

## RESEARCH

## Open Access



# meQTL and ncRNA functional analyses of 102 GWAS-SNPs associated with depression implicate *HACE1* and *SHANK2* genes

Diana M. Ciuculete<sup>1\*</sup> , Sarah Voisin<sup>2</sup>, Lara Kular<sup>3</sup>, Jörgen Jonsson<sup>1</sup>, Mathias Rask-Andersen<sup>4</sup>, Jessica Mwinyi<sup>1</sup> and Helgi B. Schiöth<sup>1,5</sup>

## Abstract

**Background:** Little is known about how genetics and epigenetics interplay in depression. Evidence suggests that genetic variants may change vulnerability to depression by modulating DNA methylation (DNAm) and non-coding RNA (ncRNA) levels. Therefore, the aim of the study was to investigate the effect of the genetic variation, previously identified in the largest genome-wide association study for depression, on proximal DNAm and ncRNA levels.

**Results:** We performed DNAm quantitative trait locus (meQTL) analysis in two independent cohorts (total  $n = 435$  healthy individuals), testing associations between 102 single-nucleotide polymorphisms (SNPs) and DNAm levels in whole blood. We identified and replicated 64 SNP-CpG pairs ( $p_{\text{adj.}} < 0.05$ ) with meQTL effect. Lower DNAm at cg02098413 located in the *HACE1* promoter conferred by the risk allele (C allele) at rs1933802 was associated with higher risk for depression ( $p_{\text{raw}} = 0.014$ , DNAm = 2.3%). In 1202 CD14+ cells sorted from blood, DNAm at cg02088412 positively correlated with *HACE1* mRNA expression. Investigation in postmortem brain tissue of adults diagnosed with major depressive disorder (MDD) indicated 1% higher DNAm at cg02098413 in neurons and lower *HACE1* mRNA expression in CA1 hippocampus of MDD patients compared with healthy controls ( $p = 0.008$  and 0.012, respectively). Expression QTL analysis in blood of 74 adolescent revealed that hsa-miR-3664-5p was associated with rs7117514 (*SHANK2*) ( $p_{\text{adj.}} = 0.015$ , mRNA difference = 5.2%). Gene ontology analysis of the miRNA target genes highlighted implication in neuronal processes.

**Conclusions:** Collectively, our findings from a multi-tissue (blood and brain) and multi-layered (genetic, epigenetic, transcriptomic) approach suggest that genetic factors may influence depression by modulating DNAm and miRNA levels. Alterations at *HACE1* and *SHANK2* loci imply potential mechanisms, such as oxidative stress in the brain, underlying depression. Our results deepened the knowledge of molecular mechanisms in depression and suggest new epigenetic targets that should be further evaluated.

**Keywords:** DNA methylation, Genetics, meQTL, MicroRNA, Depression

\* Correspondence: [diana-maria.ciuculete@neuro.uu.se](mailto:diana-maria.ciuculete@neuro.uu.se)

<sup>1</sup>Department of Neuroscience, Functional Pharmacology, Uppsala University, BMC, Box 593, Husargatan 3, 753124 Uppsala, Sweden

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



© The Author(s). 2020 **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Depression is the leading cause of disability worldwide [1]. Twin studies have shown that depression has a heritability of 30–40% [2], and genome-wide association studies (GWASs) have uncovered many single nucleotide polymorphisms (SNPs) associated with this condition [3]. The most recent and largest genome-wide meta-analysis of depression included 807,533 individuals and identified 102 such SNPs, 87 of which were replicated in an independent cohort of 1,306,354 individuals [4]. These SNPs all showed small effect sizes, and their functional roles remain unknown. The collective profile of SNPs in each individual forms a genetic background that confers a degree of predisposition to depression. Environmental factors interact with this genetic profile to trigger depression, partly via epigenetic mechanisms, such as the regulatory action of DNA methylation (DNAm) and non-coding RNAs (ncRNA). Studying the intricate relationship between genetic variation, DNA methylation, and ncRNA expression is therefore important to understand the molecular mechanisms underpinning susceptibility to depression.

DNAm is strongly involved in neurodevelopmental processes such as synaptic formation and function [5, 6], especially during adolescence, and its perturbation can lead to psychiatric disorders [7]. CpG methylation, a process characterized by the addition of a methyl group to the fifth carbon of the cytosine ring, influences the packing of the DNA and its availability for transcription factors [8]. Typically, DNAm in gene promoters is associated with transcriptional repression while is positively associated with gene expression when located in gene bodies [9]. Site-specific DNAm is extensively influenced by the genetic background [10], age [11], sex [12], and environmental exposure, such as social stress [13, 14]. ncRNAs such as microRNAs (miRNAs) operate at the post-transcriptional level to add another layer of gene expression regulation. Increasing evidence suggests miRNAs as key modulators of synaptic plasticity and activity [15, 16] and as an important mechanism linking external stimuli to changes in brain function [17].

DNAm and miRNA expression may differ depending on the genotype at a given SNP, which then may cause alterations in cell, tissue, and whole-body physiology underlying depression. Interestingly, previous studies suggested that genetic and epigenetic variations could interact [18, 19]. This emphasizes the contribution of epigenetics to human disease, especially as alterations may be linked to non-protein-coding, disease-associated genetic variants from GWAS [20, 21].

In this study, we aimed to investigate genotype-dependent DNAm and ncRNA levels in depression, by using a multi-tissue (blood and brain) and multi-layered (genetics, epigenetics, transcriptomics) approach in

different cohorts. We uncovered associations between the 102 genetic variants identified in the abovementioned GWAS and shifts in epigenetics (DNAm and ncRNA) that play a role in potential mechanisms underlying depression. First, we identified and replicated DNA methylation quantitative trait loci (meQTL) in blood and brain from two independent cohorts. We used two additional cohorts to investigate the functional consequences of those DNAm alterations at the gene expression level. Finally, we addressed association between genetic variation and ncRNA expression, using a *cis*-expression quantitative trait locus (eQTL) approach.

## Results

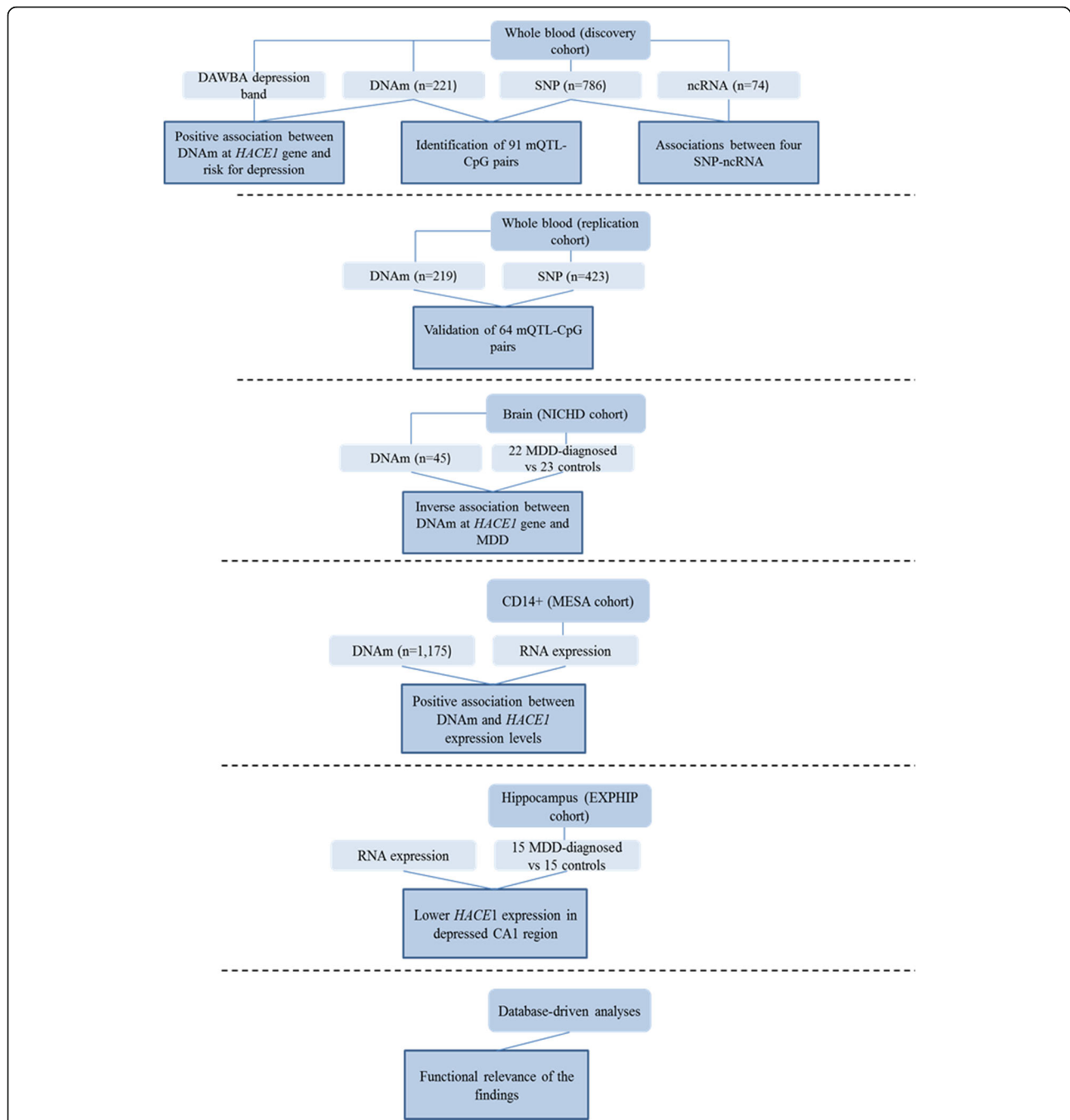
### Demographic characteristics of the cohorts

