Convergent evidence for predispositonal effects of brain volume on alcohol consumption

David AA Baranger[1], Catherine H. Demers[1], Nourhan M. Elsayed[2] , Annchen R. Knodt[3], Spenser R. Radtke[3], Aline Desmarais[1], Lauren R. Few[4], Arpana Agrawal[4], Andrew C. Heath[4], Deanna M. Barch[1,4], Lindsay M. Squeglia[5], Douglas E. Williamson[2,6], Ahmad R. Hariri[3], Ryan Bogdan[1]

[1] Department of Psychological and Brain Sciences, Washington University, St. Louis, MO
[2] Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC
[3] Department of Psychology and Neuroscience, Duke University, Durham, NC
[4] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO
[5] Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC
[6] Durham VA Medical Center, Durham, NC

Author's note: Correspondence concerning this article should be addressed to David Baranger (email: dbaranger@gmail.com, address: CB 1125 Psychological and Brain Sciences Bldg Room 453, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA) and Ryan Bogdan (email: rbogdan@wustl.edu, address: CB 1125 Psychological and Brain Sciences Bldg Room 453, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA

Funding: Data for this study were provided by the Human Connectome Project, WU-Minn Consortium (principal investigators: David Van Essen, PhD, and Kamil Ugurbil, PhD; grant 1U54MH091657) funded by the 16 National Institutes of Health institutes and centers that support the National Institutes of Health Blueprint for Neuroscience Research, as well as by the McDonnell Center for Systems Neuroscience at Washington University. The Duke Neurogenetics Study is supported by Duke University and NIDA (DA033369). The Teen Alcohol Outcomes Study (TAOS) was supported by Duke University and NIAAA (AA016274) (DEW). DAAB was supported by NIH (T32-GM008151) and NSF (DGE-1143954). LF was supported by NIAAA (AA023693). CHD was supported by NIH (T32-DA007313 and T32-GM081739). AA was supported by NIDA (5K02DA32573). ARH receives additional support from NIDA (DA031579) and NIA (AG049789). LS was supported by K23 AA025399 and U01 DA041093. RB was supported by the Klingenstein Third Generation Research and NIH (R01-AG045231, R01-HD083614, R01-AG052564).

Abstract

Alcohol consumption accounts for 5% of global disease burden. Using family and longitudinal data from three independent samples (Ns = 1,303, 897, & 223) spanning childhood/adolescence to middle age (ages 11-37), we identify replicable reductions in gray matter volumes of frontal gyri, which are genetically-conferred, and prospectively predict alcohol consumption and initiation of use. Further, gene expression in the frontal cortex is replicably associated with genetic risk for alcohol consumption. Frontal volume is a promising prognostic biomarker for alcohol consumption liability.

Alcohol use and its associated negative consequences are ubiquitous international public health concerns. Worldwide, the average person aged 15 or older consumes 6.2 liters of alcohol annually, and alcohol use accounts for 6% of deaths and 5% of disease burden.[1] Consequently, it is critically important to advance efforts for prevention and identify individual differences that can serve as prognostic biomarkers of liability.

Neuroimaging studies have shown that alcohol consumption and use disorder are associated with smaller subcortical and cortical brain volumes, particularly among regions that feature prominently in emotion, memory, reward, cognitive control, and decision making [2,3]. While there is evidence that these associations may arise as a consequence of drinking (e.g., reduced neurogenesis in non-human primate models, greater gray matter decline among adolescents following initiation of heavy drinking, gray matter normalization following abstinence among dependent individuals)[4–6], emerging data suggest that these neural signatures may reflect preexisting vulnerabilities that precede and predict drinking initiation and escalating use [7,8].

Here, we first identify replicable gray matter volume correlates of alcohol use. We then test whether reduced volume is: (1) attributable to shared predisposing factors (e.g., shared genetic influence) and/or results from alcohol use, (2) prospectively predictive of future drinking in young adulthood, and (3) predictive of drinking initiation in adolescence. To this end, we used data from 3 independent neuroimaging samples with family or longitudinal data: the Duke Neurogenetics Study (DNS; N=1,303)[9]; Human Connectome Project (HCP; N=897)[10]; and Teen Alcohol Outcomes Study (TAOS; N=223)[11] (**Supplementary Information**). Finally, we examined whether genetic risk for alcohol consumption is associated with genes and genetically-conferred differences in gene expression that are preferentially expressed in the regions identified by neuroimaging analyses and/or the brain more generally. Here, we applied gene-set enrichment,

partitioned heritability, and transcriptome-wide (TWAS)[12] analyses to genome-wide association study (GWAS) summary statistics from the UK Biobank (N=112,117)[13] and AlcGen/CHARGE+ (N = 70,460)[14] studies of alcohol consumption, and RNA-seq data from GTEX (N=81-103)[15] and the Common Mind Consortium (N=452)[16].

Consistent with prior observations in unselected samples[2], as well as those with alcohol use disorder[3], whole brain discovery analyses in the DNS revealed that greater alcohol consumption is associated with lower gray matter volume across 8 clusters (**Figure 1**; **Supplementary Table 1**), encompassing regions identified in prior studies [2,3,7]. The associations with two of these clusters (right insula, right superior/middle frontal gyrus) replicated within an ROI analysis in the HCP (**Figure 1**; **Supplementary Table 1**).

Family-based analyses in the HCP (N=804) revealed that alcohol consumption and gray matter volume of the right insula and right middle/superior frontal gyrus are moderately to largely heritable (**Figure 2A; Supplementary Table 2**). Moreover, decomposition analyses showed that phenotypic correlations between frontal and insular gray matter volume and alcohol consumption are attributable to shared genetic, but not environmental, influences (**Figure 2B; Supplementary Table 2**). Analyses within twin and sibling pairs in the HCP concordant or discordant for the extent of alcohol use revealed that, relative to siblings concordant for low alcohol use, siblings concordant for high use or discordant for use (i.e., one high use, one low use) had lower insular and frontal gray matter volumes (**Figure 2C&D; Supplementary Table 3**). Further, brain volumes 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 gray matter volume of frontal gyri and insula are preexisting vulnerability factors associated with alcohol use, as opposed to a consequence of alcohol use.

Using available longitudinal data from the DNS (N=674), we found that lower gray matter volume 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) in the United States (**Figure 3A; Supplementary Table 4**). Similarly, in the TAOS longitudinal sample of children and adolescents, lower right middle and superior frontal gyrus volume predicted the initiation of alcohol use at an earlier age in those who were nondrinkers at baseline (**Figure 3B&C; Supplementary Table 5)**.

Gene-based association and partitioned heritability enrichment analyses of the UK Biobank GWAS of alcohol consumption revealed enrichment only among brain gene-sets. Moreover, Brodmann Area 9, which is in the frontal region in which we observed a negative association between volume and alcohol consumption that is attributable to shared genetic influence and predictive of drinking initiation, was among the regions with strongest enrichment (**Supplementary Figure 1, Supplemental Data**). A transcriptome-wide association analysis (TWAS) similarly found that genetic risk for alcohol consumption was significantly associated with differences in gene expression across the brain, including Brodmann Area 9 (**Supplementary Figure 2**), which replicated in an independent dataset (**Supplementary Figure 2, Supplemental Table 6**). Notably, genetic risk for alcohol consumption was not significantly associated with the expression of any gene in the liver (**Supplementary Figure 2**).

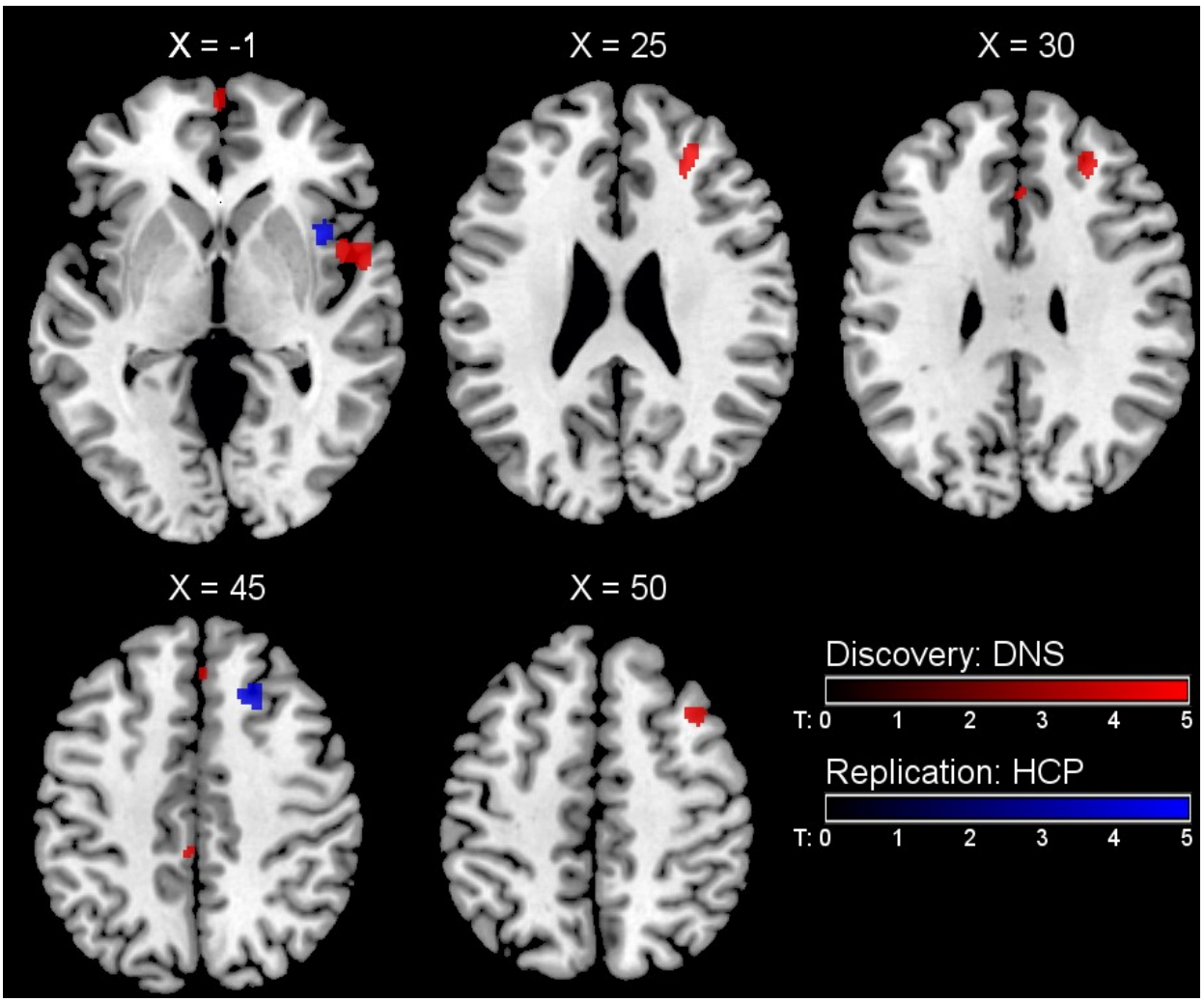
The above analyses in three independent samples provide unique convergent evidence that associations between middle/superior frontal gray matter volume and alcohol use are genetically-conferred, and predict future use and initiation. Taken alongside evidence that heavy alcohol consumption induces gray matter volume reductions[4–6], our data raise the intriguing possibility that genetically-conferred reductions in regional gray matter volumes may promote alcohol use

from adolescence to young adulthood, which may, in turn, lead to accelerated atrophy within these and other regions. Given evidence that genetic liability is shared across substance use involvement[17], our findings may generalize to other substances; this could be tested within genetically-informed and longitudinal studies enriched for other substance use, or large prospective studies, such as the recently launched Adolescent Brain Cognitive Development (ABCD) study[18]. While enrichment analyses implicate only brain pathways and TWAS 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[19]. Regardless, our convergent evidence from three independent samples with familial or longitudinal data, as well as evidence of expression enrichment, extends the literature primarily focused on alcohol-induced brain atrophy by demonstrating that lower gray matter volume in middle/superior frontal gyri and insula may represent a preexisting genetic liability for drinking that could serve as a prognostic biomarker.

References

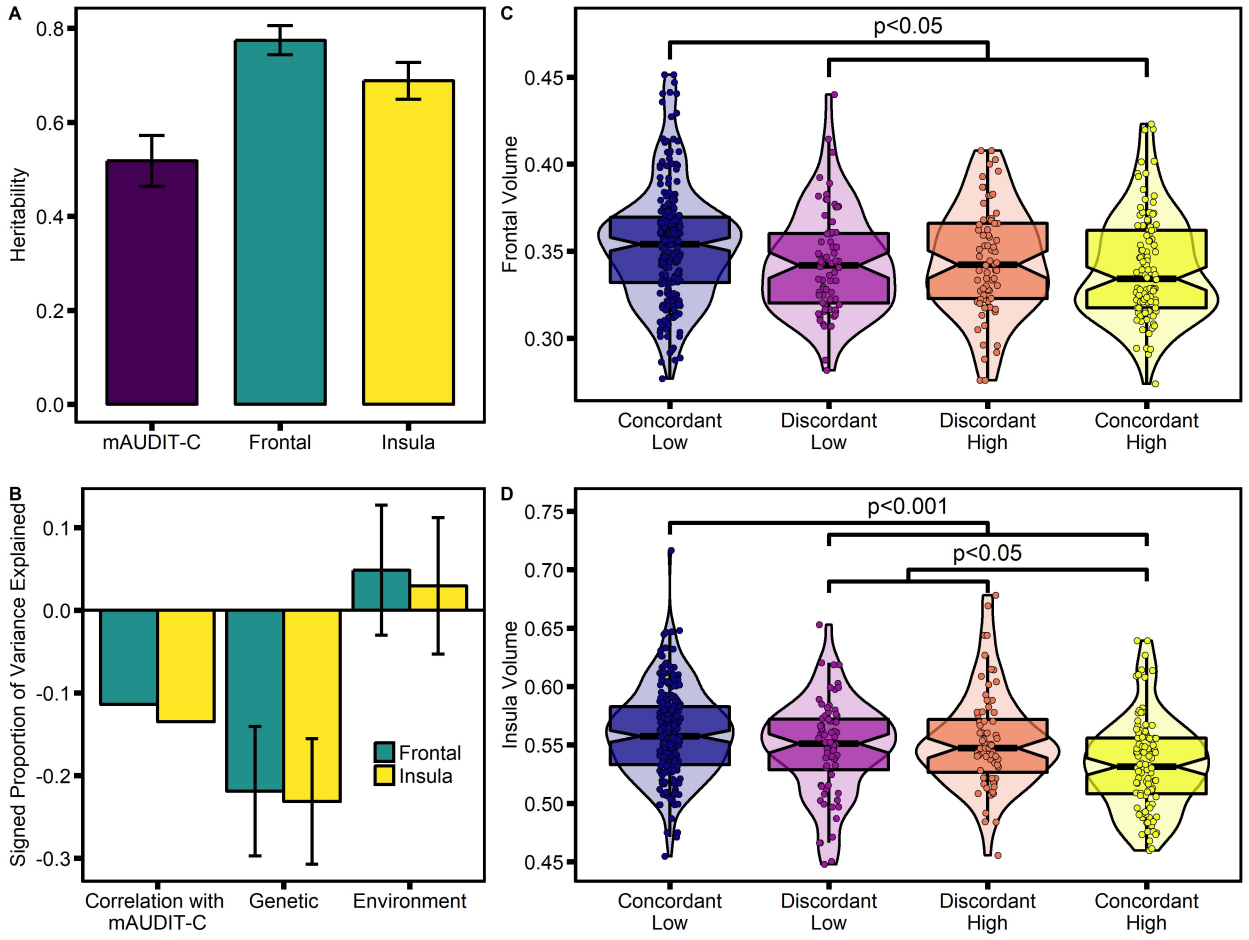
1. World Health Organization (Geneva, 2014).
2. Lange, E.H. et al. *Psychol. Med.* **47**, 655–668 (2017).
3. Yang, X. et al. *Neurosci. Biobehav. Rev.* **66**, 92–103 (2016).
4. Taffe, M.A. et al. *Proc. Natl. Acad. Sci.* **107**, 11104–11109 (2010).
5. Pfefferbaum, A. et al. *Am. J. Psychiatry* appi.ajp.2017.1 (2017).
6. Kühn, S. & Gallinat, J. *Biol. Psychiatry* **73**, 70–74 (2013).
7. Squeglia, L.M. & Gray, K.M. *Curr. Psychiatry Rep.* **18**, (2016).
8. Dager, A.D. et al. *Neuropsychopharmacology* **40**, 412–20 (2015).
9. Nikolova, Y.S., Knodt, a R., Radtke, S.R. & Hariri, a R. *Mol. Psychiatry* **21**, 1–9 (2015).
10. Van Essen, D.C. et al. *Neuroimage* **80**, 62–79 (2013).
11. Swartz, J.R., Williamson, D.E. & Hariri, A.R. *Am. J. Psychiatry* **172**, 276–283 (2015).
12. Gusev, A. et al. *Nat. Genet.* **48**, 245–252 (2016).
13. Clarke, T.-K. et al. *Mol. Psychiatry* **22**, 1376–1384 (2017).
14. Schumann, G. et al. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1611243113- (2016).
15. The GTEx Consortium et al. *Science (80-.).* **348**, 648–60 (2015).
16. Fromer, M. et al. *Nat. Neurosci.* **19**, 1442–1453 (2016).
17. Carey, C.E. et al. *Front. Genet.* **7**, 1–10 (2016).
18. Volkow, N.D. et al. *Dev. Cogn. Neurosci.* 1–4 (2017).doi:10.1016/j.dcn.2017.10.002
19. Dick, D.M. & Agrawal, A. *Alcohol Reserach Heal.* **31**, 111–118 (2008).

