

UC San Diego

UC San Diego Previously Published Works

Title

Genome-wide association studies of the self-rating of effects of ethanol (SRE).

Permalink

<https://escholarship.org/uc/item/94p1n78c>

Journal

Addiction biology, 25(2)

ISSN

1355-6215

Authors

Lai, Dongbing
Wetherill, Leah
Kapoor, Manav
et al.

Publication Date




2020-03-01

DOI

10.1111/adb.12800

Peer reviewed

Genome-wide association studies of the self-rating of effects of ethanol (SRE)

Dongbing Lai¹  | Leah Wetherill¹ | Manav Kapoor² | Emma C. Johnson³ | Melanie Schwandt⁴ | Vijay A. Ramchandani⁵ | David Goldman⁴ | Geoff Joslyn⁶ | Xi Rao¹ | Yunlong Liu¹ | Sean Farris⁷ | R. Dayne Mayfield⁷  | Danielle Dick⁸ | Victor Hesselbrock⁹ | John Kramer¹⁰ | Vivia V. McCutcheon³  | John Nurnberger^{1,11} | Jay Tischfield¹² | Alison Goate² | Howard J. Edenberg^{1,13} | Bernice Porjesz¹⁴ | Arpana Agrawal³ | Tatiana Foroud¹ | Marc Schuckit¹⁵

¹Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana

²Department of Neuroscience, Icahn School of Medicine at Mt. Sinai, New York, New York

³Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri

⁴Office of the Clinical Director, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland

⁵Section on Human Psychopharmacology, Division of Intramural Clinical and Biological Research, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland

⁶Ernest Gallo Clinic and Research Center, Emeryville, California

⁷Waggoner Center for Alcohol and Addiction Research, The University of Texas at Austin, Austin, Texas

⁸Department of Psychology, Virginia Commonwealth University, Richmond, Virginia

⁹Department of Psychiatry, University of Connecticut, Farmington, Connecticut

¹⁰Department of Psychiatry, Roy Carver College of Medicine, University of Iowa, Iowa City, Iowa

¹¹Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana

¹²Department of Genetics and the Human Genetics Institute of New Jersey, Rutgers University, Piscataway, New Jersey

¹³Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana

¹⁴Henri Begleiter Neurodynamics Lab, Department of Psychiatry, State University of New York, Downstate Medical Center, Brooklyn, New York

¹⁵Department of Psychiatry, University of California, San Diego Medical School, San Diego, California

Correspondence

Dongbing Lai, Department of Medical and Molecular Genetics, Indiana University School of Medicine, 410 W. 10th Street, HS 4000, HITS, Indianapolis, IN 46202-3002.
Email: dlai@iu.edu

Funding information

National Institute on Alcohol Abuse and Alcoholism (NIAAA), Grant/Award Number: AA020926, AA012404; National Institute on Drug Abuse (NIDA), Grant/Award Number: DA032573; National Institutes of Health, Grant/Award Numbers: HHSN268201200008I and U10AA008401; NIAAA Division of Intramural Clinical and Biological Research

Abstract

The level of response (LR) to alcohol as measured with the Self-Report of the Effects of Alcohol Retrospective Questionnaire (SRE) evaluates the number of standard drinks usually required for up to four effects. The need for a higher number of drinks for effects is genetically influenced and predicts higher risks for heavy drinking and alcohol problems. We conducted genome-wide association study (GWAS) in the African-American (COGA-AA, N = 1527 from 309 families) and European-American (COGA-EA, N = 4723 from 956 families) subsamples of the Collaborative Studies on the Genetics of Alcoholism (COGA) for two SRE scores: SRE-T (average of first five times of drinking, the period of heaviest drinking, and the most recent 3 months of consumption) and SRE-5 (the first five times of drinking). We then meta-analyzed the two COGA subsamples (COGA-AA + EA). Both SRE-T and SRE-5 were modestly heritable (h^2 : 21%-31%) and genetically correlated with alcohol dependence (AD) and DSM-IV AD criterion count (r_g : 0.35-0.76). Genome-wide significant associations

were observed (SRE-T: chromosomes 6, rs140154945, COGA-EA $P = 3.30E-08$ and 11, rs10647170, COGA-AA+EA $P = 3.53E-09$; SRE-5: chromosome13, rs4770359, COGA-AA $P = 2.92E-08$). Chromosome 11 was replicated in an EA dataset from the National Institute on Alcohol Abuse and Alcoholism intramural program. In silico functional analyses and RNA expression analyses suggest that the chromosome 6 locus is an eQTL for *KIF25*. Polygenic risk scores derived using the COGA SRE-T and SRE-5 GWAS predicted 0.47% to 2.48% of variances in AD and DSM-IV AD criterion count in independent datasets. This study highlights the genetic contribution of alcohol response phenotypes to the etiology of alcohol use disorders.

KEYWORDS

genetic correlation, genome-wide association study (GWAS), heritability, polygenic risk score, RNA expression, self-rating of the effects of ethanol (SRE)

1 | INTRODUCTION

A low level of response (LR) to alcohol, a measure of the intensity of reaction to alcohol as evaluated by either alcohol challenges, or a self-report questionnaire regarding the number of standard drinks needed for a range of effects is an important predictor of the future development of alcohol dependence (AD).¹ AD is associated with considerable morbidity and affects about 12.5% of the US population during their lifetime.² Typically, individuals who experience fewer intoxicating or sedating effects of alcohol are more likely to escalate their drinking behaviors.³ Thus, low LR has been related to future problems with alcohol primarily via heavy consumption,³ even after accounting for histories of heavy drinking⁴ and family history.⁵ Multiple putative pathways, including overly positive expectations of the effects of alcohol, heavier peer drinking, and using alcohol to cope with stress,³ have also been found to contribute to the heightened risk for AD in those with low LR. Interestingly, recent research suggests college-based prevention programs that target LR are effective in individuals with low, but not high, LR.^{6,7} Thus, even though low LR is a marker for predisposition to AD, largely via heavy consumption, individual differences in LR are also associated with unique sources of variance.

