Received: 1 February 2019 | Revised: 19 April 2019 | Accepted: 11 May 2019

DOI: 10.1111/gbb.12579

ORIGINAL ARTICLE

Genes, Brain and Behavior

Genome-wide association studies of alcohol dependence, DSM-IV criterion count and individual criteria

Dongbing Lai[1] | Leah Wetherill[1] | Sarah Bertelsen[2] | Caitlin E. Carey[3] | Chella Kamarajan[4] | Manav Kapoor[2] | Jacquelyn L. Meyers[4] | Andrey P. Anokhin[5] | David A. Bennett[6] | Kathleen K. Bucholz[5] | Katharine K. Chang[3] | Philip L. De Jager[7,8] | Danielle M. Dick[9] | Victor Hesselbrock[10] | John Kramer[11] | Samuel Kuperman[11] | John I. Nurnberger Jr[1,12] | Towfique Raj[2] | Marc Schuckit[13] | Denise M. Scott[14,15] | Robert E. Taylor[16] | Jay Tischfield[17] | Ahmad R. Hariri[18] | Howard J. Edenberg[1,19] | Arpana Agrawal[5] | Ryan Bogdan[3] | Bernice Porjesz[4] | Alison M. Goate[2] | Tatiana Foroud[1]

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

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

[3]BRAIN Lab, Department of Psychological and Brain Sciences, Washington University School of Medicine, St. Louis, Missouri

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

[5]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri

[6]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois

[7]Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts

[8]Department of Psychiatry, Brigham and Women's Hospital, Boston, Massachusetts

[9]Department of Psychology, Virginia Commonwealth University, Richmond, Virginia

[10]Department of Psychiatry, University of Connecticut, Farmington, Connecticut

[11]Department of Psychiatry, Roy Carver College of Medicine, University of Iowa, Iowa City, Iowa

[12]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana

[13]Department of Psychiatry, San Diego Medical School, University of California, San Diego, California

Abstract

Genome-wide association studies (GWAS) of alcohol dependence (AD) have reliably identified variation within alcohol metabolizing genes (eg, *ADH1B*) but have inconsistently located other signals, which may be partially attributable to symptom heterogeneity underlying the disorder. We conducted GWAS of DSM-IV AD (primary analysis), DSM-IV AD criterion count (secondary analysis), and individual dependence criteria (tertiary analysis) among 7418 (1121 families) European American (EA) individuals from the Collaborative Study on the Genetics of Alcoholism (COGA). Trans-ancestral meta-analyses combined these results with data from 3175 (585 families) African-American (AA) individuals from COGA. In the EA GWAS, three loci were genome-wide significant: rs1229984 in *ADH1B* for AD criterion count (P = 4.16E−11) and *Desire to cut drinking* (P = 1.21E−11); rs188227250 (chromosome 8, *Drinking more than intended*, P = 6.72E−09); rs1912461 (chromosome 15, *Time spent drinking*, P = 1.77E−08). In the trans-ancestral meta-analysis, rs1229984 was associated with multiple phenotypes and two additional loci were genome-wide significant: rs61826952 (chromosome 1, DSM-IV AD, P = 8.42E−11); rs7597960 (chromosome 2, *Time spent drinking*, P = 1.22E−08). Associations with rs1229984 and rs18822750 were replicated in independent datasets. Polygenic risk scores derived from the EA GWAS of AD predicted AD in two EA datasets (P < .01; 0.61%-1.82% of variance). Identified novel variants (ie, rs1912461, rs61826952) were associated with differential central evoked theta power (loss − gain; P = .0037) and reward-related ventral striatum reactivity (P = .008), respectively. This study suggests that studying individual criteria may unveil new insights into the genetic etiology of AD liability.

[14]Department of Pediatrics, Howard University, Washington, District of Columbia

[15]Department of Human Genetics, Howard University, Washington, District of Columbia

[16]Department of Pharmacology, Howard University, Washington, District of Columbia

[17]Department of Genetics, Rutgers University, Newark, New Jersey

[18]Laboratory of NeuroGenetics, Department of Psychology and Neuroscience, Duke University, Durham, North Carolina

[19]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana

Correspondence
Dongbing Lai, Health Information and Translational Sciences (HITS) 4000, 410 West, 10th Street, Indianapolis, IN 46202-3002.
Email: dlai@iu.edu

Funding information
National Institute on Alcohol Abuse and Alcoholism and National Institute on Drug Abuse (NIDA), Grant/Award Number: U10AA008401; National Institute on Drug Abuse, Grant/Award Numbers: DA032573, K01DA037914; NIDA; National Institutes of Health (NIH), Grant/Award Numbers: R01-AG049789, R01-DA031579, AG045231, HD083614, AG052564, DA038834, DA040716; National Science Foundation, Grant/Award Number: DGE-1143954; NIDA; Duke University; NIH, Grant/Award Numbers: AG045231, AG052564, DA038834, DA040716, HD083614, HHSN268201200008I, K02AA018755, R01-AG049789, R01-DA031579, R01-DA033369

KEYWORDS
alcohol dependence, DSM-IV alcohol dependence criterion, DSM-IV criterion count, DSM-IV individual criteria, event-related theta oscillations, functional magnetic resonance imaging, genome-wide association study, item response analysis, meta-analysis, polygenic risk score

1 | INTRODUCTION

Alcohol dependence (AD), characterized by excessive drinking and diagnosed using features such as loss of control over drinking and excessive consumption despite negative consequences, is one of the most common and costly public health problems worldwide.[1] In the United States, 12.5% of the population meets criteria for DSM-IV AD.[1,2] AD is a complex disease with both genetic and environmental underpinnings and an estimated heritability around 50%.[3] Identification of loci associated with AD liability could provide new insights into the biological mechanisms underlying this serious disorder and lead to new therapeutic pathways.

Individual genome-wide association studies (GWAS) of AD have been relatively modest in size (but see a recent large publication using International Classification of Disease [ICD] codes[4]) and have failed to identify consistently replicable loci,[5] with the exception of variants within the alcohol metabolizing genes, notably *ADH1B*, and to a lesser degree, *ADH1C*. A recent large GWAS meta-analysis of 14 904 AD cases and 37 944 controls, which includes some of the samples used in this study, also only detected genome-wide significant (GWS) association with rs1229984 (Europeans) and rs2066702 (African-Americans); both single nucleotide polymorphism (SNP) are in *ADH1B*.[6] However, when examining a broader definition of alcohol use disorders from medical records, loci in additional genes have recently been identified.[4] We have previously conducted GWAS of AD-related phenotypes in smaller subsets of the data used in the present study, but results have eluded replication and power to detect rs1229984 has been low (eg, for AD in a subset of 1884 unrelateds,[7] for AD, criterion count and criteria in 2010 to 2322 individuals from 118 families[8,9]).

One possible challenge to identification of novel loci contributing to AD susceptibility may be the heterogeneity underlying the diagnosis of AD. Meeting criteria for DSM-IV AD requires that an individual endorse any three (or more) of the seven DSM-IV criteria (*Tolerance*; *Withdrawal*; *Drinking more than intended*; *Desire to cut drinking*; *Giving up activities*; *Time spent drinking*; *Drinking despite problems*) during the same 12-month period. However, psychometric literature points to the differential severity and contribution of individual criteria.[10] An approach to reduce diagnostic heterogeneity may be the analysis of individual DSM-IV criteria in addition to the overall AD diagnosis.

Twin studies have suggested that the individual criteria that comprise the AD diagnosis are heritable.[11-13] For instance, Kendler et al showed the heritability of individual criteria ranged from 36% (*Desire to cut drinking*) to 59% (*Time spent drinking*).[14] Another study found that heritability of individual criteria (in a subset of the data used here) were between 29% (*Tolerance*) and 59% (*Drinking more than intended*).[9] Genomic data also support this variability with Palmer et al reporting a SNP-based heritability ranging from 13% (*Time spent drinking*) to 34% (*Tolerance*).[15] The variability across these estimates likely arises from ascertainment (eg, ascertained for addiction vs twin epidemiologic sample) and the analytic approach (eg, using SNPs vs family relatedness). In addition, in one study, the observed associations with *ADH1B* loci were also differentially attributable to *Tolerance*, *Withdrawal*, *Drinking more than intended*, and *Time spent drinking*, relative to other criteria.[16]

Another strategy to improve the ability to detect variants contributing to DSM-IV AD is to consider the severity of the AD. One approach is to analyze a quantitative variable representing the total number of criteria that a person endorses. Although multiple combinations of criteria and study characteristics may result in a similar criterion count,[17] especially when fewer criteria are endorsed,[18] this proxy for AD severity has been successfully employed in previous studies[8,19] as it makes no assumptions about the cut-off of three or more criteria as an index of "affection status" nor does it equate individuals with 1-2 criteria with those who endorse no criteria during their lifetime.

