

ARTICLE

Open Access

Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: A Monozygotic Discordant Twin Study

Yun Zhu¹, Eric Strachan², Emily Fowler², Tamara Bacus³, Peter Roy-Byrne² and Jinying Zhao¹

Abstract

DNA methylation plays an important role in major depressive disorder (MDD), but the specific genes and genomic regions associated with MDD remain largely unknown. Here we conducted genome-wide profiling of DNA methylation (Infinium MethylationEPIC BeadChip) and gene expression (RNA-seq) in peripheral blood monocytes from 79 monozygotic twin pairs (mean age 38.2 ± 15.6 years) discordant on lifetime history of MDD to identify differentially methylated regions (DMRs) and differentially expressed genes (DEGs) associated with MDD, followed by replication in brain tissue samples. Integrative DNA methylome and transcriptome analysis and network analysis was performed to identify potential functional epigenetic determinants for MDD. We identified 39 DMRs and 30 DEGs associated with lifetime history of MDD. Some genes were replicated in postmortem brain tissue. Integrative DNA methylome and transcriptome analysis revealed both negative and positive correlations between DNA methylation and gene expression, but the correlation pattern varies greatly by genomic locations. Network analysis revealed distinct gene modules enriched in signaling pathways related to stress responses, neuron apoptosis, insulin receptor signaling, mTOR signaling, and nerve growth factor receptor signaling, suggesting potential functional relevance to MDD. These results demonstrated that altered DNA methylation and gene expression in peripheral blood monocytes are associated with MDD. Our results highlight the utility of using peripheral blood epigenetic markers and demonstrate that a monozygotic discordant co-twin control design can aid in the discovery of novel genes associated with MDD. If validated, the newly identified genes may serve as novel biomarkers or druggable targets for MDD and related disorders.

Background

Major depressive disorder (MDD) affects over 350 million people worldwide and is projected to be the leading cause of disease burden by 2030¹. Although the etiology of MDD involves genetic and environmental factors², the precise underlying mechanisms remain poorly understood. A growing body of evidence has

implicated a role for DNA methylation in MDD^{3,4}, but the specific genes and biological pathways underlying MDD still remain unclear.

DNA methylation is influenced by genetic⁵, early life environment, and behavioral factors⁶, and is tissue and cell-type specific⁷. Thus, establishing the relationship between DNA methylation and mental illnesses, such as MDD requires the control of these potential confounding variables. Monozygotic (MZ) discordant twin pairs provide a powerful tool to examine the role of epigenetic mechanisms in depression, because they are matched on genotype, age, and sex. Moreover, identical twins reared together share early life familial environment⁸, which may

Correspondence: Jinying Zhao (jzhao66@ufl.edu)

¹Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, Gainesville, FL, USA

²Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA

Full list of author information is available at the end of the article.

These authors contributed equally: Yun Zhu, Eric Strachan

© The Author(s) 2019



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

contribute to the risk of depression later in life. At the time of this writing, several epigenome-wide association studies (EWAS) have been conducted to identify differentially methylated (DM) genes/regions associated with MDD using either unrelated individuals⁴ or twin pairs^{9–12}. However, almost all previous studies utilized heterogeneous cell types from blood, buccal, or post-mortem brain tissue samples. As different cell types have different DNA methylation profiles⁷, the use of homogeneous cell types such as purified monocytes should minimize confounding by cellular heterogeneity in epigenetic research. Moreover, because monocytes are key innate immune cells involved in inflammation, a mechanism known to be implicated in MDD¹³, identifying methylation changes in circulating monocytes is likely to provide mechanistic insight into disease pathogenesis. Further, the potential functional consequences of epigenetically altered genes on gene expression have not been adequately evaluated in previous studies.

Here we report findings from an integrated genome-wide profiling of DNA methylome and transcriptome in purified blood monocytes from 79 MZ twin pairs discordant for lifetime history of MDD, followed by replication in brain tissue and network analysis. Our goal is to identify key candidate genes and novel pathways associated with MDD.

Methods

Twin pairs

The current analysis included 79 MZ twin pairs discordant on lifetime history of MDD. Twins included in the current analysis were enrolled through the Mood and Methylation Study (MMS), an observational study designed to understand epigenetic and other molecular mechanisms underlying MDD using a co-twin control design. Detailed information regarding the study design, twin recruitment, clinical examination, and sample collection for the MMS has been described elsewhere¹⁴. Briefly, all twins enrolled in the MMS were members of the Washington State Twin Registry (WSTR), a community-based twin registry consisting of over 10,000 twin pairs¹⁵. Detailed methods for the construction of the registry and enrollment of twin pairs have been described previously¹⁵. All twin participants provided informed consent. Zygosity of the twin pairs included in the present study was confirmed using the 59 polymorphic SNPs included in the HumanMethylationEPIC array. Power analysis showed that we should have 80% power to detect a 1.1-fold change (FC) in DNA methylation at a genome-wide significance level¹⁶.

MDD diagnosis

To identify twin pairs for the present study, we sent an introductory letter to WSTR members via postal mail and

email that asked general questions about MDD. These letters were only sent to twins where one or both members were in Western Washington because of the need for an in-person study visit in Seattle. Interested twins were then interviewed by trained research staff using approved scripts to screen for likelihood of eligibility. After pre-screening, lifetime and current MDD diagnoses were determined using the Structured Clinical Interview for DSM-IV Research Version (SCID-4-RV). Interviews were administered via phone by a clinical psychologist (E.D.S.), who was blind to pre-screening clinical information about both the interviewee and his or her co-twin, and final diagnosis was confirmed by consulting the senior psychiatrist (P.P.R.B.). The final sample was drawn from a total of 693 clinical interviews and all eligible and interested MZ pairs who could complete the in-person visit prior to the end of the enrollment period were enrolled. A discordant pair was defined as a twin pair in which one twin met the criteria for lifetime history of MDD, and his/her co-twin did not.

Inclusion/exclusion criteria

Only complete twin pairs were eligible for the present study. As over 85% of the twins in the WSTR are Caucasian (which reflects Washington State generally), all twin pairs included in the current analysis are Caucasian. Inclusion criteria included: (1) monozygosity determined by DNA analysis; (2) pairwise discordance on lifetime history of MDD; (3) aged 18 or older; (4) reared together; (5) Beck Depression Inventory II (BDI-II)¹⁷ score < 13 for the non-depressed twin (0–13 is the range for 'minimal symptoms' recommended by the BDI-II authors); and (6) willingness to provide blood samples. The primary exclusionary conditions included schizophrenia or other psychotic disorder, bipolar disorder, current substance use disorder, cancer within the past 5 years, autoimmune disorders, uncontrolled endocrine disorders, and uncontrolled sleep apnea.

Other measures

Each twin was also asked to complete standard questionnaires regarding sociodemographic factors, lifestyle, early life experience, use of psychiatric medications, and disease history. Severity of current depressive symptoms was assessed by the BDI-II and the Quick Inventory of Depressive Symptomatology (QIDS-SR-16) at the time of the blood draw¹⁸. Self-reported early life stress was measured using the Adverse Childhood Experience (ACE)¹⁹ and Early Trauma Inventory (ETI)²⁰ questionnaires. PTSD was not exclusionary in this study due to high comorbidity with MDD, but was diagnosed and recorded during the SCID-RV interviews. Participants reported their use of medications commonly prescribed for MDD along with the approximate duration of use (ranging from 2 weeks to >10

years). They were also asked to list any additional psychiatric medications they had taken. For the current analyses, medications were first categorized into types (i.e., antidepressants, benzodiazepines, mood-stabilizers, and other (i.e. Ritalin)), and then translated into a “yes/no” variable for each of these categories for each participant. A participant was defined as having a history of antidepressant use if he/she used any of the medications.

