EPIGENETIC AND EXPRESSION-BASED MODIFICATIONS OF DEPRESSION-RELEVANT GENES AS PERIPHERAL BIOMARKERS FOR TREATMENT-RESISTANT DEPRESSION AND MAJOR DEPRESSIVE DISORDER

BY

JOHN R. PFEIFFER

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Master's Committee:

Associate Professor Monica Uddin, Chair
Assistant Professor Dave Zhao, Co-Chair

Abstract

Mechanisms contributing to treatment resistant depression (TRD) and major depressive disorder (MDD) are poorly understood. Identifying peripheral biomarkers in an easily accessible tissue, such as whole blood, will enhance the mechanistic understanding of these disorders and could potentially establish targets for personalized treatment development. It has been shown that CpG methylation-based epigenetic modifications in the brain are associated with stress-related phenotypes. Therefore, the first portion of the current study sought to identify CpG methylation-based peripheral biomarkers of TRD concordant with brain CpG methylation in the rat, in a hormone-induced model of TRD.

Many factors, both environmental and biological, can alter CpG methylation in blood and/or brain tissues in eukaryotes. This differential CpG methylation is a well-documented functional mechanism by which genes are differentially expressed. Depending on the specific gene, this functional alteration of expression can have downstream physiological and phenotypic effects. Taking into account the need to identify peripheral biomarkers of MDD in humans, the second portion of the current study sought to identify the association between lifetime MDD diagnosis status and peripheral, blood-derived methylation and expression. Finally, supplementary analyses of publically available datasets highlight the association between neuronal CpG methylation and lifetime MDD diagnosis status and using a cohort of deceased subjects with balanced prevalence of lifetime MDD cases and "healthy" controls, as well as intra-subject brain and blood CpG methylation correlations in "healthy" controls.

The current study identified one salient peripheral biomarker for TRD that is concordant in CNS tissue, as well as one potential peripheral biomarker for MDD. In the TRD rat model, the first CpG site upstream of the transcription start site (TSS) in the promoter region of *Slc6a3* was differentially methylated between case and control group animals in both mPFC and whole blood. Second, using blood drawn from human participants from the Detroit Neighborhood Health Study, *SLC6A4* mRNA expression approached a significant association with lifetime MDD diagnosis status. Efforts to elucidate additional peripheral biomarkers should continue to expand to include polygenic CpG methylation measurements in multiple tissues.

Table of Contents

Chapter 1: Introduction

Both major depressive disorder (MDD) and treatment resistant depression (TRD) are mood disorders that manifest heterogeneously in regards to risk factors, onset, symptoms, and treatment response[1]. To be diagnosed with MDD a patient must exhibit two or more weeks of five of the following symptoms as a departure from their basal state: significant unintentional weight loss/gain, significant differential sleep patterns, irritability, fatigue/energy loss, feelings of worthlessness/guilt, difficulty concentrating/thinking, or recurring thoughts of self-harm/suicide, with at least one of the symptoms being either anhedonia and/or depressed mood[2]. Lifetime prevalence of MDD in the United States (US) is estimated at ~16%, and past year prevalence of MDD is estimated at ~6% in the US[3]. MDD is commonly comorbid with anxiety, substance abuse, and personality disorders[4], and patients afflicted with MDD are at a significantly heightened risk for suicide completion compared to the general population[5]. MDD is also considered the leading cause of non-fatal disease burden in the world[6].

Numerous treatment options exist for patients with MDD, including deep brain stimulation, electroconvulsive therapy, psychotherapy, and pharmacological intervention. The most common treatment course is pharmacological intervention by way of tricyclic antidepressants (TCAs), monoamine oxidase inhibitors (MAOIs), selective serotonin (5HT) and norepinephrine (NE) reuptake inhibitors (SNRIs), or selective 5HT reuptake inhibitors (SSRIs). Although a wide range of pharmacological treatment options exist, only 30-50% of patients reach remission with appropriate use[7,8], and 10-20% of care-seeking patients remain significantly symptomatic even after two years of treatment[9]. Clinical diagnosis of TRD is typical after failed pharmacological intervention with two separate classes of antidepressants[10]. The personal and societal costs of TRD compared to MDD are great: patients with TRD spend more time in medical care, have an increased quantity of comorbid disorders, and are more likely to experience thoughts of self-harm/suicide[11].

Hypothalamic-pituitary-adrenal (HPA) axis hyperactivation is a common and well-characterized phenomenon in both TRD and MDD[12]. The HPA axis involves a neuroendocrine negative feedback loop in which corticotropin release hormone (CRH) is released from the hypothalamus in response to psychological stress. This stimulates the anterior pituitary to release adrenocorticotropin hormone (ACTH), which signals the

adrenal cortices to release cortisol (CORT). High CORT levels contribute to a negative feedback mechanism in the HPA axis, as they result in an inhibition of CRH release from the hypothalamus and ACTH from the anterior pituitary. Importantly, a meta-analytic study including over 18,000 participants across 361 original studies determined that depressed individuals displayed elevated levels of both CORT and ACTH compared to "healthy" controls[1].

The genes *SLC6A3* and *SLC6A4*, the protein products that they encode, and the neurotransmitters that the proteins interact with have also been implicated in the etiology of MDD, TRD, and other mental health disorders. *SLC6A3* and *SLC6A4* encode the dopamine (DA) and the 5HT reuptake transporter proteins, respectively (DAT and SERT). These membrane-bound sodium-dependent transport proteins function to traffic DA and 5HT from the synaptic cleft back into the presynaptic neuron for metabolism or reuse/repackaging. Buproprion, an NE and DA reuptake inhibitor, is commonly used to treat depression and other mood disorders. The mechanism of action of Buproprion is antagonism of DAT. This mechanism has been shown to increase extracellular levels of DA in the nucleus accumbens (NAC)[13] and the prefrontal cortex (PFC) in rodent models of depression[14]. SSRI's are another common class of antidepressants used to treat many types of depression and mood disorders. SSRIs act mechanistically to antagonize SERT, thereby inhibiting the reuptake of 5HT from the cleft and providing an increased quantity of 5HT for the postsynaptic neuron to utilize. Importantly, a previously validated ACTH-induced treatment-resistant depression model in the rat[15] showed significantly different extracellular 5HT and DA levels in the PFC of ACTH-exposed animals compared to saline-exposed control animals after the stress of a Porsolt forced swim test (FST). Specifically, the extracellular DA PFC levels of ACTH-exposed animals were *decreased* compared to controls, whereas the extracellular 5HT PFC levels of ACTH-exposed animals were *increased* compared to controls. This model also showed that a common TCA (Imipramine) did not rescue chronic ACTH-exposed animals immobility time in the FST compared to animals exposed chronically to ACTH, and no Imipramine, suggesting an endophenotype of depression-like behaviors that is resistant to TCA[16]. Additionally, hyper-ACTH activation in patients with Cushing's disease leads to an increased occurrence of depression resistant to TCA treatment[17,18]. These results provide strong evidence that through chronic ACTH exposure in rodents, neurotransmitter alterations occur which may mimic TRD in humans.

In humans, sequence variations in the *SLC6A3* and *SLC6A4* genes have been associated with MDD status[19]. These sequence variations have also been associated with differential expression of the mRNA transcripts mapping to these genes[20,21], although inconsistencies in the literature have been reported. Regardless, these sequence variations are static in nature and do not account for interaction of environmental factors (such as stress) that may affect these systems. One mechanism that helps to explain environmental factors affecting physiological systems is DNA methylation: the covalent addition of a methyl group to the 5' position of a cytosine residue, typically where cytosine is adjacent to a guanine residue (i.e. CpG methylation). CpG methylation functions as a regulator of transcriptional activity in both eukaryotes and prokaryotes. This function is especially salient when differential CpG methylation occurs in the region directly upstream of the transcription start site (promoter region), which can be rich in CpG residues (CpG island) compared to other regions of the genome. Increased CpG methylation of a CpG island in a promoter region can sterically inhibit the association of polymerases and transcription factors functioning to transcribe mRNA[22], thereby facilitating a reduction in expression of the downstream locus.

Taking into account the great personal and societal costs of the TRD and MDD, it is apparent that characterizing the underlying neurobiological mechanisms of both disorders is necessary. However, significant hurdles remain that inhibit this desired progress. Primarily, the inability to access central nervous system (CNS) tissue of living patients creates an inability to assay these etiological tissues. In turn, this key deficiency creates a need to identify salient peripheral biomarkers of depression in humans (such as in the blood or saliva), and to utilize animal models in which CNS biomarkers of TRD can be compared to peripheral biomarkers; however, current research has yet to determine the degree to which peripheral biomarkers correspond to relevant CNS biomarkers of the disorders[23,24].

To address these gaps in knowledge, we have undertaken a three-pronged approach in this research project. Utilizing the previously validated ACTH-induced TRD model in the rat[15,16], the first portion of the current study assays CpG methylation of *Slc6a3* and *Slc6a4* promoter regions in rat whole blood and compares it to CpG methylation in medial PFC (mPFC), in order to identify CNS-relevant peripheral biomarkers of TRD. *Our working hypothesis is that ACTH-exposed animals (case group) will exhibit significantly differential CpG methylation in the promoter regions near the TSS of Slc6a3 and Slc6a4 genes in mPFC tissue and whole*

blood, compared to control group animals. Additionally, we hypothesize that mPFC CpG methylation in both case and control group animals will significantly correlate with intra-group measurements of CpG methylation in whole blood.

