RUNNING HEAD: Alcohol use disorder dual methylation study

# Dual methylation and hydroxymethylation study in blood and brain identifies *BAIAP2* as a mediator of gene expression differences associated with alcohol use disorder

Shaunna L. Clark[1], Robin F. Chan[2], Min Zhao[2], Lin Y. Xie[2], Brenda W.J.H. Penninx[3], Karolina A. Aberg[2], Edwin J.C.G. van den Oord[2]

[1]Department of Psychiatry, Texas A&M University
[2]Center for Biomarker Research and Precision Medicine, Virginia Commonwealth University
[3] Department of Psychiatry, Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands

Corresponding Author: Shaunna L. Clark, Email: slclark@tamu.edu; Address: 8447 Riverside Parkway, Bryan, TX, 77807; Phone: (979) 436-0179

Keywords: alcohol use disorder; epigenetics; methylation; hydroxymethylation; *BAIAP2*

## ABSTRACT

**Background:** Using a three-stage, multi-tissue design we sought to characterize methylation and hydroxymethylation changes in blood and brain associated with alcohol use disorder (AUD).

**Methods:** In the discovery stage, we used epigenomic deconvolution to perform cell-type specific methylome-wide association studies within subpopulations of granulocytes/T-cells/B-cells/monocytes in 1,132 blood samples. Blood findings were examined for overlap with AUD-related methylation and hydroxymethylation in 50 human post-mortem brain samples in the brain overlap stage. Follow-up analyses investigated if overlapping findings mediated AUD-associated transcription changes in the same brain samples in the final stage.

**Results:** Cell-type specific analyses in blood identified methylome-wide significant associations in monocytes and T-cells. One of the top genic findings is located in *PLA2G4A*, a gene required for monocyte chemotaxis. Alcohol inhibits monocyte chemotaxis, thereby contributing to alcohol-induced inflammation. The monocyte findings were significantly enriched for AUD-related methylation and hydroxymethylation in brain. Hydroxymethylation in *BAIAP2* was found to mediate AUD-associated transcription in the same brain samples. *BAIAP2* regulates dendritic spine density and is linked to cognitive deficits that are clinical features of advanced AUD.

**Conclusions:** As part of the most comprehensive methylation study of AUD to date, this work involved the first cell-type specific methylation study of AUD conducted in blood, identifying methylation sites that are involved in alcohol-induced inflammation. In this first study to consider the role of hydroxymethylation in AUD, we found evidence for a novel mechanism for AUD brain-related impairments. Our results suggest promising new avenues for AUD research.

## INTRODUCTION

Alcohol use disorder (AUD) is a devastating illness characterized by excessive, uncontrolled consumption of alcohol despite its negative consequences on the drinker's health and well-being. DNA methylation (DNAm) studies of AUD offer unique opportunities to better understand and potentially treat AUD. For example, phenomena such as sensitization and tolerance suggest that repeated alcohol use creates a "biological memory" affecting future responses to the drug (1). Because drug-induced epigenetic changes can persist over time (2-9) and have lasting effects (7-9), DNAm could be a potential mechanism behind such phenomena.

DNAm studies, including previous studies of AUD, are typically conducted using bulk tissue (e.g., whole blood or brain tissue) that is composed of multiple cell-types, each with a potentially different DNAm profile. In studies using bulk tissue, estimates of cell-type proportions are typically included as covariates to protect against false positive associations (10). By focusing only on bulk tissue, however, many potentially important associations may be missed (11). For example, when effects involve a single cell type or have opposite directions in different cell-types, they may be diluted or potentially nullify each other in bulk tissue. Further, as the most common cell-types will drive the results in bulk tissue, effects from low abundance cells may be missed altogether. Additionally, knowing which specific cells harbor effects is key for the biological interpretation and crucial for designing follow-up experiments.

One way to obtain cell-type specific methylation results are to sort samples into the respective cell populations of interest and then generate methylation data in each of the sorted cell populations. It is, however, infeasible to sort and assay methylation in each cell-type for every sample in large-scale studies. An alternative is to use a deconvolution approach which uses reference methylomes from sorted cells to test cell-type specific case-control differences. This approach was initially developed for gene expression studies (12) but has been validated for methylation data (13) and successfully applied to, for example, a cell-type specific methylation study of depression (14).

While alcohol impacts many parts of the body, DNAm studies aimed at understanding the pathogenic processes of AUD are ideally performed in brain. However, brain tissue cannot be obtained from living patients. It can be collected postmortem, but obtaining the sample sizes needed for adequate statistical power is challenging. Studies have found correlations of 0.6–0.7 between methylation profiles in human blood and brain (15). Multiple factors may contribute to this overlap. Peripheral tissues such as blood may reveal methylation marks predating or resulting from the epigenetic reprogramming events affecting the germ line and embryogenesis (16), sequence variants can affect methylation levels and will be identical across tissues (17, 18), and systemic AUD disease processes such as inflammation will affect both blood and brain. This observed overlap suggests that one approach to improve statistical power of studies of postmortem brain, is to combine results with studies in (antemortem) blood samples that are easier to collect in large sample sizes.

Our three-stage, multi-tissue design improves statistical power and helps avoid potential study specific confounders. An overview of the study design is shown in Figure 1. In the discovery stage (Figure 1, top panel) of the current study, we sought to characterize AUD-linked cell-type specific (granulocytes/T-cells/B-cells/monocytes) methylation (mCG; DNA methylation occurring at CpG sites) changes in blood samples in a large discovery sample of 1,132 subjects. In the next stage (Figure 1, middle panel), we examined if corresponding AUD-related changes can be observed in the pre-frontal cortex (PFC) of AUD cases and controls from 50 independent subjects. The brain methylome, however, is more complex than the blood methylome. For example, in addition to mCG and in contrast to blood and many other tissues, hydroxymethylation (hmCG) is common in brain. Therefore, we considered both mCG and hmCG in our brain investigations. Finally, for AUD-associated changes that overlapped across blood and brain, we followed-up findings by studying whether the observed mCG or hmCG changes mediated case-control differences in transcript abundance levels in the same brain samples (Figure 1, bottom panel). To our knowledge, this is the first alcohol methylation study to

examine cell-type specific mCG in blood, the largest study of AUD methylation conducted to date both in terms of sample size and the number of methylation sites investigated, and the first study to consider mCG, hmCG and transcript abundance from the same samples in AUD.

