Genome-Wide Meta-Analysis of Longitudinal Alcohol Consumption Across Youth and Early Adulthood

Daniel E. Adkins,[1] Shaunna L. Clark,[1] William E. Copeland,[2] Martin Kennedy,[3] Kevin Conway,[4]
Adrian Angold,[2] Hermine Maes,[5] Youfang Liu,[6] Gaurav Kumar,[1] Alaattin Erkanli,[2] Ashwin A. Patkar,[2]
Judy Silberg,[5] Tyson H. Brown,[7] David M. Fergusson,[8] L. John Horwood,[8] Lindon Eaves,[5]
Edwin J. C. G. van den Oord,[1] Patrick F. Sullivan,[9] and E. J. Costello[2]

[1]*Center for Biomarker Research and Personalized Medicine, Virginia Commonwealth University, School of Pharmacy, Richmond, Virginia, USA*
[2]*Department of Psychiatry, Duke University Medical Center, Durham, North Carolina, USA*
[3]*Department of Pathology, University of Otago, Christchurch, New Zealand*
[4]*Division of Epidemiology, Services and Prevention Research, National Institute on Drug Abuse, Bethesda, Maryland, USA*
[5]*Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia, USA*
[6]*Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA*
[7]*Center for Medicine, Health and Society, Vanderbilt University, Nashville, Tennessee, USA*
[8]*Department of Psychological Medicine, University of Otago, Christchurch, New Zealand*
[9]*Departments of Genetics and Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA*

The public health burden of alcohol is unevenly distributed across the life course, with levels of use, abuse, and dependence increasing across adolescence and peaking in early adulthood. Here, we leverage this temporal patterning to search for common genetic variants predicting developmental trajectories of alcohol consumption. Comparable psychiatric evaluations measuring alcohol consumption were collected in three longitudinal community samples ($N = 2,126$, obs = 12,166). Consumption-repeated measurements spanning adolescence and early adulthood were analyzed using linear mixed models, estimating individual consumption trajectories, which were then tested for association with Illumina 660W-Quad genotype data (866,099 SNPs after imputation and QC). Association results were combined across samples using standard meta-analysis methods. Four meta-analysis associations satisfied our pre-determined genome-wide significance criterion (FDR < 0.1) and six others met our 'suggestive' criterion (FDR <0.2). Genome-wide significant associations were highly biological plausible, including associations within GABA transporter 1, SLC6A1 (solute carrier family 6, member 1), and exonic hits in LOC100129340 (mitofusin-1-like). Pathway analyses elaborated single marker results, indicating significant enriched associations to intuitive biological mechanisms, including neurotransmission, xenobiotic pharmacodynamics, and nuclear hormone receptors (NHR). These findings underscore the value of combining longitudinal behavioral data and genome-wide genotype information in order to study developmental patterns and improve statistical power in genomic studies.

∎ **Keywords:** alcohol, genome-wide, longitudinal, developmental trajectory, SLC6A1, GABA

Nearly 4% of all deaths worldwide are attributable to alcohol (WHO, 2011). In the United States, alcohol is the third leading preventable cause of death, after obesity and tobacco use (Mokdad et al., 2004). This public health burden is unevenly distributed across the life course with the highest levels of use, abuse, and dependence in early adulthood. It is well established that both alcohol use and alcohol use disorder rates increase markedly across adolescence and into the early twenties (Chen & Kandel, 1995; Harford et al., 2005), after which disorder rates decrease considerably during young adulthood (Harford et al., 2005). In the

2001 National Household Survey on Drug Abuse, 12-month prevalence rates for any alcohol disorders dropped 28% for

males and 47% for females from ages 18–23 years to ages 24–29 years (Harford et al., 2005). During early adulthood, alcohol is the leading cause of death worldwide (WHO, 2011). Clearly, adolescence and young adulthood are critical developmental stages for studying alcohol misuse and its genetic influences (Rutter et al., 1999).

Across a number of measures (e.g., drinking frequency, binge drinking, alcohol abuse, or dependence), alcohol involvement is clearly familial and moderately heritable ($h^2 \approx$ 40–60%; Kendler et al., 2008; Miles et al., 2005; Pagan et al., 2006; Silberg et al., 2003). This finding has led to extensive genome-wide linkage and candidate gene studies to identify specific susceptibility markers. Indeed, a number of risk variants have been identified (e.g., acetaldehyde and alcohol dehydrogenase genes) (Quertemont, 2004), but many others have failed to replicate these in independent samples. Genome-wide association studies (GWAS) offer comprehensive tests of common genetic variants and have proven successful in characterizing the genetic architecture of many complex diseases (The Wellcome Trust Case Control Consortium, 2007). To date, a small body of GWAS have studied problematic alcohol use and related endophenotypes and identified a number of putative risk loci (Bierut et al., 2010; Edenberg et al., 2010; Heath et al., 2011; Schumann et al., 2011; Treutlein et al., 2009). Implicated variants include polymorphisms in *PECR* (Johnson et al., 2006; Treutlein et al., 2009), a gene cluster on chromosome 11 (Edenberg et al., 2010), *PKNOX2* (Bierut et al., 2010), *KIA0040* (Bierut et al., 2010; Zuo et al., 2011), *HTR1* (Bierut et al., 2010; Zuo et al., 2011), and *AUTS2* (Schumann et al., 2011), among others. Particularly compelling evidence has been presented for rs6943555 in *AUTS2*, which was associated with alcohol consumption in a large meta-analysis ($N \approx$ 26K), and corroborated by *AUTS2* expression data in human prefrontal cortex and murine brain, as well as reduced alcohol sensitivity with *AUTS2* homolog downregulation in *Drosophila*. Such robust findings remain the exception to the rule; however, as most genetic variants implicated in GWAS studies have failed to replicate.

In response to the challenges of genomic studies of developmentally dynamic, difficult to measure substance abuse phenotypes, the NIDA-sponsored Genes, Environment and Development Initiative (GEDI) combines three longitudinal samples of psychiatric/behavioral repeated assessments with genome-wide data generated using a common platform (Illumina Human660W-Quad v1). The project offers several scientific advances. Foremost, harmonized data across three independent samples reduces the risk of study-specific artifacts, facilitating the search for cross-validated, replicated signals (Agrawal et al., 2012). Second, longitudinal data offer multiple statistical and substantive benefits (Adkins et al., 2010). Most fundamentally, statistical models of individual development are unidentifiable without multiple assessments per subject (Singer, 2003). Thus, GEDI's

structure allows genome-wide investigation of genetic risk for longitudinal patterns of alcohol abuse across its principle risk period — adolescence and young adulthood. Next, and crucially, when measuring any outcome with non-trivial measurement error (e.g., alcohol consumption), increasing the number of assessments/subject monotonically increases estimate precision, and consequently statistical power, with dramatic gains when assessments/subject are numerous, as in GEDI (Willett et al., 1998). Modeling all GEDI alcohol consumption assessments (>5 per subject) increases estimate precision >3-fold compared to using only first and last assessments (Willett et al., 1998). Finally, we employ the simple, often overlooked, strategy of improving statistical power by specifying a quantitative outcome — alcohol consumption. The power gains of quantitative versus dichotomous outcome specifications are well documented (Cohen, 1983), and it has been demonstrated that focusing solely on clinical diagnoses in psychiatric GWAS dramatically decreases power to detect effects (van der Sluis et al., 2013).

Cumulatively, these techniques help maximize power in the current study, which analyzes three longitudinal samples with repeated assessment of alcohol consumption across adolescence and young adulthood (total $N =$ 2,126, Obs = 12,166). Consumption-repeated measurements were analyzed using linear mixed models, estimating individual consumption trajectories (and a longitudinal mean consumption measure), which were then tested for association with Illumina 660W-Quad genome-wide data (866,099 SNPs after genotype imputation and QC). Association results were combined across samples using standard meta-analytic procedures (de Bakker et al., 2008). Top associations were examined for evidence of enriched association to known biological pathways.