The second portion of the current study utilizes available genomic data from the Detroit Neighborhood Health Study (DNHS) in order to investigate the relationship between peripheral measures of CpG methylation, peripheral measures of mRNA transcription, and lifetime MDD diagnosis status. The DNHS is a longitudinal epidemiological study of adults from Detroit, Michigan[26,27]. Participants were interviewed via phone, and demographic, socio-economic, and standard assessments of MDD were administered. Additionally, whole blood was collected for genome-wide methylation and expression assays, as well as other genomic data collection. *The second portion of the current study hypothesizes that CpG methylation of probes mapping to the human promoter regions near the TSS of SLC6A3 and SLC6A4 genes will be significantly associated with mRNA expression of probes mapping to their respective genes. Additionally, we hypothesize that mRNA expression of probes mapping to the human SLC6A3 and SLC6A4 genes will be significantly associated with lifetime MDD diagnosis status. Finally, we hypothesize that CpG methylation of probes mapping to the promoter regions near the TSS of human SLC6A3 and SLC6A4 genes will be significantly associated with lifetime MDD diagnosis status.*

The third portion of the current study utilizes a publically available dataset in which genome-wide neuronal CpG methylation was profiled in frontal cortex from deceased MDD patients, as well as "healthy" controls (GEO Accession No. GSE41826), in order to test the association between brain CpG methylation and lifetime MDD diagnosis status. *The third portion of the current study therefore hypothesizes that frontal cortex neuronal CpG methylation of probes mapping to the promoter regions near the TSS of the human SLC6A4 gene will be significantly associated with lifetime MDD diagnosis status.*

Finally, to complement the human-based analyses of DNHS and GEO-derived data, and to investigate whether CpG methylation levels in the blood are related to CpG methylation levels in CNS tissues within our loci and tissues of interest, we undertook additional secondary analyses of publically available data. To highlight *intra*-subject brain and blood CpG methylation correlations, we used the Blood Brain DNA Methylation Comparison Tool available at http://epigenetics.iop.kcl.ac.uk/bloodbrain[28]. *The final portion of the*

current study therefore hypothesizes that CpG methylation of probes mapping to the promoter regions of human SLC6A3 and SLC6A4 genes will show significant correlations within subjects, between blood and PFC tissue.

In summation of the aforementioned hypotheses, the current study offers a model by which CpG methylation and mRNA expression relate to depression phenotypes in both humans and rats (Figure 1). In relationship A. of the model, differential CpG methylation is associated with differential mRNA expression. In relationship B. of the model, differential mRNA expression is associated with depression phenotypes. Finally, in relationship C. of the model, differential CpG methylation is associated with depression phenotypes in both humans and rats.

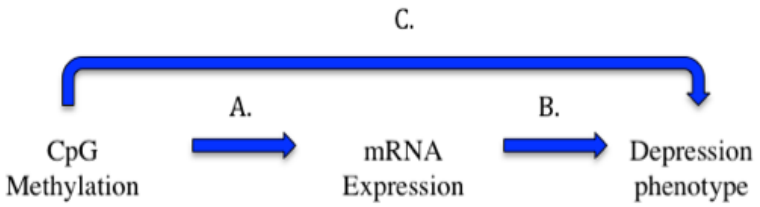

Figure 1. Schematic model of hypothesized effects between CpG methylation, mRNA expression, and depression phenotypes in human and rat.

Chapter 2: Methods

ACTH-Induced Treatment-Resistant Depression Model in the Rat

The first phase of the current study utilized brain tissue and whole blood harvested from a previously validated ACTH treatment model of TRD in rats[15,16], provided by collaborator Dr. Susannah Tye. Briefly, case animals (n = 12) were injected daily with ACTH (100ug/day) dissolved in distilled water (1mL total volume). Control animals (n = 13) were injected daily with 0.9% saline (1mL total volume). After 14 days, animals were sacrificed via i.p. injection of sodium pentobarbitone (0.4mL). 200-300uL of cardiac whole blood was collected from each subject post-mortem. Additionally, whole brains were removed from the subjects. mPFC tissue was sectioned and snap frozen. For all subjects, whole blood and mPFC tissues were shipped to the Uddin Lab Group on dry ice. Genomic DNA was then isolated from samples using the Qiagen Animal Blood and Tissue DNA kit (Catalogue No. 69504; Qiagen, Hilden, Germany). The silica-based membrane spin-column protocol was used for all tissues and samples following manufacturer's recommended protocol. After isolation, genomic DNA quantity and quality were assessed using spectrophotometry (NanoDrop) methods. NanoDrop concentrations for whole blood isolated genomic DNA ranged from 12.58 ng/uL to 95.05 ng/uL. NanoDrop concentrations for mPFC isolated genomic DNA ranged from 88.64 ng/uL to 220.80 ng/uL.

After DNA quantification and purity assessment, the isolated genomic DNA (750 ng/sample) underwent bisulfite conversion (BSC), using the Qiagen Epitect Bisulfite Kit (Catalogue No. 59104) and following the manufacturer's recommended protocol. In brief, BSC is used in order to enable the assay of CpG methylation in DNA during subsequent protocols. DNA is denatured, and unmethylated cytosine residues are converted to uracil with sodium bisulfite incubation; methylated cytosines remain uncoverted, resulting in sequence differences between methylated and unmethylated CpG sites within DNA (i.e. "Ts" vs. "Cs"). NanoDrop concentrations for whole blood isolated BSC DNA ranged from 1.88 ng/uL to 26.20 ng/uL. NanoDrop concentrations for mPFC isolated BSC DNA ranged from 2.92 ng/uL to 8.76 ng/uL.

Polymerase chain reaction (PCR) of BSC DNA was then performed to amplify locus-specific regions of DNA pertaining to *Slc6a3*[29] and *Slc6a4*[30] TSS's and CpG Islands (CGI's). Qiagen's Pyromark PCR Kit (Catalogue No. 978703) was used for all samples. All reactions were performed in duplicate using 25ng of sample BSC DNA per primer set in a 25uL total reaction volume. PCR primer set information for *Slc6a3* and

Slc6a4 are provided in Table 1. Presence of desired amplicon was visually confirmed in each sample by loading 4.5 ul of each reaction on a 2% agarose gel.

Slc6a3 primer sets	Sequence
PCR - Forward	5' - TAG GGT TAG TGG AGT AGT GGA GAT - 3'
PCR - Reverse	5' - /BIOT/ CCC AAA TTC AAT AAC CTC TAA CTC CC - 3'
SEQ - CpG 11-15 (BSC)	5' - AAG TTT TTA GTT TTT TTT TAT AGT - 3'
SEQ - CpG 16-20 (BSC)	5' - GGG GTA GGT TTG TTT AGT - 3'
Slc6a4 primer sets	Sequence
PCR - Forward	5' - GGA GAA ATA AAG TTT TTT GGT TTT AAA GT - 3'
PCR - Reverse	5' - /BIOT/ ACA CCC TAA ATC CAA ACA ATC A - 3'
SEQ - CpG 12-15 (BSC)	5' - AAG TTT TTT GGT TTT AAA GTG A - 3'
SEQ - CpG 16-23 (BSC)	5' - TGA GGA GGT TGT TTA GAG ATT AGA - 3'

Table 1. PCR and pyrosequencing primer sets of target *Slc6a3* and *Slc6a4* loci.

The total remaining volume of the PCR reactions was then used in bisulfite pyrosequencing to assay percent CpG methylation at specific CpG sites within the BSC amplified PCR fragments. For this procedure, two sequencing primers for each of *Slc6a4* and *Slc634* were generated de novo using the PyroMark Assay Design software (Version 2.0.1.15) based off of the predicted sequence of the PCR amplicon after BSC. The aforementioned sequencing primer sets enabled assay of 12 CpG sites in the *Slc6a4* amplicon, and 10 CpG sites in the *Slc6a3* amplicon. For *Slc6a4*, 12 Cpg sites within the first exon were assayed. This location lies within a CGI, with a putative transcription factor-binding site (TFBS) also present (Figure 2).

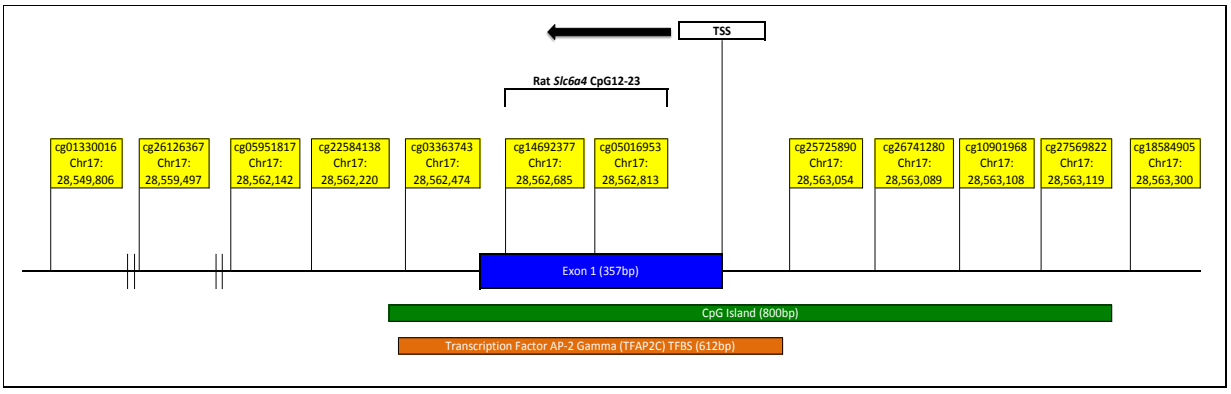

Figure 2. *SLC6A4* Mayo and DNHS CpG Methylation Probe Locations

For *Slc6a3*, the first ten CpG sites upstream of the TSS in the promoter region were assayed. Cpg1 is the first CpG site upstream of the TSS, whereas CpG10 is the tenth site upstream of the TSS (Figure 3).