## METHODS AND MATERIALS

### Samples

*Blood Samples*

The mCG data was previously generated using methyl-CG binding domain sequencing (MBD-seq) (19, 20) for DNA extracted from blood samples from 1,132 individuals from the Netherlands Study of Depression and Anxiety (NESDA) (21). Lifetime AUD was diagnosed using the DSM-IV based Composite International Diagnostic Interview (CIDI version 2.1) (22) that was administered by specially trained research staff. AUD cases (n=323) included individuals that met criteria for either abuse or dependence on the DSM-IV to match the AUD definition in the brain samples (see below). The study was approved by ethical committees of all participating locations and participants provided written informed consent. Further details about the study sample, and the demographic and clinical characteristics of participants used for the present study are in Table S1.

*Post-mortem Brain Samples*

The brain samples consist of 50 brain autopsies provided by the New South Wales Brain Tissue Resource Centre, Australia. Specifically, the post-mortem tissue samples are from the PFC from 25 AUD cases and 25 unaffected controls matched to the cases on age and sex. The AUD cases consumed ≥80g alcohol per day during the majority of their adult lives and met criteria for a diagnosis of either abuse (n=10) or dependence (n=15) on the DSM-IV and AUD on DSM-V. Additionally, AUD cases did not have liver cirrhosis, Wernicke–Korsakoff's syndrome, or multi-drug abuse history. Subjects in the control group had either abstained from

alcohol completely or were social drinkers who consumed ≤20g of alcohol per day on average. Further details about the subjects used in the present study are presented in Table S2. From each tissue sample we performed simultaneous extraction of DNA and RNA using the AllPrep Universal kit (Qiagen).

### *Methylation and Hydroxymethylation*

#### *Assaying mCG*

The same MBD-seq approach, which was previously used to generate mCG data for the blood samples (21) was also used to generate mCG data for the brain samples. Specifically, we used an optimized protocol for the MBD-seq approach (19, 20) that achieves near-complete coverage of the 28 million possible mCG sites at a cost comparable to commonly used methylation arrays that assay only 2-3% of all mCG sites (23). Briefly, genomic DNA was sheared into, on average, 150 bp fragments using the Covaris™ E220 focused ultrasonicator system. We performed enrichment with MethylMiner™ (Invitrogen) to capture the mCG fraction of the genome. Please note that MethylMiner is specific for mCG (23). Thus, in contrast to bisulfite-based approaches, MethylMiner does not detect hydroxymethylation. Next, dual-indexed sequencing library for each methylation capture was prepared using the Accel-NGS® 2S Plus DNA Library Kit (Swift Biosciences) and sequenced on a NextSeq500 instrument (Illumina).

#### *Assaying hmCG*

As hydroxymethylation is very rare in blood, hmCG was only assayed in the brain samples. To assay hmCG, selective chemical labeling and enrichment of hmC (hMe-Seal) was performed using components of the Hydroxymethyl Collector kit (Active Motif). We substituted the enzyme in the kit with T4 β-glucosyltransferase from New England Biosciences (#M0357) to improve labeling performance. In our experience, using high-quality and high-activity enzyme is

critical for successful hmC enrichment. Briefly, T4 β-glucosyltransferase was used to selectively label hmC residues with 150 μM final UDP-azide-glucose. Each azo-glucosylated hmC was next biotinylated via dibenzocyclooctyl click chemistry by addition of the provided Biotin Conjugate Solution. After biotinylation, the labeled DNA was column purified, enriched with paramagnetic streptavidin beads, washed and eluted following the vendor supplied protocol, with the substitution of end-over-end rotation with agitation on an orbital shaker at 700 rpm in a 96-well 1.2 ml square well plate. Next, dual-indexed sequencing library for each methylation capture was prepared using the Accel-NGS® 2S Plus DNA Library Kit (Swift Biosciences) and sequenced on a NextSeq500 instrument (Illumina).

*mCG and hmCG Data Processing*

First, the sequence reads were aligned to the human reference genome (hg19/GRCh37) using Bowtie2 (24). Data quality control and analyses were performed in RaMWAS (25). mCG and hmCG scores were calculated by estimating the number of fragments covering each site using a non-parametric estimate of the fragment size distribution (26). These scores provide a quantitative measure of mCG/hmCG for each individual at that specific site (23). Details on data processing and quality control are provided in the supplemental material. In total, 21,868,402 mCGs were investigated in blood, 20,823,597 mCGs and 26,153,809 hmCGs were investigated in the brain tissue.

*Methylome-wide association testing (MWAS)*

MWAS testing in blood was performed with RaMWAS (25) using multiple regression analyses that included several classes of covariates (27). These classes were measured technical variables such as batch and peak (27), demographic variables for age and sex, and estimated cell-type proportions (10). We also included smoking (yes/no; current use of

cigarettes) and depressive disorder (yes/no; DSMIV-based diagnosis) status as covariates. Principal component analysis was performed on the covariate-adjusted methylation data to capture any remaining unmeasured sources of variation. We used a scree test to select three principal components to include in the final blood MWAS. To declare methylome-wide significance we applied a false discovery rate (FDR) (28) threshold of 0.1(29). Operationally, the FDR was controlled using *q*-values that are FDRs calculated using the *p*-values of the individual tests as thresholds for declaring significance (30, 31). The brain mCG and hmCG association testing was conducted separately for mCG and hmCG, and included similar covariates to the blood MWAS.

*Cell-type Specific MWAS*

We used an epigenomic deconvolution approach to perform cell-type specific MWAS for the major nucleated cell-types in blood: T-cells (CD3+), monocytes (CD14+), granulocytes (CD15+) and B-cells (CD19+). This approach applies statistical methods in combination with MBD-seq mCG profiles from a reference set of purified cells (32) to deconvolute the cell-type specific effects from data generated in bulk tissue (33, 34). An in-depth description of this approach as applied to methylation sequencing data and the generation of the reference panels has previously been described elsewhere (14). In short, cell-type proportion estimates for each sample were obtained using Houseman's method (10, 35). These estimates were then used to test the null hypothesis that methylation of a given site is not correlated with AUD status for each cell-type (12). To declare methylome-wide significance we applied a false discovery rate (FDR) threshold of 0.1.