In this study, we sought to harness the phenotypic richness of the high density alcohol-dependent families recruited as part of the Collaborative Study on the Genetics of Alcoholism (COGA) to perform a series of complementary analyses designed to identify variation contributing to the risk of AD. Our primary GWAS focused on DSM-IV AD diagnosis, a clinically validated measure of pathological drinking that is commonly used in GWAS.[6] We also conducted secondary GWAS of AD severity defined as the count of these criteria (range 0-7), as this quantitative phenotype has been shown to facilitate identification of GWS loci over the binary diagnostic measure of DSM-IV AD. In tertiary analyses, we conducted exploratory GWAS of the seven individual DSM-IV AD criteria, in order to assess which criteria were the most significant contributors to the overall findings observed for DSM-IV AD diagnosis and criterion count, and further, examine whether novel loci emerged for individual criteria. To identify common variants associated with these phenotypes, a GWAS was performed in the European American (EA, n = 1114 families; "EA GWAS") subsample of COGA, followed by a trans-ancestral genome-wide meta-analysis of the EA and African-American (AA; N = 585 families) subsamples. GWS ($P < 5E{-}8$) findings were tested for replication in three independent datasets (Study of Addiction: Genetics and Environment [SAGE],[20] AD GWAS in European and African-Americans (Yale-Penn)[19] and the Australian Twin-family Study of Alcohol Use Disorder (OZALC),[21] which included EA (OZ-ALC, SAGE) and AA (SAGE, Yale-Penn) individuals. Polygenic risk scores (PRS) were created from the COGA EA GWAS and used to predict AD in EAs from SAGE and OZ-ALC. We also performed gene-based analyses using COGA EA GWAS. Lastly, to probe the potential neural correlates of the GWS variants associated with aspects of AD, we tested whether GWS variants identified in the primary (DSM-IV AD), secondary (AD criterion count) or tertiary (individual criteria) analyses were associated with two reward-related neural phenotypes, one within a subset of young individuals from COGA[22] and another within the independent Duke Neurogenetics Study (DNS).[23] The overall design of this study is shown in Figure 1.

2 | MATERIALS AND METHODS

2.1 | Collaborative study on the genetics of alcoholism

Sample: COGA recruited AD probands from inpatient and outpatient AD treatment facilities in seven sites. Community-based families were also recruited from a variety of sources.[24] Institutional review boards from all seven sites approved the study and all participants provided informed consent. COGA participants were administered the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA), a poly-diagnostic interview.[25,26] Individuals below age 18 were administered the child version of the SSAGA, the C-SSAGA. If an individual was interviewed more than once, data from the interview with the maximum total number of endorsed DSM-IV AD criteria were utilized.

Measures: To avoid the inclusion of individuals with high genetic risk who do not drink for personal, social or cultural reasons, only individuals who reported ever drinking at least one full drink of alcohol in their lifetime were included in analyses (EA: N = 7418; AA: N = 3175).

The *primary phenotype* in this study was diagnosis of DSM-IV AD.[2] Individuals meeting criteria for DSM-IV AD at age 15 or older were coded as affected. Individuals were coded as unaffected if they met all of the following criteria: (a) ≥21 years; (b) endorsed <2 criteria for DSM-IV dependence or abuse for alcohol and (c) endorsed <2 criteria for DSM-IV dependence or abuse for cocaine, opioids, marijuana, sedatives and stimulants. Affected individuals <15 years of age and unaffected individuals <21 years of age were excluded. Exclusions for age removed affected individuals with early onset AD who might be etiologically distinct, due to the potentially stronger role of environmental than genetic influences.[27] For unaffected individuals, exclusion of those <21 years of age removed those who may not have passed through the peak period of risk for the onset of AD.[28,29] Due to the strong evidence for shared genetic influences on alcohol and other forms of substance use disorders, individuals who did not meet criteria for AD but endorsed multiple abuse or dependence criteria for other substances were also excluded from the analysis.

The *secondary phenotype* in this study was the sum of endorsed criteria out of the seven DSM-IV AD criteria.

Tertiary phenotypes included each of the seven individual DSM-IV AD criteria. Individuals who drank alcohol but did not endorse that specific criterion were coded as unaffected.

Phenotypic analysis: Tetrachoric correlations (for binary phenotypes) and polychoric correlations (for binary and count phenotypes) were calculated using SAS9.4 (SAS Institute Inc. Cary, NC). We

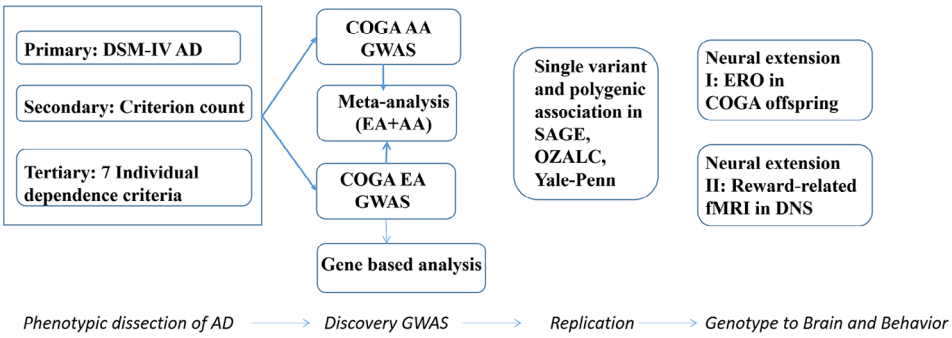

FIGURE 1 Flow chart of analyses

conducted an item response analysis in Mplusv8,[30] using a two-parameter logistic model, to confirm the uni-dimensionality underlying the seven criteria and to examine the discrimination and difficulty associated with each criterion (see Data S1).

2.1.1 | Genotyping, quality review, ancestry and imputation

Four different genome-wide genotyping arrays were used in COGA: (a) COGA case/control data were genotyped on the Illumina Human1M array (Illumina, San Diego, CA) at the Center for Inherited Disease Research (CIDR), Johns Hopkins University[7]; (b) COGA EA family data were genotyped on the Illumina Human OmniExpress 12V1 array (Illumina) at the Genome Technology Access Center, Washington University School of Medicine[9,31]; (c) COGA AA family data were genotyped on the Illumina 2.5M array (Illumina) at CIDR[32]; (d) The remaining samples were genotyped on the Smokescreen genotyping array (Biorealm LLC, Walnut, CA) at Rutgers University. Among these arrays, 2 to 127 samples were genotyped on at least two different arrays with pairwise concordance rates all >99.18%.

A set of 47 000 variants genotyped on all arrays and meeting the following four criteria: common (defined as minor allele frequency [MAF] >10% in the combined sample), independent (defined as R^2 < .5), high quality (missing rate < 2% and Hardy-Weinberg Equilibrium [HWE] *P*-values >.001), were used to assess duplicate samples included on multiple arrays and also to confirm the reported pedigree structure. Family structures were altered as needed, and genotypes were checked for Mendelian inconsistencies using PedCheck[33] with the revised family structure. Genotype inconsistencies were set to missing. The same set of 47 000 variants was also employed to calculate principal components (PCs) using Eigenstrat[34] and 1000 Genomes (Phase 3, version 5). Based on the first two PCs, each individual was then assigned a race classification (AA, EA and Other). To maximize the value of the multiplex family recruitment strategy of COGA, family-based analyses were performed. Families were assigned a family-based race, according to the majority of individual-based race in that family.

All samples were imputed to 1000 Genomes using the cosmopolitan reference panel (Phase 3, version 5, NCBI GRCh37) using SHAPEIT2[35] then Minimac3[36] within each array. Only variants with non-A/T or C/G alleles, missing rates <5%, MAF >3%, and HWE *P*-values >.0001 were used for imputation. Imputed variants with R^2 < .30 were excluded, and genotype probabilities were converted to genotypes if probabilities ≥.90. PedCheck[33] was used again to detect and clean Mendelian inconsistences for imputed variants. All genotyped and imputed variants with missing rates <25%, MAF ≥1% and HWE *P*-values >1E−6 were included in analyses. 8 021 023 and 6 832 792 genotyped and imputed variants passed QC and were included in COGA EA and trans-ancestral (EA + AA) meta-analysis, respectively.