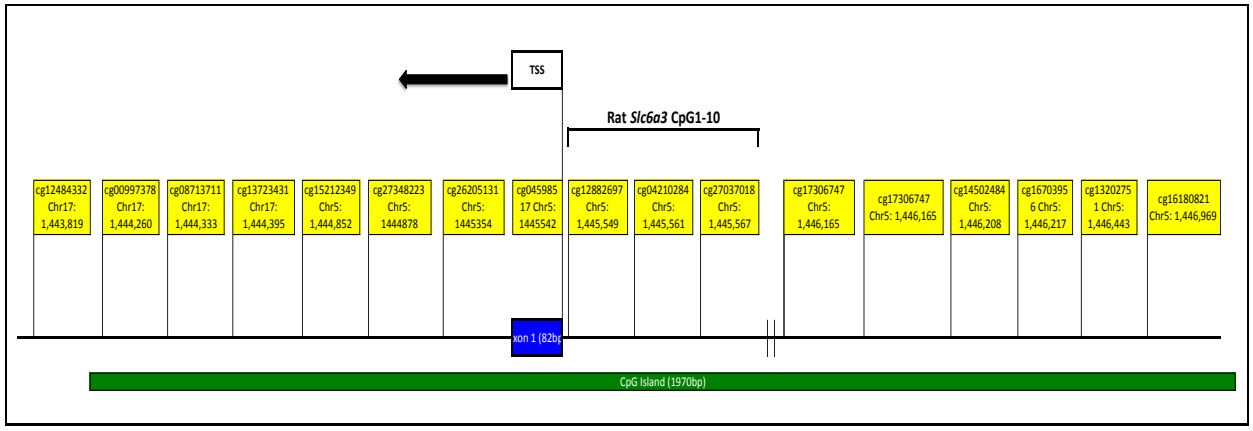

Figure 3. *SLC6A3* Mayo and DNHS CpG Methylation Probe Locations

Sequencing primer set information for *Slc6a4* and *Slc6a3* are provided in Table 1. Bisulfite pyrosequencing was performed on the Qiagen Pyrosequencing Q24 Advanced machine using the Qiagen PyroMark CpG Assay Kit (Catalogue No. 978746). Briefly, after PCR amplification, dsDNA is denatured and the biotinylated strand serves as the template for the sequencing primers in the pyrosequencing reaction. This template is incubated with DNA polymerase, ATP sulfurylase, luciferase, and apyrase, as well as the substrates adenosine 5' phosphosulfate and luciferin. Deoxyribonucleotide triphosphate (dNTP) molecules are added to the sequencing primer on the template strand by DNA polymerase if the dNTPs are complementary to the template strand. When this occurs, Pyrophosphate (PPi) is cleaved from the dNTP, and the biochemical properties of the aforementioned factors enable the generation of light signal proportional to number of dNTPs incorporated. In this way, sequencing by synthesis was performed, and with BSC DNA CpG methylation ratios can be determined[31]. Locus-specific methylation values at each CpG site were recorded by the Q24 machine and exported to Microsoft Excel.

Replicates exhibiting intra-sample differences greater than 5.4% were excluded from analysis. For each CpG site, case group versus control group means were compared using independent samples t-tests. Afterwards, t-test p-values were false discovery rate (FDR) corrected using the Benjamini-Hochberg (BH) approach[32]. Due to the exploratory nature of these analyses, unadjusted p values are presented first, followed

by FDR-corrected results. Standard error of the mean (SEM) was calculated and included as error bars in the figures. In addition, Pearson correlation analyses were performed between mPFC and whole blood methylation values within each gene, within each case or control group. P values, and correlation coefficients are presented.

Lifetime Major Depressive Disorder, Peripheral Methylation and Expression in the DNHS

DNHS Sample. The Detroit Neighborhood Health Study (DNHS) is a longitudinal epidemiological study of adults older than 18 years old from Detroit, Michigan[26,27]. The DNHS study was performed in five waves; however, all measures relevant to the current study were gathered in wave 2. One participant per household was randomly selected from a total pool of 1547 households. Structured telephone interviews were conducted for each consenting participants, and included assessments of demographic (age, biological sex, self-reported race), socio-economic (educational attainment, household income), and psychiatric symptoms, including lifetime MDD according to DSM-IV criteria[26]. A subset of DNHS participants provided blood samples via venipuncture (VP) through clinical in-home visits, at which time medication use was also recorded. From these VP blood draws, DNA was isolated from whole blood using the Qiagen QIAamp DNA Mini Kit (Catalogue No. 51306), and total RNA was isolated using the Ambion Leukolock kit (Catalogue No AM1923, Ambion, Foster City, CA).

Methylation Microarray Analyses. 1ug of the whole blood isolated genomic DNA from 192 DNHS samples was subjected to BSC using the Zymo EZ-96 DNA Methylation Kit (Catalogue No. D5003, Zymo, Irvine, CA) and applied to the Illumina HumanMethylation450 (450k) BeadChip Kit (Catalogue No. WG-314-1002, Illumina, San Diego, CA), which measures methylation at more than 485,000 CpG sites per sample at single nucleotide resolution. Raw output from the 450k microarray was background corrected using the Illumina GenomeStudio Software and exported for additional analyses in the R programming language (version 3.2.4). DNA methylation measurements are represented by beta values, with 0 representing a completely unmethylated locus and 1 representing a completely methylated locus. Using the CpGAssoc package, the beta values matrix was subjected to quality control[33]. CpG sites were removed where the proportion of missing values across all samples exceeded 10%, and where the detection p-value was less than 0.001. Cross-hybridizing probes and CpG sites near known single nucleotide polymorphisms[34] were filtered out using the wateRmelon package in R[35]. Beta mixture quantile (BMIQ) normalization by subject was then implemented to

correct for probe type bias[36]. Then, principle component analysis (PCA) was performed on the beta values matrix to identify significant associations between technical artifacts/demographic measures and variability throughout the data. The identified chip and position technical artifacts were then corrected for while preserving the effects of biological sex and lifetime MDD status using ComBAT in R[37], as implemented in the SVA package[38]. After ComBAT, PCA was again performed to confirm chip and position technical artifact corrections. After the aforementioned background correction, quality control, filtering, normalization, and technical artifact correction, beta values from CpG sites near *SLC6A4* and *SLC6A3* TSS/CGI (Tables 2, 3) were extracted for each participant and compiled.

Illumina ID	Genome Build	Chromosome	Probe Location	UCSC Gene Name	UCSC Gene Accession Number	UCSC Gene Functional Region	UCSC Associated CGI
cg01330016	37/hg19	17	28549806	*SLC6A4*	NM_001045	5'UTR	
cg26126367	37/hg19	17	28559497	*SLC6A4*	NM_001045	5'UTR	chr17:28562387-28563186
cg05951817	37/hg19	17	28562142	*SLC6A4*	NM_001045	5'UTR	chr17:28562387-28563186
cg22584138	37/hg19	17	28562220	*SLC6A4*	NM_001045	5'UTR	chr17:28562387-28563186
cg03363743	37/hg19	17	28562474	*SLC6A4*	NM_001045	5'UTR	chr17:28562387-28563186
cg14692377	37/hg19	17	28562685	*SLC6A4*	NM_001045	1stExon;5'UTR	chr17:28562387-28563186
cg05016953	37/hg19	17	28562813	*SLC6A4*	NM_001045	1stExon;5'UTR	chr17:28562387-28563186
cg25725890	37/hg19	17	28563054	*SLC6A4*	NM_001045	TSS200	chr17:28562387-28563186
cg26741280	37/hg19	17	28563089	*SLC6A4*	NM_001045	TSS200	chr17:28562387-28563186
cg10901968	37/hg19	17	28563108	*SLC6A4*	NM_001045	TSS200	chr17:28562387-28563186
cg27569822	37/hg19	17	28563119	*SLC6A4*	NM_001045	TSS200	chr17:28562387-28563186
cg18584905	37/hg19	17	28563300	*SLC6A4*	NM_001045	TSS1500	chr17:28562387-28563186

Table 2. Illumina HumanMethylation450k BeadChip probes mapping to *SLC6A4*.

Illumina ID	Genome Build	Chromosome	Probe Location	UCSC Gene Name	UCSC Gene Accession Number	UCSC Gene Functional Region	UCSC Associated CGI
cg12484332	37/hg19	5	1443819	*SLC6A3*	NM_001044	5'UTR	chr5:1444259-1444463
cg00997378	37/hg19	5	1444260	*SLC6A3*	NM_001044	5'UTR	chr5:1444259-1444463
cg08713711	37/hg19	5	1444333	*SLC6A3*	NM_001044	5'UTR	chr5:1444259-1444463
cg13723431	37/hg19	5	1444395	*SLC6A3*	NM_001044	5'UTR	chr5:1444259-1444463
cg15212349	37/hg19	5	1444852	*SLC6A3*	NM_001044	5'UTR	chr5:1444678-1446648
cg26205131	37/hg19	5	1444878	*SLC6A3*	NM_001044	5'UTR	chr5:1444678-1446648
cg27348223	37/hg19	5	1445354	*SLC6A3*	NM_001044	5'UTR	chr5:1444678-1446648
cg04598517	37/hg19	5	1445542	*SLC6A3*	NM_001044	5'UTR;1stExon	chr5:1444678-1446648
cg12882697	37/hg19	5	1445549	*SLC6A3*	NM_001044	TSS200	chr5:1444678-1446648
cg04210284	37/hg19	5	1445561	*SLC6A3*	NM_001044	TSS200	chr5:1444678-1446648
cg27037018	37/hg19	5	1445567	*SLC6A3*	NM_001044	TSS200	chr5:1444678-1446648
cg05030481	37/hg19	5	1445593	*SLC6A3*	NM_001044	TSS200	chr5:1444678-1446648
cg17306747	37/hg19	5	1446165	*SLC6A3*	NM_001044	TSS1500	chr5:1444678-1446648
cg14502484	37/hg19	5	1446208	*SLC6A3*	NM_001044	TSS1500	chr5:1444678-1446648
cg16703956	37/hg19	5	1446217	*SLC6A3*	NM_001044	TSS1500	chr5:1444678-1446648
cg13202751	37/hg19	5	1446443	*SLC6A3*	NM_001044	TSS1500	chr5:1444678-1446648
cg16180821	37/hg19	5	1446969	*SLC6A3*	NM_001044	TSS1500	chr5:1444678-1446648

Table 3. Illumina HumanMethylation450k BeadChip probes mapping to *SLC6A3*.

