Convergent Evidence for Predispositional Effects of Brain Gray Matter Volume on Alcohol Consumption

David A.A. Baranger, Catherine H. Demers, Nourhan M. Elsayed, Annchen R. Knodt, Spenser R. Radtke, Aline Desmarais, Lauren R. Few, Arpana Agrawal, Andrew C. Heath, Deanna M. Barch, Lindsay M. Squeglia, Douglas E. Williamson, Ahmad R. Hariri, and Ryan Bogdan

ABSTRACT
BACKGROUND: Alcohol use has been reliably associated with smaller subcortical and cortical regional gray matter volumes (GMVs). Whether these associations reflect shared predisposing risk factors or causal consequences of alcohol use remains poorly understood.

METHODS: Data came from 3 neuroimaging samples (N = 2423), spanning childhood or adolescence to middle age, with prospective or family-based data. First, we identified replicable GMV correlates of alcohol use. Next, we used family-based and longitudinal data to test whether these associations may plausibly reflect a predispositional liability for alcohol use or a causal consequence of alcohol use. Finally, we used heritability, gene-set enrichment, and transcriptome-wide association study approaches to evaluate whether genome-wide association study–defined genomic risk for alcohol consumption is enriched for genes that are preferentially expressed in regions that were identified in our neuroimaging analyses.

RESULTS: Smaller right dorsolateral prefrontal cortex (DLPFC) (i.e., middle and superior frontal gyri) and insula GMVs were associated with increased alcohol use across samples. Family-based and prospective longitudinal data suggest that these associations are genetically conferred and that DLPFC GMV prospectively predicts future use and initiation. Genomic risk for alcohol use was enriched in gene sets that were preferentially expressed in the DLPFC and was associated with replicable differential gene expression in the DLPFC.

CONCLUSIONS: These data suggest that smaller DLPFC and insula GMV plausibly represent genetically conferred predispositional risk factors for, as opposed to consequences of, alcohol use. DLPFC and insula GMV represent promising biomarkers for alcohol-consumption liability and related psychiatric and behavioral phenotypes.

Keywords: Alcohol, Gene expression, Heritability, Imaging, Longitudinal, Structure

https://doi.org/10.1016/j.biopsych.2019.08.029
examined whether genetic risk for alcohol consumption is associated with genes and genetically conferred differences in gene expression that are preferentially expressed in regions identified by neuroimaging analyses or the brain more generally. Here, we applied gene-set enrichment, partitioned heritability, and transcriptome-wide association study (TWAS) (27) analyses to genome-wide association study (GWAS) summary statistics from the UK Biobank (N = 112,117) (28) and Alcohol Genome-Wide Consortium and the Cohorts for Heart and Aging Research in Genomic Epidemiology Plus consortia (AlcGen/CHARGE+) (N = 70,460) (29) studies of alcohol consumption, and RNA-seq data from the Genotype-Tissue Expression (GTEx) project (n = 81–103) (30) and the Common Mind Consortium (N = 452) (31).

METHODS AND MATERIALS

Participants

Neuroimaging data were drawn from 3 independent studies—the Duke Neurogenetics Study (DNS) (n = 1303) (26), the Human Connectome Project (HCP) (n = 897) (24), and the Teen Alcohol Outcomes Study (TAOS) (n = 223) (29)—that assessed behavioral, experiential, and biological phenotypes among young adult college students (DNS sample), young to middle-aged adults (HCP sample), and children and adolescents (TAOS sample). The DNS and TAOS studies collected longitudinal data on alcohol use subsequent to the baseline scan. The family-based HCP sample is composed of twin and non-twin siblings. All studies followed protocols approved by relevant institutional review boards and remunerated participants. Additional information regarding each sample is provided in Supplement 1.

Alcohol Use Assessment

Alcohol use in the DNS was assessed at baseline (past 12-month use) and follow-ups (questions modified to reflect use following the prior assessment) using the Alcohol Use Disorders Identification Test consumption subscale (AUDIT-C) (DNS: \( \alpha = .85 \); mean = 3.76; SD = 2.64; range = 0–12) (32,33). The AUDIT-C was approximated (aAUDIT-C) in the HCP sample (\( \alpha = .786 \); mean = 3.42; SD = 2.65; range = 0–12) and TAOS (\( \alpha = .893 \); mean = 0.45; SD = 1.26; range = 0–9) using questions from Semi-Structured Assessment for the Genetics of Alcoholism (34) and Substance Use Questionnaire (35), respectively. In TAOS, the initiation of alcohol use was defined as attaining a score of \( \geq 1 \) on the aAUDIT-C (i.e., participant reports consuming \( \geq 1 \) full alcoholic beverage; n = 82 started during the study; age (in years): mean = 16.68; SD = 1.39; range 14.12–19.64). Supplement 1 contains additional details.