Monocyte isolation and DNA/RNA extraction

Monocytes were isolated from fresh peripheral blood (collected into EDTA vacutainer tubes) using the Monocyte Isolation Kit II from Miltenyi Biotec (Auburn, CA, USA). DNA/RNA was isolated from monocytes using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, CA, USA) according to manufacturer’s instructions. Double strand DNA and total RNA was quantified using PicoGreen Fluorometry. RNA integrity was assessed by capillary electrophoresis (e.g., Agilent BioAnalyzer 2100). Except for three samples with RNA integrity number (RIN) < 8 (one sample 7.0, one sample 7.6 and one sample 7.9), all other samples have RIN > 8. Of these, 88% samples have RIN > 9 (Table S1).

DNA methylation profiling

The Infinium HumanMethylationEPIC BeadChip (Illumina Inc., CA, USA)²¹ was used for DNA methylation profiling. Genomic DNA (500 ng) was bisulfite converted using the EZ-96 DNA methylation kit (Zymo Research, CA). Modified DNA was then amplified, fragmented, and hybridized, followed by fluorescent staining and scanning on a HiScan scanner (Illumina Inc.). The methylation level of each probe was represented as a β -value ranging from 0 (unmethylated) to 1 (fully methylated). To ensure accuracy of the methylation assay, each batch included two CEPH DNA samples (Coriell Institute, NJ) as positive controls. Twin pairs were hybridized on the same chip to minimize batch effect.

Methylation data pre-processing and QC

We first removed the following: (1) probes with a detection $p > 0.05$ in more than 20% of the samples; (2) probes with raw signal intensities greater or less than three SD from the mean; and (3) probes located on sex chromosomes (both X- and Y-chromosomes). Probes passing initial QC were then annotated to the UCSC (GRCh37/hg19), and those mapped to multiple locations or overlapping with known SNPs were further removed. The final analyses included 813,382 autosomal probes. All samples passed QC procedures. Prior to analysis, DNA methylation data were normalized with functional normalization using the R package *minfi*²². This method corrects for bias of different probe types and batch effects via an unsupervised approach²².

Transcriptome profiling by RNA-seq

Monocyte gene expression was quantified by paired end RNA-seq (50 PE). Using sequence-specific Ribozero capture probes, total RNA (500 ng) was globin-depleted and rRNA-depleted, fragmented, and reverse transcribed. TruSeq libraries were quantified using PicoGreen Fluorimetry and sequenced on HiSeq 2500. Sequencing reads were aligned to the UCSC (GRCh37/hg19) by bowtie2²³. Gene expression level measured as transcription per million reads (TPM) was quantified by RSEM v 1.12²⁴. Genes with expression level below an appreciable level ($\log_2(\text{TPM} + 1) < 1$) in more than 5% samples were discarded. A total of 10,329 genes in all samples was included in the statistical analysis. At least 20 million reads per sample were captured in RNA-seq analysis (Table S1).

Replication in the brain

To replicate the putative DMRs identified in blood monocytes, we downloaded brain DNA methylation data (HumanMethylation450K BeadChip) through the publicly available GEO database (GSE41826). The brain tissue was collected by the NICHD Brain Bank of Developmental Disorders²⁵, and DNA methylation data were generated in sorted neuronal nuclei from 58 postmortem brain tissue, including 29 MDD patients and 29 controls (51.7% female, mean age 32.6 ± 16.0 years old). Demographic information of the brain donors was described previously²⁵.

To replicate the putative DEGs identified in blood monocytes, we downloaded brain gene expression data (RNA-seq) through publicly available GEO database (GSE101521). The expression data were generated in 59 postmortem brain tissue (29% females, mean age 49.3 ± 20.3 years old) collected by The Division of Molecular Imaging and Neuropathology at the New York State Psychiatric Institute and Columbia University. In brief, the whole-exome gene expression was examined in non-psychiatric controls ($N = 29$), MDD suicides ($N = 21$), and MDD non-suicides ($N = 9$) in postmortem brain tissue (dorsal lateral prefrontal cortex, Brodmann Area 9). All participants were free of medications. Detailed information of the brain samples was described previously²⁶. Depression was diagnosed by the structured clinical interview (DSM-IV) in both datasets.

Statistical analysis

The primary goal of our statistical analysis is to identify differentially methylated genes/regions (DMRs) associated with MDD. The overall study design and analytical plan was shown in Fig. S1.

Identifying differentially methylated regions (DMRs) associated with MDD

As DNA methylation between adjacent probes could be functionally and/or spatially correlated, identifying

genomic regions containing biologically relevant probes should be preferable compared to single CpG analysis. To achieve this, we employed a mixed-model designed specifically for discordant twin pairs²⁷

$$\begin{aligned} & \text{Log}_2 \left[\frac{\text{Methylation level in depressed twin}}{\text{Methylation level in nondepressed cotwin}} \right] \\ & = \alpha + \sum \beta_i x_i + \sum 1 | y_j \end{aligned}$$

In this model, β_i represents regression coefficient for the i th fixed-effect variable (e.g., age, sex), and y_j denotes the random-effect of the j th covariate (e.g. batch effect). The intercept α stands for the mean FC in DNA methylation between depressed and non-depressed twins within a pair. By testing the null hypothesis of $\alpha = 0$, we were able to test whether DNA methylation alteration ($\alpha > 0$ —hypermethylation, $\alpha < 0$ —hypomethylation) at a specific CpG site was associated with MDD or not. The effect of environmental factors on DNA methylation was assessed by testing, where β equals 0 or not ($\beta > 0$ indicates that the corresponding covariable causes hypermethylation, $\beta < 0$ indicates that it causes hypomethylation at the CpG site being tested in the depressed twins). This model allows for establishing a link between environmental exposure, epigenetic alteration, and depression. In this analysis, we adjusted for covariates including twin age, sex, BMI, smoking (pack-year), alcohol consumption, family income, and education level as fixed effect, and batch was included as random effect.

Region-based analysis was performed using results from single CpG analysis (p -value and effect size) by the program DMRcate²⁸, which identifies genomic regions harboring CpG sites and accounts for correlations between adjacent probes. Here we defined a DMR as a region containing ≥ 5 correlated probes (peak probe $p < 0.01$ and correlation among probes ≥ 0.30), and a significant DMR was defined as a region with $q < 0.05$ after correcting for total number of 6858 regions. Putative DMRs were ranked and annotated to genomic features based on the UCSC database (GRCh37/hg19). Genomic features of DMRs were compared to the null distribution of CpG probes included in the MethylationEPIC array²¹.

Differentially expressed genes (DEGs) associated with MDD

Using the statistical model described above, we identified DEGs associated with lifetime history of MDD, adjusting for same covariates.

Integrated methylome and transcriptome analysis

To examine the impact of DNA methylation on gene expression in peripheral blood monocytes, we calculated partial correlation coefficients (corrected for twin age, sex, BMI, smoking, alcohol, family income, and education) between DNA methylation and *cis*-acting gene expression

for each probe in all samples. Here *cis*-acting was defined as correlation between DNA methylation of a putative gene with its own expression (± 5 kb to a tested probe). We used a conservative threshold to determine significant negative (partial correlation < -0.84) or positive (partial correlation > 0.84) correlation. This threshold was obtained by randomly permuting DNA methylation and gene expression datasets for all twins. The cutoff was based on the fifth percentile of the empirical distribution of partial correlation coefficients, assuming no correlation between methylation and gene expression across all participants (null hypothesis). To further examine the role of DNA methylation in gene regulation, we used the Fisher's Exact test implemented in the GeneOverlap software²⁹ to determine the statistical significance of genes showing both differential methylation and differential expression in relation to MDD status.

Methods used for the replication in brain

For each of the 39 DMRs identified in blood monocytes, we tested the association of each probe within each region with depression (y/n) using logistic regression, adjusting for age and sex. Similar analysis was conducted to test the association of each putative DEG with depression, adjusting for age, sex, RIN, and brain PH. Multiple testing corrected for 39 DMRs or 30 DEGs using false discovery rate (FDR).

