele of rs279858, the top AUD SNP within a GABRA2 LD block, was associated with an increase in nervous feelings, a reduced level of self-reported risk-taking behavior, reduced number of lifetime sexual partners, and shorter duration of mobile phone use. We postulate that these traits may be related to reduced novelty seeking/behavioral inhibition or anxiety that has been identified as a risk factor for alcohol misuse [43,44]. These phenotypes have been shown to underlie anxiety-related symptoms and disorders that have been involved in pathways to AUD [45]. Our analyses of the genetic correlations both within these the anxiety/inhibition phenotypes as well...
with a large number of alcohol related behaviors suggests that the phenotypic links may share overall genetic architecture with AUD in addition to their common association with rs279858.

This work builds upon the substantial body of research that has attempted to identify genetic and environmental risk factors that may inform the multifactorial architecture of AUD. The broad range of phenotypes examined in these PheWAS tools, and particularly in the UKBB, extend far beyond those that are typically captured in EHR data. It is also notable that the effect sizes for the behavioral traits are quite small (~0.005-0.05), which may explain why these associations are only seen in large samples.

The strongest associations that survived correction for multiple comparisons were for anxiety phenotypes and reduced novelty seeking/behavioral inhibition. The association between rs279858 C allele and the anxiety phenotypes in the PheWAS confirms the links between AUD and anxiety disorders, particularly social anxiety, reported in earlier family [46,47] and twin studies [48]. Furthermore, specific anxiety disorder subtypes have been shown to elevate the risk of subsequent onset of both alcohol use and dependence in community surveys worldwide [2,49]. The heterogeneity and complexity of this association has been described in several earlier reviews [50,51]. The convergence of our PheWAS findings with these genetic epidemiologic studies suggests that inquiry into potential mechanisms for the role of genetic factors indexed by this SNP may inform this potential pathway to AUD.

Our finding of the associations between anxiety related phenotypes with the C allele of rs279858 is consistent with reports of an association between externalizing behavior, including substance use and greater risk taking, with the opposite T allele of rs279858 allele [52-56]. In contrast, direct tests of association show that the C allele is associated with AUD [5,21,52]. Analyses of an adolescent sample [52] revealed indirect associations of AUD with the T allele through externalizing factors and direct associations of problematic alcohol use with the C allele. This suggests that the role of GABRA2 in AUD may be multi-faceted and developmentally expressed. In the large nationally representative sample of adults in the National Comorbidity Survey [57], anxiety and externalizing were each independently associated with higher risk of AUD, but there was a potential interactive effect between anxiety and externalizing disorders that was associated with reduced rates of AUD.

The lack of an association between rs279858 with alcohol traits assessed in the UKBB (e.g., consumption measures (alcohol intake frequency, p=0.69) and self-reported alcohol use problems (ever addicted to alcohol, p=0.04)) was not surprising given the lack of previous associations between rs279858 and AUD phenotypes in general population-based samples that were not recruited for substance use disorder studies. However, these analyses do show strong genetic correlations between our top PheWAS hits and alcohol and smoking phenotypes in UKBB. A potential explanation may be that the association of alcohol phenotypes with GABRA2 has been found in samples enriched for AUD but not in a volunteer sample of older adults such as the UKBB, which is predominantly female and among which there was a low frequency of AUD. Another consideration is that a smaller proportion of participants completed some of the alcohol use modules in the UKBB, meaning that these traits have smaller sample sizes than the ones identified in our PheWAS.

GABRA2 has been associated with a number of other traits in GWAS, including epilepsy [17,58], post bronchodilator FEV1/FVC ratio [59], Alzheimer disease and age of onset [18], Mononucleosis [60], general cognitive ability [61], general risk tolerance (MTAG) [19], adventurousness [19], protein quantitative trait loci (liver)[62], and Type 1 diabetes (age at diagnosis)[63]. It is notable that the association findings for general cognitive ability, general risk tolerance, and adventurousness were found in UKBB data. These traits are composite phenotypes formed from the UKBB traits identified in
our PheWAS that include nervous feelings and reduced risk-taking and lend credence to the theory that
the impact of GABRA2 on alcohol-related phenotypes is through anxiety or risk taking.