Covariates

Variables known to be correlated with alcohol consumption, GMV, or both were included as covariates in all analyses: age (36–38), sex (37–40), ethnicity (40,41), socioeconomic status (SES) (36–40), early-life and recent-life stress (42–44), and intracranial volume. In adult samples (DNS, HCP), the presence of any nonsubstance Axis I DSM-IV psychiatric disorder was included as a covariate. As the TAOS sample was composed of children and adolescents enriched for a family history of depression, Tanner stage and depressive symptoms were included as covariates. Supplement 1 contains additional details, including consideration of nicotine and cannabis use.

Magnetic Resonance Imaging Processing

Acquisition parameters and GMV processing for each study are described in Supplement 1.

Statistical Analyses

Discovery—DNS. A whole-brain voxel-based morphometry generalized linear model regression analysis was conducted using SPM12 to test whether alcohol consumption (AUDIT-C) is associated with differences in GMV. Covariates included sex, age, self-reported race/ethnicity (i.e., not-white/white, not-black/black, not Hispanic/Hispanic), scanner identification (2 identical scanners were used), intracranial volume, presence of a diagnosis other than alcohol or substance abuse or dependence, perceived stress, parental education level, early-life stress (assessed via the Childhood Trauma Questionnaire), and perceived SES. Analyses were thresholded at \( p < .05 \) familywise error corrected with a cluster extent threshold of 10 contiguous voxels (\( k_c = 10 \)) across the entire search volume.

Replication—HCP. Analyses examined whether alcohol consumption (aAUDIT-C) predicted GMV only within regions of interest (ROIs) where associations were observed in the discovery DNS sample (Figure 1, Table S1 in Supplement 1). ROIs were defined by the Automated Anatomic Labeling atlas (45). A voxelwise generalized linear model regression was conducted using multilevel block permutation–based nonparametric testing (FSL PALM v.alpha103; tall approximation \( p < .10 \) with 5000 permutations), which accounts for the family structure of the HCP data while correcting for multiple comparisons (46–48). Covariates included sex, age, self-reported race and/or ethnicity, intracranial volume, twin and/or sibling status ( dizygotic or not, monozygotic or not, half-sibling or not), presence of a diagnosis other than alcohol or substance abuse or dependence, perceived stress, education level, and SES. Analyses were thresholded at \( p < .05 \) familywise error corrected with a cluster extent threshold of 10 contiguous voxels (\( k_c = 10 \)).

Post Hoc Analyses

Total anatomical GMV of ROIs associated with alcohol use in both the DNS and HCP (i.e., right insula and middle and superior frontal gyr) (see Results) were extracted from both datasets for post hoc analyses. Total volumes were used so that effect sizes would not be inflated by selecting only voxels that were specifically associated with the outcome of interest (49).

Heritability. SOLAR-Eclipse software (http://solar-eclipse-genetics.org) (50), in conjunction with the R package solarius (51), which uses maximum likelihood variance decomposition methods, was used to estimate phenotypic heritability (\( h^2 \), the fraction of phenotypic variance attributable to additive genetic factors), as well as genetic (\( \rho_g \)) and unique environmental (\( \rho_n \)) correlations (i.e., the fraction of the correlation between 2 phenotypes that is attributable to either additive genetic or
Alcohol Consumption and Brain Volume

Figure 1. Identification of replicable volumetric associations with alcohol consumption. This statistical parametric map illustrates regions of reduced brain volume associated with increased alcohol consumption (Table S1 in Supplement 1), which are overlaid onto a canonical structural brain image (Montreal Neurological Institute coordinates and statistics; Duke Neurogenetics Study [DNS]; \( p < .05 \), familywise error whole-brain corrected, \( \geq 10 \) contiguous voxels; Human Connectome Project [HCP]; \( p < .05 \), familywise error region-of-interest corrected, \( \geq 10 \) contiguous voxels). Alcohol consumption was not associated with increased volume in any region. Notably, in the HCP dataset, the superior frontal gyrus cluster extended into the right middle frontal gyrus and was located relatively far (34 mm dorsal) from the original right superior frontal cluster identified in the DNS. In contrast, this peak in the HCP was located 11.6 mm away from the right middle frontal peak identified in the DNS. Thus, for the purposes of post hoc analyses, the combined volume of both the right middle and superior frontal gyri cortices was extracted from both samples. Cluster overlap at an uncorrected threshold and comparison of effect sizes are shown in Figures S2 and S3 in Supplement 1.