*Cross-tissue Overlap and Colocalization Testing*

To explore if there is overlapping AUD-associated methylation at the same CpGs across blood and brain, we tested whether the blood mCG findings were enriched for brain

mCG and hmCG top findings. Specifically, testing was conducted using circular permutations that generate the empirical test-statistic distribution under the null hypothesis while preserving the correlational structure of the data (36). We defined the "top" findings as the top 0.1% and 0.5% of sites from our MWASs and corrected for testing multiple thresholds by using the same threshold in the permutations used to generate the null distrbution. In addition, we applied the same approach to test whether the overlapping sites across tissues colocalized with UCSC Genome Browser genomic feature tracks and Roadmap Epigenomics Project chromHMM 15-state chromatin model tracks (37). Further details can be found in the Supplemental Material.

## Transcript Expression

### *Assaying Transcript Abundance*

We used high-quality RNA, with average RNA integrity number (RIN) of 9.3 (sd=0.58), to assess transcript abundance with RNA-seq. In short, for each sample 800 ng of total RNA was used with 0.8 reactions of the TruSeq Stranded Total RNA Gold sample preparation kit (Illumina) to perform bead-based depletion of cytoplasmic and mitochondrial rRNA, followed by cDNA synthesis. Next, using Truseq RNA UD indexes (Illumina) amplification was performed to generate a uniquely labeled paired-end sequencing library for each sample. Following library preparation, appropriate library size was confirmed using the Agilent 2100 Bioanalyzer with the high sensitivity DNA kit (Agilent) and fluorometric quantification of the concentration of double stranded DNA was assessed via Qubit (ThermoFisher Scientific). Following quality control 8 libraries were pooled in equal molarities and sequenced with 150 cycles (i.e. 2 x 75 base pair reads) on a NextSeq500 instrument (Illumina).

Transcriptome sequence reads were aligned using HiSat2 (38). Transcript assembly and estimation of abundance levels per transcript were performed using StringTie (39). Abundance levels were then normalized to transcript per million (TPM). Further details of transcript processing and quality control are provided in the supplemental material.

*Testing Methylation Mediation*

We used the information about transcript abundance to examine if mCG/hmCG sites that showed case-control differences across tissues were related to AUD-associated transcription changes. Specifically, in the cell-types that had significant associations with AUD in blood, we tested whether the mCG\hmCG levels in brain of sites that overlapped across tissues were associated with transcript abundance (i.e., cis-meQTL) in the same brain samples. To declare a site a potential cis-meQTL, we used a p-value threshold of 0.01. Mediation analyses were then conducted with the significant cis-meQTLs to determine if the AUD-associated variation in mCG/hmCG accounted for possible case-control differences related to AUD-associated transcription changes. Mediation testing was conducted with the mediate function of the mediation R package using a quasi-Bayesian approach with 100,000 Monte Carlo draws for the approximation of the p-values for the mediation effect (40). A Bonferroni-correction was used to declare significant mediation.

## RESULTS

### Discovery: Blood MWAS

The Quantile-Quantile (QQ) plots for whole blood and individual cell-types suggested there was no test statistic inflation with lambda (ratio of the median of the observed distribution of the test statistic to the expected median) ranging from 0.97-1.06 (Figure 2). The majority of findings involved monocytes and T-cells (Tables S3-S7). No methylome-wide significant findings were observed for whole blood, granulocytes or B-cells. The MWAS for monocytes yielded 1,397 methylome-wide significant CpGs (FDR<0.1; Table S5), with the top genic findings located in, e.g., *BAZ2B*, *PLA2G4A*, and *PTPRN2*. In T-cells, 3 CpGs passed methylome-wide significance with the top two sites ($p=3.72x10^{-10}$ and $1.24x10^{-09}$) found in the long intergenic noncoding RNA *RP11-342D11.3* between *ITGB1* and *NRP1* (Table S7).

## Brain Overlap

The exploratory methylome-wide results for both mCG and hmCG showed that no individual site reached methylome-wide significance in brain (Table S8 and S9). The top mCG finding was located in *GOSR2* ($p=3.98x10^{-08}$), a member of the golgi SNAP receptor complex gene family linked to synaptic integrity(41). The top hmCG finding was located in the spastin gene (*SPAST*; $p=1.54x10^{-07}$) which has been linked to vesicle trafficking in neurons(42).

To assess whether there was overlapping AUD-related methylation across blood and brain, we tested whether the top whole blood and cell-type specific MWAS results were enriched for the top mCG and hmCG results in brain. Top results were defined as the top 0.1% and 0.5% of findings. We observed significant overlap (Table 1) in top sites between the monocyte MWAS with both mCG (2,315 sites, p = 0.009) and hmCG (2,366 sites, p = 0.035) in brain and the B-cell MWAS with both mCG (144 sites, p =0.007) and hmCG (38 sites, p = 0.010).

## Characterization

### *Colocalization of Overlap*

To characterize the overlapping sites across tissues, we tested if they were enriched for specific genomic features and chromatin states (Tables S10 & S11). As B-cells did not have any significant results in blood, the overlapping results were not characterized. The overlap between monocytes and both mCG and hmCG were enriched at CpG islands, exons, genes, gene promotors and DNase hypersensitivity regions. The overlap between monocytes and both mCG and hmCG were enriched for the following chromatin states: strong transcription, while the overlap with monocytes and hmCG was additionally enriched for weak transcription.

### *Mediation testing of cis-MeQTLs with brain expression data*

Genic mCG and hmCG findings that were with cell-types that had significant associations in the blood MWAS and overlapped between blood and brain tissue were tested for association with transcript levels (i.e., cis-meQTL testing). Of the overlapping sites, two were associated with transcript level expression for monocyte/mCG and six for monocyte/hmCG (Table 2).

For the significant cis-meQTLs, we performed mediation analyses to test whether mCG/hmCG mediated the relationship between AUD and transcript expression. The results (Table 2) revealed one significant mediation effect after Bonferroni correction ($\alpha$=0.05/8=0.00625).

The significant mediation effect was detected between hmCG in *BAIAP2,* on chromosome 17 at location 79,062,899 and ENST00000572073, a transcript of *BAIAP2*. As Figure 3 shows, the regression coefficients between AUD and expression of transcript ENST00000572073 and hmCG at location 79,062,899 on chromosome 17 were both significant. The mediation effect was 2.09×-0.19 = 0.40 and was significant (p= 0.00142). As the regression coefficient between AUD and expression is significant in the mediation models, it suggests that hmCG partially mediates the relationship between AUD and transcript expression at this location.