Materials and Methods

Description of Individual Studies

Great smoky mountain study (GSMS). GSMS is a longitudinal, representative study of children in 11 predominantly rural counties in southeast United States begun in 1993 (Costello et al., 1996). Three cohorts of children, aged 9, 11, and 13 years, were recruited from a pool of ~20,000 children, resulting in $N =$ 1,420 participants (49% female). Annual assessments were completed with the child and primary caregiver until age 16 years and then with the participant again at ages 19, 21, and 24–26 years for 9,904 total assessments. These assessments used the Childhood and Adolescent Psychiatric Assessment (Angold & Costello, 2000) and its upward extension. Informed consent forms were completed for all aspects of data collection and the study protocol was approved by the Duke Institutional Review Board. An average of 82% of all possible interviews was completed across all waves, ranging from 75% to 94% at individual waves. Blood spots collected at each observation

TABLE 1

Characteristics of the Studies

	GSMS	CHDS	VTSABD
Representative population sample?	Y	Y	Y (twins)
Number of data waves	7–10	22	5
Age range from start to most recent wave	9–26	Birth–30	8–30
Total phenotyped subjects	1,415	1,265	2,588
Total phenotype repeated observations	9,437	6,921	5,837
Assessment instruments	CAPA; YAPA	DISC; RAPI; CIDI	CAPA; YAFU
DNA source	Dried Bloodspots	Whole Blood	Whole Blood
Subjects removed during quality control	94	22	17
Total number of subjects included in genome-wide association testing	784	739	603
Genotyping platform	Illumina 660W-Quad v1 DNA Analysis BeadChip	Illumina 660W-Quad v1 DNA Analysis BeadChip	Illumina 660W-Quad v1 DNA Analysis BeadChip

were used for DNA extraction performed at Rutgers University Cell and DNA Repository. A total of 784 GSMS subjects, contributing 5,766 repeated alcohol consumption assessments, were analyzed in genome-wide association testing. Study characteristics are summarized in Table 1.

Christchurch health and development study (CHDS).
The CHDS is a longitudinal study of a birth cohort of 1,265 children born in the Christchurch region of New Zealand in 1977 (Fergusson & Horwood, 2001). This cohort involved 97% of children born from April 15 to August 5, 1977 and has been assessed on 22 occasions to age 30 years. Data were gathered during face-to-face interviews with subjects and parents, supplemented by data from official records. Signed consent was obtained for all aspects of data collection and the study has been subject to ethical review throughout its history. The present analysis is based on data collected during assessments of the cohort in adolescence (ages 14, 15, 16, and 18 years) and adulthood (ages 21, 25, 30 years). The number of subjects assessed at these ages ranged from 953 to 1,025, representing between 76% and 82% of the surviving cohort at each age. Whole blood was collected from most subjects, but for ~9% of participants saliva samples were used instead (Oragene TM, DNA Genotek Inc., Ontario, Canada). DNA extractions were performed at the University of Otago, Gene Structure, and Function Lab. A total of 739 CHDS subjects, contributing 4,959 repeated alcohol consumption assessments, were analyzed in genome-wide association testing.

Virginia twin study on adolescent behavioral development (VTSABD). The VTSABD is an ongoing cohort-longitudinal study of twins born 1974–1983; 1,412 families were included in the first wave of data collection, with three subsequent waves of data collection occurring at approximately 1.5-year intervals, and a fifth wave when participants were in their mid-20s (Simonoff et al., 1997). A sample of 1,894 putative twin pairs was ascertained through the state school system and participating private schools in Virginia, and through families who contacted the VTSABD. A total of 1,412 families (2,775 children) participated (75%). Up to

age 18 years, children and parents completed the Childhood and Adolescent Psychiatric Assessment. Parents completed a similar assessment on both twins. After the age of 18 years, the twins alone were interviewed using an age-updated interview designed for telephone administration. Whole blood was collected from VTSABD subjects and DNA extractions were completed at Rutgers University Cell and DNA Repository. A total of 603 unrelated VTSABD subjects, contributing 1,441 repeated alcohol consumption assessments, were analyzed in genome-wide association testing.

Genotyping and Quality Control

All samples were genotyped using Illumina Human660W-Quad v1 DNA Analysis BeadChips at the Mayo Clinic. Quality control procedures were applied to each study using R, PLINK (Purcell et al., 2007) and EIGENSOFT (Patterson et al., 2006). In each study, single nucleotide polymorphisms (SNPs) with missing rate>0.01, minor allele frequency (MAF)<0.05, or extreme deviation ($p <$1E-6) from Hardy–Weinberg equilibrium were removed from further analysis. Subjects with missing rate >0.01 or unusual genome-wide homozygosity ($|$normalized homozygosity rate$|>5$) were excluded. Sex was investigated using no-call proportions of chr Y SNPs and heterozygosity proportions for chr X SNPs. Mislabeled sex information was corrected after double-checking the original data and subjects with unexplainable results were deleted. In addition, pairwise identical-by-descent (IBD) estimation was evaluated to identify unexpected duplicates and relative pairs. For VTSABD, one subject per twin pair, selected at random, was retained in the current analysis.

SNP dosages were imputed in each study using MaCH (Li et al., 2010). The imputation reference was HapMap3 CEU (Utah residents with Northern and Western European ancestry) for subjects of European ancestry. Unobserved population admixture due to ancestry is a well-known confound in GWAS (Patterson et al., 2006). To protect against false positives due to ancestry, we extracted five principal components from each sample to capture population stratification. To improve the efficiency of population stratification

principal components analysis (PCA), a subset of independent SNPs was selected using PLINK, with 77155–79517 SNPs analyzed in each study. PCA was applied to the selected SNPs using the 'smartpca' module of EIGENSOFT.

Data Analysis

Phenotype Modeling

Two longitudinal measures of over-time alcohol consumption were used as phenotypic outcomes. Both measures were derived from repeated assessments of average drinks per week. The first measure, as detailed below, was a developmental trajectory estimate of alcohol consumption (drinks per week) across adolescence and early adulthood (maximum age range: 8–30 years). Intuitively, this trajectory slope measures the rate at which alcohol consumption increases, starting with no consumption for virtually all subjects in childhood, then increasing at different rates (or remaining flat among non-drinkers) through adolescence and the transition to adulthood, before stabilizing at different levels in the late twenties. The second measure was the simple mean of all alcohol consumption (drinks per week) repeated assessments collected across adolescence and the transition to adulthood (maximum age range: 12–21 years) for each subject. We selected these two specifications as: (1) the trajectory outcome was derived from the best fitting longitudinal model of the several tested, and maximized power through including all repeated measures, while (2) the mean consumption outcome provided a simpler summary of individuals' drinking behavior, and thus greater continuity to existing literature (Agrawal et al., 2012; Grant et al., 2009). The mean consumption measure also has the benefit of focusing solely on adolescence and the transition to adulthood — a developmental period of non-normative drinking, which is associated with increased risk of concurrent comorbid psychiatric disorders and future substance abuse and dependence (Rutter et al., 1999).

To generate the developmental trajectory estimates, we applied a method to estimate longitudinal development from all available repeated observations using linear mixed models (van den Oord et al., 2009). This method first determines the optimal functional form of the developmental trajectory, then generates individual trajectory estimates based on the best-fitting model, using best linear unbiased predictors (BLUPs; Pinheiro & Bates, 2000). As this approach takes advantage of all available repeated assessments, it results in more precise estimates than traditional approaches that estimate development using two assessments (i.e., change scores; Willett et al., 1998). Further, as estimates are based on mixed model trajectory slopes, they are more robust to attrition than competing approaches (Laird & Ware, 1982). This method of estimating outcomes from longitudinal phenotypic assessments is well established in psychiatric GWAS (Adkins et al., 2011; Clark et al., 2012; McClay et al., 2011a).