Our workflow is summarized in Fig. 1. Both the 216 adolescents from the discovery cohort and the 219 young adults from the replication cohort had whole blood methylome and SNP data available. Adolescents were  $15.5 \pm 0.63$  years old, and 75% of them were female and 26% at high risk for depression, as assessed by the Development and Well-Being Assessment (DAWBA) (Table 1). The participants in the replication cohort were young adults ( $23.5 \pm 3.3$  years old), and 22% of them were females. All participants of the replication cohort, and 97.2% of the discovery cohort, were free of medication. Differences between the discovery and replication cohorts included a difference in age of approximately 8 years and in sex distribution (75% vs 22% females). Body mass index (BMI) was comparable between the two cohorts ( $21.9 \pm 3.44$  and  $23.7 \pm 3.03$ , Supplementary Figure 1). The 22 adults with major depression disorder (MDD) and the 23 matched controls from the NICHD Brain Bank of Developmental Disorders had similar age and sex distribution (Table 2).

Expression analysis in blood cells was performed in 1202 adults in the MESA cohort, and the mean age was  $60.2 \pm 9.45$  years old. Information about race and sex could not be extracted from the variable recorded incorporating race, sex, and study site. The 30 adults from the EXPHIP cohort had mRNA levels available in two hippocampal regions (CA1 and dentate gyrus) and were 60% males and 83% Caucasians (Table 3).

### Association of depression-related genetic risk variants with proximal DNAm in whole blood

We performed a *cis*-meQTL analysis in the discovery cohort (Table 1), using a dominant genetic model for each SNP, and controlling for age, sex, and batch. In our cohort, the minor allele frequency (MAF) of the 102 GWAS studied SNPs that were in a range similar to other Caucasian populations [4] (Supplementary Table 1). We identified 91 CpGs associated with 46 SNPs, forming 91 SNP-CpG pairs spread throughout the



**Fig. 1** Study design and workflow diagram. Using whole blood of the discovery cohort comprising adolescents, meQTL analyses, together with associations between ncRNA levels and genotype and DNAm and depression scores were performed. The replication cohort comprised 219 adults and 64 previously identified meQTL-CpG pairs were validated in whole blood. The NICHD cohort containing 45 adult brain tissues was used to analyze the association between the MDD diagnosis and DNAm at the HACE1 gene. A positive relationship between DNAm degree and RNA expression levels of HACE1 was identified in CD14+ cells, in the MESA cohort. The EXPHIP cohort included 15 MDD-diagnosed and 15 matched controls and helped to identify lower HACE1 mRNA levels in depressed CA1 hippocampal region. The functional relevance of the findings was investigated using different bioinformatic or molecular biological software tools. DAWBA, Development and Well-Being Assessment; DNAm, DNA methylation; SNP, single-nucleotide polymorphism; ncRNA, non-coding RNA; meQTL, methylation quantitative trait locus; MDD, major depressive disorder

**Table 1** Characteristics of the adolescents from the discovery cohort

	Low-risk (n = 160)	High-risk (n = 56)	p value**
Male: Female	47 (19.4):113 (80.6)	6 (13.1):50 (86.9)	<b>0.0089</b>
Age (years)	15.4 ± 0.63	15.6 ± 0.65	0.15
Body mass index (kg/m <sup>2</sup> )	21.61 ± 3.02	22.83 ± 4.33	0.054
DAWBA level bands*			
General band	70 (19.4)	56 (100)	<b>6.03e-14</b>
Panic disorder	9 (5.6)	5 (8.9)	<b>0.00041</b>
Posttraumatic disorder	14 (2.5)	9 (16)	<b>5.44e-6</b>
Separation anxiety disorder	5 (3.1)	6 (10.7)	<b>2.05e-5</b>
Generalized anxiety disorder	30 (18.7)	23 (41.1)	<b>2.11e-7</b>
Social phobia	17 (10.6)	14 (25)	<b>0.016</b>
Obsessive-compulsive disorder	9 (5.6)	6 (10.7)	<b>0.028</b>
Conduct disorder	10 (6.2)	9 (16.1)	<b>7.89e-6</b>
Specific phobia	9 (5.6)	7 (12.5)	<b>0.0072</b>

Variables are shown as mean ± standard deviation or number (percentage)

\*All listed DAWBA bands refer to individuals at high-risk (defined by DAWBA bands = 3, 4, or 5) of, e.g., general band, depression, and panic disorder

\*\*Two-tailed Student's *t* test for continuous variables and chi-square test for categorical variables (likelihood ratio). Significant *p* values are written in bold and italics

Individuals with depression DAWBA band risk scores <15% were considered at low-risk for depression; individuals with depression DAWBA level bands 3 (≈ 15%), 4 (≈ 50%), or 5 (> 70%) were considered at high-risk for depression

DAWBA Development and Well-Being Assessment

genome (adjusted *p* value < 0.05, Fig. 2a, Supplementary Table 1).

We then attempted to replicate our initial 91 SNP-CpG pairs in an independent cohort. Four of the 91 pairs could not be analyzed as they were excluded in the preprocessing procedure. DNA levels at the 87 remaining CpGs were linearly regressed against the previously identified meQTLs-SNPs (dominant genetic model), covarying for age and sex. In total, 64 SNP-CpG pairs were replicated in the independent cohort, with Benjamini-Hochberg (BH)-adjusted *p* values < 0.05 (Table 4, Supplementary Table 2). All but two of the 64 replicated pairs had a similar DNAm direction for SNP risk carriers in both discovery and replication cohorts. Fifty four percent of the 64 replicated CpGs were hypomethylated in carriers of the risk allele for depression (Supplementary Table 2). The rs3823624 SNP within *MAD1L1* was associated with the highest number of CpG sites (*n* = 9 CpGs). The risk allele for depression

(the C allele) was associated with lower DNAm at six CpGs, with the highest effect size at the pair rs3823624-cg19624444 (6.8% DNAm difference in the discovery cohort and 8.6% in the replication cohort between T major allele carriers and C risk allele carriers).

#### Lower DNAm at cg02098413 in whole blood is associated both with depression risk and with the risk allele at rs1933802 in the *HACE1* gene

We next tested whether blood DNAm at the 64 discovered and replicated CpGs was associated with depression risk in the discovery cohort. DNAm levels were available for 221 adolescents, who were categorized based on the DAWBA depression scores as “low-risk” or “high-risk” for depression. The binary variable was regressed against the 64 replicated CpGs together with age, sex, and DNAm batch as covariates (Supplementary Table 3). Higher DNAm at cg02098413 located upstream of HECT Domain and Ankyrin Repeat Containing E3

**Table 2** Characteristics of individuals in NICHD Brain Bank of Developmental Disorders cohort

	MDD (n = 22)	Controls (n = 23)	p value*
Male:Female	12 (54.5):10 (45.5)	12 (52.2):11 (47.8)	1
Age (years)	33.6 ± 16.4	33.4 ± 16.2	0.95

Variables are shown as mean ± standard deviation or number (percentage).

\*Student's *t* test for continuous variables and chi-square test for categorical variables (likelihood ratio)

MDD major depression disorder

**Table 3** Characteristics of individuals in the EXPHIP cohort

	MDD (n = 15)	Controls (n = 15)	p value*
Male:female	9 (60):6 (40)	9 (60):6 (40)	1
Age (years)	59.2 ± 16.7	58.4 ± 15.2	0.68
Race (C:A:Am)	14 (93.3):1 (6.7)	11 (73.3):4 (26.7)	0.32

Variables are shown as mean ± standard deviation or number (percentage).

