DNA Methylation Signatures of Depressive Symptoms in Middle-aged and Elderly Persons

Meta-analysis of Multiethnic Epigenome-wide Studies

Olivera Story Jovanova, PhD; Ivana Nedeljkovic, MD; Derek Spieler, MD; Rosie M. Walker, PhD; Chunyu Liu, PhD; Michelle Luciano, PhD; Jan Bressler, PhD; Jennifer Brody, BA; Amanda J. Drake, PhD; Kathryn L. Evans, PhD; Rahul Gondalia, MPH; Sonja Kunze, PhD; Brigitte Kuhnel, MSc; Jari Lahti, PhD; Rozenn N. Lemaitre, PhD; Riccardo E. Marioni, PhD; Brenton Swenson, MSc; Jayandra Jung Himali, PhD; Hongsheng Wu, PhD; Yun Li, PhD; Allan F. McRae, PhD; Tom C. Russ, PhD; James Stewart, MSc; Zhiying Wang, MSc; Karl-Heinz Ladwig, PhD; Andre G. Uitterlinden, PhD; Xiuling Guo, PhD; Annette Peters, PhD; Katri Räikkönen, PhD; John M. Starr, PhD; Melanie Waldenberger, PhD; Naomi R. Wray, PhD; Eric A. Whitson, MD; Nona Sotoodehnia, MD; Sudha Seshadri, MD; Daniel Levy, MD; Lifang Hou, PhD; Johan G. Eriksson, PhD; Myriam Fornage, PhD; Ian J. Deary, PhD; Andrea Baccarelli, PhD; Karl-Heinz Ladwig, PhD; Henning Tiemeier, PhD; Najaf Amin, PhD

IMPORTANCE Depressive disorders arise from a combination of genetic and environmental risk factors. Epigenetic disruption provides a plausible mechanism through which gene-environment interactions lead to depression. Large-scale, epigenome-wide studies on depression are missing, hampering the identification of potentially modifiable biomarkers.

OBJECTIVE To identify epigenetic mechanisms underlying depression in middle-aged and elderly persons, using DNA methylation in blood.

DESIGN, SETTING, AND PARTICIPANTS To date, the first cross-ethnic meta-analysis of epigenome-wide association studies (EWAS) within the framework of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium was conducted. The discovery EWAS included 7948 individuals of European origin from 9 population-based cohorts. Participants who were assessed for both depressive symptoms and whole-blood DNA methylation were included in the study. Results of EWAS were pooled using sample-size weighted meta-analysis. Replication of the top epigenetic sites was performed in 3308 individuals of African American and European origin from 2 population-based cohorts.

MAIN OUTCOMES AND MEASURES Whole-blood DNA methylation levels were assayed with Illumina-Infinium Human Methylation 450K BeadChip and depressive symptoms were assessed by questionnaire.

RESULTS The discovery cohorts consisted of 7948 individuals (4104 [51.6%] women) with a mean (SD) age of 65.4 (5.8) years. The replication cohort consisted of 3308 individuals (2456 [74.2%] women) with a mean (SD) age of 60.3 (6.4) years. The EWAS identified methylation of 3 CpG sites to be significantly associated with increased depressive symptoms: cg04987734 (P = 1.57 x 10^-08; n = 11 256; CDC42BPB gene), cg12325605 (P = 5.24 x 10^-09; n = 11 256; ARHGEF3 gene), and an intergenic CpG site cg14023999 (P = 5.99 x 10^-08; n = 11 256; chromosome = 15q26.1). The predicted expression of the CDC42BPB gene in the brain (basal ganglia) (effect, 0.14; P = 2.7 x 10^-03) and of ARHGEF3 in fibroblasts (effect, -0.48; P = 9.8 x 10^-04) was associated with major depression.

CONCLUSIONS AND RELEVANCE This study identifies 3 methylated sites associated with depressive symptoms. All 3 findings point toward axon guidance as the common disrupted pathway in depression. The findings provide new insights into the molecular mechanisms underlying the complex pathophysiology of depression. Further research is warranted to determine the utility of these findings as biomarkers of depression and evaluate any potential role in the pathophysiology of depression and their downstream clinical effects.

Published online July 11, 2018. Corrected on August 22, 2018.

© 2018 American Medical Association. All rights reserved.
Depression is one of the most common mental health disorders that is projected to play a leading role in disease burden by 2030. In later life, depression is associated with disability, increased mortality, dementia, and poor outcomes from physical illness. Furthermore, more people older than 65 years commit suicide than in any other age group, and most have major depression. Limited understanding of the molecular mechanisms underlying depression is a hindrance in the development of innovative treatment, prognostic markers, and prevention strategies.

Studying depression is challenging because it is a heterogeneous disorder with a multifactorial etiology. This heterogeneity increases with age as the incidence of chronic diseases and disability rises. The contribution of genetics to the risk of depression is moderate, with heritability estimates ranging from 40% to 50% in the elderly. Genome-wide association studies (GWAS) have identified numerous rare and common genetic variants associated with depression and related traits. However, genetic variation alone does not completely explain an individual’s risk for developing depression. Among environmental factors, adverse life events and stress are major risk factors for depression.