Expression Microarray Analyses. Total RNA was applied to the Illumina HumanHT-12 v4 Expression BeadChip (HT-12)[39], which assesses mRNA expression at > 47,000 transcripts in the human genome. Raw output from the HT-12 microarray (n = 129) was background corrected using the Illumina GenomeStudio Suite. Expression intensity values derived from HT-12 are continuous variables, which were log transformed for normality and variance stability, then filtered to remove probes with low intensity or that exhibited cross-hybridization[34]. Then, PCA was performed on the expression intensity values matrix to identify significant associations between technical artifacts/demographic measures and variability throughout the data. The identified chip technical artifact was then corrected for while preserving the effects of biological sex and lifetime MDD status using ComBAT[37] in R, and implemented in the SVA package[38]. PCA was again performed to confirm chip technical artifact correction. After the aforementioned processing steps, mRNA expression intensity values for probes mapping to *SLC6A4* and *SLC6A3* (Tables 4, 5) were extracted.

Illumina Probe ID	Chromosome	Probe Location	UCSC Gene Name	UCSC Gene Accession Number	Protein_Product
ILMN_1683694	17	25549322-25549371	*SLC6A4*	NM_001045.3	NP_001036.1

Table 4. Illumina HumanHT-12 v4 Expression BeadChip probes mapping to *SLC6A4*.

Illumina Probe ID	Chromosome	Probe Location	UCSC Gene Name	UCSC Gene Accession Number	Protein Product
ILMN_23614	5	1446076-1446125	*SLC6A3*	NM_001044.2	NP_001035.1

Table 5. Illumina HumanHT-12 v4 Expression BeadChip probes mapping to *SLC6A3*.

In addition to demographic, methylation, and expression measures taken for each participant, cell type heterogeneity[40] and ancestry[41] principle components (PCs) were derived for each participant, based on DNA methylation measurements, for inclusion as covariates in downstream analyses. Cell heterogeneity PCs were included in order to correct for cell type specific methylation and expression profiles, whereas ancestry PCs were included to correct for potential population stratification. A total of 104 wave 2 subjects were included in the methylation- and expression-based DNHS analyses.

Statistical Analyses. Main effect regression models were employed to establish the relationship between predictors of interest and outcomes of interest. Utilizing multivariate linear regression, **model one** used CpG methylation of 12 *SLC6A4* 450k probes nearest to the TSS from human blood as predictors of HT-12 mRNA expression of *SLC6A4* (ILMN_1683694) from human blood (Figure 1, Relationship A). Included in the model

as covariates were: participant age at time of blood draw, biological sex, ancestry principle components (PCs 2, 3, and 4), cell type heterogeneity estimates (CD8T, CD4T, NK, Bcell, and monocytes), and antidepressant medication status. These covariates were also included in all regression models to follow. Note, CpG methylation of the 17 *SLC6A3* 450k probes from human blood were not used with this model because the HT-12 mRNA expression probe mapping to *SLC6A3* was filtered out during initial processing steps. Next, utilizing logistic regression, **model two used** HT-12 mRNA expression of *SLC6A4* (ILMN_1683694) from human blood as the predictor of lifetime MDD diagnosis status, retaining the same covariates as the previous model (Figure 1, Relationship B). Again utilizing logic regression, **model three** used CpG methylation of 12 *SLC6A4* 450k probes nearest to the TSS from human blood as the predictor of lifetime MDD diagnosis status, retaining the same covariates as the previous model (Figure 1, Relationship C). Additionally, using the same model design, CpG methylation of 17 *SLC6A3* 450k probes nearest to the TSS from human blood served as the predictor of lifetime MDD diagnosis status, retaining the same covariates as the previous model (Figure 1, Relationship C).

Supplementary analyses of publically available datasets

In order to highlight brain CpG methylation associations with lifetime MDD diagnosis status, supplementary analyses were performed using the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO). A publically available dataset was located in which DNA methylation was profiled for both neuronal and glial nuclei[42]. These profiles were obtained from human post-mortem frontal cortex of "healthy" control subjects (n = 29) and age/sex matched major depression case subjects (n = 29) using the Illumina HumanMethylation450 (450k) Beadchip (GEO Accession No. GSE41826). According to study information, cell types were sorted using fluorescent activated cell sorting (FACS). Subsequently, genomic DNA was isolated from neuronal nuclei, bisulfite converted, fragmented, and hybridized to the 450k Beadchip. Raw output was background corrected using GenomeStudio, and beta values were normalized. 450k beta values, age, sex, race, MDD status, post-mortem interval in hours (PMI), and suicide completion status were extracted from GEO. Afterwards, cross-hybridizing probes and CpG sites near known single nucleotide polymorphisms[34] were filtered out[35]. Then, PCA was performed on the beta value matrix to identify significant associations between technical artifacts/demographic measures and variability throughout the data.

The known chip and position technical artifacts were then corrected. After the aforementioned background correction, normalization, and technical artifact correction, beta values from CpG sites near the *SLC6A4* TSS/CGI (Table 2) were extracted for each participant and compiled. The data were analyzed using the R programming language. Utilizing logic regression, **model four** used CpG methylation of the 12 *SLC6A4* 450k probes nearest to the TSS from human frontal cortex neurons as the predictor of lifetime MDD diagnosis status (Figure 1, Relationship C). Suicide completion status, biological sex, self-reported race, age at time of death, and post-mortem interval were included in this model as covariates. Afterwards, FDR correction using the BH method was employed.

In order to highlight intra-subject brain and blood CpG methylation correlations in humans, supplementary analyses were performed using the Blood Brain DNA Methylation Comparison Tool available at http://epigenetics.iop.kcl.ac.uk/bloodbrain[28]. Specifically, we investigated the CpG methylation of "healthy" control subjects (n=75) in whole blood as well as PFC. Pairwise correlation analysis for the methylation of each Illumina 450k Beadchip probe mapping nearest to the TSS of the *SLC6A4* (Table 2) and *SLC6A3* (Table 3) genes was performed between blood and PFC. Comparison results were not subjected to BH FDR corrections.

Chapter 3: Results

ACTH-Induced Treatment-Resistant Depression Model in the Rat

Slc6a3 Independent samples t-tests of CpG methylation in mPFC tissue showed significant differences between cases and controls at CpG1-9 individually ($p < 0.050$, cases > controls in all tests). The CpG10 independent samples t-tests comparison showed no significant difference (Table 6). BH FDR correction rendered only CpG6 non-significant. In whole blood, independent samples t-tests showed a significant difference between cases and controls at CpG1, however in a direction opposite to that observed in mPFC ($p = 0.041$, control > case) (Table 7). BH FDR correction rendered this comparison non-significant.

	Control Mean ± SEM	Case Mean ± SEM	t-test p value	BH FDR Correction
CpG1	4.137 ± 0.238	5.645 ± 0.273	0.001	sig.
CpG2	7.016 ± 0.802	11.441 ± 1.267	0.014	sig.
CpG3	4.944 ± 0.523	6.762 ± 0.376	0.011	sig.
CpG4	4.976 ± 0.568	8.565 ± 0.977	0.009	sig.
CpG5	2.076 ± 0.309	3.398 ± 0.308	0.009	sig.
CpG6	4.604 ± 0.571	6.845 ± 0.846	0.047	ns.
CpG7	5.427 ± 0.635	8.935 ± 0.944	0.008	sig.
CpG8	4.881 ± 0.419	6.695 ± 0.572	0.029	sig.
CpG9	2.719 ± 0.239	3.640 ± 0.259	0.022	sig.
CpG10	2.919 ± 0.239	3.417 ± 0.201	0.126	ns.
1-10 CpG AVG	4.313 ± 0.205	6.267 ± 0.299	0.0000002	sig.

Table 6. Rat *Slc6a3* mPFC methylation descriptive statistics, independent samples t-test p-values, and BH FDR correction.

	Control Mean ± SEM	Case Mean ± SEM	t-test p value	BH FDR Correction
CpG1	5.401 ± 0.422	4.192 ± 0.359	0.041	ns.
CpG2	8.252 ± 0.970	8.997 ± 1.012	0.604	ns.
CpG3	7.691 ± 0.861	7.662 ± 0.794	0.981	ns.
CpG4	7.102 ± 0.973	6.921 ± 1.188	0.908	ns.
CpG5	3.086 ± 0.365	3.202 ± 0.506	0.864	ns.
CpG6	5.998 ± 0.765	5.458 ± 0.657	0.597	ns.
CpG7	6.946 ± 0.646	7.092 ± 0.885	0.895	ns.
CpG8	6.391 ± 0.460	5.620 ± 0.696	0.366	ns.
CpG9	3.186 ± 0.247	4.434 ± 0.625	0.083	ns.
CpG10	3.421 ± 0.296	3.040 ± 0.284	0.365	ns.
1-10 CpG AVG	5.702 ± 0.265	5.572 ± 0.282	0.738	ns.

Table 7. Rat *Slc6a3* blood methylation descriptive statistics, independent samples t-test p-values, and BH FDR correction.

When comparing average case and control group CpG methylation across all 10 *Slc6a3* CpG sites in mPFC, the case group showed a significantly higher level of CpG methylation compared to the control group ($p = 2.1x10^{-6}$) (Figure 4). This aggregate comparison remained significant after BH FDR correction. When comparing average case and control group CpG methylation across all 10 *Slc6a3* CpG sites in whole blood, no significant difference was detected ($p > 0.05$) (Figure 5).