Figure 1: Identification of replicable volumetric associations with alcohol consumption.



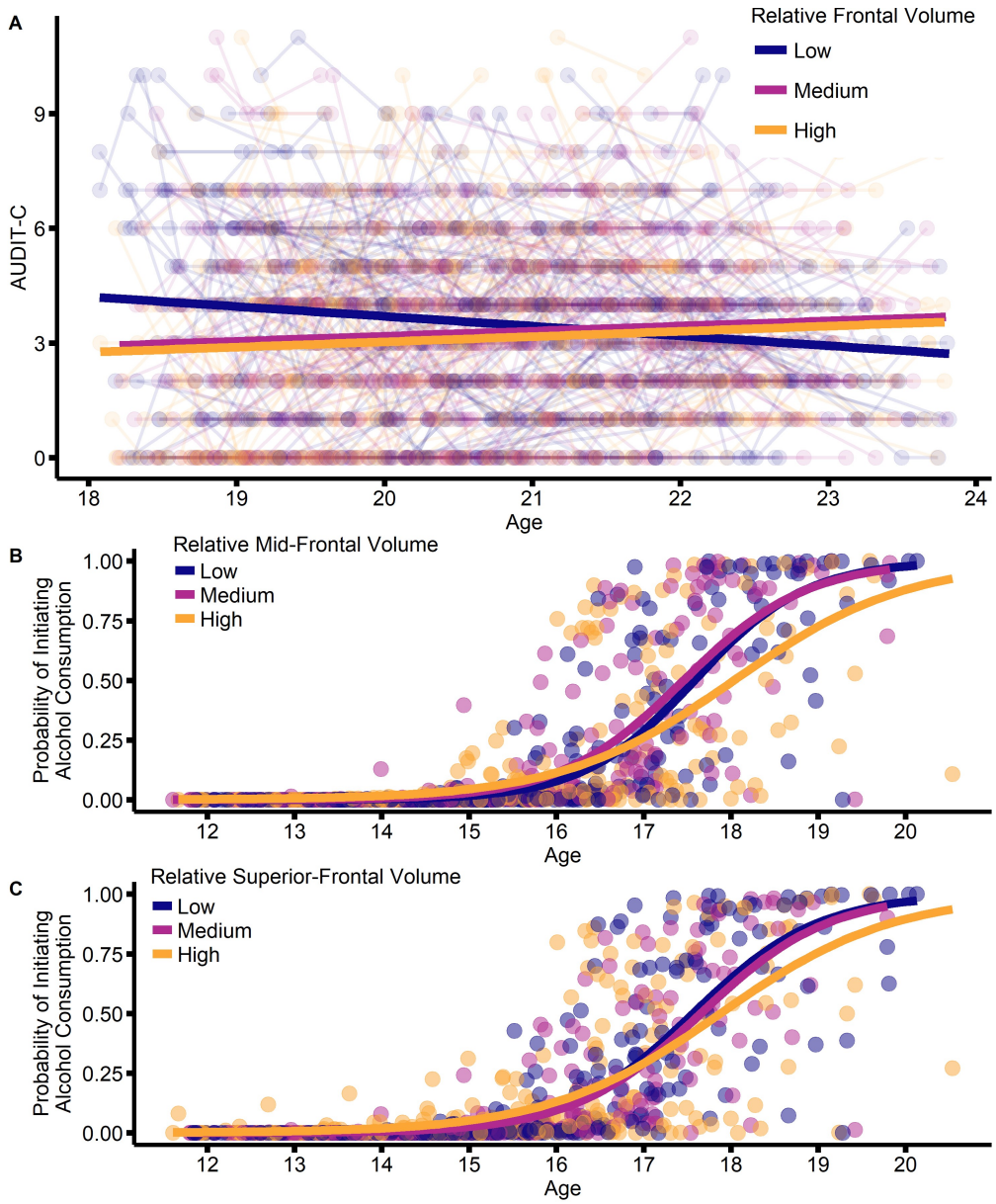
Statistical parametric map illustrating regions of reduced brain volume associated with increased alcohol consumption (**Supplemental Table 1**), overlaid onto a canonical structural brain image Montreal Neurological Institute coordinates and statistics (DNS: $p<0.05$, family-wise error whole-brain corrected, ≥10 contiguous voxels; HCP: $p<0.05$, family-wise error region-of-interest corrected, ≥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 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 gyrus cortices was extracted from both samples. Cluster overlap at an uncorrected threshold is shown in **Supplemental Figure 3**. DNS = Duke Neurogenetics Study. HCP = Human Connectome Project.

Figure 2: Shared genetic predisposition between alcohol consumption and brain volume.



HCP: A) Alcohol consumption scores (mAUDIT-C) and gray-matter volume of the right insula and right middle/superior frontal cortex were all observed to be heritable (mAUDIT-C: 51.79%, $p<2.2x10^{-16}$; insula: 68.83%, $p<2.2x10^{-16}$; frontal: 74.46%, $p<2.2x10^{-16}$; **Supplemental Table 2**). **B)** Significant phenotypic correlations between mAUDIT-C scores and volumes of the right insula and middle/superior frontal gyri are attributable to shared genetic (Insula: -0.2314, p=0.0022; Frontal: -0.2192, p=0.0054), but not environmental factors (**Supplemental Table 2**). **C&D)** Distribution of (**C**) right insula and (**D**) right middle/superior frontal volumes by alcohol exposure group. High = mAUDIT-C score > sample mean + 0.5 SD (i.e. > 4.67); Low = mAUDIT-C score < sample mean - 0.5 SD (i.e. < 1.54); Concordant = both siblings are in same group; Discordant = one sibling is High, while other is Low. Contrast comparisons found evidence for predispositonal effects of brain volume on alcohol consumption in both cases (Insula: Graded Liability: β=-0.0037, p=0.049, Predispositonal: β=0.0037, p=0.0006; Frontal: Predispositonal: β=0.0019, p=0.029; **Supplemental Table 3)**.

Figure 3: Frontal volume prospectively predicts alcohol use and initiation of consumption



A) DNS: Participants with reduced volume of the right middle/superior frontal cortex reported elevated alcohol consumption before the age of 20.85 years following the neuroimaging scan, and after accounting for baseline drinking (Frontal x Age interaction: β=0.150, p-fdr=0.008; **Supplemental Table 4**). **B&C)** TAOS: Participants with increased volume of the right middle and superior frontal cortex report initiation of alcohol consumption at an older age (Mid-Frontal x Age interaction: β=-57.042, p-fdr=0.036; Superior-Frontal x Age interaction: β=-60.74, p-fdr=0.036 **Supplemental Table 5).** Analyses were conducted with continuous data; partitioned into three equally-sized groups according to volume was done for display-purposes only.

Online Methods

Participants

Neuroimaging data were drawn from three independent samples: the Duke Neurogenetics Study (DNS; n=1,303), the Human Connectome Project (HCP; n=897), and the Teen Alcohol Outcome Study (TAOS; n=223).

Duke Neurogenetics Study (DNS): The DNS (cross-sectional; n=1334) assessed a wide range of behavioral, experiential, and biological phenotypes among young-adult (18-22 year-old) college students. Each participant provided informed written consent prior to participation in accord with the guidelines of the Duke University Medical Center Institutional Review Board and received $120 remuneration. All participants were in good general health and free of DNS exclusion criteria: (1) medical diagnosis of cancer, stroke, diabetes requiring insulin treatment, chronic kidney or liver disease or lifetime psychotic symptoms; (2) use of psychotropic, glucocorticoid or hypolipidemic medication, and (3) conditions affecting cerebral blood flow and metabolism (e.g., hypertension). DSM-IV Axis I and select Axis II disorders (Antisocial Personality Disorder and Borderline Personality Disorder) were assessed with the electronic Mini International Neuropsychiatric Interview (e-MINI)[1] and Structured Clinical Interview for the DSM-IV Axis II Personality Disorders[2]. These disorders are not exclusionary as the DNS seeks to establish broad variability in multiple behavioral phenotypes related to psychopathology. Participants were excluded from analyses due to: 1) non-completion of T1 structural scans (n=10), 2) scanner-related artifacts in MRI data (n=12), 3) incidental structural abnormalities (n=4), 4) missing or incomplete data (n=4), and genetic anomalies (e.g., Kleinfelter's syndrome; n=1). The final DNS sample consisted of 1,303 participants after quality assurance (**Supplementary Table 7**; age=19.70±1.25; 747 female; 258 with a DSM-IV Axis I disorder; **Supplementary Table 8**).

DNS participants were contacted every 3 months after initial study completion, and asked to complete a brief online assessment. Participants were entered into a lottery for a $50 gift-card following completion of each online assessment. Of the 734 participants who completed at least one online assessment (2,075 total responses), 705 completed the AUDIT questionnaire at least once, and 679 of these participants were among those included in initial DNS analyses (1,903 responses; **Supplemental Table 9**). Participants completed between 2 and 17 follow-ups (M=4.05, SD=2.62), between 28 and 1707 days after study completion (M=413.96, SD=331.22; age range: 18.33 - 23.82; **Supplementary Figure 4**).

Human Connectome Project (HCP): Data from participants contained in the HCP December 2015 public data release (N = 970), were considered for analyses. The HCP aims to recruit 1200 individuals (3-4 siblings per family, most including a twin pair) with the broad goal of examining individual differences in brain circuits and their relation to behavior and genetic background[3]. Each participant provided informed written consent prior to participation in accord with the guidelines of the Washington University in St Louis Institutional Review Board and received $400 remuneration, as well as additional winnings ($5) and travel expenses. All participants were aged 22 to 35 years and free of the following exclusionary criteria: preterm birth, neurodevelopmental, neuropsychiatric, or neurologic disorders; a full list of exclusions is available in prior publication[4]. Participants were excluded from analyses in the present study for poor quality structural MRI data (n=73), and non-completion of study questionnaires (n=3), resulting in a final sample of 894 (**Supplementary Table 7**; age=28.82±3.68; age range: 22-37; 393 males; 149 meeting criteria in a phone interview for a possible DSM-IV Axis I disorder[5] **Supplementary Table 8**).

Teen Alcohol Outcomes Study (TAOS): TAOS is a longitudinal study designed to examine the association between the development of depression and alcohol use disorders by recruiting adolescents aged 11 – 15 (N = 330) at high and low familial risk for depression (high: at least one first-degree and one second-degree relative with a lifetime history of major depression; low: no first-degree and minimal second-degree relatives (< 20%) with a lifetime history of depression). Participant diagnoses were assessed through structured clinical interviews with the adolescent and parent separately, using the Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version[6]. Participants were excluded if they met criteria for a substance use disorder or reported binge drinking at baseline (based on National Institute on Alcohol Abuse and Alcoholism guidelines). Participants were permitted to present with anxiety disorders in the high familial depression risk group (specific phobia, N=9; social phobia, N=6; panic disorder, N=1; generalized anxiety disorder, N=12). No other forms of psychopathology were present within the sample at baseline.

Participants were contacted every year to complete diagnostic interviews and questionnaires, and also underwent a follow-up MRI scanning session during the 2nd-4th year of participation (these scans were not used in the present analysis). Participants provided assent, and parents provided written informed consent following procedures approved by the Institutional Review Board at the University of Texas Health Sciences Center at San Antonio, and received $165 remuneration, as well as additional winnings ($10), travel expenses, and $40 for each annual follow-up. Participants were excluded from analyses in the present study for non-completion of the MRI study session (n=17), poor quality structural MRI data (n=13), non-completion of follow-up visits (n=24), missing baseline self-report measures (n=40), and initiation of alcohol use prior

to MRI scan (n=14), resulting in a final sample of n=223 (baseline age: 11-15; follow-up ages: 12-20; **Supplemental Table 7**).

Alcohol Use Assessment

DNS: Participants completed the 10-item Alcohol Use Disorders Identification Test (AUDIT), which was developed by the World Health Organization to screen for hazardous or dependent alcohol use patterns by assessing the frequency and nature of consumption over the past 12 months[7,8]. The AUDIT had reasonable internal consistency (α=0.81; M=5.22; SD=4.31; range 0-23). We computed the subscale score of the 3 items that correspond to the hazardous use or consumption domain of the AUDIT (AUDIT-C; Bush et al., 1998) (α=0.85; M=3.76; SD=2.64; range 0-12). Participants completed the AUDIT at baseline and during follow-up online assessments. In these follow-up assessments, the AUDIT questions were modified and instead asked about alcohol consumption following the participant's last assessment.

HCP: Participants completed the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)[5]. From the SSAGA we created a metric, the modified AUDIT-C (mAUDIT-C), to approximate the 12-month AUDIT-C (α=0.786; M=3.42; SD=2.65; range 0-12). This used questions almost identical to those contained in the AUDIT-C, but with the difference that the SSAGA asks about frequency of drinking 5+ drinks in a 24-hour period, while the AUDIT asks about 6+ drinks.