Converging evidence from animal and human studies suggests that psychosocial stressors trigger depression onset by inducing elevations in proinflammatory cytokine levels. These psychosocial stressors are also known to influence epigenetic mechanisms, such as DNA methylation, that can drive sustained changes in gene expression. The high contribution of environmental factors to depression in the elderly makes DNA methylation a candidate mechanism for studies of the molecular basis of late-life depression.

DNA methylation may be global or tissue specific. Tissues likely to be involved in complex psychiatric disorders, such as brain, are not directly accessible from living patients. The use of postmortem brain tissue to study DNA methylation is a possible solution, although obtaining a sufficient sample size is challenging. To study differential DNA methylation associated with mental health symptoms on a large scale, peripheral tissues, such as blood, constitute a useful proxy for detecting trans-tissue changes and the most appropriate tissue for biomarkers. Moderate correlation has been demonstrated between blood and brain tissues at nontissue-specific regulatory regions across the methylome.

To date, several studies have assessed the correlation between depression and DNA methylation. However, these studies have been limited to a small number of DNA methylation sites (CpG sites) and/or small samples. For instance, the largest published epigenome-wide association study (EWAS) assessed brain DNA methylation in 76 persons who died during a depressive episode and 45 controls. Moreover, these studies compared depressed cases with healthy controls. Depression represents an arbitrarily selected extreme of the continuum of varying severity and duration, whereas a broad phenotype approach can be more representative for the general population. A large study consisting of 252,503 individuals from 68 countries showed that subthreshold depressive disorders produce significant decrements in health and do not qualitatively differ from full-blown episodes of depression.

A meta-analysis in individuals older than 55 years found 2 to 3 times higher prevalence of subthreshold depressive symptoms than major depression. Use of rating scales has therefore been recommended for the assessment of depressive problems in the elderly.

In the present investigation, we performed an EWAS of depressive symptoms using whole-blood samples of 7948 individuals of European ethnicity from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. We replicated our findings in 3308 individuals of African American and European ancestry. Finally, we used publicly available expression quantitative methylation loci and expression quantitative loci databases to identify the downstream effects of the associated methylation loci.

## Methods

### Study Population

The study sample for the discovery analysis included a total of 7948 participants of European ancestry from 9 population-based cohorts of the CHARGE consortium (Table 1). Cardiovascular Health Study, Framingham Heart Study, Helsinki Birth Cohort Study, Cooperative Health Research in the Augsburg Region study, 2 subcohorts from Lothian Birth Cohort (LBC) born in 1921 (LBC1921) and 1936 (LBC1936), and Generation Scotland: Scottish Family Health Study. These cohorts included community-dwelling individuals who were not selected based on disease status. Informed consent was obtained from all participants. The same cohorts have been used successfully to identify differentially methylated sites associated with cognitive traits, inflammation, and smoking. The protocol for each study was approved by the institutional review board of each institution.

The replication sample included 3308 participants, largely of African American origin from the Atherosclerosis Risk in Communities Study and European origin from the Women’s Health Initiative—Epigenetic Mechanisms of PM-Mediated Cardiovascular Disease that joined the consortium.
Depression Subscale, which consists of 7 items (range, 0-21, t janapsychiatry.com