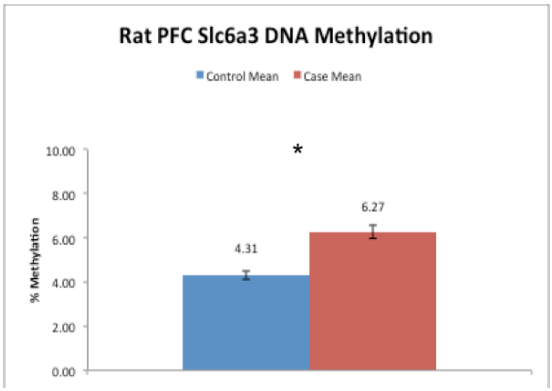

Figure 4. Rat mPFC *Slc6a3* CpG Methylation aggregate comparison of CpG 1-10 (* $p < 0.000005$).

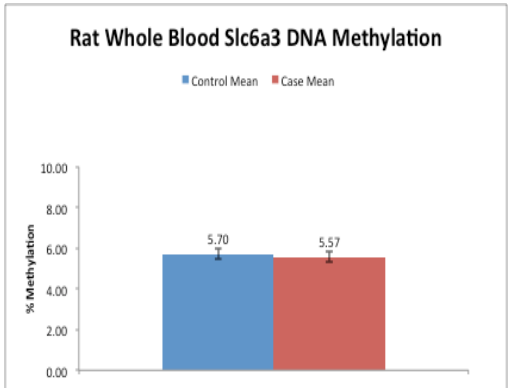

Figure 5. Rat whole blood *Slc6a3* CpG Methylation aggregate comparison of CpG 1-10 ($p > 0.05$).

Slc6a4 mPFC methylation independent samples t-tests showed no significant differences between case and control group means at any measured CpG sites (Table 8). In whole blood, independent samples t-tests showed a significant difference between case and control group means at CpG19 ($p = 0.008$, case > control) (Table 9). However, this comparison did not survive BH FDR correction as significant.

	Control Mean ± SEM	Case Mean ± SEM	t-test p value	BH FDR Correction
CpG12	3.278 ± 0.240	2.818 ± 0.164	0.121	ns.
CpG13	1.556 ± 0.156	1.856 ± 0.207	0.284	ns.
CpG14	1.981 ± 0.227	2.271 ± 0.123	0.253	ns.
CpG15	2.031 ± 0.503	1.916 ± 0.168	0.815	ns.
CpG16	4.537 ± 0.340	4.173 ± 0.220	0.361	ns.
CpG17	3.944 ± 0.294	3.852 ± 0.200	0.792	ns.
CpG18	2.828 ± 0.416	2.833 ± 0.132	0.988	ns.
CpG19	2.958 ± 0.285	2.946 ± 0.283	0.978	ns.
CpG20	2.869 ± 0.220	2.483 ± 0.144	0.143	ns.
CpG21	2.739 ± 0.258	3.022 ± 0.123	0.295	ns.
CpG22	2.503 ± 0.068	2.719 ± 0.126	0.193	ns.
CpG23	2.131 ± 0.449	1.56 ± 0.105	0.141	ns.
12-23 CpG AVG	2.845 ± 0.119	2.759 ± 0.083	0.539	ns.

Table 8. Rat *Slc6a4* mPFC methylation descriptive statistics, independent samples t-test p-values, and BH FDR correction.

	Control Mean ± SEM	Case Mean ± SEM	t-test p value	BH FDR correction
CpG12	3.633 ± 0.3050	3.99 ± 0.3890	0.474	ns.
CpG13	3.888 ± 0.3600	3.369 ± 0.4360	0.370	ns.
CpG14	3.25 ± 0.2170	2.962 ± 0.3170	0.472	ns.
CpG15	2.855 ± 0.4340	2.317 ± 0.2090	0.265	ns.
CpG16	5.67 ± 0.4730	5.53 ± 0.3780	0.820	ns.
CpG17	4.952 ± 0.3950	4.95 ± 0.3180	0.998	ns.
CpG18	3.049 ± 0.1940	3.593 ± 0.4220	0.257	ns.
CpG19	3.092 ± 0.2470	4.403 ± 0.3520	0.008	ns.
CpG20	3.493 ± 0.2080	3.992 ± 0.4150	0.282	ns.
CpG21	3.703 ± 0.1920	3.858 ± 0.3740	0.727	ns.
CpG22	3.17 ± 0.2660	3.898 ± 0.5430	0.241	ns.
CpG23	2.757 ± 0.3120	3.316 ± 0.3780	0.265	ns.
12-23 CpG AVG	3.631 ± 0.1160	3.829 ± 0.1300	0.258	ns.

Table 9. Rat *Slc6a4* blood methylation descriptive statistics, independent samples t-test p-values, and BH FDR correction.

When comparing average case and control group CpG methylation across all 12 *Slc6a4* CpG sites in mPFC (Figure 6) and in whole blood (Figure 7), no significant differences were detected ($p > 0.05$).

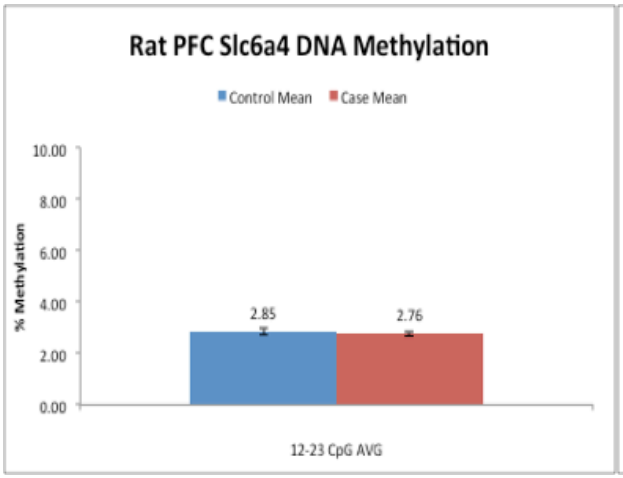

Figure 6. Rat mPFC *Slc6a4* CpG Methylation aggregate comparison of CpG 12-23 ($p > 0.05$).

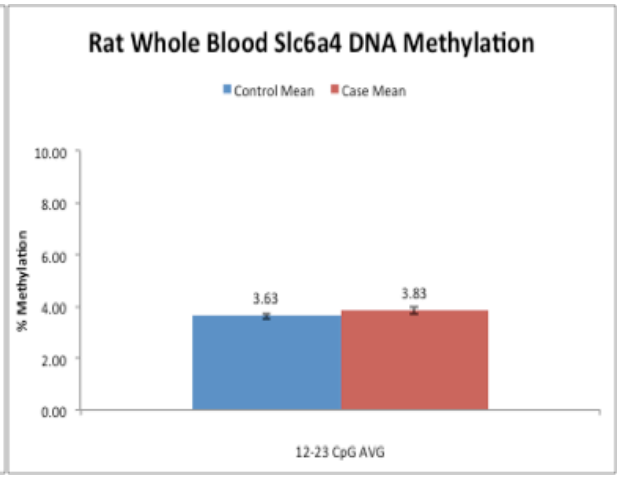

Figure 7. Rat whole blood *Slc6a4* CpG Methylation aggregate comparison of CpG 12-23 ($p > 0.05$).

Pearson correlational analyses between mPFC and whole blood methylation values were performed for each gene on a per CpG site basis, within each case and control group. No significant relationship was observed between mPFC and whole blood CpG methylation values at any site in the *Slc6a4* gene in either the case or control group. In contrast, within *Slc6a3*, significant correlations were observed for CpG methylation

values between mPFC and whole blood in CpG 2 ($p = 0.002$, $r = 0.987$, $df = 3$) and CpG 5 ($p = 0.001$, $r = 0.841$, $df = 8$) in the case group (Table 10); and in the control group for CpG methylation values between mPFC and whole blood in CpG 3 ($p = 0.023$, $r = 0.927$, $df = 3$) (Table 10). Although both significant correlations in the case group survived BH FDR correction, the correlation in the control group did not.

	Case Group mPFC v. Whole Blood	BH FDR Correction	Control Group mPFC v. Whole Blood	BH FDR Correction
CpG1	ns.	ns.	ns.	ns.
CpG2	p = 0.002, r = 0.987, df = 3	sig.	ns.	ns.
CpG3	ns.	ns.	p = 0.023, r = 0.927, df = 3	ns.
CpG4	ns.	ns.	ns.	ns.
CpG5	p = 0.001, r = 0.841, df = 8	sig.	ns.	ns.
CpG6	ns.	ns.	ns.	ns.
CpG7	ns.	ns.	ns.	ns.
CpG8	ns.	ns.	ns.	ns.
CpG9	ns.	ns.	ns.	ns.
CpG10	ns.	ns.	ns.	ns.

Table 10. Rat Slc6a3 mPFC v. whole blood Pearson correlation analysis.

Lifetime Major Depressive Disorder, Peripheral Methylation and Expression in the DNHS

Model one used CpG methylation of the 12 *SLC6A4* 450k probes from human blood as predictor variables and HT-12 expression of *SLC6A4* in human blood as the response variable. Cg05951817 methylation was significantly associated with ILMN_1683694 expression ($p = 0.004$) (Table 11). Additionally, cg10901968 methylation was shown to be significantly associated with ILMN_1683694 expression ($p = 0.025$) (Table 12). However, BH FDR correction attenuated these results to non-significance.

	beta	t value	p value
Intercept	5.771	23.215	<2e-16 *
cg05951817	0.628	2.929	0.004 *
Age	0.000	0.683	0.496
female2	0.022	0.722	0.472
Comp.2	-0.285	-1.091	0.278
Comp.3	-0.749	-1.168	0.245
Comp.4	0.135	0.591	0.556
CD8T	0.184	0.355	0.723
CD4T	0.111	0.318	0.751
NK	0.782	1.644	0.103
Bcell	0.697	1.649	0.102
Mono	-0.334	-0.641	0.523
Antidepressant	0.017	0.454	0.651

Significance Codes: *p < 0.05, #p< 0.10

Table 11. Association between cg05951817 methylation and ILMN_1683694 mRNA expression.

	beta	t value	p value
Intercept	6.065	28.738	<2e-16 *
cg10901968	4.498	2.280	0.025 *
Age	0.001	1.294	0.199
female2	-0.004	-0.167	0.868
Comp.2	-0.335	-1.259	0.211
Comp.3	-1.027	-1.583	0.117
Comp.4	0.000	-0.001	0.999
CD8T	0.367	0.700	0.486
CD4T	0.071	0.200	0.842
NK	0.676	1.397	0.166
Bcell	0.874	2.000	0.048 *
Mono	-0.045	-0.088	0.930
Antidepressant	0.013	0.336	0.738

Significance Codes: *p < 0.05, #p< 0.10

Table 12. Association between cg10901968 methylation and ILMN_1683694 mRNA expression.

