



Assessing the Role of Long Noncoding RNA in Nucleus Accumbens in Subjects With Alcohol Dependence

John Drake , Gowon O. McMichael, Eric Sean Vornholt, Kellen Cresswell, Vernell Williamson, Chris Chatzinakos, Mohammed Mamdani, Siddharth Hariharan, Kenneth S. Kendler, Gursharan Kalsi, Brien P. Riley , Mikhail Dozmorov, Michael F. Miles, Silviu-Alin Bacanu, and Vladimir I. Vladimirov

Background: Long noncoding RNA (lncRNA) have been implicated in the etiology of alcohol use. Since lncRNA provide another layer of complexity to the transcriptome, assessing their expression in the brain is the first critical step toward understanding lncRNA functions in alcohol use and addiction. Thus, we sought to profile lncRNA expression in the nucleus accumbens (NAc) in a large postmortem alcohol brain sample.

Methods: lncRNA and protein-coding gene (PCG) expressions in the NAc from 41 subjects with alcohol dependence (AD) and 41 controls were assessed via a regression model. Weighted gene coexpression network analysis was used to identify lncRNA and PCG networks (i.e., modules) significantly correlated with AD. Within the significant modules, key network genes (i.e., hubs) were also identified. The lncRNA and PCG hubs were correlated via Pearson correlations to elucidate the potential biological functions of lncRNA. The lncRNA and PCG hubs were further integrated with GWAS data to identify expression quantitative trait loci (eQTL).

Results: At Bonferroni adj. p -value ≤ 0.05 , we identified 19 lncRNA and 5 PCG significant modules, which were enriched for neuronal and immune-related processes. In these modules, we further identified 86 and 315 PCG and lncRNA hubs, respectively. At false discovery rate (FDR) of 10%, the correlation analyses between the lncRNA and PCG hubs revealed 3,125 positive and 1,860 negative correlations. Integration of hubs with genotype data identified 243 eQTLs affecting the expression of 39 and 204 PCG and lncRNA hubs, respectively.

Conclusions: Our study identified lncRNA and gene networks significantly associated with AD in the NAc, coordinated lncRNA and mRNA coexpression changes, highlighting potentially regulatory functions for the lncRNA, and our genetic (cis-eQTL) analysis provides novel insights into the etiological mechanisms of AD.

Key Words: long noncoding RNA, gene expression, eQTLs, postmortem brain, nucleus accumbens.

ALCOHOL USE DISORDER (AUD) is a chronic and debilitating disease, with an estimated heritability of around 50% (Verhulst et al., 2015). Previous postmortem brain expression studies have shown that chronic alcohol consumption leads to broad transcriptional changes in different brain regions (Warden and Mayfield, 2017). Gene

From the Center for Integrative Life Sciences Education (JD), Virginia Commonwealth University, Richmond, Virginia; Virginia Institute for Psychiatric and Behavioral Genetics, (GOM, ESV, CC, MM, KSK, BPR, MFM, S-AB, VIV), Virginia Commonwealth University, Richmond, Virginia; Department of Biostatistics, (KC, MD), Virginia Commonwealth University, Richmond, Virginia; Department of Pathology, (VW), Virginia Commonwealth University, Richmond, Virginia; Summer Research Fellowship, (SH), School of Medicine, Virginia Commonwealth University, Richmond, Virginia; Department of Psychiatry, (KSK, BPR, S-AB, VIV), Virginia Commonwealth University, Richmond, Virginia; Department of Human and Molecular Genetics, (KSK, BPR), Virginia Commonwealth University, Richmond, Virginia; Department of Social, Genetic and Developmental Psychiatry, (GK), Institute of Psychiatry, London, UK; Department of Pharmacology and Toxicology, (MFM), Virginia Commonwealth University, Richmond, Virginia; Center for Biomarker Research and Personalized Medicine, (VIV), Virginia Commonwealth University, Richmond, Virginia; and Lieber Institute for Brain Development, (VIV), Johns Hopkins University, Baltimore, Maryland.

Received for publication April 2, 2020; accepted October 1, 2020.

Reprint requests: John Drake, MS, Center for Life Sciences Education, Virginia Commonwealth University, 800 E. Leigh Street, Suite 100, Box 980126, Richmond, VA 23298-0126; Tel.: (804) 787-0156; Fax: (804) 828-1471; E-mail: jdrake3@mymail.vcu.edu and

Vladimir I. Vladimirov, MD, PhD, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, 800 E. Leigh Street, Suite 100, Box 980126, Richmond, VA 23298-0126; Tel.: (804) 828-3594; Fax: (804) 828-1471; E-mail: vladimir.vladimirov@vcuhealth.org

© 2020 The Authors. Alcoholism: Clinical & Experimental Research published by Wiley Periodicals LLC on behalf of Research Society on Alcoholism.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1111/acer.14479

expression studies in the prefrontal cortex have identified genes encoding to GABA_A receptor subunits or related to mitochondrial function (Buckley et al., 2000; Fan et al., 1999) as well as in functions related to myelination, cell cycling, oxidative stress, and transcription (Farris et al., 2015; Flatscher-Bader et al., 2005, 2006; Iwamoto et al., 2004; Kapoor et al., 2019; Lewohl et al., 2000; Mayfield et al., 2002; Ponomarev et al., 2012; Rao et al., 2019; Sutherland et al., 2014; Zhang et al., 2014). Expression studies in nucleus accumbens (NAc) and ventral tegmental area have also revealed gene expression changes related to cell architecture, signaling, vesicle formation, and synaptic transmission (Flatscher-Bader et al., 2006, 2008, 2010; Mamdani et al., 2015; Rao et al., 2019). These findings suggest that there are region-specific susceptibilities and adaptations to chronic alcohol consumption in the brain that are likely to have a distinct effect on the behavioral phenotypes comprising AUD (Flatscher-Bader et al., 2010).

Evaluation of the regulatory mechanisms underlying genetic differentiation is necessary for a better understanding of the neurobiology of alcohol addiction (Ron and Messing, 2013). The recent emergence of Long noncoding RNA (lncRNA) has provided an additional layer of transcriptional and translational control, highlighting potentially important neurobiological mechanisms of AUD that could be missed if only the protein-coding genome was studied. lncRNAs are longer than 200 base pairs with limited or no protein-coding potential (Jarroux et al., 2017), and while functionally not fully characterized yet, they have been shown to participate in chromatin remodeling (Han and Chang, 2015), transcriptional and posttranscriptional regulation (Clark and Blackshaw, 2014), and in sequestering miRNA (Du et al., 2016; Shan et al., 2018). lncRNAs have also been implicated in the plasticity of neuronal circuitry (Wang et al., 2017) and neurodegenerative and neuropsychiatric disorders, including substance abuse and alcohol dependence (AD; Bohnsack et al., 2019; Zuo et al., 2016). The potential role of lncRNA in the etiology of AUD is further supported by several recent alcohol-related genomewide association scans (GWAS; Adkins et al., 2017; Xu et al., 2015). Due to the complex nature of lncRNA functions, the relationship between lncRNA and alcohol consumption is largely unknown; while miRNA roles in the control of gene expression and functions in the brains of subjects with AUD have been reported (Li et al., 2013; Mamdani et al., 2015; Mizuo et al., 2012; Pietrzykowski et al., 2008; Yadav et al., 2011), to our knowledge the interactions between lncRNA and protein-coding gene (PCG) are still understudied.

Gene expression alone cannot explain the complex etiology of AUD, and assessing lncRNA and PCG expression in the context of available genetic data is necessary to discern the genetic basis of AUD susceptibility. This approach models associations between genetic variants and gene expression as quantitative traits, that is, expression quantitative trait loci (eQTL; Farris et al., 2010; Li, 2013; Nica and Dermitzakis, 2013). eQTLs can help discover unknown AUD risk loci or

offer specific, testable, hypotheses for the genetic impact of polymorphisms associated with AUD (Hindorff et al., 2009; Liu, 2011; Shastry, 2009). Linkage disequilibrium (LD) of eQTLs with genetic variants implicated in AUD can further establish a biological mechanism for disease-associated variants with no apparent functions; there is empirical evidence suggesting that eQTLs are overrepresented among GWAS signals (Nica et al., 2010; Nicolae et al., 2010).

We hypothesize that lncRNAs are involved in the neuropathology of AUD and that, on a molecular level, this involvement is manifested by dysregulated patterns of expression between cases with AD and controls. To further elucidate the lncRNA disease functions in subjects with AD, we performed a weighted gene coexpression network analysis (WGCNA) that led to the identification of lncRNA and PCG modules significantly correlated with AD. Key genes from these lncRNA and PCG modules (i.e., “hubs”) were then integrated to identify a set of interacting lncRNA and PCG. Since lncRNAs are not yet annotated, we used system approaches to assign biological functions for the PCG hubs interacting with the lncRNA hubs. By integrating our genomewide SNP and expression data, we also identified eQTL affecting the expression of lncRNA and PCG in NAc. Thus, the overall goals of this study were first to profile the lncRNA and PCG expressions in NAc, identify protein-coding and lncRNA networks, and finally test whether these are under the control of specific genetic elements. To our knowledge, this is the only study to comprehensively assess the lncRNA expression in NAc between subjects with AD and controls.

MATERIALS AND METHODS

Postmortem Tissue

Brain tissue from 41 cases with AD and 41 controls was received from the Australian Brain Donor Program, New South Wales Tissue Resource Centre, which is supported by The University of Sydney, National Health and Medical Research Council of Australia, Schizophrenia Research Institute, National Institute of Alcohol Abuse and Alcoholism, and the New South Wales Department of Health (<http://sydney.edu.au/medicine/pathology/trc/>; Supplemental Methods; Table S1).

RNA Isolation and Sample Selection

Total RNA was isolated from 50 mg of frozen tissue from NAc using the mirVana-PARIS kit (Thermo Fisher, Carlsbad, CA), following manufacturer's protocols. RNA concentration was measured using the Quant-iT Broad Range RNA Assay kit (Life Technologies), and the RNA integrity number (RIN) was measured on the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA).