A person's LR can be measured by administering an alcohol challenge, either through oral ingestion or intravenously, and recording the individual's level of reaction over several hours. Such alcohol-administration paradigms are time-consuming and costly, making it difficult to evaluate large numbers of subjects.¹ LR can also be measured by the retrospective Self-Report of the Effects (SRE) of Alcohol Questionnaire, in which participants estimate the number of standard drinks usually required to experience up to four different effects of alcohol (feel an effect; feel dizzy or slurring speech; stumble or walk in an uncoordinated manner; pass out or fall sleep) at three time periods (first five times of drinking, the period of heaviest drinking, and the most recent 3 months of consumption). SRE scores and LR measured during an alcohol challenge are correlated ($r = 0.3-0.6$), suggesting that the easier to collect SRE can be used in large-scale studies to measure LR.^{1,8,9} SRE

scores also have an 1 year test-retest correlation between 0.72 and 0.82 and 5-year retest of 0.66, indicating high levels of reliability.^{8,9} Their predictive validities regarding future heavy drinking and alcohol problems across different populations and sexes in studies conducted in the United States, United Kingdom, and Switzerland are also consistent.^{4,10-15} These studies cumulatively indicate that an individual's LR, as measured by the SRE, could serve as an additional risk marker for the development of problem drinking, and can be a useful tool in efforts to personalize prevention and treatment.⁶

Both AD and LR (self-report or alcohol challenge results) are heritable.^{11,16,17} While there have been numerous efforts to identify genetic variants for AD,¹⁸ there have been only a few studies examining the genetic contributions to SRE. Two linkage studies, one using 745 European ancestry individuals (included in this study)¹⁹ and another in American Indians,²⁰ detected modest evidence of linkage. Recently, a meta-analysis of genome-wide association studies (GWAS) of SRE score of the first five times of drinking (SRE-5) did not identify genome-wide significant loci.²¹ Genetic analyses of the average SRE scores across the three time periods (SRE-T) have been limited to candidate gene studies.²² A recent analysis showed that polygenic risk scores derived from a GWAS of alcohol consumption in a large UK population of older adults explained minimal variance in SRE-T²³ suggesting that even at a polygenic level, SRE scores might be associated with unique genetic variation. Given the importance of SRE as a predictor of problem drinking, and its role in intervention research, the identification of such variants is necessary.

We conducted genome-wide association studies of SRE scores in the Collaborative Study on the Genetics of Alcoholism (COGA). First, using variance decomposition approaches for family data, we estimated the heritability of the SRE scores in the African-American and European-American COGA subsamples. The covariance between SRE-T, SRE-5, AD (defined by DSM-IV), and DSM-IV AD criterion count (used as a measurement of severity of AD) was decomposed into its genetic and environmental sources. We then performed GWAS to identify genetic variants contributing to the variation in SRE-T and SRE-5 scores. Genome-wide significant loci were examined

in two independent datasets and in RNA expression analysis using brain tissue from individuals with AD and controls. Finally, we derived polygenic risk scores (PRS) from these discovery GWAS and evaluated whether the polygenic effects of SRE-related variants predicted AD and DSM-IV AD criterion count in four independent AA and EA datasets.

2 | METHODS

2.1 | Sample and assessments

AD probands from alcohol use disorder inpatient and outpatient treatment facilities and their family members were recruited from the seven participating COGA sites. Comparison individuals and their families were ascertained from the community in the same seven areas.²⁴ Institutional review boards at all sites approved the study and all participants provided informed consent.

The SRE of Alcohol Questionnaire⁹ is a self-report instrument used to measure the number of standard drinks required to produce up to four effects of alcohol at three different time periods: (a) the first five times the subject used alcohol; (b) the period of heaviest drinking; and (c) the most recent 3 months of consumption. For each time period, the subject is asked: (a) "How many (standard) drinks did it take for you to begin to feel an effect?"; (b) "How many drinks did it take for you to feel a bit dizzy or begin to slur your speech?"; (c) "How many drinks did it take you to begin to stumble or walk in an uncoordinated manner?"; (d) "How many drinks did it take you to pass out or fall asleep when you did not want to?" Extreme observations were winsorized at the mean plus two standard deviations and only individuals who drank at least two drinks on one occasion were included in the analysis. On the basis of their distributions, the square root of SRE-T and the natural logarithm of SRE-5 were used in analyses. Distributions of original and transformed SRE scores are in supplemental material.

The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) interview and an adolescent version of the SSAGA²⁴ were administered to adults and individuals age 17 or under, respectively. An AD case was defined as an individual age 15 or older who met the DSM-IV criteria for AD, to exclude early onset AD, which may have different etiology.²⁵ Control subjects were at least 21 years old and had consumed at least one full drink of alcohol in their lifetime, but who did not endorse more than one DSM-IV AD criterion. We set the minimum age as 21 to exclude individuals who had not have passed through the period of major risk for the onset of AD.^{26,27} All others were not classified. Although these unclassified individuals were not used in any analysis related to DSM-IV AD, they were included in all other analysis (eg, GWAS of SRE scores).

2.2 | Genotyping, quality review, and imputation

COGA samples were generated on four arrays: Illumina Human1M array (Illumina, San Diego, CA, USA) genotyped at the Center for Inherited Disease Research (CIDR), Johns Hopkins University; Illumina

Human OmniExpress 12 V1 array (Illumina, San Diego, CA, USA) genotyped at the Genome Technology Access Center, Washington University School of Medicine in St. Louis; Illumina 2.5M array (Illumina, San Diego, CA, USA) at CIDR; and Smokescreen genotyping array (Biorealm LLC, Walnut, CA, USA) genotyped at Rutgers University. Two to 127 samples were genotyped on at least two different arrays and pairwise concordance rates within subjects typed on more than one array exceeded 99.18% for those common SNPs, confirming limited batch or array effects. Sex and heterogeneity checks were performed and any discrepancies were removed.

To confirm pedigree structure, a set of 47 000 variants genotyped on all four arrays were used. These variants met the following criteria: (a) not in linkage disequilibrium (LD, defined as $r^2 < .5$); (b) minor allele frequency (MAF) greater than 10%; (c) Hardy-Weinberg equilibrium (HWE) P value > 0.001 ; and (d) missing rate less than 2%. Pairwise identity by descent was computed using PLINK²⁸ to confirm family relationships. When necessary, family structures were altered accordingly. The same set of 47 000 variants were also used to estimate principal components (PCs) of population stratification using Eigenstrat.²⁹ On the basis of the first two PCs, each individual was assigned an ancestry (ie, African-American, European-American, or other). Family ancestry was assigned on the basis of the individual ancestry of the greatest proportion of family members.

Prior to imputation, Mendelian inconsistencies were detected using Pedcheck,³⁰ and genotype inconsistencies were set to missing. Variants with missing rates $> 5\%$, MAF $< 3\%$, and HWE P values < 0.0001 were excluded. In addition, variants with A/T or C/G alleles were also removed to avoid ambiguities in strand designation. SHAPEIT2,³¹ which utilizes pedigree information to increase accuracy, was used to phase haplotypes of each sample. All samples were imputed to 1000 genomes the cosmopolitan reference panel (Phase 3, version 5, <http://www.internationalgenome.org/>) using Minimac3.³² Because of the differences in the variants on each array, samples were imputed separately by array. After imputation, variants with $R^2 < 0.30$ were excluded. To allow for further Mendelian error checking, genotype probabilities were converted to genotypes if they were greater than 0.90 (similar to procedures for QC and analysis of family data in Walters et al¹⁸; see also Kranzler et al that used converted genotypes³³). All genotype and imputed variants with missing rates less than 20%, MAF $\geq 1\%$, and HWE P values > 0.000001 were used in analyses. Variants that passed QC were 15 464 468 and 7 823 558 and were included in the AA and EA subsamples analyses, respectively. All variants were mapped to NCBI GRCh37.