2.1.2 | Genome-wide association studies and meta-analysis

Discovery GWAS were focused on the EA subsample and a trans-ancestral meta-analysis of GWAS summary statistics from the COGA AA and EA subsamples (EA + AA; see Figure 1). Even though a GWAS was conducted in the AA subsample, results were only used in the trans-ancestral meta-analysis. Due to the strict definition of AD controls, the individual AA subsample was too small for use as a discovery sample (both cases and controls had a sample size <1000; full results available upon request). For binary traits, association analysis was performed using a generalized estimating equation framework (with a binomial probability distribution) to control for relatedness with each family treated as a cluster. For the criterion count measure, a linear mixed effects model was fit to continuously distributed data with family relationship adjusted through a kinship matrix. The R package GWAF[37] was used to test both models. Birth cohort (birth year: 1890-1929; 1930-1949; 1950-1969; ≥1970) was a stronger predictor of AD than was age (see also Reference [38]), and hence was selected along with sex, GWAS array indicator, and the first four ancestral PCs (as in a prior study by Wetherill et al[31]) as covariates in the model. In GWS regions, conditional analyses were performed by including the most significant variant in the region as a covariate to evaluate whether a single locus explained the association signal. The trans-ancestral (EA + AA) meta-analysis was performed using inverse-variance weighting in METAL.[39] As implemented in METAL, genomic control, which was estimated by comparing the median test statistics to those expected by chance alone, was applied to the GWAS of COGA AA and COGA EA. For the trans-ancestral meta-analysis (EA + AA), genomic control was applied to the standard errors of the effect sizes. All genomic control estimations were implemented in METAL. Only GWS variants (*P* < 5E−8) were evaluated in replication samples. As we tested seven individual criteria for the tertiary

analyses, a matrix of the phenotypic correlations between these criteria in the EA participants (Table S1B) was spectrally decomposed using matSpD[40,41] resulting in three effectively independent tests and thus a revised GWS *P*-value threshold of 1.67E−8 was used for the tertiary analyses.

2.2 | Replication samples

Three independent datasets from the database of genotypes and phenotypes (dbGaP) were used to replicate significant findings from primary, secondary and tertiary analyses: SAGE (non-overlapping individuals from SAGE, phs000092.v1.p1, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000092.v1.p1), AD GWAS in European and African-Americans (Yale-Penn, phs000425.v1.p1, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000425.v1.p1) and the Australian Twin-family Study of Alcohol Use Disorder (OZALC, phs000181.v1.p1, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000181.v1.p1). Genotypic data from these samples were combined with genotypic data from the COGA samples to identify identical individuals across all datasets; overlapping subjects were retained in the discovery GWAS in COGA but excluded from the replication samples. Ancestry in the combined replication sample was determined in a manner similar to COGA. A similar definition of AD was employed where unaffected individuals with alcohol abuse, or other substance dependence were excluded. The secondary (DSM-IV AD criterion count) and tertiary (individual criteria) phenotypes were also coded in an identical manner. In each replication attempt, only the identical phenotype was tested in the replication cohort (eg, for a variant that was GWS for one criterion but not others, only association with that criterion was tested in the replication samples). Due to the small sizes of the individual AA and EA subsets of the replication datasets, only the AA subsample of SAGE (SAGE-AA), EA subsample of SAGE (SAGE-EA), AA subsample of Yale-Penn (Yale-Penn-AA) and EA subsample of OZALC (OZALC-EA) were included as replication samples. Empirical kinships were estimated from genome-wide genotypic data using the "vcf2kinship" tool as implemented in RVTESTS, then mixed models adjusting for empirical kinships were fitted to the data using RVTESTS.[42] For both SAGE-AA and SAGE-EA, sex and birth cohort (as defined in COGA) were used as covariates, while for OZALC-EA and Yale-Penn-AA, sex and age were used, as in publications of the parent studies. In addition, the first three PCs were included in all replication analyses.

2.3 | Polygenic risk scores analyses

PRS analyses were performed using PRSice-2.[43] EA summary statistics for the primary phenotype, DSM-IV AD, were used to score individuals in SAGE-EA and OZALC-EA datasets. Due to their well-known roles in AD, the alcohol dehydrogenase (ADH) gene cluster on chromosome 4 (99 985 095 bp to 100 430 930 bp) and *ALDH2* on chromosome 12 (112 196 532 bp to 112 276 464 bp) were excluded from PRS analyses to allow for estimation of polygenicity attributable to loci with smaller effects. A set of unrelated individuals was randomly selected from each replication sample (SAGE-EA: N = 1373; OZALC-EA: N = 1441) as required by PRSice-2. Variants located within 500 kb of the index variant and having $r^2 \geq .25$ with the index variant were clumped. PRS were derived by multiplying effect sizes from the EA GWAS of the primary phenotype, DSM-IV AD, with the number of effect alleles in each individual in the target dataset. These product terms were then averaged across the total number of included variants. We only used the *P*-value threshold of $P \leq .05$ (ie, SNPs associated with DSM-IV AD in the discovery EA GWAS at $P \leq .05$) in order to reduce the burden of multiple testing and included the same covariates as those used in replication analyses in each dataset.

2.4 | Gene-based analysis

MAGMA[44], which is implemented in FUMA, a web-based functional mapping and annotation tool[45] was used to perform gene-based analysis. LD was estimated using the European samples from 1000 Genomes projects.

2.5 | Neural extension I: Event-related theta oscillations analysis of GWS loci in COGA Prospective Sample

The COGA Prospective Sample includes offspring aged 12 to 34 years from COGA families, and was designed to assess multiple domains (eg, clinical, neurophysiological), at 2-year intervals.[22] Neurophysiological analyses of reward-related theta ERO (event-related theta oscillation) data from the most recent assessments were carried out in a subsample of 825 COGA AA (49.9% male, 22.12 ± 5.21 years of age) and 1726 COGA EA (48.8% male, 22.26 ± 5.21 years of age) young adults (see Data S1).

A monetary gambling task was implemented as detailed elsewhere.[46] Briefly, individuals bet 50¢ or 10¢ in each of 172 trials, with one of four possible outcomes: lose 50¢, lose 10¢, gain 50¢ or gain 10¢, with equal number of loss and gain trials (Figure S1). Evoked theta ERO power (3.5-7.5 Hz) during monetary loss and gain feedback were measured and differential reward processing ("loss – gain") was derived at frontal, central and parietal regions (Figure S2). Linear regression was applied to test the associations between the top variants and theta ERO power after adjusting for sex, age and first three PCs. We did not examine rs1229984 in *ADH1B* in either the COGA Prospective Sample or the DNS due to its well-known role in the alcohol metabolizing process. For the remaining four GWS loci (rs61826952 and rs7597960 from EA + AA meta-analysis, as well as rs188227250 and rs1912461 from the EA GWAS), three brain regions were tested; therefore, after multiple testing correction, the significance threshold was $P \leq .0042$ (ie, 12 tests). Further details on data acquisition and processing are given in Data S1.

2.6 | Neural extension II: Reward-related functional magnetic resonance imaging analyses of GWS loci in the Duke Neurogenetics Study

We examined whether GWS loci identified in analyses of alcohol-related phenotypes were associated with reward-related brain function among non-Hispanic AA (n = 118; 72% female, 19.6 ± 1.2 years of age) and EA (n = 481; 54.5% female, 19.8 ± 1.2 years of age) undergraduate students who completed the DNS[23] (see Data S1). For rs7597960, which was unavailable in DNS imputed data, we used a proxy SNP, rs2418646, which is in complete LD (ie, r^2 = 1.0, D' = 1.0) within those of African and European ancestries. The chromosome 8 and 15 loci were unavailable in DNS imputed data and no proxies were available; due to their low MAFs, they were difficult to impute in this smaller sample. A number guessing paradigm was used to elicit ventral striatum (VS) reactivity associated with positive and negative feedback linked to monetary gains and losses while blood-oxygen-level dependence functional magnetic resonance imaging (fMRI) data were acquired.[47] Statistical Parametric Mapping version 8 (SPM8 https://www.fil.ion.ucl.ac.uk/spm/software/spm8/) software was used to extract parameter estimates for the contrast of Positive Feedback > Negative Feedback from maximal voxels within left and right VS regions of interest (ROIs). Imaging acquisition protocol, task, ROIs, and preprocessing details are described in Data S1. Extracted parameter estimates from VS activity in each hemisphere were regressed on genotype (rs61826952 coded as 1 or more copies of the minor allele due to sample size; rs2418646 coded using an additive model for the number of C alleles) while covarying for sex, and three (AA) or two (EA) ancestral PCs using Full Information Maximum Likelihood in MPlus v7.3.[48] Trans-ancestral meta-analysis was conducted using METAL.[39] To adjust for multiple comparisons, we used a Bonferroni-corrected *P*-value threshold ($P < .0125$), to account for our hypothesized four tests (ie, rs61826952 and rs2418646 in both brain hemispheres in a trans-ancestral meta-analysis).