In modeling the developmental trajectories, we tested several specifications of age-based longitudinal change, including linear and quadratic functions, as well as a series of piecewise plateau models specifying linear change until a given age and flat thereafter (Bollen & Curran, 2005). Model comparisons using likelihood ratio tests for nested models, and BIC/AIC for non-nested quadratic and piecewise models (Burnham & Anderson, 2004), indicated that a piecewise model with alcohol consumption stabilizing at approximately age 28 years best fit the data. Thus, the preferred model used to generate our trajectory slope outcome can be written as

$$Y_{ti} = \beta_{00} + \beta_{10} a_{ti} + u_{0i} + u_{1i} a_{ti} + \varepsilon_{ti},$$

where i and t are subject and assessment occasions, respectively; Y is the alcohol consumption for subject i at assessment t; β_{00} is the overall sample intercept; β_{10} is the mean age slope coefficient for age specification a, which is coded 0 for the earliest age observed in the sample, 1 for second earliest age, and so forth, until reaching the plateau age (~28 years), after which its value remains constant for subsequent ages; u_{0i} is the subject-specific deviation from that overall sample intercept; u_{1i} is the subject-specific deviation from the mean age slope; ε_{ti} is the residual for subject i at assessment t. Random effects u_0 and u_{1i} were specified orthogonal. The term of interest, u_{1i}, may more intuitively be described as the subject-specific alcohol consumption age trajectory. As subject-level parameters are not directly estimated by the mixed model, they were calculated as BLUPs in a post-estimation step (Pinheiro & Bates, 2000). See Supplementary Material 1 for further phenotypic modeling details.

Genome-Wide Association and Meta-Analyses

Genome-wide association testing was performed using PROBABEL (Aulchenko et al., 2010), GWAS software for use with probabilistically imputed genotype dosages. Tests were specified as linear regressions of additive SNP effects, controlling for the top five ancestral PCA dimensions as covariates. Meta-analysis procedures proposed by de Bakker et al. (2008) were used to aggregate results across the three studies. Specifically, aggregate β coefficients and associated standard errors were computed as follows:

$$\beta = \frac{\sum_i \left[\beta_i / (SE_i)^2 \right]}{\sum_i \left[1/(SE_i)^2 \right]}$$

$$SE = \sqrt{\frac{1}{\sum_i \left[1/(SE_i)^2 \right]}},$$

where i indexes study number. Given the use of a common genotyping platform (Illumina Human660W-Quad v1 DNA Analysis BeadChips) and genotyping lab (Mayo Clinic), many common pitfalls in GWAS meta-analysis were avoided, including strand orientation mismatches, imputation reference sample heterogeneity, variable QC criteria, and so forth.

Multiple Testing

We used a false discovery rate (FDR; Benjamini & Hochberg, 1995) approach to declare statistical significance. In comparison to controlling a family-wise error rate (e.g., Bonferroni correction), FDR: (a) provides a better balance between finding true effects versus controlling false discoveries, (b) results in comparable standards for declaring significance across studies because it does not directly depend on the number of tests, and (c) is relatively robust against having correlated tests (Brown & Russell, 1997). FDR is commonly used in high-dimensional applications, including GWAS (Beecham et al., 2009; Dubois et al., 2010; McClay et al., 2011b). We set an FDR threshold of 0.1 for declaring genome-wide significance. This specifies that, on average, 10% of the SNPs declared significant are expected to be false discoveries. Additionally, we discuss suggestive associations at an FDR threshold of 0.2 to reduce the probability of Type II statistical errors, while explicitly noting reduced confidence in these associations. Operationally, FDR was controlled using q-values, which are FDRs calculated using the p-value of the markers as thresholds for declaring significance (Storey & Tibshirani, 2003).

Pathway Analyses

Beyond our primary meta-analysis of single marker GWAS findings, we also conducted secondary analyses to determine whether any known biological pathways harbored an excess of SNPs with small p-values. Specifically, to interrogate the data for evidence of enriched association to known biological pathways, we conducted pathway analysis using ConsensusPathDB (http://cpdb.molgen.mpg.de/), a human-centric meta-database of functional biological data, compiled from 30 separate public sources of biological interactions (Kamburov et al., 2011). All alcohol consumption-associated SNPs with meta-analysis $p < .01$ were matched to the closest gene ±25 kb using RefSeq (GRCh37) coordinate information. All implicated genes were assembled into a final non-redundant gene list, comprising 2,590 (mean consumption) and 2,506 (developmental trajectory) genes. For each of the 4,601 reference pathways present in ConsensusPathDB, a hypergeometric test was performed to calculate the significance of the overlap between the genes from our putative susceptibility list and those present in each reference pathway. FDR was implemented to adjust significance for the large number of tests performed.

Results

Single Marker GWAS

Table 2 provides details on those SNPs that were genome-wide significant/suggestive ($q < 0.1/q < 0.2$). Figure 1A shows Q–Q plots for meta-analyses. The plots show that the p-values from the GWAS are generally on a straight line, indicating that the p-value distribution generally conforms

TABLE 2
Strongest SNP-Longitudinal Alcohol Consumption Associations (q < 0.2)