Table 1. Descriptive Statistics of the Discovery and Replication Cohorts

<table>
<thead>
<tr>
<th>Study</th>
<th>Ethnicity</th>
<th>No.</th>
<th>Women, No. (%)</th>
<th>Age, Mean (SD), y</th>
<th>Current Smoker, No. (%)</th>
<th>Depressive Symptoms (No. of Scale Items)</th>
<th>Antidepressant Medication Use, No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHS25</td>
<td>European</td>
<td>323</td>
<td>194 (60.1)</td>
<td>75.6 (5.2)</td>
<td>173 (53.6)</td>
<td>CES-D26 (10)</td>
<td>19 (5.9)</td>
</tr>
<tr>
<td>FHS27</td>
<td>European</td>
<td>2722</td>
<td>1460 (53.6)</td>
<td>58.5 (11.6)</td>
<td>948 (34.8)</td>
<td>CES-D28 (20)</td>
<td>251 (9.2)</td>
</tr>
<tr>
<td>HBCS29</td>
<td>European</td>
<td>122</td>
<td>0</td>
<td>65.2 (2.7)</td>
<td>24 (19.7)</td>
<td>CES-D28 (20)</td>
<td>11 (9.0)</td>
</tr>
<tr>
<td>KORA30</td>
<td>European</td>
<td>1727</td>
<td>882 (51.1)</td>
<td>61.0 (8.9)</td>
<td>250 (14.5)</td>
<td>PHQ-931</td>
<td>82 (4.7)</td>
</tr>
<tr>
<td>LBC192132</td>
<td>European</td>
<td>432</td>
<td>261 (60.4)</td>
<td>79.1 (0.6)</td>
<td>194 (44.9)</td>
<td>HADS33</td>
<td>15 (3.5)</td>
</tr>
<tr>
<td>LBC193614</td>
<td>European</td>
<td>916</td>
<td>452 (49.3)</td>
<td>69.6 (0.8)</td>
<td>504 (55.0)</td>
<td>HADS33</td>
<td>30 (3.3)</td>
</tr>
<tr>
<td>RS-III35</td>
<td>European</td>
<td>722</td>
<td>391 (54.2)</td>
<td>59.8 (8.1)</td>
<td>167 (23.1)</td>
<td>CES-D28 (20)</td>
<td>38 (5.3)</td>
</tr>
<tr>
<td>RS-BIOS35</td>
<td>European</td>
<td>757</td>
<td>319 (42.1)</td>
<td>67.6 (5.9)</td>
<td>78 (10.3)</td>
<td>CES-D28 (20)</td>
<td>51 (6.7)</td>
</tr>
<tr>
<td>GS36a</td>
<td>European</td>
<td>227</td>
<td>145 (63.9)</td>
<td>52.4 (8.1)</td>
<td>46 (20.3)</td>
<td>SCID37</td>
<td>37 (16.3)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>7948</td>
<td>4104 (51.6)</td>
<td>65.4 (5.8)</td>
<td>2384 (30.0)</td>
<td>534 (6.7)</td>
<td></td>
</tr>
<tr>
<td>Replication</td>
<td></td>
<td>3308</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIC38</td>
<td>African</td>
<td>2297</td>
<td>1445 (62.9)</td>
<td>56.1 (5.7)</td>
<td>584 (25.4)</td>
<td>21-MQ39</td>
<td>74 (3.2)</td>
</tr>
<tr>
<td>WHI-EMPC40</td>
<td>European</td>
<td>1011</td>
<td>1011 (100)</td>
<td>64.6 (7.1)</td>
<td>509 (50.3)</td>
<td>CES-D/DIS41</td>
<td>61 (6.0)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3308</td>
<td>2456 (74.2)</td>
<td>60.3 (6.4)</td>
<td>1093 (33.0)</td>
<td>135 (4.1)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ARIC, Atherosclerosis Risk in Communities Study; CES-D/DIS, Centre for Epidemiologic Studies Depression scale/Diagnostic Interview Schedule; CHS, Cardiovascular Health Study; FHS, Framingham Heart Study; GS, Generation Scotland: Scottish Family Health Study; HADS, Hospital Anxiety and Depression Scale; HBCS, Helsinki Birth Cohort Study; KORA, Cooperative Health Research in the Augsburg Region; LBC, Lothian Birth Cohort; MQ, Maastricht Questionnaire; PHQ, Patient Health Questionnaire; RS, Rotterdam Study; SCID, Structured Clinical Interview for DSM-IV Disorders; WHI-EMPC, Women’s Health Initiative-Epigenetic Mechanisms of PM-Mediated Cardiovascular Disease. 

Table 1. Descriptive Statistics of the Discovery and Replication Cohorts

were obtained for DNA methylation quantification (Table 1).26,28,31,33,37,39,41 Four cohorts (Framingham Heart Study, Helsinki Birth Cohort Study, RS-III, and RS-BIOS) assessed depressive symptoms using the 20-item Centre for Epidemiologic Studies Depression (CES-D) scale,45 and the Cardiovascular Health Study used the 10-item CES-D scale. Participants could score from 0 to 60 (or 30 for Cardiovascular Health Study) points, where higher scores suggest more depressive symptoms. Women’s Health Initiative-Epigenetic Mechanisms of PM-Mediated Cardiovascular Disease used a cohort-specific instrument that combines 6 items from the CES-D scale and 2 items from the Diagnostic Interview Schedule. The score on this instrument ranges from 0 to 42 (higher scores indicate more exhaustion). In all cohorts, depressive symptoms were analyzed as continuous variables, except for Generation Scotland: Scottish Family Health Study, which studied depression status as binary trait.

**DNA Methylation Sample and Measurement**

In all cohorts, DNA was extracted from whole blood and methylation levels were assessed (Illumina-Infinium Human Methylation 450K BeadChip; Illumina Inc) using standard manufacturer’s protocols. The 450K array includes more than 450 000 CpGs and is enriched for genic regions, covering 99% of all genes. DNA methylation data preprocessing, including quality control and normalization, was conducted per cohort using study-specific methods. In all cohorts, DNA methylation levels were quantified as β values, which range from 0 to 1 and indicate the proportion of DNA strands in a sample methylated at a specific CpG. Detailed information about cohort-specific DNA extraction, bisulfite conversion, DNA methylation profiling, normalization, and quality control is described in the eAppendix in the Supplement.