3 | RESULTS

3.1 | Phenotypic analyses

Table 1 (primary and secondary phenotypes of DSM-IV AD and criterion count) and Table 2 (tertiary analysis of seven individual criteria) summarize the samples used in discovery and replication analyses. There were 7418 (1114 families) EA and 3175 (585 families) AA individuals, respectively. In total, there were 18 586 individuals evaluated for DSM-IV AD in both discovery and replication samples, with 7482 AD cases and 6169 controls. As shown in Table S1, the primary, secondary and tertiary phenotypes were highly correlated with each other in both EAs and AAs, with DSM-IV AD and DSM-IV AD criterion count having the highest correlations with each individual criterion in both AA and EA subsamples ($r > .87$). As shown in Table S2, the item response analysis demonstrated that all criteria loaded well on a single underlying AD factor. Some criteria discriminated liability at the lower end of the liability distribution (eg, *Drinking more than intended*) while others (eg, *Withdrawal*, *Time spent drinking*, *Giving up activities*) contributed at the higher end of the severity continuum (Data S1).

3.2 | GWAS findings

Regions on chromosomes 1, 2, 4, 8 and 15 reached GWS ($P \leq 5E{-}8$) for primary, secondary and tertiary phenotypes in EA and EA + AA GWAS, respectively (Table 3; Manhattan, quantile-quantile and regional association plots for GWS findings are shown in Figures S3 and S4, respectively; effect sizes, standard errors and *P*-values for EA and AA subsamples and the EA + AA analysis in Table S3). All genomic controls (lambda) are listed in Table S4.

Primary phenotype (*DSM-IV AD diagnosis*): In EA, no GWS findings were identified. In the trans-ancestral meta-analysis (EA + AA), consistent with prior GWAS, rs1229984 in *ADH1B* was significantly associated with AD (P = 1.72E−8). In addition, a novel GWS locus was also identified on chromosome 1 (rs61826952, P = 8.42E−11) in the EA + AA analysis. Both the EA (P = 7.73E−6) and AA (P = 1.50E−07; results available upon request) subsamples contributed to the finding, with the same direction of effect. Conditional analyses confirmed that there were independent associations in the *ADH1B* region but not in the chromosome 1 region (Figure S5A,C).

Secondary phenotype (*DSM-IV AD criterion count*): rs1229984 in *ADH1B* was associated at GWS levels in the EA and the EA + AA analysis.

Tertiary phenotypes (*individual criteria*): In EA, rs1229984 was associated with *Desire to cut drinking* (P = 1.21E−11). Two novel regions were GWS for two individual DSM-IV criteria: rs188227250 on chromosome 8 for *Drinking more than intended* (P = 6.72E−09); rs1912461 on chromosome 15 for *Time spent drinking* (P = 1.77E−08). For the

TABLE 1 Summary of characteristics of COGA and replication datasets

		AA			EA		
Sample		# AD case (% Male)	# AD control (% Male)	# Individuals with DSM-IV criterion (% Male)	# AD case (% Male)	# AD control (% Male)	# Individuals with DSM-IV criterion (% Male)
Discovery	COGA	880 (61.70)	951 (25.45)	3175 (46.58)	2411 (62.01)	2438 (28.47)	7418 (47.53)
Replication	Yale-Penn	1524 (60.50)	485 (29.69)	2010 (53.08)	—	—	—
	SAGE	387 (59.17)	330 (39.09)	930 (46.24)	630 (52.70)	758 (34.17)	1708 (38.82)
	OZALC	—	—	—	1650 (62.24)	1206 (46.10)	3345 (53.69)
Total		2791	1767	6115	4691	4402	12 471

TABLE 2 Summary of samples with individual DSM-IV AD criteria in all datasets

			AA		EA	
DSM-IV AD criterion number	Criterion description	Sample	# Case	# Control	# Case	# Control
1	Tolerance	COGA	1110	2024	3348	3958
		Yale-Penn	1192	818	—	—
		SAGE	353	577	777	930
		OZALC	—	—	2274	1071
2	Withdrawal	COGA	514	2616	1259	6046
		Yale-Penn	694	1316	—	—
		SAGE	200	730	257	1451
		OZALC	—	—	478	2867
3	Drinking more than intended	COGA	1317	1817	3826	3480
		Yale-Penn	1525	485	—	—
		SAGE	507	421	1074	631
		OZALC	—	—	2055	1290
4	Desire to cut drinking	COGA	1436	1701	2896	4413
		Yale-Penn	1411	599	—	—
		SAGE	425	505	601	1107
		OZALC	—	—	1420	1925
5	Giving up activities	COGA	578	2558	1437	5871
		Yale-Penn	1201	809	—	—
		SAGE	215	715	274	1434
		OZALC	—	—	246	3099
6	Time spent drinking	COGA	546	2590	1533	5776
		Yale-Penn	1004	1006	—	—
		SAGE	251	679	354	1354
		OZALC	—	—	668	2677
7	Drinking despite problems	COGA	784	2351	2163	5144
		Yale-Penn	989	1021	—	—
		SAGE	310	619	741	966
		OZALC	—	—	1180	2165

1: *Tolerance*. Need for markedly increased amounts of alcohol to achieve intoxication or desired effect; or markedly diminished effect with continued use of the same amount of alcohol.
2: *Withdrawal*. The characteristic withdrawal syndrome for alcohol or drinking (or using a closely related substance) to relieve or avoid withdrawal symptoms.
3: *Drinking more than intended*. Drinking in larger amounts or over a longer period than intended.
4: *Desire to cut drinking*. Persistent desire or one or more unsuccessful efforts to cut down or control drinking.
5: *Giving up activities*. Important social, occupational, or recreational activities given up or reduced because of drinking.
6: *Time spent drinking*. A great deal of time spent in activities necessary to obtain, to use, or to recover from the effects of drinking.
7: *Drinking despite problems*. Continued drinking despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to be caused or exacerbated by drinking.

trans-ancestral (EA + AA) analysis, rs1229984 was significantly associated with *Desire to cut drinking* (P = 6.01E−14) and *Tolerance* (P = 8.06E−9). An additional GWS region on chromosome 2 (rs7597960, P = 1.22E−8) was noted for *Time spent drinking*. The regions on chromosomes 2, 4 and 8 survived the more stringent correction for the seven criteria ($P \leq$ 1.67E−8) while the chromosome 15 variant was GWS but did not survive the additional correction for multiple testing of individual criteria (ie, P = 1.77E−8). Conditional analyses demonstrated that there was only one association signal in the chromosome 15 region; however, the possibility of a second independent signal in the chromosome 8 region could not be ruled out (P < .001) (Figure S5D,E). Conditional analyses also suggested independent associations in the chromosome 2 region (Figure S5B).

3.3 | Replication

rs1229984 in *ADH1B* was replicated in OZALC-EA for the primary AD phenotype (Table 3); in SAGE-AA for the secondary DSM-IV AD

TABLE 3 Top variants in each region that met genome-wide significance

Phenotype	CHR	BP	Variant	Gene	Minor allele	Major allele	AA MAF	EA MAF	EA *P*-value	COGA AA + EA *P*-value	Yale-Penn-AA *P*-value	SAGE-AA *P*-value	SAGE EA *P*-value	OZALC EA *P*-value	Meta *P*-value	Direction
DSM-IV AD	1	174 637 937	rs61826952	*RABGAP1L*	G	A	0.10	0.07	7.73E−06	**8.42E−11**	0.01	0.13	0.46	0.24	1.66E−04	− − ++ − −
Time spent drinking	2	123 424 651	rs7597960	*TSN,LINC01826*	T	A	0.65	0.21	4.43E−06	**1.22E−08**	0.23	8.94E−03	0.83	0.33	8.86E−05	− − − + − +
DSM-IV AD	4	100 239 319	rs1229984	*ADH1B*	T	C	0.01	0.03	2.09E−07	**1.72E−08**	0.82	0.11	0.54	0.01	**5.30E−09**	− − + − − −
DSM-IV AD criterion count	4	100 239 319	rs1229984	*ADH1B*	T	C	0.01	0.03	**4.16E−11**	**2.61E−13**	0.86	0.01	0.10	0.06	**4.09E−10**	− − + − − −
Desire to cut drinking	4	100 239 319	rs1229984	*ADH1B*	T	C	0.01	0.03	**1.21E−11**	**6.01E−14**	0.82	0.03	4.89E−03	5.52E−03	**9.79E−18**	− − + − − −
Tolerance	4	100 239 319	rs1229984	*ADH1B*	T	C	0.01	0.03	1.89E−07	**8.06E−09**	0.95	7.31E−03	0.32	0.14	**1.20E−09**	− − − − − −
Drinking more than intended	8	127 213 398	rs188227250	*LINC00861, LOC101927657*	A	G	0.02	0.03	**6.72E−09**	1.03E−06	—	—	0.91	1.96E-03	**3.71E−09**	+ − +
Time spent drinking	15	36 344 539	rs1912461	*MIR4510, C15orf41*	C	T	0.02	0.01	**1.77E−08**	1.60E−05	—	—	0.18	0.12	**2.31E−08**	+ + +