Co-methylation and co-expression networks

To examine the correlation patterns among putative DMRs and identify genes that are co-methylated, we performed network analyses using the weighted gene correlation network analysis (WGCNA). Co-methylation or co-expression networks were constructed separately in depressed twins and their non-depressed co-twins using all genes. Differential methylation or expression networks were identified by comparing the connectivity of differential genes between the two groups. This analysis included 322 DMRs showing nominal association (raw $p < 0.001$) with MDD. Similar analysis was used to identify co-expression networks (326 genes with raw $p < 0.001$ was included). Network visualization was done using CytoScape³⁰.

Functional-enrichment analysis

To explore the potential functional relevance of the identified DM or differentially expressed (DE) genes, we conducted functional enrichment analysis using the program DEPICT³¹. We first tested whether the putative genes ($p < 0.001$) are enriched in GWAS loci for major depression. A total of 746 independent SNPs previously associated with major depression ($p < 1 \times 10^{-5}$) was downloaded from the GWAS catalog³² for this analysis. We then tested whether the identified genes are enriched

in gene sets related to antidepressants by extracting 283 drug target genes from the Open Targets database³³.

Sensitivity analysis

To examine whether ACE modulates the association between DNA methylation and MDD, we further adjusted for ACE (y/n) in the above-described statistical model. Similarly, we tested the influence of antidepressants usage (y/n) or history of PTSD (y/n) on the relationship between DNA methylation and MDD.

Control for multiple comparisons

In the above-described analyses, we adjusted for multiple testing by FDR and FDR-adjusted p (i.e., q -value) < 0.05 was considered statistically significant.

Results

Table 1 shows the characteristics of the twins (mean age 38.2 ± 15.6 years, 68.4% females). Of the 79 twins with MDD, 8 twins have current MDD and 71 have MDD in the past. In addition, among the 79 depressed twins, 12 twins were under current medications, and 30 twins were taking medications in the past. Of note, 12 non-depressed co-twins were taking antidepressants prescribed for a number of different clinical indications outside of MDD, including chronic pain, anxiety, insomnia, and postpartum mood symptoms. None of the non-depressed co-twins with a history of antidepressant use met criteria for any MDD at any point in their lives. Except for current BDI-II score, PTSD, and use of antidepressants, depressed twins did not differ significantly from their non-depressed co-twins.

DMRs associated with lifetime history of MDD

We identified 39 DMRs (annotated to 36 unique genes) significantly associated with MDD at $q < 0.05$ (Table 2). Of these, 33 DMRs are hyper-methylated, and 6 are hypomethylated in relation to MDD. Figure 1 shows a Manhattan plot for these DMRs. The genomic distribution of these DMRs is shown in Fig. S2, which shows that the identified DMRs are enriched in the first exon and promoter regions but depleted in intergenic regions. In relation to CpG context, the identified DMRs are largely located within CpG Islands (CGIs).

DEGs associated with MDD

We identified 30 DEGs (14 upregulated, 16 downregulated) associated with lifetime history of MDD at $q < 0.05$ (Table 3). A Manhattan plot of these putative DEGs is shown in Fig. S3.

Replication in the brain

Of the 39 DMRs identified in blood, 10 regions contain at least one CpG probe showing significant association (same direction) with MDD ($q < 0.05$) in the brain after adjustments for covariates and total number of probes in each region (Table S2).

Of the 30 DEGs identified in blood monocytes, two genes (*NDUFA8*, *GUSBP9*) were also significantly associated with MDD ($q < 0.05$) in the brain (Table S3). While the expression level of the *NDUFA8* gene was lower in MDD patients than that in controls (i.e., downregulated) in both blood and brain, the *GUSBP9* gene was in an opposite direction (i.e., upregulated in blood, but downregulated in the brain between cases and controls).

Table 1 Clinical characteristics of twin pairs participating in the MMS

Variable	Non-depressed co-twin (N = 79)	Depressed twin (N = 79)	p-value ^a
Age, mean (SD), years	38.2 (15.6)	38.2 (15.6)	–
Female, no. (%)	54 (68.4)	54 (68.4)	–
Body mass index, mean (SD) (kg/m ²)	26.9 (6.5)	27.0 (6.9)	0.82
Smoking, mean (SD) (pack/year)	2.3 (9.4)	2.4 (8.3)	0.83
AUDIT-C score, mean (SD) ^b	3.6 (2.6)	3.4 (2.8)	0.70
Education below high school, no. (%)	6 (7.6)	4 (5.1)	0.43
Family income less than \$20,000, no. (%)	6 (7.6)	7 (8.9)	0.53
BDI-II score, mean (SD)	3.1 (5.1)	6.6 (6.5)	0.003
Exposure to ACE, no. (%)	8 (10.1)	9 (11.4)	0.66
History of PTSD, no. (%)	1 (1.3)	6 (7.6)	0.014
Use of antidepressants, no. (%)	12 (15.2)	42 (53.2)	0.006

BDI-II Beck Depressive Inventory II, ACE adverse childhood experiences, PTSD post-traumatic stress disorder

^ap-values were calculated using paired t -test

^bScore based on alcohol use disorders identification test (AUDIT-C)

Table 2 Significant DMRs associated with lifetime history of MDD in MZ discordant twin pairs