TAOS: Adolescents were assessed at baseline and each annual follow-up session with the Substance Use Questionnaire (SUQ)[10]. The SUQ assesses lifetime exposure to alcohol (e.g., have you ever had a drink, have you ever been drunk, age of first drink) and onset of regular use of a

substance (i.e., at least once per month for at least six months). The SUQ also includes items assessing the average quantity and frequency of use for alcohol, marijuana, tobacco, and six other drug classes (sedatives, stimulants, opioids, cocaine, hallucinogens, other (e.g., ecstasy)), periods of abstinence, and accessibility of each substance. The SUQ contains questions identical to those on the AUDIT-C. Thus, as in HCP analyses, a modified AUDIT-C score was created (mAUDIT-C: α=0.893; M=0.45; SD=1.26; range 0-9). Initiation of alcohol consumption was defined as attaining a score of 1 or greater on the mAUDIT-C (i.e. participant reports having consumed at least one full alcoholic beverage in the past year; N=82 initiated during the study; Age: M = 16.68, SD = 1.39, 14.12 – 19.64 yrs).

Self-report Questionnaires and Behavioral Phenotypes Covariates

Demographic and environmental variables known to be correlated with alcohol consumption were included in all analyses. In addition to age[11–13] , sex[12–15], and ethnicity[15,16], socioeconomic status (SES; income and education, or parental education) was included, as higher SES participants have been found to consume more alcohol, while lower SES is associated with higher rates of alcohol-related problems[11–15]. Self-report of early-life and recent life-stress was also included, given the well-replicated association of stress with alcohol use and alcohol-related disorders[17–19].

Socioeconomic Status – Income and Education

DNS: Participants completed three Likert-scale questions, (scores range from 1-11) where higher scores reflect having a higher socioeconomic status (more money, education, and respected jobs). Participants were asked to place themselves, their biological father during their childhood and adolescence, and their biological mother during their childhood and adolescence, on this scale.

These three responses were averaged to compute a proxy for the participant's socioeconomic status (P-SES: α=0.816; M=7.44; SD=1.70; range 1.33-11). Since all participants in the DNS were enrolled in college at the time of the study, parental education (the average of the education of the participant's female and male guardian) was used as a proxy, which ranged from 1 (some high school), to 9 (doctoral degree; MD, PhD, JD, or PharmD) (P-ED: α=0.70; M=7.09; SD=1.62).

HCP: Participants reported their total annual household income on a scale ranging from 1 (<$10,000), to 8 (>=$100,000) (SES: M=4.97; SD=2.18) as well as the number of years of education that they had completed (ED: M=14.89; SD=1.82; range=11-17).

TAOS: The parents of participants reported their annual income and their spouses annual income on a scale ranging from 0 (Less than $10,000) to 7 ($150,000 or more) (Self: M=4.27; SD=1.39; range=0-6. Spouse: M=4.12; SD=1.71; range=0-6.). Parents also reported their highest level of education and their spouses highest level of education on a 0 (Less than 9th grade) to 6 (Graduate or Professional Degree) scale (Self: M=2.67; SD=12.36; range=0-9. Spouse: M=4.03; SD=2.06; range=0-9.). An overall estimate of SES was calculated by averaging the standardized values of these variables.

Early-life stress

DNS: Participants completed the 28-item Childhood Trauma Questionnaire (CTQ)[20], which asks participants to retrospectively report on the occurrence and frequency of emotional, physical, and sexual abuse as well as emotional and physical neglect before the age of 17 (α=0.88; M=33.55; SD=8.76; range 25-76). The instrument's five subscales, each representing one type of abuse or

neglect, have robust internal consistency and convergent validity with a clinician-rated interviews of childhood abuse [21]. Total CTQ scores across the 5 subscales were used as a covariate.

HCP: The HCP did not include a measure of early-life stress.

TAOS: Participants completed the 28-item CTQ; total scores across the 5 subscales were used as a covariate in analyses (α=0.795; M=32.98; SD=8.76; range 25-64).

Perceived Stress

DNS: Participants completed the 10-item version of the Perceived Stress Scale (PSS)[22], which instructs participants to appraise how unpredictable, uncontrollable, and stressful their daily life was in the preceding week. The PSS had good internal consistency (α=0.86; M=14.66; SD=6.08; range 0-37).

HCP: HCP participants completed the same 10-item version of the PSS, which had good internal consistency (α=0.90; M=13.07; SD=5.76; range 0-35).

TAOS: TAOS participants completed the Stressful Life Events Schedule (SLES)[23], which assesses the presence of more than 80 possible stressors in the past 12 months, each rated on a 4-point scale (1-4). Each stressor is given a subjective stress score as rated by the adolescent and an objective stress rating by a consensus panel. For both subjective and objective stress, a summary score derived by summing the squares of each individual stressor. Herein we use the total subjective stress (M=31.52; SD=28.79; range 0-156).

Additional Measures in TAOS

Depressive Symptoms: As the TAOS sample was enriched for participants with a family history of depression, self-report of depressive symptoms was included as a covariate in all analyses. Participants completed the Mood and Feelings Questionnaire (MFQ)[24], a 33-item measure designed to detect clinically meaningful signs and symptoms of depressive disorders in children and adolescents (α=0.68; M=9.03; SD=8.01; range 0-52).

Tanner Stage: The tanner scale was used to assess pubertal status (Female: M=3.6, SD = 0.9; Male: M=2.97, SD=0.8)[25,26].

Magnetic Resonance Imaging: Acquisition and Processing of Gray Matter Volume Data

DNS: Two identical research-dedicated GE MR750 3 T scanners equipped with high-power high-duty-cycle 50-mT/m gradients at 200 T/m/s slew rate, and an eight-channel head coil for parallel imaging at high band-width up to 1 MHz were used to acquire data at the Duke-UNC Brain Imaging and Analysis Center (N=224, 17% of the sample, was scanned on the second scanner. Scanner is included as a covariate in all analyses). High-resolution T1-weighted images were obtained using a 3D Ax FSPGR BRAVO with the following parameters: TR = 8.148 s; TE = 3.22 ms; 162 sagittal slices; flip angle, 12°; FOV, 240 mm; matrix =256×256; slice thickness = 1 mm with no gap; and total scan time = 4 min and 13 s. Regional gray matter volumes were determined using the unified segmentation[27] and DARTEL normalization[28] modules in SPM12 (http://www.fil.ion.ucl.ac.uk/spm)[29]. Using this approach, individual T1-weighted images were segmented into gray, white, and CSF images then non-linearly registered to the existing IXI template of 550 healthy subjects averaged in standard Montreal Neurological Institute space,

available with VBM8 (http://dbm.neuro.uni-jena.de/vbm/). Subsequently, gray matter images were modulated for nonlinear effects of the high-dimensional normalization to preserve the total amount of signal from each region, and smoothed with an 8mm FWHM Gaussian kernel. The voxel size of processed images was 1.5×1.5×1.5 mm. A gray matter mask for subsequent analyses was created by thresholding the final stage (6th) IXI template at 0.1.

HCP: High-resolution (0.7-mm isotropic voxels) anatomical images were acquired using a customized Siemens Skyra 3-T scanner with a 32-channel head coil[30]. Briefly, relevant steps for this study from the HCP processing pipeline within FSL v5.0.6 [31] included: (1) Gradient distortion correction, (2) Coregistration and averaging of T1 and T2 runs, (3) Linear registration of T1 and T2 runs, (4) FSL FNIRT brain extraction, (5) Field mad distortion correction, and (6) Bias field correction. Additional processing for VBM analyses were then applied. Brain-extracted images were grey matter-segmented before being registered to the MNI 152 standard space using non-linear registration [32]. The resulting images were averaged and flipped along the x-axis to create a left-right symmetric, study-specific grey matter template. Native grey matter images were then non-linearly registered to this study-specific template and multiplied by the Jacobian of the warp field to correct for local expansion (or contraction) due to the non-linear component of the spatial transformation. These images were then smoothed with an isotropic Gaussian kernel with a sigma of 4 mm.

TAOS: Imaging data was collected using a Siemens 3T Trio scanner located at the Research Imaging Institute at the University of Texas Health Science Center, San Antonio (UTHSCSA). The study used an MRI protocol specifically optimized for GM thickness measurement [33]. The protocol was designed to collect data to resolve the cortical ribbon across to cortex using isotropic spatial resolution of 0.8mm, voxel size =0.5mm. T1-weighted contrast was achieved using a

magnetization prepared sequence with an adiabatic inversion contrast-forming pulse (scan parameters: TE/TR/TI=3.04/2100/785 ms, flip angle=11 degrees).

The processing of T1-weighted images consisted of removing non-brain tissues, global spatial normalization and radio frequency (RF) inhomogeneity correction. Non-brain tissues such as skin, muscle and fat was removed using an automated skull stripping procedure and images were corrected for radio-frequency (RF) inhomogeneity[34]. A retrospective motion-correction technique was used to reduce subject motion-related artifacts[35]. Next, images were imported into the structural analysis package, BrainVisa, and processed using its cortical extraction and parcellation pipelines[35]. This pipeline extracts the pial and grey matter and white matter (WM) interface surfaces, performs extraction, labeling and verification of sulcal surfaces[36] and segments the cortical landscape into 15 cortical regions using the primary sulcal structures.

Statistical analysis

Sample demographics and comparisons, as well as associations of self-report measures with alcohol consumption, were computed in R (3.3.2)[37]. Self-report questionnaire data were winsorized (to ± 3 SDs; DNS: AUDIT-C N=1, P-SES N=5, PSS N=6, CTQ N=25; HCP: mAUDIT-C N=3, PSS N=3; TAOS: SLES N=12, CTQ N=8; MFQC = 2) to maintain variability while limiting the influence of extreme outliers. Self-report variables with high skew (>1 or <-1) were transformed prior to analyses. Left-skewed variables (DNS: CTQ skew=1.34; TAOS: CTQ skew = 1.16, SLES skew = 1.23) were log-transformed, while right-skewed variables (DNS: P-ED skew=-1.43) were squared.

Discovery: A whole-brain voxel-based morphometry GLM regression analysis was conducted using SPM8 in the DNS sample to test whether alcohol consumption (AUDIT-C) is associated with differences in gray-matter volume. Covariates included sex, age, self-reported race/ethnicity (i.e., not-white/white, not-black/black, not Hispanic/Hispanic), scanner id, intracranial volume (ICV), presence of a diagnosis other than alcohol or substance abuse or dependence, perceived stress (PSS), parental education (P-ED), early-life stress (CTQ), and perceived economic status (P-SES). Analyses were thresholded at $p<0.05$ FWE with a cluster extent threshold of 10 contiguous voxels ($k_e=10$) across the entire search volume.

Replication: Replication analyses in the HCP sample examined whether alcohol consumption (mAUDIT-C) predicted differences in gray-matter volume only within regions of interest (ROIs) where associations were observed in the discovery DNS sample (**Figure 1, Supplemental Table 1**). ROIs were defined by the AAL atlas[38]. A voxelwise GLM regression limited to these ROIs was conducted using multi-level block permutation-based non-parametric testing (FSL PALM v.alpha103; tail approximation $p<0.10$ with 5,000 permutations), which accounts for the family-structure of the HCP data while correcting for multiple comparisons across space[39–41]. Covariates included sex, age, self-reported race/ethnicity (i.e., not-white/white, not-black/black, not Hispanic/Hispanic), intracranial volume (ICV), twin/sibling status (dizygotic/not, monozygotic/not, half-sibling/not), presence of a diagnosis other than alcohol or substance abuse or dependence, perceived stress (PSS), education (ED), and economic status (SES). Analyses were thresholded at $p<0.05$ FWE with a cluster extent threshold of 10 contiguous voxels ($k_e=10$) across the entire search volume (i.e., across all ROIs collectively).

Post-hoc analyses

The total anatomical ROI volume of regions which replicated in the HCP (right Insula and right Middle/Superior Frontal Gyrus; see **Results**) were extracted from both datasets for use in *post-hoc* analyses. The total volume was used to reduce overestimation of effect sizes that can arise from selecting only those voxels that are specifically associated with the variable of interest[42].

Heritability: Heritability analyses were conducted using a subset of participants from the HCP, which excluded singletons and half-siblings (n=804). This resulted in a sample of 293 families, including 115 MZ twin-pairs, 64 DZ twin-pairs, and 422 non-twin siblings. The SOLAR-Eclipse software package (http://solar-eclipse-genetics.org)[43], in conjunction with the R package 'Solarius'[44], was used for all heritability analyses. SOLAR (Sequential oligogenic linkage analysis routines) implements maximum likelihood variance decomposition methods to estimate phenotypic heritability (h^2; the fraction of phenotypic variance attributable to additive genetic factors), as well as genetic (ρ_g) and environmental (ρ_e) correlations (the fraction of the correlation between two phenotypes that is attributable to either additive genetic or individual-specific environmental factors, respectively). SOLAR was used to estimate the heritability of gray-matter volume and alcohol consumption, as well as the co-heritability of volume and alcohol consumption. Covariates were identical to 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.

Discordant twin analysis: Following evidence that alcohol consumption is co-heritable with volume of the right insula and middle/superior frontal gyrus (see **Main Text**), we examined whether same-sex twin and non-twin sibling pairs discordant for alcohol consumption differed from each other on brain volume. These analyses examined whether mAUDIT-C was associated with insular or middle/superior frontal volume after accounting for sibling-shared genetic

background and experience. Same-sex siblings were considered "high alcohol consumers" or "low alcohol consumers" if their mAUDIT-C score was greater than 0.5 SD above the sample mean (mAUDIT-C > 4.67, or less than 0.5 SD below the sample mean (mAUDIT-C < 1.54), respectively. Of the original 476 sibling pairs, 214 pairs were removed because they did not meet this criteria. Concordant sibling-pairs were 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 (19 pairs failed to meet this last criteria). Discordant sibling-pairs were defined as a pair where siblings were in different categories (i.e. one is high and the other is low). This resulted in 117 concordant low alcohol consumer pairs ("low concordant"; mAUDIT-C M=0.84, SD=0.77), 54 concordant high alcohol consumer pairs ("high concordant"; mAUDIT-C M=7.13, SD=1.41), and 72 discordant sibling pairs ("low discordant"; mAUDIT-C M=1.25, SD=0.73; "high discordant"; mAUDIT-C M=6.47, SD=1.67). Participants could be included in more than one pair (N=368 individuals) when considering relationships with multiple siblings. Discordancy analyses were conducted using linear mixed models, as sibling pairs are non-independent, using the 'Psych'[45] and 'lme4'[46] packages in R to account for the multiple-sibling structure within families. Covariates were identical to those used in neuroimaging analyses.

Three contrasts were entered into mixed-effect models, which modeled 3 different possible associations between brain volume, alcohol consumption, and familial/predispositonal risk[47]. The first tested whether alcohol consumption may cause reduced brain volume, which would be evidenced by a difference in brain volume between the exposed and unexposed members of discordant pairs. Both the second and third contrasts tested the hypothesis that the association between reduced brain volume and alcohol consumption is driven by a shared predisposition towards both. This would be primarily evidenced by the discordant pairs – biological siblings who

differ in their alcohol consumption – having the same volume, which would be reduced relative to concordant unexposed pairs. The second contrast tested that brain volume decreases as a function of increasing familial/predispositonal liability (i.e. graded liability), which would be additionally evidenced by reduced volume in concordant exposed pairs relative to discordant pairs. The third contrast tested whether any amount of shared familial/predispositonal risk would be reflected by the same reduction in volume, which would be additionally evidenced by no difference in volume between discordant pairs and concordant exposed siblings.