Minor allele is defined as minor allele in EA samples. Directions of effect are in terms of minor alleles with + and − meaning increasing or decreasing the chance of having AD or DSM-IV AD criterion, or the number of DSM-IV AD criterion count. Significant P values are in bold. For significant findings in COGA AA + EA, the order of directions are: COGA AA, EA, Yale-Penn-AA, SAGE-AA, SAGE-EA, OZALC-EA; for significant findings in EA, the order of directions are: COGA, EA, SAGE-EA, OZALC-EA.

criterion count as well as for tertiary phenotypes of *Desire to cut drinking* in SAGE-AA, SAGE-EA and OZALC-EA, and in SAGE-AA, for *Tolerance*. Meta-analysis of all available datasets enhanced significance across primary and tertiary phenotypes (Table 3). The association between rs188227250 and *Drinking more than intended* was replicated in OZALC-EA and a meta-analysis of EA, SAGE-EA and OZALC-EA strengthened the association (P = 3.71E−09, Table 3). Although rs1912461 on chromosome 15 was not significantly associated with *Time spent drinking* in either the SAGE-EA or OZALC-EA samples (P > .12), the direction of the effect was the same and meta-analysis across COGA and the replication samples retained significance for this variant (P = 2.31E−08, Table 3). Variants on chromosomes 1 and 2 did not replicate in any dataset (all P > .07 or opposite direction of effects; Table 3).

3.4 | Polygenic risk score analyses

PRS derived using the EA discovery GWAS of the primary phenotype (ie, DSM-IV AD) predicted 1.82% and 0.61% of the variance in AD in SAGE-EA (P = 1.32E−05) and OZALC-EA (P = 7.73E−03), respectively.

3.5 | Gene-based analyses

Table S5 lists the results of gene-based analyses. Two genes, *OTOP1* (P = 8.73E−7) for DSM-IV criterion count, and *BRINP1* (P = 7.85E−8) for *Drinking despite problem*, were GWS.

3.6 | Neural extension I: COGA Prospective Sample: theta ERO

rs1912461 on chromosome 15 for *Time spent drinking* was significantly associated with differential evoked theta power (loss − gain) in the Central ($F_{1,1370}$ = 8.4346; P = .0037) region (Table S6). The minor allele carriers of rs1912461 manifested higher differentiation of gambling outcomes (loss − gain) at the anterior region of the brain (Figure S6). Other variants did not survive the multiple testing correction.

3.7 | Neural extension II: Duke Neurogenetics Study: fMRI

Carriers of the minor (G) allele of rs61826952 had lower left, but not right, reward-related (positive feedback − negative feedback) VS activity when compared to noncarrier individuals in the combined and AA and EA samples (Left: trans-ancestral meta-analysis: β = −.041, P = .008; AA: β = −.124, P = .018; EA: β = −.033, P = .041; Right: trans-ancestral meta-analysis: β = −.01, P = .570). Reward-related VS activity was not significantly associated with rs2418646 genotype (Left: trans-ancestral meta-analysis: β = −.007, P = .560; Right: trans-ancestral meta-analysis: β = .0003, P = .97).

4 | DISCUSSION

This large, family study of AA and EA individuals utilized a multipronged approach (Figure 1) to dissect the genetic underpinnings of AD (DSM-IV AD). In addition to the primary phenotype of DSM-IV diagnosis of AD, and severity as captured by the AD criterion count, it is, to our knowledge, the largest GWAS of each DSM-IV AD criterion. We detected five regions with variants meeting traditional GWS criteria, of which four were novel (chromosomes 1, 2, 8 and 15). Notably, the chromosome 8 signal was replicated in an independent dataset, as was the well-known association with rs1229984 in *ADH1B*. Even when excluding the larger effect size associated with rs1229984, PRS derived from the EA GWAS predicted 0.61% to 1.82% of the variation in AD in independent datasets, underscoring significant polygenicity underlying liability to the disorder. Analyses of two reward-related neural phenotypes also showed associations with two GWS variants.

Consistent with several prior studies,[6] rs1229984 in *ADH1B* was associated with DSM-IV AD. Although GWS was only noted in the trans-ancestral (EA + AA) analysis, as shown in Table S7, rs1229984 was associated with the AD criterion count and criteria indexing physiological dependence and *Desire to cut drinking* at GWS levels, and with other AD criteria at nominal levels of significance. Despite the robust relationship between this functional variant and AD, its relatively low minor allele frequency necessitates fairly large samples to detect a GWS effect for a binary trait, as was shown in a recent meta-analysis of DSM-IV AD.[6] However, for DSM-IV AD criterion count, rs1229984 was GWS in both the EA and EA + AA analyses. Similar to another study,[16] we found that while rs1229984 was associated with each individual criterion (EA all P < 3.61E−04; EA + AA all P < 4.54E−05), the association was stronger with certain DSM-IV AD criteria. Consistent with Hart et al, *Tolerance* was strongly associated with rs1229984 (P = 8.06E−09 in EA + AA). However, the additional GWS associations with *Desire to cut drinking* in our study differs from the prior study which used a sequential regression approach to identify *Withdrawal* and *Drinking more than intended* as additional criteria related to rs1229984 in EA, and *Time spent drinking* in AA. However, another study of 1130 individuals of Jewish descent reported associations between rs1229984 and both *Tolerance* and *Desire to cut drinking*.[49] Across these studies, the most robust association signal for rs1229984 appears to arise from *Tolerance*, which is notably an index of excessive consumption and consistent with the role of *ADH1B* in other studies of nonproblem alcohol intake.[50] Plausibly, the strong findings with *Desire to cut drinking* might also support this as epidemiological studies have shown this criterion to index liability to less severe AD (Table S2 and Figure S7), and therefore, serve as a marker of excessive drinking, rather than severe pathology and impairment.[10,51-53] Differences in associations with other criteria could stem from the relative severity of individual criteria in each dataset or their relationship with excessive drinking.

The GWS findings for the other loci are novel and have not been previously reported for AD or related phenotypes, although these

regions have been linked to some neuropsychiatric diseases/traits. The region on chromosome 1 was previously linked to cerebrospinal fluid biomarker level,[54] migraine,[55] illegal substance dependence,[56] and neuroticism.[57] This region encompasses gene *RABGAP1L*, with many other genes nearby (Figure S4A). *RABGAP1L* is broadly expressed in brain regions and showed association with cerebrospinal fluid biomarker levels[54] and migraine.[55] Other genes near this region seem interesting too, for example, *KIAA0040*, which is downstream of this region, was associated with alcohol dependence.[58] The chromosome 2 region is in a gene desert (Figure S4B) and has been linked to cognitive test scores,[59] attention-deficit/hyperactivity disorder (ADHD) symptom count,[60] ADHD,[61] current smoking[62] and juvenile myoclonic epilepsy.[63] The region on chromosome 8 has been linked to bipolar disorder.[64] The only gene near the chromosome 8 region is *FAM84B* (Figure S4D), however, this gene does not seem to be related to any neuropsychiatric diseases. The chromosome 15 region harbors some noncoding RNAs (Figure S4E) and was previously linked to the rate of cognitive decline,[54] ADHD[65] and major depression.[66] Thus, despite our discovery of novel loci, much further study is needed to investigate the role of these variants in the etiology of AD and related traits.