Chr	Start (bp)	End (bp)	Size (bp)	Nearest gene	# of probes	Peak P^a	Region P^b	Mean FC ^c	Mean difference (%) ^d	Peak difference (%) ^e
16	2,866,834	2,868,001	1168	<i>PRSS21</i>	10	6.28×10^{-4}	1.16×10^{-9}	1.13	4.06	14.28
5	43,037,123	43,037,666	544	<i>ANXA2R</i>	7	2.99×10^{-4}	2.96×10^{-9}	1.15	0.46	1.26
1	54,411,017	54,412,009	993	<i>HSPB11</i>	18	1.90×10^{-3}	3.69×10^{-7}	1.05	0.72	4.02
2	69,870,526	69,871,424	899	<i>AAK1</i>	8	5.70×10^{-4}	4.32×10^{-7}	1.05	2.48	8.45
4	186,732,926	186,733,331	406	<i>SORBS2</i>	8	3.53×10^{-3}	4.89×10^{-7}	1.14	6.98	13.36
2	26,395,359	26,395,859	501	<i>GAREML</i>	8	5.55×10^{-3}	7.20×10^{-7}	1.08	0.29	0.81
5	140,800,398	140,800,983	586	<i>PCDHGA11</i>	10	1.85×10^{-3}	1.04×10^{-6}	1.11	8.02	11.54
1	153,940,616	153,941,285	670	<i>CREB3L4</i>	6	2.76×10^{-3}	1.16×10^{-6}	1.12	11.12	15.53
5	43,602,380	43,603,353	974	<i>NNT</i>	17	2.76×10^{-3}	1.33×10^{-6}	1.06	1.03	6.84
5	8,457,538	8,458,392	855	<i>RP11</i>	9	4.84×10^{-3}	1.63×10^{-6}	0.89	-3.24	-8.65
22	38,598,577	38,599,166	590	<i>MAFF</i>	9	5.79×10^{-3}	3.06×10^{-6}	1.05	4.46	7.30
12	64,173,610	64,174,367	758	<i>TMEM5</i>	9	6.05×10^{-3}	3.32×10^{-6}	1.04	1.73	4.48
19	11,484,448	11,485,452	1005	<i>SWSAP1</i>	14	8.08×10^{-5}	3.94×10^{-6}	1.05	0.56	1.61
2	27,485,922	27,486,460	539	<i>SLC30A3</i>	8	3.43×10^{-3}	4.49×10^{-6}	1.07	1.52	6.90
2	101,034,246	101,034,295	50	<i>CHST10</i>	6	5.93×10^{-3}	5.48×10^{-6}	1.10	7.51	8.30
11	85,779,252	85,780,378	1127	<i>PICALM</i>	10	1.85×10^{-3}	5.72×10^{-6}	1.06	5.28	13.31
13	114,814,024	114,814,401	378	<i>CHAMP1</i>	5	7.42×10^{-3}	6.81×10^{-6}	1.09	7.08	13.05
11	32,454,216	32,455,025	810	<i>WT1</i>	8	2.18×10^{-3}	8.16×10^{-6}	1.06	2.34	6.02
16	70,557,411	70,557,707	297	<i>SF3B3</i>	10	7.70×10^{-3}	8.65×10^{-6}	1.06	3.32	8.43
7	90,224,158	90,225,380	1223	<i>CDK14</i>	11	9.78×10^{-4}	9.23×10^{-6}	0.99	-1.06	-3.14
11	87,908,134	87,908,805	672	<i>RAB38</i>	7	6.00×10^{-3}	9.72×10^{-6}	1.09	1.26	5.74
12	122,019,031	122,019,117	87	<i>KDM2B</i>	5	8.20×10^{-3}	1.37×10^{-5}	0.91	-8.86	-11.79
16	85,096,632	85,097,151	520	<i>KIAA0513</i>	5	6.05×10^{-3}	1.73×10^{-5}	0.94	-3.88	-7.55
2	65,594,021	65,595,186	1166	<i>SPRED2</i>	6	9.76×10^{-5}	1.81×10^{-5}	1.09	2.29	5.55
9	98,079,646	98,080,622	977	<i>FANCC</i>	10	2.64×10^{-4}	2.02×10^{-5}	1.07	6.64	11.76
10	90,611,604	90,612,228	625	<i>ANKRD22</i>	7	2.33×10^{-3}	2.13×10^{-5}	1.07	5.22	8.13
5	140,777,344	140,777,655	312	<i>PCDHA1</i>	9	4.87×10^{-3}	2.14×10^{-5}	1.11	8.87	13.05
1	178,994,834	178,995,133	300	<i>FAM20B</i>	8	2.76×10^{-3}	2.31×10^{-5}	1.06	5.27	9.33
7	148,936,572	148,937,410	839	<i>ZNF212</i>	9	3.03×10^{-3}	2.40×10^{-5}	1.06	1.46	4.08
16	68,118,822	68,119,261	440	<i>NFATC3</i>	9	7.94×10^{-3}	2.64×10^{-5}	1.05	4.97	8.83
19	59,030,662	59,031,081	420	<i>ZBTB45</i>	7	6.62×10^{-3}	3.75×10^{-5}	1.02	0.97	2.52
16	87,351,006	87,351,824	819	<i>C16orf95</i>	10	1.63×10^{-3}	4.00×10^{-5}	1.05	5.03	13.48
17	58,499,300	58,500,186	887	<i>C17orf64</i>	9	9.07×10^{-4}	4.27×10^{-5}	0.93	-5.43	-9.35
19	19,739,060	19,739,414	355	<i>LPAR2</i>	8	6.05×10^{-3}	4.65×10^{-5}	1.04	1.81	4.80
3	179,280,056	179,280,746	691	<i>ACTL6A</i>	9	3.97×10^{-3}	4.75×10^{-5}	1.02	1.48	4.75
1	70,876,598	70,877,381	784	<i>CTH</i>	9	5.78×10^{-3}	5.33×10^{-5}	1.06	1.23	3.90
7	78,400,383	78,400,769	387	<i>MAGI2</i>	5	6.61×10^{-3}	5.39×10^{-5}	1.09	8.22	12.42
6	30,297,174	30,297,941	768	<i>TRIM39</i>	10	2.76×10^{-3}	5.57×10^{-5}	1.09	8.93	14.69
15	69,222,400	69,223,018	619	<i>NOX5</i>	7	6.22×10^{-3}	5.79×10^{-5}	0.88	-4.11	-10.37

Chr: chromosome, Size: region size, # of probes: number of probes in the region

^aAdjusted for twin age, sex, BMI, smoking, alcohol consumption, education, and family income

^bAdjusted for a total number of 6858 regions

^cMean fold change (FC) in DNA methylation level across all CpG probes in a region. FC > 1 represents hypermethylated, whereas FC < 1 represents hypomethylated (depressed twin vs. non-depressed co-twin)

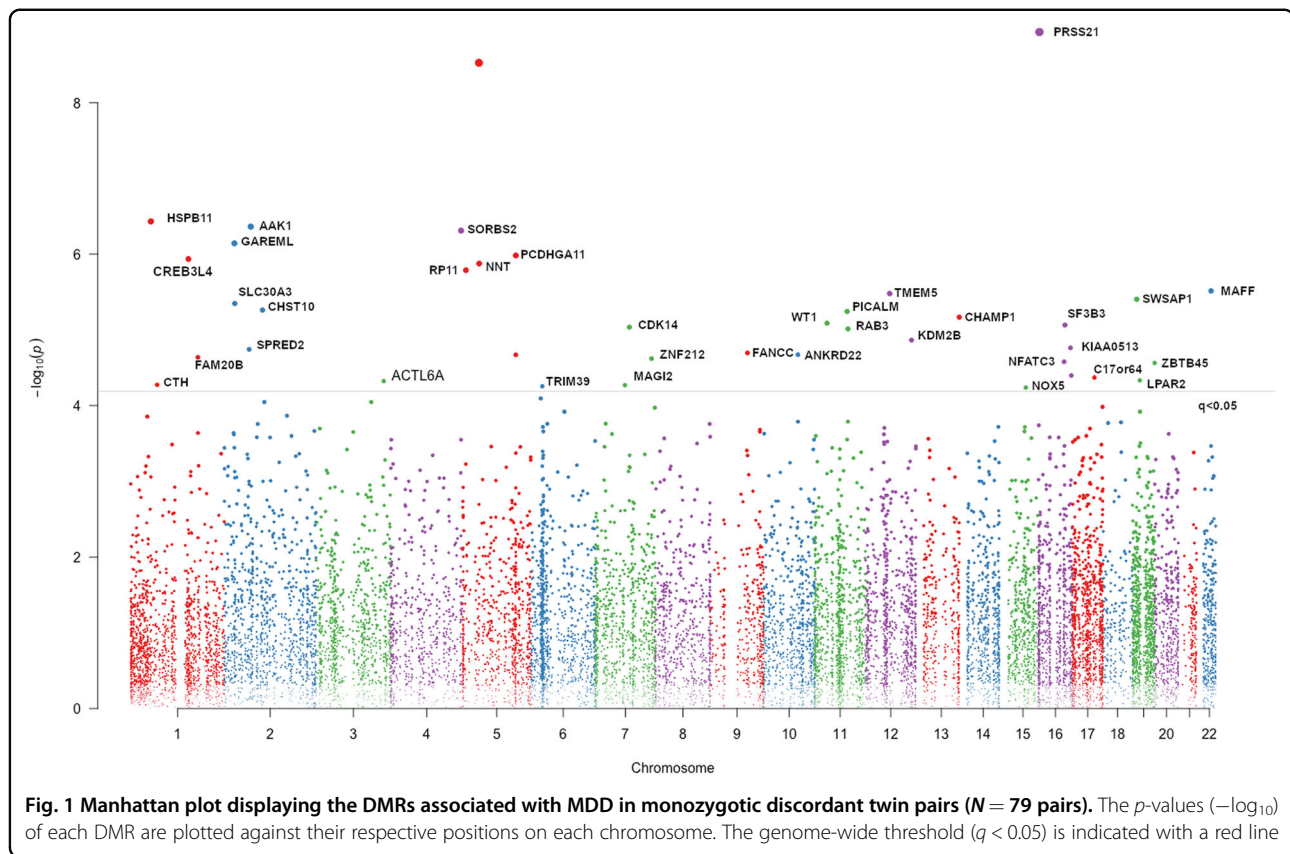
^dMean methylation difference across all probes in the region between depressed twins and their non-depressed co-twins

^eMethylation difference of the peak probe in the region between depressed twins and their non-depressed co-twins

Genome-wide integration of DNA methylome and transcriptome in blood monocytes

Figure 2 displays the genome-wide correlation patterns between DNA methylation and *cis*-acting gene expression in peripheral blood monocytes. It shows that DNA methylation was largely negatively (74% of the correlation pairs) correlated with gene expression, but positive

correlations were also observed. Interestingly, we found that the correlation patterns between DNA methylation and gene expression vary by genomic locations, with genes located about 1 kb upstream (negative) or downstream (positive) of the transcription start site (TSS) showing stronger correlation, whereas those located in between showing weak or no correlation (Fig. S4). These



correlation pairs (Table S4) involve 140 distinct methylated loci (from 64 unique genes) and 60 corresponding expression regions (representing 57 unique genes). The observed correlation between DNA methylation and gene expression appears to be independent of MDD, because after additional adjustment for MDD status (y/n) in the correlation analysis, majority (93.8%) of the correlation pairs remain statistically significant.

