

RESEARCH ARTICLE

Blood and brain gene expression signatures of chronic intermittent ethanol consumption in mice

Laura B. Ferguson^{1,2,3*}, Amanda J. Roberts⁴, R. Dayne Mayfield^{1,3}, Robert O. Messing^{1,2,3*}

1 Waggoner Center for Alcohol and Addiction Research, University of Texas at Austin, Austin, Texas, United States of America, **2** Department of Neurology, Dell Medical School, University of Texas at Austin, Austin, Texas, United States of America, **3** Department of Neuroscience, University of Texas at Austin, Austin, Texas, United States of America, **4** Animal Models Core Facility, The Scripps Research Institute, San Diego, California, United States of America

* lferguson@utexas.edu (LBF); romessing@austin.utexas.edu (ROM)



OPEN ACCESS

Citation: Ferguson LB, Roberts AJ, Mayfield RD, Messing RO (2022) Blood and brain gene expression signatures of chronic intermittent ethanol consumption in mice. *PLoS Comput Biol* 18(2): e1009800. <https://doi.org/10.1371/journal.pcbi.1009800>

Editor: Atle van Beelen Granlund, NTNU Det Medisinske Fakultet DMF: Norges Teknisk-Naturvitenskapelige Universitet Fakultet for Medisin og Helsevitenskap, NORWAY

Received: July 8, 2021

Accepted: January 3, 2022

Published: February 17, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1009800>

Copyright: © 2022 Ferguson et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Raw data (fastq files) and processed data (raw counts and log-CPM)

Abstract

Alcohol Use Disorder (AUD) is a chronic, relapsing syndrome diagnosed by a heterogeneous set of behavioral signs and symptoms. There are no laboratory tests that provide direct objective evidence for diagnosis. Microarray and RNA-Seq technologies enable genome-wide transcriptome profiling at low costs and provide an opportunity to identify biomarkers to facilitate diagnosis, prognosis, and treatment of patients. However, access to brain tissue in living patients is not possible. Blood contains cellular and extracellular RNAs that provide disease-relevant information for some brain diseases. We hypothesized that blood gene expression profiles can be used to diagnose AUD. We profiled brain (prefrontal cortex, amygdala, and hypothalamus) and blood gene expression levels in C57BL/6J mice using RNA-seq one week after chronic intermittent ethanol (CIE) exposure, a mouse model of alcohol dependence. We found a high degree of preservation (ρ range: [0.50, 0.67]) between blood and brain transcript levels. There was small overlap between blood and brain DEGs, and considerable overlap of gene networks perturbed after CIE related to cell-cell signaling (e.g., GABA and glutamate receptor signaling), immune responses (e.g., antigen presentation), and protein processing / mitochondrial functioning (e.g., ubiquitination, oxidative phosphorylation). Blood gene expression data were used to train classifiers (logistic regression, random forest, and partial least squares discriminant analysis), which were highly accurate at predicting alcohol dependence status (maximum AUC: 90.1%). These results suggest that gene expression profiles from peripheral blood samples contain a biological signature of alcohol dependence that can discriminate between CIE and Air subjects.

Author summary

Recent evidence in mice suggests that brain gene expression profiles can predict disease status as well as predict drugs effective for treating alcoholism. However, it is not possible

voom-transformed normalized data) are publicly available on Gene Expression Omnibus GSE176122. R code used to analyze data can be found at https://github.com/zeavin-ferguson/blood_brain. Data can be explored at https://lauraferguson.shinyapps.io/blood_brain/.

Funding: This work was supported by National Institute on Alcohol Abuse and Alcoholism grants F32AA028148 (LBF), R01AA012404 (RDM), U01AA020926 (RDM), U01AA013520 (ROM), and R01AA026075 (ROM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

to obtain brain specimens from human patients which limits the usefulness of this approach. This study investigated the extent to which blood can act as a surrogate for brain tissue and predict CIE-induced alcohol dependence status. This information lays critical groundwork for developing molecular-based diagnosis and treatment options for alcoholic patients and provides insights into the biological mechanisms that might contribute to the transition from recreational alcohol use to excessive drinking.

Introduction

Alcohol Use Disorder (AUD) is a highly prevalent and costly syndrome with few effective treatments [1–3]. AUD like other psychiatric disorders is diagnosed by evaluating a patient's symptoms and behaviors over time as described in the Diagnostic and Statistical Manual of Mental Disorders (DSM5) [4]. Patients meeting two or more criteria within the last year are considered to have AUD with different degrees of severity. Adding molecular-based criteria would provide useful objective data to refine diagnosis and possibly afford earlier detection of problematic drinking before detrimental medical, legal, or social consequences of AUD appear. There are currently three FDA-approved treatments for AUD: disulfiram, acamprostate, and naltrexone, none of which are effective for all patients. Moreover, there are no reliable prognostic indicators that predict responses to therapeutic intervention.

Improvements in molecular technologies, computing power, and bioinformatics have revolutionized many fields of science and are beginning to impact medicine by harnessing vast amounts of data to inform diagnosis, prognosis, and treatment. Many genome-wide gene expression datasets are available from different brain regions, multiple species, and alcohol-related phenotypes. Studies examining these datasets have revealed that alcohol use (or the genetic risk for excessive alcohol use) alters brain gene expression, and these alterations can distinguish alcohol dependent subjects from healthy individuals, as well as predict therapeutic compounds. To discriminate between AUD and control subjects, application of partial least squares discriminative analysis (PLSDA) to gene expression patterns from postmortem prefrontal cortex tissue [5] has revealed a consistent re-programming of gene expression by years of having AUD that reliably discriminates AUD from non-AUD individuals. Our group recently identified a gene expression signature of risk for binge drinking from the brains of HDID-1 mice selected for high levels of binge alcohol drinking, and then used the Library of Integrated Network-Based Cellular Signatures (L1000) database from the Broad Institute to identify drugs with opposing patterns of gene expression, hypothesizing that drugs that produce anti-correlated patterns of gene expression might reduce alcohol drinking. The top-ranking drug candidates, terreic-acid and pergolide, both reduced ethanol consumption and blood alcohol levels in HDID-1 mice [6]. While these examples are promising, they have relied on *brain* gene expression data, and it is not possible to get brain tissue from AUD patients. Thus, applying advanced computational approaches to diagnose AUD and personalize AUD treatment will require noninvasive access to biological samples, such as blood. To this end, it is important to understand the extent to which blood can be used as a surrogate tissue for brain.

Although AUD is primarily considered a brain disease, alcohol use affects multiple other tissues and systems including gut, liver, lung, muscle, bone, heart, blood vessels, pancreas, and the immune system [7]. Whole blood is readily available and is routinely obtained in the clinic. We hypothesized that blood could be a useful surrogate for brain tissue, because it contacts every organ in the body, including the brain. Blood expresses ~80% of the genes that are expressed in brain, most of which are responsive to physiological or environmental

adaptations [8]. These genes include most of those that have been linked to AUD through genome-wide association studies or to alcohol-related behaviors in preclinical rodent studies [9, 10]. Blood and brain genome-wide gene expression profiles have been compared for several brain diseases including, schizophrenia [11], bipolar disorder [12], depression, [13, 14] Huntington's disease [15] and other neurodegenerative disorders [16], autism [17], and PTSD [18, 19]. These profiles have not been compared for AUD. To fill this knowledge gap, we compared gene expression levels in whole blood to those in the amygdala, prefrontal cortex, and hypothalamus from CIE mice undergoing withdrawal. We chose these brain areas because of their importance in reward signaling and alcohol dependence [20, 21] and because gene expression responses in these brain areas have been shown to parallel those in blood under other conditions [11, 15, 19, 22–24]. We used the chronic intermittent ethanol (CIE) procedure [25] to induce alcohol dependence in male and female C57BL/6J mice. We used this strain of mice because they show significant increases in alcohol drinking after being made alcohol dependent through CIE exposure [26–28]. CIE is thought to model the alcoholic individual's experience of episodic patterns of excessive ethanol consumption with repeated withdrawals [27]. It is postulated that alcohol is initially ingested for its rewarding effects, but repeated use leads to a negative emotional state experienced in alcohol's absence (the “dark side” of addiction) [29]; this state promotes excessive alcohol drinking through negative reinforcement, and treatments targeted against this state are postulated to help prevent relapse [30]. After CIE, rodents exhibit increased stress-responsiveness, depression-like behavior, and anxiety-like behavior [31–35], and they show a persistent increase in voluntary alcohol consumption that results in high blood alcohol levels (BALs) [26, 27, 36, 37]. We assessed gene expression profiles after 1 week of alcohol withdrawal because this is when increased voluntary alcohol intake is highest [27]. Additionally, the assessment of longer-term gene expression effects of ethanol avoided the confounds of acute ethanol intoxication and the volatile gene expression changes observed in early withdrawal [38, 39]. The within-subjects design we used to compare blood and brain gene expression enabled the detection of correlated responses across both tissues during withdrawal.

Results

Chronic intermittent ethanol (CIE) effects on voluntary ethanol drinking

To induce alcohol dependence, C57BL/6J mice underwent chronic intermittent ethanol (CIE) exposure interspersed with voluntary drinking sessions as described in the Materials and Methods section. For male mice, there was a main effect of treatment ($F(1, 16) = 6.463$, $p = 0.0217$) and time ($F(4, 64) = 12.73$, $p < 0.0001$), and a time x treatment interaction ($F(4, 64) = 4.830$, $p = 0.0018$). Planned Dunnett's comparisons revealed that CIE significantly increased voluntary intake of 15% ethanol over baseline after CIE cycles 2, 3, and 4 (Fig 1). Ethanol consumption in mice exposed to air did not differ from baseline drinking levels throughout the study, except for a slight decrease in ethanol intake after the first air exposure (Fig 1).

For female mice, a two-way ANOVA showed a main effect of time ($F(4, 68) = 5.882$, $p = 0.0004$), but not of treatment ($F(1, 17) = 2.745$, $p = 0.1159$) and no time x treatment interaction ($F(4, 68) = 1.436$, $p = 0.2317$). A one-way ANOVA revealed a main effect of time for the mice receiving ethanol vapor ($F(4, 45) = 3.677$, $P = 0.0113$) but not air ($F(4, 40) = 1.466$, $P = 0.2306$). Planned Dunnett's comparisons revealed that CIE significantly increased voluntary intake of 15% ethanol over baseline after CIE cycles 2, 3, and 4 for the ethanol treatment group (Fig 1). Ethanol consumption in mice exposed to air did not differ from baseline drinking levels throughout the study.

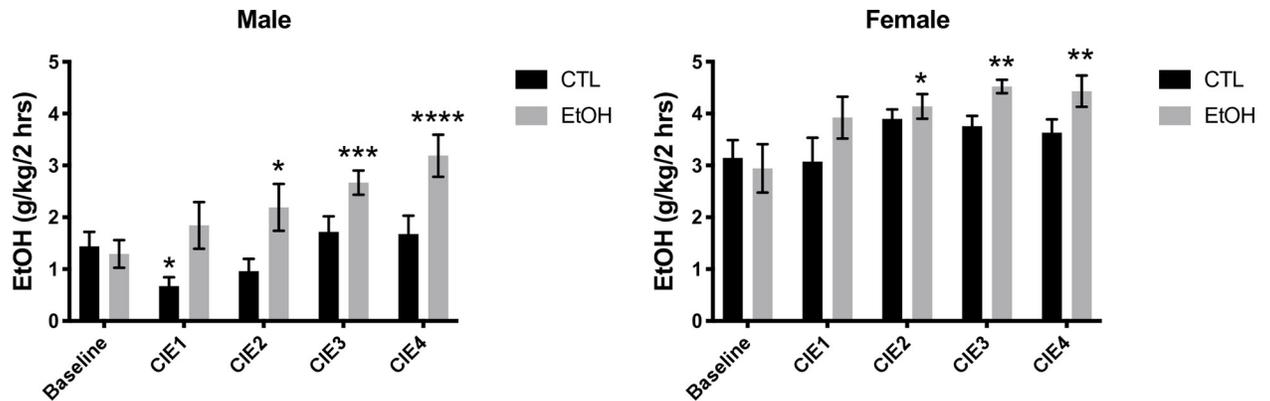


Fig 1. Effects of CIE on voluntary drinking in C57Bl/6J mice. Chronic intermittent ethanol (CIE) exposure significantly increased voluntary ethanol (15%) intake in male (but not female) mice as revealed by a two-way ANOVA. For female mice a one-way ANOVA revealed an effect of time on voluntary ethanol consumption in only the ethanol vapor group (EtOH) and not the control group (CTL). Results of Dunnett's planned comparisons are indicated above the SEM bars. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$ vs the baseline group. Values represent mean \pm SEM (n = 8–10/sex/group).

<https://doi.org/10.1371/journal.pcbi.1009800.g001>

Comparison of blood and brain gene expression

A major goal of this study was to determine the correspondence between peripheral blood and brain transcriptomes. We addressed this in several complimentary ways. First, to determine if the transcriptional response to CIE in blood are reflective of those in brain, we assessed the overlap of the genes and gene networks affected by CIE exposure across tissues. Next, to determine whether blood transcript levels can predict those in brain irrespective of treatment, we included all subjects and calculated the correlation coefficient between blood and brain gene expression levels.

CIE-responsive genes conserved between blood and brain

We found that CIE exposure dysregulated the expression levels of hundreds of genes in brain and blood (Fig 2A and S1 Table). The number of DEGs in brain was greater in males than females, while females showed a greater number of DEGs in blood relative to males (Fig 2A). To gain insight into the cellular specificity of gene expression perturbations, we determined whether any cell type-specific genes were over-represented in each of the gene sets. In females, microglial genes were up-regulated in all brain areas, endothelial genes up-regulated in amygdala and hypothalamus, and neuronal, T cell, and macrophage genes up-regulated in hypothalamus (Fig 2A and S2 Table). Neuronal and oligodendrocyte genes were down-regulated in female amygdala and hypothalamus (Fig 2A and S2 Table). In males, there were no cell type specific genes enriched in any of the up-regulated gene sets (Fig 2A and S2 Table). Astrocytic genes were down-regulated in male amygdala and hypothalamus, oligodendrocyte genes were down-regulated in amygdala, and microglial and endothelial cell genes were down-regulated in PFC (Fig 2A and S2 Table). In whole blood, T cell and B cell genes were down-regulated in females, while macrophage, neutrophil, and monocyte genes were down-regulated in males (Fig 2A and S2 Table).

We compared the DEGs in whole blood to those in each brain region. Depending on the brain region and sex, there were 45 to 88 overlapping DEGs between blood and brain. This overlap was statistically significant between whole blood and hypothalamus and PFC in male mice, and borderline significant between whole blood and hypothalamus ($p = 0.064$) and amygdala ($p = 0.046$) in female mice (Fig 2B).

regulators, diseases, and biological functions) and their relationships between one another. Immune-related entities were prominent in all the gene sets and were generally predicted to be activated in brain and inhibited in blood (except in male PFC where immune-related entities also were predicted to be inhibited). For example, *Ifng* and *Stat1* were predicted to be activated in all female brain regions but inhibited in blood, and *Il33*, *Il17a*, and *Map3k1* were predicted to be inhibited in male blood and PFC (Fig 2C). *Irf7* was a “hub” entity predicted to be activated in the amygdala of both sexes (Fig 2C). There were also several categories related to leukocyte extravasation and activation (e.g., recruitment of leukocytes, leukocyte migration, cell movement of leukocytes) which followed a similar pattern as the immune-related entities (activated in brain and inhibited in blood and male PFC) (Fig 2C). In addition to the predicted inhibition of immune-related entities in female blood, there were a number of entities that have been traditionally been associated with brain that were predicted to be activated in female blood, e.g., learning, cognition, *Bdnf* as “hub” entity (Fig 2C). Nuclear hormone signaling (LXR/RXR Activation) was predicted to be activated in male brain and blood (Fig 2C).