Discordant Twin and Nontwin Sibling Analysis. Following evidence that alcohol consumption is coherable with volume of the right insula and middle and superior frontal gyri (see Results), we examined whether same-sex twin and nontwin sibling pairs discordant for alcohol consumption differed from each other on brain volume in the HCP sample. These analyses examined whether aAUDIT-C was associated with insular or middle and superior frontal volume after accounting for sibling-shared genetic background and experience. Same-sex siblings were considered “high alcohol consumers” if their aAUDIT-C score was >0.5 SD above the sample mean (aAUDIT-C > 4.67), or “low alcohol consumers” if their score was <0.5 SD below the sample mean (aAUDIT-C < 1.54), respectively. A discordant sibling pair was defined as a pair who were both in the same category of consumption (i.e., high or low) and additionally scored within 1 SD of each other low alcohol concordant pairs: \( n = 117; \) aAUDIT-C mean = 0.84, SD = 0.77; high alcohol concordant pairs: \( n = 54; \) aAUDIT-C mean = 7.08, SD = 1.4). There were 72 discordant sibling pairs (“low discordant”; aAUDIT-C mean = 1.25, SD = 0.73; “high discordant”; aAUDIT-C mean = 6.47, SD = 1.67). Participants could be included in >1 pair (\( n = 368 \) individuals) when considering relationships with multiple siblings. Discor-
dancy analyses were conducted using linear mixed models, using the psych (52) and lm4 (53) packages in R (54) to account for the multiple-sibling structure within families. Covariates were identical to those used in neuroimaging analyses. Additional information on models tested and their interpretation are available in Supplement 1.

DNS Longitudinal Changes in Alcohol Consumption. Hierarchical density-based clustering (R dbscan package) (55) was used to detect and remove temporal outlier responses to the follow-up questionnaire (Supplement 1, Figure S5 in Supplement 1). The R nlme package (56) was used to fit a longitudinal multilevel linear model to examine whether GMV predicted AUDIT-C at follow-ups. The model included both random intercept and random slope components with a continuous autoregressive correlation structure. Time was coded as both linear and quadratic age at the date of response (baseline or follow-up). Models tested the interaction between brain volume and age (i.e., does baseline ROI volume predict a different slope of change in drinking behavior as participants age?). Covariates were z-scored, and they were identical to those used in neuroimaging analyses, with the addition of second-order interactions between covariates and primary variables (57,58). Each of the 2 ROIs was tested in a separate model, and \( p \) values were false discovery rate (FDR) corrected (i.e., 4 tests: middle \( \times \) linear age, superior \( \times \) linear age, middle \( \times \) quadratic age, superior \( \times \) quadratic age).

TAOS Longitudinal Initiation of Alcohol Use. The R lm4 package (59) was used to fit a longitudinal logistic

individual-specific environmental factors, respectively) of GMV and alcohol consumption. Post hoc analyses (Supplement 1) assessed the contribution of shared environmental factors to phenotypic correlations (\( \rho_e \)). These analyses were conducted among the subset of related participants from the HCP (\( n = 804; 293 \) families, 115 monozygotic and 64 dizygotic twin pairs and 422 nontwin siblings, excluding singletons and half-siblings). Covariates were identical to those in neuroimaging analyses. To ensure normality of measurements and accuracy of estimated parameters, an inverse normal transformation was applied to all continuous traits and covariates prior to analyses.
multilevel model, which tested whether baseline brain volume in nondrinking adolescents predicted future initiation of alcohol use. The model included both random intercept and random slope components, and time was coded as both the linear and quadratic age at the date of response. The model tested the interaction between GMV and age (i.e., does baseline ROI volume predict a different likelihood of initiation as participant’s age?). Covariates were z-scored, and they included demographic variables (age, sex, ethnicity, and SES), stress (Childhood Trauma Questionnaire and Stressful Life Events Schedule), Tanner stage, Mood and Feelings Questionnaire scores, family history of depression, age at magnetic resonance imaging scan, and intracranial volume. Second-order interactions between covariates and primary variables (e.g., middle frontal volume × sex, middle superior volume × SES, age × sex, age × SES) were also included (58). Each of 2 ROIs, right superior frontal cortex and right middle frontal cortex, was tested in a separate model. Insula volume was excluded, as it was not significant in DNS longitudinal analyses. The p values were subsequently FDR corrected (4 tests).

