Towards Molecular-Based Diagnostics for Alcohol Use Disorders: Blood and Brain Signatures of Chronic Intermittent Ethanol Consumption

Laura B. Ferguson[1,2], Amanda J. Roberts[4], Avi Ma'ayan[5], R. Dayne Mayfield[1,3], Robert O. Messing[1,2,3]

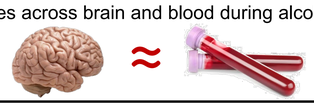
[1]Waggoner Center for Alcohol and Addiction Research, University of Texas at Austin, TX.
[2]Department of Neurology, Dell Medical School, University of Texas at Austin, TX.
[3]Department of Neuroscience, University of Texas at Austin, TX.
[4]Animal Models Core Facility, The Scripps Research Institute, La Jolla, CA.
[5]Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, NY.

Background

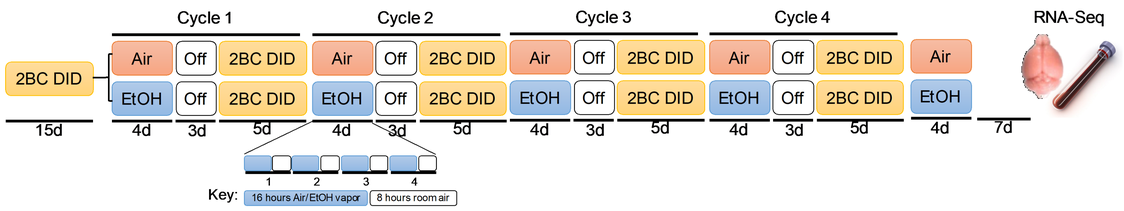
Alcohol Use Disorder (AUD) is a chronic, relapsing illness requiring a subjective diagnosis based on patient symptoms and with few effective treatment options. Early detection of AUD permits earlier intervention which could mitigate the negative impacts of AUD on the patient. AUD, like other psychiatric illnesses, is primarily considered a brain disease that arises from aberrant molecular, cellular, and neurophysiological properties in brain. Consistent with this, the application of partial least squares discriminative analysis (PLSDA) to gene expression patterns from postmortem prefrontal cortex tissue reliably discriminates AUD from non-AUD individuals (PMID: 16292326). However, it is not possible to obtain brain specimens from patients, limiting the ability of computational approaches to aid in AUD diagnostics. The brain works in harmony with the rest of the body to accomplish its functions, and alcohol use affects multiple other tissues and systems. Therefore, it might be possible to identify a surrogate for brain tissue that could be used as a more accessible transcriptome for diagnosis and longitudinal follow up. Whole blood makes contact with every organ in the body (including brain), and the majority of whole blood RNA is derived from white blood cells whose function is to generate immune responses which have been shown to be important in AUD pathophysiology (e.g., PMID: 30590091).

Hypothesis: Blood and brain transcriptomes share enough common features that the blood transcriptome can be used to diagnose AUD.

Objective: Detect correlated gene expression responses across brain and blood during alcohol withdrawal using a within-subjects design.