## DISCUSSION

In one of the most comprehensive methylation studies of AUD to date, we first used epigenomic deconvolution to characterize methylome-wide associations with AUD in a large discovery sample in whole blood and common blood cell-types. We identified several novel findings in CD14+ monocytes and CD3+ T-cells that may provide potential mechanisms for the increased susceptibility to infections and inflammation seen in alcohol dependent patients. Of particular interest is one of the top findings from monocytes, *PLA2G4A* (phospholipase A2 group IVA), which is required for monocyte recruitment and migration (i.e., chemotaxis) to

inflammatory sites in the body (43) including sites in brain as monocytes can cross the blood-brain barrier under such conditions (44). Alcohol is known to inhibit monocyte chemotaxis (45) which may contribute to alcohol-induced inflammation including neuroinflammation in brain. Our findings in monocytes suggest that alcohol-related mCG in *PLA2G4A* may be one of the mechanisms behind alcohol-induced inflammation.

We also examined if changes in blood showed corresponding AUD-related mCG and hmCG changes in brain tissue. The cross-tissue analyses indicated significant overlapping AUD-related methylation signals across blood and brain. This finding aligns with previous alcohol mCG studies in both model organisms and humans showing there are overlapping methylation sites associated with alcohol across whole blood and bulk brain tissue (46-49). This finding has potential clinical applications as it suggests that alcohol-related methylation changes in blood at specific sites may be biomarkers for mCG or hmCG changes occurring in brain.

A novel aspect of our study is the consideration of hydroxymethylation within the context of AUD. While no previous study has directly examined AUD-related hydroxymethylation in brain, several have explored differences in the expression of TET methylcytosine dioxygenase 1 (*TET1*), which produce a 10-11 translocation (TET) protein that directly converts methylation to hydroxymethylation (50). Studies conducted in human postmortem brain tissue have found differences in mRNA levels of *TET1* between AUD cases and controls in the cerebellum (51) and PFC (52). A mouse study of chronic intermittent ethanol exposure found increased expression of *Tet1* with increasing ethanol exposure (53). Further, rodent studies of other drugs of abuse have found hydroxymethylation to be associated with cocaine and methamphetamine administration (54, 55). Given that previous work shows there are expression differences in *TET1* between AUD cases and controls and there is some association between hydroxymethylation and other drugs of abuse, it is not unexpected that we see some evidence linking hydroxymethylation to AUD.

Our main finding with hydroxymethylation involved a site on chromosome 17 where hmCG was associated with AUD, and also partially mediated the relationship between AUD and abundance levels at transcript ENST00000572073. This protein coding transcript of *BAIAP2* (BAR/IMD domain containing adaptor protein 2) is highly expressed in brain (56). Although, the function of the specific isoform is unknown the gene encodes for a brain-specific angiogenesis inhibitor (BAI1)-binding protein. This gene has been previously linked to attention-deficit hyperactivity disorder (57, 58), autism (59), and schizophrenia (60). Mice lacking *Baiap2* (i.e., *IRSp53*$^{-/-}$*)* show cognitive and social deficits and hyperactivity, which are clinical features of the psychiatric disorders linked to *BAIAP2* (61). *BAIAP2* is highly expressed in brain where it regulates membrane/actin dynamics at excitatory synapses and dendritic spine density in neurons (61). *Baiap2* has been shown to be differentially expressed in the brains of alcohol exposed rats (versus controls) in both a binge-drinking model and a fetal alcohol exposure model (62). Further, another study showed that *BAIAP2* is hypomethylated in the PFC of AUD cases (63). If methylation was decreased as a result of AUD, we would expect to see an increase is hydroxymethylation as it is involved in the demethylation of a cytosine. Our findings suggest that increased hydroxymethylation in *BAIAP2* AUD cases leads to changes in expression. This change in expression of *BAIAP2* may lead to the cognitive deficits and other brain-impairments often seen in alcohol dependent patients.

Our findings must be interpreted in the context of potential limitations. All biosamples were collected after the development of AUD. Thus, some of the associated sites may have been present before AUD onset, and therefore may be susceptibility loci, while others occurred after AUD onset and instead reflect the disease state. To properly disentangle these effects, a study design that also included pre-AUD onset and potentially pre-drinking onset biosamples would be required. This design is not feasible for studies conducted in human postmortem brain tissue, and would only apply to studies conducted in human peripheral tissues or model organisms. Further, our results suggested potential mechanisms through which AUD-related

mCG and hmCG may lead to inflammation or the development of brain impairments. For findings that overlapped in brain, some sites did influence expression in brain tissue providing some evidence of functional effects. Possible next steps to tests these suggested mechanisms include testing findings in independent samples and exploring the mechanism through functional experiments. For example, clustered regularly interspaced short palindromic repeats-associated protein 9 (CRISPR-Cas9) based technologies would enable targeted alterations to test the downstream functional effects of associated mCG/hmCG experimentally (64-66).

In conclusion, to our knowledge, this work involves one of the first cell-type specific MWASs of AUD conducted in blood, identifying candidate methylation sites that are potentially involved in alcohol related susceptibility to infections and inflammation seen in AUD patients. Blood findings were further explored in post-mortem brain tissue, where we found significant overlap of AUD-related mCG and hmCG on site-specific basis between blood and brain. In one of the first studies to consider the role of hydroxymethylation in AUD, we found hydroxymethylation in *BAIAP2* mediated AUD expression in the PFC, potentially providing a novel mechanism for AUD cognitive impairments. Our results suggest promising new avenues for AUD research as associated methylation sites have profound translational potential as methylation can be modified by drugs and targeted epigenetic editing (65).

**Acknowledgments**

This study was supported by the National Institutes of Health (R01MH099110 to E.J.C.G., and 1R01AA026057 to S.L.C.), the VCU CTSA (UL1TR002649 from the National Institutes of Health's National Center for Advancing Translation Science and CCTR Endowment Fund of the Virginia Commonwealth University, and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation. The funding sources had no role in the study design, writing of the report or decision to submit the article for publication.

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organisation for Health Research and Development (ZonMw, grant number 10-000-1002) and financial contributions by participating universities and mental health care organizations (VU University Medical Center, GGZ inGeest, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum).

Tissues were received from the New South Wales Brain Tissue Resource Centre at the University of Sydney. Research reported in this publication was supported by the National Institute of Alcohol Abuse and Alcoholism of the National Institutes of Health under Award Number R28AA012725. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

**Disclosures**

The authors reported no biomedical financial interests or potential conflicts of interest.