DNS longitudinal changes in alcohol consumption: Hierarchical density-based clustering (R 'dbscan' package)[48], was used to detect and remove outlier responses to the follow-up questionnaire. A scree-plot comparison of the minimum-points parameter (minPts; minimum number of points in a cluster) and the number of outliers found an elbow at minPts=35 when time-of-response was represented as time-since-baseline, and subsequently an elbow at minPts=14 when time-of-response was represented as participant age. This resulted in the removal of 112 responses that occurred more than 1,035 days after the baseline study visit, and the removal of 17 responses that were given by participants older than 23.82 years old. The final dataset consisted of 1,756 responses from 674 participants, who gave 1-12 (M: 3.59, SD: 2.10) responses, 28-1,034 (M: 350.20, SD: 245.39) days after the baseline visit, between the ages of 18.33 and 23.82 (M: 20.93, SD: 1.24) years (**Supplemental Figure 4**).

The R 'nlme' package[49] was used to fit a longitudinal multilevel linear model, examining whether brain-volume in the DNS sample predicted follow-up AUDIT-C questionnaire responses over time. The 'nlme' package was used as it can model different classes of correlation structures between observations, though it does not include logistic models. The model included both random intercept and random slope components, with a continuous autoregressive correlation structure.

Time was coded as both the 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 participant's age?). Covariates were Z-scored, and were identical to neuroimaging analyses, with the addition of second-order interactions between covariates and primary variables[50,51]. Each of the two ROIs were tested in separate models, and p-values were FDR corrected (4 tests – middle/superior x linear-age, middle/superior x quadratic-age, etc.).

TAOS longitudinal initiation of alcohol use: The R 'lme4 package[52] was used to fit a longitudinal logistic multilevel model, which tested whether baseline brain volume in non-drinking 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 brain volume and age (i.e. does baseline ROI volume predict a different likelihood of initiation as participant's age?). Covariates were Z-scored, and included demographic variables (age, sex, ethnicity, and SES), stress (CTQ and SLES), tanner-stage, MFQ-scores, family history of depression, age at MRI scan, and intracranial volume. Second-order interactions between covariates and primary variables (e.g., Middle Frontal volume x Sex, Middle Superior volume x SES, Age x Sex, Age x SES, etc.) were also included[50]. Each of two ROIs were tested in separate models - right superior frontal cortex and right middle frontal cortex. Insula volume was excluded as it was not significant in the DNS analyses, but given the new sample we considered both right middle and right superior frontal ROIs separately. P-values were subsequently FDR corrected (4 tests).

SNP-Based Enrichment: We tested whether the SNP-based heritability of alcohol consumption is enriched in brain-expressed gene-sets and whether this enrichment is specific to any region. Stratified LD-score regression[53,54] was applied to summary statistics from the genome-wide association study of alcohol consumption in the UK Biobank (N=112,117)[55]. Tissue-enriched gene-sets, provided by the Alkes Group[56], were generated using data from the GTEx Consortium[57]. In this analysis, a gene is assigned to a gene-set if it shows greater enrichment in that tissue than 90% of genes. Gene-sets for brain regions were generated both by comparing each region to all non-brain tissues, and by comparing each brain region to all other regions. It was further tested whether the genetic associations with alcohol consumption are enriched in brain-expressed gene-sets. Gene-set analyses were conducted with MAGMA[58], implemented through FUMA[59].

Transcriptome-Wide Analysis (TWAS): Following evidence that genetic associations with alcohol consumption are enriched in brain-expressed gene-sets, and that the SNP-based heritability of alcohol consumption is enriched in these gene-sets, we tested whether genetic risk for alcohol consumption is predictive of changes in post-mortem gene expression in the human brain. Pre-computed gene-expression RNA-seq weights for nine brain regions and the liver from GTEX[57] were downloaded and analyzed using the FUSION suite[60]. We tested whether genetic risk for alcohol consumption, as determined by the results from the UK Biobank GWAS[55], is associated with differential RNA expression. Results were bonferroni-corrected for n=9,839 tests across the ten regions. Replication of TWAS results was sought using independent GWAS data from an earlier study of alcohol consumption (N = 70,460)[61] and computed gene-expression weights for the dorsolateral prefrontal cortex from the CommonMind Consortium[62]. As the gene that showed the strongest association in the discovery dataset was not present in the replication data, it was

examined whether any of the gene-expression associations at p-fdr<0.05 were significant in the replication data.

Data Availability

All data pertaining to this study are available upon request to the corresponding author. Data sets can also be accessed at the locations listed:

Duke Neurogenetics Study (DNS): https://www.haririlab.com/projects/procedures.html

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

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

ALCGEN/CHARGE+: Requests for access to summary statistics should be submitted to the study PI, Dr. Gunter Schumann - gunter.schumann@kcl.ac.uk

GTEx/CMC: Pre-computed gene expression weights were provided by the Gusev lab - http://gusevlab.org/projects/fusion/

References

1. Sheehan, D. V. et al. *J. Clin. Psychiatry* **59**, 22–33 (1998).
2. First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B. & Benjamin, L. *Am. Psychiatr. Press. Inc.* (1997).
3. Smith, S.M. et al. *Nat. Neurosci.* **18**, 1–7 (2015).
4. Van Essen, D.C. et al. *Neuroimage* **62**, 2222–2231 (2012).
5. Bucholz, K.K. et al. *J. Stud. Alcohol* **55**, 149–158 (1994).
6. Kaufman, J. et al. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 980–988 (1997).
7. Saunders, J.B., Aasland, O.G., Babor, T.F., de la Fuente, J.R. & Grant, M. *Addiction* **88**, 791–804 (1993).
8. Babor, T.F., Higgins-Biddle, J.C., Saunders, J.B. & Monteiro, M.G. *World Heal. Organ.* (2001).doi:10.1177/0269881110393051
9. Bush, K. et al. *Arch. Interal Med.* **158**, 1789–1795 (1998).
10. Molina, B.S.G. et al. *J. Am. Acad. Child Adolesc. Psychiatry* **46**, 1028–1040 (2007).
11. Kendler, K.S. et al. *J. Stud. Alcohol Drugs* **75**, 541–5 (2014).
12. Meng, Y., Holmes, J., Hill-Mcmanus, D., Brennan, A. & Meier, P.S. *Addiction* **109**, 206–215 (2014).
13. Collins, S.E. *Alcohol Res.* **38**, 83–94 (2016).
14. Grittner, U., Kuntsche, S., Gmel, G. & Bloomfield, K. *Eur. J. Public Health* **23**, 332–339 (2013).
15. Delker, E., Brown, Q. & Hasin, D.S. *Alcohol Res.* **38**, 7–15 (2016).
16. Cacciola, E.E.T. & Nevid, J.S. *Psychol. Addict. Behav.* **28**, 1278–1283 (2014).
17. Keyes, K.M., Hatzenbuehler, M.L., Grant, B.F. & Hasin, D.S. *Alcohol Res.* **34**, 391–400 (2012).
18. Enoch, M.-A. *Psychopharmacology (Berl).* **214**, 17–31 (2011).
19. Green, J.G. et al.**67**, 113–123 (2013).
20. Bernstein, D.P. et al. *Child Abus. Negl.* **27**, 169–190 (2003).
21. Scher, C.D., Stein, M.B., Asmundson, G.J.G., Mccreary, D.R. & Forde, D.R. *J. Trauma. Stress* **14**, 843–857 (2001).
22. Cohen, S., Kamarck, T. & Mermelstein, R. *J. Health Soc. Behav.* **24**, 385–396 (1983).
23. Williamson, D.E. et al. *Psychiatry Res.* **119**, 225–241 (2003).
24. Angold, A. et al. *Int. J. Methods Psychiatr. Res.* 237–249 (1995).doi:1049-8931/95/040251-12
25. Marshall, W.A. & Tanner, J.M. *Arch. Dis. Child.* 291–303 (1969).doi:10.1136/adc.44.235.291
26. Marshall, W.A. & Tanner, J.M. *Arch. Dis. Child.* **45**, 13–23 (1970).
27. Ashburner, J. & Friston, K.J. *Neuroimage* **26**, 839–851 (2005).

28. Ashburner, J. *Neuroimage* **38**, 95–113 (2007).

29. Kurth, F., Gaser, C. & Luders, E. *Nat. Protoc.* **10**, 293–304 (2015).

30. Glasser, M.F. et al. *Neuroimage* **80**, 105–124 (2013).

31. Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W. & Smith, S.M. *Neuroimage* **62**, 782–790 (2012).

32. Andersson, J.L.R., Jenkinson, M. & Smith, S. *In Pract.* 22 (2007).at <http://fmrib.medsci.ox.ac.uk/analysis/techrep/tr07ja2/tr07ja2.pdf>

33. Kochunov, P. & Davis, M.D. *Methods* **50**, 136–146 (2010).

34. Smith, S.M. et al. *Neuroimage* **31**, 1487–1505 (2006).

35. Kochunov, P. et al. *Hum. Brain Mapp.* **27**, 957–962 (2006).

36. Mangin, J.-F. et al. *IEEE Trans. Med. Imaging* **23**, 968–982 (2004).

37. R Core Team (Vienna, Austria, 2014).at <http://www.r-project.org/>

38. Tzourio-Mazoyer, N. et al. *Neuroimage* **15**, 273–289 (2002).

39. Winkler, A.M., Webster, M.A., Vidaurre, D., Nichols, T.E. & Smith, S.M. *Neuroimage* **123**, 253–268 (2015).

40. Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M. & Nichols, T.E. *Neuroimage* **92**, 381–397 (2014).

41. Winkler, A.M., Ridgway, G.R., Douaud, G., Nichols, T.E. & Smith, S.M. *Neuroimage* **141**, 502–516 (2016).

42. Vul, E., Harris, C., Winkielman, P. & Pashler, H. *Perspect. Psychol. Sci.* **4**, 274–290 (2009).

43. Kochunov, P. et al. *Neuroimage* **111**, 300–301 (2015).

44. Ziyatdinov, A. et al. *Bioinformatics* **32**, 1901–1902 (2016).

45. Revelle, W. *R Packag.* 1–358 (2015).

46. Bates, D., Maechler, M., Bolker, B. & Walker, S. *J. Stat. Softw.* **67**, 1–91 (2015).

47. Pagliaccio, D. et al. *JAMA Psychiatry* **63130**, 1–8 (2015).

48. Hahsler, M., Piekenbrock, M., Arya, S. & Mount, D. (2017).at <https://cran.r-project.org/package=dbscan>

49. Pinheiro, J., DebRoy, S., Bates, D., Sarkar, D. & R Core Team (2017).at <https://cran.r-project.org/package=nlme%7D>

50. Baranger, D.A.A. et al. *Front. Psychol.* **7**, 1–10 (2016).

51. Keller, M.C. *Biol. Psychiatry* **75**, 18–24 (2014).

52. Bates, D., Mächler, M., Bolker, B.M. & Walker, S.C. *J. Stat. Softw.* **67**, 1–48 (2015).

53. Finucane, H.K. et al. *Nat. Genet.* **47**, 1228–1235 (2015).

54. Finucane, H. et al. *Nat. Genet.* **50**, 621–629 (2018).

55. Clarke, T.-K. et al. *Mol. Psychiatry* **22**, 1376–1384 (2017).

56. Bulik-Sullivan, B. & Finucane, H.K. (2017).at <https://github.com/bulik/ldsc>

57. The GTEx Consortium et al. *Science (80-.).* **348**, 648–60 (2015).

58. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. *PLoS Comput. Biol.* **11**, 1–19 (2015).

59. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. *Nat. Commun.* **8**, 1–10 (2017).

60. Gusev, A. et al. *Nat. Genet.* **48**, 245–252 (2016).

61. Schumann, G. et al. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1611243113- (2016).

62. Fromer, M. et al. *Nat. Neurosci.* **19**, 1442–1453 (2016).

SUPPLEMENTARY INFORMATION

Convergent evidence for predispositonal effects of brain volume on alcohol consumption

David AA Baranger[1], Catherine H. Demers[1], Nourhan M. Elsayed[2] , Annchen R. Knodt[3], Spenser R. Radtke[3], Aline Desmarais[1], Lauren R. Few[4], Arpana Agrawal[4], Andrew C. Heath[4], Deanna M. Barch[1,4], Lindsay M. Squeglia[5], Douglas E. Williamson[2,6], Ahmad R. Hariri[3], Ryan Bogdan[1]

[1] Department of Psychological and Brain Sciences, Washington University, St. Louis, MO
[2] Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC
[3] Department of Psychology and Neuroscience, Duke University, Durham, NC
[4] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO
[5] Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC
[6] Durham VA Medical Center, Durham, NC