In our data, the chromosome 1 variant showed nominal association with multiple AD criteria and the criterion count, but none at GWS levels. However, a highly correlated variant (rs1890881) was associated at GWS with a phenotype representing dependence on alcohol or illicit drugs (cannabis, cocaine, sedatives, stimulants, opioids) in the same sample (see accompanying paper by Wetherill et al). It is possible that this variant is associated with overall liability to AD and dependence on other drugs but to a lesser extent with AD severity as indexed by a single continuous criterion count. Research has noted that mere summation does not capture the heterogeneity underlying AD severity, where constellations of criteria could result in meaningful individual differences.[10] Prior latent class analyses aimed to parse out such groups of individuals with unique sets of criteria including in a subset of these data.[9] However, assessment of the genomic underpinnings of such heterogeneous groups of individuals would require extremely large sample sizes. The chromosome 8 variant, rs188227250, was uniquely associated with *Drinking more than intended* (Table S7). In epidemiological studies and in COGA (Table S2), this criterion is endorsed quite frequently by individuals with AD, and also by those who do not meet criteria for DSM-IV AD and thus, might index lower severity. Indeed, in item response theory analyses, this criterion had the lowest difficulty as indicated by the item characteristic curves in Figure S7. In contrast, the finding on chromosomes 2 and 15, while GWS for *Time spent drinking* were also associated with *Giving up activities* (at nearly GWS for chromosome 15), both highly correlated criteria indicative of high difficulty, and thus, risk for DSM-IV AD.[9] In addition to *Withdrawal*, we previously found these criteria to distinguish a highly heritable high-risk group of individuals at risk for AD from those in both low- and moderate-risk groups. Thus, as shown in Figure S7, while the chromosome 8 finding potential maps to lower AD severity, the chromosome 2 and 15 findings potentially indicate greater severity. However, none of these loci were GWS for our AD criterion count measure, which is commonly used as an index of severity. These results are consistent with the argument that the validity of an individual criterion, and its impact on impairment may rely heavily on the other criteria that are endorsed alongside it.[10] Importantly, these results underscore that novel information can be gained from studying individual criteria that index differing levels of AD severity that may operate discontinuously.

Gene-based analysis identified two GWS genes for two different phenotypes. *OTOP1* was associated with DSM-IV criterion count. This gene is related to maintaining metabolic homeostasis but it is not well-studied. *BRINP1* showed association with *Drinking despite problems*. This gene is mostly expressed in brain regions and has been linked to schizophrenia[67,68]; cognition disorders[54] and Parkinson's disease.[69] Further studies are needed to test its role in AD.

Previous studies indicate that AD may be related to variations in the brain's reward system,[70] including decreased reward-network volume[71] and differential neural activity in reward circuitry.[72-74] In the COGA Prospective Sample, minor allele carriers of rs1912461 showed greater differentiation in frontal evoked theta power between loss and gain feedback trials in an EEG-based Monetary Gambling Task. Prior studies have found lower reward-related theta power in alcoholics and in high-risk offspring of alcoholics than controls performing the same task.[46,75] Frontal theta response underlies a variety of cognitive processes[76,77] including reward processing.[78-80] Moreover, it has recently been proposed that frontal theta reflects a promising mechanism through which cognitive control may be enacted by invoking a shift from habitual-based striatum responses to deliberative prefrontal-based control of behavior.[81] Furthermore, the frontal-central theta power difference between loss and gain conditions may reflect the need for cognitive control to process goal-relevant information, such as decision making and action selection, based on choice-relevant information (approach-avoidance, reward-punishment, success-failure, etc.) for optimal functioning in the environment.[81] In this study, the COGA Prospective Sample participants were included in the COGA discovery GWAS. We, therefore, examined the sensitivity of our discovery findings to exclusion of these overlapping individuals from the prospective sample. The resulting GWAS found that while statistical significance decreased in some instances due to the decrease in sample size, the overall results remained highly consistent (eg, for the EA-only finding of *Drinking more than intended*, the *P*-value decreased from 6.72E−09 to 3.61E −08; data not shown), indicating that the overlapping subjects were not solely responsible for the GWS findings from the discovery GWAS.

In the DNS, rs61826952 minor allele carriers had decreased VS activity to positive vs negative feedback in a number-guessing fMRI task. Increased VS activity and dopamine release to non-alcohol reward have been associated with substance use initiation and problematic drinking.[23,82-84] In contrast, studies of AD reported relatively reduced VS activity to non-alcohol reward[85,86] and heightened activity to alcohol cues.[87] These apparently disparate findings can be integrated with stage-based theories of addiction, which hypothesize that initial problematic use is associated with the positive reinforcing

aspects of a substance, while later compulsive use is driven by negative reinforcement and diminished cognitive control, resulting from changes in neural plasticity induced by chronic alcohol use[88] (see also Wetherill et al accompanying paper). Thus, results from the college-based DNS suggested that the minor allele of rs61826952 may protect from AD by reducing VS-related reward drive, thereby diminishing the likelihood of initiating problematic drinking behavior.

Replication of individual variants/genes other than those involved in alcohol metabolism can be challenging and notably influenced by heterogeneity across samples, ascertainment approach, definitions of affected and unaffected, and even nuanced differences in interview instruments.[17] For instance, although families ascertained for AD were included in the replication samples, OZALC had samples ascertained for heavy smoking and drinking (as well as sibships ascertained merely for large pedigree size), and SAGE included two subsamples recruited for nicotine and cocaine dependence. In addition, unlike the prior large AD GWAS by Gelernter et al,[19] we excluded individuals with ≥2 abuse or dependence criteria for alcohol or any illicit drug from our unaffected group. This may have led to a greater degree of genetic separation between affected and unaffected in the current analysis and contributed to the lack of replication. Despite these potential differences, for 2 of the 5 loci (rs1229984 and rs188227250), meta-analyses across samples yielded more significant associations. In addition, the PRS analyses found that the aggregated effect of variants in regions other than the ADH cluster and the *ALDH2* locus significantly contributed to AD liability in these diversely ascertained samples. While the proportion of explained variance is modest, it is consistent with other PRS analyses[89] and supports the generalizability of our findings at a polygenic level.

We also examined whether our analyses supported recent findings of Kranzler et al, who conducted a GWAS of alcohol use disorders defined using ICD codes derived from the electronic health records of individuals participating in the Million Veterans Project.[4] In this multiancestral sample of 274 424 predominantly male veterans, Kranzler et al identified 18 genome-wide significant loci for AUD as well as for the consumption subscale of the Alcohol Use Disorders Identification Testkit (AUDIT-C). Their signal for rs1229984 was also noted in our COGA GWAS. In addition, modest evidence for directional and statistical support was also noted for rs12639940 on chromosome 4 (P = .03; COGA-EA), and rs2961816 on chromosome 5 (P = .04; COGA EA + AA).

Our findings should be considered within the context of a few key limitations. First, despite being large, it is evident that our sample is underpowered to detect loci of modest effect. However, our sample was considerably larger than in our prior efforts in a subset of these data[7-9] and one GWS SNP from those prior studies, previously linked to a latent class representing high-risk for AD,[9] continued to be nominally associated with DSM-IV AD in the current analysis (rs17484734, prior P = 4.1E−8, current P = 8.77E−5) but two other borderline significant variants were not as strongly associated in the current larger sample (rs11035102, for *Desire to cut back*[9]: prior P = 7.3E−8, current P = .002; rs12903120, for AD criterion count[8]: prior P = 5.45E−8, current P = .03). Second, some of our GWS loci had low minor allele frequencies which may also have limited replication efforts. Third, our AA subsample, while utilized in the EA + AA analysis, was too small to report on individually, due to the strict definition of AD affected. Larger discovery GWAS of non-EA samples is much needed.

In summary, our study highlights the importance of utilizing a variety of phenotypes, including individual dependence criteria in locus discovery for AD. The heterogeneity that underlies the diagnosis of AD due to the various combinations of individual criteria that can be endorsed to meet diagnostic criteria, is also true for major depression disorder (MDD), and has been shown to hinder GWAS.[90] While significant increases in sample size can potentially overcome this heterogeneity (as has been shown in the GWAS of MDD[91]), the study of individual criteria, alongside diagnosis and severity, can provide a more detailed characterization of common and specific genetic influences on aspects of AD, especially when viewing individual criteria as psychometric indices of various cut-points of AD liability, and may eventually shape individualized treatment based on criterion profiles and other related features, over and above a mere diagnosis of AD.