\*Chi-square test for categorical variables (likelihood ratio)

MDD major depression disorder, C Caucasians, AAm African American

**Table 4** Validated meQTL-CpG pairs

meQTL SNP	Chromosome	Associated gene	No. of paired CpGs*
rs301799	1	<i>RERE</i>	2
rs1890946	1	<i>NRDC</i>	1
rs72710803	1	<i>AL122019</i>	2
rs7585722	2	<i>RNF103-CHMP3</i>	2
rs4346585	3	<i>SOC55P3, ZKSCAN7-AS1</i>	3
rs13084037	3	<i>KLHDC8B</i>	2
rs7624336	3	<i>TKT</i>	1
rs6783233	3	<i>LOC101926953</i>	1
rs7685686	4	<i>HTT</i>	5
rs34937911	4	<i>SLC30A9</i>	1
rs60157091	5	<i>KIF2A</i>	1
rs200949	6	<i>HIST1H1B</i>	7
rs7758630	6	<i>LOC107984041</i>	1
rs1933802	6	<i>HACE1--[]--LIN28B</i>	2
rs3823624	7	<i>MAD1L1</i>	9
rs2043539	7	<i>TMEM106B</i>	2
rs7030813	9	<i>PAX5</i>	1
rs2670139	9	<i>DENND1A</i>	1
rs198457	11	-	4
rs7117514	11	<i>SHANK2</i>	1
rs3213572	12	<i>SPPL3</i>	1
rs1343605	13	<i>OLFM4--[]</i>	1
rs4772087	13	<i>STK24</i>	1
rs1045430	14	<i>AREL1</i>	1
rs62091461	18	<i>RAB27B</i>	1
rs7241572	18	<i>KCNG2</i>	6
rs12624433	20	<i>SLC12A5</i>	1
rs5995992	22	<i>RBX1--[]--EP300</i>	3

\*Validation analysis was run by applying lm models between the methylation level and genetic dominant model, adjusting for age and sex. The Benjamini-Hochberg multiple-testing adjustment was applied to the  $p$  values. A CpG site was paired with the meQTL SNP if the adjusted  $p$  value < 0.05. meQTL methylation quantitative trait locus, SNP single nucleotide polymorphism

Ubiquitin Protein Ligase 1 (*HACE1*) was associated with lower depression score ( $p_{\text{raw}} = 0.014$ , odds ratio (OR) = 0.23) (Fig. 2b). Lower DNAm at cg02098413 was also associated with the risk allele at rs1933802, in both the discovery (Fig. 2c) and replication cohorts ( $p_{\text{adj.}} = 0.0029$  and 0.0063, respectively) (Supplementary Table 2). Each risk allele (C allele) was associated with a 1% decrease in DNAm at cg02098413. The MAF in adolescents at low-risk for depression was 0.54, compared with 0.42 in those at high-risk for depression.

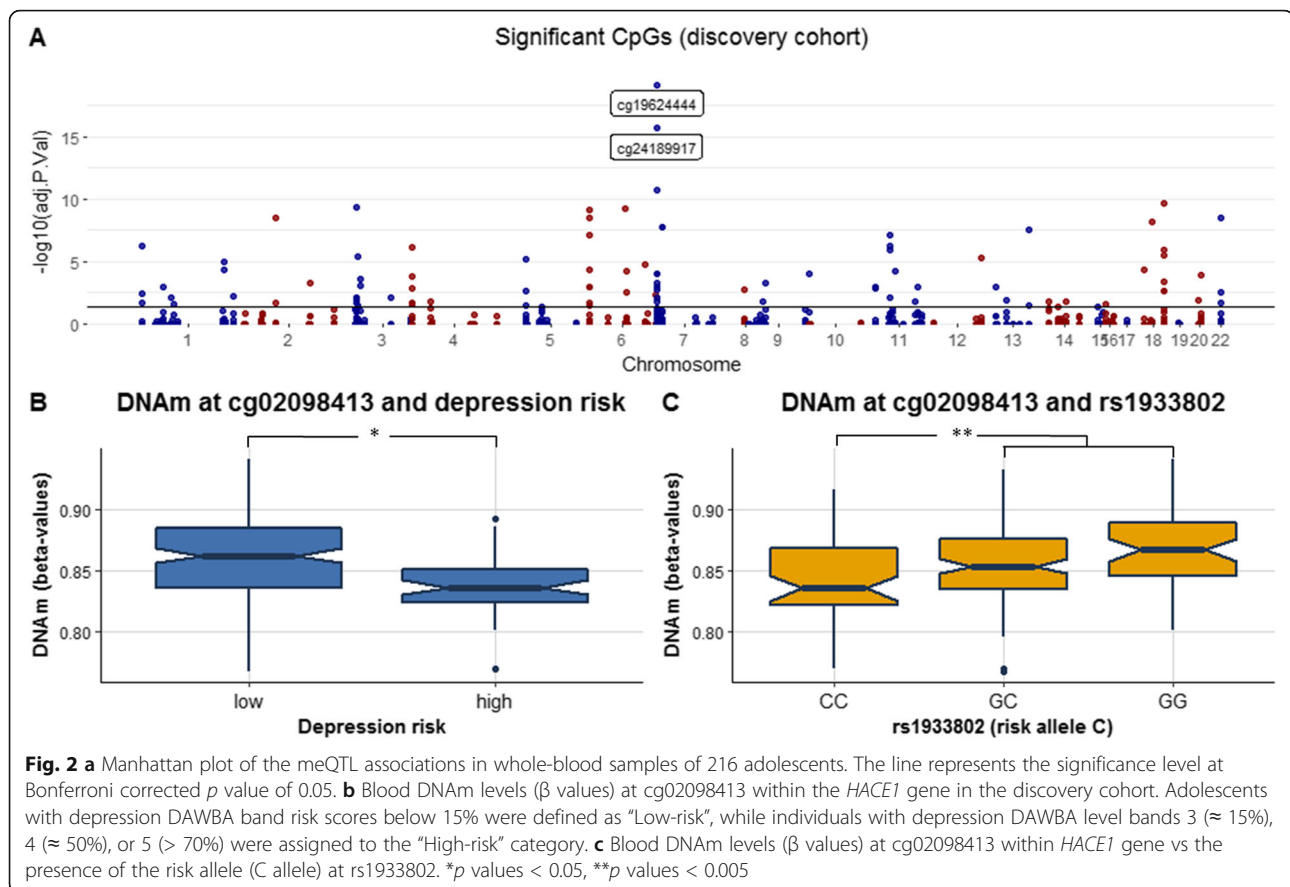
#### DNAm at cg02098413 is positively associated with *HACE1* gene expression in whole blood

DNAm levels at cg02098413 and *HACE1* mRNA levels (transcript ID: ILMN\_1740217) were investigated in 1202 CD14+ samples. A 15% increase in DNAm at cg02098413

was associated with a one unit change in *HACE1* mRNA expression ( $p = 0.00060$ ) (Supplementary Figure 2).

#### Increased DNAm at cg02098413 and reduced expression of *HACE1* in the adult brain are associated with depression

Exploration of chromatin states of this locus revealed that cg02098413 is located next to the active chromatin region in all seven investigated brain regions and peripheral blood mononuclear primary cells (PBMC) (Fig. 3). DNAm level at cg02098413 in blood was positively correlated with DNAm in Brodmann area (BA) 10 region (prefrontal cortex), but not to BA20 or BA7 regions. In addition, rs1933802 is a meQTL for cg02098413 in brain ( $p = 1.97 \times 10^{-05}$ , Spearman  $r = 0.19$ ). Transcriptome data across tissues (GTEx) and at a single-cell resolution



in the human analyses indicated that *HACE1* gene is highly expressed in the brain, particularly in neurons, followed by astrocytes (Supplementary Figure 3).

We investigated the relevance of DNAm at cg02098416 in the brain of 22 adults with MDD and 23 matched controls from the NICHD Brain Bank of Developmental Disorders. DNAm was available for both neuronal and glial cells, and adjusted for age and sex. DNAm at cg02098413 in neurons was positively associated with depression ( $p = 0.008$ , difference of 1% between MDD and controls), and we found the opposite trend in glial cells ( $p = 0.072$ ) (Fig. 4a, b).

We next examined *HACE1* expression in the postmortem hippocampal tissue of patients diagnosed with MDD compared with matched controls. CA1 regions from MDD individuals displayed lower mRNA levels of *HACE1* compared with controls ( $p = 0.012$ ). No difference in *HACE1* expression was identified in the dentate gyrus region ( $p = 0.6$ ).