Differential network analysis

Our co-methylation network analysis identified three differential co-methylation modules containing at least 50 genes. Table S5 lists the co-methylation modules along with hub genes (genes with highest module membership) and biological pathways in each module. A total of 304 genes (94.4% of the total 322 genes showing nominal association with MDD) were assigned to at least one module. The largest module (Fig. S5) comprises 167 genes involved in four biological processes. The network connectivity (measured by node degrees) of two pathways in this module significantly differs between depressed and non-depressed twins. For example, the node connectivity for a pathway named “negative regulation of neuron apoptotic process” in depressed twins was significantly higher than that in non-depressed twins (3.8 vs. 2.6, $P = 7.15 \times 10^{-5}$), whereas the node connectivity of another

pathway called “stress-activated protein kinase signaling cascade” was significantly lower in depressed twins compared to their non-depressed co-twins (2.8 vs. 4.2, $P = 2.21 \times 10^{-5}$).

Our co-expression network analysis identified five modules containing 292 genes (Table S6). The largest module comprises 94 genes involved in eight biological processes. Of these, the “positive regulation of cytokine secretion” pathway showed significant difference between depressed and non-depressed twins (Fig. S6)

Functional-enrichment analysis

The identified DM genes are significantly enriched in pathways related to stress-activated protein kinase signaling cascade, neuron apoptotic process, negative regulation of insulin receptor signaling, mTOR signaling, and nerve growth factor receptor signaling pathways (Table S7). The differentially expressed genes are enriched in biological processes related to “positive regulation of cytokine secretion” and “regulation of response to stress”. Gene overlapping analysis revealed that the putative DM genes are 2.44 times more likely to be differentially expressed ($P = 1.1 \times 10^{-4}$), and are significantly over-represented in previous GWAS loci for major depression (2.32 times, $P = 2.4 \times 10^{-4}$). Moreover, these DM genes are significantly enriched in drug targets related to

Table 3 Significant DEGs associated with lifetime history of MDD in MZ discordant twin pairs

Chr	Start (bp)	End (bp)	Size (bp)	Nearest gene	FC ^a	P ^b	q ^c
2	215,996,329	216,082,955	86,626	<i>PECR</i>	1.49	7.00×10^{-8}	7.10×10^{-3}
1	185,292,384	185,294,372	1988	<i>AL356273.3</i>	1.32	5.12×10^{-7}	2.59×10^{-2}
1	44,800,225	44,805,990	5765	<i>PLK3</i>	2.48	9.35×10^{-7}	2.89×10^{-2}
5	71,197,646	71,208,130	10,484	<i>GUSBP9</i>	1.65	1.29×10^{-6}	2.89×10^{-2}
8	63,015,079	63,039,171	24,092	<i>GGH</i>	1.27	1.78×10^{-6}	2.89×10^{-2}
6	11,538,278	11,583,524	45,246	<i>TMEM170B</i>	8.02	2.47×10^{-6}	2.89×10^{-2}
11	64,223,799	64,226,254	2455	<i>TRPT1</i>	0.75	2.90×10^{-6}	2.89×10^{-2}
16	19,701,934	19,718,235	16,301	<i>KNOP1</i>	0.78	3.36×10^{-6}	2.89×10^{-2}
16	3,222,325	3,236,221	13,896	<i>ZNF200</i>	0.82	3.39×10^{-6}	2.89×10^{-2}
6	30,617,709	30,626,395	8686	<i>MRPS18B</i>	0.57	3.48×10^{-6}	2.89×10^{-2}
7	66,682,164	66,811,464	129,300	<i>RABGEF1</i>	1.46	3.65×10^{-6}	2.89×10^{-2}
22	49,900,229	49,918,458	18,229	<i>ALG12</i>	1.19	3.71×10^{-6}	2.89×10^{-2}
12	52,076,841	52,082,084	5243	<i>AC025259.1</i>	0.69	5.78×10^{-6}	3.62×10^{-2}
19	19,668,796	19,683,509	14,713	<i>ZNF101</i>	0.61	6.16×10^{-6}	3.62×10^{-2}
1	151,156,629	151,159,749	3120	<i>TNFAIP8L2</i>	0.16	6.35×10^{-6}	3.62×10^{-2}
8	33,473,386	33,513,601	40,215	<i>TTI2</i>	0.80	6.79×10^{-6}	3.62×10^{-2}
13	41,457,559	41,470,882	13,323	<i>RGCC</i>	4.27	7.44×10^{-6}	3.74×10^{-2}
9	122,144,058	122,159,819	15,761	<i>NDUFA8</i>	0.59	7.73×10^{-6}	3.74×10^{-2}
11	93,741,591	93,764,749	23,158	<i>C11orf54</i>	0.59	8.66×10^{-6}	3.99×10^{-2}
16	85,690,084	85,751,129	61,045	<i>C16orf74</i>	1.62	9.17×10^{-6}	4.04×10^{-2}
7	80,742,538	80,922,359	179,821	<i>SEMA3C</i>	1.81	1.03×10^{-5}	4.10×10^{-2}
19	21,397,119	21,427,573	30,454	<i>ZNF493</i>	1.65	1.05×10^{-5}	4.10×10^{-2}
11	4,384,897	4,393,696	8799	<i>TRIM21</i>	0.28	1.18×10^{-5}	4.30×10^{-2}
19	52,949,379	52,962,911	13,532	<i>ZNF816</i>	0.72	1.21×10^{-5}	4.30×10^{-2}
10	43,436,841	43,483,179	46,338	<i>ZNF487</i>	1.44	1.23×10^{-5}	4.30×10^{-2}
19	57,466,663	57,477,570	10,907	<i>ZNF772</i>	0.84	1.27×10^{-5}	4.30×10^{-2}
22	42,509,968	42,519,802	9834	<i>RRP7A</i>	0.43	1.34×10^{-5}	4.31×10^{-2}
21	36,069,941	36,073,166	3225	<i>CBR1</i>	0.35	1.36×10^{-5}	4.31×10^{-2}
2	98,619,106	98,731,126	112,020	<i>MGAT4A</i>	2.80	1.60×10^{-5}	4.91×10^{-2}
1	145,911,350	145,918,837	7487	<i>PEX11B</i>	0.67	1.66×10^{-5}	4.94×10^{-2}

