

Supplementary Data 2
 DNA methylome-wide a

implicates antig Supplementary of the control
Supplicates antigraphy DNA methylome-wide association study of genetic risk for depression implicates antigen processing and immune responses

 \overline{a})
|
| Xueyi Shen* , Doretta Caramaschi*, Mark J Adams*, Rosie M Walker*, Josine L Min*
2010 - ^ مرد المسلم المعلم المسلم , , Alex Kwong1,2, Gibran Hemani2 , Genetics of DNA Methylation Consortium² , Miruna C Barbu¹ , Heather C Whalley¹ , Sarah E Harris⁵ , Ian J Deary⁵ , Stewart W Morris⁴ , Simon R Cox°, Caroline L Relton°, Riccardo E Marioni⁻, Kathryn L Evans⁻, Andrew M
...** M cintosh 1^*

¹ Division of Psychiatry, University of Edinburgh, Edinburgh, United Kingdom

- ² Medical Research Council Integrative Epidemiology Unit, Bristol Medical School, Population Health Sciences, University of Bristol, Bristol, United Kingdom
- ³ Centre for Clinical Brain Sciences, Chancellor's Building, 49 Little France Crescent, Edinburgh BioQuarter, Edinburgh, UK
- 4 Centre for Genomic and Experimental Medicine, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, United Kingdom
	- ⁵ Lothian Birth Cohorts group, Department of Psychology, The University of Edinburgh, UK

*Corresponding authors: Dr Xueyi Shen E-Mail: xueyi.shen@ed.ac.uk Address: University of Edinburgh, Royal Edinburgh Hospital, Morningside Park Edinburgh EH10 5HF United Kingdom

Prof Andrew M McIntosh E-Mail: andrew.mcintosh@ed.ac.uk Address: University of Edinburgh, Royal Edinburgh Hospital, Morningside Park Edinburgh EH10 5HF United Kingdom

Abstract

Background Depression is a disabling and highly prevalent condition where genetic and
epigenetic differences, such as DNA methylation (DNAm), contribute to prediction of disease
liability.
Method We investigated the asso

iability.
 Method We investigated the association between polygenic risk scores (PRS) for depression

and DNAm by conducting a methylome-wide association study (MWAS) in Generation

Scotland (N–8,898, mean age–49,8 vears Method
Method
Scotland
1921 ap Method We investigated the association between polygente risk scores (PRS) for depression
and DNAm by conducting a methylome-wide association study (MWAS) in Generation
Scotland (N=8,898, mean age=49.8 years) with replicat Scotland (N=8,898, mean age=49.8 years) with replication in the Lothian Birth Cohorts of
1921 and 1936 and adults in Avon Longitudinal Study of Parents and Children (ALSPAC)
(N_{combined}=2,049, mean age=79.1, 69.6 and 47.2 Scotland 1936 and adults in Avon Longitudinal Study of Parents and Children (ALSPAC)
(N_{combine}d=2,049, mean age=79.1, 69.6 and 47.2 years, respectively). We also conducted a
replication MWAS in the ALSPAC children (N=423 1921 (N_{combined}=2,049, mean age=79.1, 69.6 and 47.2 years, respectively). We also conducted
1936 replication MWAS in the ALSPAC children (N=423, mean age=17.1 years).
1921 Result Wide-spread associations were found betwe

Result Wide-spread associations were found between PRS constructed using genetic risk
variants for depression and DNAm in cytosine-guanine dinucleotide (CpG) probes that **Result** Wide-spread associations were found between PRS constructed us
variants for depression and DNAm in cytosine- guanine dinucleotide (CpG
localised to genes involved in immune responses and neural development
n and C Numarks for depression and DNAm in cytosine-guanine dinucleotide (CpG) probes that
localised to genes involved in immune responses and neural development (N_{CpG}=599,
p_{Bonferroni}<0.05, p<6.5×10⁸). The effect sizes for Variants for depression and DNAm in cytomic guantity dinutries (CpC) probes that
localised to genes involved in immune responses and neural development (N_{CpG}=599,
p_{Bonferroni}<0.05, p<6.5×10⁻⁸). The effect sizes for $p_{\text{Bonferroni}}$ <0.05, p<6.5×10⁸). The effect sizes for the significant associations were highly correlated between the discovery and replication samples in adults (r=0.83) and in adolescents (r=0.76). Additional analysis o p_{Bonferroni}<0.05, p<6.5×10°). The effect sizes for the significant associations were highly
correlated between the discovery and replication samples in adults (r=0.83) and in
adolescents (r=0.76). Additional analysis on adolescents (r=0.76). Additional analysis on the methylome-wide associations was
conducted for each lead genetic risk variant. Over 40% of the independent genetic
variants showed associations with CpG probe DNAm located in conducted for each lead genetic risk variant. Over 40% of the independent genetic
variants showed associations with CpG probe DNAm located in both the same (cis)
probes (trans) to the genetic loci (p_{Bonferroni}<0.045). S variants showed associations with CpG probe DNAm located in both the same *(cis)* and
probes *(trans)* to the genetic loci (p_{Bonferron}:<0.045). Subsequent Mendelian randomisationally sis showed that DNAm and depression a probes *(trans)* to the genetic loci ($p_{\text{Bonferroni}}$ <0.045). Subsequent Mendelian randomisation
analysis showed that DNAm and depression are mutually causal (p_{FDR} <0.039), and there is a
greater number of causal effect analysis showed that DNAm and depression are mutually causal (p_{FDR} <0.039), and there is
greater number of causal effects found from DNAm to depression (DNAm to depression: p
ranged from 0.045 to 2.06×10⁻¹²⁰; depressi greater number of causal effects found from DNAm to depression (DNAm to depression: p_{FD}
ranged from 0.045 to 2.06×10⁻¹²⁰; depression to DNAm: p_{FDR} ranged from 0.046 to 2.1×10⁻²³)
Conclusion Polygenic risk scores

ranged from 0.045 to 2.06×10⁻¹²⁰; depression to DNAm: p_{FDR} ranged from 0.046 to 2.1×10⁻²³).
Conclusion Polygenic risk scores for depression, especially those constructed from genome-
wide significant genetic risk <u>Conclusion</u> Polygenic risk scores for depression, especially those constructed from genome-
wide significant genetic risk variants, showed epigenome-wide methylation association
differences in the methylome associated wit Conclusion Polygellic risk scores for depression, especially those constructed from genome-
wide significant genetic risk variants, showed epigenome-wide methylation association
differences in the methylome associated with differences in the methylome associated with immune responses and brain development
We also found evidence from Mendelian randomisation evidence that DNAm may be ca
to depression, as well as a causal consequence of depress We also found evidence from Mendelian randomisation evidence that DNAm may be cause
to depression, as well as a causal consequence of depression.
Notice that DNAm may be caused to servel as a causal consequence of depressi We also found evidence from Mendelian randomism enterity and DNAm may be causal
to depression, as well as a causal consequence of depression. to depression, as well as a causal consequence of depression.

Introduction

Depression is a highly prevalent condition and a leading cause of global disability⁺
the underlying biological mechanisms are unclear. Genetic factors account for a si
proportion of differences in liability to depression ,
ility of
ility of
trified
ns^{4,5},
idely used the underlying biological mechanisms are unclear. There is included to the tale linking
proportion of differences in liability to depression, which has a twin-based heritability of
approximately 37% and with common genetic proximately 37% and with common genetic variants capturing around 6-10% of
phenotypic variance^{2,3}. Recent genome-wide association studies (GWAS) have identified
specific genetic risk variants for depression that implicat phenotypic variance^{2,3}. Recent genome-wide association studies (GWAS) have ider
specific genetic risk variants for depression that implicate regional brain alteration
Polygenic risk scores (PRS) derived from the results phenotypic variance^{4,3}. Recent genome-wide association studies (GWAS) have identified
specific genetic risk variants for depression that implicate regional brain alterations^{4,5}.
Polygenic risk scores (PRS) derived from Polygenic risk scores (PRS) derived from the results of GWAS studies, have been widely used
to estimate additive genetic risk⁶. PRS provides a means to identify traits that share their
genetic architecture with depressi Folygenia and and the genetic risk⁶. PRS provides a means to identify traits that share their
genetic architecture with depression, which may help to prioritise factors of biological and
mechanistic relevance for the dis

to estimate additive genetic risk". PRS provides a means to identify traits that share their
genetic architecture with depression, which may help to prioritise factors of biological an
mechanistic relevance for the disorde mechanistic relevance for the disorder⁷.
DNA methylation (DNAm) at cytosine- guanine dinucleotides (CpG) sites is one of the most
studied epigenetic markers and there is growing evidence of its role in understanding
depr mechanistic relevance for the disorder'
DNA methylation (DNAm) at cytosine- g
studied epigenetic markers and there is
depression⁸. DNAm risk scores have bee
wide association studies (MWAS)⁸. Thes
independent samples, a .
.
.
.
.e studied epigenetic markers and there is growing evidence of its role in understanding
depression⁸. DNAm risk scores have been developed from the results of DNA methylome-
wide association studies (MWAS)⁸. These can be station epigenein manute and interest growing enterties of the testils of DNA methylo
depression⁸. DNAm risk scores have been developed from the results of DNA methylo
wide association studies (MWAS)⁸. These can be use depression"
wide associa
independen
influenced k
heritability
genetic vari $\frac{1}{2}$ cross tissues $\frac{1}{2}$ and developmental stages¹². Genetic risk variants for diseases wide association studies (MWAS)°. These can be used to predict prevalent depression in
independent samples, and chronic depression that requires long-term treatment⁹. DNAm
influenced by both genetic and environmental fac independent samples, and chronic depression that requires long-term treatment". DNAm is
influenced by both genetic and environmental factors 8,10 and, in blood tissue, it has a mean
heritability of 19% across the epi heritability of 19% across the epigenome¹¹ with ~7% of its variance captured by common genetic variants¹². For the highly heritable DNA methylation probes, genetic effects are consistent across tissues¹³ and develop heritability of 19% across the epigenome²² with ~7% of its variance captured by common
genetic variants¹². For the highly heritable DNA methylation probes, genetic effects are
consistent across tissues¹³ and developm genetic variants¹². For the highly heritable DNA methylation probes, genetic effects are consistent across tissues¹³ and developmental stages¹². Genetic risk variants for disease such as schizophrenia, have been foun consistent across tissues²⁵ and developmental stages²². Genetic risk variants for diseases,
such as schizophrenia, have been found enriched in DNA methylation variation^{14–16}.
Associations between genetic risk and epi such as schizophrenia, have been found enriched in DNA methylation variation⁴⁴²⁰.
Associations between genetic risk and epigenetic changes can enrich our understant
the functional composition of genetic risk loci, and the functional composition of genetic risk loci, and thus inform the mechanisms that lead to
the onset of depression^{17,18}. However, systematic examination of the molecular genetic
associations between genetic risk of dep the onset of depression^{17,18}. However, systematic examination of the molecular genetic
associations between genetic risk of depression and DNAm has not, to the best of our
knowledge, been conducted.
In the present study, associations between genetic risk of depression and DNAm has not, to the best of our
knowledge, been conducted.
In the present study, we aim to investigate the association between PRS for depression and
genome-wide DNA met

associations between generations between and DNA mowledge, been conducted.
In the present study, we aim to investigate the association between PRS for depression
genome-wide DNA methylation. Methylome-wide association stud In the present study, we aim t
genome-wide DNA methylatic
conducted on four cohorts: G
sample, N=8,898), the Lothiar
(LBC1936). Avon Longitudinal genome-wide DNA methylation. Methylome-wide association studies (MWAS) were
conducted on four cohorts: Generation Scotland: Scottish Family Health Study (GS, discovery
sample, N=8,898), the Lothian Birth Cohort (LBC1921), genome-wide 200 minutes, part in the statement of the Party (MWA) (GS, d
conducted on four cohorts: Generation Scotland: Scottish Family Health Study (GS, d
sample, N=8,898), the Lothian Birth Cohort (LBC1921), the Lothian sample, N=8,898), the Lothian Birth Cohort (LBC1921), the Lothian Birth Cohort 1936
(LBC1936), Avon Longitudinal Study of Parents and Children (ALSPAC) adults (adult
replication sample, combined N=2,049) and ALSPAC childre sample, N=9,889), Avon Longitudinal Study of Parents and Children (ALSPAC) adults (adult replication sample, combined N=2,049) and ALSPAC children for replication (adolesce replication sample, N=423). Mendelian randomisati (LETTERT), AND LONGITUMENT COLLY TO THE THE THE THEFT (CHILDRET) TREATH
replication sample, N=423). Mendelian randomisation was used to test for causal
associations between DNAm and depression using data from the Genetics replication sample, N=423). Mendelian randomisation was used to test for causal
associations between DNAm and depression using data from the Genetics of DNA
Methylation Consortium (GoDMC) (N=25,561) and GS. associations between DNAm and depression using data from the Genetics of DNA Methylation Consortium (GoDMC) (N=25,561) and GS.

Methylation Consortium (GoDMC) (N=25,561) and GS. Methylation Consortium (GoDMC) (N=25,561) and GS.

It is made available under a [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) . **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2021.06.30.21259731;](https://doi.org/10.1101/2021.06.30.21259731) this version posted July 5, 2021. The copyright holder for this preprint

Methods

Sample descriptions

GS

Generation Scotland (GS) is a family-based population cohort with over 24,000 participants^{19,20} set-up to identify the causes of common complex disorders, such as depression. DNA methylation data and genetic data were b participants^{29,20} set-up to identify the causes of common complex disorders, such as
depression. DNA methylation data and genetic data were both collected, processed and
quality-checked for 8,898 people (mean age=49.8 ye and the state of the Matter and Batter and With Care Care and the processes paid and set 2 were 4757 (mean age 48.5 years, SD of age=14.0 years, 38.5% were men) and 4141 (mean age=51.4 years, SD of age=13.2 years, A0.90% w men) in two sets. Sample sizes for set 1 and set 2 were 4757 (mean age=48.5 years, SD of age=14.0 years, 38.5% were men) and 4141 (mean age=51.4 years, SD of age=13.2 years, 43.66% were men), respectively. Written informed age=14.0 years, 38.5% were men) and 4141 (mean age=51.4 years, SD of age=13.2 years, 43.66% were men), respectively. Written informed consent was obtained for all participants.
The study was approved by the NHS Tayside Research Ethics committee (05/s1401/89).
Lothian Birth Cohorts (LBC) 1921 and 1936
Part

The study was approved by the NHS Tayside Research Ethics committee (05/s1401/89).
Lothian Birth Cohorts (LBC) 1921 and 1936
Participants from LBC 1921 and LBC 1936^{21,22} were born in 1921 and 1936. Almost all lived in
t Lothian Birth Cohorts (LBC) 1921 and 1936