**Single Nucleotide Polymorphism-Based Enrichment.** We tested whether the single nucleotide polymorphism (SNP)-based heritability of alcohol consumption is enriched in brain-expressed gene sets and whether this enrichment is specific to any region. Stratified linkage disequilibrium–score regression (60–62) was applied to summary statistics from the GWAS of alcohol consumption in the UK Biobank (N = 112,117) (28). Tissue-enriched gene sets were generated using data from the GTEx Consortium (30,61). A gene is assigned to a gene set if it shows greater enrichment in that tissue than 90% of genes. It was further tested whether genetic associations with alcohol consumption are enriched in brain-expressed gene sets using the analysis tool MAGMA (63) implemented through the platform FUMA (64).

Transcriptome-wide Analysis. We tested whether genetic risk for alcohol consumption is predictive of differences in postmortem gene expression. Precomputed gene-expression RNA-sequencing weights for 9 brain regions and the liver from the GTEx project (30) were analyzed using the FUSION suite (27). Analyses used GWAS results for alcohol consumption from the UK Biobank (28). Results were Bonferroni-corrected for n = 9839 tests across the 10 tissues (Supplemental Data). Replication was tested using an independent alcohol-consumption GWAS (N = 70,460) (29) and dorsolateral prefrontal cortex (DLPFC) gene-expression weights from the CommonMind Consortium (31). As the gene that showed the strongest association in the discovery dataset was not present in the replication data, we examined whether any of the Brodmann area 9 (BA 9) gene-expression associations at FDR-corrected p < .05 were significant in the replication data (see Results). Replicated genes were probed for association with other GWAS phenotypes using a phenome-wide association study (PheWAS) implemented through the “GWAS Atlas” browser (65). BA 9 in the GTEx dataset and DLPFC in the CommonMind consortium dataset overlap with the prefrontal regions implicated in our neuroimaging analyses (i.e., middle and superior frontal gyri) (see Results). No postmortem insula data were available.

**RESULTS**

Whole-brain discovery analyses in DNS revealed that greater alcohol consumption is associated with lower GMV across 8 clusters (Figure 1, Table 1) that encompass regions identified in prior studies of unselected samples (5,17) and among individuals with alcohol use disorder (4,6). The associations with 2 of these clusters (right insula, right middle and superior frontal gyri) replicated within an ROI analysis in the HCP (Figure 1, Table 1). Post hoc analyses revealed that effect sizes were nearly identical in the 2 samples, that results were equivalent when excluding nondrinkers (Figure S2 in Supplement 1), and that associations between alcohol use and GMV remained largely unchanged and significant when correcting for tobacco

**Table 1. Location of Volumetric Reductions Associated With Alcohol Consumption**

<table>
<thead>
<tr>
<th>Index</th>
<th>No. of Voxels</th>
<th>p, Familywise Error Corrected</th>
<th>t</th>
<th>x, mm</th>
<th>y, mm</th>
<th>z, mm</th>
<th>AAL-Atlas Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duke Neurogenetics Study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>279</td>
<td>.003</td>
<td>5.24</td>
<td>27</td>
<td>39</td>
<td>25</td>
<td>R middle frontal</td>
</tr>
<tr>
<td>1.b</td>
<td></td>
<td>.010</td>
<td>4.95</td>
<td>32</td>
<td>50</td>
<td>23</td>
<td>R middle frontal</td>
</tr>
<tr>
<td>2</td>
<td>344</td>
<td>.004</td>
<td>5.14</td>
<td>56</td>
<td>3</td>
<td>0</td>
<td>R superior temporal</td>
</tr>
<tr>
<td>2.b</td>
<td></td>
<td>.006</td>
<td>5.05</td>
<td>48</td>
<td>6</td>
<td>–5</td>
<td>R insula</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>.005</td>
<td>5.08</td>
<td>0</td>
<td>63</td>
<td>–3</td>
<td>L medial orbital frontal</td>
</tr>
<tr>
<td>4</td>
<td>76</td>
<td>.007</td>
<td>5.03</td>
<td>38</td>
<td>18</td>
<td>51</td>
<td>R middle frontal</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>.007</td>
<td>5.03</td>
<td>–3</td>
<td>–33</td>
<td>42</td>
<td>L middle cingulum</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>.019</td>
<td>4.80</td>
<td>2</td>
<td>27</td>
<td>30</td>
<td>R middle cingulum</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>.023</td>
<td>4.75</td>
<td>29</td>
<td>62</td>
<td>8</td>
<td>R superior frontal</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>.029</td>
<td>4.59</td>
<td>2</td>
<td>33</td>
<td>45</td>
<td>R medial superior frontal</td>
</tr>
<tr>
<td>Human Connectome Project</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>42</td>
<td>.003</td>
<td>4.92</td>
<td>38</td>
<td>12</td>
<td>0</td>
<td>R insula</td>
</tr>
<tr>
<td>2</td>
<td>88</td>
<td>.008</td>
<td>4.70</td>
<td>20</td>
<td>24</td>
<td>42</td>
<td>R superior/middle frontal</td>
</tr>
</tbody>
</table>

AAL, Automated Anatomical Labeling; L, left; R, right.