Expression Arrays and Data Normalization

The RNA samples were assayed using the Arraystar Human lncRNA Array v3.0 (Rockville, MD) that is designed to profile both 30,586 lncRNAs and 26,109 PCGs. Linear Models for Microarray Data (limma) v3.42 (Ritchie et al., 2015) package in R environment (v3.6.1) was used to log₂ transform, background correct (method = “normexp”), and normalize (method = “quantile”) each

microarray. Control probes were removed prior to downstream analysis, leaving 61,464 probes remaining. As each probe recognizes a specific and unique transcript, probes were used in all subsequent analyses. The final number of probes that were used in the subsequent analyses was 61,464. The efficiency of batch effect removal and overall array quality were assessed using principal components analysis on the expression values in which each array is plotted along the first 3 principal components (PCs) to identify potential outliers. Of the 82 samples, 9 samples were excluded due to low RINs (i.e., RINs ≤ 4), leaving 73 samples for the microarray run. Of the 73 samples run on the array, 8 samples (5 cases and 3 controls) did not load on 2 of the first 3 PCs and were removed from the subsequent analysis. Finally, 2 subjects lacked corresponding genotypic data and were excluded leaving a final sample of 63 subjects for all subsequent analyses.

Statistical Analyses

Differential Expression Analysis. Differential expression analysis was performed by limma using ordinary least square. Number Cruncher Statistical Software (NCSS) v12 was used to extract the first 3 PCs for all technical (e.g., Batch, RIN, PMI, pH) and biological (e.g., sex, smoking, age, brain weight, hemisphere, neuropathology) confounds, which were then used as covariates in the subsequent analysis (Mamdani et al., 2015). Microarray reliability was validated by assessing the expression of 5 genes at the Arraystar facilities using quantitative PCR (qPCR). The assessed genes showed a high mean correlation (Pearson $r = 0.88$ [SD ± 0.049]) between the 2 platforms (Supplemental Methods and Table S2).

WGCNA. WGCNA has been heavily applied to cancer phenotypes, which are associated with dramatic changes to the transcriptome (Gibb et al., 2011; Peng et al., 2015; Uhlen et al., 2016). Given that the effect sizes for neuropsychiatric and addiction disorders are substantially lower, preselection of genes to be included in WGCNA is necessary for the identification of propitious targets without violating the assumption of scale-free topology. Thus, a relaxed nominal significance for differentially expressed genes (i.e., p -value ≤ 0.10) was chosen to: (i) include genes with smaller effect sizes, albeit true positive signals, (ii) exclude genes with low disease variance, and (iii) provide a sufficient number of genes for the network analysis, which was performed separately on the differentially expressed lncRNA and PCG using the WGCNA v1.68 package in R environment v3.6.1. While WGCNA allows for the inclusion of additional variables, it does not correct for their effects; therefore, the gene networks were built using the residuals from the regression model, that is, fitting a regression model with all potential confounds, except disease status. WGCNA relies on pairwise Pearson correlations to generate a signed similarity matrix, selecting for positive correlations only. The signed similarity matrix of the lncRNA and PCG expression data was raised to the lowest power ($\beta = 5$ and $\beta = 10$, respectively) that approximated a scale-free network topology ($R^2 > 0.85$), to generate an adjacency matrix (Fig. S1A,B). A topological overlap measure (TOM) was calculated to assess transcript interconnectedness. A dissimilarity measure was derived from TOM and was subsequently used for average linkage hierarchical clustering. The clustered genes were defined into modules using the “tree” method with a minimum module size of 25 genes and a minimum module merge height of 0.8.

Following module definition, the first principal component of each module—the module eigengene (ME)—was calculated as a synthetic gene representing the expression profile of all genes within a given module. With the exception of the gray module, which contains genes poorly defined by other modules, modules are arbitrary named after a conventional color scheme. After naming, modules are correlated to AD case status, matching demographics, and relevant covariates. Statistical significance was assessed at Bonferroni adj. $p \leq 0.05$ (corrected for the number of tested modules).

As previously outlined, for the identification of genes with a high module membership (MM) and strong phenotype association (i.e., the hub genes), we applied 2 selection criteria: (i) genes with the high intramodular connectivity ($r^2 \geq 0.7$) and (ii) significantly correlated with AD (i.e., p -value ≤ 0.05 ; Mamdani et al., 2015; Ponomarev et al., 2012).

Weighted Key Driver Analysis. As lncRNA genes can act locally, near their transcription site, and mechanically heterogeneous, there was concern that some lncRNAs would build smaller modules, which may not be captured by WGCNA (Marchese et al., 2017). To overcome this limitation, we used the weighted key driver analysis (WKDA) from the mergeomics v1.14.0 R package (Shu et al., 2016). WKDA identifies “key drivers,” which are genes whose local connectivity to neighboring genes of prespecified modules (i.e., modules significantly correlated with AD) is greater than what would be expected by chance. This connectivity score takes into account the topology and edge weights of a gene to its neighbors, which was generated by WGCNA as a TOM matrix. As the TOM matrix topology is complete, the edgefactor parameter was set to 1 to maximize the influence of the edge weights. Key drivers, which had an FDR of 5%, were identified as hub genes and merged with those found by WGCNA.

Gene Set Enrichment Analysis. Prior to running the enrichment analysis, all transcript IDs from the PCG modules significantly correlated with AD case status were converted to HUGO Gene Nomenclature Committee (HGNC; Gray et al., 2013). Our gene enrichment analysis was based on 2 approaches: (i) a brain-related enrichment and (ii) enrichment for specific gene ontology terms. In the brain-related enrichment, we created a curated list derived from both of WGCNA’s brain-related cell-type lists, PsychENCODE’s brain cell-type marker genes, and a neural gene set comprised of physiological and behavioral marker genes (Schrode et al., 2019; Wang et al., 2018). The significance of the brain-related enrichment was estimated via a hypergeometric test using WGCNA’s “userListEnrichment” function with a significant enrichment declared at FDR of 10%.

To detect a broader enrichment for known biological processes and pathways, we used the gene set enrichment analysis (GSEA) software (v4.0.1) from the Broad Institute as previously described (Hung et al., 2012; Subramanian et al., 2005). Due to the poor lncRNA annotation, in this analysis, the individual genes from each of the significant PCG modules only were generated by rank-ordering all genes by their MM to each of the significant AD modules. We derived the a priori gene sets from the Molecular Signatures Database v7.0 (MSigDB; <http://www.broadinstitute.org/gsea/msigdb>) from the Broad Institute. A total of 5,501 gene sets from the Canonical Pathways subset of the C2: Curated Pathways collection of MSigDB were assessed. Default parameters were then applied to give a minimum, and maximum a priori gene set size between 15 and 500 genes, respectively. The identified enriched pathways of both approaches were further adjusted for multiple testing at FDR of 10%.

eQTL Detection. The genotype calls were generated as part of a larger meta-GWAS study (Adkins et al., 2017) and were integrated with the lncRNA and PCG hubs expression values to identify eQTLs for the lncRNA and PCG. To reliably estimate the eQTL effects, we excluded SNPs with missingness of $\geq 20\%$ and a minor allele frequency of $\leq 10\%$, thus retaining a final set of 3,494,304 SNPs for the eQTL analysis. As previously outlined and to maintain power, we focused on using only local, cis-eQTLs (Mamdani et al., 2015). Specifically, SNPs located 500 kilobase pairs (kbp) upstream and downstream from the selected lncRNA and PCG hubs (i.e., cis-eQTL) were extracted and filtered with Plink v1.90 to exclude variants in LD ($R^2 \geq 0.7$, window size = 250 kbp, step size = 25 kbp;

Purcell et al., 2007). eQTL impact on gene expression was detected by MatrixEQTL software package in R using a linear regression framework and adjusting for the potential effects of technical and biological covariates as summarized by the first 3 PCs (Shabaln, 2012). The significant eQTL results were adjusted for multiple testing at FDR of 10%. Visualization of eQTLs and subsequent integration analysis was conducted utilizing ggplot2 package in R environment v3.6.1 (Wickham, 2016).

Integration Analyses. To increase the robustness of our integration analyses, we followed a 2-pronged approach. First, the MEs from the significant lncRNA and PCG modules were correlated using Pearson product-moment correlations, and only the significant module correlations at FDR of 10% were selected for consideration. In the second round, all individual hubs from the significant PCG and lncRNA module correlations were extracted and correlated to identify the significant individual mRNA/lncRNA pairs. All pairs surviving this second round of correction for multiple testing at an FDR of 10% were then selected for additional series of tests.

RESULTS

Chronic Alcohol Consumption Leads to Generalized Gene Expression Changes Between Cases and Controls

At the single gene expression analysis ($p \leq 0.10$), we identified a total of 6,469 probes differentially expressed between AD cases and controls, representing 3,645 lncRNA and 2,725 PCG that comprise 5.9% and 4.4% of the total gene pool assayed, respectively. Among the differentially expressed genes, we also identified 99 pseudogenes. The complete list of significant results from the univariate gene expression analysis is provided in the supplementary table (Table S3).

lncRNA and PCG Show a Disease Relevant Network Pattern

While the gene expression analysis identifies differentially expressed genes in NAc of subjects with AD, these do not provide an integrative, systemic view of the interaction between the lncRNA and mRNA genes. Therefore, to better understand the higher-order (system biology impact) of the interactions between lncRNA and PCG, we performed a WGCNA (Langfelder and Horvath, 2008; Zhang and Horvath, 2005). WGCNA identified 36 lncRNA and 12 mRNA modules (Fig. 1A,B). In the lncRNA network analysis, we identified 19 modules (M_{brown} , M_{sienna3} , $M_{\text{steelblue}}$, M_{black} , $M_{\text{darkmagenta}}$, M_{tan} , M_{violet} , M_{plum1} , M_{skyblue} , M_{salmon} , $M_{\text{lightyellow}}$, $M_{\text{saddlebrown}}$, $M_{\text{darkorange}}$, M_{green} , $M_{\text{darkgreen}}$, $M_{\text{darkolivegreen}}$, M_{cyan} , M_{orange} , and M_{gray}) significantly correlated with AD at a Bonferroni adj. $p \leq 0.001$ (Fig. 2), with M_{black} being the largest containing 200 genes. In the PCG network analysis, we identified 5 modules (M_{gray60} , M_{green} , M_{red} , M_{black} , M_{gray}) significantly correlated with AD at a Bonferroni adj. $p \leq 0.004$ (Fig. 3), with M_{black} being the largest containing 407 genes. In addition to the primary dichotomous case/control status, we also tested additional quantitative phenotypes such as the age of initiation of alcohol drinking, the amount of daily alcohol consumption, and years of drinking. We identified only 2 PCG modules (M_{gray}

and M_{red}) as significantly correlated with the amount of daily alcohol consumption at a Bonferroni adj. p -value ≤ 0.004 . Full tables with module size, correlations, and p -values for all PCG and lncRNA modules correlated with AD are provided in supplementary tables (Tables S4A,B).