Methods



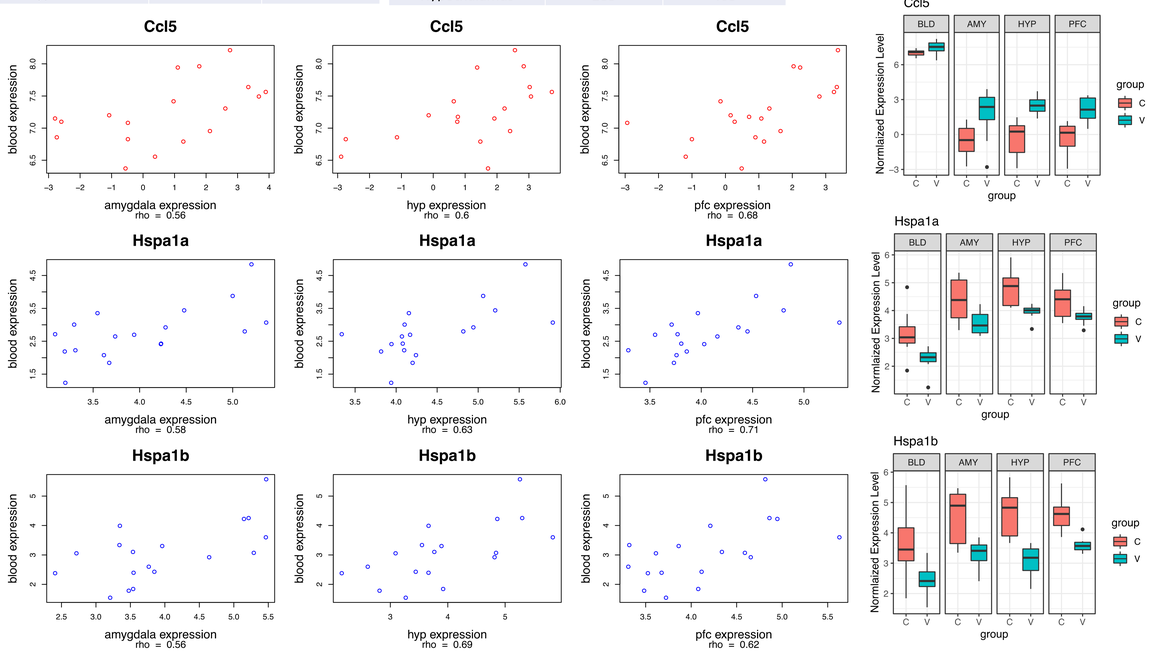
Goal	Technique
Induce alcohol dependence	Chronic intermittent ethanol (CIE) treatment with voluntary alcohol drinking tests (above) N = 40 C57Bl/6J mice; 10 per group; male and female; alcohol-dependent (ethanol) and non-dependent (air)
Assess protein-coding gene expression	RNA seq (TagSeq) of globin-depleted whole blood & Brain (prefrontal cortex (PFC), amygdala (AMY), and hypothalamus (HYP)) mRNA
Discover genes with increased or decreased expression levels during alcohol withdrawal	Differential expression analysis (EtOH vs air) limma R package version 3.42.2
Parse the transcriptional response to alcohol withdrawal into groups of genes with highly correlated expression patterns (modules)	Weighted Gene Co-expression Network Analysis (one network per tissue) WGCNA R package version 1.69
Gain insight into the functional impact and cellular specificity of gene expression perturbations	Functional Enrichment Analysis: (1) Cell types (userListEnrichment function from the WGCNA package), (2) Biological pathways (Ingenuity Pathway Analysis (IPA)), (3) Upstream regulators (transcription factors, drugs, microRNAs, or other regulators that might explain the observed changes in gene expression) (IPA) Cell type datasets: brain (astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, oligodendrocyte progenitor cells) (PMID: 29892006). Immune populations (B cells, plasma cells, monocytes, macrophages, neutrophils, NK cells, T cells) (PMID: 30266715).
Compare blood and brain transcriptional responses to alcohol withdrawal	Within-subjects pairwise Spearman correlation coefficient (brain vs blood normalized gene expression levels) (corr.test function from the psych package (version 2.0.7)) Compare DEGs between blood and each brain region (userListEnrichment) Compare modules between blood and each brain region (userListEnrichment; Cytoscape version 3.2.1)
Predict alcohol dependence status from whole blood gene expression	Machine Learning Classification Algorithms: (1) Artificial Neural Network, (2) Logistic Regression, (3) Random Forest, (4) Support Vector Machine, and (5) Partial Least Squares Repeated cross validation was used to choose the optimal parameters with the trainControl function from the caret package (version 6.0-86). Final model trained using the optimal hyperparameters. Five-fold cross validation repeated 50 times was used to estimate the model's accuracy.

Result 1. Blood gene expression levels predict brain gene expression levels for hundreds of genes