*Coordinates are provided in Montreal Neurological Institute space.*

4 Biological Psychiatry ■ ■, 2019; ■■■■ www.sobp.org/journal
Shared genetic predisposition between alcohol consumption and brain volume. In the Human Connectome Project (HCP) sample, (A) alcohol consumption scores (approximated Alcohol Use Disorders Identification Test consumption subscale scores [aAUDIT-C]) and gray matter volume of the right insula and right middle and superior frontal cortices were all observed to be heritable: aAUDIT-C: 51.79%, \( p < 2.2 \times 10^{-16} \); insula: 69.83%, \( p < 2.2 \times 10^{-16} \); frontal: 74.46%, \( p < 2.2 \times 10^{-16} \) (Table S1 in Supplement 1). (B) Significant phenotypic correlations between aAUDIT-C scores and volumes of the right insula and middle and superior frontal cortices were attributable to shared genetic factors (insula: \( -0.2314, p = .0022 \); frontal: \( -0.2192, p = .0054 \)) but not unique environmental factors (Table S1 in Supplement 1). Distribution of (C) right insula and (D) right middle and superior frontal volumes by alcohol exposure group. High = aAUDIT-C score > sample mean + 0.5 SD (i.e., > 4.67); Low = aAUDIT-C score < sample mean – 0.5 SD (i.e., < 1.54); Concordant = both siblings are in the same alcohol exposure group; Discordant = one sibling is in the high group, while the other is in the low group. Contrast comparisons found evidence for predispositional effects of brain volume on alcohol consumption in both cases (insula: graded liability: \( \beta = -0.0037, p = .049 \), predispositional: \( \beta = 0.0037, p = .0006 \); frontal: predispositional: \( \beta = 0.0019, p = .029 \) (Table S2 in Supplement 1).

and cannabis use (Table S11 in Supplement 1). Statistics are presented in the table and figure legends.

Family-based analyses in the HCP (n = 804) revealed that alcohol consumption and GMV of the right insula and right middle/superior frontal gyri are moderately to largely heritable (Figure 2A; Table S1 in Supplement 1). Moreover, decomposition analyses showed that phenotypic correlations between frontal and insular GMV and alcohol consumption are attributable to shared genetic, but not unique environmental, influences (Figure 2B, Table S1 in Supplement 1). Post hoc analyses confirmed that shared environmental factors did not significantly contribute to the correlation of alcohol consumption and GMV (Table S1 in Supplement 1). Analyses within twin and sibling pairs in the HCP sample who were concordant or discordant for alcohol use revealed that relative to siblings who were concordant for low alcohol use, siblings who were concordant for high use or discordant for use (i.e., 1 high use, 1 low use) had lower insular and frontal GMVs (Figure 2C, D; Table S2 in Supplement 1). Further, GMVs did not differ between low and high alcohol-using members of discordant pairs. As shared genetic and familial factors are matched within pairs, this pattern of results suggests that smaller frontal gyri and insula GMVs may reflect preexisting vulnerability factors associated with alcohol use, as opposed to a consequence of alcohol use.

Using available longitudinal data from the DNS (n = 674), lower GMV of the right frontal gyri, but not insula, predicted increased future alcohol consumption, over and above baseline consumption, but only in individuals who are under the legal age of drinking (i.e., younger than 21 years of age) in the United States (Figure 3A, Table S3 in Supplement 1). Similarly, in the TAOS longitudinal sample of children and adolescents, lower right middle and superior frontal gyri GMV predicted the initiation of alcohol use at an earlier age in those who were nondrinkers at baseline (Figure 3B, C; Table S4 in Supplement 1).