Participants from LBC 1921 and LBC 1936^{21,22} were born in 1921 and 1936. Almost all live

the Edinburgh and surrounding Lothian area when recruited. They are a mostly heat

com Examin Birth Cohorts (EDC) 1921 and 1936
Participants from LBC 1921 and LBC 1936²¹
the Edinburgh and surrounding Lothian a
community-dwelling sample of men and v
included 1,330 participants from both cohortata (LBC 1921: Participants from LBC 1921 and LBC 1936^{21,22} were born in 1921 and 1936. Almost all lived in
the Edinburgh and surrounding Lothian area when recruited. They are a mostly healthy,
community-dwelling sample of men and wome the Edinburgh and surrounding Estimate area when the terms are plus are a most of community-dwelling sample of men and women. The sample used in the current analysis included 1,330 participants from both cohorts combined w community-analysis sample of men and nontrolling sample asset in the sample used in the sample included 1,330 participants from both cohorts combined with genetic and DNA methylation data (LBC 1921: mean age=79.1 years, SD data (LBC 1921: mean age=79.1 years, SD of age=0.6, 39.7% were men; LBC 1936: mean age=69.6 years, SD of age=0.8, 50.6% were men; all participants were unrelated). Written informed consent was obtained from all participan age=69.6 years, SD of age=0.8, 50.6% were men; all participants were unrelated). Written
informed consent was obtained from all participants. Ethics permission for LBC1921 was
obtained from the Lothian Research Ethics Com age=69.6 years, serving critical metallicity, and participants. Ethics permission for LBC1921 was
obtained from the Lothian Research Ethics Committee (LREC/1998/4/183). Ethics permission
for LBC1936 was obtained from the M obtained from the Lothian Research Ethics Committee (LREC/1998/4/183). Ethics permission
for LBC1936 was obtained from the Multi-Centre Research Ethics Committee for Scotland
(MREC/01/0/56) and the Lothian Research Ethics of LBC1936 was obtained from the Multi-Centre Research Ethics Committee for Scotland
(MREC/01/0/56) and the Lothian Research Ethics Committee (LREC/2003/2/29)^{23,24}.
Avon Longitudinal Study of Parents and Children (ALSPAC

Avon Longitudinal Study of Parents and Children (ALSPAC)

(MREC/01/0/56) and the Lothian Research Ethics Committee (LREC/2003/2/29)^{23,24}.
Avon Longitudinal Study of Parents and Children (ALSPAC)
The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing longitu population-based study that recruited pregnant women residing in Avon (South-West of England) with expected delivery dates between 1st April 1991 and 31st December 1992^{25,26}.
The cohort consists of 13,761 mothers and the England) with expected delivery dates between 1st April 1991 and 31st December 1992^{25,26}.
The cohort consists of 13,761 mothers and their partners, and their 14,901 children, now
young adults²⁷. The study website cont young adults²⁷. The study website contains details of all the data that is available through a
fully searchable data dictionary and variable search tool
(<u>http://www.bristol.ac.uk/alspac/researchers/our-data/</u>). Ethical young adults²⁷. The study website contains details of all the data that is available through a
fully searchable data dictionary and variable search tool
(http://www.bristol.ac.uk/alspac/researchers/our-data/). Ethical ap (**http://www.bristol.ac.uk/alspac/researchers/our-data/**). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. A subsample of 719 unrelated mothers (Entertainment and the ALSPAC Ethics and Law Committee and the Local Research Ethics
Committees. A subsample of 719 unrelated mothers with DNA methylation (DNAm) data
(mean age=47.2 years, SD of age=4.6) were included in t Committees. A subsample of 719 unrelated mothers with DNA methylation (DNAm) data (mean age=47.2 years, SD of age=4.6) were included in the replication study²⁸.
Supplementary analyses were also conducted on a younger su (mean age=47.2 years, SD of age=4.6) were included in the replication study²⁸.
Supplementary analyses were also conducted on a younger subsample with DNAm
consisting of 423 young people (mean age=17.1 years, SD of age=1 consisting of 423 young people (mean age=17.1 years, SD of age=1.1 and 41% were boys).
Details of the selection of participants for these subsamples are in the study by Relton et al²⁸.
Consent for biological samples has Details of the selection of participants for these subsamples are in the study by Relton et al²⁸.
Consent for biological samples has been collected in accordance with the Human Tissue Act
(2004).
Genotyping and imputati Consent for biological samples has been collected in accordance with the Human Tissue Act (2004).
(2004).
Genotyping and imputation Consent for the egach samples has been consent in accordance with the Human Tissue Actor
(2004).
Detailed information on the quality control and genotyping methods for GS¹⁹, LBC1921,

(2004).
Genoty
Detailed

Genotyping and imputation
Detailed information on the q Detailed information on the quality control and genotyping methods for GS¹⁹, LBC1921,
Intervalse of the quality control and genotyping methods for GS¹⁹, LBC1921,
Intervalse of the quality control and genotyping methods

LBC1936²² and ALSPAC³³³ have been previously published and is described briefly below.
Analyses were conducted on European participants.
GS
Each sample was genotyped using the IlluminaHumanOmniExpressExome-8v1.0 BeadCh GS
Each sample was genotyped using the IlluminaHur
(48.8%) or Illumina HumanOmniExpressExome-8
chemistry³¹. Quality control included removing part
SNP removal of those with a minor allele freque C3
Eac
(48
Che
SNI
We (48.8%) or Illumina HumanOmniExpressExome-8 v1.2 BeadChip (51.2%) with Infinum chemistry³¹. Quality control included removing participants with genotyping call rate<98%, SNP removal of those with a minor allele frequenc chemistry³¹. Quality control included removing participants with genotyping call rate<98%,
SNP removal of those with a minor allele frequency (MAF)<1%, call rate<98%, Hardy-Weinberg equilibrium (HWE) p-value<5×10⁻⁶. I Imputation server with the Haplotype Reference Consortium reference panel (HRC.r1-1).
SNPs with INFO score<0.8 were removed from analysis.

LBC1921 and LBC1936

Genotyping was performed using the Illumina610-Quadv1 chip (Illumina, Inc., San Diego, CA, IMPs with INFO score<0.8 were removed from analysis.
IBC1921 and LBC1936
Genotyping was performed using the Illumina610-Quadv1 chip (Illumina, Inc., San Diego, CA,
USA). Participants were excluded with a call rate<95%. SNP SNPS WITH INFO SCORE THE REMOVED SCORE
SENSIFY And LBC1936
Genotyping was performed using the Illumina610-Quad
USA). Participants were excluded with a call rate<95%.
rate<98%, HWE p-value<0.001. Imputation and quality c
sa USA). Participants were excluded with a call rate<95%. SNPs were removed if MAP<5%, call
rate<98%, HWE p-value<0.001. Imputation and quality control based on INFO score were the
same as GS.
ALSPAC USA). Participants were entitled with a call rate convention to the call rate of the call rate components and the
same as GS.
ALSPAC
Genotyping arrays used were the Illumina Human660W-quad chip for mothers and Illumina
Hum

ALSPAC

rate serve, the power streamly maked with quality contribution in the streamline same as GS.
ALSPAC
Genotyping arrays used were the Illumina Human660W-quad chip for mothers and Illumina
HumanHap550-quad chip for children. *ALSPAC*
Genotyping
HumanHap5
MAF<0.01 w
genotyped s HumanHap550-quad chip for children. SNPs with missingness>0.05, HWE p-value<1×10⁻⁶,
MAF<0.01 were excluded. The above quality control steps were conducted on the entire
genotyped sample. Imputation and quality control ba HumanHap550-quad chip for children. SNPs with missingness>0.05, HWE p-value<1×10^{-o}, MAF<0.01 were excluded. The above quality control steps were conducted on the entire genotyped sample. Imputation and quality control ba genotyped sample. Imputation and quality control based on INFO score were consistent with
similar procedures used in GS.
Polygenic profiling geniar procedures used in GS.
Polygenic profiling
Polygenic risk scores (PRS) of depression were calculated using PRSice-2³² for GS, LBC1921,

Polygenic profiling

similar procedures december
Polygenic profiling
Polygenic risk scores (PRS) of
LBC1936 and ALSPAC separat LBC1936 and ALSPAC separately, using the summary statistics of a genome-wide meta-
analysis of depression by Howard *et al.*³³ excluding individuals from GS previously included in
that GWAS meta-analysis. Nine p-value c reported in the GWAS by Howard et al^{33} for supplementary analysis. Details of the PRS analysis of depression by Howard *et al.*³³ excluding individuals from GS previously included in that GWAS meta-analysis. Nine p-value cut-offs were used for thresholding SNPs in the summary statistics (pT): 1, 0.5, 0.1 summary statistics (pT): 1, 0.5, 0.1, 0.05, 0.01, 1×10⁻³, 1×10⁻⁴, 1×10⁻⁵ and 5×10⁻⁸ for clumping
and thresholding. Each set of SNPs was used to generate a depression-PRS in GS. A separate
PRS was generated using t PRS was generated using the lead genetic risk variants or their closest proxies (in LD r²>0.1)
reported in the GWAS by Howard *et al.*³³ for supplementary analysis. Details of the PRS
profiling procedures can be found

profiling procedures can be found elsewhere³³ (also see Supplementary Information and Supplementary Table 1-2).
Subsequently, using the lead risk variants reported by Howard et al³³, we tested for individual SNP-CpG a Subsequently, using the lindividual SNP-CpG associanting the 1-2).
Subsequently, using the lindividual SNP-CpG associanting the H
variants reported in the H
selected as leading risk vari Subsequently, using the lead risk variants reported by Howard et al³³, we tested for individual SNP-CpG associations in GS. Lead risk variants were selected by extracting the most significant proxy SNPs (p<5×10⁻⁸) in most significant proxy SNPs ($p<5\times10^{-8}$) in linkage disequilibrium (LD R²>0.01) with the lead variants reported in the Howard *et al.* study⁴. A total of 96 SNPs were available and thus selected as leading risk var variants reported in the Howard *et al.* study⁴
selected as leading risk variants for further ana
DNA methylation data I ysis. $\mathsf{I} \subset \mathsf{I}$

selected as leading risk variants for further analysis.
DNA methylation data DNA methylation data

GS

GS
Ger
Illu
<u>(htt.</u>
nor Illumina Infinium Methylation EPIC array

(https://emea.support.illumina.com/array/array_kits/infinium-methylationepic-beadchip-

<u>kit.html</u>). Data processing was performed separately for each set. Quality control (QC) an (https://emea.support.illumina.com/array/array_kits/infinium-methylationepic-beadchip-

kit.html). Data processing was performed separately for each set. Quality control (QC) and

normalisation were conducted using R packa array, or a bead count<3 in \geq 5 % of the total probe sample, or a detection p-value>0.05 for marries of the processing was performed separately for each set. $(25, 0)^{34}$, 'watermelon' (version 1.36.0)³⁵ and 'meffil' (version 1.1.1)³⁶. Details of the protocol are described elsewhere³⁷. In summary, quality described elsewhere³⁷. In summary, quality control procedures removed probes if there was
an outlying log median methylated signal intensity against unmethylated signal for each
array, or a bead count<3 in \geq 5 % of t an outlying log median methylated signal intensity against unmethylated signal for each array, or a bead count<3 in \geq 5 % of the total probe sample, or a detection p-value>0.05 for set 1 and p-value>0.01 for set 2 in an example, or a bead count<3 in \geq 5 % of the total probe sample, or a detection p-value>0.05 for set 1 and p-value>0.01 for set 2 in \geq 0.5% of the total sample in each respective set. Cross-
hybridising probes that set 1 and p-value>0.01 for set 2 in ≥0.5% of the total sample in each respective set. Cross-
hybridising probes that map to genetic variants at MAF>0.05 and polymorphic probes were
removed³⁸. Samples were excluded if se removed³⁸. Samples were excluded if sex prediction from methylation data was inconsistent
with self-reported data, or a detection p-value >0.05 for set 1 and p-value>0.01 for set 2
found in >1% of the overall probes for removed³⁶. Samples were excluded if sex prediction from methylation data was inconsistent
with self-reported data, or a detection p-value >0.05 for set 1 and p-value>0.01 for set 2
found in >1% of the overall probes for

set 1 and p-value>0.01 for set 2 in ≥0.5% of the total sample in each respective set. Cross-Found in >1% of the overall probes for each set respectively. The data was then normalised
using the 'dasen' method from the 'waterRmelon' R package (version 1.36.0).
The raw intensities were then transformed into M-value using the 'dasen' method from the 'waterRmelon' R package (version 1.36.0).
The raw intensities were then transformed into M-values by log-transforming the
proportional methylation intensity³⁹. The M-values were correct The raw intensities were then transformed into M-values by log-tran proportional methylation intensity³⁹. The M-values were corrected using model, controlling for relatedness using the GCTA-estimated genetic relation for proportional methylation intensity³⁹. The M-values were corrected using a linear-mixed
model, controlling for relatedness using the GCTA-estimated genetic relationship matrix⁴⁰
for set 1. This step was omitted for set proportional methylation intensity³⁹. The M-values were corrected using a linear-mixed
model, controlling for relatedness using the GCTA-estimated genetic relationship matrix⁴⁰
for set 1. This step was omitted for set model, controlling for relatedness using the GCTA-estimated genetic relationship matrix⁴⁶
for set 1. This step was omitted for set 2 as all participants were unrelated within the set and
to set 1. The residualised M-valu for set 1. The residualised M-values for over 750,000 autosomal CpG probes were then used
for further analysis.
LBC1921 and LBC1936
Genome-wide DNA methylation data was obtained from blood sample using the for further analysis.
LBC1921 and LBC1936
Genome-wide DNA methylation data was obtained from blood sample using the
HumanMethylation450K array (https://emea.illumina.com/content/dam/illumina-

LBC1921 and LBC1936

/
LBC1921 and LBC19
Genome-wide DNA
HumanMethylation²
<u>marketing/documen</u> HumanMethylation450K array (https://emea.illumina.com/content/da
marketing/documents/products/datasheets/datasheet_humanmethylation450.p
Quality_control_and_normalisation_were_performed_using_the_'minfi'_R_packa
1.38.0)⁴¹ marketing/documents/products/datasheets/datasheet humanmethylation450.pdf)^{-1,42}.
Quality control and normalisation were performed using the 'minfi' R package (ver
1.38.0)⁴¹. Probes with low call rate (<95%), outlying $(1.38.0)^{41}$. Probes with low call rate (<95%), outlying M-values (>3 SD from mean) or identified as cross-hybridising and polymorphic were removed⁴³. Participants with insufficient cell count information were exclude

 \mathbf{H} array (https://emea.illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-illumina-ill identified as cross-hybridising and polymorphic were removed⁴⁵. Participants with
insufficient cell count information were excluded from analysis.
All participants in LBC1921 and LBC1936 with methylation data were unrela All participants in LBC1921 and LBC1936 with methylation datransformation were conducted consistently with the GS sample
probes were retained for further analysis.
ALSPAC transformation were conducted consistently with the GS sample. Data for over 450,000 CpG
probes were retained for further analysis.
ALSPAC
Illumina Infinium HumanMethylation 450 Beadchip arrays were used for measuring geno

ALSPAC

Illumina Infinium HumanMethylation450 Beadchip arrays were used for measuring genome-
wide DNA methylation data from peripheral blood sample. The R package 'meffil' (version
1.1.1) was used for pre-processing, quality cont P
ALSPAC
Illumina Infinium HumanMethylation450 E
wide DNA methylation data from periphe
1.1.1) was used for pre-processing, q
reported⁴⁴. Further removal of probes included in the analyses³⁰. 1.1.1) was used for pre-processing, quality control and normalisation as previously
reported⁴⁴. Further removal of probes was conducted based on background detection
(p>0.05) and if they reached beyond the 3 times inter 1.11, the control of probes was conducted based on background detection (p>0.05) and if they reached beyond the 3 times inter-quantile range from 25% and 75% or identified as cross-hybridising or polymorphic⁴³. Related reported⁴⁴. Further removal of probes was conducted based on background detection
(p>0.05) and if they reached beyond the 3 times inter-quantile range from 25% and 75% or
identified as cross-hybridising or polymorphic⁴³ (p) and it they reached beyond the 3 times inter-quantile range from 25% and 75% or identified as cross-hybridising or polymorphic⁴³. Related (IBD>0.1) participants were not included in the analyses³⁰.
M-value transfor identified as cross-hybridising or polymorphic³³. Related (IBD>0.1) participants were not
included in the analyses³⁰.
M-value transformation was conducted and over 468,000 CpG probes remained for analysis.

included in the analyses"[.]
M-value transformation wa
. M-value transformation was conducted and over 468,000 CpG probes remained for analysis.