CIE-responsive gene networks conserved between blood and brain

We identified conserved gene coexpression modules in blood and brain. We built gene coexpression networks for each tissue individually and determined whether the different tissues had similar modules, i.e., were comprised of the same genes (if so, these modules were said to be conserved between blood and brain and are referred to as blood-brain modules).

To link the blood-brain modules to CIE exposure, we identified modules enriched with DEGs or with eigengenes correlated with alcohol preference or consumption levels in the final drinking test, or group (CIE and Air). To gain functional insight into the blood-brain modules, we performed an IPA Core Analysis on the genes within the blood-brain modules (S4 and S5 Tables). IPA analysis included predicted upstream regulators which are transcription factors, chemical compounds, microRNAs, or other regulators that might explain the observed changes in gene expression. If a predicted upstream regulator of a blood-brain module was also a member of the blood-brain module, this might indicate a particularly important role for that gene in regulating the transcriptional response to CIE-induced alcohol dependence. We highlight these genes in the following sections as we describe each blood-brain module identified in the network analyses.

We found three main groups of alcohol dependence-related blood-brain modules for female mice and four for male mice (boxes in Fig 3). The modules conserved between blood and brain were similar in males and females as revealed by the categories that emerged from IPA Core Analysis of the genes within the blood-brain modules. For example, there was a “cell-cell signaling” module (Fig 3) (e.g., endocannabinoid signaling, GABA and glutamate receptor signaling, synaptogenesis), an “immune response” module (Fig 3) (e.g., antigen presentation, communication between innate and adaptive immune systems, JAK/STAT signaling), and a “protein processing / mitochondrial function” module (Fig 3) (e.g., ubiquitination, unfolded protein responses, oxidative phosphorylation) for both males and females. Male mice also exhibited a “fatty acid metabolism / peroxisome proliferator activated receptor (PPAR)” blood-brain module (Fig 3).

Network analysis in female mice

The “cell-cell signaling” module was particularly highly conserved across blood and brain in female mice (Fig 3; edge thickness reflects degree of overlap). The genes in these modules were mostly up-regulated in blood and brain after CIE. Interestingly, this group of modules was enriched with genes already associated with alcohol-related behavior [9, 10, 40] (via mutant

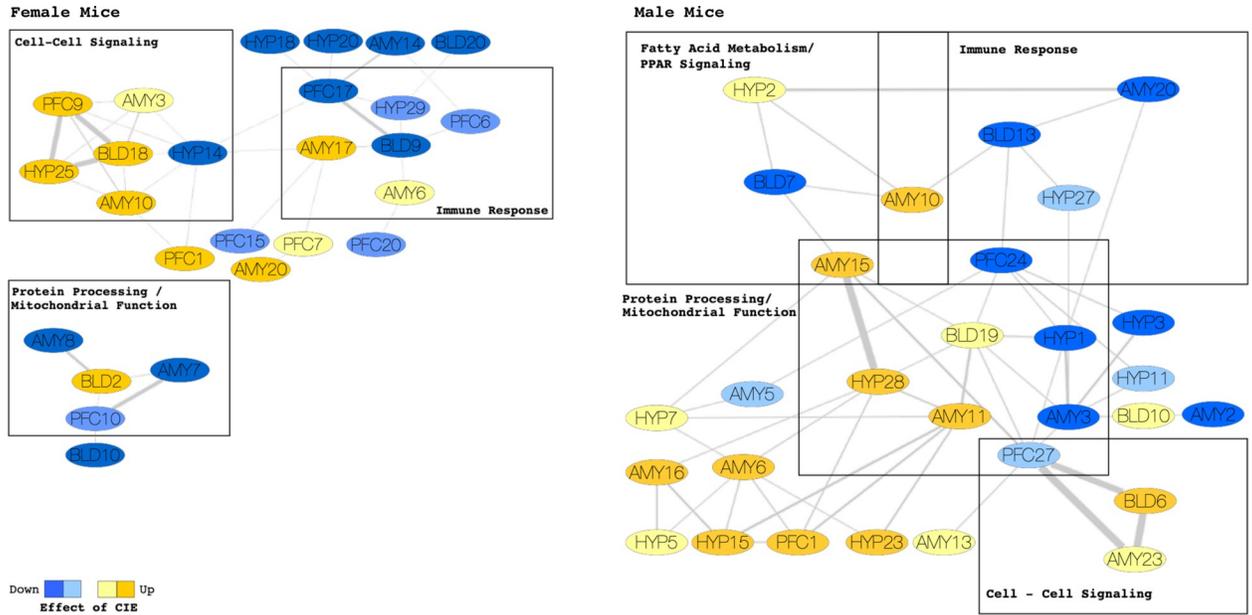


Fig 3. Blood/Brain Gene Coexpression Modules Affected by CIE-induced Alcohol Dependence. A meta-network of overlapping gene coexpression modules in blood (BLD) and brain [prefrontal cortex (PFC), amygdala (AMY), hypothalamus (HYP)] in female (left) and male (right) C57BL/6J mice. Each node represents a module of coexpressed genes. Nodes are labeled with tissue type and a module number. An edge between two nodes indicates a significant overlap of genes between two modules of different tissues. Thickness of connecting edges is proportional to the significance of the overlap. Module colors represent the direction and magnitude of regulation of CIE-induced alcohol dependence based on the significance of the enrichment with differentially expressed genes (yellow, upregulation; blue, downregulation in CIE mice; intense colors, $p < 0.00001$; light colors, $p < 0.05$). Only modules affected by CIE are shown. An overlap of a blood module and at least two modules from different brain regions indicates a cluster of highly conserved coexpression modules regulated by CIE in blood and brain (termed blood-brain modules in the text; represented by rectangular boxes in the figure). All overlapping modules within a blood-brain cluster were overrepresented with genes from a major biological category. These were the same for male and female mice apart from a Fatty Acid Metabolism and PPAR Signaling cluster unique to male mice. The top 5 enriched pathways and predicted upstream regulators from the IPA analysis for each of the broad categories (rectangular boxes) are reported in the text, but examples include GABA and Glutamate Receptor Signaling for the Cell-Cell Signaling cluster, Inflammasome Pathway and Th1 Pathway for the Immune Response cluster, Protein Ubiquitination Pathway and Oxidative Phosphorylation for the Protein Processing / Mitochondrial Function cluster, and Fatty Acid Oxidation and PPARA for the Fatty Acid Metabolism and PPAR Signaling cluster. The overlap between blood and brain modules was particularly strong for the Cell-Cell Signaling clusters in both male and female coexpression networks. The Cell-Cell Signaling cluster also contained genes known to play a role in alcohol-related behaviors in rodents (e.g., *Bdnf*, *Npy*, *Gabra1*, and *Pdyn*). Network visualization was performed using Cytoscape version 3.8.2.

<https://doi.org/10.1371/journal.pcbi.1009800.g003>

mouse studies) including: *Adcy1*, *Grin2a*, *Grm5*, *Hrh3*, *Npy1r*, *Adcy5*, *Cacna1b*, *Chrna4*, *Faah*, *Gabrb2*, *Gabrd*, *Grm4*, *Hras*, *Prkar2b*, *Prkce*, *Bdnf*, *Cckbr*, *Cnr1*, *Gabra1*, *Pdyn*, and *Prkar1b*. The top five predicted upstream regulators for this preserved module were *Rest*, *Hdac4*, *Snca*, *Fmr1*, *Htt*. Several of the predicted upstream regulators were also members of the module, suggesting these genes may be particularly important regulators of the blood-brain “cell-cell signaling” module in CIE-induced alcohol dependence (in decreasing order of significance): *Bdnf*, *Mapt*, *App*, *Tshz3*, *Slc30a3*, *Shank3*, *Nfasc*, *Kcnk9*, *Slitrk5*, *Fezf2*, *Gabra1*, *Ntrk3*, *Fbxo2*, *Mapk8ip1*, *Lhx2*, *Dmd*, *Slc9a6*, *Agrn*, *Gabbr2*, *Htr2a*, *Akap5*, *Dpp10*, *Bhlhe22*, *Kcnd3*, *Scn1b*, *Nptx1*, *Dlg3*, *Baiap2*, *Pdyn*, *Cacnb4*, *Elavl4*, *Grm3*, *Kif1b*, *Dab1*, *Dnm1*. Of these, *Bdnf*, *Gabra1*, and *Pdyn* are genes already associated with alcohol-related behavior.

The “immune response” module (Fig 3) was down-regulated in blood, PFC, and HYP and up-regulated in AMY after CIE. The top five predicted upstream regulators were: *Cst5*, mir-17, betulinic acid, *Mmp3*, miR-17-5p (and other miRNAs w/seed AAAGUGC). Several of the predicted upstream regulators were also members of the module, suggesting these genes may be particularly important regulators of the blood-brain “immune response” module in CIE-

induced alcohol dependence (in decreasing order of significance): *Nup107*, *Rbm5*, *Tlr4*, *Ly96*, *Sf3b1*, *Aim2*, *Tnrc6a*, *Abca1*.

The “protein processing / mitochondrial function” module (Fig 3) was up-regulated in blood and down-regulated in PFC and AMY after CIE. There was not a corresponding module in the HYP. The top five predicted upstream regulators were: *Abcb6*, *Hipk2*, enterotoxin B, *Torin1*, and *Fancd2*. Several of the predicted upstream regulators were also module members suggesting these genes may be particularly important regulators of the blood-brain response to CIE-induced alcohol dependence related to protein processing and mitochondrial function / oxidative phosphorylation (in decreasing order of significance): *Bnip3l*, *Gpx1*, *Irf7*, *Adipor1*, *Uros*, *Mafg*.

Network analysis in male mice

Similar to female mice, the overlap between blood and brain modules was strongest for the “cell-cell signaling” blood-brain module. The genes in the “cell-cell signaling module” in males (Fig 3) tended to be up-regulated in blood and AMY, and down-regulated in PFC after CIE. There was not a corresponding module in the HYP. Similar to female mice, the blood module was enriched with genes already associated with alcohol-related behavior including: *Adcy1*, *Adcy5*, *Adora2a*, *Bdnf*, *Cckbr*, *Cnr1*, *Gabra2*, *Gabra5*, *Gabrd*, *Grin2a*, *Homer2*, *Hrh3*, *Kcnj6*, *Npy*, *Prkce*, *Prkg2*. The top five predicted upstream regulators were *Jak1/2*, *Snca*, *Hdac4*, *Fmr1*, *Htt*. Several of the predicted regulators were also in the cell-cell signaling module (in decreasing order of significance): *Bdnf*, *Tshz3*, *Slc30a3*, *Fezf2*, *Homer2*, *Bhlhe22*, *Prmt8*, *Grm2*, *Fkbp1b*, *Baiap2*, *NpaS3*, *Grm3*, *Nfib*, *Chrm1*, *Gabbr2*, *Npy*. Of these, *Bdnf*, *Homer2*, and *Npy* are genes already associated with alcohol-related behavior.

The “immune response” module in males (Fig 3) was down-regulated in blood, PFC, HYP, and both up- and down-regulated in AMY after CIE. The top five predicted upstream regulators were: miR-124-3p (and other miRNAs with seed AAGGCAC), lipopolysaccharide, *Pgr*, *Rel*, *Btnl2*. Several of the upstream regulators were also in the immune response module (in decreasing order of significance): *Csf1r*, *Stat3*, *Vcan*, *Spi1*, and *Notch1*.

The “protein processing / mitochondrial function” module in males (Fig 3) was up-regulated in blood and both up- and down-regulated in brain after CIE. The top five predicted upstream regulators were: *Abcb6*, 1,2-dithiol-3-thione, *Klf1*, *Rictor*, *St1926*. Several upstream regulators were also module members (in decreasing order of significance): *Bnip3l*, *Fth1*, *Cdc25b*, *Grp*, *Mrpl12*, *E2f1*, *Ola1*, *Tpm1*, *Fancd2*, *Ctsb*, *E2f2*, *Uba1*, *E2f3*, *Cul1*, *Rb1cc1*, *Fzr1*, *Sub1*, *Stx2*, *Rb1*, *Gadd45a*, *Gsk3a*.

The “fatty acid metabolism / peroxisome proliferator activated receptor (PPAR)” module in males (Fig 3) was down-regulated in blood and up-regulated in AMY and HYP. There was not a corresponding module in PFC. The top five predicted upstream regulators were: pirinixic acid, bezafibrate, methotrexate, *Ppara*, *Ehhadh*. Three of the upstream regulators were also in the fatty acid metabolism / PPAR” module: *Adipor2*, *Acox1*, *Nrg4*.

Correlation of blood and brain gene expression levels

Between-subjects analysis. We studied the preservation of mean gene expression levels of the genes between brain and blood (irrespective of treatment). The pairwise scatterplots in S1 Fig related mean expression values in the three brain regions to mean expression values in blood. We found significant correlations (rho range males: [0.67, 0.67], rho range females: [0.50,0.51]) between mean expression in brain and mean expression in blood (S1 Fig). The correlation between blood and brain expression levels was notably higher in males relative to females (S6 Table).

Within-subjects analysis. To determine the genes with expression levels that are correlated between brain and blood, we calculated the within-subject correlation between gene levels in blood and brain. Expression levels of hundreds of genes were significantly correlated between blood and brain even after correcting for multiple comparisons (Tables 1 and S7). Table 1 shows the top ten genes correlated between blood and brain for each brain region. To gain insight into the cellular specificity of the correlated genes, we determined whether any cell type-specific genes were over-represented in each of the gene sets. The enriched cell types are noted in Table 1. The genes correlated between amygdala and blood in both males and females were enriched with microglial markers, the majority of which were negatively correlated between these tissues. The genes correlated between PFC and blood in both males and females were enriched with endothelial markers, with most negatively correlated between blood and PFC. Additionally, the genes correlated between PFC and blood in males were enriched with microglial markers and the genes correlated between amygdala and blood in males were enriched with T cell markers; all were negatively correlated between blood and PFC.

To gain functional insight into the genes correlated between blood and brain, we performed pathway enrichment analysis on each gene set (S8 Table). There was little overlap between the enriched pathways across the correlated gene sets between blood and the different brain regions. For example, cholesterol biosynthesis and ethanol degradation pathways were prominent in genes correlated between female blood and PFC (but no other brain region). However, there were some overlapping pathways between blood and multiple brain regions. Glucocorticoid receptor signaling and regulation pathways were among the most commonly enriched pathways, found in the genes correlated between blood and all 3 brain regions for females and between blood and hypothalamus for males. Interferon signaling and antiviral response pathways were commonly enriched in the genes correlated between blood and hypothalamus and blood and PFC in both sexes. MIF Regulation of Innate Immunity was common to male hypothalamus and female amygdala. IL-17 production and signaling pathways were common to female amygdala and hypothalamus. DNA methylation and transcriptional regulation pathways were common to female amygdala and PFC, and male amygdala. Pathways related to DNA damage responses were common to in the correlated gene sets between blood and multiple brain regions as well (e.g., base excision repair for male amygdala and female hypothalamus, nucleotide excision repair for female amygdala, and role of BRCA1 in DNA damage response for male hypothalamus). Methionine Degradation was common to the PFC of both sexes and female hypothalamus.