2.3 | Heritability, genetic correlations, and GWAS

All analyses were performed separately in COGA AA (COGA-AA) and EA (COGA-EA) subsamples, then meta-analyzed (COGA-AA + EA). Because of the complex pedigree structure in COGA, methods using genome-wide genotype data such as Genome-wide complex traits analysis³⁴ could not be used to estimate heritability or genetic correlations. Instead, using an approach similar to those applied in classical twin

studies, the Sequential Oligogenic Linkage Analysis Routines (SOLAR8.3.1)³⁵ package, which was designed to partition variance in large pedigrees, was used to estimate heritability of SRE scores (ie, proportion of total phenotypic variance attributable to latent genetic influences), as well as the genetic correlations between the SRE measures and AD and DSM-IV AD criterion count.³⁵ Sex and birth cohorts (birth year: 1890-1929; 1930-1949; 1950-1969; ≥ 1970) were important covariates as shown in previous studies³⁶; therefore both of them and the first four PCs were used as covariates in SOLAR analyses.

GWAS was conducted using linear mixed models with random effects adjusting for relatedness within families using the R package GWAF.³⁷ Sex, birth cohorts, array indicators, and the first four PCs were included as covariates. METAL³⁸ was used to perform meta-analysis with effect size of each variant weighted by the inverse of the estimated standard error from each subsample. Genomic control was applied to the results of each subsample as well as to the meta-analysis results. As SRE-T and SRE-5 were highly correlated (correlation coefficients were 0.75 and 0.78 in COGA-AA and COGA-EA, respectively), we used matSpD³⁹ to spectrally decompose the correlation matrix of SRE scores, resulting 1.44 effective test. Thus, the genome-wide significance was set at $P < 3.5 \times 10^{-8}$. In regions of genome-wide significance, conditional analyses were performed by including the most significant variant in the region as a covariate.

2.4 | Gene prioritization

FUMA,⁴⁰ a web-based tool that incorporates information from multiple public databases, was used to prioritize potentially causal genes. Three gene prioritization strategies are implemented in FUMA: positional mapping, expression Quantitative Trait Locus (eQTL) mapping, and chromatin interaction mapping. Positional mapping annotates variants based on their physical positions and functional consequences. For any region in which at least one variant met genome-wide significance, all variants in that region with P values $\leq 1.0E-04$ were annotated. If an annotated variant was in the promoter or untranslated regions (UTRs) of a gene, or in exonic or splicing regions of a gene and predicted to be deleterious, that gene was prioritized as a potential causal gene. Ensembl (build 85; <http://www.ensembl.org/>) was used to map variants to genes. For eQTL mapping, genes within 1 Mb of the most significant variant were tested, and those with significant associations (defined as $FDR < 0.05$) were considered as potential causal genes. eQTL data from brain tissues in BRAINEAC (<http://www.braineac.org/>), CommonMind Consortium (<https://www.synapse.org>), and GTEx v7 (<https://www.gtexportal.org/>) were used. Genes prioritized by any of these three strategies were considered as potential causal genes and further examined by subsequent RNA expression analysis.

2.5 | RNA expression analysis

To test the associations between the most significant variants and gene expressions, human autopsy brain samples were obtained from the New South Wales Brain Tissue Resource Centre (NSWBTRC) at

the University of Sydney (<http://sydney.edu.au/medicine/pathology/btrc/>). Fresh frozen samples of the superior frontal cortex (Brodmann area 8) were used. This region was selected because a prior enrichment analysis suggests aggregation of AD-associated GWAS signals in this brain region.⁴¹ In addition, there was an adequate sample size available for this region. The samples were sequenced on Illumina Hi-Seq (Illumina, CA, USA) at the New York Sequencing center (N = 83) and at the Waggoner Center for Alcohol and Addiction Research (N = 60). Total RNA was extracted from frozen tissue and samples with contamination or degraded quality RNA (RNA integrity numbers [RIN] < 5.0) were excluded. Samples of 138 passed quality review and were used in the analysis (65 AD and 73 age- and sex-matched controls). The KAPA Stranded RNA-Seq Kit with RiboErase was used for library preparation. The library QC included a measurement of the average size of library fragments using the Fragment Analyzer (Advanced Analytical Technologies, IA, USA), and estimation of the total concentration of DNA by PicoGreen (Thermo Fisher Scientific, MA, USA). Raw reads were aligned to human genome 19 (hg19) using STAR aligner.⁴² Quality control was assessed using RSeQC⁴³ and Picard (% GC, % duplicates, gene body coverage, unsupervised clustering, and library complexity, <http://broadinstitute.github.io/picard/>). The Picard "MarkDuplicates" option was used to flag and remove the duplicate reads. Gene quantification was performed with featureCounts using Gencode annotations. We filtered out all genes with lower expression in a substantial fraction of the cohort, with 18 463 genes with at least 1 CPM (counts per million) in at least 50% of the individuals; note that only these genes were carried forward in all subsequent analyses.

DNA was extracted from the same brain tissue samples used in the RNA analysis. Genotyping was performed on the Axiom Biobank Plus Genotyping Array (Catalog number: 000854). Initial QC was performed using PLINK²⁸ to remove variants with: (a) no alternate alleles; (b) genotyping call rate ≤ 0.98 ; (c) Hardy-Weinberg Equilibrium P value $< 5E-5$; and (d) samples with genotyping call rate < 0.90 . Phasing was performed on each chromosome using SHAPEIT2,³¹ then imputation was done using IMPUTE2⁴⁴ with the 1000 Genomes Phase 1 integrated reference panel excluding singleton variants. Analysis was performed on the imputed variants with imputation score (INFO) ≥ 0.8 and estimated MAF ≥ 0.05 .

eQTLs were computed using a linear model on the imputed genotype dosages using Matrix EQTL69.⁴⁵ If the most significant variant was filtered out in this sample because of QC, the next available variant with the lowest P value and in LD ($r^2 > .9$) with the most significant variant was tested instead. The gene expression data were adjusted for sex, age, post-mortem interval, RIN, batch, and decedent phenotype (AD or control).