#### Four of the 102 SNPs are eQTL for nearby ncRNA

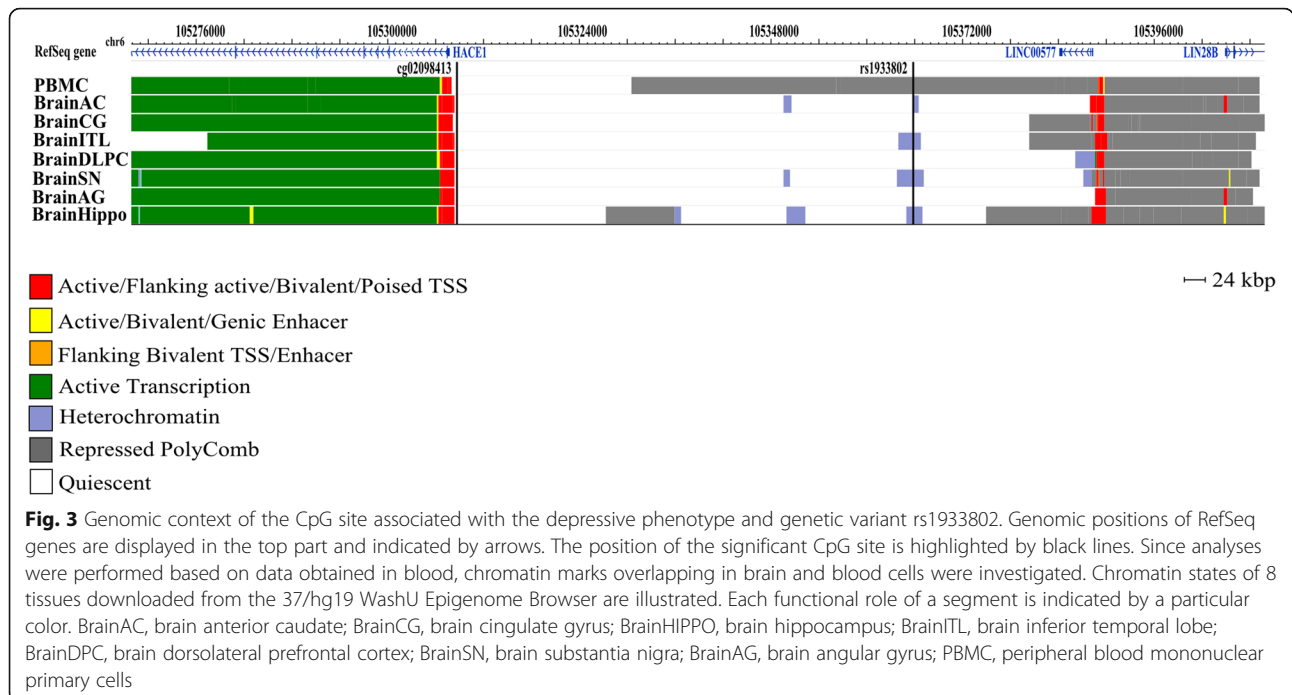
In addition to DNAm, we investigated the link between depression-related SNPs and expression of nearby ncRNAs ( $\pm$  500 kbp window around the SNP locus) in 74 adolescents from the discovery cohort. ncRNA

expression levels were regressed against the SNP, adjusting for sex and age. Out of 229 unique SNP-ncRNA tested pairs, four were significant ( $p_{\text{adj.}} < 0.05$ , Table 5), the most significant association being rs7117514-miR-3664-5p (Fig. 5a).

To infer the biological relevance of miRNA expression changes, we performed gene ontology (GO) analysis of the predicted target genes for the identified miRNAs, hsa-miR-3664-5p, hsa-miR-488-5p, and hsa-miR-6728-3p. Enriched GO terms associated with hsa-miR-3664-5p target genes were behavioral fear and defense response, presynaptic signal transductions, and presynaptic active zone organization; GO term related with hsa-miR-6728-3p target genes was social behavior, synapse organization, locomotory behavior, and muscle adaptation to stimulus, learning, or memory (Fig. 5b). No significant GO terms had gene set enrichment  $>$  10% coverage for hsa-miR-488-5p. Pathway analysis indicated neuronal system and intracellular signaling as potential processes implicating the target genes (Supplementary Table 2).

#### Discussion

We addressed the potential mechanisms underlying genetic risk for depression by intersecting several molecular



layers, namely, epigenetic and transcriptional, with genetic profiles. We investigated the association between genetic variations and DNAm as well as ncRNA expression in diverse datasets, including adolescents at high risk for depression and adults diagnosed with MDD. Among the replicated meQTLs, DNAm at cg02098413 located in the promoter of *HACE1* was associated with rs1933802. In addition, this CpG in whole blood and brain was associated with depression in two independent cohorts. This differential methylation could have a functional impact as *HACE1* expression levels were lower in CA1 hippocampal samples of depressed adults. Therefore, we showed cumulative evidence for rs1933802 to play a regulatory role in depression etiology via differential DNAm at cg02098413. Lastly, two GWAS SNPs were eQTL for four small ncRNAs implicated in neurobiological processes.

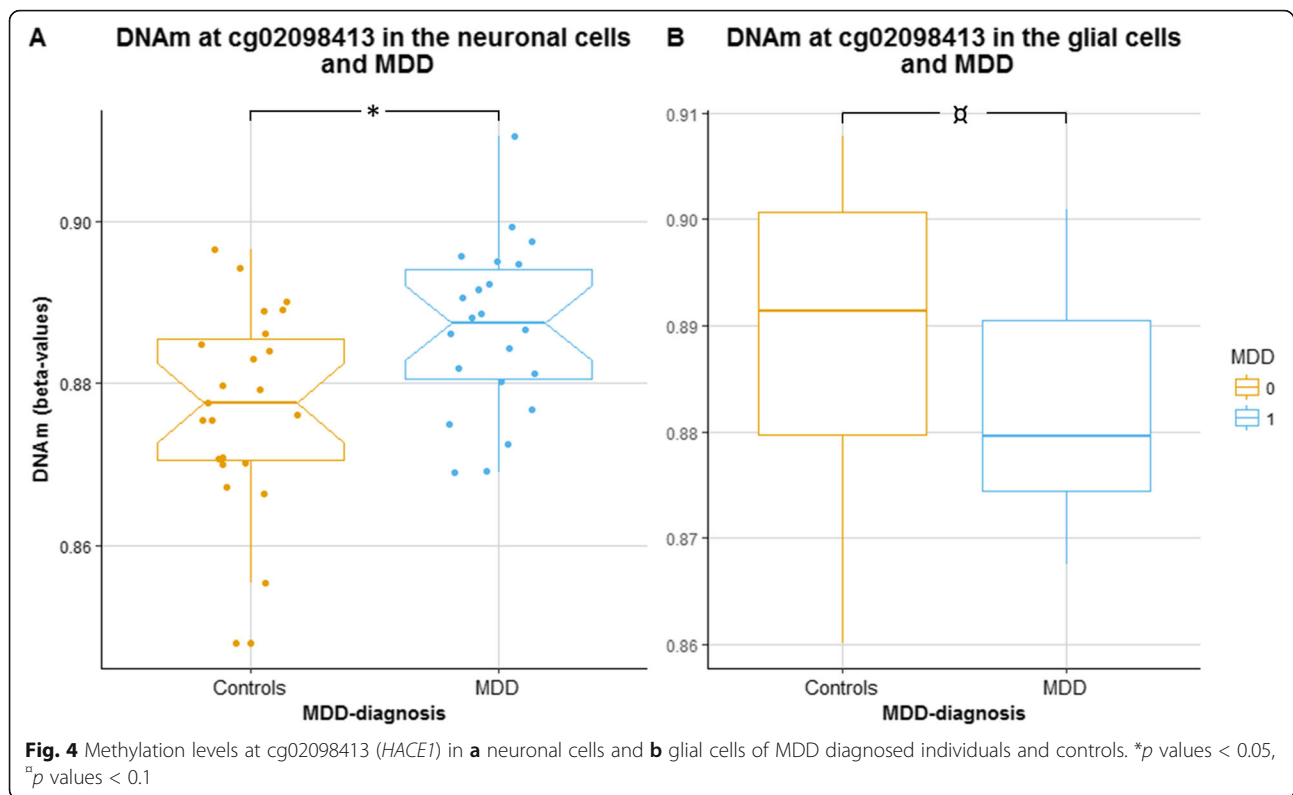
Approximately, half of the 102 associated SNPs with depression were found to be *cis*-meQTL, indicating that they may give susceptibility to depression via effects on local DNAm. We found that the risk allele at rs1933802

was associated with lower DNAm at cg02098413 in blood. In contrast, the risk allele was associated with higher DNAm levels in brain tissue. Additionally, DNAm at cg02098413 in whole blood was associated with depression risk, and DNAm at cg02098413 in the brain was associated with MDD. Interestingly, our data revealed the tissue specificity of genetically dependent DNAm levels in relation to MDD, i.e., lower DNAm in whole blood and higher DNAm in the prefrontal cortex. We hypothesize that rs1933802 exerts tissue- or cell type-specific effect on proximal DNAm levels, as meQTLs are coordinated by variation in transcription factor binding to specific genes, which is tissue-specific [22, 23]. Moreover, one cannot exclude the possibility of additional tissue- or cell type-specific mechanisms involved in the observed DNA methylation changes. In line with this, DNAm levels can vary within tissues, including across different brain regions [24], and we identified differential DNAm in neurons specifically, as opposed to glial cells, of MDD patients. Thus, even though we found a positive correlation between DNAm