Statistical models for methylome-wide association analysis (MWAS)

sets 1 and 2, and the final summary statistics were obtained by meta-analysing the two sets
of results using METAL (version released in 2011)⁴⁵. We used the default analysis scheme
without genomic control correction (gen of results using METAL (version released in 2011)⁴⁵. We used the default analysis scheme
without genomic control correction (genomic inflation factors reported in the
Supplementary Information). P-values for the meta-ana without genomic control correction (genomic inflation factors reported in the Supplementary Information). P-values for the meta-analysis were obtained from a fixed-
effect inverse-variance model. A replication analysis on Supplementary Information). P-values for the meta-analysis were obtained from a fixed-
effect inverse-variance model. A replication analysis on adults was then conducted on the
total sample from LBC1921, LBC1936 and ALSPAC total sample from LBC1921, LBC1936 and ALSPAC adults. Replication MWAS was first
conducted separately for each cohort and then meta-analysed using the same parameters
for the discovery analysis. Finally, an additional rep conducted separately for each cohort and then meta-analysed using the same parameters
for the discovery analysis. Finally, an additional replication MWAS was conducted on ALSPAC
children. All analyses were conducted using

Linear regression was used to test the associations between depression-PRS and for each for the discovery analysis. Finally, an additional replication MWAS was conducted on ALSPAC
children. All analyses were conducted using R (version 3.5.1) under Linux environment.
Linear regression was used to test the asso For the discovery analysis. Finally, an additional replication MTML with conducted children. All analyses were conducted using R (version 3.5.1) under Linux environment.
Linear regression was used to test the associations Linear regression was used to test the associations between depression-PRS and for
CpG using the R package 'limma'⁴⁶ (version 3.48.0) for GS, LBC1921 and LBC1936. The '
function was first implemented to test the associat CpG using the R package 'limma'⁴⁶ (version 3.48.0) for GS, LBC1921 and LBC1936. The 'lmFit' function was first implemented to test the association for each CpG. The inference statistics of each linear model was then adj function was first implemented to test the association for each CpG. The inference statistics
of each linear model was then adjusted using the 'eBayes' function, by which an empirical
Bayes method was used to adjust for g Bayes method was used to adjust for gene-wise variance using a shrinkage factor.
Moderated t-statistics and p-values were produced by this step. In ALSPAC, the analyses
were conducted using the 'meffil' (version 1.1.1) R

Self-reported smoking status, smoking pack years, DNAm-estimated white-blood cell Moderated t-statistics and p-values were produced by this step. In ALSPAC, the analyses
were conducted using the 'meffil' (version 1.1.1) R package, using the 'sva' option⁴⁴.
Self-reported smoking status, smoking pack y Moderated transmission and produced using the 'meffil' (version 1.1.1) R package, using the 'sva' option⁴⁴.
Self-reported smoking status, smoking pack years, DNAm-estimated white-blood cell
proportions (CD8+T, CD4+T, nat proportions (CD8+T, CD4+T, natural killer cells, B cells and granulocytes)³⁷, batch, the first 20
principal component derived from the M-values, age and sex were included as covariates for
the discovery methylome-wide as proportions (CD8+T, CD4+T, natural killer cells, B cells and granulocytes)", batch, the first 20
principal component derived from the M-values, age and sex were included as covariates for
the discovery methylome-wide assoc prime discovery methylome-wide association analysis (see Supplementary Information and
Supplementary Table 3 for details). Where possible, the same covariates were used in the
replication analyses, although only smoking st Supplementary Table 3 for details). Where possible, the same covariates were used in the replication analyses, although only smoking status (and not pack years) were available in LBC 1921, LBC 1936 and ALSPAC. Details for replication analyses, although only smoking status (and not pack years) were available in LBC
1921, LBC 1936 and ALSPAC. Details for all the covariates included in the replication analysis
can be found in the Supplementar 1921, LBC 1936 and ALSPAC. Details for all the covariates included in the replication analysis
can be found in the Supplementary Information. MWAS were conducted for the nine
depression-PRS scores separately. P-values wer can be found in the Supplementary Information. MWAS were conducted for the nine
depression-PRS scores separately. P-values were Bonferroni-corrected (p-value threshold =
6.5×10⁸ for EPIC array used in GS, 1.1×10⁻⁷ for can be found in the Supplementary information in the seculation of depression-PRS scores separately. P-values were Bonferroni-corrected (p-value threshold = 6.5×10⁸ for EPIC array used in GS, 1.1×10⁻⁷ for 450k array us 6.5×10^8 for EPIC array used in GS, 1.1×10^{-7} for 450k array used for replication analysis in
LBC1921, LBC1936 and ALSPAC adults, and 1.1×10^{-7} for replication in ALSPAC children).
Standardised regression coef 6.5×10° for EPIC array used in GS, 1.1×10′ for 450k array used for replication analysis in LBC1921, LBC1936 and ALSPAC adults, and 1.1×10^{-7} for replication in ALSPAC children).
Standardised regression coefficients are LBC1921, LBC1936 and ALSPAC adults, and 1.1×10⁷ for replication in ALSPAC children).
Standardised regression coefficients are reported as effect sizes. For the significant CpG
probes, gene symbol annotation and UCSC cla probes, gene symbol annotation and UCSC classification of CpG Island positions were
acquired from the 'UCSC_RefGene_Name' and 'Relation_to_Island' columns, respectively,
from the annotation object generated by the
'Illumi probes, acquired from the 'UCSC_RefGene_Name' and 'Relation_to_Island' columns, respectively,
from the annotation object generated by the
'IlluminaHumanMethylationEPICanno.ilm10b4.hg19' R package (version 3.13)⁴⁷.
Indivi acquired from the annotation object generated by the

"IlluminaHumanMethylationEPICanno.ilm10b4.hg19' R package (version 3.13)⁴⁷.

Individual SNP-CpG DNAm association tests were also performed, using the same covariates

'IlluminaHumanMethylationEPICanno.ilm10b4.hg19' R package (version 3.13)⁴⁷.
Individual SNP-CpG DNAm association tests were also performed, using the same covariates
and p-value corrections as used in the PRS association 'IlluminaHumanMethylationEPICanno.ilm10b4.hg19' R package (version 3.13)"'.
Individual SNP-CpG DNAm association tests were also performed, using the sam
and p-value corrections as used in the PRS association analyses.
Gene Individual SNP-CpG DNAm association tests were also performed, using the same covariates

Gene ontology analysis

and p-value corrections as used in the PRS assembly, 2008
Gene ontology analysis
function in R package missMethyl⁴⁸. Default settings were used, function in R package missMethyl⁴⁸. Default settings were used, which include correction for
the number of probes per gene. CpGs that showed significant association with depression-
PRS at $pT < 5 \times 10^{-8}$ in the discove the number of probes per gene. CpGs that showed significant association with depression-PRS at pT<5×10° in the discovery analysis were selected as CpGs of interest, 'EPIC' was

The discovery analysis were selected as CpGs of interest, 'EPIC' was

The discovery analysis were selected as CpGs of interest, 'EPI FDR-correction was applied for all analyses.
Colocalisation analysis
We used Howard et al.'s (2019) MDD GWAS⁴ for MDD-associated SNPs and GoDMC

Colocalisation analysis

FR-Colocalisation analysis.
The distribution was applied for all analysis.
Summary statistics for mQTL analysis wit |
|
|-We used Howard et al.'s (2019) MDD GWAS" for MDD-associated SNPs and GoDMC
summary statistics for mQTL analysis with PGC studies and GS study removed, which
resulted in 32 remaining studies imputed to the 1000 Genome refer resulted in 32 remaining studies imputed to the 1000 Genome reference panel¹³. We used
the package "gwasglue" (version 0.0.0.9000, <u>https://mrcieu.github.io/gwasglue/</u>) to extract
SNPs that were +/- 1Mb of each of the 10 the package "gwasglue" (version 0.0.0.9000, <u>https://mrcieu.github.io/gwasglue/</u>) to extract SNPs that were $+/-$ 1Mb of each of the 102 genome-wide significant, lead SNPs identified in Howard et al. (2019) and then extrac SNPs that were +/- 1Mb of each of the 102 genome-wide significant, lead SNPs identified in
Howard et al. (2019) and then extracted the same SNPs within those regions from the
GoDMC mQTL analysis. We used the coloc.abf fun Howard et al. (2019) and then extracted the same SNPs within those regions from the GoDMC mQTL analysis. We used the coloc.abf function with default parameters in the "coloc" package in R (version 5.1.0)⁴⁹ to perform co GoDMC mQTL analysis. We used the coloc.abf function with default parameters in the "coloc"
package in R (version 5.1.0)⁴⁹ to perform colocalisation analysis for each SNP association.
The method tests for five mutually e package in R (version 5.1.0)⁴⁹ to perform colocalisation analysis for each SNP association.
The method tests for five mutually exclusive scenarios in a genetic region: H_o: there exist no
causal variants for either tra package in R (version 5.1.0)⁴⁹ to perform colocalisation analysis for each SNP association.
The method tests for five mutually exclusive scenarios in a genetic region: H_o: there exist no
causal variants for either tra causal variants for either trait; H_1 : there exists a causal variant for trait one only; H_2 : there exists a causal variant for trait two only; H_3 : there exist two distinct causal variants, one for each trait; and exists a causal variant for trait two only; H_3 : there exist two distinct causal variants, one for each trait; and H_4 : there exists a single causal variant common to both traits.
Mendelian randomisation (MR) each trait; and H₄: there exists a single causal variant common to both traits.
Mendelian randomisation (MR)
Three MR methods, inverse-weighted median (IVW), weighted median (WM) and MR Egger,

Mendelian randomisation (MR)

each trains of the H4: there exists a single median randomisation (MR)
Three MR methods, inverse-weighted median (IVW), weighted median (WM
were used to test for causal effects between DNA methylation and depre

were used to test for causal effects between DNA methylation and depression using the

TwoSampleMR' R package (version 0.5.6)^{50,51}.

GWAS summary statistics for DNAm were from GoDMC and GS. mQTL summary statistics

from TwoSampleMR' R package (version 0.5.6)^{50,51}.
GWAS summary statistics for DNAm were from GoDMC and GS. mQTL summary statistics
from GoDMC included 32 cohorts with 25,561 participants from European ancestry¹³. The
summa from GoDMC included 32 cohorts with 25,561 participants from European ancestry¹³. The summary statistics were computed using a two-phased design. First, every study performs a full analysis of all candidate mQTL associa full analysis of all candidate mQTL associations, returning only associations at a threshold of $p<1\times10^{-5}$. All candidate mQTL associations at $p<1\times10^{-5}$ are combined to create a unique 'candidate list' of mQTL asso $p<1\times10^{-5}$. All candidate mQTL associations at $p<1\times10^{-5}$ are combined to create a unique 'candidate list' of mQTL associations. The candidate list (n=120,212,413) is then sent back to all cohorts, and the associati variance method. Details of the database can be found elsewhere¹³. mQTL summary statistics from GS (N=~10,000) included a full set of all SNPs with no p-value thresholding. Summary mQTL statistics from GS were generated all cohorts, and the association estimates are obtained for every mQTL association on the candidate list. Candidate mQTL associations were meta-analysed using fixed-effect inverse-
variance method. Details of the database candidate list. Candidate mQTL associations were meta-analysed using fixed-effect inverse-
variance method. Details of the database can be found elsewhere¹³. mQTL summary
statistics from GS (N=~10,000) included a full s details of the mQTL analysis can be found in the Supplementary Information. variance method. Details of the database can be found elsewhere²⁵. mQTL summary
statistics from GS (N=~10,000) included a full set of all SNPs with no p-value thresholding.
Summary mQTL statistics from GS were generated Summary mQTL statistics from GS were generated using the OmicS-data-based Complex
Trait Analysis package (https://cnsgenomics.com/software/osca/#eQTL/mQTLAnalysis)⁵².
Covariates were consistent with the MWAS for depress Trait Analysis package (https://cnsgenomics.com/software/osca/#eQTL/mQTLAnalysis)⁵².
Covariates were consistent with the MWAS for depression-PRS discovery analysis. Further
details of the mQTL analysis can be found in t

Covariates were consistent with the MWAS for depression-PRS discovery analysis. Furth details of the mQTL analysis can be found in the Supplementary Information.
Summary statistics for depression GWAS by Howard *et al.*³ ³².
er
be details of the mQTL analysis can be found in the Supplementary Information.
Summary statistics for depression GWAS by Howard *et al.³³* were used. A total of 807,553
unrelated, European participants were included in the Summary statistics for depression GWAS by Howard *et al.*³³ were used. A tunrelated, European participants were included in the analysis. Details for the found elsewhere³³.
GoDMC, GS and depression GWAS samples were mu

unrelated, European participants were included in the analysis. Details for the study can be
found elsewhere³³.
GoDMC, GS and depression GWAS samples were mutually exclusive. Individual cohorts that
overlapped with the found elsewhere³³.
GoDMC, GS and depression GWAS samples were mutually exclusive. Individual cohorts that
overlapped with the Howard *et al.* depression GWAS and GS were removed from the
GoDMC mQTL meta-analysis. Depress overlapped with the Howard *et al.* depression GWAS and GS were removed from the
GoDMC mQTL meta-analysis. Depression GWAS summary statistics from Howard et al (2019)
were calculated excluding GS participants³³. See Supp GoDMC mQTL meta-analysis. Depression GWAS summary statistics from Howard et al (2019)
Were calculated excluding GS participants³³. See Supplementary Information for details.
Were calculated excluding GS participants³³. God materials analysis. Depression Community statistics from Howard Et al. (2019)
were calculated excluding GS participants³³. See Supplementary Information for details. were calculated excluding GS participants33. See Supplementary Information for details.

methylation to depression. Second, we used full mQTL summary statistics from GS to
replicate the findings and to test the causal effect to DNA methylation and depression bi-
directionally. Finally, in contrast to the univa replicate the findings and to test the causal effect to DNA methylation and depression bi-
directionally. Finally, in contrast to the univariable MR analyses (that is, each risk-outcome
pair tested separately), a multi-var that showed the strongest independent casual associations with MDD. pair tested separately), a multi-variable MR analysis was conducted to test for direct causal
associations from DNAm at multiple CpGs to MDD, using the mQTL data from GS. Using this
method, all CpG probes where there was e pair tested separately), a multi-variable minimal per line conducted to the CBC. Using this
associations from DNAm at multiple CpGs to MDD, using the mQTL data from GS. Using this
method, all CpG probes where there was evi method, all CpG probes where there was evidence of a potential causal effect on depression
were entered into the two-sample MR analysis simultaneously, in order to prioritise SNPs
that showed the strongest independent casu

method, all chemical chemical causal simultaneously, in order to prioritise SNPs
that showed the strongest independent casual associations with MDD.
Exposures were selected from CpG probes found significantly associated w that showed the strongest independent casual associations with MDD.
Exposures were selected from CpG probes found significantly associated with depression-
PRS generated using the p-threshold = $5x10^8$. As the *MHC* regi Exposures were selected from CpG probes found significantly associal
PRS generated using the p-threshold = 5x10⁸. As the *MHC* region contaring highly correlated probes, we selected the four independent probes af
at leas 14 probes were selected for final analyses. PRS generated using the p-threshold = 5x10 $^{\circ}$. As the MHC region contains a large number of highly correlated probes, we selected the four independent probes after pruning (r<0.1 for at least two nearest probes, prunin at least two nearest probes, pruning window = 3M base pairs). The probes were further
removed from analyses if they did not present in the GoDMC/GS mQTL data or had <5
independent genetic instruments overlapping with the removed from analyses if they did not present in the GoDMC/GS mQTL data or had \leq 5 independent genetic instruments overlapping with the outcome summary statistics. In result, 14 probes were selected for final analyses. removed from analyses in they did not present in the GoDMC/CS magnetic instruments overlapping with the outcome summary statistics. In result, 14 probes were selected for final analyses. independent genetic instruments over a selected for final analyses.
In result, the outcome summary statistics over the outcome summary statistics. 14 probes were selected for final analyses.