The effects of the GABRA2 gene in basic animal model studies provide intriguing parallels to our
findings. Numerous phenotypes that overlap our association findings have been found in GABRA2
knockout mice [64] (http://www.informatics.jax.org/marker/phenotypes/MGI:95614). In particular,
GABRA2 has been shown to impact parallel behavioral phenotypes in mice, including increased anxiety-
related response and an abnormal conditioned emotional response, as well as medical phenotypes in
the immune and nervous systems (Supplementary Table 5). Moreover, GABRA2 has been related to
hypersensitivity to the acute effects of alcohol [65] and is a binding site for benzodiazepines known to
decrease anxiety [66]. There is further indirect evidence that neuroactive steroids may be involved in
some of the subjective effects of alcohol reflecting the role of this SNP on both immune function and
alcohol and benzodiazepine effects and metabolism [25,67-69]. It would be interesting to examine the
impact of this marker on the preliminary evidence for the efficacy of gabapentin, a gamma aminobutyric
acid (GABA) analogue, in the treatment of AUD [70].

Strengths & limitations
The main strength of this study is the application of the PheWAS approach to identify potential
influences of the GABRA2 gene that extend beyond the diagnostic code-based EHR. We include a wide
range of other biological and social phenotypes that may reflect the multifactorial etiology of AUD.

This work has several limitations that should be considered in interpreting these findings. First,
phenotypes can appear multiple times in the same database or in different databases. This leads to a
lack of independence in the analyses, which creates a challenge for appropriately adjusting for multiple
comparisons, meaning the Bonferroni method is extremely conservative in this case. Second, the wide
range of phenotypes that extend beyond medically relevant conditions to include behavior,
temperamental traits, and social variables may include variables that are neither empirical or conceptual
correlates of AUD. This may make it difficult to interpret the results in the context of AUD, as
exemplified by the associations with dietary phenotypes. Third, there is overlap in the two databases
that were used for these analyses. Although the use of the UKBB data is a strength of this study, it is also
a limitation because these findings may have overshadowed those of the smaller studies. Fourth, the
applicability of this approach for understanding mechanisms of AUD relies on the relevance of the
GABRA2 gene, which although shown in both enriched samples and basic animal model studies, has not
been identified in the large GWAS studies of AUD. Nevertheless, the similarity in the phenotypic findings
from the PheWAS and the knockout mouse studies suggest that we may be honing in on relevant
phenotypic manifestations of this gene.

Conclusion
In conclusion, we have shown that rs279858 (C effect allele) is associated with anxiety phenotypes,
which may represent pathophysiological intermediaries between GABRA2 and AUD. If replicated, these
findings would confirm the utility of identifying anxiety as a target of prevention and intervention of
AUD. Future in depth investigation of the potential genetic and biologic mechanisms for the link
between this genetic system and its correlates in the development of AUD is warranted.
Funding and Disclosure
Support for the work presented here was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401). The authors declare no conflicts of interest.

Acknowledgments
The support of Hannah L. Bocchinfuso, Resource Coordinator for the Almasy Lab, is invaluable. We are indebted to all of the researchers who deposit their data into central repositories, and those who developed the online tools used to complete this research.

The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators B. Porjesz, V. Hesselbrock, T. Foroud; Scientific Director, A. Agrawal; Translational Director, D. Dick, includes eleven different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, T. Foroud, Y. Liu, M. Plawecik); University of Iowa Carver College of Medicine (S. Kuperman, J. Kramer); SUNY Downstate Health Sciences University (B. Porjesz, J. Meyers, C. Kamarajjan, A. Pandey); Washington University in St. Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, R. Hart, J. Salvatore); The Children's Hospital of Philadelphia, University of Pennsylvania (L. Almasy); Virginia Commonwealth University (D. Dick); Icahn School of Medicine at Mount Sinai (A. Goate, P. Slesinger); and Howard University (D. Scott). Other COGA collaborators include: L. Bauer (University of Connecticut); J. Nurnberger Jr., L. Wetherill, X., Xuei, D. Lai, S. O'Connor, (Indiana University); G. Chan (University of Iowa; University of Connecticut); D.B. Chorlian, J. Zhang, P. Barr, S. Kinreich, G. Pandey (SUNY Downstate); N. Mullins (Icahn School of Medicine at Mount Sinai); A. Anokhin, S. Hartz, E. Johnson, V. McCutcheon, S. Sacccone (Washington University); J. Moore, Z. Pang, S. Kuo (Rutgers University); A. Merikangas (The Children's Hospital of Philadelphia and University of Pennsylvania); F. Aliev (Virginia Commonwealth University); H. Chin and A. Parsian are the NIAAA Staff Collaborators. We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, P. Michael Conneally, Raymond Crowe, and Wendy Reich, for their critical contributions. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA).