2.6 | Replication samples

The first replication dataset was from the San Diego Sibling Pair investigation and genotyped on the Illumina HumanCNV370-duo array (Illumina, San Diego, CA, USA). This study was approved by the

Human Subjects Protection Committee of the University of California, San Diego. The same quality review and imputation used in COGA were performed. This dataset included both unrelated and related individuals. Empirical kinships were estimated from genome-wide genotype data using the vcf2kinship tool as implemented in RVTESTS and adjusted in mixed models using RVTESTS.⁴⁶ The majority of the dataset had European ancestry; therefore, only EA samples was used (SD-EA, N = 206, 68 AD cases and 138 controls). Only sex was available and included as a covariate.

Two datasets from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) Intramural program were also used as replication. Participants were recruited under two NIH Institutional Review Board-approved screening and assessment protocols and were comprehensively assessed at the National Institutes of Health Clinical Center (Bethesda, Maryland, USA) between 2005 and 2015. All participants provided written informed consent. Genotyping was conducted at the NIAAA Laboratory of Neurogenetics (Rockville, MD, USA) using the Illumina Human OmniExpress and Human OmniExpressExome arrays. The same quality review and imputation as in COGA were performed; 169 (130 AD cases, 27 controls, and 12 not classified) and 100 (36 AD cases, 27 controls, and 12 not classified) unrelated AA (NIAAA-AA) and EA (NIAAA-EA) individuals were used in analyses, respectively. Sex and age were included as covariates. Only AA samples were adjusted for the first three PCs because unlike COGA-EA, EA samples in replication and PRS analysis did not include non EA individuals.

2.7 | Polygenic risk scores analyses

Summary statistics from both SRE scores in COGA-AA and COGA-EA were used to derive PRS using PRSice-2.⁴⁷ Four test datasets from the database of genotypes and phenotypes (dbGaP) were included: (a) AA subsample (SAGE-AA) in the *Study of Addiction: Genetics and Environment* (SAGE, phs000092.v1.p1); (b) AA subsample (Yale-Penn-AA) in *Alcohol Dependence GWAS in European and African Americans* (phs000425.v1.p1); (c) EA subsample in SAGE (SAGE-EA); and (d) EA subsample (OZALC-EA) in *The Australian Twin-family Study of Alcohol Use Disorder* (OZALC, phs000181.v1.p1). Overlapping COGA individuals were excluded from these independent datasets. PRSice-2 requires the analysis of unrelated samples; therefore, in datasets with related individuals, one individual was randomly selected from each family. Summary of each test dataset is shown in supplemental

material. Any variant that was within 500 kb from the index variant and had $r^2 \geq 0.25$, with the index variant was clumped using PRSice-2. PRS was calculated on the basis of effect size.⁴⁷ To reduce the burden of multiple testing, only variants that had P values < 0.05 were included. For SAGE, sex and the same birth cohort as defined in COGA were included as covariates, while for Yale-Penn-AA and OZALC-EA, sex and age were used. The first three PCs were also adjusted in AA samples.

3 | RESULTS

3.1 | Sample summary, heritability, and genetic correlations

Table 1 summarizes the COGA-AA and COGA-EA subsamples. Compared with those in COGA-AA, individuals in COGA-EA had higher mean DSM-IV AD criterion count in both AD cases and controls, SRE-T, and SRE-5 scores. Latent genetic influences (familial h^2) contributed to 21% to 31% of the variance in SRE scores, indicating moderate heritability (Table 2). Overall, genetic correlations (r_g) with AD and DSM-IV AD criterion count were higher for SRE-T ($r_g = 0.49$ - 0.76) than for SRE-5 ($r_g = 0.35$ - 0.48) (Table 2).

3.2 | GWAS results

Two loci for SRE-T (chromosome 6 for COGA-EA; chromosome 11 for COGA-AA + EA) and one locus for SRE-5 (chromosome 13 for COGA-AA) reached genome-wide significance (Table 3, Figure S2). Tables S1 and S2 provide a broader list of association results ($P < 1E-04$) for SRE-T and SRE-5, respectively.

SRE-T: the most significant association results were with variants on chromosome 11 in COGA-AA + EA (most significant: rs10647170, $P = 3.53E-09$), with both subsamples providing supporting evidence for the association. Evaluation of rs10647170 in independent datasets provided nominal evidence of replication in the NIAAA-AA ($P = 0.05$), but not in the SD-EA or NIAAA-EA (Table 3) samples. Meta-analysis of COGA-EA + AA and the three replication datasets resulted in a lower P value of $6.45E-10$.

In addition, a genome-wide significant result was noted with a variant on chromosome 6 (rs140154945, $P = 3.30E-08$) in

TABLE 1 Summary of COGA-AA and COGA-EA subsamples

Sample	number of family	number of individual	% AD	number of AD case ^a (%male)	number of AD control ^b (%male)	Mean AD DSM-IV criterion count (stdev) in cases	Mean AD DSM-IV criterion count (stdev) in controls	Mean SRE-T (stdev)	Mean SRE-5 (stdev)
COGA-AA	309	1527	32.74	500 (58.60)	421 (28.50)	5.05 (1.46)	0.19 (0.39)	4.55 (2.86)	3.36 (2.08)
COGA-EA	956	4723	36.35	1717 (60.51)	1336 (27.77)	5.14 (1.54)	0.21 (0.41)	5.26 (2.78)	3.73 (1.98)
Total	1265	6250							

Abbreviations: AD, alcohol dependence; AA, African-American; COGA, Collaborative Studies on the Genetics of Alcoholism; EA, European American.

^aIndividuals who met the DSM-IV criteria for AD but younger than 15 were not classified.

^bIndividuals who had two DSM-IV criteria, never drank, or younger than 21 were not classified.

TABLE 2 Heritability, genetic correlations between AD, DSM-IV AD criterion count and SRE scores

Sample	Phenotype	Heritability (P value)	AD		DSM-IV AD criterion count	
			Genetic correlation	P value	Genetic correlation	P value
COGA-AA	SRE-T	0.21 (4.10E-06)	0.76	1.25E-07	0.63	4.20E-06
COGA-AA	SRE-5	0.20 (6.35E-05)	0.48	1.88E-03	0.40	6.13E-03
COGA-EA	SRE-T	0.31 (1.14E-35)	0.61	4.55E-25	0.49	8.26E-19
COGA-EA	SRE-5	0.31 (7.01E-34)	0.41	4.52E-11	0.35	9.44E-10

Abbreviations: AD, alcohol dependence; AA, African-American; COGA, Collaborative Studies on the Genetics of Alcoholism; EA, European American.

COGA-EA (Table 3) that not supported in COGA-AA. There was also no evidence of replication for rs140154945 in the EA replication datasets.

Regional association plots demonstrated strong LD support in both the chromosomes 6 and 11 regions (Figures S3A, S3B, and S3C). Conditional analyses using the top variant in each region as a covariate confirmed that there were no additional association signals in these chromosomal regions associated with SRE-T.