Some of the correlated genes were also differentially expressed between CIE and Air mice in both blood and brain (Tables 2–4). For males, *Hsp1a* and *Hsp1b* were correlated between

Table 1. Genes Correlated Between Blood and Brain.

	Females			Males		
	Number of Corr Genes	Enriched Cell Type	Top 10 Corr Genes	Number of Corr Genes	Enriched Cell Type	Top 10 Corr Genes
Amygdala	542	mic	Pacrg, Hgf, Pofut1, Helb, Tmem131, Hpgd, Card19, Clec12a, Setx, Rbm12b2	707	mic, T cells	Zfp521, Gm4799, Ghr, Zfp385a, Kirrel3, Ceacam1, Gm32647, Dcp1a, Zfp217, Sema3f
PFC	634	endo	Smarcc2, Kcnk3, Smim101l, Agps, Rdh11, Hivep2, Raver2, Gm42984, Cbfa2t3, Cep57l1	656	endo, mic	Lurap1l, Tenm4, Mdh1, Elf4, Pyurf, Sod1, Ccdc166, Ube3b, Slc36a1, Snx20
Hypothalamus	649	None	Dglucy, Hsp90aa1, Lnpep, Btaf1, Hmox2, Usp6nl, R3hcc1l, Uap1l1l, Jade1, Taf13	686	None	Hdac11, Arf6, Sdk2, Scamp1, Gm46515, Fgd1, Wwp2, Zscan18, Fam13c, Espn

Mic = microglia, endo = endothelial cells, PFC = prefrontal cortex

<https://doi.org/10.1371/journal.pcbi.1009800.t001>

Table 2. Genes Differentially Expressed between CIE and Air C57Bl/6J Mice in Amygdala with Expression Levels that are also Correlated between Blood and Amygdala.

Male		Female	
Gene Symbol	Gene Name	Gene Symbol	Gene Name
Acaa2	Acetyl-CoA acyltransferase 2	Ccl5	C-C Motif Chemokine Ligand 5
Ccp110	Centriolar Coiled-Coil Protein 110	Cdc42ep2	CDC42 Effector Protein 2
Dennd4b	DENN/MADD domain containing 4B	Dusp10	Dual Specificity Phosphatase 10
Galm	Galactose Mutarotase	Gabpb1	GA Binding Protein Transcription Factor Subunit Beta 1
Hspa1a	Heat Shock Protein Family A (Hsp70) Member 1A	Setx	Senataxin
Hspa1b	Heat Shock Protein Family A (Hsp70) Member 1B	Trim25	Tripartite Motif Containing 25
Hspa5	Heat Shock Protein Family A (Hsp70) Member 5	Vipr1	Vasoactive Intestinal Peptide Receptor 1
Mgst1	Microsomal Glutathione S-Transferase 1		
Pou6f1	POU Class 6 Homeobox 1		
Pygl	Glycogen Phosphorylase L		
S100a6	S100 Calcium Binding Protein A6		
Vav1	Vav Guanine Nucleotide Exchange Factor 1		

<https://doi.org/10.1371/journal.pcbi.1009800.t002>

blood and each brain area and differentially expressed in blood and all brain regions, as was *Pygl* (except for hypothalamus) (Fig 4). Blood and brain levels of *Hspa1a* and *Hspa1b* tended to be negatively correlated with alcohol consumption in the last 2BC drinking test, while *Pygl* had weak correlations with alcohol intake (S2 Fig). For females, *Ccl5* was correlated between blood and each brain area and differentially expressed in blood and all brain regions, as was *Vipr1* (except for hypothalamus) (Fig 4). Blood expression levels of *Ccl5* were positively correlated with alcohol intake, while blood expression levels of *Vipr1* were negatively correlated with alcohol intake (S2 Fig). We created an R Shiny app where interested readers can see the correlation between voluntary ethanol intake levels and the expression level of any gene of interest (https://lauraferguson.shinyapps.io/blood_brain/).

Whole blood gene expression signatures can distinguish CIE and Air subjects

We determined whether CIE status could be predicted by blood gene expression profiles using three different classification techniques: logistic regression (LR) with elastic net regularization,

Table 3. Genes Differentially Expressed between CIE and Air C57Bl/6J Mice in Hypothalamus with Expression Levels that are also Correlated between Blood and Hypothalamus.

Male		Female	
Gene Symbol	Gene Name	Gene Symbol	Gene Name
Chsy1	Chondroitin Sulfate Synthase 1	Aacs	Acetoacetyl-CoA Synthetase
Clec2d	C-Type Lectin Domain Family 2 Member D	Bcl2l11	BCL2 Like 11
Dnajb11	DnaJ Heat Shock Protein Family (Hsp40) Member B11	Caln1	Calneuron 1
Fam13c	Family With Sequence Similarity 13 Member C	Card6	Caspase Recruitment Domain Family Member 6
Gm46515	predicted gene, 46515 (lncRNA)	Ccl5	C-C Motif Chemokine Ligand 5
Hspa1a	Heat Shock Protein Family A (Hsp70) Member 1A	Gpm6b	Glycoprotein M6B
Hspa1b	Heat Shock Protein Family A (Hsp70) Member 1B	Gria1	Glutamate Ionotropic Receptor AMPA Type Subunit 1
Kdsr	3-Ketodihydrospingosine Reductase	Smyd1	SET And MYND Domain Containing 1
Parp9	Poly(ADP-Ribose) Polymerase Family Member 9	Vrk2	VRK Serine/Threonine Kinase
Spred2	Sprouty Related EVH1 Domain Containing 2	Zcchc3	Zinc Finger CCHC-Type Containing 3
Sptssa	Serine Palmitoyltransferase Small Subunit A		

<https://doi.org/10.1371/journal.pcbi.1009800.t003>

Table 4. Genes Differentially Expressed between CIE and Air C57Bl/6J Mice in Blood and PFC with Expression Levels that are also Correlated between Blood and PFC.

Male		Female	
Gene Symbol	Gene Name	Gene Symbol	Gene Name
Ahsa2	Activator of heat shock 90kDa protein ATPase homolog 2	4930549G23Rik	RIKEN cDNA 4930549G23 gene (lncRNA)
Cebpb	CCAAT Enhancer Binding Protein Beta	Atp6v0e2	ATPase H+ Transporting V0 Subunit E2
Dnajb1	Dnaj Heat Shock Protein Family (Hsp40) Member B1	Btbd8	BTB Domain Containing 8
Dnlz	DNL-Type Zinc Finger	Ccl5	C-C Motif Chemokine Ligand 5
Far1	Fatty Acyl-CoA Reductase 1	Efcab5	EF-Hand Calcium Binding Domain 5
Fkbp5	FKBP Prolyl Isomerase 5	Hnrnpf	Heterogeneous Nuclear Ribonucleoprotein F
Hspa1a	Heat Shock Protein Family A (Hsp70) Member 1A	Vipr1	Vasoactive Intestinal Peptide Receptor 1
Hspa1b	Heat Shock Protein Family A (Hsp70) Member 1B		
Insig1	Insulin Induced Gene 1		
Naaladl2	N-Acetylated Alpha-Linked Acidic Dipeptidase Like 2		
Pygl	Glycogen Phosphorylase L		
Zfp329	Zinc finger protein 329		

<https://doi.org/10.1371/journal.pcbi.1009800.t004>

random forest (RF), and partial least squares discriminant analysis (PLSDA). Ideally, in addition to being highly accurate at the classification task, a good model would also be interpretable. We chose these three techniques because they included measures of variable importance which enabled interpretation of the models. Furthermore, because the dataset is small, we wanted to use noncomplex models to avoid overfitting. We used all genes to train the classifiers. Some of the coefficients in the regularized LR model go to zero and those features fall out of the model. Some features in the RF model have zero Mean Decrease in Gini metric and are not included in the final model. In this way the regularized LR and RF methods include embedded feature selection. The PLSDA model includes variable importance measures but does not include built-in feature selection and uses all genes for the model.

For female subjects, the logistic regression model performed best, and was highly accurate at identifying CIE subjects from controls (AUC 90.1%) using only 32 genes (*Gatad2a*, *ENSMUSG00000096544.2*, *Plgrkt*, *E030030I06Rik*, *ENSMUSG00000085633.1*, *Ccdc85b*, *Hsd11b1*, *Gzma*, *Utp14a*, *Zfp472*, *Foxq1*, *Cope*, *Saxo2*, *Ankle1*, *Med10*, *Gm4841*, *Ubn1*, *Sephs1*, *Smarca4*, *Phf6*, *Eri1*, *Aldh7a1*, *Il4i1*, *Pigz*, *Apol11a*, *Chchd6*, *Lrrc8c*, *Mlec*, *Cenps*, *Cct5*, *Apa2*, *ENSMUSG00000109006.2*) (Fig 5A). The partial least squares and random forest techniques were also able to identify CIE subjects from controls with a high degree of accuracy (PLSDA AUC was 80.8% and random forest AUC was 79.2%) (Fig 5A). Random forest required 807 genes to achieve its high performance (Fig 5A).

For male subjects, partial least squares performed best at identifying CIE subjects from controls (AUC 80.5%) (Fig 5C). Logistic regression also performed well (AUC 75.9%) and required only 2 genes (*Plet1* and *Hspa5*) (Fig 5C). The random forest model employed 382 genes. It was not able to distinguish between CIE and Air mice very well (AUC 58.2%) (Fig 5C).

To determine which genes were the most important for classifying the subjects as CIE and Air mice, we examined the variable importance measures from random forest, partial least squares discriminant analysis, and logistic regression. There was little overlap between the top important genes between the different models. For females *Gzma* was common to logistic regression and partial least squares discriminant analysis, *Apol11a* was common to random forest and partial least squares discriminant analysis (*Apol11a* was also one of the 32 genes with non-zero coefficients in the logistic regression model), and *Gatad2a*, *Hsd11b1*, and

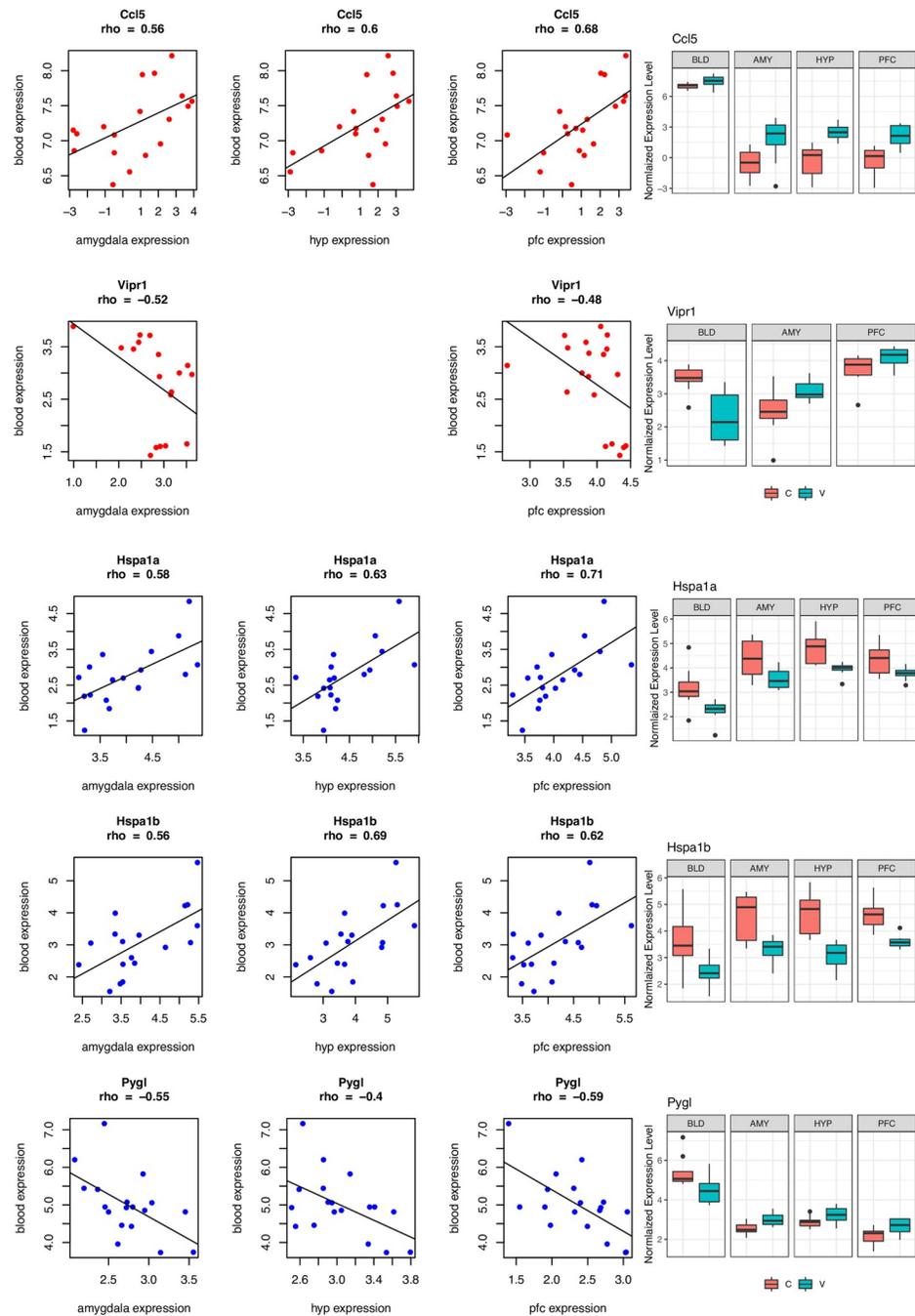


Fig 4. Genes with Expression Levels that are Correlated between Blood and Brain in C57BL/6J Mice. The scatterplots show the relationship between gene expression levels in blood (y-axis) and each brain area (x-axis). Each point in the scatterplot is a subject. Some correlated genes were also differentially expressed between CIE (V = ethanol vapor) and Air (C = control) mice in both blood and brain, which could make them ideal biomarker candidates. For females, *Ccl5* was correlated and differentially expressed in blood and all brain regions. For males, *Hspa1a* and *Hspa1b* were correlated and differentially expressed in blood and all brain regions. The correlation coefficient is shown under the gene name. The boxplots show the normalized gene expression levels by group.