**Statistical Analysis**

**Epigenome-wide Association Analysis**

In all cohorts, the association between depressive symptoms and CpG sites was assessed using linear regression analysis. In the regression analysis, DNA methylation β value at each
CpG site was specified as the dependent variable and the depressive symptoms or depression as the predictor of interest. Association analysis was adjusted for age, sex, smoking, and total cholesterol (assessed at the time of blood sampling for methylation), methylation batch effects, white blood cell composition (imputed or directly measured), principal components estimated using genome-wide genotype data to control for population stratification, and familial relationships when required. Cohort-specific details of these analyses are provided in the eAppendix in the Supplement. Furthermore, sensitivity analysis was performed by adjusting the initial model for antidepressant medication use at the time of sample collection.

To pool the results from independent studies, we performed sample size–weighted meta-analysis in METAL.49 We chose the sample size–weighted method because of the differences in the measurement scales of depressive symptoms across studies. A drawback of using the sample size–weighted method is that no pooled-effect estimates are generated. To obtain pooled-effect estimates, we additionally performed inverse variance–weighted meta-analysis for the top sites in cohorts that used the CES-D 20-item scale for the assessment of depressive symptoms. Any CpG sites missing in more than 3 of the participating cohorts were removed. In total, 484516 probes were tested for association, giving a Bonferroni-corrected genome-wide significance threshold of 0.05/484516 = 1.03 × 10−7. All CpG sites suggestive of association (P ≤ 10−5) were tested for association in the independent replication cohorts using the same model as used in the discovery EWAS. Finally, a sample size–weighted meta-analysis was performed for all cohorts included in the discovery and replication phases in METAL. To evaluate the contribution of each study to the association results, we generated posterior probabilities of the effects in each study (M-values) using the METASOFT package.50 M-value and forest plots for z scores were generated using custom-made scripts in R. For annotating CpG sites, we used the annotation provided by Illumina and the University of California Santa Cruz UCSC database (GRCh37/hg19).

Causal Inference Analysis
To help infer causal associations, we studied the cis single-nucleotide polymorphisms (SNPs) identified by the BIOS consortium51 as instrumental variables for the CpG sites as proposed by Relton and Davey Smith.57 We checked the association of these cis-SNPs with depression, smoking, and inflammation in the published GWAS of these traits. Similarly, we checked whether the SNPs associated with inflammation (C-reactive protein levels), smoking, and depression were associated with the identified CpG sites using the BIOS consortium database. We chose smoking and inflammation because these are highly correlated with both depression and DNA methylation and thus could influence the association between depression and DNA methylation.

Results
The mean (SD) age in the discovery cohorts ranged from 52.4 (8.1) years in Generation Scotland: Scottish Family Health Study to 79.1 (0.57) years in LBC1921. Of the total discovery sample, 4104 (51.6%) were women. The replication cohort comprised 2456 (74.2%) women and had a mean age of 60.3 (6.4) years (Table 1).

Epigenome-wide Association Analysis
In the meta-analysis of depressive symptoms of European ancestry, we identified 1 CpG site on chromosome 14q32.32 (cg04987734, CDC42BPB [OMIM 614062]; P = 4.93 × 10−08; n = 7948) that passed the Bonferroni threshold for significance (Table 2; eFigure 1 in the Supplement). Furthermore, suggestive association was observed at 19 additional CpG sites (Table 2). Adjusting for antidepressive medication use did not meaningfully change the results (eTable 1 in the Supplement).

Meta-analysis of discovery and replication cohorts showed nominal association (P < 0.05; n = 3308) with depressive symptoms in the validation data set (Table 2). In addition, significant association of a CpG site (cg12325605; P = 9.17 × 10−05; n = 3308) was annotated to the ARHGEPF3 [OMIM 612115] gene with depressive symptoms was observed in the replication sample.

Meta-analysis of discovery and replication cohorts showed a significant association of both cg04987734 (P = 1.57 × 10−08; n = 11 256) and cg12325605 (P = 5.24 × 10−05; n = 11 256) with depressive symptoms (Table 2, Figure 1 and Figure 2). An additional intergenic CpG site (cg14023999; P = 5.99 × 10−08; n = 11 256) at chromosome 15q26.1 locus showed genome-wide significant association with depressive symptoms (eTable 2, eFigure 3, and eFigure 4 in the Supplement). The independent contributions of each cohort to the association signals of the 3 CpGs are depicted in eFigure 5 in the Supplement and also provided in eTable 3 in the Supplement. For all 3 CpG sites, the association signals were not driven by a single cohort, but appeared to be linearly related to the sample size, suggesting stronger association in larger studies (eFigure 5 in the Supplement). Pooled-effect estimates in cohorts that used the CES-D
scale suggest that a 1-unit increase in the CES-D score increases methylation by 0.05% at cg04987734, 0.04% at cg12325605, and 0.03% at cg14023999.