| RS #	Outcome	Meta-analysis						GSMS			CHDS			VTSABD			Gene	Gene name
		Q	Z	p	MAF	CHR	MA	MAF	N	p	MAF	N	p	MAF	N	p		
rs7031417	Mean	0.01	-5.68	1.4E-08	0.07	9	C	0.07	782	1.0E-04	0.07	738	9.5E-04	0.06	531	1.2E-02	10 kb	LOC100129340
rs17053864	Mean	0.05	-5.25	1.5E-07	0.09	9	G	0.10	782	7.2E-05	0.09	738	2.4E-03	0.09	531	6.1E-02	yes	LOC100129340
rs11710497	Mean	0.05	-5.29	1.2E-07	0.11	3	A	0.10	782	7.2E-04	0.12	738	1.8E-02	0.12	531	5.3E-04	yes	SLC6A1
rs7019589	Mean	0.06	-5.17	2.3E-07	0.09	9	T	0.10	782	6.5E-05	0.09	738	3.3E-03	0.09	531	7.1E-02	yes	LOC100129340
rs12257178	Trajectory	0.14	-5.04	4.8E-07	0.39	10	A	0.37	784	5.9E-04	0.39	739	7.8E-05	0.40	575	2.8E-01	230 kb	ADRA2A
rs694424	Trajectory	0.14	-5.01	5.4E-07	0.12	11	G	0.13	784	1.9E-01	0.12	739	4.2E-02	0.11	575	1.8E-08	yes	IGSF9B
rs1984450	Trajectory	0.14	-5.15	2.6E-07	0.10	19	A	0.09	784	1.8E-03	0.11	739	4.4E-03	0.12	575	3.4E-03	yes	ZNF578
rs7253326	Trajectory	0.14	-5.14	2.8E-07	0.10	19	G	0.09	784	1.8E-03	0.11	739	4.6E-03	0.12	575	3.4E-03	yes	ZNF578
rs6778281	Mean	0.15	-4.96	7.1E-07	0.08	3	C	0.07	782	1.1E-03	0.08	738	3.2E-01	0.08	531	5.6E-06	yes	SLC6A1
rs4898641	Mean	0.19	-4.88	1.1E-06	0.06	14	G	0.08	782	8.8E-04	0.06	738	8.5E-04	0.04	531	1.1E-01	yes	MIPOL1
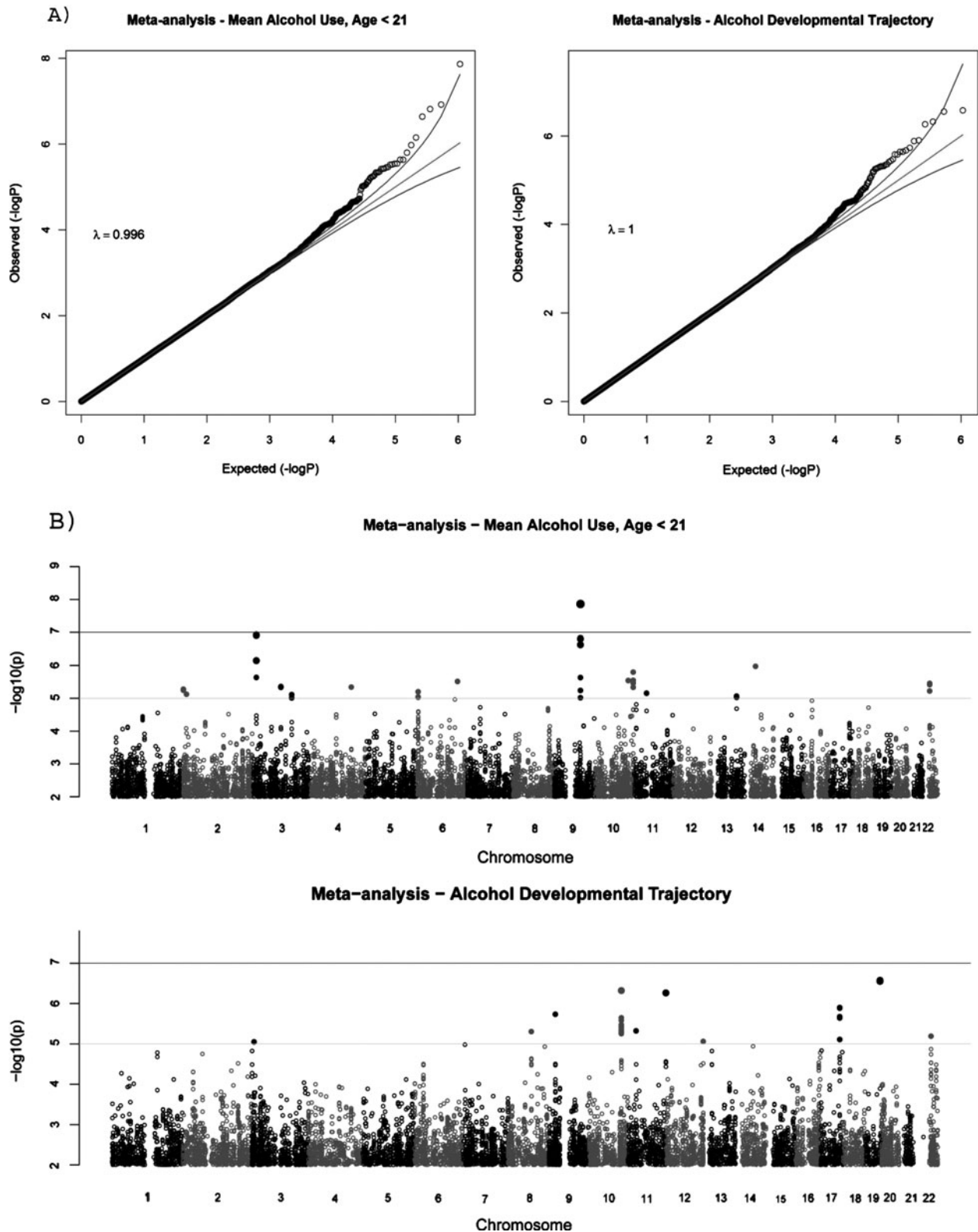
FIGURE 1

(A) QQ plots and (B) Manhattan plots for GWAS results of two longitudinal alcohol consumption measures.

to the expected null distribution assuming no effects of the markers. However, in both plots, there is also evidence that markers in the right upper corner have *p*-values smaller than would be expected under the null hypothesis, suggesting true association between these markers and the outcome. The plots also display λ values (i.e., ratio of median observed *p*-value to the median expected *p*-value under the null hypothesis) approximately equal to 1 (λ = 0.998 and 1.000), indicating no systematic test statistic inflation and adequate control of population stratification (see Supplementary Material 2 for study-specific Q–Q plots). Manhattan plots are provided in Figure 1B, and provide a low-resolution visualization of spatial clustering among association signals. The full set of GWAS *p*-values considered in this study is available for download at: http://www.pharmacy.vcu.edu/biomarker/resources/supplementary/.

The top significant finding involves three genome-wide significant SNPs (rs7031417: $q = 0.01$, $p = 1.4E-08$; rs17053864: $q = 0.05$, $p = 1.5E-07$; rs7019589: $q = 0.06$, $p = 2.3E-07$) at pseudogene *LOC100129340* (mitofusin-1-like) associated with mean adolescent alcohol consumption. These SNPs show moderately low MAF (~0.09) with minimal MAF differences across the three samples (Table 2). Study specific results show that this meta-analysis association was not driven by a single sample, with GSMS and CHDS exhibiting comparably low *p*-values (Table 2). Examination of the regional plot for the locus shows a robustly associated haplotype spanning ~100 kb, with >10 spatially clustered SNPs showing nominal association (Supplementary Material 3A). Notably, one of the genome-wide significant hits is located in the transcribed exon 4.

The second genome-wide significant association, also to mean consumption, is a locus in GABA transporter *SLC6A1* (solute carrier family 6 (neurotransmitter transporter, GABA), member 1). In addition to the genome-wide hit in this gene (rs11710497: $q = 0.05$, $p = 1.2E-07$), a second SNP met our suggestive criteria (rs6778281: $q = 0.15$, $p = 7.1E-07$). Again, MAFs were moderately low (~0.10) and highly comparable across individual samples. Also, significance was not driven by a single study, with comparably low *p*-values for GSMS and VTSABD (Table 2). Inspection of the regional plot indicates a fairly small LD block overlaying exonic regions of *SLC6A1*. Further, this small linkage disequilibrium (LD) block is highly enriched for strong associations including several nominally associated coding SNPs ($p < .001$; Supplementary Material 3B).

Top findings for the developmental consumption trajectory included a cluster of correlated SNPs about 150 kb upstream of *ADRA2A* (adrenoceptor alpha 2A), with rs12257178 exhibiting the strongest association ($q = 0.14$, $p = 4.8E-07$). MAFs were high (~0.4) and homogenous across studies. Association strength was primarily driven by GSMS and CHDS samples (Table 2). While it is questionable whether this putative susceptibility locus is in LD with *ADRA2A*, the regional plot does show a correlated SNP located ~80 kb upstream of the gene (Supplementary Material 3C). Also, further inspection of the regional plot indicates an unusually wide area of moderately enriched association surrounding the suggestive locus, spanning ~500 kb and encompassing *ADRA2A*. Thus, it is possible that this association may tag *cis* regulatory elements or genic variants for *ADRA2A*.

Other top hits included a pair of SNPs in high LD ($R^2 > 0.9$) within *ZNF578* (rs1984450 and rs7253326: $q = 0.14$; $p \approx 2.7E-07$) associated with the trajectory outcome, and an interesting spatial pattern of SNPS associated with mean consumption surrounding an intronic SNP in *MIPOL1* (rs4898641: $q = 0.19$, $p = 1.1E-06$). The *ZNF578* finding exhibits almost exactly the same level of association across studies ($p \approx 3.0E-03$) and moderately low MAF (~0.10). Correlated signals overlapped *ZNF578* and the adjacent *ZNF808* and *ZNF701* genes, and tag two nominally significant coding SNPs in *ZNF578* (Supplementary Material 3D). Next, rs4898641 in *MIPOL1* appears to tag a ~200 kb haplotype, enriched for association, which closely corresponds to the gene boundaries of *MIPOL1* (Supplementary Material 3E). Our final suggestive finding involved a genic locus in *IGSF9B* centered on rs694424 ($q = 0.14$, $p = 5.4E-07$; Supplementary Material 3F). This finding is of lower confidence, however, as its association strength is essentially driven by a single sample (VTSABD).