## REFERENCES

1. Nestler EJ. Cellular basis of memory for addiction. Dialogues Clin Neurosci. 2013;15(4):431-43.
2. Kalivas PW, O'Brien C. Drug addiction as a pathology of staged neuroplasticity. Neuropsychopharmacology. 2008;33(1):166-80.
3. Nestler EJ. Molecular basis of long-term plasticity underlying addiction. Nat Rev Neurosci. 2001;2(2):119-28.
4. Szutorisz H, DiNieri JA, Sweet E, Egervari G, Michaelides M, Carter JM, et al. Parental THC exposure leads to compulsive heroin-seeking and altered striatal synaptic plasticity in the subsequent generation. Neuropsychopharmacology. 2014;39(6):1315-23.
5. Vassoler FM, White SL, Schmidt HD, Sadri-Vakili G, Pierce RC. Epigenetic inheritance of a cocaine-resistance phenotype. Nat Neurosci. 2013;16(1):42-7.
6. Byrnes JJ, Babb JA, Scanlan VF, Byrnes EM. Adolescent opioid exposure in female rats: transgenerational effects on morphine analgesia and anxiety-like behavior in adult offspring. Behav Brain Res. 2011;218(1):200-5.
7. Nestler EJ. Epigenetic mechanisms of drug addiction. Neuropharmacology. 2014;76 Pt B:259-68.
8. Robison AJ, Nestler EJ. Transcriptional and epigenetic mechanisms of addiction. Nat Rev Neurosci. 2011;12(11):623-37.
9. Pizzimenti CL, Lattal KM. Epigenetics and memory: causes, consequences and treatments for post-traumatic stress disorder and addiction. Genes Brain Behav. 2015;14(1):73-84.
10. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.
11. Shen-Orr SS, Gaujoux R. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. Curr Opin Immunol. 2013;25(5):571-8.
12. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM, et al. Cell type-specific gene expression differences in complex tissues. Nat Methods. 2010;7(4):287-9.
13. Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. Nat Methods. 2018;15(12):1059-66.
14. Chan RF, Turecki G, Shabalin AA, Guintivano J, Zhao M, Xie LY, et al. Cell Type-Specific Methylome-wide Association Studies Implicate Neurotrophin and Innate Immune Signaling in Major Depressive Disorder. Biol Psychiatry. 2020;87(5):431-42.
15. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol. 2012;13(6):R43.
16. Efstratiadis A. Parental imprinting of autosomal mammalian genes. Curr Opin Genet Dev. 1994;4(2):265-80.
17. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, et al. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. Nat Genet. 2008;40(7):904-8.

18. Sutherland JE, Costa M. Epigenetics and the environment. Ann N Y Acad Sci. 2003;983:151-60.
19. Aberg KA, Chan RF, Shabalin AA, Zhao M, Turecki G, Heine Staunstrup N, et al. A MBD-seq protocol for large-scale methylome-wide studies with (very) low amounts of DNA. Epigenetics. 2017:0.
20. Chan RF, Shabalin AA, Xie LY, Adkins DE, Zhao M, Turecki G, et al. Enrichment methods provide a feasible approach to comprehensive and adequately powered investigations of the brain methylome. Nucleic Acids Res. 2017;epub 25 February 2017.
21. Aberg KA, Dean B, Shabalin AA, Chan RF, Han LKM, Zhao M, et al. Methylome-wide association findings for major depressive disorder overlap in blood and brain and replicate in independent brain samples. Mol Psychiatry. 2018.
22. Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. J Psychiatr Res. 1994;28(1):57-84.
23. Aberg KA, Chan RF, van den Oord E. MBD-seq - realities of a misunderstood method for high-quality methylome-wide association studies. Epigenetics. 2020;15(4):431-8.
24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.
25. Shabalin AA, Hattab MW, Clark SL, Chan RF, Kumar G, Aberg KA, et al. RaMWAS: Fast Methylome-Wide Association Study Pipeline for Enrichment Platforms. Bioinformatics. 2018.
26. van den Oord EJ, Bukszar J, Rudolf G, Nerella S, McClay JL, Xie LY, et al. Estimation of CpG coverage in whole methylome next-generation sequencing studies. BMC Bioinformatics. 2013;14(1):50.
27. Shabalin AA, Hattab MW, Clark SL, Chan RF, Kumar G, Aberg KA, et al. RaMWAS: Fast Methylome-Wide Association Study Pipeline for Enrichment Platforms. Bioinformatics. 2018.
28. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological. 1995;57(1):289-300.
29. van den Oord EJ, Sullivan PF. False discoveries and models for gene discovery. Trends Genet. 2003;19(10):537-42.
30. Storey JD. The positive false discovery rate: A Bayesian interpretation and the q-value. Annals of Statistics. 2003;31(6):2013-35.
31. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003;100(16):9440-5.
32. Hattab MW, Shabalin AA, Clark SL, Zhao M, Kumar G, Chan RF, et al. Correcting for cell-type effects in DNA methylation studies: reference-based method outperforms latent variable approaches in empirical studies. Genome Biol. 2017;18(1):24.
33. Venet D, Pecasse F, Maenhaut C, Bersini H. Separation of samples into their constituents using gene expression data. Bioinformatics. 2001;17 Suppl 1:S279-87.
34. Onuchic V, Hartmaier RJ, Boone DN, Samuels ML, Patel RY, White WM, et al. Epigenomic deconvolution of breast tumors reveals metabolic coupling between constituent cell types. Cell reports. 2016;17(8):2075-86.
35. Koestler DC, Christensen B, Karagas MR, Marsit CJ, Langevin SM, Kelsey KT, et al. Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. Epigenetics. 2013;8(8):816-26.