SRE-5: genome-wide significant associations were identified in COGA-AA on chromosome 13 (rs4770359, $P = 2.92E-08$), although it was not associated in COGA-EA (Table 3). The variant was not replicated in the NIAAA-AA. There was modest LD support in this region possibly because of unusual LD pattern as shown in 1000 genomes data using LDlink (<https://ldlink.nci.nih.gov/?tab=home>) (Figures S3D and S4). Unlike SRE-T, conditional analyses found that some variants demonstrated modest association ($P < 0.001$) after adjustment of the top variant, suggesting that additional independent associations might exist in each region, consistent with the LD patterns showing in Figure S3D. There were no genome-wide significant association findings in COGA-AA + EA and COGA-EA GWAS.

The effect sizes of the most significant variants in each locus in discovery and replication samples are shown in the forest plots in Figure S5.

3.3 | Gene prioritization and RNA expression analysis

For chromosome 11, one variant (rs2270678, COGA-AA + EA $P = 1.81E-08$) was located in the 5' UTR of *PRMT3* (protein arginine methyltransferase 3). Another variant (rs3781679, COGA-AA + EA $P = 3.58E-05$, Table S1) was located in the 3' UTR of *HTATIP2* (HIV-1 Tat interactive protein 2). Therefore, both genes were prioritized in FUMA. Chromatin interaction mapping further prioritized *DBX1* (developing brain homeobox 1) and *SLC6A5* (solute carrier family 6 member 5) in this region. On chromosome 6, three variants were located in the 3' UTR of *AFDN* (adherens junction formation factor) (Table S1) and this gene was therefore prioritized by FUMA. eQTL analysis further prioritized the gene *KIF25* (kinesin family member 25) in this region (multiple variants had $FDR < 0.01$). No genes were prioritized in the chromosome 13 region.

For the six genes prioritized by FUMA (Chromosome 11: *PRMT3*, *HTATIP2*, *DBX1*, *SLC6A5*; Chromosome 6: *AFDN*, *KIF25*), RNA expression in brain tissue obtained from NSWBRTC was examined. After adjusting for six tests, the Bonferroni correction threshold was set at $P < 0.0083$. Rs140154945 did not pass QC in this smaller postmortem tissue sample, therefore, the next most significant variant in high LD (rs3778662, COGA-EA $P = 7.24E-08$, $r^2 = 0.92$ with rs140154945) was used. After adjustment for covariates, including AD status of decedent, rs3778662 was identified as an eQTL of *KIF25* ($P = 6.62E-05$).

3.4 | PRS analyses

PRS analysis results are summarized in Table 4. For both SRE-T and SRE-5, there are four discovery-test dataset pairs (i.e., Bonferroni $P < 0.0125$). Overall, PRS from SRE-T outperformed those from SRE-5 in prediction of AD phenotypes; PRS of both SRE-T and SRE-5 predicted greater variance in DSM-IV AD criterion count than AD; prediction was also higher in EA than in AA samples. The strongest finding was for polygenic scores based on the COGA-EA SRE-T GWAS, which predicted 2.48% of the variance in AD in the SAGE-EA.

4 | DISCUSSION

This is the first GWAS to identify common variants associated with SRE-T and the first GWAS of SRE scores in AA samples. This is also the first study to report genome-wide significant findings for any SRE score. These findings are not under linkage regions previously published^{19,20} using SRE-5. Of the genome-wide significant loci, the locus on chromosome 11 was nominally replicated in one sample, although the replication sample was small. Of the genes prioritized for SRE-T in the EA sample, *KIF25* was supported by data from both in silico functional analyses and RNA expression analysis. In addition, PRS derived from these discovery GWAS of SRE predicted variance in AD-related phenotypes in independent datasets. Broadly, these GWAS analyses suggest that larger samples of SRE phenotypes, particularly SRE-T, might be highly informative in the identification of loci related to LR, which influences risk for development of AD and is emerging as a notable target for prevention.

TABLE 3 Summary of genome wide significant variants. Genome-wide significant *P* values are in bold

CHR	BP	Variants	Gene	Function	Minor allele	COGA-AA MAF	COGA-EA MAF	SRE-T <i>P</i> values					Replication dataset				
								Discovery dataset									
								COGA-AA	COGA-EA	COGA-AA+EA	NIAAA-AA	NIAAA-EA	SD-EA	Meta ^a			
6	168,329,922	rs140154945	AFDN	Intronic	TA	0.23	0.31	0.22	3.30E-08	7.19E-06	NA	0.14	0.24	1.82E-07			
11	20,401,835	rs7934001	HTATIP2	Intronic	T	0.20	0.05	2.46E-06	1.66E-03	4.49E-08	0.73	0.02	0.40	4.53E-08			
11	20,401,940	rs7936865	HTATIP2	Intronic	T	0.20	0.05	3.47E-06	9.34E-04	2.97E-08	0.74	0.03	0.41	2.64E-08			
11	20,402,702	rs2055244	HTATIP2	Intronic	T	0.20	0.05	2.64E-06	1.78E-03	4.35E-08	0.92	0.06	0.47	4.88E-08			
11	20,402,886	rs2055246	HTATIP2	Intronic	C	0.13	0.05	7.51E-08	3.21E-03	1.57E-08	0.16	0.05	0.47	8.31E-09			
11	20,409,169	rs2270678	PRMT3	UTR5	G	0.21	0.04	3.18E-06	6.73E-04	1.81E-08	0.38	0.11	0.37	8.20E-09			
11	20,411,516	rs2403592	PRMT3	Intronic	A	0.13	0.04	2.72E-07	1.42E-03	9.17E-09	0.03	0.19	0.39	1.64E-09			
11	20,420,005	rs7938365	PRMT3	Intronic	G	0.12	0.05	4.46E-07	2.04E-03	2.65E-08	0.05	0.16	0.57	7.45E-09			
11	20,422,730	rs7480236	PRMT3	Intronic	T	0.13	0.04	5.45E-07	1.04E-03	1.12E-08	0.03	0.16	0.41	1.96E-09			
11	20,422,731	rs7480181	PRMT3	Intronic	T	0.13	0.04	5.45E-07	1.04E-03	1.12E-08	0.03	0.16	0.41	1.96E-09			
11	20,424,851	rs11025551	PRMT3	Intronic	G	0.13	0.05	4.10E-07	1.18E-03	9.66E-09	0.03	0.18	0.40	1.72E-09			
11	20,427,324	rs12221482	PRMT3	Intronic	C	0.12	0.05	1.66E-07	1.09E-03	5.49E-09	0.05	0.16	0.42	1.26E-09			
11	20,432,046	rs10647170	PRMT3	Intronic	TAGTC	0.12	0.04	1.69E-07	7.25E-04	3.53E-09	0.05	0.26	0.42	6.45E-10			
11	20,432,797	rs74761974	PRMT3	Intronic	A	0.12	0.04	2.24E-07	6.93E-04	3.93E-09	0.05	0.15	0.42	9.04E-10			
11	20,435,147	rs74183993	PRMT3	Intronic	C	0.12	0.04	2.36E-07	7.04E-04	4.69E-09	0.05	0.14	0.42	9.84E-10			
SRE-5 <i>P</i> values																	
13	23,552,108	rs4770359	LINC00621,SGCG	Intergenic	G	0.19	0.46	2.92E-08	0.75	0.08	0.82	NA	NA	6.33E-08			

^aFor meta-analysis, if genome-wide significant findings were from COGA-AA, then COGA-AA, NIAAA-AA were meta-analyzed; if genome-wide significant findings were from COGA-EA, then COGA-EA, NIAAA-EA, and SD-EA were meta-analyzed; otherwise, all datasets were meta-analyzed.