Pathway Analyses

Results from ConsensusPathDB indicated significant involvement of several intuitive biological pathways (Table 3). As ConsensusPathDB is a meta-database bioinformatic tool, it mines multiple individual databases — in the current analysis significant ($q < 0.1$) pathways were found in Reactome, PharmGKB, Signlink, and Wikipathways databases (Kamburov et al., 2011). The top pathways, 'neuronal system' ($p = 2.5E-07$, $q = 4.5E-04$) and the related 'transmission across chemical synapses' ($p = 3.7E-05$, $q = 3.3E-02$), indicated an association between genetic variation in neurotransmission and individual differences in mean adolescent alcohol consumption, with disproportionate representation of GABAergic and glutamatergic genes in both of these pathways. Other significant pathways for mean adolescent consumption including 'NHR' ($p = 2.6E-04$, $q = 9.4E-02$), 'celecoxib pharmacodynamics', and 'adherens junctions interactions' ($p = 2.1E-04$, $q = 9.4E-02$; Table 3). The three significant pathways for the developmental consumption trajectory indicated more general developmental themes, with specific findings including 'developmental biology' ($p = 7.6E-05$, $q = 8.1E-02$), 'G protein signaling pathways' ($p = 1.1E-04$, $q = 8.1E-02$) and 'axon guidance' ($p = 1.3E-04$, $q = 8.1E-02$).

TABLE 3

ConsensusPathDB Pathway Results for All SNPs Within Genes Nominally Associated ($p < .01$) With Longitudinal Alcohol Consumption

Pathway	p	q	Database	Associated genes from GWAS in pathway
Mean consumption				
Neuronal system	2.5E-07	4.5E-04	Reactome	SLC6A1; KCNC1; KCNC2; GABRG3; GABRG2; HSPA8; KCNG4; GRIK5; GLUL; SLC5A7; GABRB3; AP2B1; GRIP2; CALM2; CALM3; KCNG1; CACNB2; CACNA1B; CHRNB2; KCNMA1; PRKCB; GAD1; GRIA2; KCNS3; ABCC8; GABBR2; RIMS1; PLCB1; KCNJ11; GRIA1; VAMP1; KCNQ5; KCNAB1; RPS6KA2; KCNQ1; CHRNA4; KCNH8; KCNN1; CACNG4; CACNA2D3; CACNA2D1; KCNG3; KCNJ6; ADCY4; ACTN2; KCNH7; KCNMB2; KCNH1; KCNMB4; RASGRF2; GRIK1; HCN4; GJD2; GRIN2B; GNG7; CACNA1E; GNG2; SLC1A2; GRIP1; CACNG3; SYN2; SYN3; GNG8
Transmission across chemical synapses	3.7E-05	3.3E-02	Reactome	SLC6A1; GABRG3; GABRG2; HSPA8; GLUL; SLC5A7; GABBR2; GAD1; GABRB3; CALM2; CALM3; CACNB2; CACNA1B; CHRNB2; CACNA1E; AP2B1; GRIA2; RIMS1; PLCB1; GRIA1; PRKCB; CHRNA4; CACNA2D3; CACNA2D1; ADCY4; ACTN2; RPS6KA2; GRIP2; KCNJ6; RASGRF2; SYN2; GRIK5; GRIN2B; GNG7; GNG2; SLC1A2; GRIK1; GRIP1; CACNG3; CACNG4; SYN3; GNG8
Celecoxib pathway, pharmacodynamics	1.9E-04	9.4E-02	PharmGKB	PDK1; CACNA2D1; TBXAS1; PTGIR; CACNA1I; CACNB2; CACNA1B; KCNQ5; PTGER1; CACNA1E; KCNQ1; CTNNB1; CACNA1C; PTGS2; CACNA2D3; PLA2G4A; CACNA1S
Adherens junctions interactions	2.1E-04	9.4E-02	Reactome	CDH11; CDH10; CDH13; CDH12; CTNNB1; PVRL2; CDH2; CDH6; CADM2; CDH4; CADM1
Nuclear hormone receptors	2.6E-04	9.4E-02	Signalink	NR3C2; ESRRG; RARG; ESR1; HNF4A; HNF4G; NR4A1; NR2F1; RXRA; ESRRB; THRB; NR2C1; RORA; SMAD3; NR5A2
Developmental trajectory				
Developmental biology	7.6E-05	8.1E-02	Reactome	ITGA9; ROBO2; CDH15; NRP2; SCN5A; NRCAM; ITGA1; HRAS; DOCK1; PCSK6; TRIO; CTNNB1; MAP2K1; CSNK2A1P; MED6; DNM3; CHL1; STIP1; CACNA1G; PITPNA; NCAM1; SPTBN1; FURIN; NCOA3; PRNP; BOC; CACNB2; CACNA1C; CACNA1D; TLN1; MED15; SLIT1; SCN4A; CDH2; CDH4; EPHB2; LAMA1; ACVR2B; RAC2; SIAH1; CRMP1; UNC5A; SCN8A; COL9A1; PPARG; CLTA; COL5A1; ABLIM1; FES; ABLIM2; CTNNA2; KIAA1598; SEMA5A; ROBO1; PLXNA2; VAV2; ST8SIA2; ROCK2; PLXNA4; MEF2C; CDON; ANK3; NTN1; CNTN6; PAK7; SCN3B; MYO10
G protein signaling pathways	1.1E-04	8.1E-02	Wikipathways	GNAO1; HRAS; PDE1C; PRKCA; PRKCH; GNB5; PRKCE; GNB2; PRKCZ; PRKAR1A; ITPR1; AKAP1; KCNJ3; AKAP5; AKAP6; AKAP7; ADCY1; AKAP2; ADCY9; GNA14; GNGT2; PPP3CA; PDE4B; PDE4D
Axon guidance	1.3E-04	8.1E-02	Reactome	ITGA9; ROBO2; SCN5A; NRCAM; ITGA1; HRAS; DOCK1; MAP2K1; CSNK2A1P; CNTN6; DNM3; CHL1; STIP1; TLN1; PITPNA; NCAM1; SPTBN1; CACNB2; CACNA1C; PRNP; CACNA1G; SLIT1; SCN4A; EPHB2; LAMA1; TRIO; RAC2; SIAH1; UNC5A; SCN8A; CACNA1D; CRMP1; COL5A1; ABLIM1; FES; ABLIM2; COL9A1; NRP2; KIAA1598; SEMA5A; ROBO1; NTN1; VAV2; ST8SIA2; ROCK2; PLXNA4; ANK3; PLXNA2; CLTA; PAK7; SCN3B; MYO10

Discussion

Developmental perspectives have long stressed the importance of accounting for temporality and life course variation in models of substance abuse. A primary insight of such perspectives is that the importance of various risk factors fluctuates across developmental trajectories (Adkins et al., 2012; Willett et al., 1998). The current study endeavors to wed this perspective to genomic approaches to alcohol consumption. While psychiatric genetics has made advances toward elucidating the link between genetic variation and alcohol involvement, this research has largely been atemporal. The weakness of such static perspectives on the genetic determinants of alcohol involvement is highlighted not only by developmental perspectives, but also by research within genetics showing that epigenetic mechanisms (i.e., methylation, histone modification, and chromatin remodeling) regulate gene expression in response to developmental and environmental cues (Whitelaw & Whitelaw, 2006). Using the GEDI GWAS samples, this longitudinal GWAS meta-analysis has addressed the issue of developmental variation in genetic influences across adolescence and young adulthood through genome-wide testing of common variant effects on longitudinal measures of alcohol consumption, employing FDR methods to control the risk of false discoveries.