36. Cabrera CP, Navarro P, Huffman JE, Wright AF, Hayward C, Campbell H, et al. Uncovering networks from genome-wide association studies via circular genomic permutation. G3 (Bethesda). 2012;2(9):1067-75.
37. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518:317.
38. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-60.
39. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290-5.
40. Imai K, Keele L, Tingley D. A General Approach to Causal Mediation Analysis. Psychological Methods. 2010;15(4):309-34.
41. Praschberger R, Lowe SA, Malintan NT, Giachello CNG, Patel N, Houlden H, et al. Mutations in Membrin/GOSR2 Reveal Stringent Secretory Pathway Demands of Dendritic Growth and Synaptic Integrity. Cell Rep. 2017;21(1):97-109.
42. Plaud C, Joshi V, Marinello M, Pastre D, Galli T, Curmi PA, et al. Spastin regulates VAMP7-containing vesicles trafficking in cortical neurons. Biochim Biophys Acta Mol Basis Dis. 2017;1863(6):1666-77.
43. Kundu S, Roome T, Bhattacharjee A, Carnevale KA, Yakubenko VP, Zhang R, et al. Metabolic products of soluble epoxide hydrolase are essential for monocyte chemotaxis to MCP-1 in vitro and in vivo. J Lipid Res. 2013;54(2):436-47.
44. Prinz M, Priller J. The role of peripheral immune cells in the CNS in steady state and disease. Nat Neurosci. 2017;20(2):136-44.
45. Imhof A, Blagieva R, Marx N, Koenig W. Drinking modulates monocyte migration in healthy subjects: a randomised intervention study of water, ethanol, red wine and beer with or without alcohol. Diab Vasc Dis Res. 2008;5(1):48-53.
46. Barker JM, Zhang Y, Wang F, Taylor JR, Zhang H. Ethanol-induced Htr3a promoter methylation changes in mouse blood and brain. Alcohol Clin Exp Res. 2013;37 Suppl 1:E101-7.
47. Clark SL, Costin BN, Chan RF, Johnson AW, Xie L, Jurmain JL, et al. A Whole Methylome Study of Ethanol Exposure in Brain and Blood: An Exploration of the Utility of Peripheral Blood as Proxy Tissue for Brain in Alcohol Methylation Studies. Alcohol Clin Exp Res. 2018;42(12):2360-8.
48. Gatta E, Grayson DR, Auta J, Saudagar V, Dong E, Chen Y, et al. Genome-wide methylation in alcohol use disorder subjects: implications for an epigenetic regulation of the cortico-limbic glucocorticoid receptors (NR3C1). Mol Psychiatry. 2019.
49. Meng W, Sjoholm LK, Kononenko O, Tay N, Zhang D, Sarkisyan D, et al. Genotype-dependent epigenetic regulation of DLGAP2 in alcohol use and dependence. Mol Psychiatry. 2019.
50. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell. 2011;145(3):423-34.
51. Gatta E, Auta J, Gavin DP, Bhaumik DK, Grayson DR, Pandey SC, et al. Emerging Role of One-Carbon Metabolism and DNA Methylation Enrichment on delta-Containing GABAA Receptor Expression in the Cerebellum of Subjects with Alcohol Use Disorders (AUD). Int J Neuropsychopharmacol. 2017;20(12):1013-26.

52. Guidotti A, Dong E, Gavin DP, Veldic M, Zhao W, Bhaumik DK, et al. DNA methylation/demethylation network expression in psychotic patients with a history of alcohol abuse. Alcohol Clin Exp Res. 2013;37(3):417-24.
53. Finegersh A, Ferguson C, Maxwell S, Mazariegos D, Farrell D, Homanics GE. Repeated vapor ethanol exposure induces transient histone modifications in the brain that are modified by genotype and brain region. Front Mol Neurosci. 2015;8:39.
54. Feng J, Shao N, Szulwach KE, Vialou V, Huynh J, Zhong C, et al. Role of Tet1 and 5-hydroxymethylcytosine in cocaine action. Nat Neurosci. 2015;18(4):536-44.
55. Jayanthi S, Gonzalez B, McCoy MT, Ladenheim B, Bisagno V, Cadet JL. Methamphetamine Induces TET1- and TET3-Dependent DNA Hydroxymethylation of Crh and Avp Genes in the Rat Nucleus Accumbens. Mol Neurobiol. 2018;55(6):5154-66.
56. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. Nucleic Acids Res. 2020;48(D1):D682-D8.
57. Hasler R, Preti MG, Meskaldji DE, Prados J, Adouan W, Rodriguez C, et al. Inter-hemispherical asymmetry in default-mode functional connectivity and BAIAP2 gene are associated with anger expression in ADHD adults. Psychiatry Res Neuroimaging. 2017;269:54-61.
58. Ribases M, Bosch R, Hervas A, Ramos-Quiroga JA, Sanchez-Mora C, Bielsa A, et al. Case-control study of six genes asymmetrically expressed in the two cerebral hemispheres: association of BAIAP2 with attention-deficit/hyperactivity disorder. Biol Psychiatry. 2009;66(10):926-34.
59. Toma C, Hervas A, Balmana N, Vilella E, Aguilera F, Cusco I, et al. Association study of six candidate genes asymmetrically expressed in the two cerebral hemispheres suggests the involvement of BAIAP2 in autism. J Psychiatr Res. 2011;45(2):280-2.
60. McKinney B, Ding Y, Lewis DA, Sweet RA. DNA methylation as a putative mechanism for reduced dendritic spine density in the superior temporal gyrus of subjects with schizophrenia. Transl Psychiatry. 2017;7(2):e1032.
61. Kang J, Park H, Kim E. IRSp53/BAIAP2 in dendritic spine development, NMDA receptor regulation, and psychiatric disorders. Neuropharmacology. 2016;100:27-39.
62. Lussier AA, Stepien KA, Neumann SM, Pavlidis P, Kobor MS, Weinberg J. Prenatal alcohol exposure alters steady-state and activated gene expression in the adult rat brain. Alcohol Clin Exp Res. 2015;39(2):251-61.
63. Johansson S. Post mortem studies on the human alcoholic brain : DNA methylation and molecular responses. Stockholm, Sweden: Karolinksa University; 2008.
64. Kwon DY, Zhao YT, Lamonica JM, Zhou Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. Nat Commun. 2017;8:15315.
65. Vojta A, Dobrinić P, Tadić V, Bočkor L, Korać P, Julg B, et al. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic acids research. 2016.
66. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, et al. Editing DNA Methylation in the Mammalian Genome. Cell. 2016;167(1):233-47.e17.