Gene-based association and partitioned heritability enrichment analyses of the UK Biobank GWAS of alcohol consumption revealed enrichment only among brain gene sets
Moreover, BA 9, which overlaps with the frontal region identified in neuroimaging analyses, was among the regions with strongest enrichment (Figure S4 in Supplement 1, Supplemental Data). A TWAS analysis of these GWAS data similarly found that genetic risk for alcohol consumption was significantly associated with differences in gene expression across the brain within the GTEx dataset, including expression of C16orf93 within BA 9 (Figure 5, Table 2, Supplemental Data). C16orf93 was not available in the TWAS replication dataset [i.e., the dataset of Schumann et al. (29) and the CommonMind Consortium (31)] (Table 2). Three additional genes survived FDR correction in BA 9, two of which (i.e., CWF19L1 and C18orf8) were available in the TWAS replication dataset (Figure 5, Table 2)1. Genomic risk for alcohol consumption was significantly predictive of differential expression of CWF19L1 and C18orf8 within the DLPFC of our TWAS replication dataset, in the same direction as was observed in the discovery dataset (Table 2, Supplemental Data). Notably, genetic risk for alcohol consumption was not significantly associated with the expression of any gene in the liver (Figure 5). A phenome-wide association study using the GWAS Atlas revealed evidence that both CWF19L1 and C18orf8 have been implicated in a host of phenotypes, including psychiatric conditions and related traits such as executive function and schizophrenia (CWF19L1), and substance use (C18orf8) (Supplemental Data).

Behavioral variables that might mediate links between brain structure and alcohol consumption (i.e., IQ, delay discounting, self-reported impulsivity, negative urgency, and neuroticism) and that were available in the DNS and HCP data sets and were tested for association with GMV. Despite nominally significant associations in the DNS between delay discounting and IQ and GMV of the right frontal cortex, none of these associations replicated within the HCP data or were robust to

---

1Notably, correcting for only tests within BA 9 based on our neuroimaging results, CWF19L1 remains significant following Bonferroni correction.

---

Figure 4. Tissue-specific enrichment of alcohol-consumption genomic risk. Enrichment of alcohol-consumption genome-wide association study (UK Biobank, N = 112,117) (A) associations and (B, C) heritability, in gene sets defined by the relative expression of genes (A, B) across all tissues and (C) within the brain, in the Genotype-Tissue Expression project dataset (Supplemental Data). The x-axis and color scale represent the significance of the enrichment (negative logarithmic scale of the p value). Solid, dashed, and dotted lines represent Bonferroni-corrected, false discovery rate-corrected, and nominally significant p values, respectively. BA, Brodmann area; EBV, Epstein-Barr virus.
Alcohol Consumption and Brain Volume

Figure 5. Transcriptome-wide association study of alcohol consumption predicting gene expression. Genetic risk for alcohol consumption according to the UK Biobank genome-wide association study ($n = 112,117$) is associated with differences in human postmortem gene expression (Genotype-Tissue Expression project; $n = 81–103$), including frontal cortex Brodmann area (BA) 9 (Supplemental Data). Notably, associations in the liver (far-right panel) do not survive Bonferroni correction for multiple comparisons, though 4 are significant at a less-stringent false discovery rate–based correction. The y-axis represents the significance of the association. Solid, dashed, and dotted lines represent Bonferroni-corrected, false discovery rate–corrected, and nominally significant p values, respectively.

DISCUSSION

We report convergent evidence that smaller GMVs of the right insula and DLPFC (i.e., middle and superior frontal gyri) plausibly represent genetically conferred liabilities that promote early alcohol use. First, we show that smaller GMVs of the right insula and DLPFC were replicably associated with alcohol use in 2 large neuroimaging samples. Second, family-based data provide evidence that these associations are attributable to shared genetic factors with no evidence of a causal association, or that shared or unique environmental factors contribute to this association. Third, reduced DLPFC volume prospectively predicted future alcohol use among young adults as well as alcohol use initiation during adolescence among children and adolescents who were unexposed to alcohol at baseline. Finally, we found evidence that genomic risk for alcohol use is enriched among genes that are preferentially expressed within the DLPFC and is replicably predictive of gene expression in the DLPFC. Collectively, these convergent data suggest that lower GMVs in the middle and superior frontal gyri and insula may represent a preexisting genetic liability for drinking that could serve as a prognostic biomarker. Further, these data suggest that the alcohol use in the general population does not induce reductions in GMV, at least as measured using magnetic resonance imaging, as has been previously hypothesized (5,7,9). It is possible that reduced GMVs in the middle and superior frontal gyri and insula may promote alcohol use, increasing the likelihood of heavy use, which may then further potentiate GMV loss in these regions and others (9,11,12).