<https://doi.org/10.1371/journal.pcbi.1009800.g004>

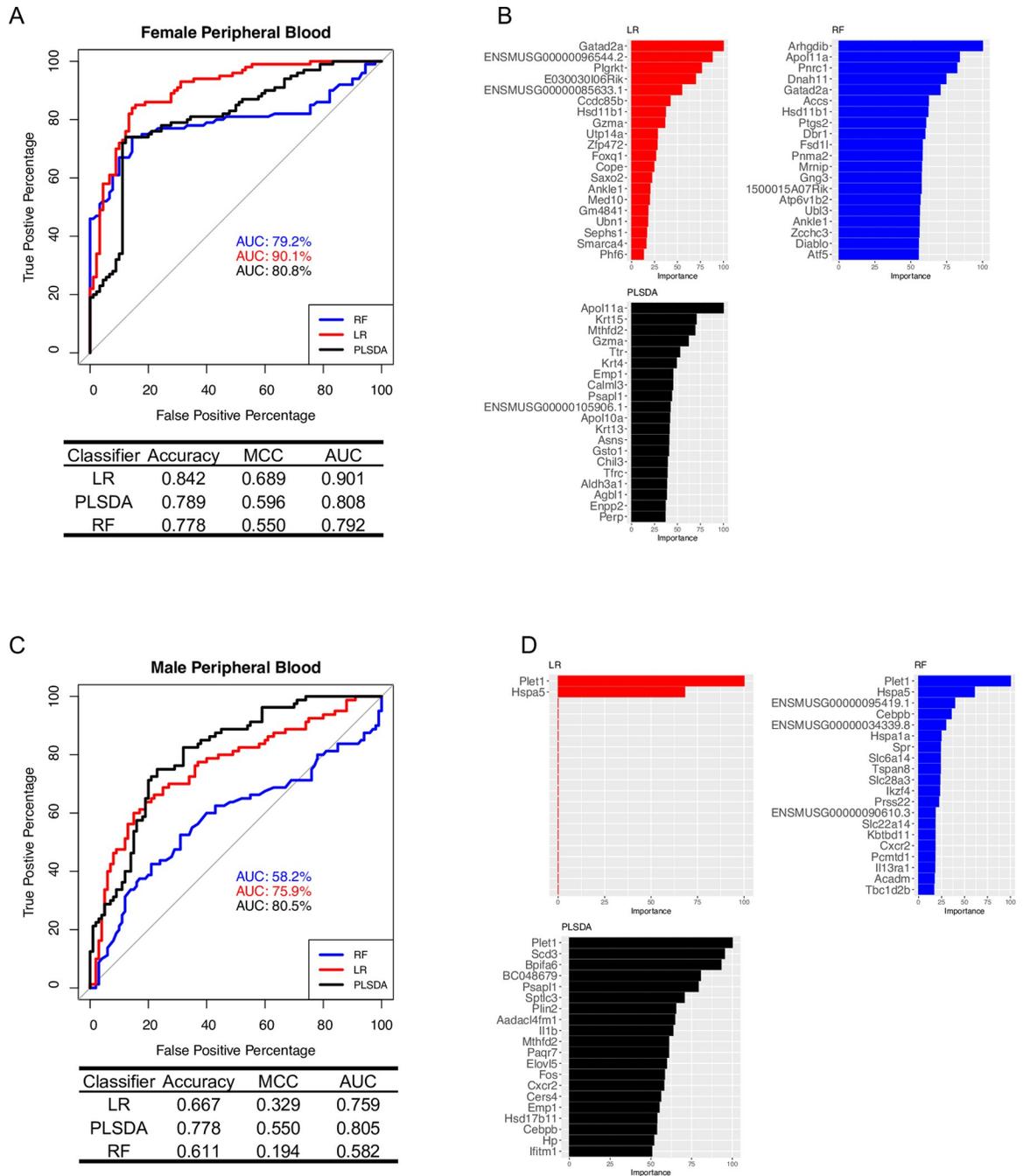


Fig 5. Classification Performance Using Blood Gene Expression. Machine learning classifiers were trained using peripheral whole blood gene expression data to predict alcohol dependence status. Receiver Operating Characteristic (ROC) curves are shown for the classification techniques indicated by the line type in the legend for (A) female and (C) male C57BL/6J mice. The ROC curve shows the relationship between the True Positive Rate (y-axis) and the False Positive Rate (x-axis) resulting from a set of binary classification tests based on each possible decision threshold value. The area under the curve (AUC) for each classifier is displayed in the graph in a corresponding color. The diagonal represents chance levels which corresponds to an AUC of 0.50. A table below the ROC curve shows different performance metrics of the classifiers. The top 20 important genes for the classification task are shown for (B) females and (D) males for the separate models. RF = Random Forest, LR = Logistic Regression, PLSDA = Partial Least Squares Discriminant Analysis, MCC = Matthews Correlation Coefficient.

<https://doi.org/10.1371/journal.pcbi.1009800.g005>

Ankle1 were common to the logistic regression and random forest (Fig 5B). For males, *Plet1* was common to all three classifiers, *Hspa5* was common to logistic regression and random forest, and *Cebpb* and *Cxcr2* were common to random forest and partial least squares discriminant analysis (Fig 5D).

Discussion

Routine blood testing has long been a part of medical care. Genomic profiles in blood could be used to provide data-driven diagnosis of AUD, stratify the heterogeneous AUD patient population for clinical trials, select optimal therapy, determine whether a patient has AUD risk, and monitor the efficacy of the patient's medications. Here we performed a well-controlled animal model study to conduct whole genome profiling of brain and blood with ethanol treatment. An important question this study addressed was whether blood gene expression signatures could predict CIE and Air mice. The highly predictive performance of the classifiers suggests that there is valuable diagnostic information in blood gene expression for CIE-induced alcohol dependence, even one week since the last alcohol exposure. Different features were selected by the different classification methods, which is consistent with the literature [41] and expected because the methods employ different techniques for the classification task and the features in gene expression datasets are highly correlated. These results confirm that good performance in the classification task can be achieved using different sets of features. That (and the dynamic nature of gene expression) might make it challenging to select a consistent panel of biomarkers and it will likely be necessary to incorporate other sources of information into a diagnostics screen. There were however several features ranked as highly important for achieving good performance in multiple models. This included *Gzma*, *Apol11a*, *Gatad2a*, *Hsd11b1*, and *Ankle1* for females and *Plet1*, *Hspa5*, *Cebpb*, and *Cxcr2* for males. *Cebpb* and *Cxcr2* are especially interesting considering the importance of immune molecules in responses to alcohol (discussed below). However, immune-responsive molecules may not be ideal biomarkers given that immune responses can be initiated by infections or other environmental perturbations.

AUROC measures the quality of the model's predictions irrespective of what classification threshold is chosen. The classification threshold strikes a balance between false positive and false negative rates, and the optimal choice would likely depend on the setting and purpose (e.g. primary care screening for problematic drinking versus criminal justice settings) [42]. Moreover, the predictive value of these classifiers will depend on the prevalence of the disorder in the population tested. A population with low prevalence of the disorder (for example, in primary care screening for AUD) results in an increase in false positive tests, whereas a higher prevalence rate (e.g., in an addiction clinic) yields more false-negative tests [42]. Here in the controlled experimental conditions the prevalence was 50/50. Thus, our findings represent important first steps in identifying novel biomarkers for AUD, but more research (including human validation) will be required before genes selected as discriminatory features could become viable biomarkers for AUD.

The mice in this study have been exposed to weeks of high dose ethanol vapor and would likely represent severe cases of AUD. These cases are usually clinically evident by patient history and physical exam and would not necessitate a molecular diagnostic tool. However, social stigma surrounding alcoholism (like most psychiatric disorders) remains a barrier to treatment. An objective molecular diagnostic tool (e.g., a "blood test") would potentially mitigate stigma and encourage patients to seek diagnosis and treatment for this disorder. Additionally, blood expression profiles might detect milder forms of AUD or those at high-risk for developing AUD. The former could be investigated in future studies that include an additional

ethanol-naïve control group or a larger number of animals to allow analysis of low versus high drinkers. The latter could be addressed in future studies by profiling the blood of ethanol-naïve genetic rodent models of AUD-risk traits (e.g., high alcohol preference or consumption, propensity for binge-like drinking), or by analyzing the blood transcriptome before and after CIE treatment and correlating gene expression changes with the amount of drinking escalation within subjects. The blood transcriptome varies between people, but it is relatively stable for an individual across time [43] which is encouraging and suggests that repeated blood sampling could be used for personalized medicine approaches. Studies that analyze gene expression profiles in blood sampled at multiple time points throughout the addiction cycle (including before alcohol exposure) will be critical in exploring the clinical utility of blood gene expression profiles for AUD.

Another question that this study examined was whether expression levels of the genes between brain and blood are preserved. While some studies have shown significant but weak correlations between blood and brain (e.g., [44]), we found that blood and brain average gene expression levels were highly correlated. The difference is likely attributable to the fact that we compared blood and brain samples from the same subjects instead of comparing blood and brain samples from different individuals. Strikingly, our within-subjects design revealed that the expression levels of hundreds of genes were significantly correlated between blood and brain. This was irrespective of treatment, which suggests that the blood might be useful for gaining insight into brain functioning in general. Indeed, there are many compelling examples from the literature that support the idea that whole blood transcriptomes can be informative for a number of brain diseases [45–47].

The transcriptional response to alcohol dependence in brain showed stronger conservation at the level of coexpression than at the level of individual genes. We identified an “immune response” network similarly perturbed in blood and brain after CIE. Immune functions also dominated the IPA Core Analysis results for blood and brain DEGs. This finding adds to the importance of “neuroimmune” signaling in alcohol dependence and reflects the key role white blood cells play in immune responses as they are the source for most of whole blood mRNA. The relationship between alcohol and the immune system has been the subject of intensive research, and much insight has emerged in the last 15 years pointing to a bi-directional relationship between alcohol consumption and immune signaling molecules, whereby alcohol ingestion increases peripheral and central cytokine levels. Conversely, manipulation of cytokines and other immune signaling molecules can increase alcohol consumption and craving [48–52], leading to further increases in cytokine levels and emergence of an out-of-control, positive-feedback cycle. This cycle demonstrates the systemic nature of AUD and the involvement of peripheral and central immune signaling in AUD pathophysiology [7].

The details of peripheral-central immune signaling crosstalk involved in alcohol dependence remain to be fully understood. One way alcohol triggers inflammatory responses is via reactive oxygen species produced during ethanol metabolism, which can occur both peripherally and centrally. Alcohol also increases intestinal permeability (“leaky gut”), which permits gut-derived bacterial products to enter the circulation where they are recognized by immune cells in blood or target organs, resulting in the release of pro-inflammatory cytokines [53]. Peripheral cytokines can trigger central “immune” responses via vagal afferents or by crossing the blood brain barrier to enter the brain. Although these are plausible mechanisms, it remains to be shown the exact peripheral-central communication that must occur, which likely involves multiple mechanisms. Our IPA analysis predicted an increase in leukocyte migration and number in brain and a decrease in blood. The cell type enrichment analysis of the DEGs revealed an up-regulation of leukocyte genes in brain and a down-regulation in blood. Also, the genes correlated between blood and brain were enriched with endothelial, microglia, and

leukocyte genes and tended to be negatively correlated between blood and brain. Taken together, these findings suggest that white blood cells could be recruited into the CNS during alcohol dependence. Future studies are planned to study this novel hypothesis as an unexplored mechanism of peripheral-central immune crosstalk in alcohol dependence.

One of the most surprising findings of this study was that the most highly conserved co-expression networks for both sexes were related to cell-cell signaling and included such “brain-related” categories as glutamate and GABA receptor signaling. The importance of these systems in the CNS has been long-known for alcohol dependence (reviewed in [54]), but our results suggest that *blood* GABA and glutamatergic signaling are also perturbed in alcohol dependence. This has been partially validated by previous studies that found that GABA serum levels are lower and glutamate serum levels are higher in alcoholic patients compared with non-dependent controls during alcohol withdrawal [55, 56]. GABA regulates the secretion of cytokines from PBMCs in a concentration-dependent manner [57] providing another novel potential link between alcohol and immune regulation suggested by this study. Moreover, glutamate serum levels upon hospital admission are predictive of developing an alcohol withdrawal syndrome 12 h later [55]. Glutamate serum levels are correlated with brain levels [58] and are predictive of other brain diseases such as multiple sclerosis [59], schizophrenia [60], and autism [61]. The cell-cell signaling blood-brain module also contained numerous genes that modulate alcohol consumption (as determined by mutant mouse studies). Of these *Adcy1*, *Bdnf*, *Cckbr*, *Cnr1*, *Gabra5*, *Gabrd*, *Grin2a*, and *Hrh3* were common to both sexes and *Bdnf* was also a hub in the Graphical Summary for the female blood DEGs in Fig 2C. Because most of the mutant mouse studies were global knockouts and these genes are ubiquitously expressed, this result calls into question the brain-specificity of the knockout findings. Perhaps there is a larger peripheral component to the causal effects of these genes on alcohol consummatory behavior than previously appreciated.

We identified a “protein processing / mitochondrial function” blood-brain module in both sexes. This included genes involved in protein ubiquitination (including ubiquitin B, numerous ubiquitin conjugating enzymes, ubiquitin specific peptidase, heat shock factors), mitochondrial dysfunction (including *Pink1*, *Fis1*, and many NADH:ubiquinone oxidoreductase subunits), and ethanol degradation (*Cat*, *Aldh1a1*). It is possible that this conserved co-expression network represents a response to ethanol-induced cellular stress. For males, genes coding for the Heat Shock Protein Family A (Hsp70) subunits, *Hspa1a* and *Hspa1b*, were upregulated and correlated between blood and all three brain areas tested. Another Hsp70 member, *Hspa5*, was downregulated in all male tissues, correlated between blood and brain, and also an important gene for the classification tasks as discussed above. *Hspa1a*, *Hspa1b*, and *Hspa5* code for potent anti-inflammatory proteins that can initiate protective responses to stress. It has been postulated that the cardioprotective effect of alcohol consumption is due in part to increased intracellular HSPA1A [62]. Acute alcohol exposure induces *Hspa1a* in human monocytes and is required for inhibition of TLR4/MyD88 (but not TLR4/TRIF) signaling in macrophages [63]. *Hspa1a* and *Hspa1b* transcripts are also dysregulated after ethanol exposure in rodent in total brain homogenates [64–66], astrocytes [67, 68], and microglia [68]. Depending on the time point at which gene expression was assayed and perhaps other differences between the protocols employed, these studies have shown increases or decreases in their transcript abundance. Patients with alcoholic hepatitis exhibit lower gene expression levels of HSPA1A in liver compared with healthy controls [69]. *In utero* exposure to ethanol increases *Hspa1a* in cortical tissue from humans and mice [70]. *Hspa5* is linked to alcohol consumption and withdrawal in rodents [71–75] and induced by chronic alcohol in cell lines [76]. In addition to their roles as protein chaperones, *Hspa1a* and *Hspa1b* are known splicing factors [77]. Another splicing factor in the Hsp70 family, *Hspa6*, is drastically increased in postmortem frontal

cortex and amygdala samples from AUD patients and thought to be important for the observed genome-wide changes in splicing observed in these brain regions with AUD [77]. From these examples, it is clear that Hsp70 subunits are responsive to ethanol in both brain and blood, and our analysis suggests that levels in blood are reflective of those in brain and might be useful biomarkers of alcohol dependence.

In females, transcripts for C-C Motif Chemokine Ligand 5 (*Ccl5*) were increased with CIE and levels of expression were correlated between blood and brain in all three brain regions. *Ccl5* transcripts are also increased in total homogenate and astrocytes in male mouse cortex after ethanol vapor exposure [68]. CCL5 transcripts are reduced in the central amygdala of human alcoholic subjects compared with controls [78] as well as in the brains of ethanol-naïve rodent lines that drink high amounts of alcohol [79]. In females, transcripts for the Vasoactive Intestinal Peptide Receptor 1, *Vipr1*, were reduced in blood and increased in the amygdala and PFC, with *Vipr1* expression levels negatively correlated between blood and these brain regions. VIPRI transcripts are also increased in the frontal cortex of human alcoholic subjects compared with controls [78] and reduced in microglia from mouse PFC after CIE [68]. There is a single nucleotide polymorphism in *VIPRI* that is associated with bipolar disorder [80]. *VIPRI* is also a hub gene that acts as a prognosis and progression biomarker for hepatocellular carcinoma [81].