Our top genome-wide finding involved three SNPs (rs7031417, rs17053864, rs7019589) at pseudogene *LOC100129340* (mitofusin-1-like). Very little is known regarding this pseudogene, which was only added in the most recent, hg19, gene track (thus, its absence in the regional plot). However, variants within *LOC100129340* were found to be associated with adult human height in a large GWAS meta-analysis (Gudbjartsson et al., 2008). Also, one of the genome-wide significant hits fell within exon 4 of the pseudogene. Given that pseudogenes, by definition, lack protein-coding function, the prior probability of this locus being a true discovery is somewhat diminished. However, considering (a) the association signal is in a transcribed sequence, (b)

pseudogene transcripts have demonstrated trans-regulation of homologous coding genes (Hirotsune et al., 2003), and (c) endogenous siRNA are sometimes derived from pseudogene transcription (Tam et al., 2008), it would be rash to dismiss this result as a false discovery.

Our second genome-wide significant locus was tagged by rs11710497, located within GABA transporter 1—*SLC6A1* (solute carrier family 6 (neurotransmitter transporter, GABA), member 1). Given that alcohol functions primarily through binding to GABA receptors, and that this gene encodes protein GAT-1, which removes GABA from the synaptic cleft, there is a notably high prior probability of this being a true discovery. Furthermore, while the two genome-wide significant/suggestive hits were in intronic regions, they tag strongly associated exonic variants in *SLC6A1*. Although the network biology of the gene suggests obvious links to alcohol consumption, there has been surprisingly little research studying this relationship. Conversely, studies have implicated the gene in anxiety disorders and ADHD (e.g., Lasky-Su et al., 2008; Thoeringer et al., 2009; 2010). This finding should focus future research on the study of *SLC6A1* variants in alcohol abuse and other psychiatric phenotypes related to cortical excitability/inhibition.

The strong association of a large region near *ADRA2A* also provides a promising potential biological mechanism. Alpha-2-adrenergic receptors are members of the G protein-coupled receptor superfamily. These receptors have a critical role in regulating neurotransmitter release from sympathetic nerves and from adrenergic neurons in the CNS. Alpha2A and alpha2C subtypes are required for normal presynaptic control of transmitter release from sympathetic nerves in central noradrenergic neurons, and the alpha2A subtype inhibits transmitter release at high neural stimulation frequencies. This gene encodes alpha2A subtype and it contains no introns in either its coding or untranslated sequences. There is an extensive literature on the gene and many findings suggesting association to ADHD (e.g., Kim et al., 2010; Roman et al., 2003; 2005; Waldman et al., 2006), as well as metabolic traits (e.g., glycemic phenotypes, blood pressure; for meta-analysis, see Talmud et al., 2011), and substance abuse (Clarke et al., 2012; Merenakk et al., 2011; Prestes et al., 2007). Given the neurobiological implications, this gene should be considered a strong candidate in future alcohol abuse studies. The final three genome-wide suggestive ($q < 0.2$) loci were for intronic SNPs in *ZNF578*, *MIPOL1,* and *IGSF9B* and have no reported evidence linking them to alcohol use/abuse phenotypes. Thus, future research will be required to adjudicate whether they are novel susceptibility loci or false discoveries.

Pathway analyses indicated several intuitive biological mechanisms. Top results for mean adolescent consumption ('neuronal system' and 'transmission across chemical synapses') suggested extensive involvement of GABAergic (e.g., *SLC6A1, GABRG3, GABRG2, GABBR2*) and glutamatergic (e.g., *GRIK5, GLUL, GRIK1, GRIN2B*) neurotransmission. Given that GABA and glutamate neurotransmitter systems are primary pharmacodynamic targets of alcohol (for review, see Vengeliene et al., 2008), these results are consistent with extant knowledge of the drug's biological mechanisms. 'Celecoxib pharmacodynamics' and 'NHRs' also offer intriguing biological mechanisms for genetic variation in alcohol metabolism. Celecoxib acts by inhibiting prostaglandin synthesis (Penning et al., 1997). Prostaglandins are lipid autocrine/paracrine mediators with a variety of physiological effects, including inflammation modulation (Hata & Breyer, 2004), which have been shown to moderate alcohol 'hangover' intensity (Wiese et al., 2000) and reduce alcohol-induced liver toxicity (Nanji et al., 1993). 'NHR, are intracellular proteins responsible for sensing steroid and thyroid hormones and some other signaling molecules. Multiple NHR genes implicated in the current analysis are specifically expressed in liver (e.g., *HNF4A, HNF4G*) and several others function as thyroid hormone receptors, which have been linked to severity of alcohol craving, consumption, and withdrawal (Alfos et al., 1996; Ozsoy et al., 2006).

The three biological pathways significantly associated with the developmental consumption trajectory were broader and more explicitly developmentally oriented, particularly 'developmental biology' and 'axon guidance', which shared several implicated genes related to nervous system development (e.g., *ROBO1, ROBO2, ANK3*). The final significant pathway associated with the developmental trajectory, G-protein signaling, functions as an intermediary messenger system for many hormones, neurotransmitters, and other signaling molecules. While G-protein signaling is involved in many biological processes, it is known that alcohol intoxication is directly connected to the actions of four G-protein gated inwardly-rectifying potassium (GIRK) channel subunits, one of which was among the genes implicated in this pathway analysis (i.e., *KCNJ3*; Aryal et al., 2009; Lewohl et al., 1999). Despite the general plausibility of these pathway results, it is worth noting that given the incomplete, low-resolution nature of current biological databases, findings should be regarded as tentative pending future tests of replication as database content and accuracy continue to improve (Khatri et al., 2012).

Clearly, it is premature to suggest direct clinical applications of these findings for therapeutic target or biomarker discovery. On the contrary, actualizing the promise of translational genomics and converting academic findings into clinical applications will require a cumulative process of aggregating and jointly considering large bodies of evidence using meta-analytic and data integration techniques. Thus, it is important to rapidly disseminate genome-wide association results from longitudinal behavioral studies with genomic data, such as the GEDI samples. To facilitate this process, we provide all *p*-values

(http://www.pharmacy.vcu.edu/biomarker/resources/supplementary/) as a resource for investigators with the requisite samples to carry out replication or further meta-analysis.

As with any genetic associations, our findings will require additional replication and functional validation. However, the present study shows the potential of GWAS meta-analyses of longitudinal data to discover genes and pathways that mediate developmental trends in alcohol consumption. A better understanding of these biological mechanisms and the roles of specific polymorphisms may facilitate the development of improved biomarker-based approaches to personalize substance abuse treatment. It is hoped that this research will eventually contribute to reducing the global health burden of substance abuse by facilitating identification of novel mechanisms, biomarkers, and therapeutic targets.

Acknowledgments

Supplementary Material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/thg.2015.36

References

Adkins, D. E., Aberg, K., McClay, J. L., Bukszar, J., Zhao, Z., Jia, P., . . . van den Oord, E. J. (2011). Genomewide pharmacogenomic study of metabolic side effects to antipsychotic drugs. *Molecular Psychiatry*, 16, 321–332.

Adkins, D. E., Aberg, K., McClay, J. L., Hettema, J. M., Kornstein, S. G., Bukszar, J., . . . van den Oord, E. J. (2010). A genomewide association study of citalopram response in major depressive disorder —- A psychometric approach. *Biological Psychiatry*, 68, e25–27.