A few notable points within our data require additional interpretation. In the longitudinal child and adolescent sample of baseline nonusers (i.e., the TAOS sample), we found that DLPFC GMV prospectively predicts an early age of drinking initiation. In the DNS longitudinal prospective data of young adults, reduced GMV in these regions also predicted future alcohol use, even after accounting for the extent of baseline alcohol use. However, this characteristic was only predictive up to 20.85 years of age. It is possible that risk conferred by reduced GMV in the DLPFC is developmentally constrained or may be minimized by environmental differences in permissivity or legality, as the legal drinking age in the United States is 21 years (66).

We found no compelling evidence that behaviors that have been speculated to contribute to alcohol use (e.g., executive function, negative urgency, and impulsivity) are associated with prefrontal or insula GMV, leaving the behavioral mechanisms through which these GMVs may influence alcohol use unclear. DLPFC GMV was negatively correlated with delay discounting and using alcohol to cope with stress in our young adult sample (DNS sample) at nominal levels of significance, while these behavior characteristics and DLPFC GMV were unlinked in our young and middle-age adult sample (Tables S9, S10 in Supplement 1).

Table 2. TWAS Discovery and Replication

<table>
<thead>
<tr>
<th>Gene</th>
<th>CHR</th>
<th>Discovery—GTEX Frontal BA 9</th>
<th>Replication—CMC DLPFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Locus Start</td>
<td>Locus End</td>
<td>TWAS z</td>
</tr>
<tr>
<td>C18orf93</td>
<td>16</td>
<td>30772519</td>
<td>30772656</td>
</tr>
<tr>
<td>CWF19L1</td>
<td>10</td>
<td>102000000</td>
<td>102000000</td>
</tr>
<tr>
<td>PHBP9</td>
<td>10</td>
<td>102000000</td>
<td>102000000</td>
</tr>
<tr>
<td>C18orf98</td>
<td>18</td>
<td>21083473</td>
<td>21110576</td>
</tr>
</tbody>
</table>

Summary of transcriptome-wide association study (TWAS) results in the frontal cortex, conducted with FUSION. Discovery analysis: UK Biobank genome-wide association study ($N = 112,117$) and Genotype-Tissue Expression (GTEX) project gene expression ($N = 92$). Replication analysis: Alcohol Genome-Wide Consortium and the Cohorts for Heart and Aging Research in Genomic Epidemiology Plus consortia (AlcGen/CHARGE+); CommonMind Consortium (CMC) gene expression ($N = 452$). Empty rows in replication data indicate that the gene was not present in replication dataset. The p values are all uncorrected—all associations listed survive false discovery rate correction. Expression quantitative trait loci are available in the Supplemental Data.

BA, Brodmann area; CHR, chromosome; DLPFC, dorsolateral prefrontal cortex.
S10 in Supplement 1). This finding suggests that these behavioral factors may represent mechanisms through which GMV influences alcohol use in adolescence and young adulthood, potentially contributing to continued use, while GMV is uncorrelated with these behaviors as measured in later life. Nonetheless, it is also plausible that these nominally significant findings represent false-positives.

While both the discovery analysis in the DNS sample and the replication analysis in the HCP sample showed that alcohol use was significantly correlated with reduced GMV, the voxels of strongest association only partially overlap. Post hoc analyses found that the effect sizes of the association with atlas-defined ROIs were nearly identical in the 2 samples (Figure S2 in Supplement 1), further supporting the interpretation that identified GMV correlations with alcohol use are replicable. The limited overlap of peaks between the samples likely reflects the lower power in the HCP sample, which is a result of its smaller sample size and the family structure of the data, which resulted in even fewer independent observations. Several GMV findings in the DNS did not replicate, an outcome that may be attributable to differences between the samples (e.g., age), though the possibility that they are false-positives, or that null findings in the HCP sample are false-negatives, cannot be ruled out.

Substantiating the idea that it is biologically plausible that reduced GMV in the DLPFC represents a preexisting genetic liability for drinking, genomic risk for alcohol use was enriched only within brain gene sets. BA 9, which overlaps with the DLPFC regions identified in our neuroimaging analyses, was among the regions of strongest enrichment (Figure 4). Further, TWAS analyses revealed replicable evidence that genomic risk for alcohol use is associated with differential expression of CWF19L1 and C18orf8 within BA 9. While the function of these genes is not understood, both have been previously implicated in psychopathology and related traits, including schizophrenia, substance use, and cognition (Supplement 1), with rare mutations in CWF19L1 causing autosomal recessive cerebellar ataxia (66,67), which is characterized by a loss of control of bodily movements, as well as developmental delay and mental retardation. Additional discussion of these findings and their limitations is presented in Supplement 1.