ACKNOWLEDGMENTS

COGA: The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators: B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut, includes 11 different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, J. Nurnberger Jr., T. Foroud, Y. Liu); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz); Washington University in St. Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, A. Brooks); 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. Almasy), Virginia Commonwealth University (D. Dick), Icahn School of Medicine at Mount Sinai (A. Goate) and Howard University (R. Taylor). Other COGA collaborators include: L. Bauer (University of Connecticut); J. McClintick, L. Wetherill, X. Xuei, D. Lai, S. O'Connor, M. Plawecki, S. Lourens (Indiana University); G. Chan (University of Iowa; University of Connecticut); J. Meyers, D. Chorlian, C. Kamarajan, A. Pandey, J. Zhang (SUNY Downstate); J.-C. Wang, M. Kapoor, S. Bertelsen (Icahn School of Medicine at Mount Sinai); A. Anokhin, V. McCutcheon, S. Saccone (Washington University); J. Salvatore, F. Aliev, B. Cho (Virginia Commonwealth University) and Mark Kos (University of Texas Rio Grande Valley). A. Parsian and H. Chen 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).

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.

The Duke Neurogenetics Study is supported by Duke University and the National Institutes of Health (NIH) (R01-DA033369).

A.A. receives additional funding support from NIDA (DA032573). A.P.A. receives additional funding support from NIH (HD083614, DA040716, DA038834). R.B. receives additional funding support from NIH (AG052564, HD083614, AG045231). C.E.C. received additional funding support from the National Science Foundation (DGE-1143954) and the Mr and Mrs Spencer T. Olin Fellowship Program. D.M.D. receives additional funding from National Institute on Alcohol Abuse and Alcoholism (K02AA018755). A.R.H. receives additional funding support from NIH (R01-DA031579 and R01-AG049789). J.L.M. receives additional funding support from NIDA (K01DA037914).

A.M.G. is listed as an inventor on issued U.S. Patent 8080371, "Markers for Addiction" covering the use of certain variants in determining the diagnosis, prognosis and treatment of addiction. J.I.N. receives support as an investigator for Janssen. No other authors report biomedical financial interests or potential conflicts of interest.

AUTHOR CONTRIBUTIONS

D.L., L.W., S.B., C.E.C., C.K., M.K., K.K.C., T.R., R.B. and A.A. conducted analyses. D.L., L.W., R.B., A.A. and T.F. drafted and revised the manuscript. T.F., R.B., B.P. and A.A. supervised all aspects of the study. A.R.H. provided feedback on phenotyping and analysis of the DNS sample. J.L.M., A.P.A., D.A.B., K.K.B., P.L.D.J., D.M.D., V.H., J.K., S.K., J.I.N., M.C., D.M.S., R.E.T., J.T., H.J.E. and B.P. provided feedback on development of phenotypes and in statistical analyses. A.P.A., A.R.H., J.L.M., C.K. and B.P. provided support with neural extension I. C.E.C., K.K.C., A.R.H. and R.B. provided support with neural extension II. All authors reviewed and approved the submission of the manuscript.

ORCID

Dongbing Lai https://orcid.org/0000-0001-7803-580X
Leah Wetherill https://orcid.org/0000-0003-2888-9051
Danielle M. Dick https://orcid.org/0000-0002-1636-893X

REFERENCES

1. 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.
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. 4th ed. Washington, DC: American Psychiatric Association Publishing; 1994.
3. Verhulst B, Neale MC, Kendler KS. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med*. 2015;45(5):1061-1072.
4. 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.
5. Hart AB, Kranzler HR. Alcohol dependence genetics: lessons learned from genome-wide association studies (GWAS) and post-GWAS analyses. *Alcohol Clin Exp Res*. 2015;39(8):1312-1327.
6. Walters RK, Polimanti R, Johnson EC, et al. Trans-ancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*. 2018;21:1656-1669.
7. Edenberg HJ, Koller DL, Xuei X, et al. Genome-wide association study of alcohol dependence implicates a region on chromosome 11. *Alcohol Clin Exp Res*. 2010;34(5):840-852.
8. Wang JC, Foroud T, Hinrichs AL, et al. A genome-wide association study of alcohol-dependence symptom counts in extended pedigrees identifies C15orf53. *Mol Psychiatry*. 2013;18(11):1218-1224.
9. Wetherill L, Kapoor M, Agrawal A, et al. Family-based association analysis of alcohol dependence criteria and severity. *Alcohol Clin Exp Res*. 2014;38(2):354-366.
10. Lane SP, Sher KJ. Limits of current approaches to diagnosis severity based on criterion counts: an example with DSM-5 alcohol use disorder. *Clin Psychol Sci*. 2015;3(6):819-835.
11. Johnson EO, van den Bree MB, Pickens RW. Indicators of genetic and environmental influence in alcohol-dependent individuals. *Alcohol Clin Exp Res*. 1996;20(1):67-74.
12. Pickens RW, Svikis DS, McGue M, Lykken DT, Heston LL, Clayton PJ. Heterogeneity in the inheritance of alcoholism. A study of male and female twins. *Arch Gen Psychiatry*. 1991;48(1):19-28.
13. Slutske WS, True WR, Scherrer JF, et al. The heritability of alcoholism symptoms: "indicators of genetic and environmental influence in alcohol-dependent individuals" revisited. *Alcohol Clin Exp Res*. 1999;23(5):759-769.
14. Kendler KS, Aggen SH, Prescott CA, Crabbe J, Neale MC. Evidence for multiple genetic factors underlying the DSM-IV criteria for alcohol dependence. *Mol Psychiatry*. 2012;17(12):1306-1315.
15. Palmer R, Brick L, Chou Y-L, et al. The etiology of DSM-5 alcohol use disorder: evidence of shared and non-shared additive genetic effects. 2019 (in press).
16. Hart AB, Lynch KG, Farrer L, Gelernter J, Kranzler HR. Which alcohol use disorder criteria contribute to the association of ADH1B with alcohol dependence? *Addict Biol*. 2016;21(4):924-938.
17. Lane SP, Steinley D, Sher KJ. Meta-analysis of DSM alcohol use disorder criteria severities: structural consistency is only 'skin deep'. *Psychol Med*. 2016;46(8):1769-1784.
18. Moss HB, Chen CM, Yi HY. DSM-IV criteria endorsement patterns in alcohol dependence: relationship to severity. *Alcohol Clin Exp Res*. 2008;32(2):306-313.
19. Gelernter J, Kranzler HR, Sherva R, et al. Genome-wide association study of alcohol dependence: significant findings in African-and European-Americans including novel risk loci. *Mol Psychiatry*. 2014;19(1):41-49.
20. Bierut LJ, Agrawal A, Bucholz KK, et al. A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci U S A*. 2010;107(11):5082-5087.
21. Heath AC, Whitfield JB, Martin NG, et al. A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry*. 2011;70:513-518.
22. Bucholz KK, McCutcheon VV, Agrawal A, et al. Comparison of parent, peer, psychiatric, and cannabis use influences across stages of offspring alcohol involvement: evidence from the COGA prospective study. *Alcohol Clin Exp Res*. 2017;41(2):359-368.
23. Nikolova YS, Knodt AR, Radtke SR, Hariri AR. Divergent responses of the amygdala and ventral striatum predict stress-related problem drinking in young adults: possible differential markers of affective and impulsive pathways of risk for alcohol use disorder. *Mol Psychiatry*. 2016;21(3):348-356.