Supplementary Information Table of Contents

Supplemental Figure 1-4: Pages 3-7

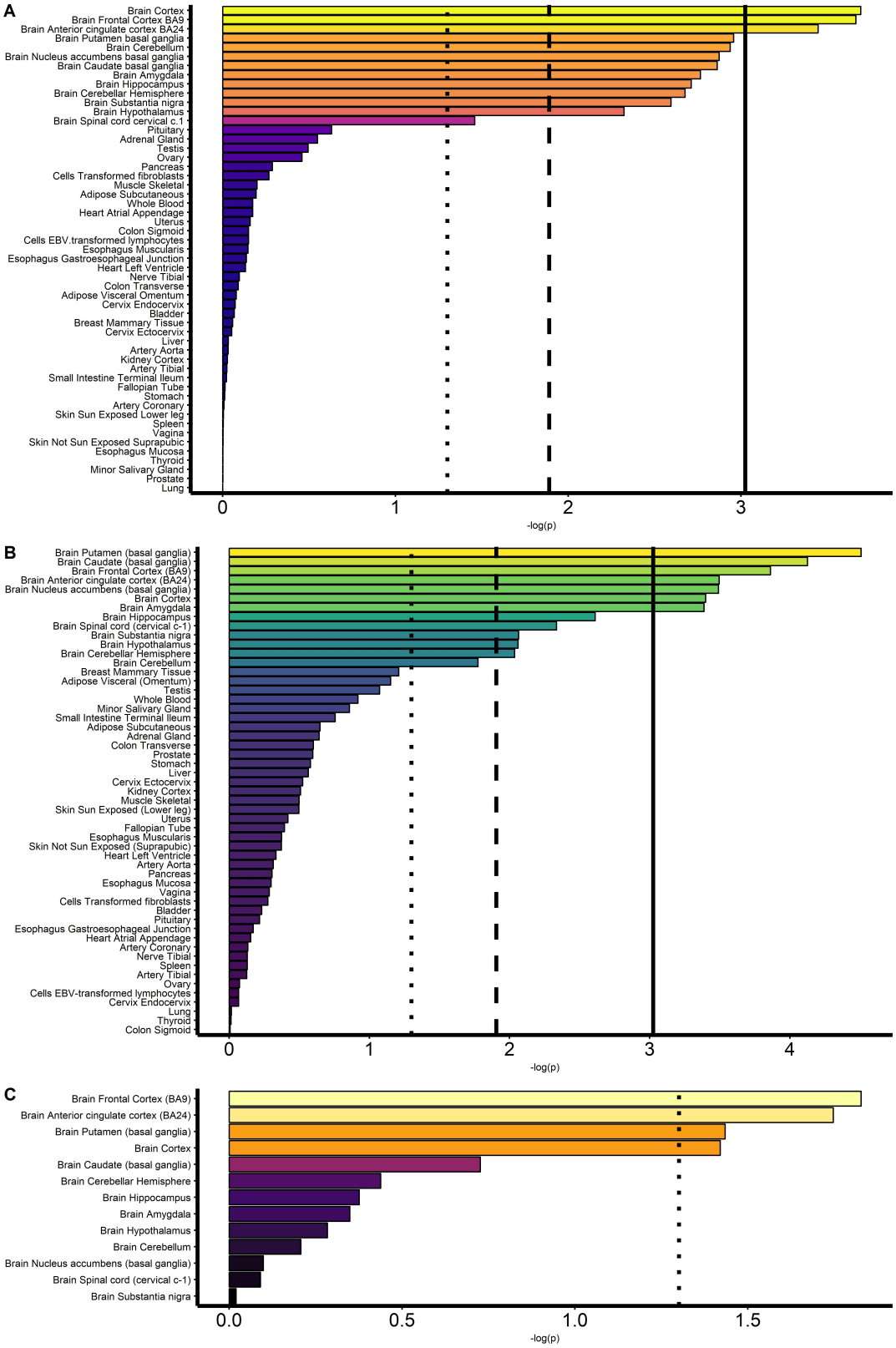
Supplemental Tables 1-11: Pages 8-21

Supplemental Results: Pages 22 – 26

Supplemental Analyses: Pages 27 - 30

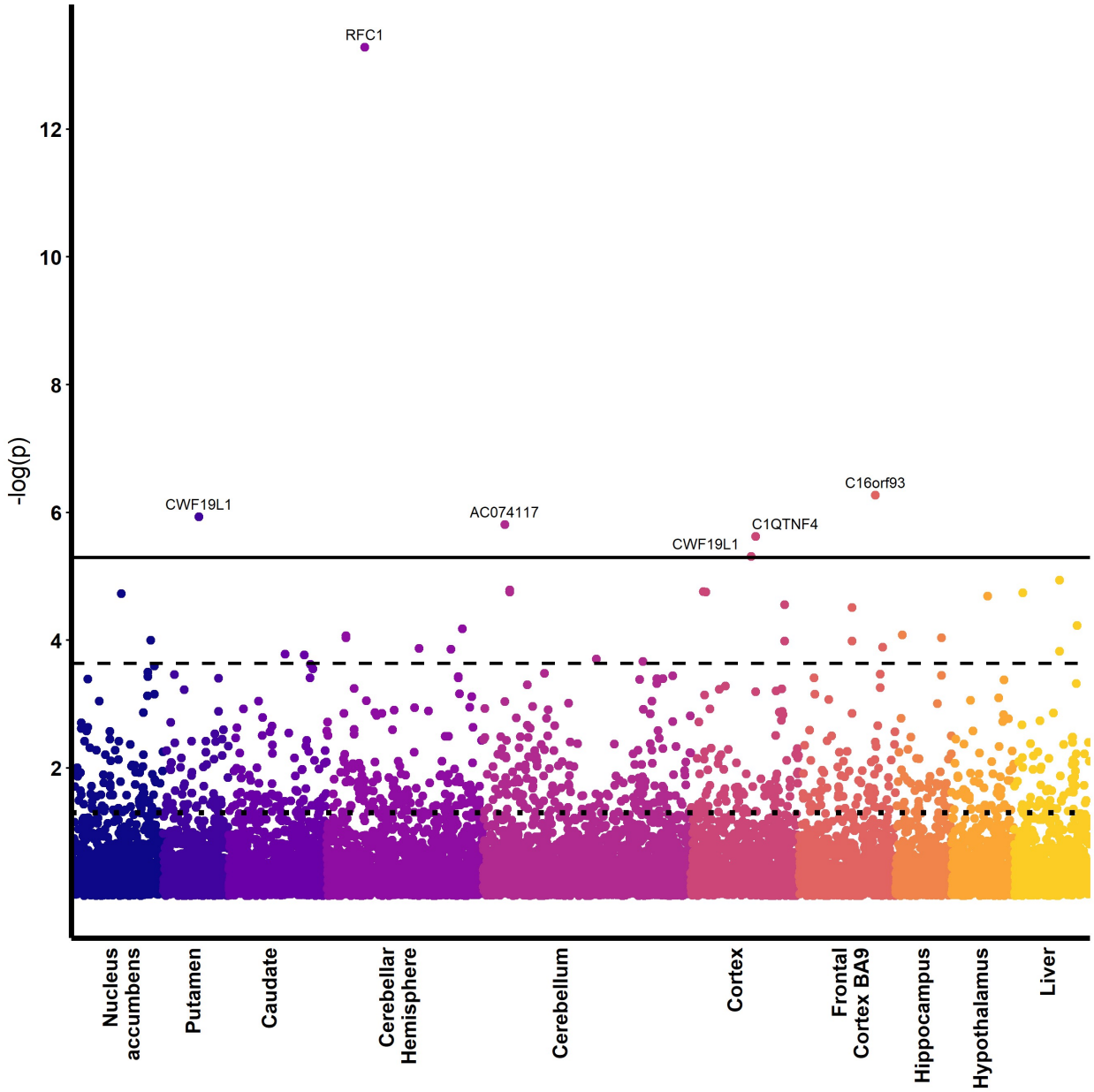
Supplemental References: Pages 31-32

Supplemental Figure 1: Tissue-specific Enrichment of Alcohol Consumption Genomic Risk



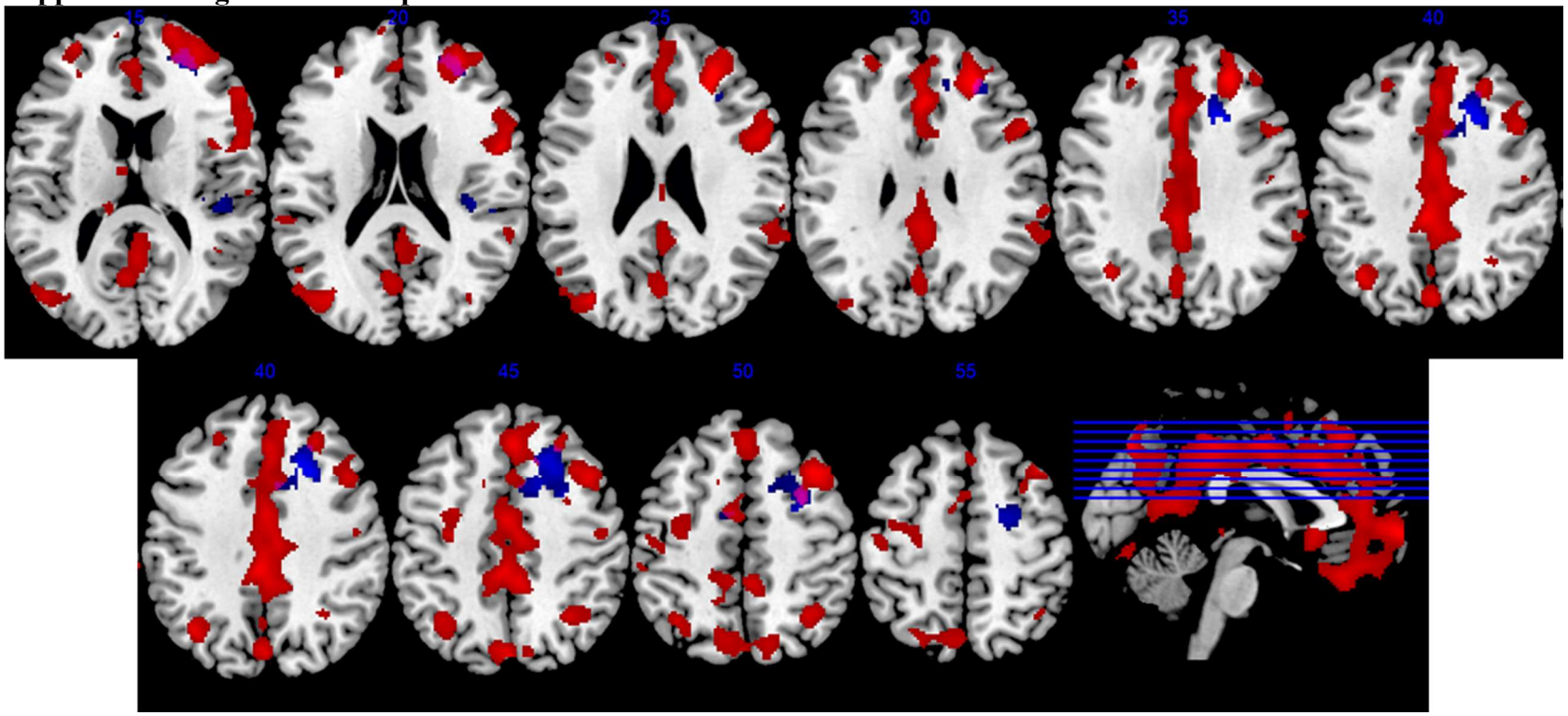
Enrichment of alcohol consumption GWAS (UK Biobank, N=112,117) **A)** associations and **B&C)** heritability, in gene-sets defined by the relative expression of genes across **A&B)** all tissues, and **C)** within the brain, in the GTEX data set (**Supplemental Data**). X-axis and color-scale represent the significance of the enrichment. Solid, dashed, and dotted lines represent Bonferroni-corrected, FDR-corrected, and nominally significant p-values, respectively.

Supplemental Figure 2: TWAS of alcohol consumption predicting gene expression



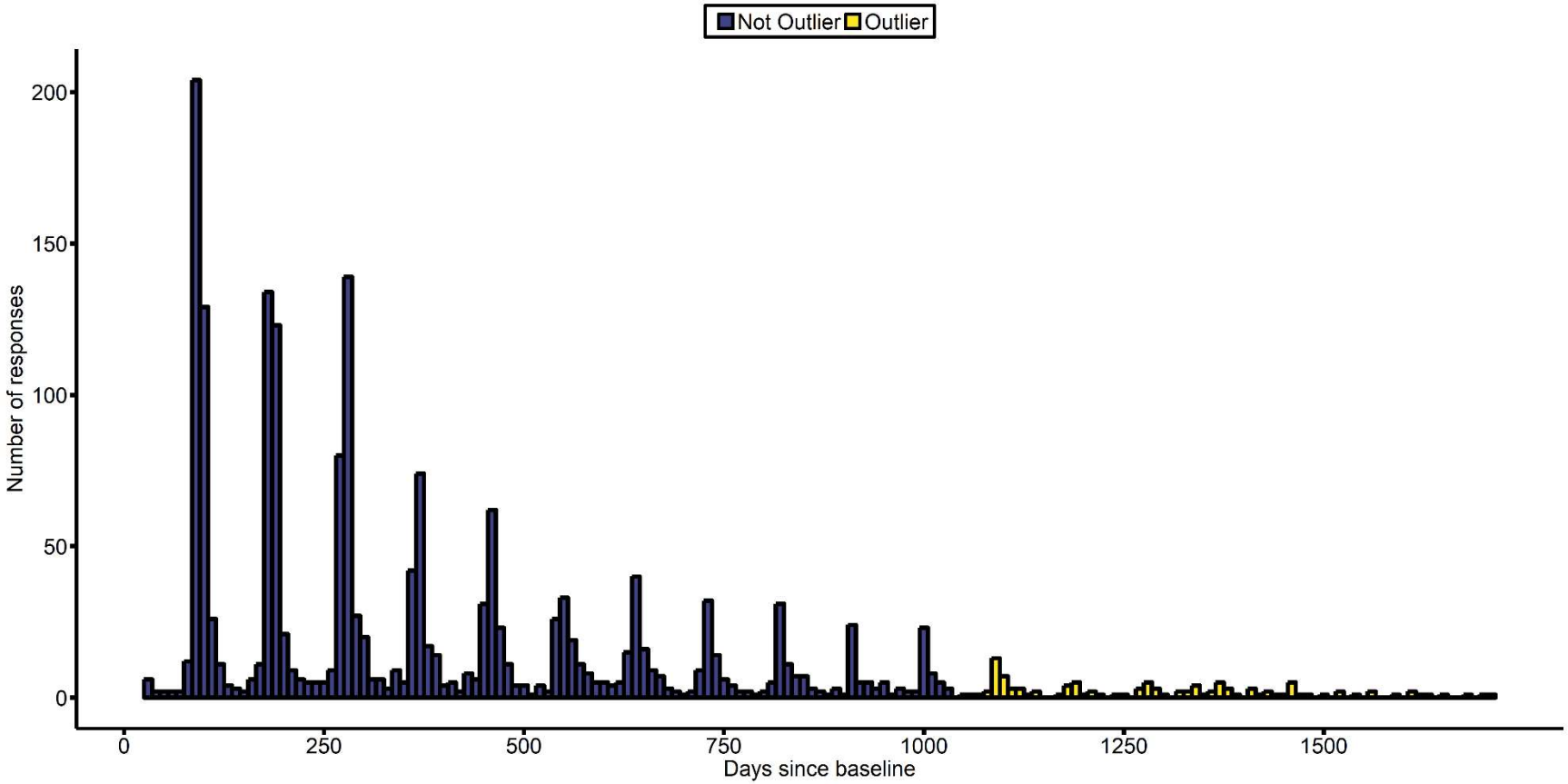
Genetic risk for alcohol consumption according to the UK Biobank GWAS (N=112,117) is associated with differences in human post-mortem gene expression (GTEx; Ns = 81 - 103), including frontal cortex BA9 (**Supplemental Data**). Notably, associations in the liver (far right) do not survive bonferroni-correction for multiple comparisons, though four are significant at a less-stringent FDR-based correction. Y-axis represents the significance of the association. Solid, dashed, and dotted lines represent Bonferroni-corrected, FDR-corrected, and nominally significant p-values, respectively.

Supplemental Figure 3: Overlap of HCP and DNS clusters at an uncorrected threshold



Overlap of volumetric associations in the frontal cortex with alcohol consumption in the DNS (red) and HCP (blue) samples (purple=overlap). Associations are displayed at an uncorrected threshold (DNS: $p<0.001$; HCP: $p<0.05$). For display purposes, different statistical thresholds were used due to the differential power of the two studies.

Supplemental Figure 4: Distribution of responses to DNS follow-up questionnaire



Histogram of responses to the DNS online follow-up questionnaire, which was emailed to participants every 3 months. Color indicates whether the clustering algorithm indicated the response to be an outlier or not – outliers occurred > 1053 days after the baseline visit and were excluded from analyses. Bins have a width of 10 days.

Supplemental Table 1. Location of volumetric reductions associated with alcohol consumption

DNS							
Index	**# voxels**	**p-FWE**	**T**	**x (mm)**	**y (mm)**	**z (mm)**	**AAL-Atlas Location**
1	279	0.003	5.24	27	39	25	R Middle Frontal
1.b		0.01	4.95	32	50	23	R Middle Frontal
2	344	0.004	5.14	56	3	0	R Superior Temporal
2.b		0.006	5.05	48	6	-5	R Insula
3	44	0.005	5.08	0	63	-3	L Medial Orbital Frontal
4	76	0.007	5.03	38	18	51	R Middle Frontal
5	64	0.007	5.03	-3	-33	42	L Middle Cingulum
6	23	0.019	4.8	2	27	30	R Middle Cingulum
7	17	0.023	4.75	29	62	8	R Superior Frontal
8	12	0.029	4.59	2	33	45	R Medial Superior Frontal
HCP							
Index	**# voxels**	**p-FWE**	**T**	**x (mm)**	**y (mm)**	**z (mm)**	**AAL-Atlas Location**
1	42	0.003	4.92	38	12	0	R Insula
2	88	0.008	4.7	20	24	42	R Superior/Middle Frontal

DNS = Duke Neurogenetics Study; HCP = Human Connectome Project. Coordinates are provided in MNI space.