**Table 1.** Enrichment of Top Blood MWAS findings for brain MWAS findings

| | Brain Methylation | | | | Brain Hydroxymethylation | | | |
|---|---|---|---|---|---|---|---|---|
| Blood MWAS | Gene Overlap | Top Threshold | Odds Ratio | P-value | Gene Overlap | Top Threshold | Odds Ratio | P-value |
| Whole Blood | 525 | 0.5;0.5 | 1.019 | 0.667 | 97 | 0.5;0.1 | 0.883 | 0.987 |
| T-cells (CD03+) | 98 | 0.5;0.1 | 0.997 | 0.814 | 99 | 0.5;0.1 | 0.905 | 0.977 |
| Monocytes (CD14+) | 149 | 0.1;0.5 | 1.378 | 0.009 | 615 | 0.5;0.5 | 1.128 | 0.035 |
| Granulocytes (CD15+) | 96 | 0.1;0.5 | 0.982 | 0.858 | 108 | 0.1;0.5 | 0.988 | 0.846 |
| B-cells (CD19+) | 144 | 0.1;0.5 | 1.368 | 0.007 | 38 | 0.1;0.1 | 1.741 | 0.011 |

MWAS, methylome-wide association study

**Table 2.** Significant cis-MeQTLs and mediation results

| Methylation Type | Gene | Chr | BP | Transcript Name | MeQTL p-value | Mediation Effect | Mediation P-value |
|---|---|---|---|---|---|---|---|
| hmCG | *ANKRD13C* | 1 | 70775032 | ENST00000498735 | 0.004 | -0.067 | 0.624 |
| hmCG | *ANKRD13C* | 1 | 70775035 | ENST00000498735 | 0.004 | -0.071 | 0.596 |
| mCG | *ERC2* | 3 | 56104924 | ENST00000466358 | 0.007 | 0.065 | 0.672 |
| hmCG | *ERICH1-AS1* | 8 | 843888 | ENST00000524139 | 0.007 | 0.065 | 0.534 |
| hmCG | *PRMT7* | 16 | 68350898 | ENST00000441236 | 0.002 | 0.042 | 0.73 |
| hmCG | *BAIAP2* | 17 | 79062899 | ENST00000572073 | 0.007 | -0.407 | 0.00142 |
| hmCG | *NEDD4L* | 18 | 55923590 | ENST00000590248 | 0.002 | -0.087 | 0.558 |
| mCG | *NFATC1* | 18 | 77193667 | ENST00000545796 | 0.009 | 0.268 | 0.266 |

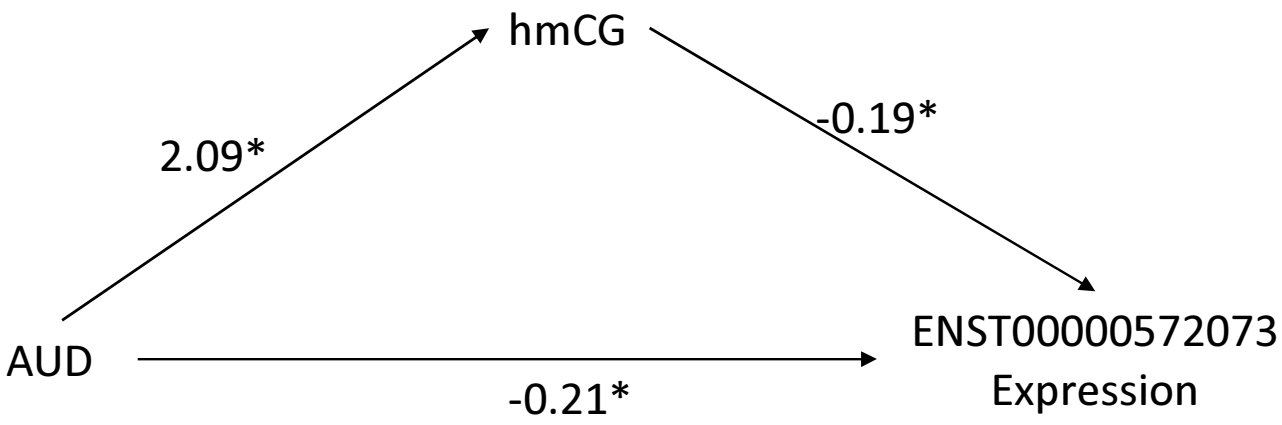
## FIGURE LEGEND

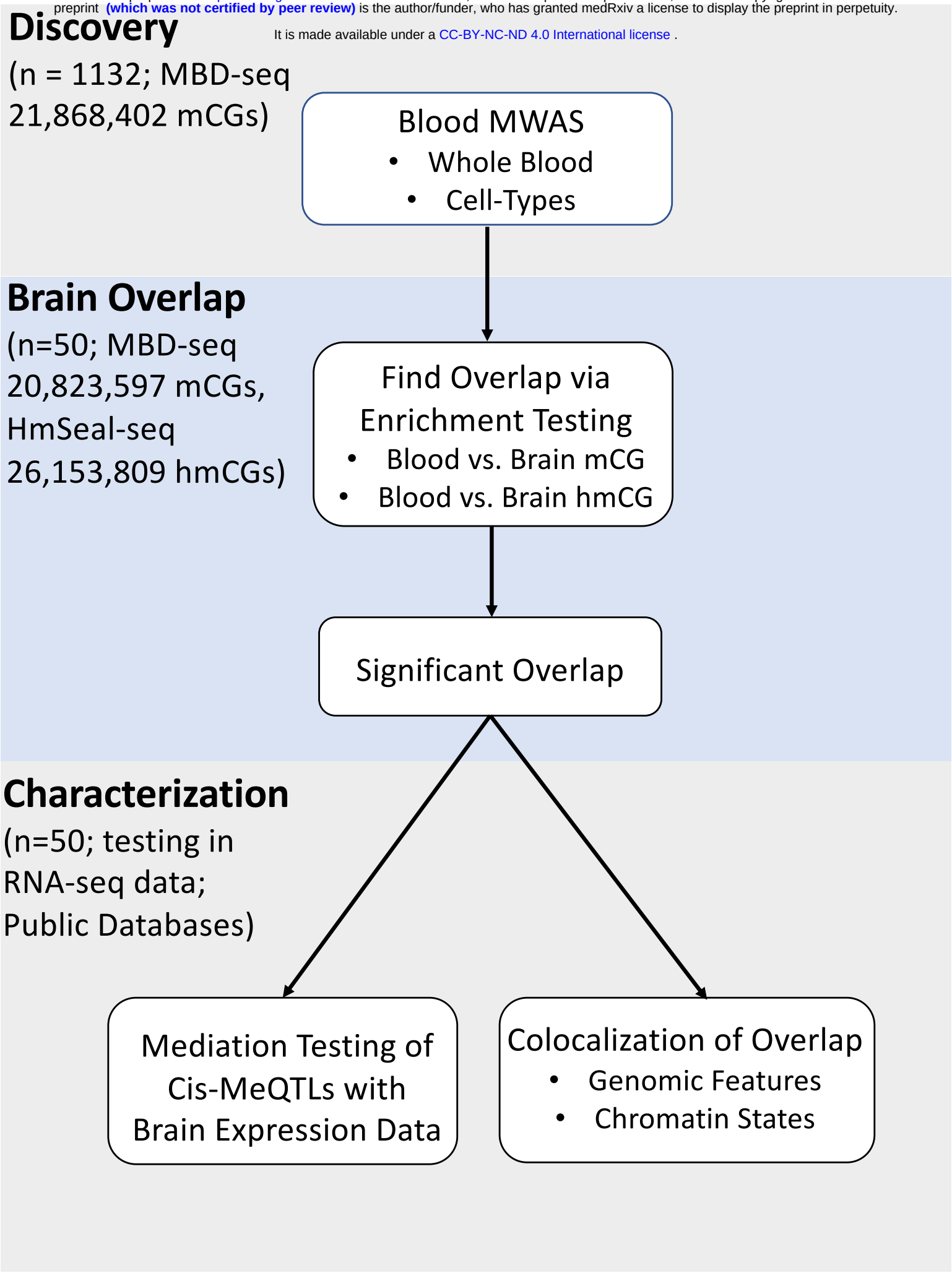
**Figure 1.** Study Overview. In this study we present a three-stage, multi-tissue design: **Top**) a Discovery stage involving a MWAS of alcohol use disorder in a large set of blood samples in whole blood and granulocytes/T-cells/B-cells/monocytes; **Middle)** a Brain Overlap stage where we test for enrichment of MWAS overlap between blood and brain; **Bottom)** a Characterization stage where we examine if sites that overlapped across tissue were associated with transcript expression (i.e., cis-MeQTLs) and mediated AUD-associated transcript expression, and determined if overlapping sites colocalized with genomic feature and chromatin states using information from public databases.