To ensure the robustness of our gene coexpression networks and reduce the potential influence of outlier samples on network structure, we used the robust “bootstrapped” version of WGCNA (rWGNCA). We performed 100 iterations, that is, generated 100 networks after randomly subsetting two-third of the total samples, as previously suggested (Gandal et al., 2018) with the resulting networks then merged into 1 large, final consensus network. The individual subnetworks showed reasonably high consistency with the final lncRNA and mRNA networks (Fig. S2A,B).

In scale-free network topology, “hubs” are the most highly connected genes (of which there are relatively few among all the nodes within a network). Hubs were identified by examining modules that were significantly associated with disease status and extracting genes with a high MM ($\text{MM} \geq 0.70$) and a significant correlation with disease status ($p \leq 0.05$). Among the significant PCG and lncRNA modules, we identified a total of 86 and 20 hubs, respectively. The smaller number of detected lncRNA hubs could be due to the lower expression values that are typical for lncRNA genes, which may lead to the formation of weaker network structures. Thus, we applied the wKDA analysis to increase our power to identify a larger set of lncRNA hubs based on their connectivity to neighboring genes within a module. At FDR of 5%, we identified 295 lncRNA hubs significantly correlated with AD ($p \leq 0.05$), which, in the wKDA analysis, are also designated as key drivers (Table S5). The 295 lncRNA hubs were then combined with the initially identified 20 hubs for a total of 315 lncRNA hubs.

AD Gene Modules Are Enriched in Alcohol-Related Processes

Previous studies have shown that coexpressed genes are enriched for specific cell types and biologically relevant functions (Guttman and Rinn, 2012; Long et al., 2017; Mamdani et al., 2015; Marty and Spigelman, 2012). As lncRNA are poorly annotated for the enrichment analyses, we focused on the PCG modules significantly associated with AD. The brain-related enrichment analysis was conducted in WGCNA using a curated list of brain-related gene markers. At an FDR of 10%, 3 separate modules (M_{black} , M_{red} , and M_{green}) showed significant brain-related enrichment (Table S6). The M_{black} module was enriched for microglia, while the M_{red} module was enriched for gene markers specific to the Raphe nuclei. The M_{green} module was enriched for cell attributes such as postsynaptic membrane and glutamatergic synapse. Using the default parameters in GSEA, at FDR of 10%, we identified 111 a priori gene sets significantly enriched in the M_{black} module and 1 a priori gene set significantly enriched in the M_{green} module (Table S7). Specifically, M_{black} was enriched for genes up-regulated in basal neuron

processes (i.e., dendritic maturation) and immune-related systems (i.e., innate immunity, interferon signaling, etc.).

lncRNA and PCG Show a Complex Pattern of Interactions

Since lncRNA can function as potential molecular “scaffolds” (Guttman et al., 2011; Guttman and Rinn, 2012), that is, by bringing different sets of genes near each other, we were interested in identifying PCG hubs significantly correlated with lncRNA hubs. Limited studies have shown that lncRNA can affect gene functions by interfering with the expression of their specific gene targets (Long et al., 2017). For maximum power, the MEs from the significant AD modules were correlated with each other (Fig. S3), and all hubs within the significantly correlated PCG and lncRNA modules were extracted. At FDR of 10%, we observed 4,985 correlations, of which the positive ($N = 3,125$) significantly exceeded the negative ($N = 1,860$) correlations (Mann–Whitney U -test $p = 2.2E-16$). The strongest positive correlation (Pearson $r = 0.95$, $q = 8.34E-27$) was observed between RP11-23J18.1 in the $M_{\text{steelblue}}$ lncRNA module and IFITM1 in the mRNA M_{black} module, while the strongest negative

correlation (Pearson $r = 0.80$, $q = 6.52E-12$) was observed between RP11-356J5.12 in the $M_{\text{lightyellow}}$ lncRNA module and YBX3 in the mRNA M_{black} module. The entire list of positive and negative lncRNA/PCG hub correlations is provided in the supplementary table (Table S8A,B).

Genetic Variants Affect lncRNA and PCG Expression in a Disease-Specific Manner

To understand the underlying genetic mechanisms of AD risk polymorphisms, we further integrated the lncRNA and PCG hub expression data from our significant modules with previously GWAS data collected on the sample (Adkins et al., 2017). To maintain sufficient power, we focused on testing cis-eQTLs only. At FDR of 10%, we identified a total of 150 eQTLs impacting hubs’ expression, with a higher number of eQTLs for the lncRNA ($n = 135$) versus PCG ($n = 15$) hubs (Table S9A,B). The most significant PCG eQTL was between Interferon regulatory factor 7 (IRF7) and rs11246386 (Fig. 4). The most significant lncRNA eQTL was between AK128400 and rs1669681 (Fig. 5). In a follow-up analysis, to identify potential disease risk eQTLs, we further tested the significant lncRNA and PCG eQTLs for an interaction (SNP \times AD) term between genotype and AD status. A significant genotype/disease interaction for a SNP/gene pair indicates that the effect of genotype on expression is significantly different in AD cases versus controls. At FDR of 10%, we identified 93 eQTL with significant (SNP \times AD) interaction terms that mediate the expression of 24 PCG hubs and 69 lncRNA hubs (Table S10A,B). The most significantly interacting PCG eQTL was between FK506 binding protein 51 (FKBP5) and rs2766532 (Fig. 6A,B), whereas the most significantly interacting lncRNA eQTL was between G006838 and rs12142153 (Fig. 7A,B). We further tested for enrichment of our 243 significant eQTLs among the significant GWAS hits of addiction phenotypes via the Cauchy and Simes test. These are powerful, mutually complementing, tests that compute the p -values to control for type I error due to background enrichment, and are appropriate for assessing eQTL enrichment in GWAS significant signals with either numerous and smaller (i.e., Cauchy) or fewer and larger (i.e., Simes) effect sizes (Suppl. Methods). We observed our eQTLs to be significantly enriched for cigarettes per day [Cauchy; p -value = 0.036, Simes; p -value = 0.047] and smoking initiation [Simes; p -value = 0.031]. We further observed a marginal significance for smoking cessation [Cauchy; p -value = 0.072, Simes; p -value = 0.073] and drinks per week [Cauchy; p -value = 0.099]. Finally, as lncRNA and mRNA are capable of interacting with each other, we tested whether any of the significant eQTLs are also capable of modulating the lncRNA and PCG interaction. We applied this test to the negatively interacting lncRNA/PCG pairs, and while we observed nominally significant genotype effects modulating lncRNA and PCG interactions (Table S11), none of these achieved statistical significance at FDR of 10%, most likely due to the limited sample size.

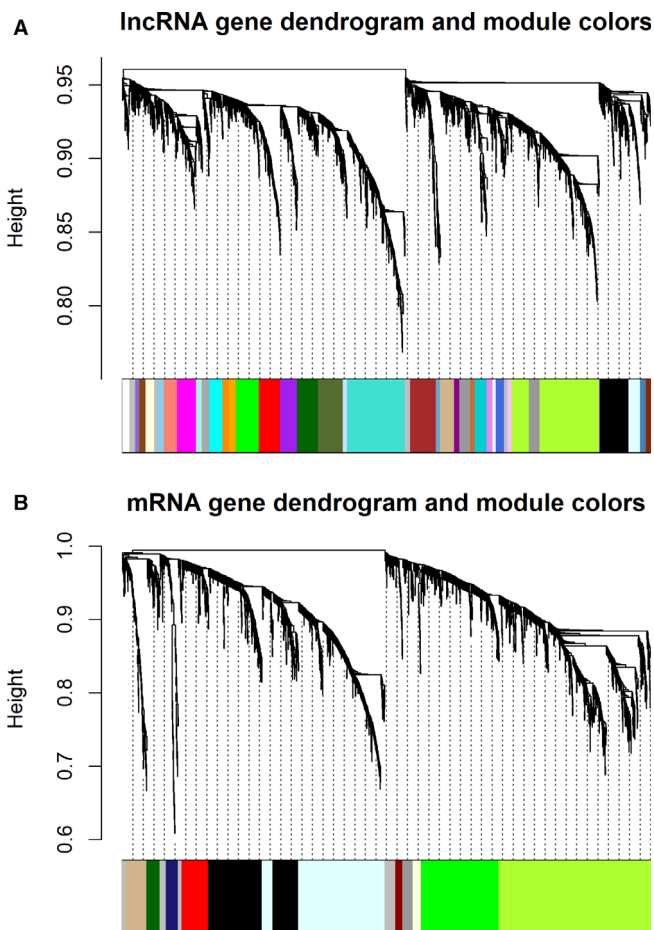


Fig. 1. Gene dendrogram of lncRNA (A) and mRNA (B) merged modules, defined by the “tree” method, minimum module size of 25 genes, and minimum module merge height of 0.8.