Next, **model two** used HT-12 expression of *SLC6A4* in human blood as the predictor, and lifetime MDD diagnosis status as the outcome variable. ILMN_1683694 mRNA expression approached significance as a predictor of lifetime MDD diagnosis status (p = 0.076) (Table 13). In this model, age (p = 0.012) and antidepressant medication status (p = 0.003) were also shown to be significantly associated with lifetime MDD diagnosis status.

	beta	z value	p value
Intercept	-20.148	-1.445	0.149
ILMN_1683694	3.867	1.775	0.076 #
Age	-0.058	-2.518	0.012 *
female2	-0.528	-0.868	0.385
Comp.2	-3.357	-0.638	0.523
Comp.3	19.413	1.429	0.153
Comp.4	0.483	0.102	0.919
CD8T	-15.158	-1.419	0.156
CD4T	-3.810	-0.521	0.602
NK	-9.113	-0.878	0.380
Bcell	5.263	0.621	0.535
Mono	15.321	1.505	0.132
Antidepressant2	-2.224	-2.962	0.003 *

Significance Codes: *p < 0.05, #p< 0.10

Table 13. Association between ILMN_1683694 mRNA expression and lifetime MDD diagnosis status.

Finally, in **model three**, CpG methylation of the 12 *SLC6A4* 450k probes from human blood were used as predictors of lifetime MDD. Cg01330016 methylation was significantly associated with lifetime MDD diagnosis status (p = 0.008) (Table 14). Cg18584905 methylation was also significantly associated with lifetime MDD diagnosis status (p = 0.015) (Table 15). It should be noted, CpG methylation of the 17 *SLC6A3* 450k probes from human blood were used as predictors of the lifetime MDD diagnosis outcome, and these analyses rendered no significant results. BH FDR correction attenuated the two significant results to non-significance.

	beta	z value	p value
Intercept	-61.750	-2.466	0.014 *
cg01330016	68.796	2.624	0.008 *
Age	-0.052	-2.139	0.032 *
female2	-0.334	-0.530	0.596
Comp.2	-7.140	-1.290	0.197
Comp.3	18.352	1.422	0.154
Comp.4	3.750	0.729	0.465
CD8T	-15.401	-1.415	0.157
CD4T	-0.970	-0.133	0.894
NK	-4.646	-0.485	0.628
Bcell	9.579	1.124	0.260
Mono	20.396	1.863	0.063
Antidepressant2	-1.953	-2.729	0.006 *

Significance Codes: *p < 0.05, #p< 0.10

Table 14. Association between cg01330016 methylation and lifetime MDD diagnosis status.

	beta	z value	p value
Intercept	8.759	2.055	0.040 *
cg18584905	-47.773	-2.440	0.015 *
Age	-0.056	-2.395	0.017 *
female2	-0.940	-1.396	0.162
Comp.2	-2.935	-0.544	0.586
Comp.3	18.981	1.420	0.155
Comp.4	0.575	0.121	0.903
CD8T	-8.505	-0.808	0.419
CD4T	1.262	0.167	0.867
NK	-0.576	-0.055	0.956
Bcell	9.861	1.102	0.270
Mono	16.633	1.607	0.108
Antidepressant2	-2.327	-3.090	0.002 *

Significance Codes: *p < 0.05, #p< 0.10

Table 15. Association between cg18584905 methylation and lifetime MDD diagnosis status.

Supplementary analyses of publically available datasets

In **model four,** logistic regression was implemented using the Guintivano 450k neuron CpG methylation dataset. The 12 *SLC6A4* 450k methylation probes were used as predictors of the lifetime MDD diagnosis status response variable (Table 2). FDR uncorrected results from this main effects model are as follows. Cg27569822 methylation was shown to approach nominally significant association with lifetime MDD diagnosis status (p = 0.053). In this model, suicide completion status was shown to be significantly associated with lifetime MDD diagnosis status (p = 0.0001) along with age (p = 0.0353) (Table 16). Importantly, FDR correction of these results using the BH method rendered no remaining significant results.

	beta	z value	p value
Intercept	-17.07	-2.252	0.024 *
cg27569822	174.42	1.936	0.053 #
suicide_status1	4.972	3.858	0.001 *
sex2	1.226	1.269	0.204
race2	-0.182	-0.18	0.857
age	0.063	2.104	0.035 *
pmi	-0.028	-0.394	0.693

Significance Codes: *p <0.05, #p < 0.10

Table 16. Association between cg27569822 frontal cortex neuronal methylation and lifetime MDD diagnosis status

Our final analyses utilized the publically available Blood Brain DNA Methylation Comparison Tool[28] to highlight intra-subject brain and blood CpG methylation correlations. Using this tool, the blood derived CpG methylation value of each probe (Tables 2, 3) was compared to the CpG methylation values of that same probe, derived from PFC. Out of the 12 450k probes mapping to the *SLC6A4* genic region, 2 probes' methylation values were shown to be significantly correlated between blood and PFC. Cg14692377 blood and PFC exhibited the strongest positive correlation amongst any comparison of *SLC6A4* between tissues (p = 0.001, r = 0.367, df > 71) (Table 17). Out of the 17 450k probes mapping to the *SLC6A3* genic region, 5 probes' methylation values were significantly correlated between blood and PFC. The strongest correlation was that of cg16180821 (p = 0.004, r = 0.329, df > 71). (Table 18).

Illumina 450k Probe	Blood v. PFC
cg14692377	p = 0.001, r = 0.367
cg05016953	p = 0.047, r =0.232

Table 17. Hannon et al., 2015 Human Blood Brain CpG Methylation Correlation Analysis (*SLC6A4* probes, df > 71). Significant intra-subject correlations between blood and PFC tissues.

Illumina 450k Probe	Blood v. PFC
cg08713711	p = 0.02, r = 0.262
cg27348223	p = 0.045, r = 0.234
cg17306747	p = 0.009, r = 0.301
cg13202751	p = 0.040, r = 0.239
cg16180821	p = 0.004, r = 0.329

Table 18. Hannon et al., 2015 Human Blood Brain CpG Methylation Correlation Analysis (*SLC6A3* probes, df > 71). Significant intra-subject correlations between blood and PFC tissues.

Chapter 4: Discussion

Identification of salient peripheral biomarkers of TRD and MDD is critically important due to the great burden these disorders carry, and because of the inability to access etiological CNS tissues in living humans. Such biomarkers can help to elucidate the neurobiological underpinnings of treatment resistance in certain patients, as well as provide a biological indicator of disease state rather than a subjective symptom indicator as frequently relied upon in clinical settings. Methylation of a CpG dinucleotide is regarded as a stable epigenetically induced mark, which is often associated with an increase or decrease in transcription of associated mRNA transcripts[42] and for this reason is the main focus of the current study.

Using an ACTH-induced TRD model in the rat, we identified differentially methylated CpG sites between our case and control groups, partially confirming the initially posed hypothesis that ACTH-exposed animals would exhibit significantly differential CpG methylation in the regions near the TSS of *Slc6a3* and *Slc6a4* genes in mPFC tissue and whole blood, compared to control group animals. Of particular interest within *Slc6a3*, CpG1 exhibited a significant difference between cases and controls in mPFC, and importantly in whole blood as well. However, in mPFC tissue, the case group had a higher mean methylation value than the control group, whereas the opposite direction of effect was observed in blood. *Slc6a4* analyses also indicated a significant difference between case and control groups at CpG19 in blood, where the case group exhibited higher methylation values than the control group. Furthermore, *Slc6a3* exhibited significant positive correlations between methylation levels detected in brain and blood, within both the case and control groups. Interestingly, *Slc6a4* exhibited no significant correlations between tissues within either the case group of the control group. Results from correlation analyses provided a modest degree of evidence towards confirmation of the rat intra-subject blood brain correlation hypothesis. .

Taken together, results from our ACTH-induced TRD model in the rat indicate that CpG methylation alterations in the promoter region of *Slc6a3* occurring in the brain and blood are a strong candidate for functional relevance to the pathology of TRD. Although we were unable to test this in the current study, a mechanism by which CpG methylation is related to the TRD phenotype likely involves alteration to mRNA expression of the downstream *Slc6a3* gene. Based on current genomic understanding, an increase in CpG methylation of the promoter region is likely to be associated with decreased mRNA expression due to steric

exclusion of polymerases and transcription factors[22]. This decrease in mRNA expression of the *Slc6a3* transcript could then result in a decrease of its corresponding protein, the dopamine reuptake transporter. With less of this protein available for embedding into the pre-synaptic neuron's terminal membrane, more DA would remain in the synaptic cleft for longer periods of time, increasing its utilization by the post-synaptic neuron. However, accounting for the potential biological effect of chronic ACTH exposure and *Slc6a3* CpG methylation in both the CNS and periphery is more challenging, and currently the role of *Slc6a3* mRNA expression in the blood is not well characterized. Nevertheless, our findings suggest that CpG1 of *Slc6a3* in the blood represents the most salient peripheral biomarker of TRD, due to its corresponding association with TRD detected in mPFC as well (albeit in opposite directions). Additionally, results indicate that CpG methylation alterations occurring in the first exon of *Slc6a4* in the blood also represent a potential peripheral biomarker of TRD in the current model, although evidence for this conclusion is less robust because CpG19 is not significant in mPFC. The juxtaposition of results pertaining to both *Slc6a3* and *Slc6a4* in the ACTH-induced rat model of TRD are of great significance. It is not surprising to see no alteration of *Slc6a4* methylation in the mPFC in TRD, as one of the most common methods of treatment for initial MDD is by way of SSRI's. Ineffective treatment with SSRI's implies the targeted system may not be aberrantly regulated, as potentially in the case of the ACTH-induced rat model of TRD. On the other hand, the ineffective initial treatment of SSRI's logically suggests alternative disrupted pathologies, and evidence from the current study points to the aberrant regulation of *Slc6a3*.