Results

Discovery MWAS of depression-PRS in GS

Methylome-wide association (MWAS) with depression-PRS at all p thresholds

(pT) at 5×10^{-8} (p<6.5×10⁻⁸ to reach significance after Bonferroni correction). In contrast to many other studies that use polygenic risk profiling at different thresholds to predict depression⁵³, both the number o depression⁵³, both the number of significant associations and the effect sizes decreased as
PRSs were calculated at increasingly lenient thresholds (Figure 1). For pT of 1×10^{-6} , 1×10^{-4}
and 0.001, 414, 57 and 17 PRSs were calculated at increasingly lenient thresholds (Figure 1). For pT of 1×10^{-6} , 1×10^{-4} and 0.001, 414, 57 and 17 CpGs were associated with depression-PRS, respectively (Supplementary Figures 1). No signifi and 0.001, 414, 57 and 17 CpGs were associated with depression-PRS, respectively (Supplementary Figures 1). No significant associations were found for PRS using p-value thresholds greater than or equal to 0.01. Quantile-qu (Supplementary Figures 1). No significant associations were found for PRS using p-value
thresholds greater than or equal to 0.01. Quantile-quantile plot and statistics for genomic
inflation factors (ranged from 0.954 to 1. (Supprementary Figure 1). The significant association interesting to the statistics for genomic
inflation factors (ranged from 0.954 to 1.011) can be found in Supplementary Figure 2 and
Supplementary Table 4. Results using thresholds greater than or equal to 0.011) can be found in Supplementary Figure 2 and
Supplementary Table 4. Results using the depression-PRS calculated using only genome-wide
significant variants are presented below.
DNAm Supplementary Table 4. Results using the depression-PRS calculated using only genome-wide
significant variants are presented below.
DNAm association with depression-PRS at a GWAS p-value association threshold of 5×10⁸

DNAm association with depression-PRS at a GWAS p-value association threshold of 5×10-8

The most significant associations of DNAm with depression-PRS were found in the major histocompatibility complex (MHC) region (25-35 Mb on Chromosome 6, Figure 2), with significant variants are presented below.
DNAm association with depression-PRS at
The most significant associations of DN.
histocompatibility complex (MHC) regio |
|-
|histocompatibility complex (*MHC*) region (25-35 Mb on Chromosome 6, Figure 2), with
575/599 (96.0%) of significant associations within this region (p_{Bonferroni} ranged from 0.05 to
1.37×10⁻¹¹¹). The top ten probes tha FIFT 1980 (96.0%) of significant associations within this region ($p_{\text{Bonferroni}}$ ranged from 0.05 to 1.37×10⁻¹¹¹). The top ten probes that showed the greatest associations are listed in Supplementary Table 5 (all $p_{\text{Bonf$ 1.37×10⁻¹¹¹). The top ten probes that showed the greatest associations are listed in Supplementary Table 5 (all p_{Bonferroni} < 2.78×10⁻⁸³). The majority of probes were in Open Sea (47.9%) and CpG shores (28.2%), and Supplementary Table 5 (all $p_{\text{Bonferroni}} < 2.78 \times 10^{-83}$). The majority of probes were in Open Sea (47.9%) and CpG shores (28.2%), and they were particularly enriched in the CpG shores compared to non-significant CpGs (18% Supplementary Table 5 (all p_{Bonferroni}<2.78×10⁻⁵⁵). The majority of probes were in Open Sea (47.9%) and CpG shores (28.2%), and they were particularly enriched in the CpG shores compared to non-significant CpGs (18% f compared to non-significant CpGs (18% for non-significant CpGs, $\chi^2=41.0$, df=1, p=1.53×10⁻¹⁰, see Supplementary Table 6 and Supplementary Information). After pruning (r<0.1 for at least two nearest probes, window = 3 compared to non-significant CpGs (18% for non-significant CpGs, χ²=41.0, df=1, p=1.53×10²⁰,
see Supplementary Table 6 and Supplementary Information). After pruning (r<0.1 for at least
two nearest probes, window = 3Mb) see Supplementary Table 6 and Supplementary Table 6 and Supplementary

September 2.40×10

See Supplementary Table 6 and Supplementary Table 5 and Supplementary

Data 1. the MHC region: cg03270340, cg14345882, cg10046620, cg08116408 (all p_{Bonferroni}<2.40×10⁵⁸). UCSC gene database annotation shows genes that are nearest to the probes are *TRIM27,* HIST1H2AI, *TRIM40* and *BTN3A2*. See

²⁰). UCSC gene database annotation shows genes that are nearest to the probes are *TRIM27,*
HIST1H2AI, *TRIM40* and *BTN3A2*. See Table 1-2, Supplementary Table 5 and Supplementary
Data 1.
Supplementary MWAS were conduct HIST1H2AI, TRIM40 and BTN3A2. See Table 1-2, Supplementary Table 5 and Supplementary
Data 1.
Supplementary MWAS were conducted on two additional depression-PRSs to investigate the
associations found within the *MHC* region Supplen
associat
indeper
wider p
and (2) Supplementary Muslim the MHC region. Two additional PRSs were calculated using (1) the independent genetic risk variants reported in the depression GWAS by Howard *et al.* with a wider pruning window of 3M base pairs and r associations found within the MHC region. Two additional PRSS were calculated using (1) the
independent genetic risk variants reported in the depression GWAS by Howard *et al.* with a
wider pruning window of 3M base pairs independent genetic risk variants reported in the depression GWAS by Howard et al. With a
wider pruning window of 3M base pairs and retaining only one variant in the MHC region
and (2) SNPs located outside of *MHC* region, and (2) SNPs located outside of *MHC* region, respectively. Analysis (1) was conducted to identify if the associations found in the *MHC* region was due to the additive effect from a large number of genetic variants inclu and (2) SNPs located outside of MHC region, respectively. Analysis (1) was conducted to identify if the associations found in the *MHC* region was due to the additive effect from a large number of genetic variants include large number of genetic variants included in the MHC region was due to the additive effect from a
large number of genetic variants included in the MHC region. Analysis (2) was conducted to
find out if the associations fou find out if the associations found within the *MHC* region. Analysis (2) was conducted to find out if the associations found within the *MHC* region was contributed by genetic variants located *trans* to this region. The find out if the associations found within the MHC region was contributed by genetic variants
located *trans* to this region. The number of significant associations found within the MHC
region for the PRS calculated using i located *trans* to this region. The number of significant associations round within the MHC
region for the PRS calculated using independent genetic risk variants reduced from 599 to 25,
at the MDD-GWAS PRS p-value threshol at the MDD-GWAS PRS p-value threshold of 5×10^{-8} . No CpGs within the *MHC* region were
found to be significantly associated with the PRS generated from variants mapping without
the MHC region. See Supplementary Figure found to be significantly associated with the PRS generated from variants mapping without the MHC region. See Supplementary Figure 3. Found to be significantly associated with the PRS generation the PRS mapping with the MHC region. See Supplementary Figure 3.
Outside of the *MHC* region, 24 probes showed significant associations with depression-PRS

Outside of the *MHC* region, 24 probes showe Outside of the MHC region, 24 probes showed significant associations with depression-PRS estimated across the genome at pT of 5×10^{-8} ($p_{\text{Bonferroni}}$ ranged from 0.048 to 4.74×10⁻⁷⁰). The
top ten probes are listed in Supplementary Table 5. Genes mapping near to the top probes
were associated with histone were associated with histone deacetylase, DNA binding and transcriptional processes (such
as *MAD1L1, TCF4, RERE* and *ZSCAN31*), and neuronal plasticity and growth (for example,
NEGR1).
The effect sizes for the signific

as *MAD1L1, TCF4, RERE* and *ZSCAN31*), and neuronal plasticity and growth (for example, *NEGR1*).
NEGR1).
The effect sizes for the significant CpG probes showed high correlations between the two
data sets (r=0.90), and as *MAD1L1, TCF4, NENE* and ZSCANS1), and neuronal plasticity and growth (for example, *NEGR1*).
The effect sizes for the significant CpG probes showed high correlations between the two
data sets (r=0.90), and direction f NEGAT).
The effect
data sets
For thes
significan
standard data sets (r=0.90), and direction for all significant associations was consistent between sets.
For these significant probes, the distance to the nearest depression risk locus was
significantly lower than those that were For these significant probes, the distance to the nearest depression risk locus was significantly lower than those that were not significant (significant versus not significant: standardised Cohen's d=0.920, $p<1\times10^{-32}$ Example significantly lower than those that were not significant (significant versus not significant:
standardised Cohen's d=0.920, p<1×10⁻³²). There were 21.2% of all significant CpGs located
outside of 1Mb boundaries standardised Cohen's d=0.920, $p<1\times10^{-32}$). There were 21.2% of all significant CpGs located
outside of 1Mb boundaries of genetic risk loci for depression and outside of the region
consisted of SNPs in LD (R²>0.1) wi outside of 1Mb boundaries of genetic risk loci for depression and outside of the region consisted of SNPs in LD (R^2 >0.1) with the genetic risk loci (see Supplementary Data 1).

Replication depression-PRS MWAS in LBC1921, LBC1936 and ALSPAC

consisted of SNPs in LD (R²>0.1) with the genetic risk loci (see Supplementary Data 1).
Replication depression-PRS MWAS in LBC1921, LBC1936 and ALSPAC
MWAS of depression-PRS on pT of 5×10⁸ on adult and adolescent samp |
|
| MWAS of depression-PRS on pT of 5×10° on adult and adolescent samples (LBC1921,
. LBC1936 and ALSPAC)

and found that the standardised effect sizes were correlated between the discovery and
replication meta-MWAS of LBC1921, LBC1936 and ALSPAC adults, with $(N_{probe}=507, r=0.83)$
or without the probes located in the MHC region $(N$ replication meta-MWAS of LBC1921, LBC1936 and ALSPAC adults, with $(N_{probe} = 507, r = 0.83)$
or without the probes located in the MHC region $(N_{probe} = 14, r = 0.78)$. There were 98.1%
associations found in the discovery MWAS remaine or without the probes located in the MHC region ($N_{probe}=14$, r=0.78). There were 98.1%
associations found in the discovery MWAS remained in the same direction and 71.8%
remained significant after Bonferroni-correction withi associations found in the discovery MWAS remained in the same direction and 71.8%

remained significant after Bonferroni-correction within the replication analysis. See Figure 3.
Standardised effect sizes for the significant CpG probes found in the discovery MWAS were
highly correlated with those in the Standardised effect sizes for the significant CpG probes found in the discovery MWAS were
highly correlated with those in the MWAS on adolescents from ALSPAC (all CpG probes:
 N_{probe} =506, r=0.76; no MHC region: N_{probe} =14, highly correlated with those in the MWAS on adolescents from ALSPAC (all CpG probes:
N_{probe}=506, r=0.76; no MHC region: N_{probe}=14, r=0.79.
MWAS for depression-PRS on all p thresholds on adult samples

MWAS for depression-PRS on all p thresholds on adult samples

N_{probe}=506, r=0.76; no MHC region: N_{probe}=14, r=0.79.
MWAS for depression-PRS on all p thresholds on adult samples
Meta-analysis of the methylome-wide association analysis of depression-PRS for replication Nprose 500, region of the region region region.

MWAS for depression-PRS on all p thresholds on adult

Meta-analysis of the methylome-wide association an

cohorts (LBC1921, LBC1936 and ALSPAC adults) sho cohorts (LBC1921, LBC1936 and ALSPAC adults) showed that, for depression-PRS of pT at 5×10^{-8} , 1×10^{-6} , 1×10^{-4} and 0.001, the number of significant CpG probes were 265, 203, 59 and 18, respectively. Similar to t 5×10^{-8} , 1×10^{-6} , 1×10^{-4} and 0.001, the number of significant CpG probes were 265, 203, 59 5×10°, 1×10°, 1×10° and 0.001, the number of significant CpG probes were 265, 203, 59
and 18, respectively. Similar to the discovery analysis, no significant associations were found
for PRS of pT≥0.01.
A full list of resu

A full list of results for replication analysis can be found in Supplementary Data 1 and Supplementary Figures 4-7. $\frac{1}{2}$

Pathway enrichment analysis

Gene ontology (GO) terms and KEGG pathways were assessed for the genes associated with Pathway enrichment analys
Sene ontology (GO) terms a
depression-PRS of pT at 5×1 |
|
| (depression-PRS of pT at 5×10⁻⁸. There were 119 enriched GO terms and 29 KEGG pathways significant after FDR correction (all $p_{FDR} < 0.044$). The majority of the significant GO terms and 19 terms significant after FDR correction (all p_{FDR} <0.044). The majority of the significant GO terms significant after FDR correction (all pFDR<0.044). The majority of the significant GO terms were associated with antigen processing and the immune response. Enriched KEGG
pathways include antigen processing and immune disease, for example, allograft rejection,
autoimmune thyroid disease and human T-cell leukemia

pathology must be the top ten GO autoimmune thyroid disease and human T-cell leukemia virus 1 infection. The top ten GO terms and KEGG pathways are listed in Tables 1-2.
A supplementary analysis without probes in the MHC r terms and KEGG pathways are listed in Tables 1-2.
A supplementary analysis without probes in the MHC region showed no significant GO terms
or KEGG pathways after FDR correction. Top GO terms that were nominally significant A supplementary analysis without probes in the M
or KEGG pathways after FDR correction. Top G
include neural development and metabolic p
phosphocreatine metabolic process, cerebellum
cerebellar Purkinje cell layer maturati For KEGG pathways after FDR correction. Top GO terms that were nominally significant
include neural development and metabolic process, for example, phosphagen and
phosphocreatine metabolic process, cerebellum and cerebell include neural development and metabolic process, for example, phosphagen and
phosphocreatine metabolic process, cerebellum and cerebellar cortex maturation and
cerebellar Purkinje cell layer maturation (p_{uncorrected} ra phosphocreatine metabolic process, cerebellum and cerebellar cortex maturation and
cerebellar Purkinje cell layer maturation (p_{uncorrected} ranged from 0.049 to 1.54×10⁻³). No
nominally significant KEGG pathway was fou eerebellar Purkinje cell layer maturation (p_{uncorrected} ranged from 0.049 to 1.54×10^{-3}). No
nominally significant KEGG pathway was found (p_{uncorrected}>0.063).
SNP-CpG mapping for the depression risk loci cerebellar Purkinje cell layer maturation (p_{uncorrected} ranged from 0.049 to 1.54×10°). No
nominally significant KEGG pathway was found (p_{uncorrected}>0.063).
SNP–CpG mapping for the depression risk loci
SNP–CpG probe