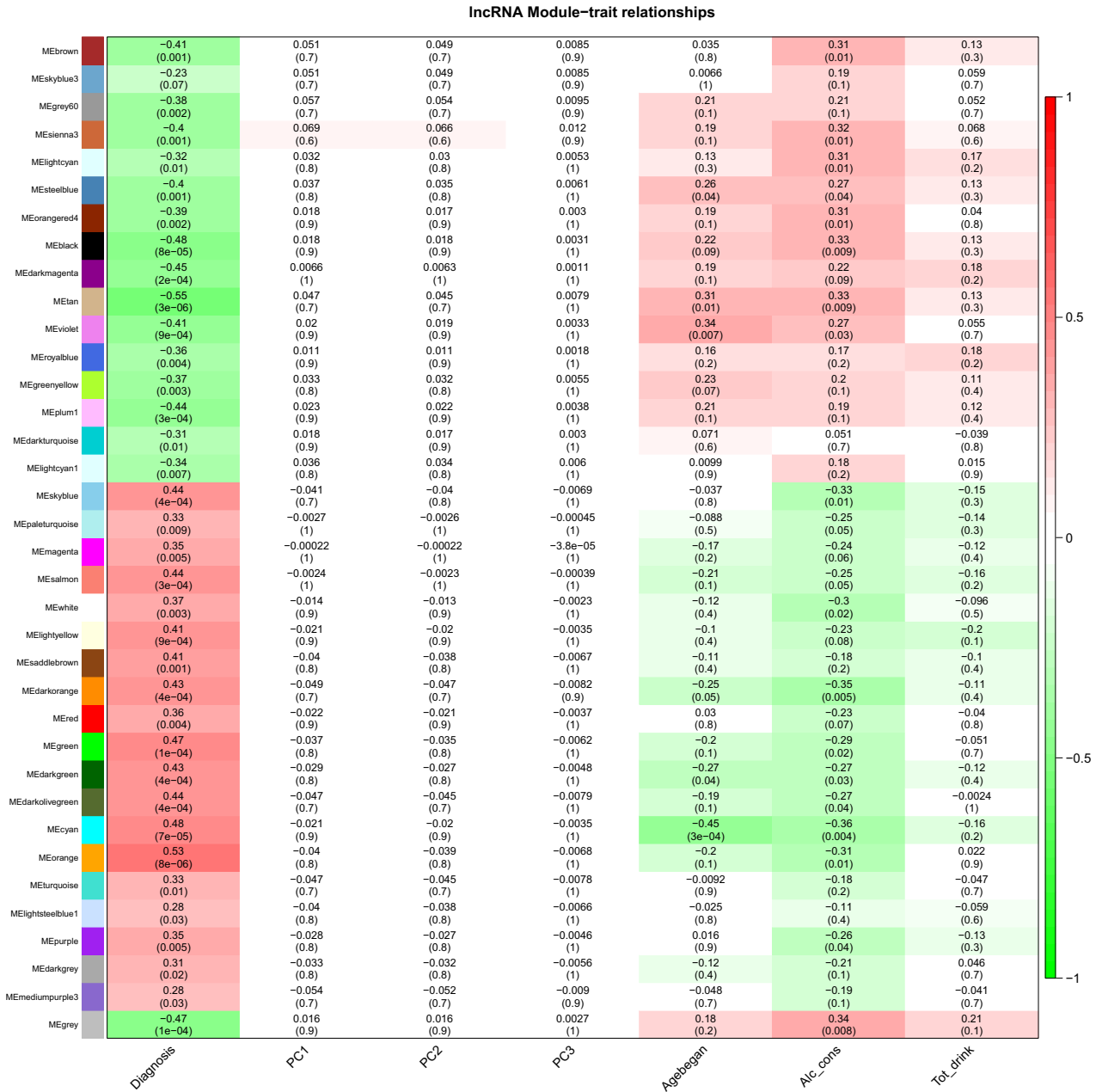


Fig. 2. Module-trait relationships for lncRNA. The residuals of the expression values used to generate the lncRNA module MEs are correlated (Pearson) to the dichotomous AD case/control status (Diagnosis) and to quantitative alcohol measures such as daily alcohol consumptions (Alc-Cons), total amount of drinks (Tot_drinks), and initial age of drinking (Age_began). The lncRNA modules were also correlated to the first 3 PCs to assess for confounding. *p*-values shown are unadjusted for multiple testing. After adjusting for number of modules tested, 19 modules (*M*_{brown}, *M*_{sienna3}, *M*_{steelblue}, *M*_{black}, *M*_{darkmagenta}, *M*_{tan}, *M*_{violet}, *M*_{plum1}, *M*_{skyblue}, *M*_{salmon}, *M*_{lightyellow}, *M*_{saddlebrown}, *M*_{darkorange}, *M*_{green}, *M*_{darkgreen}, *M*_{darkolivegreen}, *M*_{cyan}, *M*_{orange}, and *M*_{gray}) were significantly correlated with AD case status.

DISCUSSION

The main goal of this study is to profile the expression patterns of the noncoding and coding transcriptome of NAc in a large sample of subjects with AD and controls. The NAc is a central component of the mesocorticolimbic system and is known to be involved in addictive behaviors (Marty and Spigelman, 2012; Russo et al., 2010; Tabakoff and Hoffman, 2013). By integrating GWAS data on the sample (Adkins

et al., 2017) with our expression data, we further attempted to reveal hitherto unknown relationships between lncRNA and PCG and to assess how genetic variants modulate this relationship in the brain. In agreement with previous post-mortem brain expression studies, we identified differentially expressed PCG, supporting the notion that chronic alcohol consumption is associated with widespread transcriptomic changes in the brain. We further identified numerous lncRNA differentially expressed between AD cases and

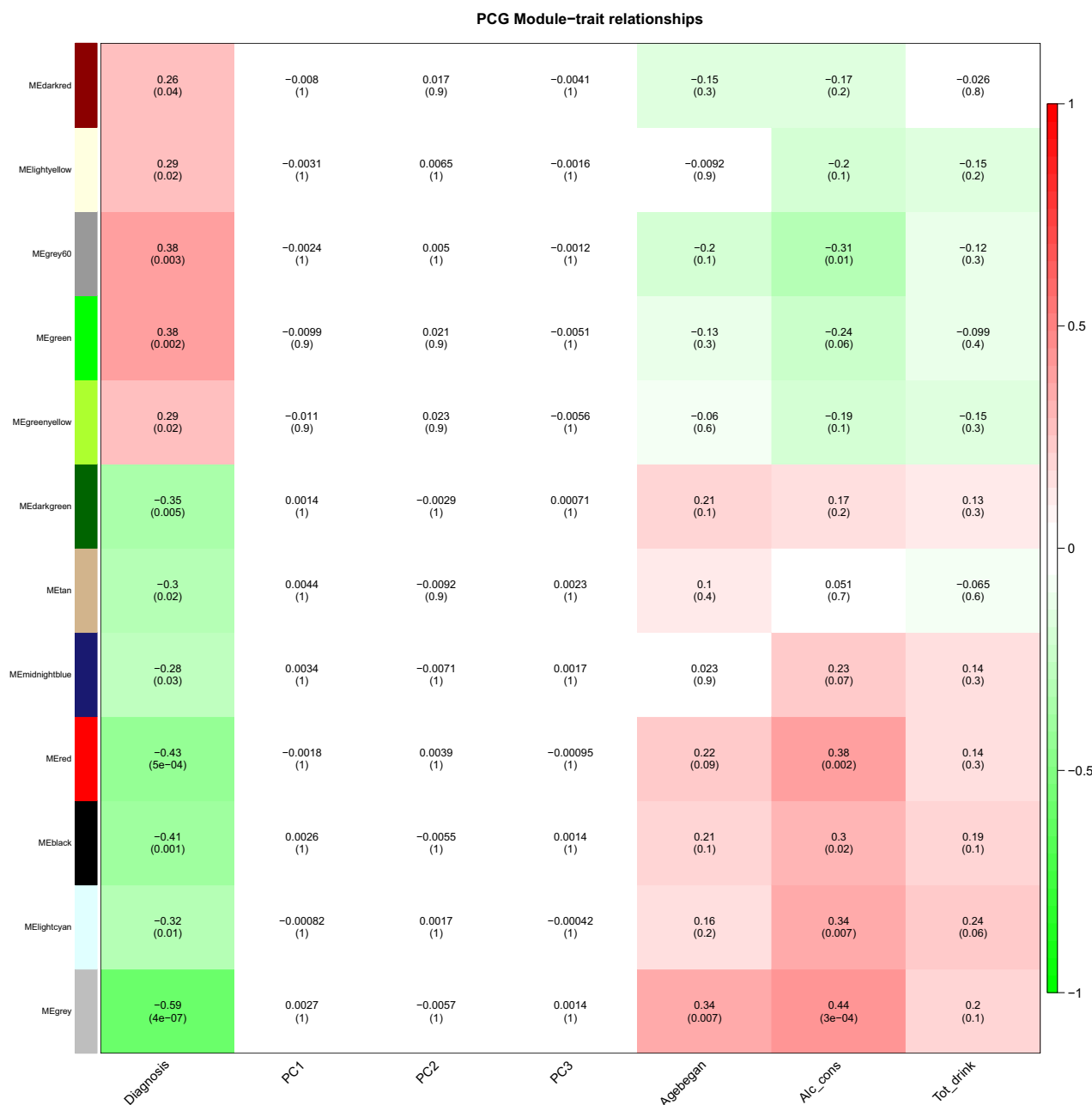


Fig. 3. Module-trait relationships for PCG. The PCG module MEs are correlated (Pearson) to the dichotomous AD case/control status (Diagnosis) and to quantitative alcohol measures such as daily alcohol consumptions (Alc-Cons), total amount of drinks (Tot_drinks), and initial age of drinking (Age_began). The lncRNA modules were also correlated to the first 3 PCs to assess for confounding. p -values shown are unadjusted for multiple testing. After adjusting for number of modules tested, M_{grey60} , M_{green} , M_{red} , M_{black} , M_{grey} are significantly correlated with AD case status (Class) and M_{grey} and M_{red} also with daily alcohol consumption.

controls, supporting their role in the neuroadaptation processes associated with chronic alcohol consumption. Among the most significant differentially expressed PCG were serpin peptidase inhibitor clade A (alpha-1 antitrypsin, antitrypsin) member 3 (SERPINA3), interferon-induced transmembrane protein 2 and 3 (IFITM2 and IFITM3), solute carrier family 14 (urea transporter), member 1 (Kidd blood group; SLC14A1) that were also identified in previous alcohol postmortem brain gene expression reports (Lewohl et al.,

2000; McClintick et al., 2013). While not the main topic of the study, interestingly, we also detected a limited number of differentially expressed pseudogenes, such as the annexin A2 pseudogenes 2 and 3 (ANXA2P2 and ANXA2P3) and interferon-induced transmembrane protein 4 pseudogene (IFITM4P) on chromosomes 9, 10, and 6, respectively. Previous postmortem brain expression studies have shown ANXA2P2 to be involved in stress response pathways in chronic alcoholics (McClintick et al., 2013); however, little is

known about the functions of ANXA2P3 and IFITM4P; it appears that ANXA2P3 is involved in reducing lipid levels in response to statins (Barber et al., 2010) and chronic obstructive pulmonary disease (Wilk et al., 2007), while no known functions were reported for IFITM4P.