Adkins, D. E., Daw, J. K., McClay, J. L., & Van den Oord, E. (2012). The influence of five monoamine genes on trajectories of depressive symptoms across adolescence and young adulthood. *Development and Psychopathology*, 24, 267–285.

Agrawal, A., Freedman, N. D., Cheng, Y. C., Lin, P., Shaffer, J. R., Sun, Q., . . . Beirut, L., for the GENEVA Consortium. (2012). Measuring alcohol consumption for genomic meta-analyses of alcohol intake: Opportunities and challenges. *American Journal of Clinical Nutrition*, 95, 539–547.

Alfos, S., Higueret, P., Pallet, V., Higueret, D., Garcin, H., & Jaffard, R. (1996). Chronic ethanol consumption increases the amount of mRNA for retinoic acid and triiodothyronine receptors in mouse brain. *Neuroscience Letters*, 206, 73–76.

Angold, A., & Costello, E. (2000). The child and adolescent psychiatric assessment (CAPA). *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 39–48.

Aryal, P., Dvir, H., Choe, S., & Slesinger, P. A. (2009). A discrete alcohol pocket involved in GIRK channel activation. *Nature Neuroscience*, 12, 988–995.

Aulchenko, Y. S., Struchalin, M. V., & van Duijn, C. M. (2010). ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics*, 11, 134.

Beecham, G. W., Martin, E. R., Li, Y. J., Slifer, M. A., Gilbert, J. R., Haines, J. L., . . . Pericak-Vance, M. A. (2009). Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. *American Journal of Human Genetics*, 84, 35–43.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300.

Bierut, L. J., Agrawal, A., Bucholz, K. K., Doheny, K. F., Laurie, C., Pugh, E., . . . Rice, J. P. & as part of the gene, environment association studies (GENEVA) Consortium (2010). A genome-wide association study of alcohol dependence. *Proceedings of the National Academy of Sciences*, 107, 5082–5087.

Bollen, K. A., & Curran, P. J. (2005). *Latent curve models: A structural equation perspective*. Hoboken, NJ: John Wiley & Sons.

Brown, B. W., & Russell, K. (1997). Methods correcting for multiple testing: Operating characteristics. *Statistics in Medicine*, 16, 2511–2528.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference —- Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261–304.

Chen, K., & Kandel, D. B. (1995). The natural history of drug use from adolescence to the mid-thirties in a general population sample. *American Journal of Public Health*, 85, 41–47.

Clark, S. L., Adkins, D. E., Aberg, K., Hettema, J. M., McClay, J. L., Souza, R. P., . . . van den Oord, E. (2012). Pharmacogenomic study of side-effects for antidepressant treatment options in STAR*D. *Psychological Medicine*, 42, 1151–1162.

Clarke, T. K., Dempster, E., Docherty, S. J., Desrivieres, S., Lourdsamy, A., Wodarz, N., . . . Schumann, G. (2012). Multiple polymorphisms in genes of the adrenergic stress system confer vulnerability to alcohol abuse. *Addiction Biology*, *17*, 202–208.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, *7*, 249–253.

Costello, E. J., Angold, A., Burns, B., Stangl, D., Tweed, D., Erkanli, A., . . . Worthman, C. (1996). The Great Smoky Mountains Study of Youth: Goals, designs, methods, and the prevalence of DSM-III-R disorders. *Archives of General Psychiatry*, *53*, 1129–1136.

de Bakker, P. I. W., Ferreira, M. A. R., Jia, X., Neale, B. M., Raychaudhuri, S., & Voight, B. F. (2008). Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Human Molecular Genetics*, *17*, R122–R128.

Dubois, P. C. A., Trynka, G., Franke, L., Hunt, K. A., Romanos, J., Curtotti, A., . . . van Heel, D. A. (2010). Multiple common variants for celiac disease influencing immune gene expression. *Nature Genetics*, *42*, 295–302.

Edenberg, H. J., Koller, D. L., Xuei, X., Wetherill, L., McClintick, J. N., Almasy, L., . . . Foroud, T. (2010). Genome-wide association study of alcohol dependence implicates a region on chromosome 11. *Alcoholism: Clinical and Experimental Research*, *34*, 840–852.

Fergusson, D., & Horwood, L. (2001). The Christchurch health and development study: Review of findings on child and adolescent mental health. *Australian and New Zealand Journal of Psychiatry*, *35*, 287–296.

Grant, J. D., Agrawal, A., Bucholz, K. K., Madden, P. A. F., Pergadia, M. L., Nelson, E. C., . . . Heath, A. C. (2009). Alcohol consumption indices of genetic risk for alcohol dependence. *Biological Psychiatry*, *66*, 795–800.

Gudbjartsson, D. F., Walters, G. B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V., Zusmanovich, P., . . . Stefansson, K. (2008). Many sequence variants affecting diversity of adult human height. *Nature Genetics*, *40*, 609–615.

Harford, T. C., Grant, B. F., Yi, H. Y., & Chen, C. M. (2005). Patterns of DSM-IV alcohol abuse and dependence criteria among adolescents and adults: Results from the 2001 National Household Survey on Drug Abuse. *Alcoholism Clinical and Experimental Research*, *29*, 810–828.

Hata, A. N., & Breyer, R. M. (2004). Pharmacology and signaling of prostaglandin receptors: Multiple roles in inflammation and immune modulation. *Pharmacology & Therapeutics*, *103*, 147–166.

Heath, A. C., Whitfield, J. B., Martin, N. G., Pergadia, M. L., Goate, A. M., Lind, P. A., . . . Montgomery, G. W. (2011). A quantitative-trait genome-wide association study of alcoholism risk in the community: Findings and implications. *Biological Psychiatry*, *70*, 513–518.

Hirotsune, S., Yoshida, N., Chen, A., Garrett, L., Sugiyama, F., Takahashi, S., . . . Yoshiki, A. (2003). An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature*, *423*, 91–96.

Johnson, C., Drgon, T., Liu, Q.-R., Walther, D., Edenberg, H., Rice, J., Foroud, T., . . . Uhl, G. R. (2006). Pooled association genome scanning for alcohol dependence using 104,268 SNPs: Validation and use to identify alcoholism vulnerability loci in unrelated individuals from the collaborative study on the genetics of alcoholism. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *141B*, 844–853.

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H., & Herwig, R. (2011). ConsensusPathDB: Toward a more complete picture of cell biology. *Nucleic Acids Research*, *39*, D712–D717.

Kendler, K. S., Schmitt, E., Aggen, S. H., & Prescott, C. A. (2008). Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Archives of General Psychiatry*, *65*, 674–682.

Khatri, P., Sirota, M., Butte, A. J. (2012). Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges. *PLoS Computational Biology*, *8*, e1002375.

Kim, B. N., Kim, J. W., Kang, H., Cho, S. C., Shin, M. S., Yoo, H. J., Hong, S. B., . . . Lee, D. S. (2010). Regional differences in cerebral perfusion associated with the alpha-2A adrenergic receptor genotypes in attention deficit hyperactivity disorder. *Journal of Psychiatry & Neuroscience*, *35*, 330–336.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*, 963–974.

Lasky-Su, J., Neale, B. M., Franke, B., Anney, R. J., Zhou, K., Maller, J. B., . . . Faraone, S. V. (2008). Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *147B*, 1345–1354.

Lewohl, J. M., Wilson, W. R., Mayfield, R. D., Brozowski, S. J., Morriset, R. A., & Harris, R. A. (1999). G-protein-coupled inwardly rectifying potassium channels are targets of alcohol action. *Nature Neuroscience*, *2*, 1084–1090.

Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010). MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology*, *34*, 816–834.

McClay, J. L., Adkins, D. E., Aberg, K., Bukszar, J., Khachane, A. N., Keefe, R. S. E., . . . van den Oord, E. (2011a). Genome-wide pharmacogenomic study of neurocognition as an indicator of antipsychotic treatment response in schizophrenia. *Neuropsychopharmacology*, *36*, 616–626.