**Gene Expression Analyses**

There was a significant association between cg04987734 and increased expression of the CDC42BPB gene (false discovery rate, $P = 7.7 \times 10^{-04}$; $n = 2101$) and cg14023999 was significantly associated with decreased expression of SEMA4B (OMIM 617029) (false discovery rate, $P = 4.7 \times 10^{-03}$; $n = 2101$) in blood (eTable 4 in the Supplement). No significantly associated gene expression probes were identified for cg12325605 in blood. Furthermore, the predicted expression of the CDC42BPB gene in the brain (basal ganglia) (effect, 0.14; $P = 2.7 \times 10^{-03}$) and of ARHGEF3 in fibroblasts (effect, $-0.48; P = 9.8 \times 10^{-04}$) was associated with major depression (eTable 5 in the Supplement). No association was observed with either smoking or inflammation.

**Blood and Brain Correlation**

We checked the correlation between methylation in blood and various brain regions at the 3 identified sites using a web-based tool (BECon) and a blood-brain DNA methylation comparison tool (http://epigenetics.essex.ac.uk/bloodbrain/). BECon showed correlation between blood and brain DNA methylation, for example, methylation at cg04987734 in the CDC42BPB gene was highly correlated ($r = 0.81$) between blood and the Brodmann area 7 that spans the medial and lateral walls of the parietal cortex (eFigure 6 in the Supplement). Methylation at the other 2 sites was negatively correlated with methylation in the Brodmann area 10 that spans anterior prefrontal cortex (cg12325605, $r = -0.39$; cg14023999, $r = -0.42$), suggesting strong but reverse methylation patterns in blood and brain (eFigure 7 and eFigure 8 in the Supplement). However, the blood-brain DNA methylation comparison tool that contrasts DNA methylation between blood and prefrontal cortex, entorhinal cortex, superior temporal gyrus, and cerebellum showed only modest correlations. For instance, methylation in blood at cg04987734 showed the strongest correlation with methylation in superior temporal gyrus ($r = 0.16$; http://epigenetics.essex.ac.uk/bloodbrain/?probenameg=cg04987734), while methylation in blood at cg12325605 (http://epigenetics.essex.ac.uk/bloodbrain/?probenameg=.cg12325605) and cg14023999 (http://epigenetics.essex.ac.uk/bloodbrain/?probenameg=cg14023999) showed the strongest correlation with methylation in cerebellum ($r = 0.16$; $r = 0.19$, respectively). Nevertheless, the findings from the 2 databases suggest some degree of correlation between methylation in blood and methylation in brain for the 3 identified CpG sites.

**Causal Inference**

In the BIOS database we identified 2 cis-SNPs for cg04987734, 4 cis-SNPs for cg12325605 (eTable 6 in the Supplement), and none for cg14023999. We took the most significant cis-SNP as the proxy for the CpG sites if available. For cg04987734, we used rs751837 as a proxy and for cg12325605 we used rs3821412 as a proxy (top cis-SNP rs9880416 was not available in the GWAS of depression, smoking, or inflammation). There was suggestive association between rs751837 and major depression ($P = 0.07$; albeit in the opposite direction) (eTable 7 in the Supplement); rs3821412 was not associated with any of the 3 tested phenotypes. None of the SNPs associated with depression, inflammation, or smoking was associated with any of the 3 CpG sites.

### Table 2. Top DNA Methylation Sites Associated With Depressive Symptoms in the Discovery Epigenome-wide Association Studies

| CpG Site ID | Chromosome | Location | Gene Symbol | $P$ Value
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cg04987734</td>
<td>14</td>
<td>101415873</td>
<td>CDC42BPB</td>
<td>$4.93 \times 10^{-04}$</td>
</tr>
<tr>
<td>cg07012687</td>
<td>17</td>
<td>80195180</td>
<td>SLC16A3</td>
<td>$3.47 \times 10^{-07}$</td>
</tr>
<tr>
<td>cg08796240</td>
<td>16</td>
<td>70733832</td>
<td>VAC14</td>
<td>$7.43 \times 10^{-07}$</td>
</tr>
<tr>
<td>cg06096336</td>
<td>2</td>
<td>231989800</td>
<td>PSMD1; HTR2B</td>
<td>$8.06 \times 10^{-07}$</td>
</tr>
<tr>
<td>cg16745930</td>
<td>10</td>
<td>100220809</td>
<td>HPSE2</td>
<td>$1.34 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg09849319</td>
<td>5</td>
<td>14949831</td>
<td>LPCAT1</td>
<td>$1.81 \times 10^{-05}$</td>
</tr>
<tr>
<td>cg17237086</td>
<td>22</td>
<td>40814966</td>
<td>MRL1</td>
<td>$3.44 \times 10^{-08}$</td>
</tr>
<tr>
<td>cg3985718</td>
<td>2</td>
<td>105924245</td>
<td>TGFBRAP1</td>
<td>$3.61 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg21098005</td>
<td>20</td>
<td>44538605</td>
<td>PLTP</td>
<td>$4.36 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg16466652</td>
<td>19</td>
<td>6271960</td>
<td>MLT1</td>
<td>$4.39 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg07884764</td>
<td>11</td>
<td>64107517</td>
<td>CCDC88B</td>
<td>$5.03 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg01541347</td>
<td>7</td>
<td>47292920</td>
<td>FOX1</td>
<td>$5.64 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg02341197</td>
<td>21</td>
<td>34185927</td>
<td>C21orf62</td>
<td>$5.84 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg01947751</td>
<td>3</td>
<td>196728969</td>
<td>Intergenic</td>
<td>$6.23 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg13747876</td>
<td>17</td>
<td>80195402</td>
<td>SLC16A3</td>
<td>$6.32 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg12764201</td>
<td>1</td>
<td>105101223</td>
<td>CORT; APITD1</td>
<td>$7.15 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg08295111</td>
<td>5</td>
<td>133866097</td>
<td>PHF15</td>
<td>$7.87 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg18030453</td>
<td>3</td>
<td>45506216</td>
<td>LAR52</td>
<td>$9.16 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg12325605</td>
<td>3</td>
<td>56810151</td>
<td>ARHGEF3</td>
<td>$9.62 \times 10^{-06}$</td>
</tr>
<tr>
<td>cg23284441</td>
<td>10</td>
<td>7533927</td>
<td>C1orf54; CDH23</td>
<td>$9.69 \times 10^{-06}$</td>
</tr>
</tbody>
</table>