Consistent with previous studies of female mice and rats [82–87], we found a less robust escalation of voluntary ethanol intake after CIE in females compared with males. This finding could be partially due to a ceiling effect (the female mice begin drinking at levels about three times higher than males before CIE). Nevertheless, our study revealed that female mice display a strong molecular phenotype after CIE. Notably, the transcriptome signature in female peripheral blood was able to discriminate between CIE and Air female mice with very high accuracy, even more-so than in male mice. Therefore, although female mice did not exhibit a robust escalation in voluntary drinking after CIE, they showed transcriptional changes that were particularly strong. Network analysis and IPA analysis showed that females and males had similar blood-brain coexpression modules affected by alcohol dependence, suggesting that this molecular phenotype could be comparable between the sexes, at least at the level of gene networks. One exception was the nuclear hormone receptor signaling pathways and blood-brain modules we observed to be affected by CIE across all male tissues which was not observed in female mice. There have been very few studies investigating the transcriptional response to alcohol dependence in females [88–90], and while the focus of these studies as well as the present study is not on sex differences, further investigation into the sex-specific responses to CIE is warranted.

There were a number of methodological choices we made for this study, such as the tissue type (whole blood), molecular measurement (mRNA), and model of alcohol dependence (C57BL/6J mice undergoing CIE) that impose limitations and should be considered. There are several accessible tissues that have been compared to brain in previous studies that we could have assessed, including saliva [91] and blood fractions (e.g., exosomes [92, 93], PBMCs [94, 95], or plasma/serum [96–98]). We chose to measure whole blood because it has been shown to contain disease-relevant information for other brain diseases, has less processing steps than for blood fractions, and contains both cellular and extracellular RNAs so would likely capture a signal from each of the aforementioned blood fractions. However, the various blood cell types could have differential responses to alcohol dependence which could be explored using single cell RNA-seq. We chose to measure mRNA, but there are other molecular measurements, including other RNAs such as miRNAs [91, 92, 97, 99, 100], histone modifications (epigenome) [101, 102], microbiome [103, 104], and the metabolome [105–108]. Furthermore, we assayed the transcriptome at a single time point, and gene expression is a highly dynamic

process. Future studies that profile other molecular responses and at different time points will undoubtedly contribute to our understanding of alcohol dependence and incorporating these additional molecular markers into the machine learning models could further improve their predictive ability. In this study we subjected C57BL/6J mice to CIE which is a commonly used animal model of alcohol dependence. There are also genetic rodent models of behavioral characteristics that put one at risk for developing AUD (such as binge-like drinking, or high alcohol preference), that could be analyzed to determine whether blood and brain responses are conserved in those models and if blood could be used to as an objective molecular tool to predict individuals at risk of developing AUD. We employed an animal model in this study to enable the comparison of blood and brain samples from the same subjects (which is not possible in humans). This study established a link between blood and brain responses in an animal model of alcohol dependence. Going forward, we will investigate whether there is a similar signature of alcohol dependence in the blood of human patients with AUD.

Our data indicate that a molecular signature of CIE-induced alcohol dependence can be detected in peripheral blood at least one week after the last alcohol exposure. This signature partially reflects brain pathophysiology and can be used to discriminate CIE and Air mice with a high degree of accuracy. Although it is not possible to draw any clinical conclusions from this dataset, our study establishes a link between blood and brain responses to CIE exposure and demonstrates that blood profiles can distinguish CIE and Air animals, which suggests that blood samples could contain information relevant to alcohol use. These findings represent important first steps in identifying novel biomarkers for AUD and provide critical context for future blood-based biomarker studies. However, more research, including human validation, will be required before genes selected as discriminatory features could become viable biomarkers. We hope this information helps drive objective diagnosis, medication development, and personalized medicine approaches for AUD and other diseases where brain is the primary affected tissue.

Materials and methods

Ethics statement

All procedures were IACUC approved and met the guidelines of the National Institute of Health detailed in the Guide for the Care and Use of Laboratory Animals.

Animals

Adult male and female C57BL/6J mice (Jackson Laboratories, ME) were used in this study (N = 10/sex/group; 40 mice total). One female mouse in the control group and two male mice in the ethanol vapor group died during the experiment leaving N = 19 female mice and N = 18 male mice as the final subjects. Mice were housed at The Scripps Research Institute, four per cage (except during the 2-hr drinking sessions), separated by sex in standard plastic cages under a reversed 12-h light/dark period (lights on at 8:00 PM), with food (Teklad Global 18% Protein Rodent Diet, Envigo) and water available ad libitum.

Chronic Intermittent Ethanol (CIE) model of alcohol dependence-induced escalation of drinking

CIE was implemented as described in previous studies [25–27, 32, 36, 37, 109–112]. Thirty minutes before the dark cycle, mice were singly housed for two hours with access to two drinking tubes, one containing 15% ethanol and the other containing water. Ethanol and water consumption during these 2-hour periods was recorded. Following this baseline period of two-

bottle choice (2BC) drinking, which lasted 20 days (5 days per week for 4 weeks), mice were divided into two balanced groups with similar distributions of ethanol and water consumption. One group was exposed to intermittent ethanol vapor and the other to control air in identical chambers. The ethanol vapor group was administered 1.75 g/kg ethanol plus 68.1 mg/kg pyrazole to inhibit alcohol dehydrogenase, and then was placed in the chambers to receive intermittent vapor for 4 days (16 hours vapor on, 8 hours vapor off). Once per week, immediately following a 16-hour bout of vapor, mice were removed, and tail blood was sampled for blood alcohol determination. Target blood alcohol levels were 175–250 mg%. Following the fourth day of exposure, mice were allowed 72 hours of undisturbed time. The mice were then given 5 days of 2BC drinking. The control group was injected with 68.1 mg/kg pyrazole in saline and placed in chambers delivering air for the same periods as the ethanol vapor group and then received 2BC testing at the same time as the vapor groups. Mice were subjected to four cycles of vapor or air exposure followed by 5 days of 2BC drinking. All mice then received one final 4-day vapor or air exposure (without 2BC testing). One week later, mice were euthanized and tissue collected for RNA sequencing.

Blood ethanol analysis for vapor-exposed mice

Approximately 40 μ l blood was obtained by cutting 0.5 mm from the tip of each mouse's tail with a clean surgical blade. Blood was collected in capillary tubes and emptied into Eppendorf tubes containing evaporated heparin and kept on ice. Samples were centrifuged and plasma decanted into fresh Eppendorf tubes. Plasma (5 μ L) was injected into an Agilent 7820A GC coupled to a 7697A (headspace-flame-ionization). Results were compared with and calibrated using a 6-point serial diluted calibration curve of 300 mg/dl ethanol (Cerilliant E-033).

Statistical analysis of 2-bottle choice drinking

Average ethanol intake (g/kg) was calculated across 5 drinking days of each week during the baseline-drinking period. During the testing cycles, mice also drank for 5 days; therefore, average drinking across these 5 days was used to represent drinking during each CIE cycle. Differences in drinking were determined by Two Way ANOVA (treatment x time (i.e., cycle)) followed by planned Dunnett's tests to determine whether drinking during each testing cycle was different from baseline drinking levels. We also used a One Way ANOVA (time (i.e., cycle)) for the ethanol vapor group and air group separately followed by planned Dunnett's tests to determine whether drinking during each testing cycle was different from baseline drinking levels. Statistical analysis was implemented with GraphPad Prism 8.3.0 (GraphPad Software, San Diego, CA, USA).

Tissue collection

Tissue was collected one week following the final vapor/air exposure between 10 AM and 12:30 PM. Mice were anesthetized with isoflurane and approximately 200–250 μ l blood was collected from the retroorbital sinus in capillary tubes and emptied into Eppendorf tubes containing 3x volume DNA/RNA Shield (Zymo Research, Irvine, CA, USA). Tubes were immediately vortexed on the highest speed for 30 seconds and remained at room temperature for about one hour before being placed at -80°C until further processing. Mice were transcardially perfused with ice-cold phosphate buffered saline (40 mL over 15 min). The mice were then decapitated, and the brains removed and flash frozen in liquid nitrogen. Brain and blood samples were kept at -80°C before being placed on dry ice and shipped to the Dell Medical School (Austin, TX) for further processing.

Dissection of brain areas

Brain samples were frozen in Optimal Cutting Temperature (OCT) media in isopentane on dry ice and were stored at -80°C until sectioning. On the day of sectioning, brains were transferred to a cryostat set at -6 to -10°C for at least 2 h. Sections ($300\ \mu\text{m}$) were collected from $+3.00$ to -4.00 mm (AP) relative to bregma and transferred to glass slides that had been pre-cooled on dry ice. The slices between the following coordinates, relative to bregma, were used for prefrontal cortex (PFC): $+3.0$ mm to $+1.70$ mm (AP), amygdala (AMY): -1.00 mm to -2.30 mm (AP), hypothalamus (HYP): -1.30 mm to -2.30 mm (AP). We performed bilateral micro-punch sampling on a frozen stage (-20 to -25°C) using a 1.00 mm diameter punch (Stoelting Co., item #57397, Wood Dale, IL, USA) to include all sub-regions of the brain areas according to the stereotaxic atlas of Paxinos and Franklin [113]. Punches were stored at -80°C until RNA extraction.

RNA extraction

Total RNA was extracted from the whole blood samples using the Quick-RNA Whole Blood kit according to manufacturer instructions (Zymo, Irvine, CA, USA). Total RNA was extracted from the brain micropunches using the PureLink RNA Mini Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to manufacturer instructions except that 2-mercaptoethanol was not added to the lysis buffer as this was found to reduce reads in downstream analyses during optimization tests. RNA purity, concentration, and integrity were determined using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), Qubit 4 (Thermo Fisher Scientific, Waltham, MA, USA), and TapeStation 4150 (Agilent Technologies, Santa Clara, CA, USA). For quantitative assessment of RNA sequencing, External RNA Controls Consortium (ERCC) RNA Spike-In Mix 1 (Thermo Fisher) were added to samples. We submitted $300\ \text{ng}$ total RNA/sample for brain samples and $2000\ \text{ng}$ total RNA/sample for whole blood samples to the Genomic Sequencing and Analysis Facility at UT Austin for sequencing. The extracted RNA was of sufficient amount and quality (S9 Table), and subsequent sequencing accurately quantified the RNA as assessed by strong correlations between the actual concentration and observed counts of the ERCC spike-in control ($R^2 = .90$ (PFC), 0.91 (HYP), 0.92 (AMY), 0.85 (blood)).

RNA sequencing

We surveyed gene expression using TagSeq by constructing libraries directed at the 3' ends of mRNA according to established protocols [114]. About 70% of mRNA from whole blood is globin mRNA, which limits the ability to detect other genes. Therefore, we depleted globin mRNA in the whole blood samples before library construction using the QIAseq Fast Select RNA removal kit (Qiagen, Germantown, MD, USA) according to manufacturer instructions. Before globin depletion, the blood samples were purified using an AMPure XP (Agencourt) bead clean-up as our optimization tests revealed this improved the blood library concentrations. Purified libraries were size-selected to obtain 350 – 550 bp target cDNA and pooled in equal proportions following relative quantification using a qPCR assay. Pooled sets of libraries were sequenced using an Illumina NovaSeq 6000 generating 1×100 bp single-end reads. Illumina adapters were removed, and lane pools were de-multiplexed according to the barcode assignments by the sequencing facility prior to return of raw data.

Bioinformatics analysis

The mean read count was approximately 5 million reads per sample. This read depth provides full-coverage detection with TagSeq (which is $\sim 12,000$ – $15,000$ unique genes, depending on the tissue). The reads were trimmed, de-duplicated, and quality filtered using custom Perl scripts

from https://github.com/z0on/tag-based_RNAseq. The data were mapped to the GRCm38/mm10 mus musculus reference transcriptome supplemented with the sequences of the ERCC spike-in control mRNAs using bowtie2 [115]. Then samtools [116] and htseq-count [117] were used to sort the mapped reads and convert mapped reads to counts. Read distributions were characterized using the read_distribution.py program from RSeQC package v2.6.2. The genome annotations for the analyses were taken from GENCODE using the main file containing comprehensive gene annotation on the reference chromosomes only. Raw counts were converted to log 2-counts per million (log-CPM) to account for library size differences and then transformed using the voom function from the limma package (version 3.42.2) to account for the heteroscedastic nature of count data [118].

Differentially expressed genes (DEGs) between CIE and Air mice were identified using the limma package (version 3.42.2) [118]. Next the transcriptional responses were parsed into groups of genes whose expression patterns are highly correlated (modules) using the Weighted Gene Coexpression Network Analysis (WGCNA) package (version 1.69) [119]. We built coexpression networks for each tissue individually as we have done previously (e.g., [40, 120, 121]). All detected genes within a tissue were used for network construction. The Pearson correlation coefficient between all pairs of probes across all samples was calculated, and a signed gene coexpression similarity matrix between genes was generated: $S_{ij} = (1 + \text{cor}(x_i, x_j)) / 2$. Then, an adjacency matrix, $a_{ij} = S_{ij}^\beta$, was used to assess gene connections. Power (β) was chosen so that the resulting network exhibited approximate scale-free topology (for male PFC, AMY, HYP, and blood $\beta = 6, 8, 7$, and 10, and for female PFC, AMY, HYP, and blood $\beta = 5, 9, 11$, and 6). Next, the topological overlap measure (TOM) was used to calculate the relative interconnectedness of a gene pair. Average linkage hierarchical clustering was applied to produce a dendrogram based on the topological overlap dissimilarity ($1 - \text{TOM}$). Branches of the tree were cut using a dynamic tree cut algorithm to detect modules (deepSplit = T, minimum module size = 100, cut height = 0.99). Similar modules were merged together (cutHeight = 0.25). The terms “modules,” “clusters,” and “gene networks” are used interchangeably in the manuscript and refer to groups of genes with highly correlated expression levels across samples.

Coexpression modules were said to be related to alcohol dependence if: (1) the module contained more DEGs than expected by chance (hypergeometric test $p < 0.05$); or (2) the module eigengene was correlated with alcohol dependence status, alcohol consumption, or alcohol preference during the voluntary drinking sessions. The module eigengene is defined as the first principal component of the expression matrix of the corresponding module and can be thought of as the summary of gene expression within each module.

Blood and brain are heterogenous tissues comprised of multiple cell types. To gain insight into functional roles for particular cell types, we determined whether cell type-specific genes were enriched in the DEGs and alcohol-related modules using cell type signatures from the literature and the userListEnrichment function from the WGCNA package in R. The brain cell type markers included six major cell types in the brain: astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and oligodendrocyte progenitor cells (OPCs) [122]. The immune cell type markers included seven immune populations: B cells, plasma cells, monocytes, macrophages, neutrophils, NK cells, and T cells [123]. Terms with Bonferroni-corrected hypergeometric $p < 0.05$ were considered significantly enriched within the dataset.

To determine what functional impact the observed transcriptional responses might have, we performed enrichment analysis on the DEGs and alcohol-related modules using the Ingenuity Pathways Analysis knowledgebase (IPA, Ingenuity Systems, www.ingenuity.com), a web-based software application. We performed a core analysis for the DEGs and modules of interest in IPA using default settings, except that expressed transcripts were used as the

background population for the right-tailed Fisher exact test (FET) calculations. Terms with FET $p < 0.05$ were considered significantly enriched within the dataset.