Chr: Chromosome, Size: length of gene in base pair

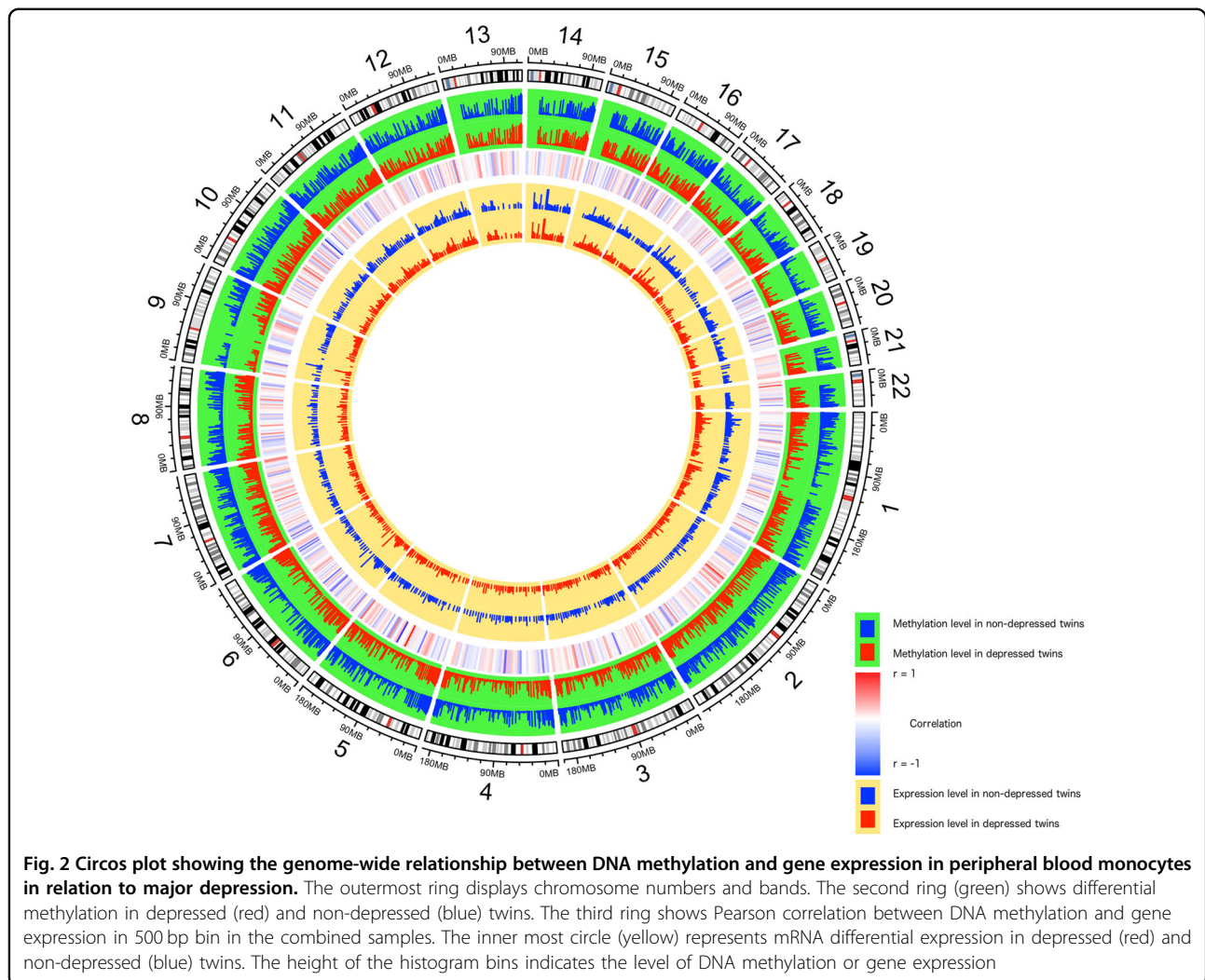
^aFold change (FC) in gene expression level between depressed twins and their non-depressed co-twins^bAdjusted for twin age, sex, BMI, smoking, alcohol consumption, education, and family income^cAdjusted for a total number of 10,329 genes

antidepressants (2.83 times, $P = 7.6 \times 10^{-5}$), and are highly expressed in tissues/cell types related to the nervous, endocrine, and urogenital systems (Fig. S7). Together, these results suggest a potential functional role of altered DNA methylation in MDD pathogenesis.

Sensitivity analysis

Additionally adjustment for childhood traumatic experience (ACE) slightly attenuated the association between

DNA methylation and MDD, but results remained largely unchanged. After further correction for use of antidepressants or PTSD, the association of the *MAFF* gene with MDD disappeared, but other genes remained statistically significant. Results for sensitivity analysis are shown in Table S8. It appears that additionally adjustments for adverse childhood events (ACE), PTSD, or use of antidepressants did not have an appreciable effect on the association between gene expression and MDD (Table S9).



Discussion

Using a MZ discordant co-twin control design, we identified 39 DMRs (33 hypermethylated, 6 hypomethylated) and 30 differentially expressed genes (14 upregulated, 16 downregulated) associated with lifetime history of MDD, after accounting for clinical covariates and multiple testing. These DM or expressed genes are significantly enriched in biological processes related to neuronal function, stress response, insulin regulation, mTOR signaling, and cytokine secretion, suggesting potential relevance to MDD pathogenesis. Moreover, the identified DMR genes are overrepresented in GWAS loci and drug targets related to antidepressants. Integrated DNA methylome and transcriptome analysis revealed that DNA methylation was both negatively and positively correlated with gene expression in peripheral blood monocytes. To the best of our knowledge, this is the first integrated DNA methylome and transcriptome analysis in purified blood monocytes for lifetime history of MDD

using a relatively large number of MZ discordant pairs from a community-based population cohort.

In a recent meta-analysis examining the association between blood DNA methylation and depressive symptoms among middle-aged and elderly persons⁴, DNA methylation at three CpG sites were significantly associated with depressive symptoms. One of these probes (cg12325605) could be replicated in our study at a nominal level ($P = 2.28 \times 10^{-4}$). Multiple prior studies have also reported possible associations of DNA methylation in stress-related genes³⁴ or genome-wide associations using unrelated individuals or twins in various tissues (peripheral or postmortem brain)^{10,35,36}. However, due to the heterogeneity of MDD phenotypes and many other differences between studies (e.g., racial/ethnic groups, age ranges of participants, tissue/cell types, risk factor profiles, etc.), it is challenging to compare results across different scenarios, and no conclusive genes have been identified so far. Of the limited number of existing EWAS on MDD,

the present study represents the first to utilize purified blood monocytes in a relatively large sample of MZ twin pairs discordant on MDD.

Of the 36 annotated DMR genes, the *PRSS21* gene showed the most significant association with MDD. The methylation level of this gene in depressed twins is on average 1.13-fold as high as that in their non-depressed co-twins. This gene (also known as testisin) encodes a glycosylphosphatidylinositol (GPI)-linked serine protease, which is a member of the trypsin family of serine proteases. A growing body of evidence demonstrates that the brain can co-opt the activities of these serine proteases and their receptors/inhibitors to regulate various processes including synaptic activity, learning, and social behavior³⁷. Aberrant activity of these molecules may contribute to neurological disorders such as Alzheimer's disease, Parkinson's disease, traumatic brain injury, and stroke³⁷. Another significantly hypermethylated gene is *HSPB11*, which encodes a family member of the heat shock proteins (HSPs) that are produced in response to stressful conditions³⁸. Although the exact mechanisms behind the association of *HSPB11* hypermethylation with MDD are unknown, HSPs are involved in protein misfolding and aggregation, which have been implicated in neurodegenerative and neuropsychiatric disorders³⁹. Other top-ranked DMR genes, such as *AAKI*, *SORBS2*, and *GAREM2*, are also abundantly expressed in the brain and may affect MDD susceptibility through a variety of biological processes. For example, the *AAKI* gene is involved in intracellular vesicle trafficking, a mechanism that is essential for neurotransmitter release and recycling of synaptic vesicle proteins⁴⁰. Inhibition of *AAKI* activity may provide a novel therapeutic target for treating neuropsychiatric and neurodegenerative disorders⁴¹. The *SORBS2* gene encodes the Arg protein tyrosine kinase-binding protein 2 (ArgBP2), downregulation of which was previously associated with mood disorders⁴². The *GAREM2* gene encodes an adapter protein that regulates MAPK/ERK signaling, which is involved in neuronal plasticity and resilience in psychiatric disorders⁴³. Further, the identified DMR genes are enriched in neuron apoptosis⁴⁴, nerve growth factor⁴⁵, stress-activated protein kinase signaling⁴⁶, insulin receptor regulation⁴⁷, and mTOR signaling⁴⁸, suggesting potential relevance to MDD pathogenesis.