This table shows the top DNA methylation sites associated with depressive symptoms in the discovery epigenome-wide association studies. The sites are ranked by their significance in the discovery phase ($P$ value). The table includes the CpG site ID, chromosome, location, gene symbol, and $P$ values for the discovery and replication phases, as well as the meta-analysis $P$ value.
In this large-scale EWAS of depressive symptoms, we identified methylation at 3 CpG sites (cg04987734, cg12325605, and cg14023999) associated with depressive symptoms in middle-aged and elderly persons. cg04987734 is annotated to the CDC42BPB gene, cg12325605 to the ARHGEF3 gene, and cg14023999 lies in an intergenic region on chromosome 15q26.1 locus. The predicted expression of CDC42BPB and ARHGEF3...

Figure 1. Regional Association Plot for the Top DNA Methylation (CpG) Site cg04987734 – log10 P Value

On the top graph, the x-axis depicts the position in base pair (bp) (hg19) for the entire CDC42BPB gene region. The y-axis indicates the strength of association in terms of negative logarithm of the association P value. Each circle represents a CpG site. Red dashed line within the graph indicates the genome-wide significance threshold. The regulatory information and correlation matrix of other CpG sites in the region with the top hit are shown below the x-axis. Color intensity marks the strength of the correlation and color indicates the direction of the correlation. The figure was made using web-based plotting tool and R-based package CoMET (http://epigen.kcl.ac.uk/comet/). Ensembl is a genome database resource (http://ensemblgenomes.org/). SNP indicates single-nucleotide polymorphism; UCSC, University of California Santa Cruz.

Discussion

In this large-scale EWAS of depressive symptoms, we identified methylation at 3 CpG sites (cg04987734, cg12325605, and cg14023999) associated with depressive symptoms in middle-aged and elderly persons. cg04987734 is annotated to the CDC42BPB gene, cg12325605 to the ARHGEF3 gene, and cg14023999 lies in an intergenic region on chromosome 15q26.1 locus. The predicted expression of CDC42BPB and ARHGEF3...
genes associated with major depression in brain and fibroblasts, respectively.

CDC42BPB (CDC42 binding protein kinase beta) encodes a member of the serine/threonine protein kinase family, which is a downstream effector of CDC42 and plays a role in the regulation of cytoskeleton reorganization, cell migration, and regulation of neurite outgrowth.\(^{58}\) CDC42BPB is highly expressed in the brain (https://www.proteinatlas.org/ENSG00000198752-CDC42BPB/tissue). Hypermethylation of cg04987734 has been associated with increased expression of CDC42BPB in blood.\(^{51}\) Methylation levels at this CpG site (cg04987734) were also previously associated with C-reactive protein levels in blood.\(^{43}\) and smoking.\(^{44}\) In our study, however, we adjusted for smoking in the regression model; therefore, the association...
between depression and DNA methylation of this CpG site may be independent of smoking habits. Also, our causal inference analyses provide no support for the possibility that smoking or inflammation explained the observed association with depressive symptoms or that the predicted expression of the gene showed an association with smoking or inflammation.