After investigating the relationship between peripheral and CNS tissue methylation in the rodent model of TRD to index CNS differential methylation with peripheral methylation, peripheral methylation and expression measures were investigated in human blood in an attempt to characterize potential peripheral biomarkers characteristic of MDD, and to potentially predict what is occurring in the brain. Using peripheral genome-wide methylation and expression data derived from the DNHS to investigate lifetime MDD in humans, the current study highlighted that two 450k probes mapping near the TSS of *SLC6A4* (cg05951817 and cg10901968) were significantly associated with mRNA expression of the *SLC6A4* transcript (ILMN_1683694). Additionally, the *SLC6A4* mRNA expression probe approached a significant association with lifetime MDD, and two 450k probes (cg01330016 and cg 18584905) were significantly associated with

lifetime MDD. Hypotheses pertaining to this portion of the current study predicted significant associations between: *SLC6A4* 450k probe methylation and *SLC6A4* HT12 probe mRNA expression, *SLC6A4* HT12 probe mRNA expression and lifetime MDD diagnosis, and *SLC6A4* 450k probe methylation and lifetime MDD diagnosis. Evidence exists in favor of all three hypotheses, and provides a potentially mechanistic explanation to the model presented in Figure 1. Cg05951817 is located in the first intron of the gene, whereas cg10901968 is located in the promoter region of the gene about 70bp upstream of the TSS in the same CGI (Figure 2). Previous work from our lab showed that a differentially methylated intronic *SLC6A4* CpG site modified the association between number of traumatic events and PTSD diagnosis within the DNHS cohort[43]. Additionally, differential intronic CpG methylation at the cg22584138 locus has previously been associated with silencing of mRNA transcription[44], and the role of non-coding intronic DNA is becoming more salient in regards to transcriptional regulation[45]. On the other hand, cg10901968 has been shown to be a significant predictor of depressive symptoms in African American women[46], and differential blood-derived CpG methylation of sites in this region have been significantly associated with in vivo measures of human brain 5HT synthesis[47], pointing to potential functional significance of both CpG sites and their respective methylation values in humans. Furthermore, our nominally significant association between mRNA expression of blood *SLC6A4* transcript and lifetime MDD status provides evidence for the potentially functional effect of differential methylation of CpG sites within and near the *SLC6A4* gene. Transcription of this gene is regulated by both genetic and epigenetic factors, and CpG methylation is well documented to be one of them. It is therefore possible that differential CpG methylation of *SLC6A4* in the blood is related to differential *SLC6A4* mRNA expression in the blood, and that these factors together are related to indicators of pathology of lifetime MDD. What further strengthens this argument are the significant associations between cg01330016 and cg18584905 to lifetime MDD status. These CpG sites are located in intron one and the promoter region of *SLC6A4* respectively, the same functional regions as the CpG mRNA expression associations (Figure 2).

Overall, the second phase of the current study showed that in peripheral blood, CpG methylation and mRNA expression of *SLC6A3* are not relevant as indicators of lifetime MDD pathology moving forward. On the other hand, it appears that *SLC6A4* CpG methylation of the first intron and the promoter regions are of functional significance in regards to serving as an indicator of the disorder. The CpG sites are most likely sites

of transcriptional regulation, with the functional differences in indicators of pathology stemming directly from differential mRNA expression. Further inquiry is necessary. Note, no Illumina CpG methylation 450k probes near the TSS of *SLC6A3* were significantly associated with lifetime MDD diagnosis, and the mRNA expression probe mapping to *SLC6A3* was removed during quality control and pre-processing. Therefore, results pertaining to *SLC6A3* indicate that although implicated in the neurological etiology of MDD, *SLC6A3* CpG methylation near the promoter region potentially plays no role as a biomarker of MDD as measured in the blood.

The third portion of the current study hypothesized that frontal cortex neuronal CpG methylation of probes mapping to *SLC6A4* would be significantly associated with lifetime MDD diagnosis status. Using a supplementary and publically available neuronal genome-wide methylation data set, one inter-subject CpG site methylation value was shown to approach significant association with lifetime MDD diagnosis status, and it did not survive FDR BH correction. This can be potentially attributed to a lack of comprehensive covariates included in regression modeling as well as small sample size. This null result lends no evidence to the initial hypothesis.

The final portion of the current study hypothesized that CpG methylation of probes mapping to the human *SLC6A3* and *SLC6A4* gene would show significant correlations within subjects, between blood and PFC tissue. Using the publically available Blood Brain DNA Methylation Comparison Tool, 2 out of 12 *SLC6A4* intra-subject blood versus PFC correlation comparisons were shown to reach significance, and 5 out of 17 *SLC6A3* intra-subject blood versus PFC correlation comparisons were shown to reach significance. All significant relationships were positive. The moderate quantity of significant correlation indicates that within subjects, brain and blood methylation levels are potentially related in mostly "healthy" subjects. This finding, along with the empirical findings from the research originating the DNA Methylation Comparison Tool[28], lends moderate support to the idea that in some cases peripheral CpG methylation can be used as surrogate for CNS CpG methylation.

As in any study, limitations are present. However, these limitations present opportunities for improvement in regards to future directions. The first limitation of the ACTH-induced TRD animal model study is the use of an exclusively male sample of animals in the ACTH rodent study. Notable differences

between sexes in regards to TRD, MDD, and antidepressant efficacy have been widely reported[48,49]. In the future, a cohort of animals comprised of both male and female rodents would be ideal. The next limitation of mention is the lack of mRNA expression data for the candidate genes of interest in this study. Although CpG methylation has been shown to functionally alter mRNA expression in some scenarios, this cannot be assumed for any given CpG site without empirical validation. Additionally, mRNA expression data would provide further insight as to the pathology of ACTH-induced TRD. In this same line of thought, expression of other types of RNA (e.g. long non-coding RNA, micro RNA) could also be investigated in the future, as some have been shown to function in regulatory roles[50]. The next limitation to mention is that *Slc6a3* and *Slc6a4* were selected as candidate genes for the current study due to previous empirical findings and based on previous research. However, current trends in the field suggest the importance of polygenic investigation at many levels of biology. Ongoing international consortia are conducting genome-wide investigations of MDD that integrate DNA methylation, mRNA expression, and SNP data to further characterize the genomic basis of this disorder.

Regarding the second portion of our study, one limitation of the DNHS study is the lack of validation of genome-wide methylation and expression data from the 450k and HT-12 platforms. Although these platforms are widely accepted and well regarded in the field, validation through bisulfite pyrosequencing and RT-PCR would strengthen our preliminary conclusions. One future direction is of special interest to the current study. Inclusion of human *SLC6A4* 5HTTLPR genotype would potentially strengthen the fit of the regression models that were employed, as genotype of this locus has been shown to significantly affect expression of the *SLC6A4* mRNA transcript, and to be significantly associated with MDD diagnosis status[21]. Inclusion of these factors may increase the amount of variability explained by the model and, in doing so, help to increase understanding of the pathology involved in this disorder. Next, the neuronal Illumina 450k dataset utilized in the current study[42] included methylation values, age, sex, race, PMI, lifetime MDD diagnosis status, and suicide completion status from frontal cortex tissue of 29 "healthy" control subjects and 29 lifetime MDD diagnosis affirmative case subjects. Although the aforementioned covariates were included in the model, many potential covariates were not available for inclusion. Stratified ancestry principle components, antidepressant information, and comorbid disorder information would have strengthened the model, and could have explained more of the variability. Finally, although our supplementary analyses of human data using the Blood Brain

DNA Methylation Comparison Tool offered some insight as to the relationship between CpG methylation in blood versus PFC within a given subject, the ability to control for demographic and technical artifacts is not possible within the context of the publically available tool.

In conclusion, the current study has preliminarily identified blood-derived, DNA-methylation based peripheral biomarkers of TRD in *Slc6a3* at CpG1, and in *Slc6a4* at CpG19. In humans, methylation of two CpG sites in the first intron and promoter region of *SLC6A4* have also been identified preliminarily as peripheral biomarkers of MDD. Additionally, *SLC6A4* mRNA expression in blood has also been identified preliminarily as a peripheral biomarker of MDD in humans. Furthermore, significant correlation analyses in both rat and human showed moderate concordance between blood and brain CpG methylation levels at select CpG sites in our two loci of interest. These potential peripheral biomarkers, identified in easily accessible tissues, offer the opportunity for biologically based diagnosis efforts of TRD or MDD, and give insight as to the potential underlying mechanisms of these disorders. However, efforts to elucidate additional peripheral biomarkers should continue to expand, as polygenic factors at multiple levels of physiology are likely more explanatory than candidate gene based research.