**Table 5** Significant SNP-ncRNA pairs after the Bonferroni multiple-testing adjustment

SNP	RefGene	Position (hg19)	ncRNA	Host gene	Bonferroni-adj. <i>p</i> value	% ncRNA mRNA levels difference
rs7117514	<i>SHANK2</i>	11:70544937	hsa-miR-3664-5p	<i>SHANK2</i>	0.015	5.2
rs72710803	<i>AL122019</i>	1:177428018	hsa-miR-488-5p	<i>ASTN1</i>	0.026	10
rs301799	<i>RERE</i>	1:8489302	hsa-miR-6728-3p	<i>ENO1</i>	0.032	- 8.9
rs7659414	-	4:177350956	ENSG00000201516	<i>WDR17</i>	0.037	- 3.9

Linear models with ncRNA expression as outcome and SNP, age, and sex as predictors. The difference in mRNA levels is shown between major allele carriers and heterozygous and homozygous for minor allele  
 SNP single nucleotide polymorphism, ncRNA non-coding RNA



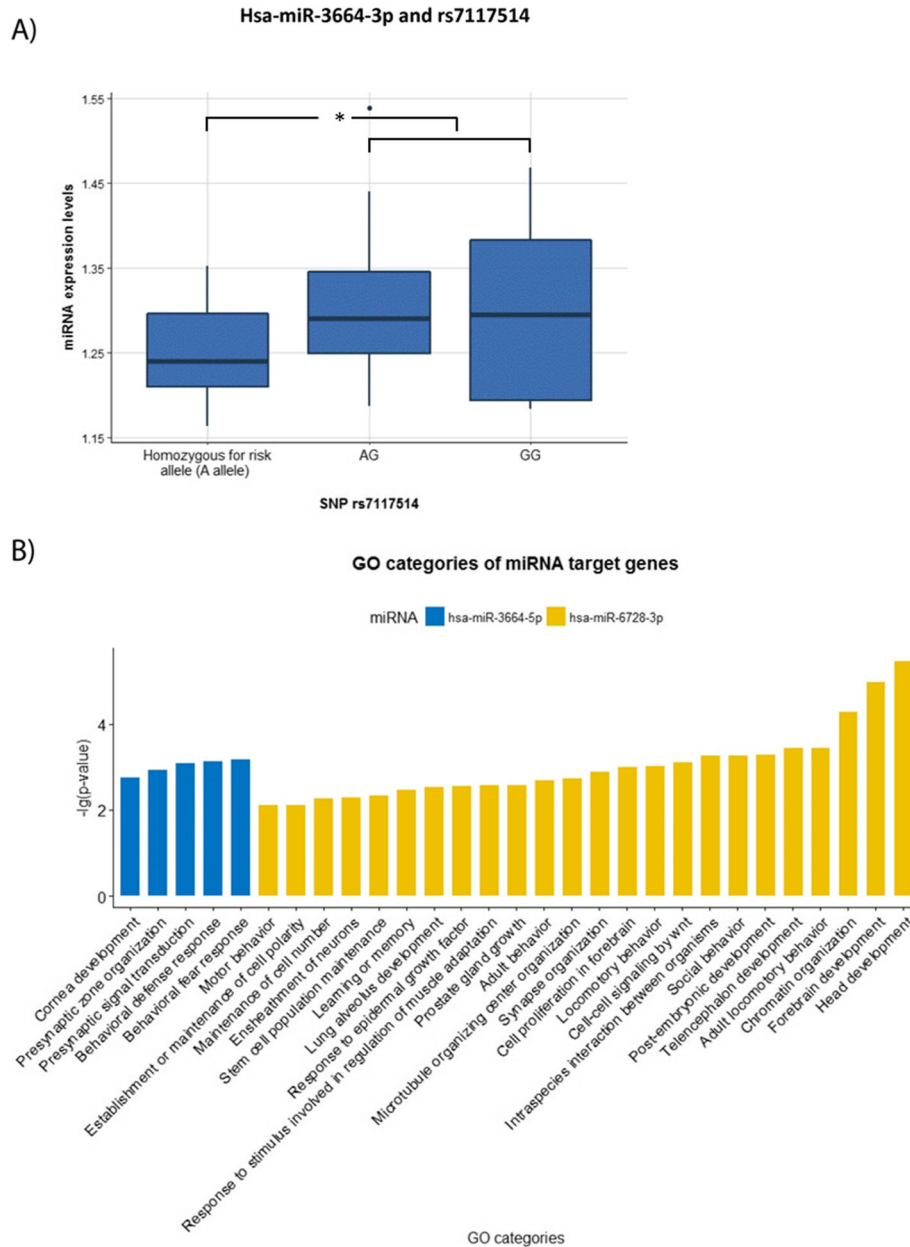
levels in BA10 within the prefrontal cortex and blood, we cannot assume the direction of the correlation between blood and the entire prefrontal cortex region.

DNAm differences at cg02098413 may have functional consequences at the cellular and molecular levels (e.g., differential transcriptional effect or splicing across tissues). The positive relationship between DNAm and *HACE1* expression we found in CD14+ cells cannot be extrapolated to a similar direction in the brain. On the contrary, the high DNAm and low *HACE1* expression observed in MDD brain, although identified in distinct brain regions and cohorts, concurs to the conventional view of anti-correlated promoter DNAm and transcriptional changes. This is supported by the diverse functions of *HACE1* in cellular processes such as tumorigenesis, oxidative stress response, and neurodevelopmental disorders. Importantly, dysfunctional *HACE1* has been implicated in structural and functional brain defects [25, 26]. Previous findings in *Hace1* knock-out mice showed that *HACE1* is thought to modulate reactive oxygen species (ROS) via its target protein RAC1 [27]. Our results are in line with the growing body of evidence highlighting the involvement of ROS in relation to depression [28, 29] and suicidal behavior [30]. Moreover, oxidative stress, resulting from an altered production of oxidant and antioxidant reactions, can negatively affect dendritic arbors and spines. Our findings suggest that lower *HACE1* expression levels in the CA1

hippocampal region of adults diagnosed with major depression might lead to dendritic damage via increased ROS. Reduced *HACE1* levels might rely, at least partly, on DNAm as MDD-diagnosed individuals had higher DNAm levels in prefrontal cortex compared with matched controls.

In our study, the expression of three miRNAs was associated with a differential expression of their host genes. The rs7117514 SNP within *SHANK2* was associated with levels of hsa-miR-3664-5p that is located in the intronic region of the gene. Recently, hsa-miR-3664-5p was found to suppress the proliferation of gastric cancer through its target MTDH [31]. While little is known about the functional role of this miRNA, especially in a disease context, previous studies highlighted the importance of intronic miRNAs as their expression may be coordinated with the host gene's expression [32, 33], as well as functional processes involving regulation of the hosts [34] or genes acting on related pathways [35, 36]. Hsa-miR-3664-5p host gene *SHANK2* encodes the largest isoform of a family of scaffold proteins and is found abundantly at the postsynaptic site of excitatory synapses in the CNS [37]. Accordingly, *SHANK2* is thought to act as an initial "organizer" of the postsynaptic density, which further recruits *SHANK3*. Human genetic studies have implicated aberrant expression of *SHANK2* in phenotypes such as intellectual disability, schizophrenia, and bipolar disorders [38–40]. Moreover,





**Fig. 5 a** Hsa-miR-3664-5p expression levels and rs7117514 genotype in blood samples of 74 adolescents. **b** Gene ontology (GO) categories for the target genes of the microRNAs hsa-miR-3664-5p and hsa-miR-6728-3p. \**p* values < 0.05

*SHANK2* and *SHANK3* have been shown to be causative genes for autism spectrum disorders [41, 42], possibly through the regulation of *RAC1* function [43]. Importantly, a decrease in SHANK2 protein at hippocampal CA1 synapses leads to dysregulation of *N*-methyl-d-aspartate receptors (NMDARs) in mice [44]. This is of interest since NMDAR plays a critical role in the fast synaptic glutamate neurotransmission, and its dysfunction contributes to the pathophysiology of depression [45]. A new generation of pharmacological treatments

for depression includes NMDAR antagonists, such as ketamine [46].

In our study, hsa-miR-6728-3p was differentially expressed in carriers of the rs301799 allele. This intronic miRNA is located in *ENO1* that has been previously related to depression and diabetes [47]. Interestingly, GO analysis of hsa-miR-6728-3p target genes revealed processes related to motor behavior, regulation of muscle adaptation in response to stimulus, and locomotory behavior. Major depression entails significant disturbances

in psychomotor function [48], such as slower motor response times [49] and decreased overall motor activity [50]. Two other GO terms relate to social behavior and learning, supporting the evidence that the genetic dependent miRNA expression could be involved in the negative response in social stimuli processing [51] and memory [52] associated with depression.