The *ARHGEF3* gene encodes for rho guanine nucleotide exchange factor 3 protein. The gene is highly expressed (https://www.proteinatlas.org/ENSG00000163947-ARHGEF3/tissue) in adrenal glands, brain, and uterus. Both *ARHGEF3* and *CDC42BPP* are coexpressed with several members of the rho subfamily (RHOA, RHOB, and RHOC) (eFigure 9 and eFigure 10 in the Supplement) of the rho guanosine triphosphatase family that also includes CDC42.59 The rho family of guanosine triphosphatases is a family of small signaling G-proteins involved in p75 neurotrophin receptor-mediated signaling60 and semaphorin-signaling pathways.61 p75 Neurotrophin receptor is a transmembrane receptor for neurotrophic factors of the neurotrophin family, which includes the brain-derived neurotrophic factor.62 p75 Neurotrophin receptor is widely expressed in the developing central and peripheral nervous systems during the period of synaptogenesis and developmental cell death.63 Both p75 neurotrophin receptor and semaphorins are implicated in axon guidance.64,65 In this context, the third associated CpG site, cg14023999, that lies in an intergenic region on chromosome 15q26.1 is also relevant as it associated with CpG site, cg14023999, that lies in an intergenic region associated with depression. The replication may also have resulted in false-negatives owing to different genetic background. Third, in these analyses we mostly used quantitative measures of depressive symptoms. Quantitative endophenotypes provide powerful alternatives for several complex outcomes, for example, hypertension.75 This is likely to be especially true for a trait such as depressive symptoms, for which the severity and duration of illness can be highly heterogeneous.22 Genome-wide studies of depressive traits, using quantitative endophenotypes, have been suggested to improve statistical power.22 However, the use of different phenotypic measures by different cohorts means that there may be some loss of statistical power owing to the heterogeneity in the phenotype assessment. Nevertheless, the top 3 sites in our study were robustly associated with depressive symptoms independent of the depressive symptom measure used.

Fourth, although we adjusted for potential confounders, the possibility of residual confounding cannot be excluded. Antidepressant medication indicates treated depression but itself may result in epigenetic modifications involved in depression pathophysiology.76 Antidepressants can thus mediate or confound the association between DNA methylation and depression. However, in sensitivity analysis additionally adjusted for antidepressant medication, our results did not change. Fifth, most cohorts included in this EWAS were composed of elderly persons. The cause of depression is more heterogeneous in elderly people than in younger people and often hidden behind somatic symptoms, either because of somatization of the disorder or because of accentuation of symptoms of concomitant illness.77 This heterogeneity may affect the generalizability of the results to younger populations. Finally, we made an attempt to disentangle cause and consequence using SNPs associated with the identified CpG sites and depression, inflammation, and smoking as instrumental variables. The results did not support a causative role, yet the association of the predicted gene expression of *CDC42BPP* in brain and *ARHGEF3* in fibroblasts with major depression does suggest a possible causal role of the regulatory effects of these genes.52

Strengths and Limitations

To our knowledge, this is the largest EWAS of depressive symptoms reported to date. Our major strength is the sample size that enabled detection of a replicable epigenome-wide significant locus, which suggests that in blood, DNA methylation signatures associated with depression may be subtle and will require large samples to be detected.

However, this study has several limitations. First, the study was using peripheral blood tissue for DNA methylation profiling, as DNA methylation is known to be tissue specific.71 Although peripheral blood is not considered to be the most relevant tissue for the pathophysiology of depression, some sites show correlated methylation profiles between tissues.15,71 The 3 sites identified in our study show some degree of correlation between methylation in blood and various brain regions. Second, although replication in African American samples suggests that some depressive symptom-related differences in DNA methylation may be similar across races/ethnicities,72-74 the replication may also have resulted in false-negatives owing to different genetic background. Third, in these analyses we mostly used quantitative measures of depressive symptoms. Quantitative endophenotypes provide powerful alternatives for several complex outcomes, for example, hypertension.75 This is likely to be especially true for a trait such as depressive symptoms, for which the severity and duration of illness can be highly heterogeneous.22 Genome-wide studies of depressive traits, using quantitative endophenotypes, have been suggested to improve statistical power.22 However, the use of different phenotypic measures by different cohorts means that there may be some loss of statistical power owing to the heterogeneity in the phenotype assessment. Nevertheless, the top 3 sites in our study were robustly associated with depressive symptoms independent of the depressive symptom measure used.

Conclusions

To our knowledge, we report the first EWAS of depressive symptoms. We identified and replicated association of 2 methylation sites in the genome with depressive symptoms. A third site was identified in the meta-analysis of discovery and replication cohorts, which requires further replication. All 3 findings point toward axon guidance as the common disrupted pathway in depression (http://www.genome.jp/kegg/pathway/hsa/hsa04360.html). Our findings provide new insights into the molecular mechanisms underlying the complex pathophysiology of depression.
ARTICLE INFORMATION
Accepted for Publication: May 17, 2018.
Correction: This article was corrected on August 22, 2018, to fix an error in the author byline and in the end matter Affiliations.
Published Online: July 11, 2018.