Females		
	+	-
Amygdala	229	313 (mic)
PFC	367	267 (endo)
Hypothalamus	305	344

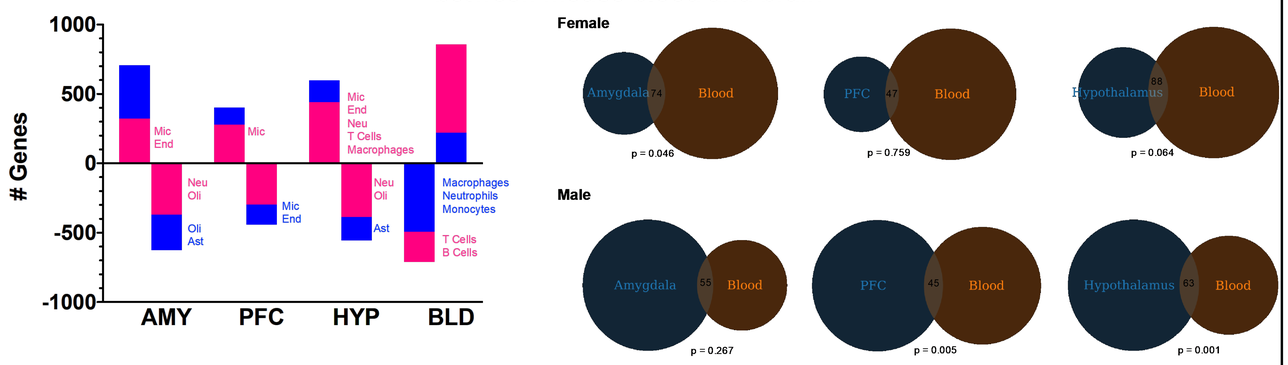
Males		
	+	-
Amygdala	364	343 (mic, T)
PFC	297	359 (endo, mic)
Hypothalamus	283	403

Number of genes positively and negatively correlated with blood gene expression in the brain area indicated by the row name (Holm Bonferroni-corrected p values < 0.05). Cell type enrichment is indicated in parentheses. endo = endothelial cells, mic = microglia, T=T cells.



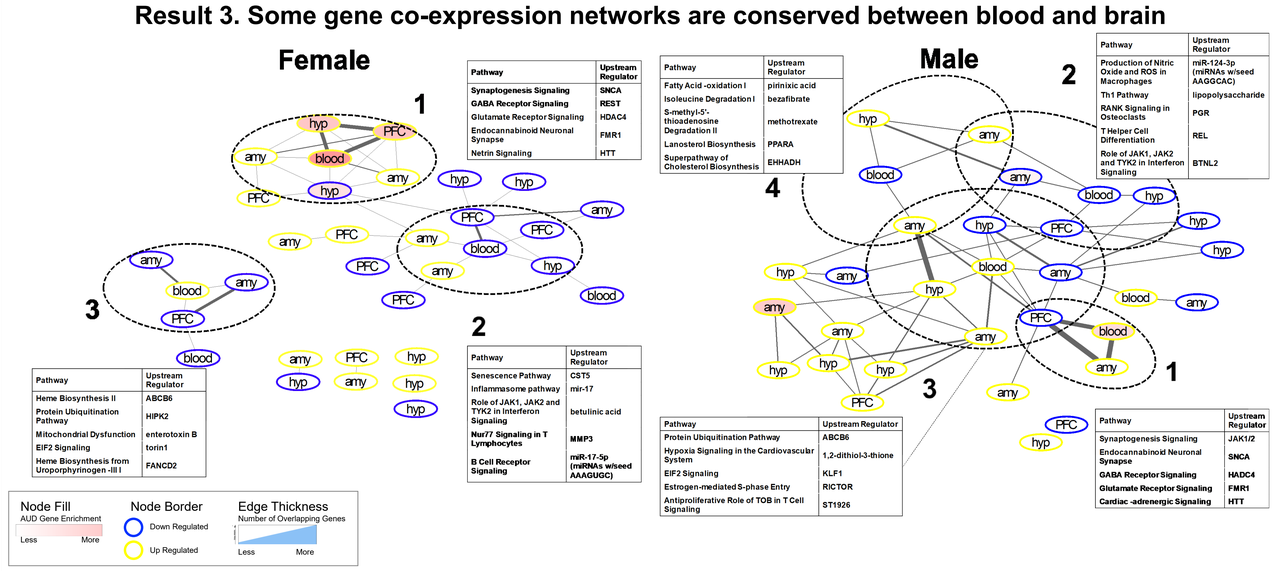
Some correlated genes were also differentially expressed between dependent (V = ethanol vapor) and non-dependent (C = control) mice in both blood and brain, which could make them ideal biomarker candidates. For females, Ccl5 was correlated and differentially expressed in blood and all brain regions. For males, Hsp1a and Hsp1b were correlated and differentially expressed in blood and all brain regions.