A major limitation of the single gene expression analysis to identify and prioritize a set of genes associated with a given phenotype is its inability to consider the complex molecular interactions between such identified genes (Wu et al., 2013); this is especially relevant in studying the non-coding transcriptome, which is poorly annotated, with many lncRNA still of unknown function. To address this limitation, we performed a WGCNA, which allowed us to identify a set of coregulated PCG and lncRNA. We identified 19 lncRNA modules and 5 PCG modules that were significantly correlated with AD. Among the significantly correlated with AD status lncRNA and PCG modules was also M_{gray} . Genes in the gray module are frequently considered as “noise,” although there are suggestions that this interpretation may not always be accurate (Saelens et al., 2018). In general, observations have indicated that the genes in the gray module are usually ubiquitously expressed, exhibit oscillating, or highly variable patterns of expression, or in some cases could be simply assigned to a wrong module (Greenfest-Allen et al., 2017). Regardless, a potential explanation for M_{gray} being significant in our analyses could be that certain lncRNA and mRNA loci may impact the neuropathology of AD in isolation and not as a part of a gene network. Indeed, other studies have also reported similar observations, where, despite their unassigned status, genes in M_{gray} have shown significant disease associations (Li et al., 2018; Reinhold et al., 2017). Using the available quantitative measures of alcohol consumption, we further identified 2 PCG modules significantly correlated with the amount of daily alcohol consumption. Interestingly, while we observed the same PCG modules to be significantly correlated with both AD status and quantitative alcohol phenotype, no such relationship

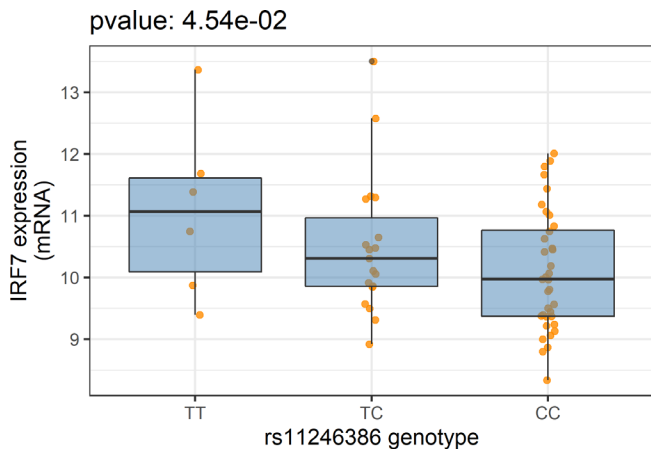


Fig. 4. Plot showing the most significant mRNA eQTL between IRF7 and rs11246386. The alternative allele of rs11246386 confers a lower IRF7 expression.

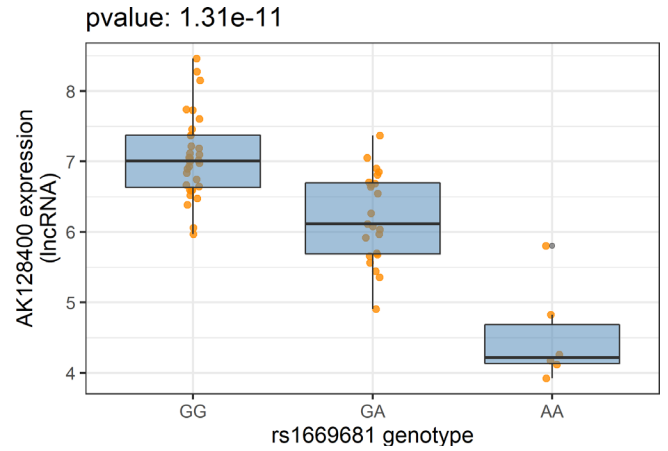


Fig. 5. Plot showing the most significant lncRNA eQTL between AK128400 and rs1669681. The alternative allele of rs1669681 confers a lower AK128400 expression.

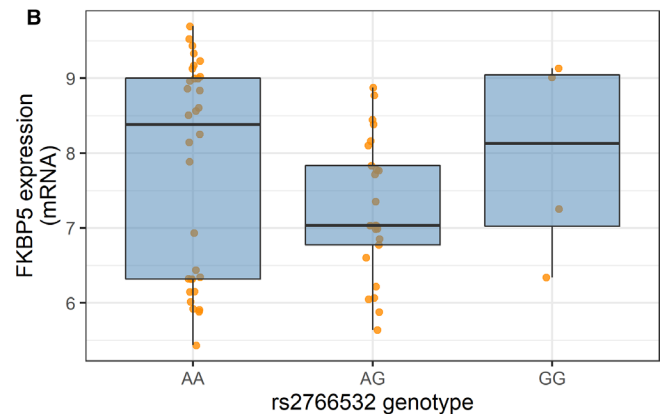
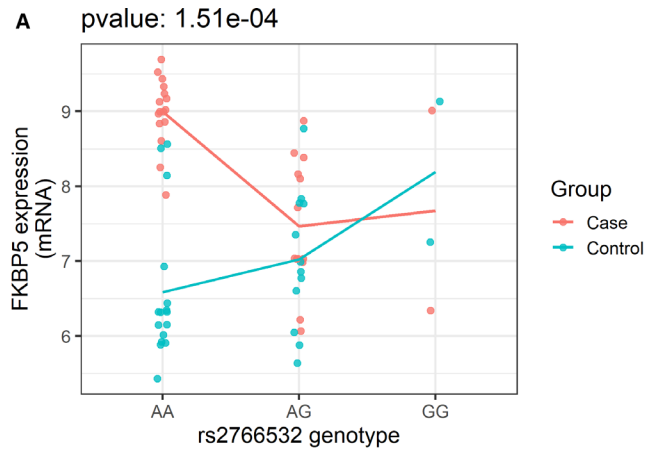


Fig. 6. Plot showing the most significant mRNA eQTL between FKBP5 and rs2766532 with (A) and without (B) the group interaction term. The alternative allele of rs2766532 confers a higher expression of FKBP5 in controls and a lower expression in cases.

was revealed for the lncRNA module. This suggests that different sets of lncRNA genes may impact these related, yet distinct, measures of alcohol-related phenotypes.

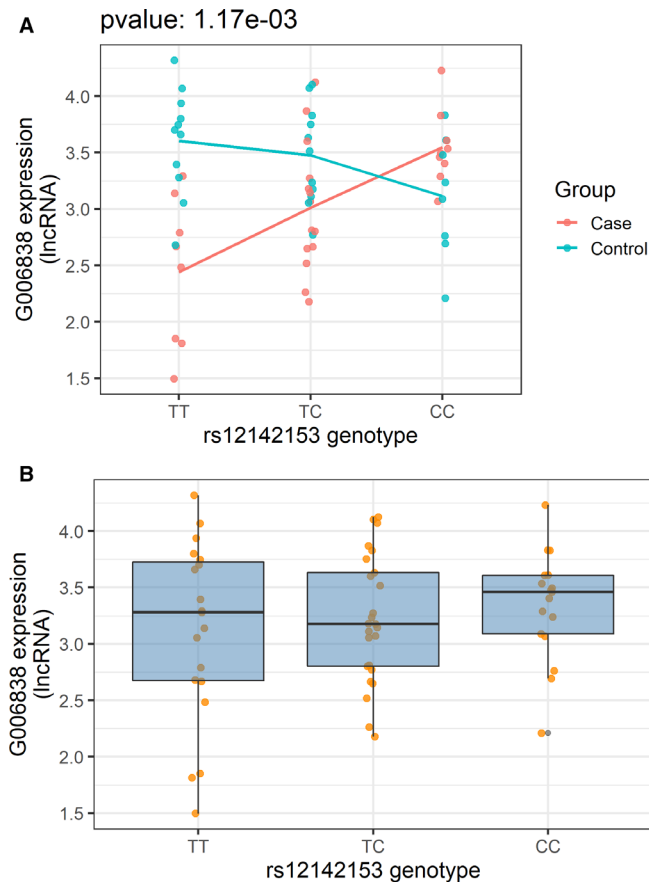


Fig. 7. Plot showing the most significant lncRNA eQTL between G006838 and rs12142153 with **(A)** and without **(B)** the group interaction term. The alternative allele rs12142153 confers a higher expression of G006838 in cases and a lower expression in controls.

To further understand the potential regulatory functions of lncRNA, we performed a series of correlation analyses between the hub genes identified from the significant lncRNA and PCG modules. In these analyses, we observed highly significant positive and negative correlations, in which the positive correlations significantly outnumbered the negative correlations. From our GSEA, we observed that similar to other studies (Mamdani et al., 2015; Osterndorff-Kahanek et al., 2015; Ponomarev et al., 2012), some of the PCG modules are enriched for immune-related or neurodegenerative processes. Activation of immune-related processes in the brain as a result of chronic alcohol consumption has been well documented (Crews et al., 2013; Kelley and Dantzer, 2011). Likewise, the neurodegenerative processes observed for some of our modules may reflect the nature of our post-mortem sample composed exclusively from chronic alcoholics, that is, subjects that have been drinking for over 2 decades. Both animal and human genetic models have demonstrated that prolonged exposure to alcohol leads to activation of the microglial population with the concurrent activation of immune processes in the brain (Cui et al., 2014; Robinson et al., 2014). It has been well known, for over

several decades now, that prolonged and excessive alcohol drinking leads to loss of a brain matter (Harper and Kril, 1985; Kubota et al., 2001) and, consequently, activation of genes involved in neurodegenerative disorders.