Author Affiliations: Department of Epidemiology, Erasmus MC-University Medical Center Rotterdam, Rotterdam, the Netherlands (Story Jovanova, Nedjeljkovic, Uitterlinden, Tiemeier, Amin); Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany (Spierer, Kunzle, Kühnel, Ladwig, Peters, Waldenberger); Klinik und Poliklinik für Psychosomatische Medizin und Psychotherapie des Klinikums Rechts der Isar der Technischen Universität München, Munich, Germany (Spierer, Ladunic, Baccarelli, van der Meulen, van der Laan); Center for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, United Kingdom (Walker, Luciano, Evans, Marioni, Russ, Starr, Porteous, Mcintosh, Deary); Centre for Genomic and Experimental Medicine, MRC Institute of Genetics and Molecular Medicine, The University of Edinburgh, Western General Hospital, Edinburgh, United Kingdom (Walker, Evans, Porteous); The Framingham Heart Study, Framingham, Massachusetts (Liu, Himali, Seshadri, Mendelson, Levy); The Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, Maryland (Liu, Mendelson, Levy); Boston University School of Public Health, Boston, Massachusetts (Liu, Himali); Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom (Luciano, Marioni, Deary); Human Genetics Center, University of Texas Health Science Center at Austin (Bressler, Brody, Wang, Porteous); Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle (Brodys, Lemaitre, Swenson, Seshadri); University/British Heart Foundation Centre for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, United Kingdom (Drape); Department of Epidemiology, University of North Carolina at Chapel Hill (Gondalia, Stewart, Whitsell); Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany (Kunzle, Kühnel, Peters, Waldenberger); Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland (Lahti, Räikkönen); Institute for Public Health Genetics, School of Public Health, University of Washington, Seattle (Swenson); Boston University School of Medicine (Baccarelli, van der Meulen, van der Laan); Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Queensland, Brisbane, Australia (McIntosh, Wray); Queensland Brain Institute, The University of Queensland, Brisbane, Australia (McIntosh, Wray); Alzheimer Research and Dementia Research Centre, Edinburgh, United Kingdom (Russ, Starr); Centre for Dementia Prevention, University of Edinburgh, Edinburgh, United Kingdom (Russ); Carolina Population Center, University of North Carolina at Chapel Hill (Stewart); Department of Internal Medicine, Erasmus Medical Center, Rotterdam, the Netherlands (Uitterlinden, van Meurs); The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Harbor-University of California Los Angeles (UCLA) Medical Center (Guo); Department of Medicine, University of North Carolina at Chapel Hill (Whitsell); MIND Center, University of Mississippi Medical Center, Jackson (Mosley); Division of Psychiatry, The University of Edinburgh, Royal Edinburgh Hospital, University of Edinburgh, Edinburgh, United Kingdom (Mcdintosh); Department of Cardiology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts (Mendelson); Feinberg School of Medicine, Northwestern University, Chicago, Illinois (Hou); Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland (Eriksson); Institute of Molecular Medicine, University of Texas Health Science Center at Houston (Fornage); Department of Environmental Health Sciences, Harvard T. H. Chan School of Public Health, Harvard University, Boston, Massachusetts (Baccarelli); Department of Child and Adolescent Psychiatry, Erasmus MC-University Medical Center Rotterdam, Rotterdam, the Netherlands (Tiemeier); Department of Social and Behavioral Science, Harvard T. H. Chan School of Public Health, Boston, Massachusetts (Tiemeier).

Author Contributions: Drs Story Jovanova, Nedjeljkovic, Spierer, Walker, Liu, and Luciano were equal contributors to the study. Drs Tiemeier and Amin had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Story Jovanova, Nedjeljkovic, Räikkönen, Starr, Mcintosh, Mendelson, Tiemeier, Amin. Acquisition, analysis, or interpretation of data: Story Jovanova, Nedjeljkovic, Spierer, Walker, Liu, Luciano, Bressler, Brody, Drake, Evans, Gondalia, Kunzle, Kühnel, Lahti, Lemaitre, Marioni, Swenson, Himali, Wu, Li, Mcrae, Russ, Stewart, Wang, Zhang, Ladwig, Uitterlinden, Guo, Peters, Waldenberger, Whitell, Seshadri, Porteous, van Meurs, Mosley, Mcintosh, Mendelson, Levy, Hou, Eriksson, Fornage, Deary, Baccarelli, Tiemeier, Amin.


Other—generated part of the methylation data (KORA): Kunze.

Conflict of Interest Disclosures: None reported.

Additional Contributions: We acknowledge the essential role of the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) Consortium in development and support of this manuscript. CHARGE members include the Netherlands’ Rotterdam Study, Framingham Heart Study, Cardiovascular Health Study, the Helsinki Birth Cohort Study, Cooperative Health Research in the Augsburg Region, Lothian Birth Cohort, Generation Scotland Study, Atherosclerosis Risk in Communities Study, and the Women’s Health Initiative–Epigenetic Mechanisms of PM-Mediated Cardiovascular disease. Detailed acknowledgments are included in the eAppendix in the Supplement.

REFERENCES


43. Tobacco GC, Tobacco and Genetics Consortium. Genome-wide meta-analyses identify...
DNA Methylation Signatures of Depressive Symptoms in Middle-aged and Elderly Persons

Original Investigation Research

Multiple loci associated with smoking behavior. *Nat Genet.* 2010;42(5):441-447. doi:10.1038/ng.1571


64. Fiore R, Püschel AW. The function of semaphorins during nervous system development. *Front Biosci.* 2003;8:s484-s499. doi:10.2741/1080