24. Reich T, Edenberg HJ, Goate A, et al. Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet*. 1998;81(3):207-215.
25. 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.
26. Hesselbrock M, Easton C, Bucholz KK, Schuckit M, Hesselbrock V. A validity study of the SSAGA—a comparison with the SCAN. *Addiction*. 1999;94(9):1361-1370.
27. 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.
28. 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.
29. Hingson RW, Heeren T, Winter MR. Age at drinking onset and alcohol dependence: age at onset, duration, and severity. *Arch Pediatr Adolesc Med*. 2006;160(7):739-746.
30. Muthen B, Muthen L. *Mplus User's Guide*. 8th ed. Los Angeles, CA: Muthen & Muthen; 2017.
31. Wetherill L, Agrawal A, Kapoor M, et al. Association of substance dependence phenotypes in the COGA sample. *Addict Biol*. 2015;20(3):617-627.
32. Meyers JL, Zhang J, Wang JC, et al. An endophenotype approach to the genetics of alcohol dependence: a genome wide association study of fast beta EEG in families of African ancestry. *Mol Psychiatry*. 2017;22(12):1767-1775.
33. 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.
34. 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.
35. Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype estimation using sequencing reads. *Am J Hum Genet*. 2013;93(4):687-696.
36. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287.
37. Chen MH, Yang Q. GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics*. 2010;26(4):580-581.
38. 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.
39. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191.
40. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)*. 2005;95(3):221-227.
41. 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.
42. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics*. 2016;32(9):1423-1426.
43. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics*. 2015;31(9):1466-1468.
44. De Leeuw CA, Mooij JM, Heskes T & Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*. 2015;11:e1004219.
45. Watanabe K, Taskesen E, Van Bochoven A & Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8:1826.
46. Kamarajan C, Rangaswamy M, Manz N, et al. Topography, power, and current source density of theta oscillations during reward processing as markers for alcohol dependence. *Hum Brain Mapp*. 2012;33(5):1019-1039.
47. Delgado MR, Miller MM, Inati S, Phelps EA. An fMRI study of reward-related probability learning. *Neuroimage*. 2005;24(3):862-873.
48. Muthen LK, Muthen BO. *Mplus User's Guide*. Los Angeles, CA: Muthen & Muthen; 1998 –2011.
49. Kilcoyne B, Shmulewitz D, Meyers JL, et al. Alcohol consumption mediates the relationship between ADH1B and DSM-IV alcohol use disorder and criteria. *J Stud Alcohol Drugs*. 2014;75(4):635-642.
50. Clarke TK, Adams MJ, Davies G, et al. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N = 112 117). *Mol Psychiatry*. 2017;22(10):1376-1384.
51. Saha TD, Stinson FS, Grant BF. The role of alcohol consumption in future classifications of alcohol use disorders. *Drug Alcohol Depend*. 2007;89(1):82-92.
52. Langenbucher JW, Labouvie E, Martin CS, et al. An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *J Abnorm Psychol*. 2004;113(1):72-80.
53. Baillie AJ, Teesson M. Continuous, categorical and mixture models of DSM-IV alcohol and cannabis use disorders in the Australian community. *Addiction*. 2010;105(7):1246-1253.
54. Li QS, Parrado AR, Samtani MN, Narayan VA. Alzheimer's disease neuroimaging I. Variations in the FRA10AC1 fragile site and 15q21 are associated with cerebrospinal fluid Abeta1-42 level. *PLoS One*. 2015;10(8):e0134000.
55. Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet*. 2013;45(8):912-917.
56. Johnson C, Drgon T, Walther D, Uhl GR. Genomic regions identified by overlapping clusters of nominally-positive SNPs from genome-wide studies of alcohol and illegal substance dependence. *PLoS One*. 2011;6(7):e19210.
57. Nagel M, Jansen PR, Stringer S, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet*. 2018;50(7):920-927.
58. Hill SY, Jones BL, Zezza N, Stiffler S. Family-based association analysis of alcohol dependence implicates KIAA0040 on chromosome 1q in multiplex alcohol dependence families. *Open J Genet*. 2013;3(4):243-252.
59. Cirulli ET, Kasperaviciute D, Attix DK, et al. Common genetic variation and performance on standardized cognitive tests. *Eur J Hum Genet*. 2010;18(7):815-820.
60. Lasky-Su J, Neale BM, Franke B, et al. Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *Am J Med Genet B Neuropsychiatr Genet*. 2008;147B(8):1345-1354.
61. Rommelse NN, Arias-Vasquez A, Altink ME, et al. Neuropsychological endophenotype approach to genome-wide linkage analysis identifies susceptibility loci for ADHD on 2q21.1 and 13q12.11. *Am J Hum Genet*. 2008;83(1):99-105.
62. Vink JM, Smit AB, de Geus EJ, et al. Genome-wide association study of smoking initiation and current smoking. *Am J Hum Genet*. 2009;84(3):367-379.
63. Wight JE, Nguyen VH, Medina MT, et al. Chromosome loci vary by juvenile myoclonic epilepsy subsyndromes: linkage and haplotype analysis applied to epilepsy and EEG 3.5-6.0 Hz polyspike waves. *Mol Genet Genomic Med*. 2016;4(2):197-210.
64. Zandi PP, Zollner S, Avramopoulos D, et al. Family-based SNP association study on 8q24 in bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet*. 2008;147B(5):612-618.

65. Mick E, Todorov A, Smalley S, et al. Family-based genome-wide association scan of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2010;49(9):898-905.e893.
66. Rietschel M, Mattheisen M, Frank J, et al. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. *Biol Psychiatry*. 2010;68(6):578-585.
67. Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res*. 2010;124(1–3):192-199.
68. Goes FS, McGrath J, Avramopoulos D, et al. Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am J Med Genet B Neuropsychiatr Genet*. 2015;168(8):649-659.
69. Edwards TL, Scott WK, Almonte C, et al. Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet*. 2010;74(2):97-109.
70. Hommer DW, Bjork JM, Gilman JM. Imaging brain response to reward in addictive disorders. *Ann N Y Acad Sci*. 2011;1216:50-61.
71. Makris N, Oscar-Berman M, Jaffin SK, et al. Decreased volume of the brain reward system in alcoholism. *Biol Psychiatry*. 2008;64(3):192-202.
72. de Greck M, Supady A, Thiemann R, et al. Decreased neural activity in reward circuitry during personal reference in abstinent alcoholics—a fMRI study. *Hum Brain Mapp*. 2009;30(5):1691-1704.
73. Luijten M, Schellekens AF, Kuhn S, Machielse MW, Sescousse G. Disruption of reward processing in addiction: an image-based meta-analysis of functional magnetic resonance imaging studies. *JAMA Psychiat*. 2017;74(4):387-398.
74. Wrase J, Schlagenhauf F, Kienast T, et al. Dysfunction of reward processing correlates with alcohol craving in detoxified alcoholics. *Neuroimage*. 2007;35(2):787-794.
75. Kamarajan C, Pandey AK, Chorlian DB, et al. Deficient event-related theta oscillations in individuals at risk for alcoholism: a study of reward processing and impulsivity features. *PLoS One*. 2015;10(11):e0142659.
76. Basar E, Schurmann M, Sakowitz O. The selectively distributed theta system: functions. *Int J Psychophysiol*. 2001;39(2–3):197-212.
77. Kahana MJ, Seelig D, Madsen JR. Theta returns. *Curr Opin Neurobiol*. 2001;11(6):739-744.
78. Christie GJ, Tata MS. Right frontal cortex generates reward-related theta-band oscillatory activity. *Neuroimage*. 2009;48(2):415-422.
79. Crowley MJ, van Noordt SJ, Wu J, et al. Reward feedback processing in children and adolescents: medial frontal theta oscillations. *Brain Cogn*. 2014;89:79-89.
80. Kamarajan C, Rangaswamy M, Chorlian DB, et al. Theta oscillations during the processing of monetary loss and gain: a perspective on gender and impulsivity. *Brain Res*. 2008;1235:45-62.
81. Cavanagh JF, Frank MJ. Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci*. 2014;18(8):414-421.
82. Heitzeg MM, Villafuerte S, Weiland BJ, et al. Effect of GABRA2 genotype on development of incentive-motivation circuitry in a sample enriched for alcoholism risk. *Neuropsychopharmacology*. 2014;39(13):3077-3086.
83. Stuke H, Gutwinski S, Wiers CE, et al. To drink or not to drink: harmful drinking is associated with hyperactivation of reward areas rather than hypoactivation of control areas in men. *J Psychiatry Neurosci*. 2016;41(3):E24-E36.
84. Weiland BJ, Zucker RA, Zubieta JK, Heitzeg MM. Striatal dopaminergic reward response relates to age of first drunkenness and feedback response in at-risk youth. *Addict Biol*. 2017;22(2):502-512.
85. Balodis IM, Potenza MN. Anticipatory reward processing in addicted populations: a focus on the monetary incentive delay task. *Biol Psychiatry*. 2015;77(5):434-444.
86. Beck A, Schlagenhauf F, Wustenberg T, et al. Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. *Biol Psychiatry*. 2009;66(8):734-742.
87. Kareken DA, Claus ED, Sabri M, et al. Alcohol-related olfactory cues activate the nucleus accumbens and ventral tegmental area in high-risk drinkers: preliminary findings. *Alcohol Clin Exp Res*. 2004;28(4):550-557.
88. Koob GF, Le Moal M. Plasticity of reward neurocircuitry and the 'dark side' of drug addiction. *Nat Neurosci*. 2005;8(11):1442-1444.
89. Bogdan R, Baranger DAA, Agrawal A. Polygenic risk scores in clinical psychology: bridging genomic risk to individual differences. *Annu Rev Clin Psychol*. 2018;14:119-157.
90. Wray NR, Pergadia ML, Blackwood DH, et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol Psychiatry*. 2012;17(1):36-48.
91. Wray NR, Ripke S, Mattheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*. 2018;50(5):668-681.

SUPPORTING INFORMATION

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

How to cite this article: Lai D, Wetherill L, Bertelsen S, et al. Genome-wide association studies of alcohol dependence, DSM-IV criterion count and individual criteria. *Genes, Brain and Behavior*. 2019;1–14. https://doi.org/10.1111/gbb.12579