Integrating our expression and genetic data led to the identification of eQTLs that affect lncRNA and PCG expression. Having an integrated view on the expression of lncRNA and PCG, we were also interested in identifying eQTL that can impact lncRNA and PCG expressions. We focused our attention on eQTL impacting negatively correlated lncRNA/mRNA hub pairs, due to the likely regulatory nature of the negative interactions as well as to increase our statistical power. While we observed some interesting genotype effects on lncRNA/PCG interactions, none of these, however, survived correction for multiple testing at FDR of 10%. For instance, the lncRNA G081251 locus was negatively correlated with 2 genes, leukocyte elastase inhibitor (SERPINB1) and S100 calcium-binding protein A11 (S100A11; Fig. S4A, B). SERPINB1 has been linked to neuroinflammation (Hou et al., 2019), whereas S100A11 exhibits one of the most significant coexpression differences between schizophrenia and bipolar disorder (Baumont et al., 2015). Nevertheless, our results demonstrate that the complex interactions between the coding and noncoding transcriptome observed in the neuropathology of AD are, at least partially, modulated by the genetic factors (Mamdani et al., 2015; Sartor et al., 2012).

Among the eQTLs affecting PCG expression, we also detected strong genetic interactions between PCG expression and disease status, with the strongest interaction effect observed between FKBP5 and rs2766532. Several studies have implicated FKBP5 in the severity of alcohol withdrawal (Huang et al., 2014), alcohol drinking patterns in rodents (Qiu et al., 2016), problematic drinking (Dragan et al., 2018), as well as other psychiatric ailments, such as posttraumatic stress disorder (PTSD; Wilker et al., 2014). We further identified eQTLs impacting the expression of PCG with potential functions linked to neuroinflammation (SERPINB1), alcohol-induced neuroinflammation (interferon regulatory factor 7, IRF7 [Coleman et al., 2017; Erickson et al., 2019]), neurodegenerative diseases (neural precursor cell expressed, developmentally down-regulated 9, or NEDD9 [Li et al., 2008]), nicotine addiction (DNA methyltransferase 3 beta or DNMT3B [Hancock et al., 2015]), and ethanol exposure (solute carrier organic anion transporter family member 4A1 or SLCO4A1 [Osterndorff-Kahanek et al., 2015]). Among the eQTLs associated with lncRNA disease expression differences, many appear to be novel discoveries, likely due to the limited number of reports studying lncRNA role in alcohol addiction. For instance, within our results, we found only 1 lncRNA locus (i.e., AK128400), whose expression was previously reported to be associated with autism (Tang et al., 2017).

The importance of our postmortem brain expression study to assess lncRNA expression in NAc is further supported by 3 recent GWAS of AUD that has reported genome-wide signals near lncRNA genes: (i) a study conducted by Gelernter and colleagues (2014) that implicated the ADH gene cluster

on chromosome 4 and the LOC100507053 locus, (ii) a study conducted by the COGA group (Wetherill et al., 2015) that identified a polymorphism near LOC151121 on chromosome 2, and (iii) a study conducted by our group that has reported several polymorphisms in LOC339975 (Adkins et al., 2017). However, in agreement with a previous report from our group, none of these 3 loci showed evidence for differential expression, although the expression of LOC339975 was nominally affected by the most significantly associated with AD polymorphism (rs11726136) in NAc (Adkins et al., 2017).

In summary, the main goal of this study was to test the hypothesis that lncRNA, as a novel class of non coding RNAs, contribute to the neuropathology of AD. While our study is novel in its approach to address the potential role of lncRNA in the neuropathology of AD by integrating both genetic and molecular data generated in postmortem brains from subjects with AD, it is not without limitations. First, postmortem brain studies are observational as manipulation of living human subjects is not possible. Although the cross-sectional nature of these studies limits the causal inference we can make, we believe the eQTL analysis is a major step toward clarifying the directionality of these observations. Secondly, although our sample size ($N = 63$) is small, we believe with our careful experimental design and implementation of integrative multivariate approaches, we can circumvent some of these limitations and further broaden our understanding of alcohol addiction processes. And finally, we acknowledged that examining AD through a lens of lncRNA's interaction with mRNA overlooks epigenetic mechanisms (Bohnsack et al., 2019; Marchese et al., 2017). We hope that our study will provide additional molecular targets that will help translate these advances into effective therapeutic strategies for patients suffering from alcohol addiction. In conclusion, we believe our study makes an important contribution to the field of alcohol research. To the best of our knowledge, ours is among the first studies to date to pave the road for future lncRNA research in the neuropathology of alcohol addiction.

ACKNOWLEDGMENTS

This work was supported by NIAAA R21AA022749-01A1 to V.I.V., NIAAA P50AA022537 to B.P.R., K.S.K., and M.F.M.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Adkins AE, Hack LM, Bigdeli TB, Williamson VS, McMichael GO, Mamdani M, Edwards AC, Aliev F, Chan RF, Bhandari P, Raabe RC, Alaimo JT, Blackwell GG, Moscati A, Poland RS, Rood B, Patterson DG, Walsh D, Whitfield JB, Zhu G, Montgomery GW, Henders AK, Martin NG, Heath AC, Madden PAF, Frank J, Ridinger M, Wodarz N, Soyka M, Zill P, Ising M, Nothen MM, Kiefer F, Rietschel M, Gelernter J, Sherva R, Koesterer R, Almasy L, Zhao H, Kranzler HR, Farrer LA, Maher BS, Prescott CA, Dick DM, Bacanu SA, Mathies LD, Davies AG, Vladimirov VI, Grotewiel M, Bowers MS, Bettinger JC, Webb BT, Miles MF, Kendler KS, Riley BP (2017) Genomewide association study of alcohol dependence identifies risk loci altering ethanol-response behaviors in model organisms. *Alcohol Clin Exp Res* 41:911–928.
- Barber MJ, Mangravite LM, Hyde CL, Chasman DI, Smith JD, McCarty CA, Li X, Wilke RA, Rieder MJ, Williams PT, Ridker PM, Chatterjee A, Rotter JJ, Nickerson DA, Stephens M, Krauss RM (2010) Genome-wide association of lipid-lowering response to statins in combined study populations. *PLoS One* 5:e9763.
- Bohnsack JP, Teppen T, Kyzar EJ, Dzitoyeva S, Pandey S (2019) The lncRNA BDNF-AS is an epigenetic regulator in the human amygdala in early onset alcohol use disorders. *Transl Psychiatry* 9:1–11.
- Buckley ST, Eckert AL, Dodd PR (2000) Expression and distribution of GABAA receptor subtypes in human alcoholic cerebral cortex. *Ann N Y Acad Sci* 914:58–64.
- Clark BS, Blackshaw S (2014) Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease. *Front Genet* 5:1–19.
- Coleman LG Jr, Zou J, Crews FT (2017) Microglial-derived miRNA let-7 and HMGB1 contribute to ethanol-induced neurotoxicity via TLR7. *J Neuroinflamm* 14:1–15.
- Crews FT, Qin L, Sheedy D, Vetreno RP, Zou J (2013) High mobility group box 1/Toll-like receptor danger signaling increases brain neuroimmune activation in alcohol dependence. *Biol Psychiatry* 73:602–612.
- Cui C, Shurtleff D, Harris RA (2014) Neuroimmune mechanisms of alcohol and drug addiction. *Int Rev Neurobiol* 118:1–12.
- de Baumont A, Maschietto M, Lima L, Carraro DM, Olivieri EH, Fiorini A, Barreta LAN, Palha JA, Belmonte-de-Abreu P, Moreira Filho CA, Brentani H (2015) Innate immune response is differentially dysregulated between bipolar disease and schizophrenia. *Schizophr Res* 161:215–221.
- Dragan WL, Domozych W, Czerski PM, Dragan M (2018) Positive metacognitions about alcohol mediate the relationship between FKBP5 variability and problematic drinking in a sample of young women. *Neuropsychiatr Dis Treat* 14:2681–2688.
- Du Z, Sun T, Hacisuleyman E, Fei T, Wang X, Brown M, Rinn JL, Lee MG, Chen Y, Kantoff PW, Liu XS (2016) Integrative analyses reveal a long noncoding RNA-mediated sponge regulatory network in prostate cancer. *Nat Commun* 7:1–10.
- Erickson EK, Blednov YA, Harris RA, Mayfield RD (2019) Glial gene networks associated with alcohol dependence. *Sci Rep* 9:1–13.
- Fan L, van der Brug M, Chen W, Dodd PR, Matsumoto I, Niwa S, Wilce PA (1999) Increased expression of mitochondrial genes in human alcoholic brain revealed by differential display. *Alcohol Clin Exp Res* 23:408–413.
- Farris SP, Arasappan D, Hunnicke-Smith S, Harris RA, Mayfield RD (2015) Transcriptome organization for chronic alcohol abuse in human brain. *Mol Psychiatry* 20:1438–1447.
- Farris SP, Wolen AR, Miles MF (2010) Using expression genetics to study the neurobiology of ethanol and alcoholism. *Int Rev Neurobiol* 91:95–128.
- Flatscher-Bader T, Harrison E, Matsumoto I, Wilce PA (2010) Genes associated with alcohol abuse and tobacco smoking in the human nucleus accumbens and ventral tegmental area. *Alcohol Clin Exp Res* 34:1291–1302.
- Flatscher-Bader T, van der Brug M, Hwang JW, Gochee PA, Matsumoto I, Niwa S, Wilce PA (2005) Alcohol-responsive genes in the frontal cortex and nucleus accumbens of human alcoholics. *J Neurochem* 93:359–370.
- Flatscher-Bader T, van der Brug MP, Landis N, Hwang JW, Harrison E, Wilce PA (2006) Comparative gene expression in brain regions of human alcoholics. *Genes Brain Behav* 5(Suppl. 1):78–84.
- Flatscher-Bader T, Zuvella N, Landis N, Wilce PA (2008) Smoking and alcoholism target genes associated with plasticity and glutamate transmission in the human ventral tegmental area. *Hum Mol Genet* 17:38–51.