TABLE 4 Results of PRS derived using COGA SRE GWAS summary statistics. $P < 0.0125$ are in bold

Discovery dataset	Test dataset	AD		DSM-IV AD Criterion Count	
		R^2	P value	R^2	P value
SRE-T COGA-EA	SAGE-EA	2.48%	4.13E-07	1.37%	8.55E-07
SRE-T COGA-EA	OZALC-EA	0.34%	0.05	0.60%	1.14E-03
SRE-T COGA-AA	SAGE-AA	1.49%	6.10E-03	1.04%	1.72E-03
SRE-T COGA-AA	Yale-Penn-AA	0.36%	0.03	0.05%	0.33
SRE-5 COGA-EA	SAGE-EA	0.88%	2.37E-03	0.61%	1.08E-03
SRE-5 COGA-EA	OZALC-EA	0.14%	0.21	0.47%	3.85E-03
SRE-5 COGA-AA	SAGE-AA	0.53%	0.10	0.34%	0.07
SRE-5 COGA-AA	Yale-Penn-AA	0.0007%	0.92	0.03%	0.43

SRE-T and SRE-5 represent self-reported assessments of LR. SRE-T and SRE-5 were moderately heritable, but our estimates were lower than some prior reports. For instance, a smaller ($n = 101$) sibling study of nonalcohol-dependent individuals estimated the heritability of SRE-T at 67% to 71%,¹¹ although the corresponding estimate for an “early drinking period” was lower and statistically nonsignificant. Similar high heritability estimates have been noted for other assessments of subjective ratings (eg, 40%-60%).^{16,48} It is possible that our estimates are less precise than those from larger twin studies, as we relied on pedigree data. We did not utilize genomic data to estimate the SNP-heritability of SRE scores as our large pedigrees, when reduced to a smaller cohort of unrelated individuals, is not well-powered to make such estimates. Nonetheless, our estimates suggest that future larger efforts should be successful in estimating the extent to which common variants explicate the heritability of SRE.⁴⁹

Our discovery GWAS in both COGA-AA and COGA-EA identified genome-wide significant loci. The association signal on chromosome 11 was supported by both COGA-AA and COGA-EA. Genes prioritized in this region have been supported by other GWAS. In a recent GWAS of alcohol dependence in a Thai population, the region between *DBX1* and *PRMT3* showed suggestive association with AD.⁵⁰ Another study reported an association between *PRMT3* and smoking.⁵¹ Variants upstream of *PRMT3* demonstrated suggestive association with marijuana-dependence criterion count in an African-American cohort.⁵² Although *HTAPI2* in this region was also prioritized by FUMA, it was not reported to be related to any neuropsychiatric diseases. Two additional genes in this region, *DBX1* and *SLC6A5*, were prioritized by chromatin interaction mapping. *DBX1* has been linked to educational achievement,⁵³ which is negatively correlated with genetic liability to alcohol dependence,¹⁸ while *SLC6A5* was reported to be related to schizophrenia,⁵⁴ which is positively genetically correlated with AD.¹⁸ Thus, there is increasing evidence for the involvement of variants in this region to liability to problem drinking.

Analyses of RNA expression data in brain tissue identified *KIF25* on chromosome 6 to be a gene of interest for SRE-T in the EA sample. The minor alleles of the associated variants in this region increased the SRE-T score, risk of AD, and expression of *KIF25* in both curated brain

tissue data (eg, from GTEx) and RNA expression analysis that accounted for AD status. Because of the small sample size, there was insufficient statistical power to analyze the effects of variants stratified by cases and controls. This gene encodes a protein that belongs to the kinesin-like protein family. A region that is approximately 40 kb away from *KIF25* was shown to be associated with Attention Deficit Hyperactivity Disorder.⁵⁵ However, there are few functional studies of this gene and its exact role in alcohol-related behaviors remains unknown. Nonetheless, these results highlight the importance of *KIF25* expression in LR.

Individuals with low LR, and consequently, higher SRE scores are at increased likelihood of drinking heavily and developing problems AD.¹ The present study found that a preponderance of this correlation is attributable to genetic factors, as indexed by strong correlations in latent heritable factors influencing SRE and AD phenotypes. However, these high correlations should be interpreted within the context of our ascertainment strategy, which oversampled for individuals with family history of AD. It should also be distinguished from more recent SNP-based genetic correlations that examine the extent of similarity between genome-wide effect sizes for two traits. In contrast to the high r_g from the pedigree data, our study also added a novel exploration of the extent to which polygenic liability to SRE relates to AD. Overall, PRS mostly explained modest proportions of variance in AD phenotypes. Nearly 2.5% and 1.4% of the variance in AD diagnosis and criterion count, respectively, were explained by polygenic liability to SRE-T in the SAGE-EA data. Predictions were less significant in other target cohorts, potentially because of the similar ascertainment for alcohol dependence of SAGE and COGA. For example, even though Yale-Penn-AA was ascertained for alcohol dependence, it has a high rate of other illicit drug dependence which may have impacted predictions. Polygenic liability to SRE-5 was poorly related to AD phenotypes, consistent with its lower genetic correlation with AD diagnosis and criterion count.

Our results suggest that some polygenic liability underlying SRE-T and AD is common but also highlight that each phenotype is likely to be influenced by unique genetic factors. For instance, the current strongest signal for AD, rs1229984 in *ADH1B*,¹⁸ was not associated at genome-wide significant levels to SRE-T (EA $P = 0.00027$; AA

$P = 0.019$; rs2066702, strongest signal in AA, $P = 0.012$). Similarly, the most significant variant for SRE-T in chromosome 11, rs10647170, was not associated with AD ($P > 0.18$ in COGA-AA, COGA-EA, and COGA-AA-EA). Thus, the genetic variants associated with SRE and those relating to AD are likely to only modestly overlap. For example, the AD risk is associated with high impulsivity and sensation seeking, low conscientiousness, and the risk for some psychiatric disorders (eg, schizophrenia), characteristics that have little, if any, links to LR.