References

1. Stetler, C., & Miller, G. E. (2011). Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosomatic medicine*, *73*(2), 114-126.
2. American Psychiatric Association. (2013). *DSM 5*. American Psychiatric Association.
3. Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., ... & Wang, P. S. (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Jama*, *289*(23), 3095-3105.
4. Melartin, T. K., Rytsala, H. J., Leskela, U. S., Lestela-Mielonen, P. S., Sokero, T. P., & Isometsa, E. T. (2002). Current comorbidity of psychiatric disorders among DSM-IV major depressive disorder patients in psychiatric care in the Vantaa Depression Study. *Journal of Clinical Psychiatry*, *63*(2), 126-134.
5. Lesage, A. D., Boyer, R., Grunberg, F., Vanier, C., Morissette, R., & Loyer, M. (1994). Suicide and mental disorders: a case-control study of young men. *The American journal of psychiatry*, *151*(7), 1063.
6. Üstün, T. B., Ayuso-Mateos, J. L., Chatterji, S., Mathers, C., & Murray, C. J. (2004). Global burden of depressive disorders in the year 2000. *The British journal of psychiatry*, *184*(5), 386-392.
7. Frank, E., & Karp, J. (1993). Rush AJ. *Efficacy of treatments for major depression. Psychopharmacol Bull*, *29*, 457-75.
8. Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Nierenberg, A. A., Warden, D., Ritz, L., ... & Shores-Wilson, K. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR* D: implications for clinical practice. *American journal of Psychiatry*, *163*(1), 28-40.
9. Keller, M. B., Shapiro, R. W., Lavori, P. W., & Wolfe, N. (1982). Recovery in major depressive disorder: analysis with the life table and regression models. *Archives of general psychiatry*, *39*(8), 905-910.
10. Mathew, S. J. (2008). Treatment-resistant depression: recent developments and future directions. *Depression and anxiety*, *25*(12), 989-992.
11. Nemeroff, C. B. (2007). Prevalence and management of treatment-resistant depression. *Journal of Clinical Psychiatry*, *68*(8), 17.
12. Pariante, C. M., & Lightman, S. L. (2008). The HPA axis in major depression: classical theories and new developments. *Trends in neurosciences*, *31*(9), 464-468.

13. Nomikos GC, Damsma G, and Wenkstern D. et al. Acute effects of bupropion on extracellular dopamine concentrations in rat striatum and nucleus accumbens studied by in vivo microdialysis. Neuropsychopharmacology. 1989 2:273–279.
14. Li SX, Perry KW, Wong DT. Influence of fluoxetine on the ability of bupropion to modulate extracellular dopamine and norepinephrine concentrations in three mesocorticolimbic areas of the rat. Neuropharmacology. 2002;42:181–190.
15. Kitamura, Y., Araki, H., & Gomita, Y. (2002). Influence of ACTH on the effects of imipramine, desipramine and lithium on duration of immobility of rats in the forced swim test. *Pharmacology Biochemistry and Behavior*, *71*(1), 63-69.
16. Walker, A. J., Burnett, S. A., Hasebe, K., McGillivray, J. A., Gray, L. J., McGee, S. L., ... & Tye, S. J. (2013). Chronic adrenocorticotrophic hormone treatment alters tricyclic antidepressant efficacy and prefrontal monoamine tissue levels. *Behavioural brain research*, *242*, 76-83.
17. Kelly, W. F., Checkley, S. A., & Bender, D. A. (1980). Cushing's syndrome, tryptophan and depression. *The British Journal of Psychiatry*, *136*(2), 125-132.
18. Sonino, N., Boscaro, M., Ambroso, G., Merola, G., & Mantero, F. (1986). Prolonged treatment of Cushing's Disease with metyrapone and aminoglutethimide. IRCS Medical Science-Biochemistry, *14*(5), 485-486.
19. Dong, C., Wong, M. L., & Licinio, J. (2009). Sequence variations of ABCB1, SLC6A2, SLC6A3, SLC6A4, CREB1, CRHR1 and NTRK2: association with major depression and antidepressant response in Mexican-Americans. *Molecular psychiatry*, *14*(12), 1105-1118.
20. Michelhaugh SK, Fiskerstrand C, Lovejoy E, Bannon MJ, Quinn JP. The dopamine transporter gene (SLC6A3) variable number of tandem repeats domain enhances transcription in dopamine neurons. *J Neurochem*. 2001;**79**:1033–1038.
21. Philibert, R. A., Sandhu, H., Hollenbeck, N., Gunter, T., Adams, W., & Madan, A. (2008). The relationship of 5HTT (SLC6A4) methylation and genotype on mRNA expression and liability to major depression and alcohol dependence in subjects from the Iowa Adoption Studies. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *147*(5), 543-549.

22. Tate, P. H., & Bird, A. P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Current opinion in genetics & development*, *3*(2), 226-231.

23. Uddin, M. (2014). Blood-based biomarkers in depression: emerging themes in clinical research. *Molecular diagnosis & therapy*, *18*(5), 469-482.

24. Lopresti, A. L., Maker, G. L., Hood, S. D., & Drummond, P. D. (2014). A review of peripheral biomarkers in major depression: the potential of inflammatory and oxidative stress biomarkers. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *48*, 102-111.

25. Nestler, E. J., & Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. *Nature neuroscience*, *13*(10), 1161-1169.

26. Uddin, M., Aiello, A. E., Wildman, D. E., Koenen, K. C., Pawelec, G., de Los Santos, R., ... & Galea, S. (2010). Epigenetic and immune function profiles associated with posttraumatic stress disorder. *Proceedings of the National Academy of Sciences*, *107*(20), 9470-9475.

27. Goldmann, E., Aiello, A., Uddin, M., Delva, J., Koenen, K., Gant, L. M., & Galea, S. (2011). Pervasive exposure to violence and posttraumatic stress disorder in a predominantly African American Urban Community: the Detroit Neighborhood Health Study. *Journal of traumatic stress*, *24*(6), 747-751.

28. Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, *10*(11), 1024-1032.

29. Paternain, L., Batlle, M. A., De la Garza, A. L., Milagro, F. I., Martinez, J. A., & Campion, J. (2012). Transcriptomic and epigenetic changes in the hypothalamus are involved in an increased susceptibility to a high-fat-sucrose diet in prenatally stressed female rats. *Neuroendocrinology*, *96*(3), 249-260.

30. Siuda, D., Wu, Z., Chen, Y., Guo, L., Linke, M., Zechner, U., ... & Li, H. (2014). Social isolation-induced epigenetic changes in midbrain of adult mice. *J. Physiol. Pharmacol*, *65*, 247-255.

31. Tost, J., & Gut, I. G. (2007). DNA methylation analysis by pyrosequencing. *Nature protocols*, *2*(9), 2265-2275.

32. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289-300.

33. Barfield, R. T., Kilaru, V., Smith, A. K., & Conneely, K. N. (2012). CpGassoc: an R function for analysis of DNA methylation microarray data. *Bioinformatics*, *28*(9), 1280-1281.

34. Chen, Y. A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., ... & Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*, *8*(2), 203-209.

35. Pidsley, R., Wong, C. C., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC genomics*, *14*(1), 1.

36. Teschendorff, A. E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., & Beck, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*, *29*(2), 189-196.

37. Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, *8*(1), 118-127.

38. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., Storey, J.D., 2012. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28(6), 882-883.

39. Logue, M. W., Amstadter, A. B., Baker, D. G., Duncan, L., Koenen, K. C., Liberzon, I., ... & Smith, A. K. (2015). The Psychiatric Genomics Consortium Posttraumatic Stress Disorder Workgroup: posttraumatic stress disorder enters the age of large-scale genomic collaboration. *Neuropsychopharmacology*, *40*(10), 2287-2297.

40. Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*(10), 1363-1369.

41. Barfield, R. T., Almli, L. M., Kilaru, V., Smith, A. K., Mercer, K. B., Duncan, R., ... & Ressler, K. J. (2014). Accounting for population stratification in DNA methylation studies. *Genetic epidemiology*, *38*(3), 231-241.

42. Guintivano, J., Aryee, M. J., & Kaminsky, Z. A. (2013). A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics*, *8*(3), 290-302.

43. Sutherland, J. E., & Costa, M. A. X. (2003). Epigenetics and the environment. *Annals of the New York Academy of Sciences*, *983*(1), 151-160.

44. Koenen, K. C., Uddin, M., Chang, S. C., Aiello, A. E., Wildman, D. E., Goldmann, E., & Galea, S. (2011). SLC6A4 methylation modifies the effect of the number of traumatic events on risk for posttraumatic stress disorder. *Depression and anxiety*, *28*(8), 639-647.

45. Zhang, X., Wu, M., Xiao, H., Lee, M. T., Levin, L., Leung, Y. K., & Ho, S. M. (2010). Methylation of a single intronic CpG mediates expression silencing of the PMP24 gene in prostate cancer. *The Prostate*, *70*(7), 765-776.

46. Mercer, T. R., Dinger, M. E., & Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews Genetics*, *10*(3), 155-159.

47. Lei, M. K., Beach, S. R., Simons, R. L., & Philibert, R. A. (2015). Neighborhood crime and depressive symptoms among African American women: Genetic moderation and epigenetic mediation of effects. *Social Science & Medicine*, *146*, 120-128.

48. Wang, D., Szyf, M., Benkelfat, C., Provençal, N., Turecki, G., Caramaschi, D., ... & Booij, L. (2012). Peripheral SLC6A4 DNA methylation is associated with in vivo measures of human brain serotonin synthesis and childhood physical aggression. *PloS one*, *7*(6), e39501.

49. Dalla, C., Pitychoutis, P. M., Kokras, N., & Papadopoulou‐Daifoti, Z. (2010). Sex differences in animal models of depression and antidepressant response. *Basic & clinical pharmacology & toxicology*, *106*(3), 226-233.

50. Marsh, W. K., & Deligiannidis, K. M. (2010). Sex-related differences in antidepressant response: When to adjust treatment. *Current Psychiatry*, *9*(5), 25-30.

51. Smalheiser, N. R., Lugli, G., Rizavi, H. S., Torvik, V. I., Turecki, G., & Dwivedi, Y. (2012). MicroRNA expression is down-regulated and reorganized in prefrontal cortex of depressed suicide subjects. *PloS one*, *7*(3), e33201.