Supplemental Table 2. Heritability and genetic correlation between gray-matter volume and alcohol consumption in the HCP

	mAUDIT-C	**Right Insula**	**Right Middle/Superior Frontal Cortex**
h^2 (SE)	0.5179 (0.0541)	0.6883 (0.0392)	0.7746 (0.0307)
p	**5.25×10^{-18}**	**1.97×10^{-32}**	**3.62×10^{-40}**
ρ_p	-	-0.1349	-0.114
p	-	**0.0006**	**0.0033**
ρ_g (SE)	-	-0.2314 (0.076)	-0.2192 (0.0784)
p	-	**0.0022**	**0.0054**
ρ_e (SE)	-	0.0294 (0.0825)	0.0483 (0.0787)
p	-	0.7214	0.541

h^2 = heritability; ρ_p = phenotypic correlation; ρ_g = genetic correlation; ρ_e = environmental correlation

Supplemental Table 3. Discordant sibling analysis in the HCP

Region	Variable	Estimate	t	p
Right Insula	**ICV**	-5.78×10^{-8}	-4.629	**4.78×10^{-6}**
	Sex	0.0052	0.892	0.3733
	Age	-0.0030	-5.571	**4.27×10^{-8}**
	MZ	-0.0057	-1.596	0.1111
	DZ	-0.0124	-2.659	**0.0081**
	W	0.0179	2.041	**0.0419**
	B	-0.0026	-0.253	0.8007
	A	0.0066	0.525	0.5998
	SES	0.0002	0.289	0.7726
	ED	1.20×10^{-5}	0.362	0.7176
	DX – Non Substance	-0.0054	-1.29	0.1978
	PSS	0.0002	0.967	0.3343
	Causal	0.0009	0.412	0.6806
	Graded	-0.0037	-1.974	**0.0491**
	Predispositonal	0.0037	3.479	**0.0006**
Region	**Variable**	**Estimate**	**t**	**p**
Right Middle/Superior Frontal Cortex	**ICV**	1.38×10^{-8}	1.315	0.1891
	Sex	0.0124	2.584	**0.0103**
	Age	-0.0029	-6.438	**3.01×10^{-10}**
	MZ	-4.89×10^{-5}	-0.016	0.9871
	DZ	-0.0059	-1.505	0.1329
	W	0.0139	1.899	0.0583
	B	0.0208	2.387	**0.0175**
	A	0.0217	2.058	**0.0402**
	SES	0.0008	1.06	0.2896
	ED	-1.87×10^{-5}	-0.665	0.5067
	DX – Non Substance	0.0092	2.557	**0.0109**
	PSS	0.0002	1.042	0.2980
	Causal	-0.0020	-1.037	0.3006
	Graded	0.0006	0.411	0.6810
	Predispositonal	0.0020	2.193	**0.0290**

SES = Socioeconomic status, PSS = Perceived Stress Scale, ED = education

Supplemental Table 4. Regression analyses of the association between brain volume and longitudinal alcohol consumption in DNS

Variable	Value	Std.Error	DF	t-value	p-value
(Intercept)	**0.114**	**0.056**	**1726**	**2.022**	**0.043**
Age-linear	**0.085**	**0.037**	**1726**	**2.266**	**0.024**
Age-quadratic	-0.019	0.031	1726	-0.618	0.537
Baseline Age	-0.026	0.048	646	-0.546	0.585
Sex	**-0.164**	**0.044**	**646**	**-3.698**	**2.35×10^{-4}**
W	**0.161**	**0.062**	**646**	**2.615**	**0.009**
B	0.032	0.045	646	0.714	0.476
A	-0.047	0.058	646	-0.809	0.419
H	**0.093**	**0.039**	**646**	**2.360**	**0.019**
DX - Non Substance	0.016	0.035	646	0.440	0.660
PSS	-0.039	0.038	646	-1.025	0.306
P-SES	0.066	0.041	646	1.594	0.111
CTQ	0.014	0.038	646	0.375	0.708
P-ED	0.074	0.040	646	1.845	0.065
Scanner	0.029	0.032	646	0.932	0.351
ICV	0.107	0.059	646	1.795	0.073
Frontal	**-0.115**	**0.053**	**646**	**-2.185**	**0.029**
Age-linear x Frontal	**0.151**	**0.038**	**1726**	**3.976**	**7.29×10^{-5}**
Age-quadratic x Frontal	0.007	0.025	1726	0.275	0.783
Age-quadratic x Baseline Age	**-0.046**	**0.018**	**1726**	**-2.490**	**0.013**
Age-quadratic x Sex	**-0.057**	**0.022**	**1726**	**-2.602**	**0.009**
Age-quadratic x W	-0.001	0.031	1726	-0.022	0.982
Age-quadratic x B	-0.009	0.022	1726	-0.411	0.681
Age-quadratic x A	0.009	0.028	1726	0.322	0.747
Age-quadratic x H	-0.009	0.021	1726	-0.413	0.679
Age-quadratic x DX - Non Substance	0.017	0.018	1726	0.953	0.341
Age-quadratic x PSS	0.023	0.018	1726	1.269	0.205
Age-quadratic x P-SES	0.019	0.021	1726	0.883	0.377
Age-quadratic x CTQ	0.015	0.018	1726	0.799	0.424
Age-quadratic x P-ED	-0.006	0.020	1726	-0.284	0.777
Age-quadratic x Scanner	0.009	0.018	1726	0.479	0.632
Age-quadratic x ICV	-0.026	0.028	1726	-0.936	0.349
Age-linear x Baseline Age	-0.019	0.048	1726	-0.390	0.696
Age-linear x Sex	0.057	0.031	1726	1.876	0.061
Age-linear x W	-0.014	0.042	1726	-0.325	0.745
Age-linear x B	-0.021	0.030	1726	-0.702	0.483
Age-linear x A	-0.010	0.039	1726	-0.256	0.798
Age-linear x H	0.042	0.028	1726	1.519	0.129
Age-linear x DX - Non Substance	**-0.053**	**0.025**	**1726**	**-2.153**	**0.031**
Age-linear x PSS	-0.015	0.025	1726	-0.595	0.552
Age-linear x P-SES	-0.047	0.028	1726	-1.654	0.098

Age-linear x CTQ	0.004	0.026	1726	0.160	0.873
Age-linear x P-ED	0.024	0.027	1726	0.889	0.374
Age-linear x Scanner	0.018	0.024	1726	0.764	0.445
Age-linear x ICV	**-0.090**	**0.040**	**1726**	**-2.243**	**0.025**
Baseline Age x Frontal	-0.016	0.041	646	-0.386	0.700
Sex x Frontal	0.073	0.051	646	1.442	0.150
W x Frontal	-0.043	0.065	646	-0.669	0.503
B x Frontal	0.035	0.046	646	0.758	0.449
A x Frontal	-0.056	0.060	646	-0.924	0.356
H x Frontal	-0.018	0.043	646	-0.429	0.668
DX - Non Substance x Frontal	0.011	0.034	646	0.322	0.748
PSS x Frontal	0.023	0.038	646	0.605	0.546
P-SES x Frontal	0.009	0.043	646	0.213	0.832
CTQ x Frontal	-0.052	0.037	646	-1.407	0.160
P-ED x Frontal	0.034	0.042	646	0.819	0.413
Scanner x Frontal	0.001	0.035	646	0.038	0.970
ICV x Frontal	0.008	0.044	646	0.195	0.846

Volume of the right middle/superior frontal cortex is associated with future change in alcohol consumption (p-fdr=$4x10^{-4}$). Standardized effects and uncorrected p-values are presented.

Supplemental Table 5. Regression analyses of the association between brain volume and future alcohol use initiation in TAOS.

	Middle Frontal				Superior Frontal			
	Estimate	SE	z	p	Estimate	SE	z-value	p
(Intercept)	-6.64	156.30	-0.04	0.97	-6.35	132.19	-0.05	0.96
Age-linear	170.77	5767.82	0.03	0.98	157.94	2198.72	0.07	0.94
Age-quadratic	-27.21	4154.01	-0.01	0.99	-25.34	1750.80	-0.01	0.99
Sex	-0.04	0.94	-0.04	0.97	0.15	0.80	0.19	0.85
W	0.99	1.68	0.59	0.56	0.69	1.38	0.50	0.62
H	-0.18	1.63	-0.11	0.91	-0.23	1.32	-0.17	0.86
High-risk	0.30	0.82	0.37	0.71	0.05	0.73	0.07	0.95
GenPop	-2.70	1217.35	0.00	1.00	-3.35	1029.61	0.00	1.00
SLES	0.83	0.93	0.89	0.37	0.94	0.78	1.20	0.23
SES	-0.10	0.83	-0.13	0.90	-0.34	0.70	-0.48	0.63
CTQ	1.15	0.77	1.50	0.13	1.02	0.68	1.51	0.13
Tanner	-0.51	0.83	-0.61	0.54	-0.28	0.73	-0.39	0.70
MFQ	-0.03	0.85	-0.03	0.97	-0.11	0.71	-0.16	0.88
Baseline Age	-0.01	0.82	-0.01	0.99	0.10	0.71	0.15	0.88
Intracranial Volume (ICV)	-0.81	1.18	-0.68	0.50	-1.07	1.05	-1.02	0.31
Brain	1.58	114.78	0.01	0.99	2.18	118.72	0.02	0.99
Age-linear x Brain	**-57.04**	**24.04**	**-2.37**	**0.02**	**-60.74**	**24.99**	**-2.43**	**0.02**
Age-quadratic x Brain	-11.91	22.47	-0.53	0.60	11.99	21.43	0.56	0.58
Age-linear x Sex	-16.07	23.59	-0.68	0.50	-19.74	20.14	-0.98	0.33
Age-linear x W	-6.29	44.08	-0.14	0.89	-5.61	35.53	-0.16	0.87
Age-linear x H	2.52	42.77	0.06	0.95	3.35	34.36	0.10	0.92
Age-linear x High Risk	-12.31	20.19	-0.61	0.54	-6.61	18.37	-0.36	0.72
Age-linear x GenPop	-24.67	44924.09	0.00	1.00	-3.90	17124.75	0.00	1.00
Age-linear x SLES	-9.99	24.93	-0.40	0.69	-15.30	21.04	-0.73	0.47
Age-linear x SES	25.87	21.32	1.21	0.23	26.53	17.57	1.51	0.13
Age-linear x CTQ	1.92	19.66	0.10	0.92	4.47	17.83	0.25	0.80
Age-linear x Tanner	16.91	19.24	0.88	0.38	15.46	18.07	0.86	0.39
Age-linear x MFQ	-11.20	19.34	-0.58	0.56	-12.20	16.73	-0.73	0.47
Age-linear x Baseline Age	10.07	21.73	0.46	0.64	3.79	18.30	0.21	0.84
Age-linear x ICV	20.43	28.01	0.73	0.47	21.73	25.25	0.86	0.39
Age-quadratic x Sex	12.71	18.66	0.68	0.50	20.88	15.54	1.34	0.18
Age-quadratic x W	15.10	30.63	0.49	0.62	4.74	24.89	0.19	0.85
Age-quadratic x H	6.69	27.90	0.24	0.81	1.36	22.93	0.06	0.95
Age-quadratic x High Risk	18.42	15.16	1.21	0.22	12.60	13.16	0.96	0.34
Age-quadratic x GenPop	51.69	32354.53	0.00	1.00	41.45	13636.19	0.00	1.00
Age-quadratic x SLES	-2.61	20.93	-0.12	0.90	0.63	17.27	0.04	0.97
Age-quadratic x SES	-7.67	16.53	-0.46	0.64	-10.84	13.83	-0.78	0.43
Age-quadratic x CTQ	1.29	17.12	0.08	0.94	-2.24	15.00	-0.15	0.88
Age-quadratic x Tanner	-2.06	15.21	-0.14	0.89	0.47	13.83	0.03	0.97
Age-quadratic x MFQ	-20.32	16.51	-1.23	0.22	-18.46	14.09	-1.31	0.19

Age-quadratic x Baseline Age	-11.03	15.32	-0.72	0.47	-5.90	13.37	-0.44	0.66
Age-quadratic x ICV	18.31	21.99	0.83	0.41	7.89	20.98	0.38	0.71
Sex x Brain	-0.42	0.76	-0.55	0.58	-0.76	0.67	-1.14	0.25
W x Brain	-0.07	1.10	-0.06	0.95	0.39	0.89	0.43	0.67
H x Brain	0.25	1.06	0.23	0.81	0.52	0.87	0.59	0.56
High Risk x Brain	-0.01	0.59	-0.01	0.99	0.31	0.52	0.60	0.55
GenPop x Brain	1.95	893.93	0.00	1.00	2.23	924.63	0.00	1.00
SLES x Brain	0.34	0.60	0.57	0.57	-0.08	0.53	-0.16	0.88
SES x Brain	0.64	0.64	1.00	0.32	0.67	0.54	1.25	0.21
CTQ x Brain	-0.12	0.66	-0.18	0.86	-0.11	0.57	-0.20	0.84
Tanner x Brain	0.52	0.65	0.80	0.43	0.25	0.58	0.44	0.66
MFQ x Brain	-0.20	0.70	-0.28	0.78	-0.15	0.58	-0.26	0.79
Baseline Age x Brain	0.73	0.68	1.07	0.29	0.21	0.51	0.41	0.68
ICV x Brain	-0.83	0.77	-1.09	0.28	-1.00	0.68	-1.48	0.14

Volume of the right middle and superior frontal cortices are both associated with future initiation of alcohol consumption (p-fdr=0.035, 0,035). Standardized effects and uncorrected p-values are presented.

Supplemental Table 6. TWAS Discovery and Replication

		Discovery – GTEX Frontal BA 9				Replication – CMC DLPFC			
Gene	**CHR**	**Locus Start**	**Locus End**	**TWAS Z**	**TWAS P**	**Locus Start**	**Locus End**	**TWAS Z**	**TWAS P**
C16orf93	16	30772519	30772656	5.0152	5.30E-07	-	-	-	-
CWF19L1	10	102000000	102000000	-4.1674	3.08E-05	102000000	102000000	-2.11116	0.0348
PHBP9	10	102000000	102000000	-3.8862	1.02E-04	-	-	-	-
C18orf8	18	21083473	21110576	3.831	1.28E-04	21083433	21111771	2.1613	0.0307

* Empty rows in replication data indicate that gene was not present in replication dataset. P-values are all uncorrected – all associations listed survive FDR correction.