Blood and brain gene expression comparisons

A major goal of this study was to compare the transcriptional response to CIE-induced alcohol dependence between blood and brain. We compared DEGs in whole blood and each brain area and identified overlapping DEGs between tissues. We determined whether a greater number of DEGs were shared between blood and brain than expected by chance using the hypergeometric test (hypergeometric $p < 0.05$ was considered significant). In addition to comparing the DEGs, we also compared the gene coexpression modules that were related to CIE between blood and brain following previously described approaches [78, 120, 121, 124, 125]. Briefly, for each pair of networks, the overlap between all possible pairs of modules was calculated, and the significance of module overlap was assessed using a one-sided hypergeometric test. Blood and brain modules sharing a significant number of genes (hypergeometric $p < 0.05$), were considered conserved and we refer to them as blood-brain modules throughout the manuscript. The software Cytoscape (<http://www.cytoscape.org/>) was used to visualize the comparisons and create a meta-network of highly overlapping CIE modules.

Another goal of the study was to determine whether blood gene expression levels are correlated with (i.e., predictive of) brain gene expression levels. Most previous studies comparing blood and brain gene expression in humans have used blood and brain data from different individuals (brain samples collected postmortem and blood samples collected from living patients). To make our results more comparable to previous studies, we analyzed our data in a similar manner (between-subjects design). For the between-subjects correlation, each point in the scatterplot represented a gene, and we plotted the normalized expression level of the gene averaged across subjects. We then calculated the Spearman correlation coefficient between blood and brain normalized gene expression levels using the `corr.test` function from the `psych` package (version 2.0.7) in R Studio. We compared the correlations between brain and blood for males and females using the `cocor` package (version 1.1–3) in R Studio [126].

We measured gene expression levels for brain and blood for the same animal which permits a within-subjects comparison. For the within-subjects analysis, we plotted the normalized expression level for a gene in brain and blood for the same subject, where each point in the scatterplot represented a subject. We then calculated the Spearman correlation coefficient between blood and brain normalized gene expression levels for each gene using the `corr.test` function from the `psych` package (version 2.0.7) in R Studio. There were over 10,000 correlation calculations performed, one for each gene. To account for multiple comparisons, we used Holm-Bonferroni correction, and corrected $p < 0.05$ was considered significant [127].

Classification algorithms and parameter selection

We determined how well blood gene expression levels could discriminate between CIE and Air subjects using three classification models that are exemplary at identifying patterns or trends in ‘omic’ data and include measures of variable importance to enable interpretation of the models: logistic regression (LR) with elastic net regularization, random forest (RF), and partial least squares discriminant analysis (PLSDA). Repeated cross validation was used to choose the optimal hyperparameters (5-fold cross validation repeated 10 times) for each model. The optimal value for the RF parameter `mtry` (the number of variables available for splitting at each tree node) was chosen to be 638 for males and 197 for females. The LR regularization parameters are `alpha` (this parameter balances the amount of emphasis given to minimizing Residual Sum of Squares versus minimizing sum of square of coefficients) and `lambda`

(regularization penalty). Alpha was 0.1 and lambda was 0.135 for males. Alpha was 1 and lambda was 0.341 for females. The PLS parameter *ncomp* (number of components to include in the model) was three for males and four for females. Model training and evaluation was implemented with the MLSeq package (version 2.8.0) in R version 4.0.3 [128].

Classifier performance evaluation

To evaluate the performance of the classifiers in assigning the correct label to each subject (CIE versus Air) we used 5-fold cross validation. MLSeq takes a matrix of raw counts as the input and performs normalization within-fold so that the normalization of the test fold is performed using coefficients estimated from the training folds. The process of randomly splitting samples into training/test folds, training the models, and then testing performance was repeated 10 times to obtain estimates of model performance. The performance metrics were averaged across the repeated folds. The performance metrics we calculated were accuracy (percentage of correct assignments), area under the Receiver Operating Characteristic curve (AUROC), and Matthews correlation coefficient (MCC). Given a test set and a specific classifier, each decision can be categorized as one of the following: (1) a positive example classified as positive (true positive; TP), (2) a positive example misclassified as negative (false negative; FN), (3) a negative example classified as negative (true negative; TN), (4) a negative example misclassified as positive (false positive; FP). After creating a Contingency Table, a 2×2 matrix with the columns as true classes and the rows as the hypothesized classes, we calculated accuracy (the fraction of predictions the model assigned correctly) as follows: $\frac{TP+TN}{TP+TN+FP+FN}$, sensitivity (the True Positive Rate, or the proportion of positives that are correctly identified): $\frac{TP}{TP+FN}$, specificity (the True Negative Rate, or the proportion of negatives that are correctly identified): $\frac{TN}{TN+FP}$, and the False Positive Rate: $1 - \frac{TN}{TN+FP}$. To plot the ROC curve, we calculated the False Positive Rate and True Positive Rate under different classification thresholds, then quantified the performance of the classifiers by calculating the area under the ROC curves (AUROC) using the *roc* function from the *pROC* package (version 1.16.2) in R. The higher the AUROC, the better the classifier performance. AUROC of 0.5 is random, 1.0 is perfect, and 0.7–0.8 is generally considered high performance [129].

We also calculated the Matthews correlation coefficient (MCC) as: $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$. MCC is essentially a correlation coefficient between the observed and predicted binary classifications; it returns a value between -1 and +1. A coefficient of +1 represents a perfect prediction, 0 no better than random prediction and -1 indicates total disagreement between prediction and observation.

Variable importance measures

Logistic regression, random forest, and partial least squares discriminant analysis provide measures of feature importance, which enable the identification of which genes were the most useful for discriminating between the CIE and Air subjects. When building a random forest, useful genes will split the mixed labeled nodes into pure single class nodes. This measure is termed Gini importance. In PLS-DA, the loading vectors (which are coefficients assigned to each gene to define each component) are obtained so that the covariance between a linear combination of the genes and the class label is maximized. The variable importance measure for PLS-DA is based on the weighted sums of the absolute regression coefficients (the loading vectors). The weights are proportional to the reduction in the sums of squares. For LR, the importance measure is the regression coefficient. Importance measures were extracted from the final models using the *varImp* function from the *caret* package (version 6.0–86).

Supporting information

S1 Fig. Comparison of blood and brain mean gene expression levels. The scatterplots display the relationship between blood (x-axis) and brain (y-axis) mean gene expression levels for male (left) and female (right) mice. Each point in the scatterplot represented a gene, and the normalized expression level of the gene averaged across subjects (irrespective of treatment) are plotted. We then calculated the Spearman correlation coefficient between blood and brain normalized gene expression levels, and this value is displayed in the plots (ρ).
(TIF)

S2 Fig. Correlation of blood and brain gene expression levels with alcohol consumption. The scatterplots display the relationship between alcohol consumption in the final limited access two bottle choice drinking test (y-axis) and blood or brain gene expression levels (x-axis) for the genes presented in Fig 4 for female (A) and male (B) mice. Each point in the scatterplot represented a subject. We then calculated the Pearson correlation coefficient between the normalized gene expression levels and alcohol intake and this value and associated p-value is displayed under the x-axis.
(TIF)

S1 Table. The full differential expression results.
(XLSX)

S2 Table. The cell type enrichment analysis results of the DEGs.
(XLSX)

S3 Table. The full IPA pathway enrichment analysis and upstream regulator analysis for the differentially expressed genes (CIE versus Air, $p < 0.05$).
(XLSX)

S4 Table. The full IPA upstream regulator results for the genes in the blood-brain modules.
(XLSX)

S5 Table. The full IPA pathway enrichment results for the genes in the blood-brain modules.
(XLSX)

S6 Table. Comparison of the correlation coefficients between blood and brain mean gene expression levels in females and males.
(XLSX)

S7 Table. List of genes whose expression levels are correlated between brain and blood (Holm-corrected $p < 0.05$).
(XLSX)

S8 Table. The full IPA upstream regulator results for the genes whose expression levels are correlated between brain and blood.
(XLSX)

S9 Table. RNA quality and quantity assessments.
(XLSX)

Acknowledgments

We thank Dr. Dennis Wylie for providing feedback on the machine learning component of the analysis. We would also like to acknowledge The Scripps Research Institute's Animal Models Core for providing the research space, tools, and equipment needed for running the chronic intermittent ethanol exposure procedure.

Author Contributions

Conceptualization: Laura B. Ferguson.

Formal analysis: Laura B. Ferguson.

Funding acquisition: Laura B. Ferguson, R. Dayne Mayfield, Robert O. Messing.

Investigation: Laura B. Ferguson, Amanda J. Roberts.

Methodology: Laura B. Ferguson, Amanda J. Roberts.

Project administration: Laura B. Ferguson, Amanda J. Roberts, R. Dayne Mayfield, Robert O. Messing.

Supervision: R. Dayne Mayfield, Robert O. Messing.

Visualization: Laura B. Ferguson.

Writing – original draft: Laura B. Ferguson.

Writing – review & editing: Laura B. Ferguson, Amanda J. Roberts, R. Dayne Mayfield, Robert O. Messing.

References

1. Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, et al. Epidemiology of DSM-5 Alcohol Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA psychiatry*. 2015.
2. SAMHSA. Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-48. Rockville, MD2014.
3. Degenhardt L, Glantz M, Evans-Lacko S, Sadikova E, Sampson N, Thornicroft G, et al. Estimating treatment coverage for people with substance use disorders: an analysis of data from the World Mental Health Surveys. *World Psychiatry*. 2017; 16(3):299–307. <https://doi.org/10.1002/wps.20457> PMID: 28941090
4. American Psychiatric Association., American Psychiatric Association. DSM-5 Task Force. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Washington, D.C.: American Psychiatric Association; 2013. xlv, 947 p. p.
5. Liu J, Lewohl JM, Harris RA, Iyer VR, Dodd PR, Randall PK, et al. Patterns of gene expression in the frontal cortex discriminate alcoholic from nonalcoholic individuals. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2006; 31(7):1574–82. <https://doi.org/10.1038/sj.npp.1300947> PMID: 16292326
6. Ferguson LB, Ozburn AR, Ponomarev I, Metten P, Reilly M, Crabbe JC, et al. Genome-Wide Expression Profiles Drive Discovery of Novel Compounds that Reduce Binge Drinking in Mice. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2017. <https://doi.org/10.1038/npp.2017.301> PMID: 29251283
7. Gonzalez-Reimers E, Santolaria-Fernandez F, Martin-Gonzalez MC, Fernandez-Rodriguez CM, Quintero-Platt G. Alcoholism: a systemic proinflammatory condition. *World J Gastroenterol*. 2014; 20(40):14660–71. <https://doi.org/10.3748/wjg.v20.i40.14660> PMID: 25356029
8. Liew CC, Ma J, Tang HC, Zheng R, Dempsey AA. The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. *J Lab Clin Med*. 2006; 147(3):126–32. <https://doi.org/10.1016/j.lab.2005.10.005> PMID: 16503242

9. Crabbe JC, Phillips TJ, Harris RA, Arends MA, Koob GF. Alcohol-related genes: contributions from studies with genetically engineered mice. *Addiction biology*. 2006; 11(3–4):195–269. <https://doi.org/10.1111/j.1369-1600.2006.00038.x> PMID: 16961758
10. Mayfield J, Arends MA, Harris RA, Blednov YA. Genes and Alcohol Consumption: Studies with Mutant Mice. *International review of neurobiology*. 2016; 126:293–355. <https://doi.org/10.1016/bs.irm.2016.02.014> PMID: 27055617
11. Glatt SJ, Everall IP, Kremen WS, Corbeil J, Sasik R, Khanlou N, et al. Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(43):15533–8. <https://doi.org/10.1073/pnas.0507666102> PMID: 16223876
12. Le-Niculescu H, Kurian SM, Yehyawi N, Dike C, Patel SD, Edenberg HJ, et al. Identifying blood biomarkers for mood disorders using convergent functional genomics. *Mol Psychiatry*. 2009; 14(2):156–74. <https://doi.org/10.1038/mp.2008.11> PMID: 18301394
13. Wittenberg GM, Greene J, Vertes PE, Drevets WC, Bullmore ET. Major Depressive Disorder Is Associated With Differential Expression of Innate Immune and Neutrophil-Related Gene Networks in Peripheral Blood: A Quantitative Review of Whole-Genome Transcriptional Data From Case-Control Studies. *Biological psychiatry*. 2020; 88(8):625–37. <https://doi.org/10.1016/j.biopsych.2020.05.006> PMID: 32653108
14. Cattaneo A, Ferrari C, Turner L, Mariani N, Enache D, Hastings C, et al. Whole-blood expression of inflammasome- and glucocorticoid-related mRNAs correctly separates treatment-resistant depressed patients from drug-free and responsive patients in the BIODEP study. *Transl Psychiatry*. 2020; 10(1):232. <https://doi.org/10.1038/s41398-020-00874-7> PMID: 32699209
15. Hensman Moss DJ, Flower MD, Lo KK, Miller JR, van Ommen GB, t Hoen PA, et al. Huntington's disease blood and brain show a common gene expression pattern and share an immune signature with Alzheimer's disease. *Scientific reports*. 2017; 7:44849. <https://doi.org/10.1038/srep44849> PMID: 28322270
16. Iturria-Medina Y, Khan AF, Adewale Q, Shirazi AH, Alzheimer's Disease Neuroimaging I. Blood and brain gene expression trajectories mirror neuropathology and clinical deterioration in neurodegeneration. *Brain*. 2020; 143(2):661–73. <https://doi.org/10.1093/brain/awz400> PMID: 31989163
17. Lombardo MV, Pramparo T, Gazestani V, Warriar V, Bethlehem RAI, Carter Barnes C, et al. Large-scale associations between the leukocyte transcriptome and BOLD responses to speech differ in autism early language outcome subtypes. *Nature neuroscience*. 2018; 21(12):1680–8. <https://doi.org/10.1038/s41593-018-0281-3> PMID: 30482947
18. Yang R, Daigle BJ Jr., Muhie SY, Hammamieh R, Jett M, Petzold L, et al. Core modular blood and brain biomarkers in social defeat mouse model for post traumatic stress disorder. *BMC Syst Biol*. 2013; 7:80. <https://doi.org/10.1186/1752-0509-7-80> PMID: 23962043
19. Daskalakis NP, Cohen H, Cai G, Buxbaum JD, Yehuda R. Expression profiling associates blood and brain glucocorticoid receptor signaling with trauma-related individual differences in both sexes. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(37):13529–34. <https://doi.org/10.1073/pnas.1401660111> PMID: 25114262
20. Koob GF. Neurocircuitry of alcohol addiction: synthesis from animal models. *Handb Clin Neurol*. 2014; 125:33–54. <https://doi.org/10.1016/B978-0-444-62619-6.00003-3> PMID: 25307567
21. Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2010; 35(1):217–38.
22. de Jong S, Boks MP, Fuller TF, Strengman E, Janson E, de Kovel CG, et al. A gene co-expression network in whole blood of schizophrenia patients is independent of antipsychotic-use and enriched for brain-expressed genes. *PloS one*. 2012; 7(6):e39498. <https://doi.org/10.1371/journal.pone.0039498> PMID: 22761806
23. Mina E, van Roon-Mom W, Hettne K, van Zwet E, Goeman J, Neri C, et al. Common disease signatures from gene expression analysis in Huntington's disease human blood and brain. *Orphanet J Rare Dis*. 2016; 11(1):97. <https://doi.org/10.1186/s13023-016-0475-2> PMID: 27476530
24. van Heerden JH, Conesa A, Stein DJ, Montaner D, Russell V, Illing N. Parallel changes in gene expression in peripheral blood mononuclear cells and the brain after maternal separation in the mouse. *BMC Res Notes*. 2009; 2:195. <https://doi.org/10.1186/1756-0500-2-195> PMID: 19781058
25. Becker HC. Alcohol dependence, withdrawal, and relapse. *Alcohol Res Health*. 2008; 31(4):348–61. PMID: 23584009
26. Becker HC, Lopez MF. Increased ethanol drinking after repeated chronic ethanol exposure and withdrawal experience in C57BL/6 mice. *Alcoholism, clinical and experimental research*. 2004; 28(12):1829–38. <https://doi.org/10.1097/01.alc.0000149977.95306.3a> PMID: 15608599