Of the differentially expressed genes associated with MDD, the peroxisomal *trans*-2-enoyl-CoA reductase (*PECR*) gene showed the strongest association. This gene is involved in mitochondrial energy production by catalyzing the reduction of enoyl-CoAs to acyl-CoAs, and genetic polymorphisms in this gene were associated with alcohol dependence⁴⁹. The differentially expressed genes are enriched in regulation of cytokine secretion and oxidative stress responses, lending further support for the

critical roles of inflammation and oxidative stress in major depression¹³.

Previous studies have demonstrated that zinc plays an essential role in synaptic activity and neuronal plasticity, and that zinc deficiency was associated with behavioral impairments⁵⁰, neurodegenerative disorders, and mood disorders including depression⁵¹. In line with these findings, we found that several zinc family genes are DM (e.g., *SLC30A3*, *ZNF212*, *ZBTB45*, *SWSAPI*, *WT1*, *TRIM39*) or differentially expressed (e.g., *ZNF200*, *ZNF101*, *ZNF493*, *ZNF816*, *ZNF487*, *ZNF772*) between depressed twins and their non-depressed co-twins. These findings provide further support for a potential important role of zinc dysregulation in depression, and suggest that DNA methylation may modulate the effect of zinc function on depression susceptibility. Together, our results may unravel novel molecular pathways underlying MDD pathogenesis.

Although DNA methylation is generally believed to cause gene silencing, a global analysis of the extent and pattern of DNA methylation with gene expression in human blood monocytes is still lacking. Here we demonstrated that monocyte DNA methylation can be both positively and negatively correlated with gene expression, although there appears to be a trend that the correlations are predominantly negative in promoter regions. These results are in agreement with previous studies reporting both positive and negative correlations between DNA methylation and gene expression in blood and brain⁵². Interestingly, we found that the relationship between DNA methylation and *cis*-acting gene expression appears to vary by genomic locations, with methylation of putative genes located upstream of TSS showing predominantly negative correlations, whereas those located downstream of TSS showing largely positive correlations with gene expression. While the negative correlation may result from the interference with transcription factor binding or recruiting repressors such as histone deacetylases⁵¹, the positive correlation between DNA methylation and gene expression may be attributed to the high level of DNA methylation in the gene body of highly transcribed genes. In addition, it has been shown that DNA methylation of a promoter or an enhancer can activate transcription of a target gene and therefore is positively correlated with gene expression⁵³. Together, our integrated DNA methylome and transcriptome analysis confirmed the importance of DNA methylation in gene regulation and identified key candidate genes whose role in depression are modulated by epigenetic changes.

MDD is a highly heterogeneous disorder involving the joint or interactive effects of many genes in multiple pathways. Traditional methods that model the effect of a single gene cannot capture the complicated biological pathways implicated in MDD. Using a network-based

approach, we identified co-methylated or co-expressed modules containing coordinated genes across different genomic loci. These findings support the hypothesis that altered DNA methylation is interdependent and that the blood monocytes methylome and transcriptome comprise a complex network of interacting processes.

Our study has several limitations. First, in spite of using 79 MZ discordant twin pairs, our study is still underpowered and thus we might have missed important disease-related genes, especially those with small individual effect size. As such, the current analysis focused on region-based rather than single probe analysis. Our results should be considered as a proof of concept rather than conclusive. Second, although we used purified blood monocytes for DNA methylation and gene expression profiling, our sample still includes other cell types, such as macrophages or dendritic cells, each of which may have different epigenetic profiles. However, our sample has nearly 97% purity and we believe confounding by cellular heterogeneity should not be a major concern for our study. Moreover, given the cell-type-specific nature of DNA methylation, it is unclear to what extent our results derived from peripheral blood could reflect methylation changes in the brain. However, monocytes are one of the key components of the innate immunity system, dysfunction of which has been implicated in neuropsychiatric disorders including depression⁵⁴, and accumulating evidence indicated that epimutations may not be limited to the affected organ (e.g., brain) but could also be detected in peripheral blood⁵⁵. Further, many of the identified DMR genes are abundantly expressed and some could be replicated in the brain. These putative genes detected in easily accessible tissues, such as blood are suitable for biomarkers. Third, the twin participants were evaluated for lifetime history of MDD and some non-depressed co-twins might ultimately develop an episode of MDD. The mean age of our sample was 38, which suggests that a large portion of the participants had already passed the peak risk period of young adulthood⁵⁶. Fourth, due to the uneven distribution of the CpG probes on the Illumina methylation arrays, we cannot entirely exclude the possibility that the observed correlation patterns between monocytes DNA methylation and *cis*-acting gene expression is likely reflecting the way the Illumina methylation array was designed. Whether the observed correlation patterns between DNA methylation and gene expression is due to true biology or artifact could be verified when whole genome bisulfite sequencing data become available in the near future. Moreover, like all other observational studies, establishing causality between DNA methylation and MDD is impossible in the present study. Fifth, of the putative CpG probes identified in this study, there is a considerable likelihood that some CpG sites, especially those with smaller effect sizes, could be

due to technical artifact but not reflect true biology. Other potential limitations include the limited genome coverage of the EPIC array used in our study, and the uncertainty to generalize our findings to other racial/ethnic groups.

Our study has several strengths. First, as MZ twin pairs share almost identical genotypes, age, sex, and early familial environment (e.g., in utero environment), as well as many unknown or unmeasured factors, the use of a MZ discordant co-twin control design minimizes or eliminates potential confounding by these factors. Second, we profiled DNA methylome and transcriptome in purified blood monocytes, which minimizes confounding by cellular heterogeneity. Moreover, we used the structured clinical interview (DSM-IV) for depression diagnosis.

In summary, our results demonstrated a critical role of altered DNA methylation and gene expression in MDD, and identified key candidate genes and pathways underlying MDD pathogenesis. If validated, the newly identified genes and pathways may serve as novel therapeutic targets for MDD and related disorders.

Acknowledgements

The authors would like to thank the twin members for their participation. Without their contribution, this research would not have been possible. This study was supported by NIH grant R01MH097018. These funding have no role in design, analysis, interpretation, or paper writing of this manuscript.

Author contributions

E.S., E.F., T.B., and P.R.-B. collected clinical data and biospecimen. Y.Z. performed data analysis and draft the manuscript. J.Z. and E.S. designed the research and wrote the manuscript. P.R.-B. and E.F. provided critical evaluation of the manuscript. All authors read and approved the final manuscript.

Author details

¹Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, Gainesville, FL, USA.

²Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA. ³Department of Pediatrics, University of Washington, Seattle, WA, USA

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41398-019-0550-2>).

Received: 5 October 2018 Revised: 11 April 2019 Accepted: 29 April 2019

Published online: 02 September 2019

References

1. World Health Organization. *Mental Health and Older Adults* (World Health Organization, Geneva, 2013).
2. Sullivan, P. F., Neale, M. C. & Kendler, K. S. Genetic epidemiology of major depression: review and meta-analysis. *Am. J. Psychiatry* **157**, 1552–1562 (2000).