Our results are strengthened by our comprehensive bioinformatics analysis that is integrating multiple lines of evidence point at a role for *HACE1* and *SHANK2* in the pathophysiology of depression. We combined five independent cohorts and multiple online tools to investigate the biological relevance of genetics, DNAm, ncRNA, and mRNA expression in depression. However, the eQTL analysis of ncRNA has not been replicated in an independent cohort. In addition, depression risk in the discovery cohort was assessed by the self-reported DAWBA questionnaire, without being validated by a clinician. The heterogeneous condition of depression is more likely to be endorsed by more individuals than those that would meet the diagnostic criteria for MDD. It should be mentioned that the DAWBA questionnaire was validated in a Brazilian sample [53]. Another limitation of our study was that age ranges and sex distribution across cohorts were different. However, while genetics appears to be especially important in regulating DNAm levels, other covarying factors were taken into account in our analyses. Importantly, the statistical models were corrected for age and sex, enabling us to adjust for the effect of age and sex on DNAm. Moreover, meQTLs were shown to be stable over the time [54], where most of age-related DNAm changes accumulate faster during the adolescence than in adulthood [55]. Sex differences in DNAm provide a potential biologic mechanism to explain the increased risk of depression in females [56]. Therefore, additional studies are required to further ascertain the findings.

## Conclusions

In conclusion, our study explores potential molecular mechanisms and biological processes underlying causal genetic variants in depression. We present a genotype-dependent DNAm at the *HACE1* gene that associates with depression in brain and whole blood across diverse datasets, specifically adolescents at risk for depression and adults diagnosed with MDD. Moreover, we found altered gene expression in neuronal cells of depressed adults, which may lead to dendrite density alteration. Furthermore, *hsa-miR-3664-5p* was differentially expressed in risk allele carriers of a SNP in *SHANK2*. Altered expression of miRNA and its host may contribute to the pathophysiology of depression via dysregulation of NMDARs. Together, our results underscore DNAm,

ncRNA expression and genetic risk factors as an interplay modulating mechanism behind depression.

## Methods

### Discovery cohort of the meQTL analyses - DNAm, SNP imputation, ncRNA and depression risk

The discovery cohort of 221 individuals was a sample from 786 unrelated Swedish adolescents aged 17–19, recruited between 2012 and 2017 in Uppsala, Sweden. Subjects were followed up approximately 1 year after enrollment, with the aim to investigate potential psychiatric risk factors in youth. Individuals with any severe cognitive impairment were excluded. Out of the 221 adolescents, six participants were under medication for depression. Height, age, and medication intake were self-reported, but body weight and height were measured and used for BMI calculation. The risk for depression was assessed using the DAWBA web-based interview designed for individuals aged 5–17 years to generate Diagnostic and Statistical Manual of Mental Disorders IV and International Statistical Classification of Diseases and Related Health Problems-10 based diagnoses [57]. Percentage of risk for depression was categorized in six band scores, i.e., 0 (< 0.1%), 1 ( $\approx$  0.5 %), 2 ( $\approx$  3%), 3 ( $\approx$  15%), 4 ( $\approx$  50%), and 5 (> 70%). The DAWBA depression band consists of questions related to the occurrence of depression symptoms in the last 4 months (e.g., level of sadness, irritability, sleep activity, wish to die, and to what extent general mood is affected). Individuals with a score  $\geq$  3 are considered at high risk for depression [58].

All 786 were genotyped using the Infinium Global Screening Array at the SNP&SEQ SciLife Platform at Uppsala University, interrogating 700,078 genetic variants. SNPs were imputed as previously described [59], and all 102 selected SNPs had imputation scores > 0.8. Blood DNAm profiles of the 221 adolescents were assessed with the Illumina 450 K methylation Beadchip. Data preprocessing was performed as described in [20, 58]. It should be noted that the methylomes were profiled at two different time points within a time range of 1 year [20]. This time difference in DNAm assessment at baseline is further referred to as “batch” and taken into consideration in the meQTL analyses. We performed a *cis*-meQTL analysis of CpG sites within  $\pm$  500 kbp of each SNP.

A subset of participants ( $n = 74$ ) had available ncRNA measurements. The RNA measurements were performed with the Affymetrix<sup>®</sup> miRNA 4.1 Array Plates, using 120  $\mu$ l of the prepared biotinylated RNA from each sample. The arrays were then hybridized, washed, stained, and scanned with the GeneTitan<sup>®</sup> Multi-Channel instrument. The raw data were normalized in Expression Console provided by Affymetrix, using the robust multi-

array average method [60]. A *cis*-eQTL analysis ncRNAs within  $\pm 500$  kbp of each SNP was performed.

#### Replication cohort of the mQTL analyses - DNAm and SNP imputation

Participants in this study ( $n = 423$ ) are healthy individuals of white European descent aged 18–34 years, recruited in Uppsala, Sweden. The replication cohort consisted of a subset of 219 individuals free of medication for psychiatric disorders, who had available both genetic and DNAm data in blood. Blood samples were taken in the fasted state.

All participants were genotyped using the Infinium Global Screening Array, interrogating 700,078 genetic variants. Prior to imputation, quality control steps were performed using Plink 1.9 [61]: samples were excluded based on genotyping call rate ( $< 95\%$ ), high heterozygosity  $> 2$  standard deviations (SD), and missingness rate  $> 0.05$  ( $n = 8$  samples); probes were removed based on MAF ( $< 1\%$ ), Hardy-Weinberg equilibrium (HWE) ( $< 1e-6$ ), missingness rate ( $> 5\%$ ), and strand ambiguousness ( $n = 219,868$  probes). A total of 480,210 genotypes for 415 individuals were included in the pre-phasing approach implemented in SHAPEIT version 2.7 (r904) [62] and then imputed using the IMPUTE2 software (version 2.3.2) [63], using the 1000 Genome Project phase 3 integrated variant set as a reference (release October 2014).

Preprocessing of the DNAm data included background correction (“noob” method) [64], data normalization (all-sample mean normalization method) [65], and sample and probe exclusion. No samples were excluded based on the detection  $p < 5e-5$ . We excluded probes with missing  $\beta$  values, probes having  $\leq 75\%$  of samples with detection  $p < 0.01$ , and probes located on the sex chromosomes. Using the annotation generated by Chen et al. [66], cross-reactive probes and probes containing SNPs with minor allele frequency  $> 1\%$  in European populations were removed as well. In total, 397,620 probes remained. Batch effects were corrected with the ComBat function [67]. All data processing was performed using R version 3.1.3.

#### Brain DNAm profiles of MDD-diagnosed individuals and controls from the NICHD Brain Bank of Developmental Disorders

We tested associations between the newly discovered meQTLs and depression in a publicly available, processed DNAm dataset of 58 postmortem cortical tissues from 29 individuals with MDD, and 29 age- and sex-matched controls. In these samples, neuronal and glial cell proportions were measured by fluorescence activated cell sorting, and DNAm levels were separately analyzed in neurons and glia (GSE41826). To avoid confounding of DNAm signals by ethnicity, only data

from Caucasian individuals were included in further analyses ( $n = 22$  MDD diagnosed individuals and  $n = 23$  controls) [68].

#### Whole blood DNAm and mRNA expression from the Multi-Ethnic Study of Atherosclerosis (MESA)

We investigated the functional relevance of the discovered meQTLs by testing associations between DNAm levels at the meQTLs and mRNA expression in an open-access dataset (E-GEOD-56047). Methylomes were profiled with the Illumina 450 K methylation Beadchip and transcriptomes with the Illumina HumanHT-12 V4.0 expression Beadchip. Details about DNAm data preprocessing and transcriptome quality control can be found elsewhere [69]. All 1202 participants had information available on age (44–83 years old), race/ethnicity, study site, sex, DNAm/expression chip, and proportions of B-cells, T-cells, neutrophils, and natural killer cell.

#### Brain mRNA levels in two hippocampal regions of MDD-diagnosed and matched controls from EXPHIP

To evaluate differential mRNA expression of target genes with depression, we compared mRNA expression levels in 15 patients with MDD with 15 age-, sex-, and race-matched controls in the dentate gyrus and CA1 subregions of the hippocampus (ArrayExpress accession E-GEOD-24095). Ten adults diagnosed with MDD were taking anti-depressants, while the controls were not under medication treatment. Data was analyzed using the 48 K human HEEBO whole-genome microarrays and further followed by normalization using linear-log normalization to stabilize the variance of low expressing genes and removal of spatial and intensity related biases, using the MAANOVA library [70].

#### Functional annotation of meQTLs

We used data from the Roadmap Epigenomics Project (WashU Epigenome Browser <http://epigenomegateway.wustl.edu/legacy/>, human genome version hg19) to determine chromatin states at the discovered meQTL in seven brain tissues (brain angular gyrus (BrainAG), brain anterior caudate (BrainAC), brain cingulate gyrus (BrainCG), brain hippocampus (BrainHIPPO), brain inferior temporal lobe (BrainITL), brain substantia nigra (BrainSN), brain dorsolateral prefrontal cortex (BrainDLPC)) and PBMC. For easier visualization, the 18-state model was reduced to five states, i.e., active/flanking active/bivalent/poised TSS, active/bivalent/genic enhancer, flanking bivalent TSS/enhancer, active transcription, and repressed Polycomb state.