Supplemental Table 7. Comparison of samples

	DNS (N=1303)	HCP (N=897)	TAOS+ (N=223)	t, F, χ^2	*p*
Age	19.7 (1.25) 18-22	28.82 (3.68) 22-37	13.42 (0.96) 11.61-15.33	F = 5420.39	$<1x10^{-300}$
Sex (% Female)	747 (57.33%)	504 (56.19%)	96 (43.05%)	$\chi 2$ = 15.93	**0.000347**
m/AUDIT-C	3.76 (2.64) 0-11.69	3.42 (2.65) 0-11.38	1.26 (1.95) 0-9	F = 89.43	**$3.38x10^{-38}$**
European/ European American	579 (44.44%)	597 (66.56%)	126 (56.5%)	$\chi 2$ = 105.31	**$1.35x10^{-23}$**
African/African American	148 (11.36%)	155 (17.28%)	3 (1.35%)	$\chi 2$ = 45.22	**$1.51x10^{-10}$**
Asian/Asian American	353 (27.09%)	45 (5.02%)	4 (1.79%)	$\chi 2$ = 225.94	**$8.68x10^{-50}$**
Hispanic	83 (6.37%)	73 (8.14%)	73 (32.74%)	$\chi 2$ = 157.54	**$6.18x10^{-35}$**
Multi-racial/Native American/Other	140 (10.74%)	27 (3.01%)	17 (7.62%)	$\chi 2$ = 45.29	**$1.46x10^{-10}$**
CTQ	33.39 (8.12) 25-59.81	-	32.98 (7.09) 25-64	t = 0.5	0.478993
PSS	14.65 (6.05) 0-32.89	13.07 (5.76) 0-35	-	t = 37.59	**$1.03x10^{-09}$**

Comparisons of measures that overlap across samples (DNS, HCP, TAOS). Continuous variables are presented in the format of: Mean (Standard deviation) Range. Dichotomous variables show the number of participants in that category, and the total percent of the sample that they comprise. Variables were winsorized prior to comparisons.

+= TAOS comparisons of mAUDIT-C scores were done with the maximum score attained by each participant.

-=Variable not present in sample.

DNS = Duke Neurogenetics Study; HCP = Human Connectome Project; TAOS = Teen Alcohol Outcomes Study; AUDIT = Alcohol Use Disorder Identification Test; PSS = Perceived Stress Scale; CTQ = Childhood Trauma Questionnaire.

Supplemental Table 8. Frequency of Psychiatric Diagnosis in the DNS and HCP samples.

Dataset	Disorder	N
DNS	Any	258
	Depression	66
	Bipolar	35
	Anxiety	50
	OCD	16
	PTSD	2
	Alcohol Abuse	143
	Substance Abuse	47
	Eating Disorder	11
	Autism Spectrum	1
	Psychosis	3
HCP	Any	311
	Agoraphobia	55
	Panic Disorder	53
	Depression	75
	Alcohol Abuse/Dependence	176
	Marijuana Abuse/Dependence	103

80 participants in the DNS have more than one diagnosis.
103 participants in the HCP have more than one diagnosis.

Supplemental Table 9. Comparison of DNS subjects who did or did-not complete follow-up questionnaires

	No Follow-Up (N=624)	**Follow-Up (N=679)**	**t/χ^2**	***p***
AUDIT-C	4.0492 (2.6632)	3.4993 (2.595)	3.7694	**0.0002**
AUDIT Total	5.5913 (4.4163)	4.8778 (4.1872)	2.9869	**0.0029**
Age	19.8125 (1.2739)	19.592 (1.2265)	3.1765	**0.0015**
P-SES	7.5065 (1.6722)	7.3884 (1.7094)	1.2603	0.2078
P-ED	7.0232 (1.6025)	7.0751 (1.7291)	-0.562	0.5742
PSS	14.4852 (5.9435)	14.7976 (6.1404)	-0.9328	0.3511
CTQ	33.7902 (8.3233)	33.016 (7.9077)	1.7179	0.0861
Frontal Volume	0.4834 (0.0559)	0.4815 (0.0564)	0.6047	0.5455
Insula Volume	0.5498 (0.0539)	0.5505 (0.0547)	-0.2484	0.8038
ICV	1.4877 (0.1451)	1.4754 (0.1384)	1.5573	0.1196
Sex (number of female respondents)*	N=320 (51.28%)	N=427 (62.89%)	17.429	**2.98×10^{-05}**
Diagnosis (non substance-related) *	N=77 (12.34%)	N=72 (10.6%)	0.8037	0.3700
Caucasian*	N=253 (40.54%)	N=326 (48.01%)	7.0435	**0.0080**
African/African American*	N=86 (13.78%)	N=62 (9.13%)	6.5319	**0.0106**
Asian/Asian American*	N=171 (27.4%)	N=182 (26.8%)	0.0327	0.8564
Hispanic*	N=42 (6.73%)	N=41 (6.04%)	0.1582	0.6908
Multi-racial/Native American/Other*	N=72 (11.54%)	N=68 (10.01%)	0.6364	0.4250

*=Analysis was run as a chi-squared test, all others were run as t-tests.

AUDIT = Alcohol Use Disorder Identification Test; CTQ = Childhood Trauma Questionnaire; P-ED = Parental Education; P-SES = Perceived Socioeconomic Status; PSS = Perceived Stress Scale. ; ICV = Intracranial Volume

Supplemental Table 10. Association of Alcohol Consumption/Initiation with Covariates

Study	**DNS**		**HCP**		**TAOS***	
Variable	**statistic (t, F, r)**	***p***	**statistic (t, F, r)**	***p***	**statistic (t, χ^2)**	***p***
Sex (Female vs Male)	t = 9.1624	**2.70×10^{-19}**	t = 8.7447	**1.72×10^{-17}**	$\chi 2 = 6.24 \times 10^{-31}$	1.0000
Diagnosis (non Substance-related) (vs none)	t = -0.3185	0.7505	t = -0.6746	0.5007	-	-
Ethnicity	F = 29.7684	**5.82×10^{-8}**	F = 2.6516	0.1038	$\chi 2 = 6.5217$	0.5890
European/ European American (vs not)	t = -7.6998	**2.84×10^{-14}**	t = -3.2474	**0.0012**	$\chi 2 = 0.0424$	0.8368
African/African American (vs not)	t = 3.1538	**0.0019**	t = 4.1931	**3.89×10^{-5}**	-	-
Asian/Asian American (vs not)	t = 5.0911	**4.54×10^{-7}**	t = 1.4615	0.1502	-	-
Hispanic (vs not)	t = -0.0295	0.9765	t = -0.9598	0.3398	$\chi 2 = 0.0908$	0.7631
Multi-racial/Native American/Other (vs not)	t = 2.2086	**0.0285**	t = -0.3249	0.7477	-	-
Age	r = 0.126	**5.10×10^{-6}**	r = -0.1222	**0.0002**	t = -2.8750	**0.0045**
P-SES/SES	r = 0.1747	**2.15×10^{-10}**	r = 0.0102	0.7604	t = -1.7684	0.0789
PSS/SLES	r = -0.0383	0.1669	r = 0.023	0.4908	t = -0.8601	0.3908
P-ED/ED	r = 0.0883	**0.0014**	r = -0.0153	0.6465	-	-
CTQ	r = -0.0966	**0.0005**	-	-	t = -1.6842	0.0945
Scanner (1 vs 2)	t = -0.5507	0.5822	-	-	-	-
High-risk for depression (TAOS-only)	-	-	-	-	t = 0.5780	0.4471
Tanner stage	-	-	-	-	t = -2.0118	**0.0459**
MFQ	-	-	-	-	t = 0.2417	0.8094

Association of model covariates with alcohol consumption (AUCIT-C/mAUDIT-C).
* = For analyses in TAOS, comparisons are between baseline measurements of participants who initiate, and those who do not.
Comparisons were run as t-tests, anovas, pearson correlations, or chi-squared tests. SES = Socioeconomic status, PSS = Perceived Stress Scale, ED = education, CTQ = Childhood Trauma Questionnaire, MFQ = Mood and Feelings Questionnaire; SLES = Stressful Life Events Schedule. DNS = Duke Neurogenetics Study; HCP = Human Connectome Project; TAOS = Teen Alcohol Outcomes Study
- = Not present in sample, not applicable, or insufficient numbers to run test.

Supplemental Table 11: Brain volume does not predict impulsivity, negative urgency, or intelligence.

		DNS					HCP				
Y	**X**	**Beta**	**Δ r^2**	**P**	**P-fdr**	**N**	**Beta**	**Δ r^2**	**P**	**P-fdr**	**N**
Coping	Frontal	-0.0876	0.0033	0.0293	0.1529	1303	-	-	-	-	-
	Insula	-0.0915	0.0036	0.0237	0.1529	1303	-	-	-	-	-
Self-report Impulsivity	Frontal	0.0487	0.0010	0.1909	0.2864	1303	-0.0844	0.0053	0.0103	0.0824	897
	Insula	0.0680	0.0020	0.0692	0.1529	1303	-0.0549	0.0046	0.1045	0.3739	897
IQ	Frontal	0.0760	0.0025	0.0514	0.1529	1275	0.0060	-5.24E-05	0.8295	0.9271	897
	Insula	0.0699	0.0021	0.0764	0.1529	1275	-0.0026	-2.71E-05	0.9271	0.9271	897
DDT - K	Frontal	-0.0891	0.0034	0.0436	0.1529	1157	-0.0142	3.62E-05	0.6877	0.9169	891
	Insula	-0.0320	0.0004	0.4720	0.5664	1157	0.0534	-0.0089	0.1402	0.3739	891
NEO - N5	Frontal	-0.0636	0.0018	0.1024	0.1756	1302	-	-	-	-	-
	Insula	-0.0478	0.0010	0.2232	0.2976	1302	-	-	-	-	-
NEO - N	Frontal	-0.0008	2.64E-07	0.9789	0.9789	1302	-0.0212	-0.0001	0.3848	0.6157	891
	Insula	-0.0080	2.73E-05	0.7879	0.8595	1302	-0.0308	0.0014	0.2337	0.4674	891

Parameters from linear regression models of brain volume predicting behavior. Covariates were the same as in whole-brain analyses. Coping = substance-use subscale of the brief COPE. Self-report impulsivity: DNS - Barratt Impulsiveness Scale; HCP - Achenbach Adult Self-Report Impulsivity. IQ: DNS: WASI-II; HCP: NIH Toolbox. DDT – K: Delayed discounting task. NEO-N: Neuroticism. NEO-N5: Neuroticism Impulsivity subscale. Standardized effect-sizes (ie beta values) are presented. Δ r^2: Change in model-fit when the x-variable is added (negative values indicate the model-fit decreased).

Supplemental Table 12: Brain volume clusters do not predict impulsivity, negative urgency, or intelligence.

Study	Y	X	Beta	Δ r^2	P	P-fdr	N
DNS	BCOPE - Sub	Insula_48_6_5	-0.0767	0.0044	0.0127	0.1110	1303
		MidFrontal_27_39_26	-0.0888	0.0048	0.0089	0.1110	1303
		MidFrontal_38_18_51	-0.0291	0.0006	0.3762	0.5424	1303
		SupFrontal_29_62_8	-0.0706	0.0030	0.0392	0.1880	1303
	BIS	Insula_48_6_5	0.0313	0.0007	0.2726	0.5424	1303
		MidFrontal_27_39_26	0.0402	0.0010	0.2010	0.4825	1303
		MidFrontal_38_18_51	-0.0006	2.44×10^{-07}	0.9840	0.9924	1303
		SupFrontal_29_62_8	0.0325	0.0006	0.3053	0.5424	1303
	IQ	Insula_48_6_5	0.0232	0.0004	0.4388	0.5850	1275
		MidFrontal_27_39_26	0.0793	0.0038	0.0163	0.1110	1275
		MidFrontal_38_18_51	0.0137	0.0001	0.6669	0.8003	1275
		SupFrontal_29_62_8	0.0329	0.0007	0.3219	0.5424	1275
	DDT - k	Insula_48_6_5	-0.0315	0.0007	0.3509	0.5424	1157
		MidFrontal_27_39_26	-0.0223	0.0003	0.5521	0.6974	1157
		MidFrontal_38_18_51	-0.0321	0.0007	0.3709	0.5424	1157
		SupFrontal_29_62_8	-0.0644	0.0025	0.0864	0.2593	1157
	NEO - N5	Insula_48_6_5	-0.0441	0.0014	0.1396	0.3721	1302
		MidFrontal_27_39_26	-0.0287	0.0005	0.3842	0.5424	1302
		MidFrontal_38_18_51	-0.0749	0.0037	0.0185	0.1110	1302
		SupFrontal_29_62_8	-0.0646	0.0025	0.0515	0.2059	1302
	NEO - N	Insula_48_6_5	0.0063	2.98×10^{-05}	0.7788	0.8608	1302
		MidFrontal_27_39_26	0.0002	3.45×10^{-08}	0.9924	0.9924	1302
		MidFrontal_38_18_51	-0.0431	0.0012	0.0730	0.2501	1302
		SupFrontal_29_62_8	0.0067	2.71×10^{-05}	0.7890	0.8000	1302
HCP	DDT - k	Frontal_20_24_42	0.0545	-0.0058	0.1084	0.4279	891
		Insula_38_12_0	0.0258	-0.0012	0.4520	0.5668	891
	ASR_IMP	Frontal_20_24_42	-0.0453	0.0008	0.1605	0.4279	897
		Insula_38_12_0	-0.0328	0.0034	0.3124	0.5668	897
	NEO - N	Frontal_20_24_42	0.0166	-0.0003	0.4960	0.5668	891
		Insula_38_12_0	-0.0198	0.0010	0.4195	0.5668	891
	IQ	Frontal_20_24_42	-0.0446	-0.0005	0.0747	0.4279	897
		Insula_38_12_0	-0.0115	1.84×10^{-05}	0.6506	0.6506	897

Parameters from linear regression models of brain volume predicting behavior. Covariates were the same as in whole-brain analyses. BCOPE - Sub = substance-use subscale of the brief COPE. BIS = Barratt Impulsiveness Scale; ASR IMP = Achenbach Adult Self-Report Impulsivity. IQ: DNS: WASI-II; HCP: NIH Toolbox. DDT – K: Delayed discounting task. NEO-N: Neuroticism. NEO-N5: Neuroticism Impulsivity subscale. Standardized effect-sizes (ie beta values) are presented. Δ r^2: Change in model-fit when the x-variable is added (negative values indicate the model-fit decreased).

Figure Creation

Figures 2 &3, and Supplemental Figures 1-3 were made in R (3.4.2) (R Core Team, 2014). The packages 'ggplot2'[1] and 'ggpubr'[2] were used to generate the figures, and the package 'viridis'[3] was used for the color-palette. 'ggrepel'[4] was used in Supplemental Figure 2 for gene-labels.

Supplemental Results

Comparison of discovery and replication samples

Sample comparisons are presented in **Supplemental Table 7**. The samples differed by sex, which was driven by fewer female participants in the TAOS sample (the DNS and HCP do not differ). Consistent with the recruitment of non-overlapping aged samples, HCP participants were all older than DNS participants, who were all older than TAOS participants. A significantly larger proportion of the participants from the HCP were Caucasian or African/African-American, while the DNS had a larger portion of Asian/Asian-American and Multi-racial/Native-American/Other participants. TAOS had the largest proportion of Hispanic participants. DNS participants had significantly higher AUDIT-C scores, which is unsurprising given observations of elevated hazardous alcohol consumption among younger college-aged populations[5]. Similarly, the higher levels of self-reported perceived stress among DNS participants is consistent with prior observations of increased stress in college-student samples[6].