Result 2. A small but significant number of genes differentially expressed during withdrawal are conserved between mouse blood and brain



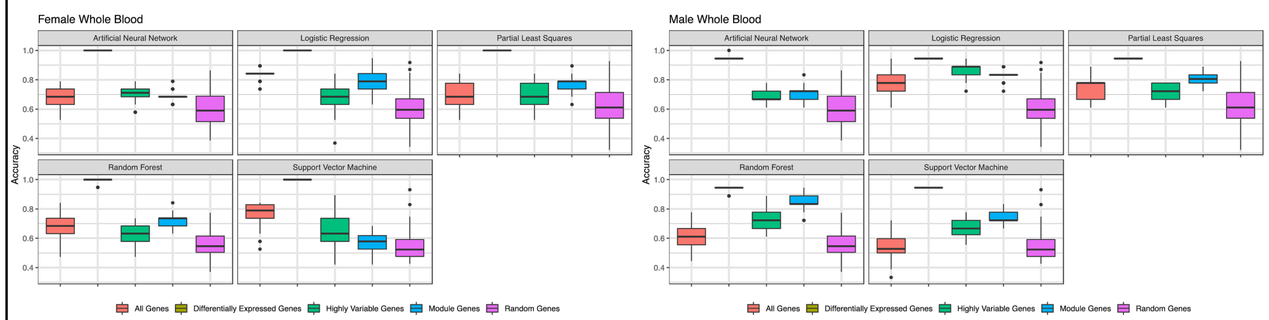
The number of up and down regulated genes are shown for males (blue) and females (pink). The number of differentially expressed genes in brain was greater in males than females, while females showed a greater number of differentially expressed genes in blood relative to males. Cell type enrichment noted next to the corresponding geneset. Venn diagrams show the number of shared DEGs between brain and blood. The overlap is small but significant for male PFC and hypothalamus and borderline significant for female amygdala and hypothalamus. Ast = astrocytes, End = endothelial cells, Mic = microglia, Oli = oligodendrocytes.

Result 3. Some gene co-expression networks are conserved between blood and brain



A network depiction of the conserved gene co-expression modules between PFC, AMY, HYP, and blood in female and male mice. Gene co-expression networks were constructed for each tissue separately. The similarity of the modules (nodes) between tissues was assessed by calculating the number of overlapping genes between each module from the separate tissue networks (represented by edge width). The dashed lines encircle the main groups of alcohol dependence-related modules that are shared between blood and brain. The top 5 enriched pathways and predicted upstream regulators of the encircled modules are shown in the adjacent tables. The preserved modules are numbered according to a corresponding match in the opposite sex.

Result 4. Whole blood gene expression signatures can distinguish alcohol-dependent and non-dependent subjects with perfect or near-perfect accuracy



Machine learning classifiers were trained using peripheral whole blood gene expression data from female and male C57Bl/6J mice to predict alcohol dependence status. Classification accuracy (percentage of correct assignments) is shown on the y-axis for the five classification techniques (indicated by the plot titles). We used 5-fold cross validation repeated 50 times to estimate classification accuracy because of the small sample size. The boxplots show the range of accuracies attained across the 50 repeats. The fill color indicates the feature set used to train the classifiers: (1) all expressed genes, (2) the genes with the highest amount of variance across samples (top 10%), (3) the top genes differentially expressed between alcohol dependent and non-dependent subjects (1 to 150), and (4) differentially expressed genes between dependent and non-dependent subjects that were also hub genes within modules associated with alcohol dependence (to incorporate co-expression network information into feature selection), and (5) 50 random genesets of 50 genes (to approximate random guessing). The top differentially expressed genes in whole blood could perfectly discriminate between alcohol dependent and non-dependent subjects for females and nearly perfectly for males.

Summary

1. Hundreds of genes in blood can predict brain levels.
2. There is a small but significant overlap between the genes differentially expressed between alcohol-dependent and non-dependent mice in peripheral blood and brain.
3. Several co-expressed gene modules are highly conserved between blood and brain related to the following broad categories:
    1. cell-cell signaling (e.g., GABA and glutamate receptor signaling, endocannabinoid signaling)
    2. immune responses (e.g., antigen presentation, communication between innate and adaptive immune systems, JAK/STAT signaling)
    3. protein processing (e.g., ubiquitination, unfolded protein responses)
4. Whole blood transcriptome can distinguish between alcohol dependent and non-dependent animals.

Conclusion

Our study suggests that **gene expression profiles from peripheral blood samples can be used to identify a biological signature of AUD**. While these results need to be validated in larger and more heterogeneous datasets, this study lays important groundwork for diagnosing AUD using blood, a much more accessible transcriptome than brain. Molecular-based diagnostics could facilitate more objective diagnoses and possibly earlier detection of AUD which would lessen the negative impacts of AUD on the patient.