Adkins AE, Hack LM, Bigdeli TB, Williamson VS, McMichael GO, Mamdani M, Edwards AC, Aliev F, Chan RF, Bhandari P, Raabe RC, Alaimo JT, Blackwell GG, Moscati A, Poland RS, Rood B, Patterson DG, Walsh D, Whitfield JB, Zhu G, Montgomery GW, Henders AK, Martin NG, Heath AC, Madden PAF, Frank J, Ridinger M, Wodarz N, Soyka M, Zill

- Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, Schork AJ, Appadurai V, Buil A, Werge TM, Liu C, White KP, Horvath S, Geschwind DH (2018) Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* 359:693–697.
- Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, Anton R, Preuss UW, Ridinger M, Rujescu D, Wodarz N, Zill P, Zhao H, Farrer LA (2014) Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry* 19:41–49.
- Gibb EA, Brown CJ, Lam WL (2011) The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* 10:1–17.
- Gray KA, Daugherty LC, Gordon SM, Seal RL, Wright MW, Bruford EA (2013) Genenames.org: the HGNC resources in 2013. *Nucleic Acids Res* 41:D545–D552.
- Greenfest-Allen E, Cartailleur JP, Magnuson MA, Stoeckert CJ (2017) iterativeWGCNA: iterative refinement to improve module detection from WGCNA co-expression networks. *bioRxiv*.
- Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477:295–300.
- Guttman M, Rinn JL (2012) Modular regulatory principles of large non-coding RNAs. *Nature* 482:339–346.
- Han P, Chang CP (2015) Long non-coding RNA and chromatin remodeling. *RNA Biol* 12:1094–1098.
- Hancock DB, Reginnson GW, Gaddis NC, Chen X, Saccone NL, Lutz SM, Qaiser B, Sherva R, Steinberg S, Zink F, Stacey SN, Glasheen C, Chen J, Gu F, Frederiksen BN, Loukola A, Gudbjartsson DF, Bruske I, Landi MT, Bickeboller H, Madden P, Farrer L, Kaprio J, Kranzler HR, Gelernter J, Baker TB, Kraft P, Amos CI, Caporaso NE, Hokanson JE, Bierut LJ, Thorgeirsson TE, Johnson EO, Stefansson K (2015) Genome-wide meta-analysis reveals common splice site acceptor variant in *CHRNA4* associated with nicotine dependence. *Transl Psychiatry* 5:1–11.
- Harper C, Kril J (1985) Brain atrophy in chronic alcoholic patients: a quantitative pathological study. *J Neurol Neurosurg Psychiatry* 48:211–217.
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 106:9362–9367.
- Hou L, Rao DA, Yuki K, Cooley J, Henderson LA, Jonsson AH, Kaiserman D, Gorman MP, Nigrovic PA, Bird PI, Becher B, Remold-O'Donnell E (2019) SerpinB1 controls encephalitogenic T helper cells in neuroinflammation. *Proc Natl Acad Sci USA* 116:20635–20643.
- Huang MC, Schwandt ML, Chester JA, Kirchhoff AM, Kao CF, Liang T, Tapocik JD, Ramchandani VA, George DT, Hodgkinson CA, Goldman D, Heilig M (2014) FKBP5 moderates alcohol withdrawal severity: human genetic association and functional validation in knockout mice. *Neuropsychopharmacology* 39:2029–2038.
- Hung JH, Yang TH, Hu Z, Weng Z, DeLisi C (2012) Gene set enrichment analysis: performance evaluation and usage guidelines. *Brief Bioinform* 13:281–291.
- Iwamoto K, Bundo M, Yamamoto M, Ozawa H, Saito T, Kato T (2004) Decreased expression of NEFH and PCP4/PEP19 in the prefrontal cortex of alcoholics. *Neurosci Res* 49:379–385.
- Jarroux J, Morillon A, Pinskaya M (2017) History, discovery, and classification of lincRNAs. *Adv Exp Med Biol* 1008:1–46.
- Kapoor M, Wang J-C, Farris SP, Liu Y, McClintick J, Gupta I, Meyers JL, Bertelsen S, Chao M, Nurnberger J, Tischfield J, Harari O, Zeran L, Hesselbrock V, Bauer L, Raj T, Porjesz B, Agrawal A, Foroud T, Edenberg HJ, Mayfield RD, Goate A (2019) Analysis of whole genome-transcriptomic organization in brain to identify genes associated with alcoholism. *Transl Psychiat* 9:1–11.
- Kelley KW, Dantzer R (2011) Alcoholism and inflammation: neuroimmunology of behavioral and mood disorders. *Brain Behav Immun* 25 (Suppl. 1):S13–S20.
- Kubota M, Nakazaki S, Nakazaki S, Saeki N, Yamaura A, Kusaka T (2001) Alcohol consumption and frontal lobe shrinkage: study of 1432 non-alcoholic subjects. *J Neurol Neurosurg Psychiatry* 71:104–106.
- Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform* 9:1–13.
- Lewohl JM, Wang L, Miles MF, Zhang L, Dodd PR, Harris RA (2000) Gene expression in human alcoholism: microarray analysis of frontal cortex. *Alcohol Clin Exp Res* 24:1873–1882.
- Li H (2013) Systems genetics in "omics" era: current and future development. *Theory Biosci* 132:1–16.
- Li J, Li J, Liu X, Qin S, Guan Y, Liu Y, Cheng Y, Chen X, Li W, Wang S, Xiong M, Kuzhikandathil EV, Ye JH, Zhang C (2013) MicroRNA expression profile and functional analysis reveal that miR-382 is a critical novel gene of alcohol addiction. *EMBO Mol Med* 5:1402–1414.
- Li J, Zhou D, Qiu W, Shi Y, Yang JJ, Chen S, Wang Q, Pan H (2018) Application of weighted gene co-expression network analysis for data from paired design. *Sci Rep* 8:1–8.
- Li Y, Grupe A, Rowland C, Holmans P, Segurado R, Abraham R, Jones L, Catanese J, Ross D, Mayo K, Martinez M, Hollingworth P, Goate A, Cairns NJ, Racette BA, Perlmutter JS, O'Donovan MC, Morris JC, Brayne C, Rubinsztein DC, Lovestone S, Thal LJ, Owen MJ, Williams J (2008) Evidence that common variation in *NEDD9* is associated with susceptibility to late-onset Alzheimer's and Parkinson's disease. *Hum Mol Genet* 17:759–767.
- Liu C (2011) Brain expression quantitative trait locus mapping informs genetic studies of psychiatric diseases. *Neurosci Bull* 27:123–133.
- Long Y, Wang X, Youmans DT, Cech TR (2017) How do lincRNAs regulate transcription? *Sci Adv* 3:1–13.
- Mamdani M, Williamson V, McMichael GO, Blevins T, Aliev F, Adkins A, Hack L, Bigdeli T, der Vaart DV, Web BT, Bacanu SA, Kalsi G, Kendler KS, Miles MF, Dick D, Riley BP, Dumur C, Vladimirov VI (2015) Integrating mRNA and miRNA weighted gene co-expression networks with eQTLs in the nucleus accumbens of subjects with alcohol dependence. *PLoS One* 10:e0137671.
- Marchese FP, Raimondi I, Huarte M (2017) The multidimensional mechanisms of long noncoding RNA function. *Genome Biol* 18:1–13.
- Marty VN, Spigelman I (2012) Effects of alcohol on the membrane excitability and synaptic transmission of medium spiny neurons in the nucleus accumbens. *Alcohol* 46:317–327.
- Mayfield RD, Lewohl JM, Dodd PR, Herlihy A, Liu JW, Harris RA (2002) Patterns of gene expression are altered in the frontal and motor cortices of human alcoholics. *J Neurochem* 81:802–813.
- McClintick JN, Xuei X, Tischfield JA, Goate A, Foroud T, Wetherill L, Ehringer MA, Edenberg HJ (2013) Stress-response pathways are altered in the hippocampus of chronic alcoholics. *Alcohol* 47:505–515.
- Mizuo K, Katada R, Okazaki S, Tateda K, Watanabe S, Matsumoto H (2012) Epigenetic regulation of MIR-124 under ethanol dependence and withdrawal. *Nihon Arukoru Yakubutsu Igakkai Zasshi* 47:155–163.
- Nica AC, Dermitzakis ET (2013) Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci* 368:1–6.
- Nica AC, Montgomery SB, Dimas AS, Stranger BE, Beazley C, Barroso I, Dermitzakis ET (2010) Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genet* 6:e1000895.
- Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ (2010) Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 6:e1000888.
- Osterndorff-Kahanek EA, Becker HC, Lopez MF, Farris SP, Tiwari GR, Nunez YO, Harris RA, Mayfield RD (2015) Chronic ethanol exposure produces time- and brain region-dependent changes in gene coexpression networks. *PLoS One* 10:e0121522.
- Peng L, Bian XW, Li DK, Xu C, Wang GM, Xia QY, Xiong Q (2015) Large-scale RNA-Seq transcriptome analysis of 4043 cancers and 548 normal tissue controls across 12 TCGA cancer types. *Sci Rep* 5:1–18.
- Pietrzykowski AZ, Friesen RM, Martin GE, Puig SI, Nowak CL, Wynne PM, Siegelmann HT, Treistman SN (2008) Posttranscriptional regulation