Association of covariates with alcohol consumption

Association of covariates with alcohol consumption are presented in **Supplemental Table 10**. In the DNS and HCP men reported higher levels of alcohol consumption than women, consistent with prior reports of sex-differences in alcohol use[7]. Notably, in TAOS, sex did not differ between initiators and non-initiators. Presence of a non-substance-related diagnosis was not associated with alcohol use in either the DNS or HCP. Consistent with prior reports, participants of European descent reported increased levels, while participants of African descent reported decreased levels [8,9]. In the DNS participants of Asian descent and of Multi-racial or Native-American descent reported the lowest amount of alcohol consumption [9]. These associations did not reach significance

in the HCP, likely due to the smaller number of participants in these groups in this sample. Age was associated with alcohol consumption in all samples, though in differing directions, which is consistent with observations in North American samples that alcohol consumption increases once young adults turn 21 (legal drinking age), and subsequently decreases as participants age[7,10]. Of note, in TAOS alcohol-consumption initiators were younger at baseline. In the DNS, perceived SES (P-SES) and parental education (P-ED) were both positively correlated with alcohol consumption, while no association was observed with these phenotypes in the HCP or in TAOS. Perceived stress was not associated with alcohol consumption in any sample, which is consistent with some prior reports [11,12]. Childhood trauma was associated with alcohol consumption in the DNS, and was trending in TAOS, consistent with an extensive literature indicating that early trauma increases risk for substance use[13].

Lower gray matter volume associated with alcohol consumption: Clusters in the discovery DNS analysis

The analysis in the DNS discovery sample identified eight clusters of lower gray-matter volume, extending across eight different regions, associated with increased AUDIT-C scores ($p<0.05$ FWE-corrected, **Figure 1, Supplemental Table 1**). Significant clusters included two clusters in the right middle frontal cortex, a large cluster extending across the right superior temporal cortex and right insula, and clusters in the left medial orbital frontal cortex, bilateral middle cingulum, right superior frontal cortex, and right medial superior frontal cortex.

Longitudinal Data Results

DNS: Differences between longitudinal responders and non-responders are presented in **Supplemental Table 9**. Participants who completed at least one follow-up questionnaire were younger, had lower AUDIT-C and total AUDIT scores. Further, responders were more likely to be female and white and less likely to be black. Of the four brain volume x age interactions modeled (Frontal x Linear-Age, Frontal x Quadratic-Age, etc.), only the interaction between right middle/superior frontal volume and the linear change in age was significant after fdr-correction for multiple comparisons (β=0.151, t=3.976, p=0.0007, p-fdr=0.008; **Figure 3, Supplemental Table 4**). Examination of regions of significance found that lower frontal volume predicted greater drinking before the age of 20.85 years.

TAOS: Of the n=223 adolescents who were non-drinkers at baseline that were included in analyses, n=82 (36%) reported having consumed at least one full alcoholic beverage during one of their follow-up interviews (i.e. they initiated consumption). Of the four brain volume x age interactions modeled (Mid-Frontal x Linear-Age, Mid-Frontal x Quadratic-Age, etc.), the interactions between both right middle and superior frontal volume and the linear change in age were significant after fdr-correction for multiple comparisons (Middle frontal: β=-57.04, t= -2.37, p-fdr = 0.036; Superior frontal: β=-60.74, t = -2.43, p-fdr=0.036; **Figure 3, Supplemental Table 5**).

Tissue-specific Enrichment of Alcohol Consumption

Enrichment analyses found evidence of significant (bonferroni-corrected) enrichment of brain-enriched gene-sets in the genetic associations of alcohol consumption and the heritability of alcohol consumption (**Supplemental Figure 1A&B; Supplemental Data)**, relative to non-brain

tissues. However, when gene-sets were generated by comparing brain-tissues only to each other, no gene-set survived correction for multiple comparisons (**Supplemental Figure 1C; Supplemental Data**). Notably, Brodmann Area 9 (BA9), which is included in the frontal ROI observed in our discovery and replication analyses, was among the significant regions in both analyses.

TWAS of Alcohol Consumption and Gene expression in the brain

TWAS identified several genes whose expression in the human brain was correlated with genetic risk for alcohol consumption (**Supplemental Figure 2; Supplemental Data**). Six genes survived bonferroni-correction for multiple comparison, one of which in the Frontal Cortex – *C16orf93*. This gene was not present in the replication data set, but two of the three other genes which passed fdr-correction in the discovery data were - *CWF19L1* and *C18orf8*. Expression of both genes showed significant effects of genetic risk for alcohol consumption – fdr-correction for n=2 comparisons – in the same direction as the discovery dataset (**Supplemental Table 6**).

Post-hoc Associations of volume with behavior

Methods

Following the observation that frontal and insula volume are predictive of future alcohol use and initiation, are predispositional to alcohol consumption, and that genetic risk for alcohol consumption is associated with changes in gene expression in the frontal cortex (**Supplemental Results**), *post-hoc* exploratory analyses sought to test whether the effects of volume on alcohol consumption are mediated by cognitive or behavioral measures. Associations between total volume of the ROI, as well as volume of each significant cluster, and three classes of outcomes were tested: impulsivity, negative urgency, and intelligence. As measures of negative urgency were not collected in the HCP, associations with neuroticism were additionally examined. The association between impulsivity and alcohol consumption is well established[14,15], and measures of impulsivity have been shown to be associated with the structure of both the insula and frontal cortex in large samples[16–19]. Negative urgency, a facet of impulsivity characterized by risky decision making when one is experiencing negative emotions, is associated with problematic drinking[20–22], and shows correlations with neural responses to reward[23], as well as with structure of the frontal cortex[24]. Childhood intelligence has been observed to be predictive of adult alcohol consumption[25], with evidence of genetic correlations between proxy-measures of IQ (i.e. educational attainment) and alcohol consumption[26], and brain structure has been repeatedly linked to intelligence, with evidence that the two share genetic underpinnings[27–29].

Delay Discounting

In delay discounting tasks participants choose between hypothetical amounts of money available immediately or after a delay. By varying the amount of money available immediately and the number of days that one would have to wait for the delayed money, participant 'indifference points' can be identified wherein the participant is equally likely to choose a smaller reward sooner versus a larger reward later. A preference for a smaller reward sooner (i.e., delay discounting) is considered a behavioral index of impulsivity [30]. The DNS and HCP used different protocols which are described below.

DNS: All combinations of immediate reward (varying from $0.10 to $105) and delay intervals [0, 7, 30, 90, 180, 365, or 1825 (i.e., 5 years) days] for the delayed reward of $100 were presented on a computer screen in randomized order[31,32]. 146 participants did not complete the delay discounting task, resulting in a final DNS sample of 1,157 for delay discounting analyses.

HCP: The delayed-reward amount was set to $200, with delays of 1, 6, 12 (1 year), 36 (3 years), 60 (5 years), or 120 months (10 years) presented in the order: 6, 36, 1, 60, 120, 12 months[33]. Immediate reward amounts were adjusted on a trial-by-trial basis based upon participant response. Six participants did not complete the delayed-discounting task.

Impulsivity

DNS participants completed the 30-item self-report Barratt Impulsiveness Scale (BIS)[34]. HCP participants completed the Achenbach Adult Self-Report (ASR) for Ages 18-59[35]. As in prior reports[36], a coarse measure of impulsivity (ASR-Imp) was computed from three questions in the ADHD subscale of the ASR. **DNS**: Barratt Impulsiveness Scale assesses the personality/behavioral construct of impulsiveness and had good internal consistency (α=0.84; M=61.69; SD=9.55; range

37-113). **HCP**: Our Achenbach Adult Self-Report (ASR) impulsivity composite had acceptable internal consistency (ASR-Imp; α=0.79; M=1.29; SD=1.24; range 0-6).

Neuroticism

Participants in the DNS completed the 240-item NEO Personality Inventory-Revised[37] (Neuroticism: α=0.85; M=86.04; SD=22.65; range: 37-113). Six participants did not complete the NEO. Participants in the HCP completed the 60-item NEO Five-Factor Inventory[38] (Neuroticism: α=0.83; M=16.43; SD=7.17; range: 0 - 43).

Negative Urgency

Participants completed two measures of negative urgency – the Impulsivity sub-scale of the NEO-PI-R[37] (α=0.71; M=17.07; SD=4.60; range: 3-32), and the substance-sue subscale of the brief COPE (BCOPE-sub) inventory. The BCOPE-sub consists of two items that assess how frequently respondents use drugs and alcohol as a coping mechanism[39] (α = 0.92; M=2.53; SD=1.11; range: 0-8).

Intelligence

DNS: Intelligence was assessed using the Wechsler Abbreviated Scale of Intelligence Second Edition (WASI-II) 2-subtest version[40], consisting of the Vocabulary and Matrix Reasoning subtests. The total score was computed as the sum of age-adjusted performance on the two sub-tests. 28 participants did not complete the WASI-II. **HCP:** Intelligence was assessed using the NIH toolbox[41]. Fluid intelligence was assessed using the Flanker Inhibitory Control and Attention Test, Picture Sequence Memory Test, List Sorting Test, Pattern Comparison Test, and

Dimensional Change Card Sort Test. Crystallized intelligence was assessed using the Oral Reading and Picture Vocabulary tests. Age-adjusted scores were averaged for fluid and crystallized intelligence, respectively, which were then averaged to form a measure of intelligence[42].

Statistical Analysis

Analyses were conducted in R (3.4.2)[43]. Covariates were identical to those used in neuroimaging analyses. As in other analyses, variables were transformed to correct for skew, and were winsorized as needed. All variables were z-scored. Linear regression models were used to test whether extracted brain volume predicted behavioral and self-report outcomes. Analyses in the HCP used linear mixed-effects models[44], which controlled for family as a random effect, and sibling-status (i.e. DZ twin, MZ twin, half-sibling) was entered as fixed-effect covariates. FDR-correction for multiple comparisons was applied the entire set of analyses for each study.

Results

Extracted volume of the frontal and insula ROI, and volume of each significant cluster, were not significantly associated with any of the behavioral measures, in either sample, after correcting for multiple testing (**Supplemental Tables 11&12**). This suggests that negative urgency, impulsivity, or IQ do not fully mediate the association between brain volume and alcohol consumption. Given nominally significant results for some of these analyses within our DNS sample, it is possible that different behavioral mechanisms might represent risk at different ages and that these null results may represent type II error that larger samples might detect.

Supplemental References

1. Wickham, H. (Springer-Verlag New York: 2016).at <http://ggplot2.org>
2. Kassambara, A. (2017).at <https://cran.r-project.org/package=ggpubr>
3. Garnier, S. (2018).at <https://cran.r-project.org/package=viridis>
4. Slowikowski, K. (2017).at <https://cran.r-project.org/package=ggrepel>
5. Wechsler, H. et al. *J. Am. Coll. Heal.* **50**, 203–217 (2002).
6. Astin, A.W. *Rev. High. Educ.* **21**, 115–135 (1998).
7. Wilsnack, R.W., Wilsnack, S.C., Kristjanson, A.F., Vogeltanz-Holm, N.D. & Gmel, G. *Addiction* **104**, 1487–1500 (2009).
8. Galvan, F.H. & Caetano, R. *Alcohol Res. & Heal.* **27**, 87–95 (2003).
9. Lotfipour, S. et al. *BMC Emerg. Med.* **15**, 24 (2015).
10. Fromme, K., Wetherill, R.R. & Neal, D.J. *J. Am. Coll. Heal.* **59**, 21–27 (2010).
11. Tavolacci, M.P. et al. *BMC Public Health* **13**, 724 (2013).
12. Esper, L.H. & Furtado, E.F. *J Alcohol. Drug Depend* **1**, 2 (2013).
13. Baranger, D.A.A. et al. *Front. Psychol.* **7**, 1–10 (2016).
14. Dick, D.M. et al. *Addict. Biol.* **15**, 217–26 (2010).
15. Amlung, M., Vedelago, L., Acker, J., Balodis, I. & MacKillop, J. *Addiction* **112**, 51–62 (2017).
16. Pehlivanova, M. et al. *J. Neurosci.* 230888 (2018).doi:10.1101/230888
17. Holmes, A.J., Hollinshead, M.O., Roffman, J.L., Smoller, J.W. & Buckner, R.L. *J. Neurosci.* **36**, 4038–4049 (2016).
18. Mackey, S. et al. *Biol. Psychiatry* **82**, 275–282 (2017).
19. Churchwell, J.C. & Yurgelun-Todd, D.A. *Dev. Cogn. Neurosci.* **6**, 80–86 (2013).
20. Stamates, A.L. & Lau-Barraco, C. *Exp. Clin. Psychopharmacol.* **25**, 521–533 (2017).
21. Coskunpinar, A., Dir, A.L. & Cyders, M. a *Alcohol. Clin. Exp. Res.* **37**, 1441–50 (2013).
22. Labrie, J.W., Kenney, S.R., Napper, L.E. & Miller, K. *Addict. Behav.* (2013).doi:10.1016/j.addbeh.2013.09.018
23. Corral-Fríasa, N.S. et al. *Psychol. Med.* **45**, 1–13 (2015).
24. Muhlert, N. & Lawrence, A.D. *Neuroimage* **115**, 138–146 (2015).
25. Kubička, L., Matějček, Z., Dytrych, Z. & Roth, Z. *Addiction* **96**, 1615–1628 (2001).
26. Clarke, T.-K. et al. *Mol. Psychiatry* **22**, 1376–1384 (2017).
27. Toga, A.W. & Thompson, P.M. *Annu. Rev. Neurosci.* **28**, 1–23 (2005).

28. Noble, K.G. et al. *Nat. Neurosci.* **18**, 773–778 (2015).

29. Brouwer, R.M. et al. *Neuroimage* **100**, 676–683 (2014).

30. Green, L., Myerson, J. & Vanderveldt, A. *Wiley Blackwell Handb. Operant Class. Cond.* 307–337 (2014).doi:10.1002/9781118468135.ch13

31. Hariri, A. et al. *J. Neurosci.* **26**, 13213–13217 (2006).

32. Nikolova, Y.S., Knodt, a R., Radtke, S.R. & Hariri, a R. *Mol. Psychiatry* **21**, 1–9 (2015).

33. Estle, S.J., Green, L., Myerson, J. & Holt, D.D. *Mem. Cognit.* **34**, 914–28 (2006).

34. Patton, J.H., Stanford, M.S. & Barratt, E.S. *J. Clin. Psychol.* **51**, 768–774 (1995).

35. Achenbach, T. (University of Vermont Research Center of Children Youth & Families: 2009).

36. Pagliaccio, D. et al. *JAMA Psychiatry* **63130**, 1–8 (2015).

37. Costa, P.T. & McCrae, R.R. *Odessa FL Psychol. Assess. Resour.* **3**, 101 (1992).

38. Costa, P.T. & McCrae, R.R. *Odessa FL Psychol. Assess. Resour.* (1992).

39. Ullman, S.E., Filipas, H.H., Townsend, S.M. & Starzynski, L.L. *J. Stud. Alcohol* **66**, 610–619 (2005).

40. Wechsler, D. *Wechsler Abbreviated Scale Intell. Ed.* (NCS Pearson: San Antonio, Tx, 2011).doi:citeulike-article-id:6135820

41. Gershon, R.C. et al. *Neurology* **80**, S2–S6 (2013).

42. Casaletto, K.B. et al. *J. Int. Neuropsychol. Soc.* **22**, 364–374 (2016).

43. R Core Team (Vienna, Austria, 2014).at <http://www.r-project.org/>

44. Bates, D., Mächler, M., Bolker, B.M. & Walker, S.C. *J. Stat. Softw.* **67**, 1–48 (2015).