27. Lopez MF, Becker HC. Effect of pattern and number of chronic ethanol exposures on subsequent voluntary ethanol intake in C57BL/6J mice. *Psychopharmacology*. 2005; 181(4):688–96. <https://doi.org/10.1007/s00213-005-0026-3> PMID: 16001125
28. Griffin WC 3rd, Lopez MF, Becker HC. Intensity and duration of chronic ethanol exposure is critical for subsequent escalation of voluntary ethanol drinking in mice. *Alcoholism, clinical and experimental research*. 2009; 33(11):1893–900. <https://doi.org/10.1111/j.1530-0277.2009.01027.x> PMID: 19673744
29. Koob GF. The Dark Side of Addiction: The Horsley Gantt to Joseph Brady Connection. *J Nerv Ment Dis*. 2017; 205(4):270–2. <https://doi.org/10.1097/NMD.0000000000000551> PMID: 27356121
30. Koob GF, Mason BJ. Existing and Future Drugs for the Treatment of the Dark Side of Addiction. *Annu Rev Pharmacol Toxicol*. 2016; 56:299–322. <https://doi.org/10.1146/annurev-pharmtox-010715-103143> PMID: 26514207
31. Bosse KE, Chiu VM, Lloyd SC, Conti AC. Neonatal alcohol exposure augments voluntary ethanol intake in the absence of potentiated anxiety-like behavior induced by chronic intermittent ethanol vapor exposure. *Alcohol*. 2018. <https://doi.org/10.1016/j.alcohol.2018.10.011> PMID: 30385201
32. Becker HC. Effects of alcohol dependence and withdrawal on stress responsiveness and alcohol consumption. *Alcohol Res*. 2012; 34(4):448–58. PMID: 23584111
33. Van Skike CE, Diaz-Granados JL, Matthews DB. Chronic intermittent ethanol exposure produces persistent anxiety in adolescent and adult rats. *Alcoholism, clinical and experimental research*. 2015; 39(2):262–71. <https://doi.org/10.1111/acer.12617> PMID: 25684048
34. Valdez GR, Roberts AJ, Chan K, Davis H, Brennan M, Zorrilla EP, et al. Increased ethanol self-administration and anxiety-like behavior during acute ethanol withdrawal and protracted abstinence: regulation by corticotropin-releasing factor. *Alcoholism, clinical and experimental research*. 2002; 26(10):1494–501. <https://doi.org/10.1097/01.ALC.0000033120.51856.F0> PMID: 12394282
35. Pleil KE, Lowery-Gionta EG, Crowley NA, Li C, Marcinkiewicz CA, Rose JH, et al. Effects of chronic ethanol exposure on neuronal function in the prefrontal cortex and extended amygdala. *Neuropharmacology*. 2015; 99:735–49. <https://doi.org/10.1016/j.neuropharm.2015.06.017> PMID: 26188147
36. Griffin WC 3rd, Haun HL, Hazelbaker CL, Ramachandra VS, Becker HC. Increased extracellular glutamate in the nucleus accumbens promotes excessive ethanol drinking in ethanol dependent mice. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2014; 39(3):707–17. <https://doi.org/10.1038/npp.2013.256> PMID: 24067300
37. Griffin WC 3rd, Lopez MF, Yanke AB, Middaugh LD, Becker HC. Repeated cycles of chronic intermittent ethanol exposure in mice increases voluntary ethanol drinking and ethanol concentrations in the nucleus accumbens. *Psychopharmacology*. 2009; 201(4):569–80. <https://doi.org/10.1007/s00213-008-1324-3> PMID: 18791704
38. Osterndorff-Kahanek EA, Becker HC, Lopez MF, Farris SP, Tiwari GR, Nunez YO, et al. Chronic ethanol exposure produces time- and brain region-dependent changes in gene coexpression networks. *PloS one*. 2015; 10(3):e0121522. <https://doi.org/10.1371/journal.pone.0121522> PMID: 25803291
39. Smith ML, Lopez MF, Archer KJ, Wolen AR, Becker HC, Miles MF. Time-Course Analysis of Brain Regional Expression Network Responses to Chronic Intermittent Ethanol and Withdrawal: Implications for Mechanisms Underlying Excessive Ethanol Consumption. *PloS one*. 2016; 11(1):e0146257. <https://doi.org/10.1371/journal.pone.0146257> PMID: 26730594
40. Ferguson LB, Most D, Blednov YA, Harris RA. PPAR agonists regulate brain gene expression: Relationship to their effects on ethanol consumption. *Neuropharmacology*. 2014. <https://doi.org/10.1016/j.neuropharm.2014.06.024> PMID: 25036611
41. Saarela M, Jauhiainen S. Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*. 2021; 3(2):272.
42. Litten RZ, Bradley AM, Moss HB. Alcohol biomarkers in applied settings: recent advances and future research opportunities. *Alcoholism, clinical and experimental research*. 2010; 34(6):955–67. <https://doi.org/10.1111/j.1530-0277.2010.01170.x> PMID: 20374219
43. Tebani A, Gummesson A, Zhong W, Koistinen IS, Lakshminanth T, Olsson LM, et al. Integration of molecular profiles in a longitudinal wellness profiling cohort. *Nature communications*. 2020; 11(1):4487. <https://doi.org/10.1038/s41467-020-18148-7> PMID: 32900998
44. Cai C, Langfelder P, Fuller TF, Oldham MC, Luo R, van den Berg LH, et al. Is human blood a good surrogate for brain tissue in transcriptional studies? *BMC Genomics*. 2010; 11:589. <https://doi.org/10.1186/1471-2164-11-589> PMID: 20961428
45. Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet*. 2006; 141B(3):261–8. <https://doi.org/10.1002/ajmg.b.30272> PMID: 16526044

46. Tylee DS, Kawaguchi DM, Glatt SJ. On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *Am J Med Genet B Neuropsychiatr Genet*. 2013; 162B(7):595–603. <https://doi.org/10.1002/ajmg.b.32150> PMID: 24132893
47. Davies MN, Lawn S, Whatley S, Fernandes C, Williams RW, Schalkwyk LC. To What Extent is Blood a Reasonable Surrogate for Brain in Gene Expression Studies: Estimation from Mouse Hippocampus and Spleen. *Frontiers in neuroscience*. 2009; 3:54. <https://doi.org/10.3389/neuro.15.002.2009> PMID: 20582281
48. Crews FT, Lawrimore CJ, Walter TJ, Coleman LG, Jr. The role of neuroimmune signaling in alcoholism. *Neuropharmacology*. 2017.
49. Crews FT, Sarkar DK, Qin L, Zou J, Boyadjieva N, Vetreno RP. Neuroimmune Function and the Consequences of Alcohol Exposure. *Alcohol Res*. 2015; 37(2):331–41, 44–51. PMID: 26695754
50. Robinson G, Most D, Ferguson LB, Mayfield J, Harris RA, Blednov YA. Neuroimmune pathways in alcohol consumption: evidence from behavioral and genetic studies in rodents and humans. *Int Rev Neurobiol*. 2014; 118:13–39. <https://doi.org/10.1016/B978-0-12-801284-0.00002-6> PMID: 25175860
51. Mayfield J, Ferguson L, Harris RA. Neuroimmune signaling: a key component of alcohol abuse. *Current opinion in neurobiology*. 2013; 23(4):513–20. <https://doi.org/10.1016/j.conb.2013.01.024> PMID: 23434064
52. Erickson EK, Grantham EK, Warden AS, Harris RA. Neuroimmune signaling in alcohol use disorder. *Pharmacology, biochemistry, and behavior*. 2019; 177:34–60. <https://doi.org/10.1016/j.pbb.2018.12.007> PMID: 30590091
53. Leclercq S, de Timary P, Delzenne NM, Starkel P. The link between inflammation, bugs, the intestine and the brain in alcohol dependence. *Transl Psychiatry*. 2017; 7(2):e1048. <https://doi.org/10.1038/tp.2017.15> PMID: 28244981
54. Most D, Ferguson L, Harris RA. Molecular basis of alcoholism. *Handb Clin Neurol*. 2014; 125:89–111. <https://doi.org/10.1016/B978-0-444-62619-6.00006-9> PMID: 25307570
55. Brousse G, Arnaud B, Vorspan F, Richard D, Dissard A, Dubois M, et al. Alteration of glutamate/GABA balance during acute alcohol withdrawal in emergency department: a prospective analysis. *Alcohol and alcoholism*. 2012; 47(5):501–8. <https://doi.org/10.1093/alcalc/ags078> PMID: 22791370
56. Aliyev NA, Aliyev ZN, Aliguliyev AR. Amino acid neurotransmitters in alcohol withdrawal. *Alcohol and alcoholism*. 1994; 29(6):643–7. PMID: 7695778
57. Bhandage AK, Jin Z, Korol SV, Shen Q, Pei Y, Deng Q, et al. GABA Regulates Release of Inflammatory Cytokines From Peripheral Blood Mononuclear Cells and CD4(+) T Cells and Is Immunosuppressive in Type 1 Diabetes. *EBioMedicine*. 2018; 30:283–94. <https://doi.org/10.1016/j.ebiom.2018.03.019> PMID: 29627388
58. Alfredsson G, Wiesel FA, Tylec A. Relationships between glutamate and monoamine metabolites in cerebrospinal fluid and serum in healthy volunteers. *Biological psychiatry*. 1988; 23(7):689–97. [https://doi.org/10.1016/0006-3223\(88\)90052-2](https://doi.org/10.1016/0006-3223(88)90052-2) PMID: 2453224
59. Al Gawwam G, Sharquie IK. Serum Glutamate Is a Predictor for the Diagnosis of Multiple Sclerosis. *ScientificWorldJournal*. 2017; 2017:9320802. <https://doi.org/10.1155/2017/9320802> PMID: 28676865
60. Madeira C, Alheira FV, Calcia MA, Silva TCS, Tannos FM, Vargas-Lopes C, et al. Blood Levels of Glutamate and Glutamine in Recent Onset and Chronic Schizophrenia. *Front Psychiatry*. 2018; 9:713. <https://doi.org/10.3389/fpsy.2018.00713> PMID: 30618883
61. Shimmura C, Suda S, Tsuchiya KJ, Hashimoto K, Ohno K, Matsuzaki H, et al. Alteration of plasma glutamate and glutamine levels in children with high-functioning autism. *PloS one*. 2011; 6(10):e25340. <https://doi.org/10.1371/journal.pone.0025340> PMID: 21998651
62. Guisasola MC. [Role of heat shock proteins in the cardioprotection of regular moderate alcohol consumption]. *Medicina clinica*. 2016; 146(7):292–300. <https://doi.org/10.1016/j.medcli.2015.12.011> PMID: 26902796
63. Muralidharan S, Lim A, Catalano D, Mandrekar P. Human Binge Alcohol Intake Inhibits TLR4-MyD88 and TLR4-TRIF Responses but Not the TLR3-TRIF Pathway: HspA1A and PP1 Play Selective Regulatory Roles. *Journal of immunology (Baltimore, Md: 1950)*. 2018; 200(7):2291–303.
64. McClintick JN, McBride WJ, Bell RL, Ding ZM, Liu Y, Xuei X, et al. Gene expression changes in the ventral hippocampus and medial prefrontal cortex of adolescent alcohol-preferring (P) rats following binge-like alcohol drinking. *Alcohol*. 2018; 68:37–47. <https://doi.org/10.1016/j.alcohol.2017.09.002> PMID: 29448234
65. McClintick JN, McBride WJ, Bell RL, Ding ZM, Liu Y, Xuei X, et al. Gene Expression Changes in Glutamate and GABA-A Receptors, Neuropeptides, Ion Channels, and Cholesterol Synthesis in the Periaqueductal Gray Following Binge-Like Alcohol Drinking by Adolescent Alcohol-Preferring (P) Rats.

- Alcoholism, clinical and experimental research. 2016; 40(5):955–68. <https://doi.org/10.1111/acer.13056> PMID: 27061086
66. Farris SP, Tiwari GR, Ponomareva O, Lopez MF, Mayfield RD, Becker HC. Transcriptome Analysis of Alcohol Drinking in Non-Dependent and Dependent Mice Following Repeated Cycles of Forced Swim Stress Exposure. *Brain Sci.* 2020; 10(5). <https://doi.org/10.3390/brainsci10050275> PMID: 32370184
 67. Pignataro L, Varodayan FP, Tannenholz LE, Protiva P, Harrison NL. Brief alcohol exposure alters transcription in astrocytes via the heat shock pathway. *Brain and behavior.* 2013; 3(2):114–33. <https://doi.org/10.1002/brb3.125> PMID: 23533150
 68. Erickson EK, Blednov YA, Harris RA, Mayfield RD. Glial gene networks associated with alcohol dependence. *Scientific reports.* 2019; 9(1):10949. <https://doi.org/10.1038/s41598-019-47454-4> PMID: 31358844
 69. Affò S, Dominguez M, Lozano JJ, Sancho-Bru P, Rodrigo-Torres D, Morales-Ibanez O, et al. Transcriptome analysis identifies TNF superfamily receptors as potential therapeutic targets in alcoholic hepatitis. *Gut.* 2013; 62(3):452–60. <https://doi.org/10.1136/gutjnl-2011-301146> PMID: 22637703
 70. Hashimoto-Torii K, Kawasawa YI, Kuhn A, Rakic P. Combined transcriptome analysis of fetal human and mouse cerebral cortex exposed to alcohol. *Proceedings of the National Academy of Sciences of the United States of America.* 2011; 108(10):4212–7. <https://doi.org/10.1073/pnas.1100903108> PMID: 21368140
 71. Metten P, Iancu OD, Spence SE, Walter NA, Oberbeck D, Harrington CA, et al. Dual-trait selection for ethanol consumption and withdrawal: genetic and transcriptional network effects. *Alcoholism, clinical and experimental research.* 2014; 38(12):2915–24. <https://doi.org/10.1111/acer.12574> PMID: 25581648
 72. Tarantino LM, McClearn GE, Rodriguez LA, Plomin R. Confirmation of quantitative trait loci for alcohol preference in mice. *Alcoholism, clinical and experimental research.* 1998; 22(5):1099–105. PMID: 9726281
 73. Belknap JK, Atkins AL. The replicability of QTLs for murine alcohol preference drinking behavior across eight independent studies. *Mammalian genome: official journal of the International Mammalian Genome Society.* 2001; 12(12):893–9. <https://doi.org/10.1007/s00335-001-2074-2> PMID: 11707775
 74. Walter NA, Denmark DL, Kozell LB, Buck KJ. A Systems Approach Implicates a Brain Mitochondrial Oxidative Homeostasis Co-expression Network in Genetic Vulnerability to Alcohol Withdrawal. *Frontiers in genetics.* 2016; 7:218. <https://doi.org/10.3389/fgene.2016.00218> PMID: 28096806
 75. Bell RL, Kimpel MW, McClintick JN, Strother WN, Carr LG, Liang T, et al. Gene expression changes in the nucleus accumbens of alcohol-preferring rats following chronic ethanol consumption. *Pharmacology, biochemistry, and behavior.* 2009; 94(1):131–47. <https://doi.org/10.1016/j.pbb.2009.07.019> PMID: 19666046
 76. Romero AM, Renau-Piqueras J, Marin MP, Esteban-Pretel G. Chronic alcohol exposure affects the cell components involved in membrane traffic in neuronal dendrites. *Neurotoxicity research.* 2015; 27(1):43–54. <https://doi.org/10.1007/s12640-014-9484-x> PMID: 25022897
 77. Van Booven D, Mengying L, Sunil Rao J, Blokhin IO, Dayne Mayfield R, Barbier E, et al. Alcohol use disorder causes global changes in splicing in the human brain. *Transl Psychiatry.* 2021; 11(1):2. <https://doi.org/10.1038/s41398-020-01163-z> PMID: 33414398
 78. Ponomarev I, Wang S, Zhang L, Harris RA, Mayfield RD. Gene coexpression networks in human brain identify epigenetic modifications in alcohol dependence. *The Journal of neuroscience: the official journal of the Society for Neuroscience.* 2012; 32(5):1884–97. <https://doi.org/10.1523/JNEUROSCI.3136-11.2012> PMID: 22302827
 79. Mulligan MK, Ponomarev I, Hitzemann RJ, Belknap JK, Tabakoff B, Harris RA, et al. Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. *Proceedings of the National Academy of Sciences of the United States of America.* 2006; 103(16):6368–73. <https://doi.org/10.1073/pnas.0510188103> PMID: 16618939
 80. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet.* 2019; 51(5):793–803. <https://doi.org/10.1038/s41588-019-0397-8> PMID: 31043756
 81. Xu B, Lv W, Li X, Zhang L, Lin J. Prognostic genes of hepatocellular carcinoma based on gene coexpression network analysis. *Journal of cellular biochemistry.* 2019. <https://doi.org/10.1002/jcb.28441> PMID: 30775801
 82. Zamudio PA, Gioia DA, Lopez M, Homanics GE, Woodward JJ. The escalation in ethanol consumption following chronic intermittent ethanol exposure is blunted in mice expressing ethanol-resistant GluN1 or GluN2A NMDA receptor subunits. *Psychopharmacology.* 2020. <https://doi.org/10.1007/s00213-020-05680-z> PMID: 33052417