SNP–CpG mapping for the depression risk loci

nominally significant Christianus SNP-CpG mapping for the depression risk loci
SNP-CpG probe associations were investigated by conducting
independent genetic risk loci for depression. There were 6,009 SNP-CpC probe accounts in the interespared by continuing interesting the same independent genetic risk loci for depression. There were 6,009 CpG probes that showed significant associations with at least one leading genetic

independent general interests to represent these trees types up to probe that shortening
indicant associations with at least one leading genetic risk variant after Bonferroni
correction.
There were 94 of the 96 genetic ris struction.
There were 94 of the 96 genetic risk variants tested showed significant *cis* association with
CpGs within 1 Mb distance (see Figure 4). There were 88 genetic risk variants (91.6% of all
variants tested) that sh There were
CpGs withir
variants tes
(85.4% of a
different ch There were 94 of the 96 genetic risk variants tested showed significant cis association with
CpGs within 1 Mb distance (see Figure 4). There were 88 genetic risk variants (91.6% of all
variants tested) that showed *trans* variants tested) that showed *trans* associations outside of their 1Mb window and 82 variants (85.4% of all variants tested) that had *trans* associations with CpGs located on at least one different chromosome.
Five geneti

(85.4% of all variants tested) that had *trans* associations with CpGs located on at least one different chromosome.
Five genetic loci showed associations with methylation levels at CpGs located in more than half of the d (85.4% of all variants tested) that had trans associations with CpGs located on at least one
different chromosome.
Five genetic loci showed associations with methylation levels at CpGs located in more than
half of the dist Five genetic loci showe
half of the distal autosome.
variants were involved
nucleic acid binding (ZM
that colocalise with tra Five general contrivers are strained in minimizing and the species of these genetic risk
variants were involved in, for example, nucleic acid transcription activities, which includes
nucleic acid binding (ZNF179 and ESR2), variants were involved in, for example, nucleic acid transcription activities, which includes
nucleic acid binding (ZNF179 and ESR2), mitotic assembly (MAD1L1) and encoding proteins
that colocalise with transcription facto nucleic acid binding (*ZNF179* and *ESR2*), mitotic assembly (*MAD1L1*) and encoding proteins
that colocalise with transcription factors (*RERE*). Regional association plots showing genes
within 1Mb distance from the five nucleic acid binding (ZWF179 and ESR2), innotic assembly (MAD1L1) and encoding proteins
that colocalise with transcription factors (RERE). Regional association plots showing genes
within 1Mb distance from the five genetic within 1Mb distance from the five genetic variants can be found in Supplementary Figure 8.

Colocalisation analysis

We hypothesised that SNPs influencing MDD risk and those influencing DNAm would be shared. Colocalisation analysis, however, indicated that there was no strong evidence shared. Colocalisation analysis, however, indicated that there was no strong evidence
(PP₄>0.8) for shared genetic factors between MDD loci and DNAm. The posterior probab
for one region was supportive of a convincing co (PP₄>0.8) for shared genetic factors between MDD loci and DNAm. The posterior prob-
for one region was supportive of a convincing co-localised association signal for both N
and DNAm in that region (PP₄=0.710). Within (PP4) and provides a convincing co-localised association signal for both MDD
and DNAm in that region (PP₄=0.710). Within this region, three SNPs were indicated to have
a posterior probability of being shared SNPs for th and DNAm in that region (PP₄=0.710). Within this region, three SNPs were indicated to have
a posterior probability of being shared SNPs for the two traits > 0.1: rs73163796=66%,
rs73163779=17%, and rs11710605=11%. All t a posterior probability of being shared SNPs for the two traits > 0.1: rs73163796=66%,
rs73163779=17%, and rs11710605=11%. All three loci colocalised with genetic variation
influencing a smoking-associated CpG probe, cg150 rs73163779=17%, and rs11710605=11%. All three loci colocalised with genetic variation
influencing a smoking-associated CpG probe, cg15099418⁵⁴. Supplementary Data 2 con
results for all 102 regions investigated in colocal rstatute and manufacture and manufacture entertainment with generations in
influencing a smoking-associated CpG probe, cg15099418⁵⁴. Supplementary Data 2 conta
results for all 102 regions investigated in colocalisation a influencing a smoking-associated CpG probe, cg15099418⁵⁴. Supplementary Data 2 contains
results for all 102 regions investigated in colocalisation analysis.
Mendelian randomisation (MR) results for all 102 regions investigated in colocalisation analysis.
Mendelian randomisation (MR)
.

\overline{a} Mendelian randomisation (MR)

Discovery MR: causal effect of DNA methylation on depression using GoDMC data

Consistent direction of effects between methods: $cg08344181$ on chromosome3,
cg17862947 on chromosome 1 and cg17841099 on chromosome 12 (absolute $\beta_{\text{I/W}}$ ranged
from 0.042 to 0.103, p_{FDR} ranged from 2.99×10⁴⁷to 3 cg17862947 on chromosome 1 and cg17841099 on chromosome 12 (absolute $\beta_{\text{I/W}}$ ranged
from 0.042 to 0.103, p_{FDR} ranged from 2.99×10⁴⁷to 3.60×10⁻¹²⁰; absolute $\beta_{\text{MR-Enger}}$ ranged
from 0.052 to 0.122, p_{FDR} ran From 0.042 to 0.103, p_{FDR} ranged from 2.99×10⁴⁷to 3.60×10⁻¹²⁰; absolute β_{MR-Egger} ranged
from 0.052 to 0.122, p_{FDR} ranged from 0.045 to 1.23×10⁻⁶, p_{FDR} for MR-Egger intercept ranged
from 0.797 to 0.228; from 0.042 to 0.103, $ρ_{FDR}$ ranged from 2.99×10⁻¹ to 3.60×10⁻²; absolute β_{MR-Egger} ranged
from 0.052 to 0.122, $ρ_{FDR}$ ranged from 0.045 to 1.23×10⁻⁶, $ρ_{FDR}$ for MR-Egger intercept ranged
from 0.797 to 0.228; a

from 0.052 to 0.122, p_{FDR} ranged from 0.045 to 1.23×10[°], p_{FDR} for MR-Egger intercept ranged
from 0.797 to 0.228; absolute β_{WM} ranged from 0.047 to 0.107, p_{FDR} ranged from 1.23×10^{°6}
to 2.54×10⁷⁶, p_{FDR} for from 0.797 to 0.228; absolute β_{WM} ranged from 0.047 to 0.107, p_{FDR} ranged from 1.23×10⁵
to 2.54×10⁷⁶, p_{FDR} for Q-statistics ranged from 0.917 to 0.038).
Seven other probes: cg07519229, cg13813247, cg14159747, c to
Se
Cg
m(g
Sta 2.54×10⁻⁷, p_{FDR} for Q-statistics ranged from 0.917 to 0.038).
ven other probes: cg07519229, cg13813247, cg1415974
19800032 and cg23275840, showed significant causal eff
ethods (absolute β_{IWW} ranged from 0.019 t Cg19800032 and cg23275840, showed significant causal effect using the IVW and WM
methods (absolute β_{IVW} ranged from 0.019 to 0.046, p_{FDR} ranged from 4.2×10⁻³ to 1.10×10⁻⁸;
β_{WM} ranged from 0.018 to 0.048, p_{FDR} CHERE THE TRANSTROND DURNAL TRANSTROND TO THE TRANSTRONDUCT THE MET DRIVIDE BUT THE ANDMONDUCT BUT THE AND THE BUT THE SIGNIFICANT CONSULTS Tanged from 0.917 to 3.7×10⁻²⁸). Effect sizes were consistent for the above pr β_{WM} ranged from 0.018 to 0.048, ρ_{FDR} ranged from 1.70×10⁻ to 4.40×10⁻⁺, ρ_{FDR} for Q-statistics ranged from 0.917 to 3.7×10⁻²⁸). Effect sizes were consistent for the above probes were consistent between t statistics ranged from 0.917 to 3.7×10⁻²⁰). Effect sizes were consistent for the above probes
were consistent between the IVW and WM methods. No significant causal effect on
depression was found using the MR-Egger metho depression was found using the MR-Egger method for these probes (absolute $\beta_{MR\text{-Egger}}$
ranged from 4.7×10⁻⁷ to 0.003, p_{FDR} ranged from 0.739 to 0.127). However, the direction of
effects remained the same with the IVW ranged from 4.7×10⁻⁷ to 0.003, p_{FDR} ranged from 0.739 to 0.127). However, the direction of
effects remained the same with the IVW and weighted-median methods and the MR-Egger
intercept tests showed no evidence of hor effects remained the same with the IVW and weighted-median methods and the MR-Egger intercept tests showed no evidence of horizontal pleiotropy p_{fDR} ranged from 0.797 to 0.163).

Results are also shown in Figure 5 and Supplementary Figure 9 and Supplementary Table 7.

Replication MR: causal effect of DNA methylation on depression using GS data

All of the causal effects of DNA methylation to depression found in the discovery MR analysis showed consistent direction in the replication analysis and across all three MR methods. For all three MR methods, the effect si /
/
c
(analysis showed consistent direction in the replication analysis and across all three MR
methods. For all three MR methods, the effect sizes were highly correlated between
discovery and replication analyses (r ranged from methods. For all three MR methods, the effect sizes were highly correlated between
discovery and replication analyses (r ranged from 0.855 to 0.934). Eight out of ten significant
effects found in the discovery MR analysis discovery and replication analyses (r ranged from 0.855 to 0.934). Eight out of ten significant
effects found in the discovery MR analysis were significant for all three MR methods in the
replication analyses (absolute $\$ effects found in the discovery MR analysis were significant for all three MR methods in the
replication analyses (absolute β_{IVW} ranged from 0.081 to 1.022, $p_{\text{IVW-FDR}}$ ranged from 1.20×10⁷
to 4.21×10⁻¹¹⁹, abso replication analyses (absolute β_{IVW} ranged from 0.081 to 1.022, $p_{IWW-FDR}$ ranged from 1.20×10⁷ to 4.21×10⁻¹¹⁹; absolute β_{IVM} ranged from 0.102 to 1.059, p_{WM-FDR} ranged from 1.82×10⁻⁷ to 2.38×10⁻⁵⁸). The two \overline{a} 2.38×10^{-58}). The two other probes showed significant causal effect at two MR methods (cg14159747 significant for MR-Egger and WM and cg19800032 significant for IVW and WM, see statistics in Supplementary Table 3). MR 2.38×10⁻⁰). The two other probes showed significant causal effect at two MR methods (cg14159747 significant for MR-Egger and WM and cg19800032 significant for IVW and WM, see statistics in Supplementary Table 3). MR-Egge (egainated in eigenvalue of MR-Egger and WM and galacterial algorithm of the MR-Congression
see statistics in Supplementary Table 3). MR-Egger intercepts were not significantly deviated
from 0 for all replication MR (p_{FDR} see statistical experimentally taken and the supplementary terms in Supplementary Table 3).
The showed no evidence of horizontal
pleiotropy. See Supplementary Table 8.
Multi-variable MR: independent causal effect of DNA me

from 0 for all reports in the production of the literation of the station of pleiotropy. See Supplementary Table 8.
Multi-variable MR: independent causal effect of DNA methylation on depression using GS
data Multi-variable MR: independent causal effect of DNA methylation on depression using GS data

We next tested for causal associations between DNAm at multiple CpGs from the discovery analysis to MDD. The significant probes were entered into the two-sample MR analysis ノノ にくらい Manalysis to MDD. The significant probes were entered into the two-sample MR analysis
simultaneously, to identify the set of independent SNPs that showed the strongest and
independent casual associations with MDD using the simultaneously, to identify the set of independent SNPs that showed the strongest and
independent casual associations with MDD using the IVW method. Four probes showed
causal effects when all CpGs were considered simultane independent casual associations with MDD using the IVW method. Four probes showed
causal effects when all CpGs were considered simultaneously. They included cg2327584
chromosome 4, cg19624444 on chromosome 7, cg17925084 on causal effects when all CpGs were considered simultaneously. They included cg23275840 on
chromosome 4, cg19624444 on chromosome 7, cg17925084 on chromosome 14 and chromosome 4, cg19624444 on chromosome 7, cg17925084 on chromosome 14 and
 $\frac{1}{2}$
 $\frac{1}{2}$ chromosome 4, cg19624444 on chromosome 7, cg17925084 on chromosome 14 and

C.040 to 1.97×10⁻⁶, see Supplementary Figure 10 and Supplementary Table 9). Genes
annotated with these CpG probes are *SLC39A1, CORIN, KLC1* and *MAD1L1*. These genes are
involved in signalling receptor binding in the b 0.040 to 1.97×10°, see Supplementary Figure 10 and Supplementary Table 9). Genes
annotated with these CpG probes are *SLC39A1, CORIN, KLC1* and *MAD1L1*. These gen
involved in signalling receptor binding in the brain and h

MR: causal effects of depression liability on DNA methylation

involved in signalling receptor binding in the brain and hormonal regulation.
MR: causal effects of depression liability on DNA methylation
MR provided evidence of a causal effect to canopise in liability on DNA methylatio ingled in the bookstandard MR: causal effects of depression liability on DNA methylation.
The brain signal relation of a causal effect of depression liability on DNA methy
CpG probes. Other than the effect to cg09256413 wa /
|
(
-CpG probes. Other than the effect to cg09256413 was significant for all three MR methods,
effects on other probes were significant for two MR methods (for significant causal effects:
absolute $\beta_{\text{V/W}}$ ranged from 0.057 effects on other probes were significant for two MR methods (for significant causal effects:
absolute β_{IVW} ranged from 0.057 to 0.172, $p_{\text{IVW-FDR}}$ ranged from 6.95×10⁻³ to 1.48×10⁻⁵;
absolute β_{WM} ranged absolute β_{IWM} ranged from 0.057 to 0.172, $p_{\text{IWM-FDR}}$ ranged from 6.95×10³ to 1.48×10⁵;
absolute β_{WM} ranged from 0.0237 to 0.058, $p_{\text{WM-FDR}}$ ranged from 0.046 to 8.24×10³, and
absolute $\beta_{\text{MR-Egger}}$

absolute β_{IVM} ranged from 0.057 to 0.172, $p_{\text{I/WF-DR}}$ ranged from 6.95×10⁻³ to 1.48×10⁻³;
absolute β_{WM} ranged from 0.0237 to 0.058, $p_{\text{WM-FDR}}$ ranged from 0.046 to 8.24×10⁻³, and
absolute $\beta_{\text{MR-E$ absolute $\beta_{MR\text{-Egger}}$ ranged from 0.251 to 6.376, $p_{MR\text{-Egger-FDR}}$ ranged from 0.024 to 2.05×10⁻²³).
Among all the CpG probes that showed significant causal effects from depression to DNA
methylation, only cg17862947 on ch absolute $\beta_{MR\text{-}Egger}}$ ranged from 0.251 to 6.376, $p_{MR\text{-}Egger\text{-}DR}}$ ranged from 0.024 to 2.05×10²²).
Among all the CpG probes that showed significant causal effects from depression to DN
methylation, only cg17862947 methylation, only cg17862947 on chromosome 12 showed significant bidirectional effects
between depression and DNA methylation (β_{IVW} =-0.093, p_{FDR} =1.15×10⁴; β_{WM} =-0.008,
 p_{FDR} =0.506; β_{MREgger} between depression and DNA methylation (β_{IVW} =-0.093, p_{FDR} =1.15×10⁻⁴; β_{WM} =-0.008,
 p_{FDR} =0.506; $\beta_{MR\text{-Egger}}$ =-0.813, p_{FDR} =0.045). However, reversed-MR for this probe showed

significant heterogeneity of between depression and DNA methylation (β_{IVW} =-0.093, p_{FDR} =1.15×10⁻; β_{WM} =-0.008, p_{FDR} =0.506; $\beta_{\text{MR-Egger}}$ =-0.813, p_{FDR} =0.045). However, reversed-MR for this probe showed significant heter significant heterogeneity of instrumental effects (p_{FDR} for Q statistics < 1×10^{-16}). None of the other CpG probes showed consistently significant bi-directional effect with depression in the discovery and replicati significant heterogeneity of instrumental effects (p_{FDR} for Q statistics < 1×10^{-6}). None of the other CpG probes showed consistently significant bi-directional effect with depression in the discovery and replicatio of the CpC probes showed consistently significant bi-directional effects with depression in the discovery and replication MR. See Figure 6, Supplementary Figure 11 and Supplementary Table 10. discovery and replication MR. See Figure 5, Supplementary Figure 11 and Supplementary figure
Table 10.