McClay, J. L., Adkins, D. E., Aberg, K., Stroup, S., Perkins, D. O., Vladimirov, V. I., . . . van den Oord, E. (2011b). Genome-wide pharmacogenomic analysis of response to treatment with antipsychotics. *Molecular Psychiatry*, *16*, 76–85.

Merenakk, L., Maestu, J., Nordquist, N., Parik, J., Oreland, L., Loit, H. M., . . . Harro, J. (2011). Effects of the serotonin transporter (5-HTTLPR) and alpha(2A)-adrenoceptor (C-1291G) genotypes on substance use in children and adolescents: A longitudinal study. *Psychopharmacology*, *215*, 13–22.

Miles, D., Silberg, J., Pickens, R., & Eaves, L. (2005). Familial influences for alcohol use in adolescent female twins:

Genetic and environmental interactions. *Journal of Studies on Alcohol*, 66, 445–451.

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the United States, 2000. *JAMA*, 291, 1238–1245.

Nanji, A. A., Khettry, U., Sadrzadeh, S. M. H., & Yamanaka, T. (1993). Severity of liver-injury in experimental alcoholic liver-disease —- Correlation with plasma endotoxin, prostaglandin-E2, leukotriene-B4, and thromboxane-B2. *American Journal of Pathology*, 142, 367–373.

Ozsoy, S., Esel, E., Izgi, H. B., & Sofuoglu, S. (2006). Thyroid function in early and late alcohol withdrawal: Relationship with aggression, family history, and onset age of alcoholism. *Alcohol and Alcoholism*, 41, 515–521.

Pagan, J. L., Rose, R. J., Viken, R. J., Pulkkinen, L., Kaprio, J., & Dick, D. M. (2006). Genetic and environmental influences on stages of alcohol use across adolescence and into young adulthood. *Behavior Genetics*, 36, 483–497.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, 2, e190.

Penning, T. D., Talley, J. J., Bertenshaw, S. R., Carter, J. S., Collins, P. W., Docter, S., . . . Isakson, P. C. (1997). Synthesis and biological evaluation of the 1,5-diarylpyrazole class of cyclooxygenase-2 inhibitors: Identification of 4- 5-(4-methylphenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl benzenesulfona mide (SC-58635, Celecoxib). *Journal of Medicinal Chemistry*, 40, 1347–1365.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-plus*. New York, NY: Springer.

Prestes, A. P., Marques, F. Z. C., Hutz, M. H., Roman, T., & Bau, C. H. D. (2007). Tobacco smoking and the ADRA2A C-1291G polymorphism. *Journal of Neural Transmission*, 114, 1503–1506.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., . . . Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559–575.

Quertemont, E. (2004). Genetic polymorphism in ethanol metabolism: Acetaldehyde contribution to alcohol abuse and alcoholism. *Molecular Psychiatry*, 9, 570–581.

Roman, T., Polanczyk, G. V., Zeni, C., Genro, J. P., Rohde, L. A., & Hutz, M. H. (2005). Further evidence of the involvement of alpha-2A-adrenergic receptor gene (ADRA2A) in inattentive dimensional scores of attention-deficit//hyperactivity disorder. *Molecular Psychiatry*, 11, 8–10.

Roman, T., Schmitz, M., Polanczyk, G. V., Eizirik, M., Rohde, L. A. & Hutz, M. H. (2003). Is the alpha-2A adrenergic receptor gene (ADRA2A) associated with attention-deficit/hyperactivity disorder? *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics*, 120B(1), 116–20.

Rutter, M., Silberg, J., O'Connor, T., & Simonoff, E. (1999). Genetics and child psychiatry: II —- Empirical research findings. *Journal of Child Psychology and Psychiatry*, 40, 19–55.

Schumann, G., Coin, L. J., Lourdusamy, A., Charoen, P., Berger, K. H., Stacey, D., . . . Elliott, P. (2011). Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. *Proceedings of the National Academy of Sciences*, 108, 7119–7124.

Silberg, J., Rutter, M., D'Onofrio, B., & Eaves, L. (2003). Genetic and environmental risk factors in adolescent substance use. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 44, 664–676.

Simonoff, E., Pickles, A., Meyer, J. M., Silberg, J. L., Maes, H. H., Loeber, R., . . . Eaves, L. J. (1997). The Virginia twin study of adolescent behavioral development: Influences of age, sex, and impairment on rates of disorder. *Archives of General Psychiatry*, 54, 801–808.

Singer, J. D. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, New York: Oxford University Press.

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 9440–9445.

Talmud, P. J., Cooper, J. A., Gaunt, T., Holmes, M. V., Shah, S., Palmen, J., . . . Humphries, S. E. (2011). Variants of ADRA2A are associated with fasting glucose, blood pressure, body mass index and type 2 diabetes risk: Meta-analysis of four prospective studies. *Diabetologia*, 54, 1710–1719.

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., . . . Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, 453, 534–538.

The Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447, 661–678.

Thoeringer, C. K., Erhardt, A., Sillaber, I., Mueller, M. B., Ohl, F., Holsboer, F., . . . Keck, M. E. (2010). Long-term anxiolytic and antidepressant-like behavioural effects of tiagabine, a selective GABA transporter-1 (GAT-1) inhibitor, coincide with a decrease in HPA system activity in C57BL/6 mice. *Journal of Psychopharmacology*, 24, 733–743.

Thoeringer, C. K., Ripke, S., Unschuld, P. G., Lucae, S., Ising, M., Bettecken, T., . . . Erhardt, A. (2009). The GABA transporter 1 (SLC6A1): A novel candidate gene for anxiety disorders. *Journal of Neural Transmission*, 116, 649–657.

Treutlein, J., Cichon, S., Ridinger, M., Wodarz, N., Soyka, M., Zill, P., . . . Rietschel, M. (2009). Genome-wide association study of alcohol dependence. *Archives of General Psychiatry*, 66, 773–784.

van den Oord, E. J., Adkins, D. E., McClay, J., Lieberman, J., & Sullivan, P. F. (2009). A systematic method for estimating individual responses to treatment with antipsychotics in CATIE. *Schizophrenia Research*, 107, 13–21.

van der Sluis, S., Posthuma, D., Nivard, M. G., Verhage, M., & Dolan, C. V. (2013). Power in GWAS: Lifting the curse of the clinical cut-off. *Molecular Psychiatry*, 18, 2–3.

Vengeliene, V., Bilbao, A., Molander, A., & Spanagel, R. (2008). Neuropharmacology of alcohol addiction. *British Journal of Pharmacology*, *154*, 299–315.

Waldman, I. D., Nigg, J. T., Gizer, I. R., Park, L., Rappley, M. D., & Friderici, K. (2006). The adrenergic receptor alpha-2A gene (ADRA2A) and neuropsychological executive functions as putative endophenotypes for childhood ADHD. *Cognitive Affective & Behavioral Neuroscience*, *6*, 18–30.

Whitelaw, N. C., & Whitelaw, E. (2006). How lifetimes shape epigenotype within and across generations. *Human Molecular Genetics*, *15*, R131–R137.

WHO. (2011). *Global status report on alcohol and health*. Geneva: Author.

Wiese, J. G., Shlipak, M. G., & Browner, W. S. (2000). The alcohol hangover. *Annals of Internal Medicine 132*, 897–902.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology*, *10*, 395–426.

Zuo, L., Gelernter, J., Zhang, C. K., Zhao, H., Lu, L., Kranzler, H. R., . . . Luo, X. (2011). Genome-wide association study of alcohol dependence implicates KIAA0040 on chromosome 1q. *Neuropsychopharmacology*, *37*, 557–566.