83. Giacometti LL, Chandran K, Figueroa LA, Barker JM. Astrocyte modulation of extinction impairments in ethanol-dependent female mice. *Neuropharmacology*. 2020; 179:108272. <https://doi.org/10.1016/j.neuropharm.2020.108272> PMID: 32801026
84. Morales M, McGinnis MM, McCool BA. Chronic ethanol exposure increases voluntary home cage intake in adult male, but not female, Long-Evans rats. *Pharmacology, biochemistry, and behavior*. 2015; 139(Pt A):67–76. <https://doi.org/10.1016/j.pbb.2015.10.016> PMID: 26515190
85. Lopez MF, Miles MF, Williams RW, Becker HC. Variable effects of chronic intermittent ethanol exposure on ethanol drinking in a genetically diverse mouse cohort. *Alcohol*. 2017; 58:73–82. <https://doi.org/10.1016/j.alcohol.2016.09.003> PMID: 27793543
86. Jury NJ, Radke AK, Pati D, Kocharian A, Mishina M, Kash TL, et al. NMDA receptor GluN2A subunit deletion protects against dependence-like ethanol drinking. *Behavioural brain research*. 2018; 353:124–8. <https://doi.org/10.1016/j.bbr.2018.06.029> PMID: 29953905
87. Jury NJ, DiBerto JF, Kash TL, Holmes A. Sex differences in the behavioral sequelae of chronic ethanol exposure. *Alcohol*. 2017; 58:53–60. <https://doi.org/10.1016/j.alcohol.2016.07.007> PMID: 27624846
88. Finn DA, Hashimoto JG, Cozzoli DK, Helms ML, Nipper MA, Kaufman MN, et al. Binge Ethanol Drinking Produces Sexually Divergent and Distinct Changes in Nucleus Accumbens Signaling Cascades and Pathways in Adult C57BL/6J Mice. *Frontiers in genetics*. 2018; 9:325. <https://doi.org/10.3389/fgene.2018.00325> PMID: 30250478
89. Wolstenholme JT, Mahmood T, Harris GM, Abbas S, Miles MF. Intermittent Ethanol during Adolescence Leads to Lasting Behavioral Changes in Adulthood and Alters Gene Expression and Histone Methylation in the PFC. *Front Mol Neurosci*. 2017; 10:307. <https://doi.org/10.3389/fnmol.2017.00307> PMID: 29018328
90. van der Vaart AD, Wolstenholme JT, Smith ML, Harris GM, Lopez MF, Wolen AR, et al. The allostatic impact of chronic ethanol on gene expression: A genetic analysis of chronic intermittent ethanol treatment in the BXD cohort. *Alcohol*. 2017; 58:93–106. <https://doi.org/10.1016/j.alcohol.2016.07.010> PMID: 27838001
91. Di Pietro V O'Halloran P, Watson CN, Begum G, Acharjee A, Yakoub KM, et al. Unique diagnostic signatures of concussion in the saliva of male athletes: the Study of Concussion in Rugby Union through MicroRNAs (SCRUM). *Br J Sports Med*. 2021.
92. Lugli G, Cohen AM, Bennett DA, Shah RC, Fields CJ, Hernandez AG, et al. Plasma Exosomal miRNAs in Persons with and without Alzheimer Disease: Altered Expression and Prospects for Biomarkers. *PloS one*. 2015; 10(10):e0139233. <https://doi.org/10.1371/journal.pone.0139233> PMID: 26426747
93. Agliardi C, Guerini FR, Zanzottera M, Bianchi A, Nemni R, Clerici M. SNAP-25 in Serum Is Carried by Exosomes of Neuronal Origin and Is a Potential Biomarker of Alzheimer's Disease. *Molecular neurobiology*. 2019; 56(8):5792–8. <https://doi.org/10.1007/s12035-019-1501-x> PMID: 30680692
94. Gazestani VH, Pramparo T, Nalabolu S, Kellman BP, Murray S, Lopez L, et al. A perturbed gene network containing PI3K-AKT, RAS-ERK and WNT-beta-catenin pathways in leukocytes is linked to ASD genetics and symptom severity. *Nature neuroscience*. 2019; 22(10):1624–34. <https://doi.org/10.1038/s41593-019-0489-x> PMID: 31551593
95. Hicks SD, Lewis L, Ritchie J, Burke P, Abdul-Malak Y, Adackapara N, et al. Evaluation of cell proliferation, apoptosis, and DNA-repair genes as potential biomarkers for ethanol-induced CNS alterations. *BMC neuroscience*. 2012; 13:128. <https://doi.org/10.1186/1471-2202-13-128> PMID: 23095216
96. Preische O, Schultz SA, Apel A, Kuhle J, Kaeser SA, Barro C, et al. Serum neurofilament dynamics predicts neurodegeneration and clinical progression in presymptomatic Alzheimer's disease. *Nature medicine*. 2019; 25(2):277–83. <https://doi.org/10.1038/s41591-018-0304-3> PMID: 30664784
97. Ignacio C, Hicks SD, Burke P, Lewis L, Szombathyne-Meszáros Z, Middleton FA. Alterations in serum microRNA in humans with alcohol use disorders impact cell proliferation and cell death pathways and predict structural and functional changes in brain. *BMC neuroscience*. 2015; 16:55. <https://doi.org/10.1186/s12868-015-0195-x> PMID: 26341662
98. Skog J, Wurdinger T, van Rijn S, Meijer DH, Gainche L, Sena-Esteves M, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat Cell Biol*. 2008; 10(12):1470–6. <https://doi.org/10.1038/ncb1800> PMID: 19011622
99. Balakathiresan NS, Chandran R, Bhomia M, Jia M, Li H, Maheshwari RK. Serum and amygdala microRNA signatures of posttraumatic stress: fear correlation and biomarker potential. *J Psychiatr Res*. 2014; 57:65–73. <https://doi.org/10.1016/j.jpsychires.2014.05.020> PMID: 24998397
100. Allen L, Dwivedi Y. MicroRNA mediators of early life stress vulnerability to depression and suicidal behavior. *Mol Psychiatry*. 2020; 25(2):308–20. <https://doi.org/10.1038/s41380-019-0597-8> PMID: 31740756

101. Nassiri F, Chakravarthy A, Feng S, Shen SY, Nejad R, Zuccato JA, et al. Detection and discrimination of intracranial tumors using plasma cell-free DNA methylomes. *Nature medicine*. 2020; 26(7):1044–7. <https://doi.org/10.1038/s41591-020-0932-2> PMID: 32572265
102. Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MP, van Eijk K, et al. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol*. 2012; 13(10):R97. <https://doi.org/10.1186/gb-2012-13-10-r97> PMID: 23034122
103. Temko JE, Bouhlal S, Farokhnia M, Lee MR, Cryan JF, Leggio L. The Microbiota, the Gut and the Brain in Eating and Alcohol Use Disorders: A 'Menage a Trois'? *Alcohol and alcoholism*. 2017; 52(4):403–13. <https://doi.org/10.1093/alcalc/agx024> PMID: 28482009
104. Leclercq S, Le Roy T, Furguiele S, Coste V, Bindels LB, Leyrolle Q, et al. Gut Microbiota-Induced Changes in beta-Hydroxybutyrate Metabolism Are Linked to Altered Sociability and Depression in Alcohol Use Disorder. *Cell Rep*. 2020; 33(2):108238. <https://doi.org/10.1016/j.celrep.2020.108238> PMID: 33053357
105. Setoyama D, Kato TA, Hashimoto R, Kunugi H, Hattori K, Hayakawa K, et al. Plasma Metabolites Predict Severity of Depression and Suicidal Ideation in Psychiatric Patients-A Multicenter Pilot Analysis. *PloS one*. 2016; 11(12):e0165267. <https://doi.org/10.1371/journal.pone.0165267> PMID: 27984586
106. Pu J, Liu Y, Zhang H, Tian L, Gui S, Yu Y, et al. An integrated meta-analysis of peripheral blood metabolites and biological functions in major depressive disorder. *Mol Psychiatry*. 2020. <https://doi.org/10.1038/s41380-020-0645-4> PMID: 31959849
107. He Y, Yu Z, Giegling I, Xie L, Hartmann AM, Prehn C, et al. Schizophrenia shows a unique metabolomics signature in plasma. *Transl Psychiatry*. 2012; 2:e149. <https://doi.org/10.1038/tp.2012.76> PMID: 22892715
108. Hashimoto K. Metabolomics of Major Depressive Disorder and Bipolar Disorder: Overview and Future Perspective. *Adv Clin Chem*. 2018; 84:81–99. <https://doi.org/10.1016/bs.acc.2017.12.005> PMID: 29478517
109. Griffin WC, Ramachandra VS, Knackstedt LA, Becker HC. Repeated cycles of chronic intermittent ethanol exposure increases basal glutamate in the nucleus accumbens of mice without affecting glutamate transport. *Front Pharmacol*. 2015; 6:27. <https://doi.org/10.3389/fphar.2015.00027> PMID: 25755641
110. Lopez MF, Griffin WC 3rd, Melendez RI, Becker HC. Repeated cycles of chronic intermittent ethanol exposure leads to the development of tolerance to aversive effects of ethanol in C57BL/6J mice. *Alcoholism, clinical and experimental research*. 2012; 36(7):1180–7. <https://doi.org/10.1111/j.1530-0277.2011.01717.x> PMID: 22309159
111. Melendez RI, McGinty JF, Kalivas PW, Becker HC. Brain region-specific gene expression changes after chronic intermittent ethanol exposure and early withdrawal in C57BL/6J mice. *Addiction biology*. 2012; 17(2):351–64. <https://doi.org/10.1111/j.1369-1600.2011.00357.x> PMID: 21812870
112. Repunte-Canonigo V, Chen J, Lefebvre C, Kawamura T, Kreifeldt M, Basson O, et al. MeCP2 regulates ethanol sensitivity and intake. *Addiction biology*. 2014; 19(5):791–9. <https://doi.org/10.1111/adb.12047> PMID: 23448145
113. Paxinos G, Franklin KBJ. *Paxinos and Franklin's the mouse brain in stereotaxic coordinates*. 4th ed. ed. Amsterdam,: Elsevier Academic Press; 2013.
114. Lohman BK, Weber JN, Bolnick DI. Evaluation of TagSeq, a reliable low-cost alternative for RNAseq. *Mol Ecol Resour*. 2016; 16(6):1315–21. <https://doi.org/10.1111/1755-0998.12529> PMID: 27037501
115. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9(4):357–9. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
116. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
117. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31(2):166–9. <https://doi.org/10.1093/bioinformatics/btu638> PMID: 25260700
118. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*. 2015; 43(7):e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
119. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*. 2008; 9:559. <https://doi.org/10.1186/1471-2105-9-559> PMID: 19114008
120. Ferguson LB, Zhang L, Kircher D, Wang S, Mayfield RD, Crabbe JC, et al. Dissecting Brain Networks Underlying Alcohol Binge Drinking Using a Systems Genomics Approach. *Molecular neurobiology*. 2018. <https://doi.org/10.1007/s12035-018-1252-0> PMID: 30062672

121. Ferguson LB, Zhang L, Wang S, Bridges C, Harris RA, Ponomarev I. Peroxisome Proliferator Activated Receptor Agonists Modulate Transposable Element Expression in Brain and Liver. *Front Mol Neurosci*. 2018; 11:331. <https://doi.org/10.3389/fnmol.2018.00331> PMID: 30283300
122. McKenzie AT, Wang M, Hauberg ME, Fullard JF, Kozlenkov A, Keenan A, et al. Brain Cell Type Specific Gene Expression and Co-expression Network Architectures. *Scientific reports*. 2018; 8(1):8868. <https://doi.org/10.1038/s41598-018-27293-5> PMID: 29892006
123. Nirmal AJ, Regan T, Shih BB, Hume DA, Sims AH, Freeman TC. Immune Cell Gene Signatures for Profiling the Microenvironment of Solid Tumors. *Cancer immunology research*. 2018; 6(11):1388–400. <https://doi.org/10.1158/2326-6066.CIR-18-0342> PMID: 30266715
124. Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, et al. Functional organization of the transcriptome in human brain. *Nature neuroscience*. 2008; 11(11):1271–82. <https://doi.org/10.1038/nn.2207> PMID: 18849986
125. Miller JA, Horvath S, Geschwind DH. Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(28):12698–703. <https://doi.org/10.1073/pnas.0914257107> PMID: 20616000
126. Diedenhofen B, Musch J. cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS one*. 2015; 10(3):e0121945. <https://doi.org/10.1371/journal.pone.0121945> PMID: 25835001
127. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*. 1979; 6(2):65–70.
128. Goksuluk D, Zararsiz G, Korkmaz S, Eldem V, Zararsiz GE, Ozcetin E, et al. MLSeq: Machine learning interface for RNA-sequencing data. *Comput Methods Programs Biomed*. 2019; 175:223–31. <https://doi.org/10.1016/j.cmpb.2019.04.007> PMID: 31104710
129. Crow M, Lim N, Ballouz S, Pavlidis P, Gillis J. Predictability of human differential gene expression. *Proceedings of the National Academy of Sciences of the United States of America*. 2019. <https://doi.org/10.1073/pnas.1802973116> PMID: 30846554