3. Mill, J. & Petronis, A. Molecular studies of major depressive disorder: the epigenetic perspective. *Mol. Psychiatry* **12**, 799–814 (2007).
4. Story Jovanova, O. et al. DNA methylation signatures of depressive symptoms in middle-aged and elderly persons: meta-analysis of multiethnic epigenome-wide studies. *JAMA Psychiatry* **7**, 949–959 (2018).
5. Chong, S. & Whitelaw, E. Epigenetic germline inheritance. *Curr. Opin. Genet. Dev.* **14**, 692–696 (2004).
6. Szyf, M. DNA methylation, the early-life social environment and behavioral disorders. *J. Neurodev. Disord.* **3**, 238–249 (2011).
7. Houseman, E. A. et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinforma.* **13**, 86 (2012).
8. Bell, J. T. & Spector, T. D. A twin approach to unraveling epigenetics. *Trends Genet.* **27**, 116–125 (2011).
9. Davies, M. N. et al. Hypermethylation in the ZBTB20 gene is associated with major depressive disorder. *Genome Biol.* **15**, R56 (2014).
10. Cordova-Palomera, A. et al. Genome-wide methylation study on depression: differential methylation and variable methylation in monozygotic twins. *Transl. Psychiatry* **5**, e557 (2015).
11. Dempster, E. L. et al. Genome-wide methylomic analysis of monozygotic twins discordant for adolescent depression. *Biol. Psychiatry* **76**, 977–983 (2014).
12. Malki, K. et al. Epigenetic differences in monozygotic twins discordant for major depressive disorder. *Transl. Psychiatry* **6**, e839 (2016).
13. Miller, A. H. & Raison, C. L. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* **16**, 22–34 (2016).
14. Strachan, E., Zhao, J., Roy-Byrne, P. P., Fowler, E. & Bacus, T. Study design and rationale for the Mood and Methylation Study: a platform for multi-omics investigation of depression in twins. *Twin Res. Hum. Genet.* **21**, 507–513 (2018).
15. Strachan, E. et al. University of Washington Twin Registry: poised for the next generation of twin research. *Twin Res. Hum. Genet.* **16**, 455–462 (2013).
16. Tsai, P.-C. & Bell, J. T. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *Int. J. Epidemiol.* **44**, 1429–1441 (2015).
17. Beck, A. T., Steer, R. A., Brown, G. K. *BDHI. Beck Depression Inventory*. 2nd edn. (The Psychological Corporation, San Antonio, TX, 1996).
18. Rush, A. J. et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS) Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol. Psychiatry* **54**, 573–583 (2003).
19. Felitti, V. J. et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. *Am. J. Prev. Med.* **14**, 245–258 (1998).
20. Bremner, J. D., Bolus, R. & Mayer, E. A. Psychometric properties of the Early Trauma Inventory-self report. *J. Nerv. Ment. Dis.* **195**, 211–218 (2007).
21. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389–399 (2016).
22. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
23. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
24. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinforma.* **12**, 323 (2011).
25. Guintivano, J., Aryee, M. J. & Kaminsky, Z. A. 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**, 290–302 (2013).
26. Pantazatos, S. P. et al. Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity. *Mol. Psychiatry* **22**, 760–773 (2017).
27. Tan, Q. Epigenetic epidemiology of complex diseases using twins. *Med. Epigenet.* **1**, 46–51 (2013).
28. Peters, T. J. et al. De novo identification of differentially methylated regions in the human genome. *Epigenet. Chromatin* **8**, 6 (2015).
29. Shen, L., Sinai, M. GeneOverlap: Test and visualize gene overlaps. 2013.
30. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
31. Pers, T. H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
32. MacArthur, J. et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**(D1), D896–D901 (2017).
33. Koscielny, G. et al. Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res.* **45**(D1), D985–D994 (2017).
34. Roy, B., Shelton, R. C. & Dwivedi, Y. DNA methylation and expression of stress related genes in PBMC of MDD patients with and without serious suicidal ideation. *J. Psychiatr. Res.* **89**, 115–124 (2017).
35. Córdova-Palomera, A., Palma-Gudiel, H. & Forés-Martos, J. Tabarés-Seisdedos R2, Fañanás L: Epigenetic outlier profiles in depression: a genome-wide DNA methylation analysis of monozygotic twins. *PLoS ONE* **13**, e0207754 (2018).
36. Dempster, Emma L. et al. Genome-wide methylomic analysis of monozygotic twins discordant for adolescent depression. *Biol. Psychiatry* **76**, 977–983 (2014).
37. Almonte, A. G. & Sweatt, J. D. Serine proteases, serine protease inhibitors, and protease-activated receptors: roles in synaptic function and behavior. *Brain Res.* **1407**, 107–122 (2011).
38. Feder, M. E. & Hofmann, G. E. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu. Rev. Physiol.* **61**, 243–282 (1999).
39. Polajnar, M. & Zerovnik, E. Impaired autophagy: a link between neurodegenerative and neuropsychiatric diseases. *J. Cell. Mol. Med.* **18**, 1705–1711 (2014).
40. Nakazawa, T. et al. Emerging roles of ARHGAP33 in intracellular trafficking of TrkB and pathophysiology of neuropsychiatric disorders. *Nat. Commun.* **7**, 10594 (2016).
41. Kuai, L. et al. AAK1 identified as an inhibitor of neuregulin-1/ErbB4-dependent neurotrophic factor signaling using integrative chemical genomics and proteomics. *Chem. Biol.* **18**, 891–906 (2011).
42. Lee, S. E., Kim, J. A. & Chang, S. nArgBP2-SAPAP-SHANK, the core postsynaptic triad associated with psychiatric disorders. *Exp. Mol. Med.* **50**, 2 (2018).
43. Taniguchi, T. et al. A brain-specific Grb2-associated regulator of extracellular signal-regulated kinase (Erk)/mitogen-activated protein kinase (MAPK) (GAREM) subtype, GAREM2, contributes to neurite outgrowth of neuroblastoma cells by regulating Erk signaling. *J. Biol. Chem.* **288**, 29934–29942 (2013).
44. Jarskog, L. F., Glantz, L. A., Gilmore, J. H. & Lieberman, J. A. Apoptotic mechanisms in the pathophysiology of schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **29**, 846–858 (2005).
45. Rao, S. et al. Peripheral blood nerve growth factor levels in major psychiatric disorders. *J. Psychiatr. Res.* **86**, 39–45 (2017).
46. Paul, A. et al. Stress-activated protein kinases: activation, regulation and function. *Cell. Signal.* **9**, 403–410 (1997).
47. Pearson, S. et al. Depression and insulin resistance: cross-sectional associations in young adults. *Diabetes Care* **33**, 1128–1133 (2010).
48. Ignacio, Z. M. et al. New perspectives on the involvement of mTOR in depression as well as in the action of antidepressant drugs. *Br. J. Clin. Pharmacol.* **82**, 1280–1290 (2016).
49. Treutlein, J. et al. Genome-wide association study of alcohol dependence. *Arch. Gen. Psychiatry* **66**, 773–784 (2009).
50. Hagemeyer, S., Haderspeck, J. C. & Grabrucker, A. M. Behavioral impairments in animal models for zinc deficiency. *Front. Behav. Neurosci.* **8**, 443 (2014).
51. Doboszewska, U. et al. Zinc in the Monoaminergic Theory of Depression: its relationship to neural plasticity. *Neural Plast.* **2017**, 3682752 (2017).
52. Bell, J. T. et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* **12**, R10 (2011).
53. Bahar Halpern, K., Vana, T. & Walker, M. D. Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development. *J. Biol. Chem.* **289**, 23882–23892 (2014).
54. Beumer, W. et al. The immune theory of psychiatric diseases: a key role for activated microglia and circulating monocytes. *J. Leukoc. Biol.* **92**, 959–975 (2012).
55. Menke, A. & Binder, E. B. Epigenetic alterations in depression and antidepressant treatment. *Dialog- Clin. Neurosci.* **16**, 395–404 (2014).
56. Weissman, M. M. Cross-national epidemiology of major depression and bipolar disorder. *JAMA* **276**, 293 (1996).