Given evidence that genetic liability is shared across substance use involvement (67) and other forms of psychopathology (68), our findings may generalize to other substances and overall psychopathology risk. While enrichment analyses implicate only brain pathways and TWASs identify replicable associations between genetic risk for alcohol consumption and gene expression in the frontal cortex, we cannot rule out the possibility that our observed effects are partially mediated by altered functioning of other pathways, such as alcohol metabolism in the liver (69). Moreover, the present results do not distinguish between reduced GMV as part of the mechanism by which genetic risk affects drinking behavior (10,70) and a pleiotropic effect of genetic risk on multiple outcomes (20).

We must note assumptions of heritability analyses including random mating and equal environments (71). On one hand, violations of the random-mating assumption would result in downwardly biased estimates of coheritability (72), though there is evidence that this bias, when present, is modest (73,74). An additional limitation of heritability analyses is that the statistical power required to parse the role of overlapping genetic and shared environmental factors is substantial, and it is beyond the scope of the current analysis. However, given that we found little evidence that shared environment contributes to the correlation of GMV and alcohol use, shared environment is an unlikely confound.

While our study is limited by our sample size, particularly of discordant siblings and longitudinal analyses, a major strength of our results is the convergent evidence provided by the different study designs (75). We note that our cross-sectional analyses of alcohol consumption and longitudinal analyses of adolescent use initiation are the largest to date that we know of. A primary limitation of our gene-expression analyses is that both of the gene-expression datasets included alcohol-exposed donors. Given the wide prevalence of alcohol use across the world (76), it will likely be impossible to ever definitively confirm in human adults that alcohol use is not confounding these results. Notably, none of the identified genes have been found to be differentially expressed in the frontal cortex of donors with alcoholism (77,78). Our analyses are also limited by the omission of the insula from the gene-expression data, precluding a comparison of the gene-expression correlates between the insula and frontal cortex.

Limitations notwithstanding, our study provides convergent evidence that smaller GMV in the insula and DLPFC associated with alcohol use may represent a genetically conferred liability that promotes early alcohol use. While early alcohol use may in turn lead to accelerated volume loss within these and other regions, these findings challenge predominant interpretations that smaller brain volumes tied to alcohol use emerge primarily from the atrophy-inducing effects of alcohol. As larger prospective samples are acquired (e.g., via the Adolescent Brain Cognitive Development study) (79), it will be interesting to examine the interplay of genetic risk and substance use on the trajectories of brain development.

ACKNOWLEDGMENTS AND DISCLOSURES

Data for this study were provided by the Human Connectome Project, WU-Minn Consortium (Grant No. 1U54MH091657; principal investigators David Van Essen, Ph.D., and Kamil Ugurbil, Ph.D.), which was funded by the 16 National Institutes of Health (NIH) institutes and centers that support the NIH Blueprint for Neuroscience Research and the McDonnell Center for Systems Neuroscience at Washington University. The Duke Neurogenetics Study was supported by Duke University and the National Institute on Drug Abuse (Grant No. DA033369 [to ARH]). The Teen Alcohol Outcomes Study (TAOS) was supported by Duke University and the National Institute on Alcohol Abuse and Alcoholism (Grant No. AA016274 [to DEWI]). DAAB was supported by the NIH (Grant No. T32-GM008151) and the National Science Foundation (Grant No. DGE-1143954). LRF was supported by the National Institute on Alcohol Abuse and Alcoholism (Grant No. AA023693). CHD was supported by the NIH (Grant Nos. T32-DA007313 and T32-GM081739). AA was supported by the National Institute on Drug Abuse (Grant No. SK02DA1172573). ARH received additional support from the National Institute on Drug Abuse (Grant No. DA031579) and the National Institute on Aging (Grant No. AG049789). LMS was supported by the NIH (Grant Nos. K23...
Alcohol Consumption and Brain Volume

AA025399 and U01 DA041093. RB was supported by the Klingenstein Third Generation Research and the NIH (Grant Nos. R01-AG045231, R01- HD083814, and R01-AG052564).

This article was published as a preprint on bioRxiv: https://doi.org/10.1101/299149.

All data pertaining to this study are available on request to the corresponding author. Datasets may also be accessed at the locations listed: Duke Neurogenetics Study (DNS): https://www.harrilab.com/projects/procedures.html

Human Connectome Project (HCP): https://www.humanconnectome.org/

Tea Alchon Onset Study (TAOS): Requests for data access should be submitted to the study principal investigator, Douglas Williamson, Ph.D.: douglas.williamson@duke.edu

UK Biobank: http://www.ukbiobank.ac.uk


References


Alcohol Consumption and Brain Volume