Discussion

associations with peripheral blood DNAm across the epigenome in Generation Scotland:
Scottish Family Health Study Cohort (GS, N=8,898). The strongest DNAm associations
showed highly consistent results in the replication a Scottish Family Health Study Cohort (GS, N=8,898). The strongest DNAm associations
showed highly consistent results in the replication analysis in adults (N=2,049, r_β = 0.83) and
in adolescents (N=423, r_β = 0.76). S showed highly consistent results in the replication analysis in adults (N=2,049, r_β = 0.83) and
in adolescents (N=423, r_β = 0.76). Significant CpG probes are enriched in immunological
processes in the human leukocyt in adolescents (N=423, r_β = 0.76). Significant CpG probes are enriched in immunological
processes in the human leukocyte antigen (HLA) system and neuronal maturation and
development. Influence from the genetic risk of processes in the human leukocyte antigen (HLA) system and neuronal maturation and
development. Influence from the genetic risk of depression was both local (*cis*) and distal
(*trans*). Five genetic risk loci showed wide-s provide the genetic risk of depression was both local (*cis*) and distal
(*trans*). Five genetic risk loci showed wide-spread *trans* effect across half of the autosomal
chromosomes. Finally, using Mendalian randomisation, (*trans*). Five genetic risk loci showed wide-spread *trans* effect across half of the autosomal
chromosomes. Finally, using Mendalian randomisation, we found evidence of a mutually
causal effect of DNAm on liability to MD (trans). Tive genetic risk foot showed wide spread trans effect across half of the autosomal
chromosomes. Finally, using Mendalian randomisation, we found evidence of a mutually
causal effect of DNAm on liability to MDD at

causal effect of DNAm on liability to MDD at CpG probes associated with polygenic risk
scores for depression.
The probes associated with genetic risk variants for depression map to genes including
TRIM27, BTN3A2 and HIST1H scores for depression.
The probes associated with genetic risk variants for depression map to genes including
TRIM27, BTN3A2 and HIST1H2AI. These HLA-related genes have been widely found
associated with psychiatric conditi The probes associated
TRIM27, BTN3A2 and
associated with psych
genes that located of
associated with neuro TRIM27, BTN3A2 and HIST1H2AI. These HLA-related genes have been widely found
associated with psychiatric conditions such as schizophrenia and bipolar disorder⁶. Other
genes that located outside of the MHC region, such a TRIMIZ7, BTN3A2 and HIST1H2AI. These TLA-related genes have been widely found
associated with psychiatric conditions such as schizophrenia and bipolar disorder⁶. Other
genes that located outside of the MHC region, such associated with psychiatric conditions such as schizophrenia and bipolar disorder⁹
genes that located outside of the MHC region, such as *BDNF*, *RERE* and *ZSCAN*
associated with neuronal development and guidance of neu 31, are
iptional
ing and
Gisland genes that located outside of the MHC region, such as *BDM*, *RERE* and *236AN31*, are
associated with neuronal development and guidance of neuronal growth⁵⁵, transcriptional
processes^{56,57}, as well as other risk facto

associated with neuronal development and guidance of neuronal growth⁵⁵, transcriptional
processes^{56,57}, as well as other risk factors for depression, for example, obesity, smoking and
abnormal physical development⁵⁸. processes^{36,37}, as well as other risk factors for depression, for example, obesity, smoking and
abnormal physical development⁵⁸.
A large proportion of identified CpG probes were localised in open sea areas and CpG isla abnormal physical development²⁰.
A large proportion of identified Cp
shores. Previous studies found tha
CpG island shores⁵⁹, along with hig
SNPs¹³. Methylation at these class
extent than those that map outs Shores. Previous studies found that probes showed greater individual variations within the
CpG island shores⁵⁹, along with higher heritability and greater number of *trans* effect from
SNPs¹³. Methylation at these clas CpG island shores⁵⁹, along with higher heritability and greater number of *trans* effect from SNPs¹³. Methylation at these classes of CpGs may be subject to genetic effects to a greater extent than those that map outs extent than those that map outside such regions. The genes related to the depression
genetic risk variants in CpG island shores have previously been linked to DNA, RNA and
chromatin binding⁶⁰. Differentially methylation genetic risk variants in CpG island shores have previously been linked to DNA, RNA and
chromatin binding⁶⁰. Differentially methylation in these regions may reflect a regulatory role
for genetic risk variants through infl

generation in Comparison in Comparison in these regions may reflect a regulatory role
for genetic risk variants through influencing DNA methylation in regions adjacent to CpG
islands⁵⁹.
Mendelian randomisation provided e probes associated with lead SNPs on depression. After controlling for functional pleiotropy
shared between CpG probes, four probes showed an independent causal effect on
depression. Genes annotated with these independent p islands"
Mendelia
probes as
shared b
depressio
phenotyp probes associated with lead SNPs on depression. After controlling for functional pleiotropy
shared between CpG probes, four probes showed an independent causal effect on
depression. Genes annotated with these independent p shared between CpG probes, four probes showed an independent causal effect on
depression. Genes annotated with these independent probes are associated with
phenotypes such as lower total brain volume⁶¹, excessive C-react depression. Genes annotated with these independent probes are associated with phenotypes such as lower total brain volume⁶¹, excessive C-reactive protein⁶², obesity⁶³ and adverse lifestyle factors such as smoking⁶⁴ phenotypes such as lower total brain volume⁶¹, excessive C-reactive protein⁶², obesity⁶³ and
adverse lifestyle factors such as smoking⁶⁴, which are implicated in both patients with
depression and those who have bee adverse lifestyle factors such as smoking⁶⁴, which are implicated in both patients with depression and those who have been exposed to early environmental risks, such as childhood trauma⁶⁵. These phenotypes have also b depression and those who have been exposed to early environmental risks, such as childhood trauma⁶⁵. These phenotypes have also been shown to have causal effects on depression in previous MR studies^{4,5}.
Although eviden

childhood trauma⁶⁵. These phenotypes have also been shown to have causal effects on
depression in previous MR studies^{4,5}.
Although evidence was found for bidirectional causal effects between DNA methylation and
depres depression, statistical evidence was stronger for the causal effect from DNA methylation to depression in previous MR studies^{4,5}.
Although evidence was found for bid
depression, statistical evidence was s
depression compared to the opposi
used for the reversed causal effect, a
depression ranged from 5 to 18 an depression, statistical evidence was stronger for the causal effect from DNA methylation to
depression compared to the opposite direction, despite that more genetic variants were
used for the reversed causal effect, and th depression compared to the opposite direction, despite that more genetic variants were
used for the reversed causal effect, and thus statistical power was greater (N_{SNP} for DNAm to
depression ranged from 5 to 18 and N used for the reversed causal effect, and thus statistical power was greater (N_{SNP} for DNAm to depression ranged from 5 to 18 and N_{SNP} for depression to DNAm was 112). The highly consistent methylome-wide association depression ranged from 5 to 18 and N_{SNP} for depression to DNAm was 112). The highly consistent methylome-wide associations found across adults and adolescents may indicate that early genetic influence on DNAm result in consistent methylome-wide associations found across adults and adolescents may indicate
that early genetic influence on DNAm result in a predominantly directional effect from
DNAm to depression⁶⁶. that early genetic influence on DNAm result in a predominantly directional effect from
DNAm to depression⁶⁶.
 $T_{\rm H}$ and $T_{\rm H}$ generically direction $\frac{1}{2}$ predominantly directional effect from $\frac{1}{2}$ predominantly directional effect from $\frac{1}{2}$ produced from $\frac{1}{2}$ produced from $\frac{1}{2}$ produced from $\frac{1}{2}$ DNAm to depression

results. One limitation for interpretating the current findings is that DNA methylation data
was collected from blood samples, that may not reflect the most relevant cell types in
depression. Nevertheless, studies have sho was collected from blood samples, that may not reflect the most relevant cell types in
depression. Nevertheless, studies have shown that the genetic drivers of DNAm have similar
effects across multiple cell types^{13,67}. T depression. Nevertheless, studies have shown that the genetic drivers of DNAm have similar effects across multiple cell types^{13,67}. The greater accessibility of DNAm from whole blood also has clear sample size and other effects across multiple cell types^{13,67}. The greater accessibility of DNAm from whole blood
also has clear sample size and other methodological advantages compared to measures
obtained from neural tissue post-mortem, and also has clear sample size and other methodological advantages compared to measures obtained from neural tissue post-mortem, and it is more likely that these measures could be used in future clinical applications. Future s obtained from neural tissue post-mortem, and it is more likely that these measures could be

used in future clinical applications. Future studies could further expand the scope by
including other cell and tissue types.
In the current study we demonstrate that genome-wide genetic risk variants for depression
show w including other cell and tissue types.

In the current study we demonstrate that genome-wide genetic risk variants for depression

show wide-spread epigenome-wide DNAm associations both individually and when

combined in a In the current study we demonstrate
show wide-spread epigenome-wid
combined in a risk score. These cha
responses and may provide clues to t In the current study and the current show wide-spread epigenome-wide DNAm associations both individually and when
combined in a risk score. These changes implicate antigen processing and immune system
responses and may pro combined in a risk score. These changes implicate antigen processing and immune system responses and may provide clues to the underlying mechanisms of depression. responses and may provide clues to the underlying mechanisms of depression.

It is made available under a [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) . **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2021.06.30.21259731;](https://doi.org/10.1101/2021.06.30.21259731) this version posted July 5, 2021. The copyright holder for this preprint

Acknowledgments

Generation Scotland received is currently supported by the Wellcome Trust Investigator
Award in Science 01/06/2021 to 31/5/26 'Exploiting genomic approaches to identify the
environmental basis of depression'. (Reference: 2 Award in Science 02/07/2022 to 21/06/2021 to McIntosh AM (PI). The
DNA methylation profiling and data preparation was supported by Wellcome Investigator
Award 220857/2/20/2 and Grant 104036/2/14/2 (PI for both grants is Pr ENA methylation profiling and data preparation was supported by Wellcome Investigator
Award 220857/Z/20/Z and Grant 104036/Z/14/Z (PI for both grants is Prof McIntosh AM)
and through funding from NARSAD (Ref 27404, Dr Howa Award 220857/2/20/2 and Grant 104036/2/14/2 (PI for both grants is Prof McIntosh AM)
and through funding from NARSAD (Ref 27404, Dr Howard DM) and the Royal College of
Physicians of Edinburgh (SIM Fellowship, Dr Whalley HC And through funding from NARSAD (Ref 27404, Dr Howard DM) and the Royal College of
Physicians of Edinburgh (SIM Fellowship, Dr Whalley HC). Genotyping of the GS:SFHS
samples was funded by the MRC and Wellcome Trust [104036 Physicians of Edinburgh (SIM Fellowship, Dr Whalley HC). Genotyping of the GS:SFHS
samples was funded by the MRC and Wellcome Trust [104036/Z/14/Z]. Generation Scotland
also receives support from the Chief Scientist Office Samples was funded by the MRC and Wellcome Trust [104036/Z/14/Z]. Generation Scotland
also receives support from the Chief Scientist Office of the Scottish Government Health
Directorates [CZD/16/6] and the Scottish Funding

samples was funded by the MRC and Wellcome Trust [104036/Z/14/Z]. Generation Scotland Directorates [CZD/16/6] and the Scottish Funding Council [HR03006].
We thank the LBC1936 participants and team members who contributed to the LBC1921 and
LBC1936 studies. Further study information can be found at <u>https://</u> Methank the LBC1936 participants and team members who contributed UBC1936 studies. Further study information can be found at https://
Dirth-cohorts. The LBC1921 was supported by the United Kingdor
Biological Sciences Resea LBC1936 studies. Further study information can be found at https://www.ed.ac.uk/lothian-
birth-cohorts. The LBC1921 was supported by the United Kingdom's Biotechnology and
Biological Sciences Research Council (BBSRC), The MR/M013111/1, MR/R024065/1), and the University of Edinburgh. Methylation typing of
LBC1921 and LBC1936 was supported by Centre for Cognitive Ageing and Cognitive Biological Sciences Research Council (BBSRC), The Royal Society and The Chief Scientist
Office of the Scottish Government. The LBC1936 is supported by Age UK (Disconnected Mind
project, which supports Dr Harris SE), the Me Office of the Scottish Government. The LBC1936 is supported by Age UK (Disconnected Mind
project, which supports Dr Harris SE), the Medical Research Council (G0701120, G1001245,
MR/M013111/1, MR/R024065/1), and the Univers project, which supports Dr Harris SE), the Medical Research Council (G0701120, G1001245,
MR/M013111/1, MR/R024065/1), and the University of Edinburgh. Methylation typing of
LBC1921 and LBC1936 was supported by Centre for C project, which supports Dr Harris Septembury and the University of Edinburgh. Methylation typing of
LBC1921 and LBC1936 was supported by Centre for Cognitive Ageing and Cognitive
Epidemiology (Pilot Fund award), Age UK, Th LBC1921 and LBC1936 was supported by Centre for Cognitive Ageing and Cognitive
Epidemiology (Pilot Fund award), Age UK, The Wellcome Trust Institutional Strategic Support
Fund, The University of Edinburgh, and The Universi Epidemiology (Pilot Fund award), Age UK, The Wellcome Trust Institutional Strategic Support
Fund, The University of Edinburgh, and The University of Queensland. Dr Cox SR and Prof
Deary IJ were supported by a National Inst Fund, The University of Edinburgh, and The University of Queensland. Dr Cox SR and Prof
Deary IJ were supported by a National Institutes of Health (NIH) research grant
R01AG054628, and Dr Cox SR is supported by a Sir Henry Deary IJ were supported by a National Institutes of Health (NIH) research grant
R01AG054628, and Dr Cox SR is supported by a Sir Henry Dale Fellowship jointly funded by
the Wellcome Trust and the Royal Society (Grant Numbe

Pearly II were supported by a Sir Henry Dale Fellowship jointly funded by
R01AG054628, and Dr Cox SR is supported by a Sir Henry Dale Fellowship jointly funded by
the Wellcome Trust and the Royal Society (Grant Number 2218 ROT THE TERRIT SUPPORTLY THE TERRIT SUPPORTLY THE TERRIT SUPPORTLY THE TERRIT PRINCIPLE THE URLY AND THE URLY MEDICINE THE UNIVERSIT The UK Medical Research Council and Wellcome (Grant Ref: 217065,
University of Bristol provide core support for ALSPAC. A comprehensive list
is available on the ALSPAC
(http://www.bristol.ac.uk/alspac/external/documents/gr University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding

is available on the ALSPAC website

(http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). GWAS

data is available on the ALSPAC website
(http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). GWAS
data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger
Institute and Lab (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). GWAS
data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger
Institute and LabCorp (Laboratory Corporation of Ame (http://www.britanumum.com/mental/documents/mental/documents/mental/documents/mental/documents/mental/documents/mental/documents/mental/documents/weare-grateful to all the families who took part in this study, the midwives Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.
We are grateful to all the families who took part in this study, the midwives for their help in
recruiting them, and the whole ALSPAC te Institute and Laboratory (Laboratory Corporation of Ministry) and Corporation Corporation
Instituting them, and the whole ALSPAC team, which includes interviewers, computer and
laboratory technicians, clerical workers, res The families the families who took part in the family, the middle study and the families precruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, rese laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. Part of this data was collected using REDCap, see the REDCap website for details https://projectredcap.org/reso receptionists and nurses. Part of this data was collected using REDCap, see the REDCap receptionists and nurses. Part of this data was converted and going using the Rep. website for details https://projectredcap.org/resources/citations/). website for details https://projectredcap.org/resources/citations/).