Author Contributions
Conceptualized study: AKM, LA; Conducted analyses: AKM, RLK; Drafted manuscript: AKM, LA; Reviewed and edited manuscript: LA, LB, CG, HJE, CK, RLK, JK, AKM, JLM, JN, MHP, BP.

Figure legends
Figure 1: OpenGWAS PheWAS plot
Phenome-wide association study results for rs279858 from the OpenGWAS PheWAS tool (a) and the Open Targets PheWAS tool (b). Traits with a p-value p<0.005 are shown, with the down arrow representing a negative beta, and the up arrow representing a positive beta (C effect allele). In (a) the dashed red line represents a Bonferroni-corrected phenome-wide significance threshold of p<4.00x10^-6. In (b) the dashed red line represents a Bonferroni-corrected phenome-wide significance threshold of p<9.77x10^-5.
Figure 2: Genetic correlations among top UKBB PheWAS hits
The genetic correlations between the top rs279858-associated phenotypes, alcohol, and smoking phenotypes in the UKBB are shown. A single asterisk represents a nominally significant p-value $p<0.05$, and a double asterisk represents a p-value that meets the Bonferroni corrected threshold of p-value $p<2.92 \times 10^{-4}$. The darker blue represents a negative correlation, while the darker red represents a positive correlation.

Tables
Table 1: Top ten associations of rs279858 in the OpenGWAS Analysis
Table 2: Top ten associations of rs279858 in the Open Targets Analysis
Table 3: Heritability of alcohol phenotypes in the UKBB from the MRC IEU OpenGWAS PheWAS

Supplemental Material
Table S1: OpenGWAS rs279858 PheWAS results for $p<0.005$
Table S2: Open Targets rs279858 PheWAS results for $p<0.005$
Table S3: OpenGWAS and Open Targets PheWAS data sources
Table S4: Genetic correlations between the statistically significant UKBB phenotypes from the PheWAS, and UKBB alcohol, tobacco, and substance use phenotypes
Table S5: GABRA2 Mammalian Phenotype Ontology Annotations (MGI:95614)
Figures

Figure 1: OpenGWAS and Open Targets rs279858 PheWAS plots

Phenome-wide association study results for rs279858 from the OpenGWAS PheWAS tool (a) and the Open Targets PheWAS tool (b). Traits with a p-value p<0.005 are shown, with the down arrow representing a negative beta, and the up arrow representing a positive beta (C effect allele). In (a) the dashed red line represents a Bonferroni-corrected phenome-wide significance threshold of p<4.00x10^{-6}. In (b) the dashed red line represents a Bonferroni-corrected phenome-wide significance threshold of p<9.77x10^{-5}. 
Figure 2: Genetic correlations among top UKBB rs279858 PheWAS hits, alcohol, and smoking phenotypes.

The genetic correlations between the top rs279858-associated phenotypes, alcohol, and smoking phenotypes in the UKBB are shown. A single asterisk represents a nominally significant p-value p<0.05, and a double asterisk represents a p-value that meets the Bonferroni corrected threshold of p-value p<2.92x10^-4. The darker blue represents a negative correlation, while the darker red represents a positive correlation.
References


26 Moe JS, Bolstad I, Morland JG, Bramness JG. GABA subunit single nucleotide polymorphisms show sex-specific association to alcohol consumption and mental distress in a Norwegian population-based sample. Psychiatry Res. 2022;307:114257.


Anker JJ, Kushner MG. Co-Occurring Alcohol Use Disorder and Anxiety: Bridging Psychiatric, Psychological, and Neurobiological Perspectives. Alcohol Res. 2019;40(1).


a. rs279858 OpenGWAS PheWAS

Risk taking

Lifetime # of sexual partners

Nervous feelings

Length of mobile phone use

Phenotype Category

Anthropomorphic
Behavioral
Cardiovascular
Chromotype
Cognitive
Demographic
Dietary
Environmental
Immune system
Metabolites
Musculoskeletal
Neurologic
Other Medical
Psychiatric
Respiratory
Substance Use

b. rs279858 Open Targets PheWAS

Lifetime # sexual partners

Risk taking

Feeling nervous

Nervous feelings

Length of mobile phone use

Never/rarely have milk

Phenotype Category

Anthropometric
Behavioral
Cancer
Cardiovascular
Chromotype
Demographic
Dietary
ENT
Environmental
Medication
Musculoskeletal
Other Medical
Psychiatric
Reproductive
Substance Use