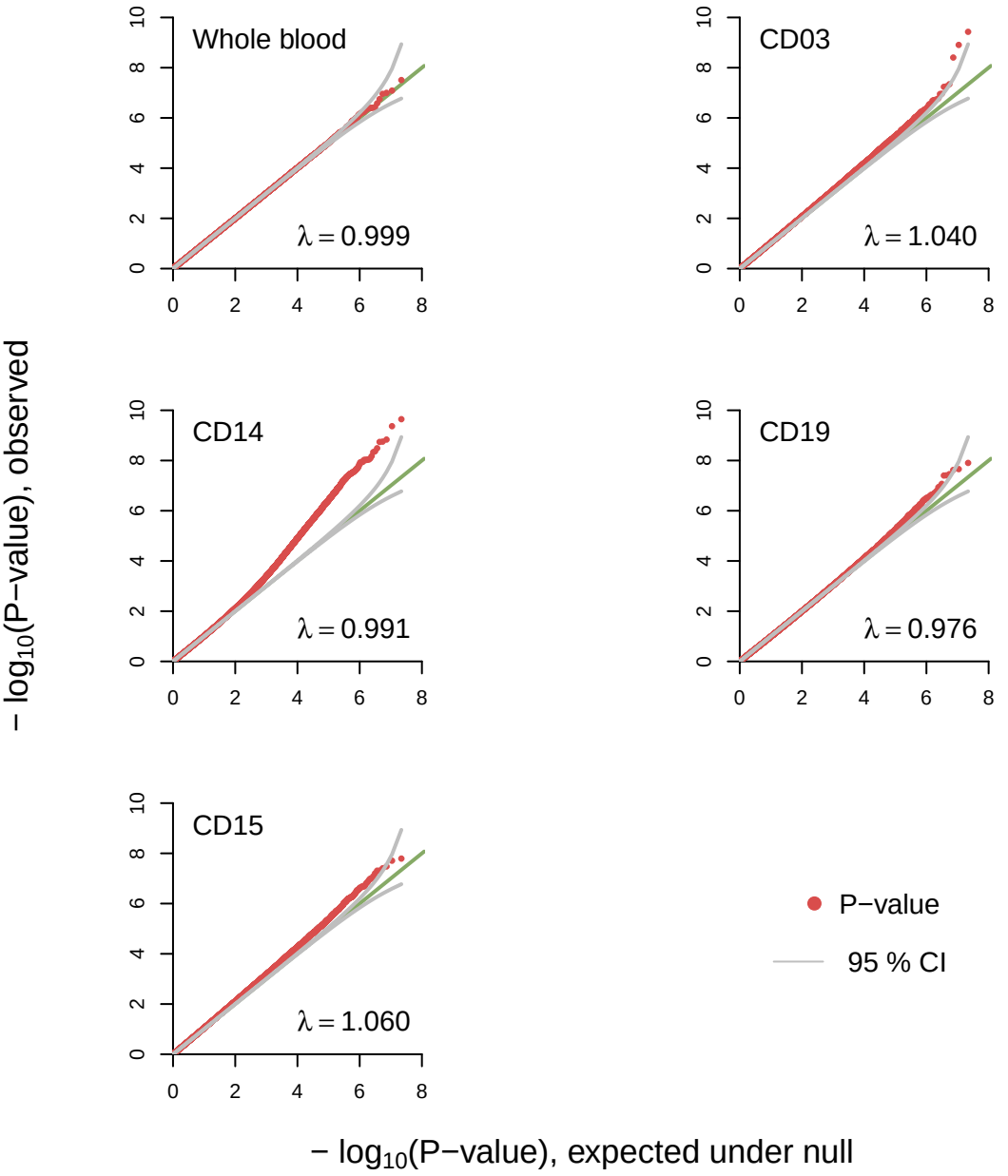
**Figure 2.** Quantile-Quantile plots for whole blood and cell-type specific MWAS of AUD
The observed p-values (red dots), on a -log10 scale, are plotted against their expected values (grey main diagonal line) under the null hypothesis assuming none of the CpGs have an effect. Grey lines indicate the 95% confidence bands (CI). A deviation of the observed p-values from the main diagonal indicates that there are CpGs associated with AUD. Coefficient lambda ($\lambda$) will be close to one if the vast majority of CpGs behave as expected under the null hypothesis. Analyses were performed in whole blood as well as blood cell subpopulations using a reference panel that was sorted using cell surface antigen molecules CD3, CD14, CD15, and CD19 that are expressed on the surface of T-cells, monocytes, granulocytes, and B-cells, respectively.

**Figure 3**. Unstandardized regression coefficients for the relationship between AUD and expression of transcript ENST00000572073 as mediated by hydroxymethylation (hmCG).

*p < 0.05; hmCG is the hydroxymethylation level at location 79,062,899 on chromosome 17






Whole blood
λ = 0.999
CD03
λ = 1.040
CD14
λ = 0.991
CD19
λ = 0.976
CD15
λ = 1.060
P−value
95 % CI
− log10(P−value), observed
− log10(P−value), expected under null