References

-
-
- 1. Marcus, M., Yasamy, M. T., Ommeren, M. Van, Chisholm, D. & Saxena, S. Depression:

4 Global Public Health Concern. (2012).

2. Sullivan, P. F., Neale, M. C. & Kendler, K. S. Genetic epidemiology of major depression:

re A Global Public Health Concern. (2012).
Sullivan, P. F., Neale, M. C. & Kendler, K.
Sullivan, P. F., Daly, M. J. & O'Donov
disorders: the emerging picture and it
(2012). 2. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architectures of psychiatric
3. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architectures of psychiatric
3. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Gene review and meta-analysis. Am J J Sychiatry 157, 1552–1562 (2000).
Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architecture
disorders: the emerging picture and its implications. Nat. Rev. G
(2012).
Howard, D. M. *e*
- 3. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet.* **13**, 537–551 (2012).

4. Howard, D. M. *et al.* Genome-wide met disorders: the emerging picture and its implications. *Nat. Rev.* Genet. 13, 537–531
(2012).
Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102
independent variants and highlights the importance 4. Howard, D. M. et al. Genome-wide meta-analysis of depression identifies 102
independent variants and highlights the importance of the prefrontal brain regions.
Nat. Neurosci. 22, 343–352 (2019).
5. Wray, N. R. *et al.*
- Nat. Neurosci. 22, 343–352 (2019).
Wray, N. R. et al. Genome-wide association analyses identify 44 risk variants and
refine the genetic architecture of major depression. Nat. Genet. 50, 668–681 (2018). Nat. Neurosci. 22, 343–352 (2019).
Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and
refine the genetic architecture of major depression. Nat. Genet. 50, 668–681 (2018).
International, T.
- schizophrenia and bipolar disorder. Nature 10, 748–752 (2009).
Shen, X. et al. A phenome-wide association and Mendelian Randomisation study of
- 5. Wray, N. R. et al. Genome-wide association analyses identify 44 11sk variants and
refine the genetic architecture of major depression. Nat. Genet. 50, 668–681 (2018).
6. International, T. *et al.* Common polygenic varia refine the genetic architecture of major depression. Nat. Genet. 50, 668–681 (2018).
International, T. *et al.* Common polygenic variation contributes to risk o
schizophrenia and bipolar disorder. Nature 10, 748–752 (2009) 6. International, T. et al. Common polygenic variation contributes to risk or
schizophrenia and bipolar disorder. Nature 10, 748–752 (2009).
7. Shen, X. et al. A phenome-wide association and Mendelian Randomisation study o
- polygenic risk for depression in UK Biobank. Nat. Commun. 11, 1–16 (2020).
Barbu, M. C. *et al.* Epigenetic prediction of major depressive disorder. *Mol. Psychiatry*
1–12 (2020). doi:10.1038/s41380-020-0808-3
Clark, S. L. 9. Shen, X. et al. A phenome-wide association and Mendelian Randomisation study of
polygenic risk for depression in UK Biobank. Nat. Commun. 11, 1–16 (2020).
8. Barbu, M. C. et al. Epigenetic prediction of major depressiv polygenic risk for depression in OK Biobank. Nat. Commun. 11, 1–16 (2020).
Barbu, M. C. *et al.* Epigenetic prediction of major depressive disorder. *Mol. 1*
1–12 (2020). doi:10.1038/s41380-020-0808-3
Clark, S. L. *et al.*
-
- 8. Barbu, M. C. et al. Epigenetic prediction of major depressive disorder. Mol. Psychiatry
1–12 (2020). doi:10.1038/s41380-020-0808-3
9. Clark, S. L. et al. A methylation study of long-term depression risk. Mol. Psychiatr 1334–1343 (2020).
Zeng, Y. *et al.* Parent of origin genetic effects on methylation in humans are common
and influence complex trait variation. *Nat. Commun.* 10, 1–13 (2019). 9. Clark, S. L. et al. A methylation study of long-term depression risk. Mol. Psychiatry 25, 1334–1343 (2020).
10. Zeng, Y. *et al.* Parent of origin genetic effects on methylation in humans are common
and influence compl
- in shaping the human methylome. *Nat. Commun.* 7, (2016).
Van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex 10. Zeng, *P. et al. Parent of origin* genetic effects on methylation in humans are common
and influence complex trait variation. *Nat. Commun.* **10**, 1–13 (2019).
11. Van Dongen, J. *et al.* Genetic and environmental inf Van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex
in shaping the human methylome. *Nat. Commun.* 7, (2016).
Van Dongen, J. *et al.* Genetic and environmental influences interact with ag
- 11. Van Dongen, J. et al. Genetic and environmental influences interact with age and sex
in shaping the human methylome. Nat. Commun. 7, (2016).
12. Van Dongen, J. et al. Genetic and environmental influences interact with In shaping the human methylome. Nat. Commun. 7, (2010).
Van Dongen, J. *et al.* Genetic and environmental influences
in shaping the human methylome. Nat. Commun. 7, 1–13 (2
Min, J. L. *et al.* Genomic and phenomic insights
- 12. Van Dongen, J. et al. Genetic and environmental influences interact with age and sex
in shaping the human methylome. Nat. Commun. 7, 1–13 (2016).
13. Min, J. L. *et al.* Genomic and phenomic insights from an atlas of In shaping the human methylome. Nat. Commun. 7, 1–13 (2016).
Min, J. L. *et al.* Genomic and phenomic insights from an atlas o
DNA methylation. *medRxiv* 25, 81 (2020).
Hannon, E. *et al.* Methylation QTLs in the developin
-
- 13. Min, J. L. et al. Genomic and phenomic insights from an atlas of genetic effects on
DNA methylation. *medRxiv* **25**, 81 (2020).
14. Hannon, E. *et al.* Methylation QTLs in the developing brain and their enrichment in
1 BINA methylation. *methalv* 25, 81 (2020).
Hannon, E. *et al.* Methylation QTLs in the
schizophrenia risk loci. *Nat. Neurosci*. 19, *i*
Hannon, E. *et al.* An integrated genetic-ep
for co-localization of genetic associati 14. Hannon, E. et al. Metriyiation QTLs in the developing brain and their emiciment in
schizophrenia risk loci. *Nat. Neurosci*. **19**, 48–54 (2015).
15. Hannon, E. et al. An integrated genetic-epigenetic analysis of schi schizophrenia risk loci. Nat. Neurosci. 19, 48–54 (2019).
Hannon, E. *et al.* An integrated genetic-epigenetic analy
for co-localization of genetic associations and differenti
Biol. 17, 176 (2016).
Hannon, E. *et al.* El 15. Hannon, E. et al. An integrated genetic epigenetic analysis of schizophrenia. Evidence
for co-localization of genetic associations and differential DNA methylation. *Genome*
Biol. 17, 176 (2016).
16. Hannon, E. *et al.*
- For co-localization of genetic associations and differential DNA methylation. Genome
Biol. 17, 176 (2016).
DNA methylation at birth. *Genome Med.* 10, 1–13 (2018).
Sullivan, P. F. & Posthuma, D. Biological pathways and net Biol. 17, 176 (2010).
Hannon, E. *et al.* Ele
DNA methylation at
Sullivan, P. F. & Pe
Sullivan, P. F. & Ge
-
- 16. Hannon, E. et al. Elevated polygenic burden for autism is associated with differential

16. Sullivan, P. F. & Posthuma, D. Biological pathways and networks implicated in

17. Sullivan, P. F. & Posthuma, D. Biological p DNA methylation at birth. *Genome Med.* 10, 1–13 (2018).
Sullivan, P. F. & Posthuma, D. Biological pathways and networks implicated in
psychiatric disorders. *Curr. Opin. Behav. Sci.* 2, 58–68 (2015).
Sullivan, P. F. & Ges 17. Sullivan, P. F. & Geschwind, D. H. Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. Cell 177, 162–183 (2019). psychiatric disorders. Carr. Opin. Behav. Sci. 2, 58–68 (2015).
Sullivan, P. F. & Geschwind, D. H. Defining the Genetic,
Diagnostic Architectures of Psychiatric Disorders. Cell 177, 16 Diagnostic Architectures of Psychiatric Disorders. Cell 177, 162-183 (2019). Diagnostic Architectures of Psychiatric Disorders. Cell 177, 162–183 (2019).
-
- 19. Smith, B. H. et al. Generation Scotland: The Scottish Family Health Study, a new
resource for researching genes and heritability. *BMC Med. Genet.* **7**, 74 (2006).
20. Smith, B. H. *et al.* Cohort profile: Generation s resource for researching genes and neritability. *BMC Med. Genet.* 7, 74 (2000).
Smith, B. H. *et al.* Cohort profile: Generation scotland: Scottish family health
(GS: SFHS). The study, its participants and their potential
- 20. Smith, B. H. et al. Cohort prome. Generation section. Section family health study (GS: SFHS). The study, its participants and their potential for genetic research on health and illness. *Int. J. Epidemiol.* **42**, 689–7 health and illness. *Int. J. Epidemiol*. **42**, 689–700 (2013).
Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J. & Fox, H. C. The Impact of
Childhood Intelligence on Later Life: Following Up the Scottish Mental S health and illness. *Int. J. Epidermol.* 42, 689–700 (2013).
Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L.
Childhood Intelligence on Later Life: Following Up the
1932 and 1947. Journal of Personality and Social
- 21. Childhood Intelligence on Later Life: Following Up the Scottish Mental Surveys of
22. Deary, I. J., Whalley, L. J. & Starr, J. M. A lifetime of intelligence: Follow-up studies of
22. Deary, I. J., Whalley, L. J. & Star 1932 and 1947. Journal of Personality and Social Psychology **86**, 130–147 (2004).
Deary, I. J., Whalley, L. J. & Starr, J. M. A lifetime of intelligence: Follow-up studies of
the Scottish mental surveys of 1932 and 1947. A 1932 and 1947. Journal of Personality and Social Psychology 66, 130–147 (2004).
Deary, I. J., Whalley, L. J. & Starr, J. M. A lifetime of intelligence: Follow-up studi
the Scottish mental surveys of 1932 and 1947. (2009). 22. Beary, I. 3., Whalley, E. 3. & Starr, 3. W. A lifetime of intelligence: Follow-up
the Scottish mental surveys of 1932 and 1947. A lifetime of intelligence: Follow-up
studies of the Scottish mental surveys of 1932 and trudies of the Scottish mental surveys of 1932 and 1947. A lifetime of intelligence: Follow-up
studies of the Scottish mental surveys of 1932 and 1947. (2009). doi:10.1037/11857-
000
Cox, S. R. et al. Brain white matter in
-
- Studies of the Scottish mental surveys of 1992 and 1947. (2009). doi:10.1097/11897
Cox, S. R. et al. Brain white matter integrity and cortisol in older men: The Lothian
Birth Cohort 1936. Neurobiol. Aging 36, 257–264 (2015 Cox,
Birth
Hilla
infla
doi:: 23. Cox, 3. R. et al. Brain white matter integrity and cortisol in older men: The Lothian
Birth Cohort 1936. Neurobiol. Aging 36, 257–264 (2015).
24. Hillary, R. et al. Integrative omics approach to identify the molecular Birth Cohort 1936. Neurobiol. Aging 36, 297–264 (2019).
Hillary, R. *et al.* Integrative omics approach to identify t
inflammatory protein levels in healthy older
doi:10.1101/2020.02.17.952135
Fraser, A. *et al.* Cohort pr 24. Hillary, R. et al. Integrative omics approach to identify the molecular architecture of
inflammatory protein levels in healthy older adults. 1–15 (2020).
doi:10.1101/2020.02.17.952135
25. Fraser, A. *et al.* Cohort pro
- doi:10.1101/2020.02.17.952135
Fraser, A. *et al.* Cohort profile: The avon longitudinal study of parents and children:
ALSPAC mothers cohort. *Int. J. Epidemiol*. **42**, 97–110 (2013).
Boyd, A. *et al.* Cohort profile: The Fraser, A. *et al.* Cohort profile: 1
ALSPAC mothers cohort. *Int. J. Ep*
Boyd, A. *et al.* Cohort profile: T
avon longitudinal study of parent
Northstone. K. *et al.* The Avon L
-
- 25. Fraser, A. et al. Cohort profile: The avon longitudinal study or parents and children.
ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
26. Boyd, A. *et al.* Cohort profile: The 'Children of the 90s'-Th ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
Boyd, A. *et al.* Cohort profile: The 'Children of the 90s'-The index offspring of the
avon longitudinal study of parents and children. *Int. J. Epidemiol.* 26. Boyd, A. et al. Cohort profile: The 'Children *Int. J. Epidemiol.* 42, 111–127 (2013).
27. Northstone, K. *et al.* The Avon Longitudinal Study of Parents and Children (ALSPAC):
28. Northstone of the enrolled sample of Northstone, K. *et al.* The Avon Longitudinal Study of Parents and Children (ALSPAC):
an update on the enrolled sample of index children in 2019 [version 1; peer review: 2
approved]. Wellcome Open Res. 4, (2019).
Relton, C 27. Northstone, K. et al. The Avon Longitudinal Study of Parents and Children (ALSPAC).

an update on the enrolled sample of index children in 2019 [version 1; peer review: 2

approved]. *Wellcome Open Res.* **4**, (2019).