- of BK channel splice variant stability by miR-9 underlies neuroadaptation to alcohol. *Neuron* 59:274–287.
- Ponomarev I, Wang S, Zhang L, Harris RA, Mayfield RD (2012) Gene co-expression networks in human brain identify epigenetic modifications in alcohol dependence. *J Neurosci* 32:1884–1897.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575.
- Qiu B, Luczak SE, Wall TL, Kirchoff AM, Xu Y, Eng MY, Stewart RB, Shou W, Boehm SL, Chester JA, Yong W, Liang T (2016) The FKBP5 gene affects alcohol drinking in knockout mice and is implicated in alcohol drinking in humans. *Int J Mol Sci* 17:1–17.
- Rao X, Thapa KS, Chen AB, Lin H, Gao H, Reiter JL, Hargreaves KA, Ipe J, Lai D, Xuei X, Wang Y, Gu H, Kapoor M, Farris SP, Tischfield J, Foroud T, Goate AM, Skaar TC, Mayfield RD, Edenberg HJ, Liu Y (2019) Allele-specific expression and high-throughput reporter assay reveal functional genetic variants associated with alcohol use disorders. *Mol Psychiatry* 1–10. <https://doi.org/10.1038/s41380-019-0508-z>.
- Reinhold D, Morrow JD, Jacobson S, Hu J, Ringel B, Seibold MA, Hersh CP, Kechris KJ, Bowler RP (2017) Meta-analysis of peripheral blood gene expression modules for COPD phenotypes. *PLoS One* 12:e0185682.
- Ritchie M, Phipson B, Wu D, Hu Y, Law C, Shi W, Smyth G (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47.
- Robinson G, Most D, Ferguson LB, Mayfield J, Harris RA, Blednov YA (2014) Neuroimmune pathways in alcohol consumption: evidence from behavioral and genetic studies in rodents and humans. *Int Rev Neurobiol* 118:13–39.
- Ron D, Messing RO (2013) Signaling pathways mediating alcohol effects. *Curr Top Behav Neurosci* 13:87–126.
- Russo SJ, Dietz DM, Dumitriu D, Morrison JH, Malenka RC, Nestler EJ (2010) The addicted synapse: mechanisms of synaptic and structural plasticity in nucleus accumbens. *Trends Neurosci* 33:267–276.
- Saelens W, Cannoodt R, Saeyns Y (2018) A comprehensive evaluation of module detection methods for gene expression data. *Nat Commun* 9:1–12.
- Sartor GC, St. Laurent G, Wahlestedt C (2012) The emerging role of non-coding RNAs in drug addiction. *Front Genet* 3:1–8.
- Schrode N, Ho SM, Yamamuro K, Dobbay A, Huckins L, Matos MR, Cheng E, Deans PJM, Flaherty E, Barretto N, Topol A, Alganem K, Abadali S, Gregory J, Hoelzli E, Phatnani H, Singh V, Girish D, Aronow B, Mccullumsmith R, Hoffman GE, Stahl EA, Morishita H, Sklar P, Brenand KJ (2019) Synergistic effects of common schizophrenia risk variants. *Nat Genet* 51:1475–1485.
- Shabalin AA (2012) Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 28:1353–1358.
- Shan Y, Ma J, Pan Y, Hu J, Liu B, Jia L (2018) LncRNA SNHG7 sponges miR-216b to promote proliferation and liver metastasis of colorectal cancer through upregulating GALNT1. *Cell Death Dis* 9:1–13.
- Shastri BS (2009) SNPs: impact on gene function and phenotype. *Methods Mol Biol* 578:3–22.
- Shu L, Zhao Y, Kurt Z, Byars SG, Tukiainen T, Kettunen J, Orozco LD, Pellegrini M, Lusi AJ, Ripatti S, Zhang B, Inouye M, Mäkinen VP, Yang X (2016) Mergeomics: multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genom* 17:1–16.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102:15545–15550.
- Sutherland GT, Sheedy D, Sheahan PJ, Kaplan W, Kril JJ (2014) Comorbidities, confounders, and the white matter transcriptome in chronic alcoholism. *Alcohol Clin Exp Res* 38:994–1001.
- Tabakoff B, Hoffman PL (2013) The neurobiology of alcohol consumption and alcoholism: an integrative history. *Pharmacol Biochem Behav* 113:20–37.
- Tang J, Yu Y, Yang W (2017) Long noncoding RNA and its contribution to autism spectrum disorders. *CNS Neurosci Ther* 23:645–656.
- Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhorji G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, Lundberg E, Hober S, Nilsson P, Mattsson J, Schwenk JM, Brunnström H, Glimelius B, Sjöblom T, Edqvist P-H, Djureinovic D, Mücke P, Lindskog C, Mardinoglu A, Ponten F (2016) A pathology atlas of the human cancer transcriptome. *Science* 357:eaan2507.
- Verhulst B, Neale MC, Kendler KS (2015) The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med* 45:1061–1072.
- Wang A, Wang J, Liu Y, Zhou Y (2017) Mechanisms of long non-coding RNAs in the assembly and plasticity of neural circuitry. *Front Neural Circuits* 11:1–14.
- Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, Clarke D, Gu M, Emani P, Yang YT, Xu M, Gandal MJ, Lou S, Zhang J, Park JJ, Yan C, Rhie SK, Manakongtreecheep K, Zhou H, Nathan A, Peters M, Mattei E, Fitzgerald D, Brunetti T, Moore J, Jiang Y, Girdhar K, Hoffman GE, Kalayci S, Gumus ZH, Crawford GE, Roussos P, Akbarian S, Jaffe AE, White KP, Weng Z, Sestan N, Geschwind DH, Knowles JA, Gerstein MB (2018) Comprehensive functional genomic resource and integrative model for the human brain. *Science* 362:eaat8464.
- Warden A, Mayfield R (2017) Gene expression profiling in the human alcoholic brain. *Neuropharmacology* 122:161–174.
- Wetherill L, Agrawal A, Kapoor M, Bertelsen S, Bierut LJ, Brooks A, Dick D, Hesselbrock M, Hesselbrock V, Koller DL, Le N, Nurnberger JI Jr, Salvatore JE, Schuckit M, Tischfield JA, Wang JC, Xuei X, Edenberg HJ, Porjesz B, Bucholz K, Goate AM, Foroud T (2015) Association of substance dependence phenotypes in the COGA sample. *Addict Biol* 20:617–627.
- Wickham, H (2016) ggplot2: Elegant Graphics for Data Analysis. 2, New York: Springer-Verlag, 1–267.
- Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT (2007) Framingham Heart Study genome-wide association: results for pulmonary function measures. *BMC Med Genet* 8(Suppl. 1):1–10.
- Wilker S, Pfeiffer A, Kolassa S, Elbert T, Lingenfelder B, Ovuga E, Papanotiropoulos A, de Quervain D, Kolassa IT (2014) The role of FKBP5 genotype in moderating long-term effectiveness of exposure-based psychotherapy for posttraumatic stress disorder. *Transl Psychiatry* 4:e403.
- Wu C, Zhu J, Zhang X (2013) Network-based differential gene expression analysis suggests cell cycle related genes regulated by E2F1 underlie the molecular difference between smoker and non-smoker lung adenocarcinoma. *BMC Bioinform* 14:1–16.
- Xu K, Kranzler HR, Sherva R, Sartor CE, Almasy L, Koesterer R, Zhao H, Farrer LA, Gelernter J (2015) Genomewide Association Study for maximum number of alcoholic drinks in European Americans and African Americans. *Alcohol Clin Exp Res* 39:1137–1147.
- Yadav S, Pandey A, Shukla A, Talwelkar SS, Kumar A, Pant AB, Parmar D (2011) miR-497 and miR-302b regulate ethanol-induced neuronal cell death through BCL2 protein and cyclin D2. *J Biol Chem* 286:37347–37357.
- Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4.
- Zhang H, Wang F, Xu H, Liu Y, Liu J, Zhao H, Gelernter J (2014) Differentially co-expressed genes in postmortem prefrontal cortex of individuals with alcohol use disorders: influence on alcohol metabolism-related pathways. *Hum Genet* 133:1383–1394.
- Zuo L, Tan Y, Wang Z, Wang KS, Zhang X, Chen X, Li CS, Wang T, Luo X (2016) Long noncoding RNAs in psychiatric disorders. *Psychiatr Genet* 26:109–116.

SUPPORTING INFORMATION

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

Fig. S1A. Scale-free topology fit graphs of (A) lncRNA and (B) mRNA. Showing the scale free topology fitting index (R^2) or mean connectivity against various power (β) settings.

Fig. S1B. Scale-free topology fit graphs of (A) lncRNA and (B) mRNA. Showing the scale free topology fitting index (R^2) or mean connectivity against various power (β) settings.

Fig. S2A. Robust, bootstrapped version of WGCNA (rWGCNA).

Fig. S2B. Robust, bootstrapped version of WGCNA (rWGCNA).

Fig. S3. Heatmap of the Pearson correlations between modules, which were significantly correlated with disease status.

Fig. S4A. Evidence of genotype (A:rs2446448 B:rs1496042) modulating lncRNA (G081251) and PCG (A:SERPINB1 B:S100A11) hubgenes, which were from significantly correlated AD modules and significantly negatively correlated with each other.

Fig. S4B. Evidence of genotype (A:rs2446448 B:rs1496042) modulating lncRNA (G081251) and PCG (A:SERPINB1 B:S100A11) hubgenes, which were from significantly correlated AD modules and significantly negatively correlated with each other.

Fig. S5. Pearson Correlation of t -values between limma's Empirical Bayes methodology with robust model and ordinary least square.

Table S1. Description of sample demographics per diagnostic group.

Table S2. Microarray expression data validation using quantitative PCR.

Table S3. Results from the univariate gene expression analysis at $p \leq 0.10$.

Table S4. Description of PCG(A) and lncRNA(B) with module membership, correlation with AD, significance of correlation, and module associations.

Table S5. Results from wkDA analysis of lncRNA within modules significantly correlated with disease status.

Table S6. Gene set enrichment analysis of PCG utilizing WGCNA's userListEnrichment function and curated list.

Table S7. Gene set enrichment analysis of PCG utilizing GSEA.

Table S8. Positive(A) and negative(B) significant correlations of hubgenes from modules significantly associated with AD.

Table S9. Results from the PCG(A) and lncRNA(B) eQTL analysis at FDR $p \leq 0.10$.

Table S10. Results from the PCG(A) and lncRNA(B) eQTL with group interaction analysis at FDR $p \leq 0.10$.

Table S11. Genotype modulation of negatively correlated lncRNA and PCG hubs.

Table S12. Primer set for lncRNA validation.

Methods S1. Supplementary methods