We used correlated DNAm between blood and three brain regions (BA10 (prefrontal cortex), BA20 (temporal cortex), and BA7 (parietal cortex)) in 16 individuals, using Spearman correlations in the online BECon tool

<https://redgar598.shinyapps.io/BECon/> [71]. More details about the preprocessing of this DNA methylation dataset are available in [71]. Moreover, the meQTL relationship was assessed in the brain using the xQTL server [72].

Single cell RNA-seq profiles from Darmanis et al. [73] were used to investigate RNA expression levels of candidate genes in cell clusters, such as neurons, astrocytes, microglia, oligodendrocytes, OPC, and endothelial cells. Expression values were log-transformed and centered using the mean.

We determined the target genes of the identified miRNAs using microT-CDS release 5.0 [74]. GO biological process categories and Reactome cell signaling pathways with a gene set enrichment > 10% were explored using Consensus path [75].

### Statistical analyses

We used linear models as implemented in the *limma* package [76] to find meQTL in the discovery cohort, with DNAm levels as the dependent variable and SNP (dominant genetic model with wild-type coded with “0” and heterozygous and homozygous for the minor allele coded with “1”), age, sex, and batch as independent variables. CpGs were deemed significant at a Bonferroni-adjusted  $p$  value < 0.05.

We used linear regression to find eQTL in the discovery cohort, with ncRNA expression levels as the dependent variable and SNP dominant model, age, and sex as independent variables. The Bonferroni multiple-testing adjustment was considered for each analyzed SNP.

In an independent replication cohort, linear regressions were applied on the identified CpGs against the SNP dominant model. Analyses were adjusted for age and sex. A BH-adjusted  $p$  value < 0.05 was considered significant. In the cohort from NICHD Brain Bank of Developmental Disorders, DNAm levels of significant CpG sites were regressed against the yes/no depression diagnosis, adjusting for age and sex. Analyses were run separately for neurons and glia cells. Using the MESA cohort, associations between DNAm and mRNA expression levels at the associated genes were investigated. DNAm data were firstly regressed against age, race, study site, gender, methylation chip, and the sample contamination with non-targeted cells, i.e., non-monocytes. The resulting residuals were included in linear regression models as an independent variable, together with expression values as dependent variables. Covariates were age, race, study site, gender, expression chip, and the sample contamination with non-targeted cells. Expression data of the EXPHIP cohort were analyzed using a two-sided binomial test, assuming a 0.5 probability threshold.

### Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13148-020-00884-8>.

**Additional file 1: Figure S1.** BMI distribution in the discovery and replication cohorts.

**Additional file 2: Figure S2.** DNAm at cg02098413 and expression levels of *HACE1* in 1,202 CD14+ samples.

**Additional file 3: Figure S3.** RNA expression of *HACE1* in cell types of blood and brain.

**Additional file 4: Figure S1-S4.**

### Abbreviations

GWAS: Genome-wide association studies; SNP: Single nucleotide polymorphism; DNAm: DNA methylation; ncRNA: Non-coding RNA; microRNA: miRNA; meQTL: Methylation quantitative trait locus; eQTL: Expression quantitative trait locus; DAWBA: Development And Well-Being Assessment; BMI: Body mass index; MDD: Major depression disorder; MESA: Multi-Ethnic Study of Atherosclerosis; BH: Benjamini-Hochberg; *HACE1*: HECT Domain and Ankyrin Repeat Containing E3 Ubiquitin Protein Ligase 1; MAF: Minor allele frequency; TSS: Transcription start site; BA: Brodmann area; GO: Gene ontology; ROS: Reactive oxygen species; NMDAR: *N*-Methyl-d-aspartate receptor; HWE: Hardy-Weinberg equilibrium

### Acknowledgements

Genotyping was performed by the SNP&SEQ Technology Platform in Uppsala. The platform is part of the Science for Life Laboratory at Uppsala University and is supported as a national infrastructure by the Swedish Research Council.

### Authors' contributions

DMC and HBS conceived and designed the study. DMC analyzed and interpreted the data and wrote the manuscript. JM interpreted the data and helped writing the manuscript. SV interpreted the results and helped analyzing and writing the manuscript. LK and MRA helped writing the manuscript and analyzing the data. JJ helped at interpretation of the results. HBS helped at the data interpretation. The authors read and approved the final manuscript.

### Funding

The studies were supported by the Swedish Research Foundation, the Åhlens Foundation, the Novo Nordisk Foundation, and the Swedish Brain Research Foundation. LK was supported by a fellowship from the Margaretha af Ugglas Foundation. Open access funding provided by Uppsala University.

### Availability of data and materials

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

### Ethics approval and consent to participate

The study was approved by the Regional Ethics Committee in Uppsala, corresponding to the Dnr 2011/446. All participants gave their written informed consent to participate.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Department of Neuroscience, Functional Pharmacology, Uppsala University, BMC, Box 593, Husargatan 3, 753124 Uppsala, Sweden. <sup>2</sup>Institute for Health and Sport (iHeS), Victoria University, Footscray, VIC 3011, Australia. <sup>3</sup>Department of Clinical Neuroscience, Center for Molecular Medicine, Karolinska Institutet, 171 76 Stockholm, Sweden. <sup>4</sup>Department of Immunology, Genetic and Pathology, Uppsala University, Uppsala, Sweden. <sup>5</sup>Institute for Translational Medicine and Biotechnology, Sechenov First Moscow State Medical University, Moscow, Russia.