This study has several limitations. Despite the size of the COGA discovery samples, power to detect association with low frequency variants is modest. Likewise, sample sizes of the replication datasets were small, raising the possibility that true findings were not replicated. Despite this, one association was confirmed by the NIAAA samples, and meta-analyses of all datasets resulted in increased significance of genome-wide significant signals. Additionally, we used hard-called genotypes as necessitated by our family data requiring rigorous Mendelian error checking, which is not straightforward with dosage data (see also Walters et al for a similar approach to family data¹⁸). Nonetheless, results for our top loci are similar when dosage data were used for a confirmatory analysis (eg, rs10647170 1.53E-08). Another limitation inherent in both in silico functional analyses and RNA expression analysis is the source of the data, which are largely from subjects of European ancestry. As a result, conclusions from these functional data may be limited when analyzing associated variants identified in the COGA-AA sample. Nonetheless, this represents an important first step towards characterizing genetic risk in a key index of AD liability. This study is also further strengthened by investigation of both EA and AA individuals, who are known to differ in SRE scores.

We identified novel loci associated with SRE scores in both EA and AA samples. As additional samples with SRE data, specifically SRE-T, become available, individual loci and polygenic risk scores from such GWAS might explain even greater variance in LR. It will set the stage for future interventions that might aim to interrupt progression to AD.^{6,56} This boost in power necessitates further collection of SRE data. Since the assessment is brief and easily administered, and is included the PhenX tool kit (<https://www.phenxtoolkit.org/>), large biobanks and nation-wide data collection efforts are encouraged to gather data on SRE to capture greater variability to susceptibility to future problem drinking.

ACKNOWLEDGEMENTS

COGA: The Collaborative Study on the Genetics of Alcoholism (COGA), principal investigators B.P., V.H., H.E., L.B., includes 11 different centers: University of Connecticut (V.H.); Indiana University (H.J. E., J.N. Jr., T.F., and Y. L.); University of Iowa (S.K. and J.K.); SUNY Downstate (B.P.); Washington University in St. Louis (L.B., J.R., K.B., and A.A.); University of California at San Diego (M.S.); Rutgers University (J.T. and A.B.); Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia; Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA (L.A.), Virginia Commonwealth University (D.D.), Icahn School of Medicine at Mount Sinai (A.G.), and Howard University (R.T.). Other COGA

collaborators include: L.B. (University of Connecticut); J.M., L.W., X.X., D.L., S.O'C., M.P., S.L. (Indiana University); G.C. (University of Iowa; University of Connecticut); J.M., D.C., C.K., A.P., J.Z. (SUNY Downstate); J.-C.W., M.K., S.B. (Icahn School of Medicine at Mount Sinai); A.A., V.M., S.S. (Washington University); J.S., F.A., B.C. (Virginia Commonwealth University); and M.K. (University of Texas Rio Grande Valley). A.P. and H.C. are the NIAAA Staff Collaborators.

We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, P. Michael Conneally, Raymond Crowe, and Wendy Reich, for their critical contributions. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA).

FUNDING AND DISCLOSURE

COGA is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). A. Agrawal receives additional funding support from NIDA (DA032573). Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268201200008I. D. Goldman, M.L. Schwandt, and V.A. Ramchandani are supported by the NIAAA Division of Intramural Clinical and Biological Research. R. Dayne Mayfield is supported from NIH/NIAAA: AA020926 and AA012404. Alison Goate is listed as an inventor on issued U.S. Patent 8080,371, "Markers for Addiction" covering the use of certain variants in determining the diagnosis, prognosis, and treatment of addiction. John Nurnberger receives support as an investigator for Janssen. All other authors have no potential conflicts of interest.

AUTHORS CONTRIBUTION

DL, AA, TF, and MS were responsible for the study concept and design. MK, MS, VAR, DG, GJ, XR, YL, SF, and RDM contributed to the acquisition of data. DL, LW, MK, and EJ performed the analysis. DL drafted the manuscript. DD, VH, JK, VVM, JN, JT, AG, HJE, BP, AA, TF, and MS provided critical revision of the manuscript for important intellectual content. All authors critically reviewed content and approved final.

ORCID

Dongbing Lai  <https://orcid.org/0000-0001-7803-580X>

R. Dayne Mayfield  <https://orcid.org/0000-0002-8045-1789>

Vivia V. McCutcheon  <https://orcid.org/0000-0002-2694-5438>

REFERENCES

- Schuckit MA, Smith TL, Tipp JE. The Self-Rating of the Effects of Alcohol (SRE) form as a retrospective measure of the risk for alcoholism. *Addiction*. 1997;92(8):979-988.

2. Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(11):1609-1640.
3. Schuckit MA, Kalmijn JA, Smith TL, Saunders G, Fromme K. Structuring a college alcohol prevention program on the low level of response to alcohol model: a pilot study. *Alcohol Clin Exp Res*. 2012;36(7):1244-1252.
4. Schuckit MA, Smith TL, Trim RS, et al. The self-rating of the effects of alcohol questionnaire as a predictor of alcohol-related outcomes in 12-year-old subjects. *Alcohol Alcohol*. 2008;43(6):641-646.
5. Schuckit MA, Smith TL. The relationships of a family history of alcohol dependence, a low level of response to alcohol and six domains of life functioning to the development of alcohol use disorders. *J Stud Alcohol*. 2000;61(6):827-835.
6. Schuckit MA, Smith TL, Clausen P, et al. The low level of response to alcohol-based heavy drinking prevention program: one-year follow-up. *J Stud Alcohol Drugs*. 2016;77(1):25-37.
7. Savage JE, Neale Z, Cho SB, et al. Level of response to alcohol as a factor for targeted prevention in college students. *Alcoholism-Clinical and Experimental Research*. 2015;39(11):2215-2223.
8. Schuckit MA, Smith TL, Trim R, Fukukura T, Allen R. The overlap in predicting alcohol outcome for two measures of the level of response to alcohol. *Alcohol Clin Exp Res*. 2009;33(3):563-569.
9. Schuckit MA, Tipp JE, Smith TL, Wiesbeck GA, Kalmijn J. The relationship between self-rating of the effects of alcohol and alcohol challenge results in ninety-eight young men. *J Stud Alcohol*. 1997;58(4):397-404.
10. Daeppen JB, Landry U, Pecoud A, Decrey H, Yersin B. A measure of the intensity of response to alcohol to screen for alcohol use disorders in primary care. *Alcohol Alcohol*. 2000;35(6):625-627.
11. Kalu N, Ramchandani VA, Marshall V, et al. Heritability of level of response and association with recent drinking history in nonalcohol-dependent drinkers. *Alcohol Clin Exp Res*. 2012;36:1034-1041.
12. Pedersen SL, McCarthy DM. Differences in acute response to alcohol between African Americans and European Americans. *Alcohol Clin Exp Res*. 2013;37(6):1056-1063.
13. Schuckit MA, Smith TL, Danko GP, et al. The Ability of the self-rating of the effects of alcohol (SRE) scale to predict alcohol-related outcomes five years later. *J Stud Alcohol Drugs*. 2007;68(3):371-378.
14. Schuckit MA, Smith TL, Trim R, et al. The performance of elements of a 'level of response to alcohol'-based model of drinking behaviors in 13-year-olds. *Addiction*. 2008;103(11):1786-1792.
15. Wall TL, Johnson ML, Horn SM, Carr LG, Smith TL, Schuckit MA. Evaluation of the self-rating of the effects of alcohol form in Asian Americans with aldehyde dehydrogenase polymorphisms. *J Stud Alcohol*. 1999;60(6):784-789.
16. Viken RJ, Rose RJ, Morzorati SL, Christian JC, Li TK. Subjective intoxication in response to alcohol challenge: heritability and covariation with personality, breath alcohol level, and drinking history. *Alcohol Clin Exp Res*. 2003;27(5):795-803.
17. Heath AC, Madden PA, Bucholz KK, et al. Genetic differences in alcohol sensitivity and the inheritance of alcoholism risk. *Psychol Med*. 1999;29(5):1069-1081.
18. Walters RK, Polimanti R, Johnson EC, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018; 21, 1656-1669
19. Schuckit MA, Edenberg HJ, Kalmijn J, et al. A genome-wide search for genes that relate to a low level of response to alcohol. *Alcohol Clin Exp Res*. 2001;25(3):323-329.
20. Ehlers CL, Gizer IR, Schuckit MA, Wilhelmsen KC. Genome-wide scan for self-rating of the effects of alcohol in American Indians. *Psychiatr Genet*. 2010;20(5):221-228.
21. Edwards AC, Deak JD, Gizer IR, et al. Meta-analysis of genetic influences on initial alcohol sensitivity. *Alcohol Clin Exp Res*. 2018;42(12):2349-2359.
22. Schuckit MA. A critical review of methods and results in the search for genetic contributors to alcohol sensitivity. *Alcohol Clin Exp Res*. 2018;42(5):822-835.
23. Johnson EC, St Pierre CL, Meyers JL, et al. The genetic relationship between alcohol consumption and aspects of problem drinking in an ascertained sample. *Alcohol Clin Exp Res*. 2019; 43, 1113-1125.
24. Bucholz KK, Cadoret R, Cloninger CR, et al. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol*. 1994;55(2):149-158.
25. Dick DM, Cho SB, Latendresse SJ, et al. Genetic influences on alcohol use across stages of development: GABRA2 and longitudinal trajectories of drunkenness from adolescence to young adulthood. *Addict Biol*. 2014;19(6):1055-1064.
26. Grant BF, Dawson DA, Stinson FS, Chou SP, Dufour MC, Pickering RP. The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991-1992 and 2001-2002. *Drug Alcohol Depend*. 2004;74(3):223-234.
27. Hingson RW, Heeren T, Winter MR. Age at drinking onset and alcohol dependence. *Arch Pediatr Adolesc Med*. 2006;160(7):739-746.
28. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575.
29. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909.
30. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet*. 1998;63(1):259-266.
31. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10(1):5-6.
32. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287.
33. Kranzler HR, Zhou H, Kember RL, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat Commun*. 2019;10(1):1499.
34. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88(1):76-82.
35. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet*. 1998;62(5):1198-1211.
36. Grucza RA, Bucholz KK, Rice JP, Bierut LJ. Secular trends in the lifetime prevalence of alcohol dependence in the United States: a re-evaluation. *Alcohol Clin Exp Res*. 2008;32(5):763-770.
37. Chen MH, Yang Q. GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics*. 2010;26(4):580-581.
38. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191.
39. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet*. 2004;74(4):765-769.
40. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826.

41. Kapoor M, Farris SP, Wang JC, et al. Analysis of wholegenome-transcriptomic organization in brain to identify genes associated with alcoholism. *Transl Psychiatry*. 2019;9, 89.
42. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
43. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics*. 2012;28(16):2184-2185.
44. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39(7):906-913.
45. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics*. 2012;28(10):1353-1358.
46. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data: Table 1. *Bioinformatics*. 2016;32(9):1423-1426.
47. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics*. 2015;31(9):1466-1468.
48. Heath AC, Martin NG. Intoxication after an acute dose of alcohol: an assessment of its association with alcohol consumption patterns by using twin data. *Alcohol Clin Exp Res*. 1991;15(1):122-128.
49. Ray LA, Mackillop J, Monti PM. Subjective responses to alcohol consumption as endophenotypes: advancing behavioral genetics in etiological and treatment models of alcoholism. *Subst Use Misuse*. 2010;45(11):1742-1765.
50. Gelernter J, Zhou H, Nunez YZ, Mutirangura A, Malison RT, Kalayasiri R. Genomewide association study of alcohol dependence and related traits in a Thai population. *Alcohol Clin Exp Res*. 2018;42(5):861-868.
51. Park SL, Carmella SG, Chen M, et al. Mercapturic acids derived from the toxicants acrolein and crotonaldehyde in the urine of cigarette smokers from five ethnic groups with differing risks for lung cancer. *PLoS ONE*. 2015;10(6):e0124841.
52. Sherva R, Wang Q, Kranzler H, et al. Genome-wide association study of cannabis dependence severity, novel risk variants, and shared genetic risks. *JAMA Psychiat*. 2016;73(5):472-480.
53. Rietveld CA, Esko T, Davies G, et al. Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc Natl Acad Sci U S A*. 2014;111(38):13790-13794.
54. Deng X, Sagata N, Takeuchi N, et al. Association study of polymorphisms in the neutral amino acid transporter genes SLC1A4, SLC1A5 and the glycine transporter genes SLC6A5, SLC6A9 with schizophrenia. *BMC Psychiatry*. 2008;8(1):58.
55. Ebejer JL, Duffy DL, van der Werf J, et al. Genome-wide association study of inattention and hyperactivity-impulsivity measured as quantitative traits. *Twin Res Hum Genet*. 2013;16(2):560-574.
56. Roche DJ, Ray LA. Subjective response as a consideration in the pharmacogenetics of alcoholism treatment. *Pharmacogenomics*. 2015;16(7):721-736.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Lai D, Wetherill L, Kapoor M, et al. Genome-wide association studies of the self-rating of effects of ethanol (SRE). *Addiction Biology*. 2020;25:e12800. <https://doi.org/10.1111/adb.12800>