Received: 17 April 2020 Accepted: 11 June 2020

Published online: 02 July 2020

## References

- Organization WH. Depression and other common mental disorders. 2017.
- Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry*. 2000;157(10):1552–62.
- Ormel J, Hartman CA, Snieder H. The genetics of depression: successful genome-wide association studies introduce new challenges. *Transl Psychiatry*. 2019;9:114.
- Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*. 2019;22(3):343–52.
- Sultan FA, Day JJ. Epigenetic mechanisms in memory and synaptic function. *Epigenomics*. 2011;3(2):157–81.
- Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, et al. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci*. 2010;13(4):423–30.
- Henikoff S, Gready JM. Epigenetics, cellular memory and gene regulation. *Curr Biol*. 2016;26(14):R644–8.
- Zhu H, Wang G, Qian J. Transcription factors as readers and effectors of DNA methylation. *Nat Rev Genet*. 2016;17(9):551–65.
- Jones MJ, Moore SR, Kobor MS. Principles and challenges of applying epigenetic epidemiology to psychology. *Annu Rev Psychol*. 2018;69:459–85.
- McRae AF, Powell JE, Henders AK, et al. Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol*. 2014;15:R73.
- Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.
- Xia Y, Dai R, Wang K, Jiao C, Zhang C, Xu Y, et al. Sex-differential DNA methylation and associated regulation networks in human brain implicated in the sex-biased risks of psychiatric disorders. *Mol Psychiatry*. 2019.
- Abdolmaleky HM, Smith CL, Faraone SV, Shafa R, Stone W, Glatt SJ, et al. Methyloomics in psychiatry: Modulation of gene-environment interactions may be through DNA methylation. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*. 2004;127b(1):51–9.
- Szyf M, Weaver IC, Champagne FA, Diorio J, Meaney MJ. Maternal programming of steroid receptor expression and phenotype through DNA methylation in the rat. *Front Neuroendocrinol*. 2005;26(3-4):139–62.
- Saab BJ, Mansuy IM. Neuroepigenetics of memory formation and impairment: the role of microRNAs. *Neuropharmacology*. 2014;80:61–9.
- Schratt G. microRNAs at the synapse. *Nat Rev Neurosci*. 2009;10(12):842–9.
- Olejniczak M, Kotowska-Zimmer A, Krzyzosiak W. Stress-induced changes in miRNA biogenesis and functioning. *Cellular and molecular life sciences : CMLS*. 2018;75(2):177–91.
- Bjornsson HT, Fallin MD, Feinberg AP. An integrated epigenetic and genetic approach to common human disease. *Trends in genetics : TIG*. 2004;20(8):350–8.
- Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H, et al. Intra-individual change over time in DNA methylation with familial clustering. *Jama*. 2008;299(24):2877–83.
- Ciuculete DM, Bostrom AE, Voisin S, Philipps H, Titova OE, Bandstein M, et al. A methylome-wide mQTL analysis reveals associations of methylation sites with GAD1 and HDAC3 SNPs and a general psychiatric risk score. *Transl Psychiatry*. 2017;7(1):e1002.
- Voisin S, Almén MS, Zheleznyakova GY, Lundberg L, Zarei S, Castillo S, et al. Many obesity-associated SNPs strongly associate with DNA methylation changes at proximal promoters and enhancers. *Genome medicine*. 2015;7:103.
- Banovich NE, Lan X, McVicker G, van de Geijn B, Degner JF, Blischak JD, et al. Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. *PLoS Genet*. 2014;10(9):e1004663.
- Sonawane AR, Platig J, Fagny M, Chen CY, Paulson JN, Lopes-Ramos CM, et al. Understanding tissue-specific gene regulation. *Cell Rep*. 2017;21(4):1077–88.
- Ladd-Acosta C, Pevsner J, Sabuncyan S, Yolken RH, Webster MJ, Dinkins T, et al. DNA methylation signatures within the human brain. *Am J Hum Genet*. 2007;81(6):1304–15.
- Nagy V, Hollstein R, Pai TP, Herde MK, Buphamalai P, Moeseneder P, et al. HACE1 deficiency leads to structural and functional neurodevelopmental defects. *Neurology Genetics*. 2019;5(3):e330.
- Ehrnhoefer DE, Southwell AL, Sivasubramanian M, Qiu X, Villanueva EB, Xie Y, et al. HACE1 is essential for astrocyte mitochondrial function and influences Huntington disease phenotypes in vivo. *Hum Mol Genet*. 2018;27(2):239–53.
- Rotblat B, Southwell AL, Ehrnhoefer DE, Skotte NH, Metzler M, Franciosi S, et al. HACE1 reduces oxidative stress and mutant Huntingtin toxicity by promoting the NRF2 response. *Proc Natl Acad Sci U S A*. 2014;111(8):3032–7.
- Palta P, Samuel LJ, Miller ER 3rd, Szanton SL. Depression and oxidative stress: results from a meta-analysis of observational studies. *Psychosom Med*. 2014;76(1):12–9.
- Grases G, Colom MA, Fernandez RA, Costa-Bauza A, Grases F. Evidence of higher oxidative status in depression and anxiety. *Oxidative Med Cell Longev*. 2014;2014:430216.
- Vargas HO, Nunes SO, Pizzo de Castro M, Bortolasci CC, Sabbatini Barbosa D, Kaminami Morimoto H, et al. Oxidative stress and lowered total antioxidant status are associated with a history of suicide attempts. *J Affect Disord*. 2013;150(3):923–30.
- Jiao Y, Yang H, Qian J, Gong Y, Liu H, Wu S, et al. miR36645P suppresses the proliferation and metastasis of gastric cancer by attenuating the NFkappaB signaling pathway through targeting MTDH. *Int J Oncol*. 2019;54(3):845–58.
- Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. *Genome Res*. 2004;14(10a):1902–10.
- Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA (New York, NY)*. 2005;11(3):241–7.
- Dill H, Linder B, Fehr A, Fischer U. Intronic miR-26b controls neuronal differentiation by repressing its host transcript, *ctdsp2*. *Genes Dev*. 2012;26(1):25–30.
- Hinske LC, Galante PA, Kuo WP, Ohno-Machado L. A potential role for intragenic miRNAs on their hosts' interactome. *BMC Genomics*. 2010;11:533.
- Lutter D, Marr C, Krumsiek J, Lang EW, Theis FJ. Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. *BMC Genomics*. 2010;11:224.
- Sheng M, Kim E. The Shank family of scaffold proteins. *J Cell Sci*. 2000;113(Pt 11):1851–6.
- Noor A, Lionel AC, Cohen-Woods S, Moghimi N, Rucker J, Fennell A, et al. Copy number variant study of bipolar disorder in Canadian and UK populations implicates synaptic genes. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*. 2014;165b(4):303–13.
- Homann OR, Misura K, Lamas E, Sandrock RW, Nelson P, McDonough SI, et al. Whole-genome sequencing in multiplex families with psychoses reveals mutations in the SHANK2 and SMARCA1 genes segregating with illness. *Mol Psychiatry*. 2016;21(12):1690–5.
- Costas J. The role of SHANK2 rare variants in schizophrenia susceptibility. *Mol Psychiatry*. 2015;20(12):1486.
- Berkel S, Marshall CR, Weiss B, Howe J, Roeth R, Moog U, et al. Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nat Genet*. 2010;42(6):489–91.
- Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, et al. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet*. 2007;39(1):25–7.
- Sarowar T, Grabrucker S, Fohr K, Mangus K, Eckert M, Bockmann J, et al. Enlarged dendritic spines and pronounced neophobia in mice lacking the PSD protein RICH2. *Molecular brain*. 2016;9:28.
- Pappas AL, Bey AL, Wang X, Rossi M, Kim YH, Yan H, et al. Deficiency of Shank2 causes mania-like behavior that responds to mood stabilizers. *JCI insight*. 2017;2:20.
- Cull-Candy SG, Leszkiewicz DN. Role of distinct NMDA receptor subtypes at central synapses. *Science's STKE : signal transduction knowledge environment*. 2004;2004(255):re16.
- Pochwat B, Nowak G, Szewczyk B. An update on NMDA antagonists in depression. *Expert Rev Neurother*. 2019;19(11):1055–67.
- Jiang G, Ma Y, An T, Pan Y, Mo F, Zhao D, et al. Relationships of circular RNA with diabetes and depression. *Sci Rep*. 2017;7(1):7285.

48. Canbeyli R. Sensorimotor modulation of mood and depression: an integrative review. *Behav Brain Res.* 2010;207(2):249–64.
49. Schwartz F, Carr AC, Munich RL, Glauber S, Lesser B, Murray J. Reaction time impairment in schizophrenia and affective illness: the role of attention. *Biol Psychiatry.* 1989;25(5):540–8.
50. Wolff EA 3rd, Putnam FW, Post RM. Motor activity and affective illness. The relationship of amplitude and temporal distribution to changes in affective state. *Arch Gen Psychiatry.* 1985;42(3):288–94.
51. Bistricky SL, Ingram RE, Atchley RA. Facial affect processing and depression susceptibility: cognitive biases and cognitive neuroscience. *Psychol Bull.* 2011;137(6):998–1028.
52. Dillon DG, Pizzagalli DA. Mechanisms of memory disruption in depression. *Trends Neurosci.* 2018;41(3):137–49.
53. Fleitlich-Bilyk B, Goodman R. Prevalence of child and adolescent psychiatric disorders in southeast Brazil. *J Am Acad Child Adolesc Psychiatry.* 2004;43(6):727–34.
54. Gaunt TR, Shihab HA, Hemani G, Min JL, Woodward G, Lyttleton O, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biol.* 2016;17:61.
55. Kananen L, Marttila S, Nevalainen T, Kummola L, Junttila I, Mononen N, et al. The trajectory of the blood DNA methylome ageing rate is largely set before adulthood: evidence from two longitudinal studies. *Age (Dordr).* 2016;38(3):65.
56. Uddin M, Sipahi L, Li J, Koenen KC. Sex differences in DNA methylation may contribute to risk of PTSD and depression: a review of existing evidence. *Depress Anxiety.* 2013;30(12):1151–60.
57. Goodman A, Heiervang E, Collishaw S, Goodman R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Soc Psychiatry Psychiatr Epidemiol.* 2011;46(6):521–32.
58. Ciuculete DM, Bostrom AE, Tuunainen AK, Sohrabi F, Kular L, Jagodic M, et al. Changes in methylation within the STK32B promoter are associated with an increased risk for generalized anxiety disorder in adolescents. *J Psychiatr Res.* 2018;102:44–51.
59. Ciuculete DM, Voisin S, Kular L, Welihinda N, Jonsson J, Jagodic M, et al. Longitudinal DNA methylation changes at MET may alter HGF/c-MET signalling in adolescents at risk for depression. *Epigenetics.* 2019:1–18.
60. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A.* 2001;98(1):31–6.
61. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559–75.
62. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods.* 2012;10:5.
63. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 2009;5(6):e1000529.
64. Triche TJ Jr, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res.* 2013;41(7):e90.
65. Decker A, Yousefi P. *asnm: All sample mean normalization.* R package version 1.0.0. 2014.
66. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8(2):203–9.
67. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics (Oxford, England).* 2012;28(6):882–3.
68. Guintivano J, Aryee MJ, Kaminsky ZA. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics.* 2013;8(3):290–302.
69. Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat Commun.* 2014;5:5366.
70. Duric V, Banasr M, Stockmeier CA, Simen AA, Newton SS, Overholser JC, et al. Altered expression of synapse and glutamate related genes in post-mortem hippocampus of depressed subjects. *Int J Neuropsychopharmacol.* 2013;16(1):69–82.
71. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry.* 2017;7(8):e1187.
72. Ng B, White CC, Klein HU, Sieberts SK, McCabe C, Patrick E, et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nat Neurosci.* 2017;20(10):1418–26.
73. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A.* 2015;112(23):7285–90.
74. Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, et al. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res.* 2013;41(Web Server issue):W169–W173.
75. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res.* 2013;41(Database issue):D793–800.
76. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