- approved]. Wellcome Open Res. 4, (2019).
Relton, C. L. et al. Data Resource Profile: Accessible Resource for Integrated
Epigenomic Studies (ARIES). *Int. J. Epidemiol*. 44, 1181–1190 (2015).
Davies, G. et al. Genome-wide a approved₁. Wencome Open Res. 4, (2019).
Relton, C. L. *et al.* Data Resource Pro
Epigenomic Studies (ARIES). *Int. J. Epidemic*
Davies, G. *et al.* Genome-wide association
highly heritable and polygenic. *Mol. Psychi*
Kw
-
- 28. Relton, C. L. et al. Data Resource Prome. Accessible Resource for Integrated
Epigenomic Studies (ARIES). *Int. J. Epidemiol.* **44**, 1181–1190 (2015).
29. Davies, G. *et al.* Genome-wide association studies establish th Epigenomic Studies (ARIES). *Int. S. Epidemiol.* 44, 1181–1150 (2015).
Davies, G. *et al.* Genome-wide association studies establish that hun
highly heritable and polygenic. *Mol. Psychiatry* 16, 996–1005 (2011).
Kwong, A. 29. Bavies, G. et al. Genome-wide association studies establish that human intelligence is
highly heritable and polygenic. Mol. Psychiatry 16, 996–1005 (2011).
30. Kwong, A. S. F. et al. Polygenic risk for depression, anxi highly heritable and polygenic. Mol. Psychiatry 16, 996–1009 (2011).
Kwong, A. S. F. *et al.* Polygenic risk for depression, anxiety and
associated with the severity and rate of change in depressive s
adolescence. J. Child 30. Kwong, A. S. F. et al. Polygelic risk for depression, anxiety and neuroticism are
associated with the severity and rate of change in depressive symptoms across
adolescence. J. Child Psychol. Psychiatry Allied Discip. (
- adolescence. J. Child Psychol. Psychiatry Allied Discip. (2021). doi:10.11111/jcpp.13422
Gunderson, K. L. Whole-genome genotyping on bead arrays. Methods Mol. Biol. 529,
197–213 (2009).
Choi, S. W., Mak, T. S. H. & O'Reill
- adolescence. J. Child Psychol. Psychiatry Allied Discip. (2021). doi:10.1111/jcpp.13422
Gunderson, K. L. Whole-genome genotyping on bead arrays. Methods Mol. Biol. 529,
197–213 (2009).
Choi, S. W., Mak, T. S. H. & O'Reilly
- 31. Gunderson, K. L. Whole-genome genotyping on bead arrays. Methods Mol. Biol. 329,
32. Choi, S. W., Mak, T. S. H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk
33. Howard, D. et al. Genome-wide meta-an Choi, S. W., Mak,
score analyses. N
Howard, D. *et*
independent vari
Nat. Neurosci. 22 33. Howard, D. *et al.* Genome-wide meta-analysis of depression identifies 102
33. Howard, D. *et al.* Genome-wide meta-analysis of depression identifies 102
34. Mat. Neurosci. **22**, 343–352 (2019).
34. Fortin, J. P., Fert score analyses. *Nature Protocols* 15, 2759–2772 (2020).
Howard, D. *et al.* Genome-wide meta-analysis of
independent variants and highlights the importance of
Nat. Neurosci. 22, 343–352 (2019).
Fortin, J. P., Fertig, E. 33. Howard, D. et al. Genome-wide meta-analysis of depression identifies 102
independent variants and highlights the importance of the prefrontal brain regions.
Nat. Neurosci. 22, 343–352 (2019).
34. Fortin, J. P., Fertig
- Nat. Neurosci. 22, 343–352 (2019).
Fortin, J. P., Fertig, E. & Hansen, K. shinyMethyl: Interactive quality control of Illumina
450k DNA methylation arrays in R. F1000Research 3, 1–11 (2014).
Pidsley, R. *et al.* A data-dri Nat. Neurosci. 22, 343–352 (2019).
Fortin, J. P., Fertig, E. & Hansen, K. shinyMethyl: Interactive quality control of Illumina
450k DNA methylation arrays in R. F1000Research 3, 1–11 (2014).
Pidsley, R. *et al.* A data-dri
- 35. Fortherm, Forting, E. Et al. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics 14, 293 (2013). 450k DNA methylation arrays in R. 71000Research 5, 1–11 (2014).
Pidsley, R. *et al.* A data-driven approach to preprocessing Illumin.
array data. *BMC Genomics* 14, 293 (2013). $35.$ Pidsley, R. et al. A data-driven approach to preprocessing Indiana 450K methylation array data. *BMC Genomics* 14, 293 (2013). array data. BMC Genomics $\overline{14}$, 293 (2013).
- 36. 136. normalization and analysis of very large DNA methylation datasets. *Bioinformatics* 34, 3983–3989 (2018).
37. Walker, R. M. *et al.* Assessment of dried blood spots for DNA methylation profiling
[version 1; peer revie
- normalization and analysis of very large DNA methylation datasets. *Biomformaties* 34,
3983–3989 (2018).
Walker, R. M. *et al.* Assessment of dried blood spots for DNA methylation profiling
[version 1; peer review: 2 appro 3983–3989 (2018).
Walker, R. M. *et al.* Assessment of dried blood spots for DNA methylation profiling
[version 1; peer review: 2 approved]. *Wellcome Open Res.* 4, 44 (2019).
Jaffe, A. E. *et al.* Mapping DNA methylation
-
- S7. Wanker, R. M. et al. Assessment of dried blood spots for DNA methylation proming
[version 1; peer review: 2 approved]. Wellcome Open Res. 4, 44 (2019).
38. Jaffe, A. E. et al. Mapping DNA methylation across development [version 1; peer review: 2 approved]. Wellcome Open Res. 4, 44 (2019).
Jaffe, A. E. *et al.* Mapping DNA methylation across development, ₁
schizophrenia in the human frontal cortex. N*at. Neurosci*. **19**, 40–47 (20
Du, P 38. Jaffe, A. E. et al. Mapping DNA methylation across development, genotype and
schizophrenia in the human frontal cortex. Nat. Neurosci. 19, 40–47 (2015).
39. Du, P. et al. Comparison of Beta-value and M-value methods f Schizophrenia in the human frontal cortex. Nat. Neurosci. 19, 40–47 (2015).
Du, P. *et al.* Comparison of Beta-value and M-value methods for q
methylation levels by microarray analysis. *BMC Bioinformatics* 11, 587 (2010)

-
- 39. Bu, P. et al. Companson of Beta-value and M-value methods for quantifying
methylation levels by microarray analysis. *BMC Bioinformatics* 11, 587 (2010).
40. Yang, J. *et al.* Genome partitioning of genetic variation methylation levels by microarray analysis. *BMC Biomformatics* 11, 587 (2010).
Yang, J. *et al.* Genome partitioning of genetic variation for complex trait
common SNPs. *Nat. Genet.* 43, 519–525 (2011).
Zhang, Q. *et al.* 40. Tang, J. et al. Genome partitioning of genetic variation for complex traits using
common SNPs. Nat. Genet. 43, 519–525 (2011).
41. Zhang, Q. et al. Genotype effects contribute to variation in longitudinal methylome
pat common SNPs. Nat. Genet. 43, 319–525 (2011).
Zhang, Q. *et al.* Genotype effects contribute to
patterns in older people. *Genome Med.* 10, 75 (2
Shah, S. *et al.* Genetic and environmental expo
the human life course. *Geno*
-
- 41. Zhang, Q. et al. Genotype effects contribute to variation in longitudinal methylome
patterns in older people. *Genome Med.* 10, 75 (2018).
42. Shah, S. *et al.* Genetic and environmental exposures constrain epigenetic patterns in older people. *Genome Med.* 10, 75 (2018).

Shah, S. *et al.* Genetic and environmental exposures constrain epigenetic drift over

the human life course. *Genome Res.* **24**, 1725–1733 (2014).

McCartney, D. L. 42. Shah, S. et al. Genetic and environmental exposures constrain epigenetic drift over
the human life course. Genome Res. 24, 1725–1733 (2014).
43. McCartney, D. L. *et al.* Identification of polymorphic and off-target pr the human life course. Oenome Res. 24, 1723–1733 (2014).
McCartney, D. L. *et al.* Identification of polymorphic and
sites on the Illumina Infinium MethylationEPIC BeadChip.
(2016).
Min, J. L., Hemani, G., Davey Smith, G.,
- 43. McCartney, D. L. et al. Identification of polymorphic and on-target probe binding
sites on the Illumina Infinium MethylationEPIC BeadChip. *Genomics Data* 9, 22–24
(2016).
44. Min, J. L., Hemani, G., Davey Smith, G., R sites on the Illumina Infinium MethylationEPIC Beauting. Oenomics Data 9, 22–24
(2016).
Min, J. L., Hemani, G., Davey Smith, G., Relton, C. & Suderman, M. Meffil: efficient
normalization and analysis of very large DNA meth (2010).
Min, J.
normali
3983–3
Willer, (
of geno
- mormalization and analysis of very large DNA methylation datasets. *Bioinformatics* **34**, 3983–3989 (2018).

45. Willer, C. J., Li, Y., Abecasis, G. R. & Overall, P. METAL: fast and efficient meta-analysis of genomewide as normalization and analysis of very large DNA methylation datasets. *Bioinformaties* 34,
3983–3989 (2018).
Willer, C. J., Li, Y., Abecasis, G. R. & Overall, P. METAL: fast and efficient meta-analysis
of genomewide associati Willer, C. J., Li, Y., A
of genomewide ass
Ritchie, M. E. *et*
sequencing and mic
Hansen KD. Illum
-
- 144. Willer, C. J., M. M. J. Limma, Bioinformatics 26, 2190–2191 (2010).

46. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-

sequencing and microarray studies. Nucleic Acids Res. 43, e47 (201 or genomewide association scans. *Biomformatics* 20, 2190–2191 (2010).
Ritchie, M. E. *et al.* Limma powers differential expression analys
sequencing and microarray studies. *Nucleic Acids Res.* 43, e47 (2015).
Hansen KD. 46. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-
sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
47. Hansen KD. IlluminaHumanMethylationEPICanno.ilm10b4.hg19: Annot sequencing and microarray studies. Nucleic Acids Res. 43, e47 (2013).

Hansen KD. IlluminaHumanMethylationEPICanno.ilm10b4.hg19: /

Illumina's EPIC methylation arrays. *R package version 0.6.0* (2017

https://bitbucket.com
- 11. Humina's EPIC methylation arrays. *R package version 0.6.0* (2017). Available at:
https://bitbucket.com/kasperdanielhansen/Illumina_EPIC. (Accessed: 18th June 2021)
48. Phipson, B., Maksimovic, J. & Oshlack, A. MissMet Illumina's EFIC methylation arrays. *R package version 0.0.0* (2017). Available at:
https://bitbucket.com/kasperdanielhansen/Illumina_EPIC. (Accessed: 18th June 2021)
Phipson, B., Maksimovic, J. & Oshlack, A. MissMethyl: A https://bittps://bittps://bittps://bittps://bittps://bittps://phipson, B., Maksimovic, J. & Oshlack, A. MissMethyl: An R package for analyzing
data from Illumina's HumanMethylation450 platform. *Bioinformatics* 32, 286–288 48. Phipson, Bioinformatics, Philadelphia Methylation 450 platform. Bioinformatics 32, 286–288
(2016).
49. Wallace, C. et al. Statistical colocalization of monocyte gene expression and genetic
risk variants for type 1 diab
-
- (2016).
(2016).
Wallace, C. *et al.* Statistical colocalization of monocyte gene expression and genetic
risk variants for type 1 diabetes. Hum. Mol. Genet. **21**, 2815–2824 (2012).
Hemani, G., Tilling, K. & Davey Smith, G. (2016).
Wallace, C. *et al.* Statistical colocalization of monocyte gene expression and genetic
risk variants for type 1 diabetes. *Hum. Mol. Genet.* **21**, 2815–2824 (2012).
Hemani, G., Tilling, K. & Davey Smith, G. Orien 49. Wanace, C. et al. Statistical colocalization of monocyte gene expression and generic
risk variants for type 1 diabetes. Hum. Mol. Genet. 21, 2815–2824 (2012).
50. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the risk variants for type 1 diabetes. *Hum. Mol.* Genet. 21, 2815–2824 (2012).
Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationsh
imprecisely measured traits using GWAS summary data. *PLoS Genet.* 13
(2 51. Hemani, G. et al. The MR-base platform supports systematic causal inference across
the human phenome. Elife 7, (2018).
52. Zhang, F. et al. OSCA: A tool for omic-data-based complex trait analysis. Genome Biol.
- imprecisely measured traits using GWAS summary data. *PLOS Genet.* 13, e1007001
(2017).
Hemani, G. *et al.* The MR-base platform supports systematic causal inference across
the human phenome. *Elife* 7, (2018).
Zhang, F. (2017).
Hemani
the hun
Zhang, I
20, 107
Rigdeli
- 51. Hemani, G. et al. The MR-base platform supports systematic causal inference across
the human phenome. Elife 7, (2018).
52. Zhang, F. et al. OSCA: A tool for omic-data-based complex trait analysis. Genome Biol.
53. Big the human phenome. *Enje 1*, (2016).
Zhang, F. *et al.* OSCA: A tool for omic
20, 107 (2019).
Bigdeli, T. B. *et al.* Genetic effects in 52. Zhang, F. et al. OSCA: A tool for omic-data-based complex trait analysis. Genome Biol.
20, 107 (2019).
53. Bigdeli, T. B. *et al.* Genetic effects influencing risk for major depressive disorder in
- 20, 107 (2015).
Bigdeli, T. B. *et* 53. Bigdeli, T. B. et al. Genetic effects influencing risk for major depressive disorder in

-
- China and Europe. Transl. Psychiatry 7, 01 (2017).

Sehanes, R. *et al.* Epigenetic Signatures of Cigarett

9, 436–447 (2016).

Sandrini, L. *et al.* Association between obes

neurotrophic factor (BDNF) levels: Systematic 54. Joehanes, R. et al. Epigenetic Signatures of Cigarette Smoking. Circ. Caranovasc. Genet.

9, 436–447 (2016).

55. Sandrini, L. *et al.* Association between obesity and circulating brain-derived

neurotrophic factor (BD 9, 436–447 (2010).
Sandrini, L. *et al*
neurotrophic factor
International Journe
Fagerberg, L. *et al.*
Integration of Trans
- SS. Sandrini, L. et al. Association between obesity and chedualing brain-derived
neurotrophic factor (BDNF) levels: Systematic review of literature and meta-analysis.
International Journal of Molecular Sciences 19, (2018). International Journal of Molecular Sciences 19, (2018).
Fagerberg, L. et al. Analysis of the Human Tissue-specific Expression by Genome-wide
Integration of Transcriptomics and Antibody-based Proteomics. *Mol. Cell. Proteom* International Journal of Molecular Sciences 13, (2016).

Fagerberg, L. *et al.* Analysis of the Human Tissue-specif

Integration of Transcriptomics and Antibody-based Pro
 13, 397–406 (2014).

Plaster, N., Sonntag, C., S
- Integration of Transcriptomics and Antibody-based Proteomics. *Mol. Cell. Proteomics*
13, 397–406 (2014).
57. Plaster, N., Sonntag, C., Schilling, T. F. & Hammerschmidt, M. REREa/atrophin-2
interacts with histone deacetyla 13, 397–406 (2014).

Plaster, N., Sonntag, C., Schilling, T. F. & Hammerschmidt, M. REREa/atrophin-2

interacts with histone deacetylase and Fgf8 signaling to regulate multiple processes

of zebrafish development. *Dev. Dy* 13, 397–400 (2014).
Plaster, N., Sonntag
interacts with histor
of zebrafish developi
Mansur, R. B. *et al.*
results from the glob
- 57. Plaster, N., Sonntag, C., Schilling, T. F. & Hammerschmidt, M. REREa/atrophin-2
interacts with histone deacetylase and Fgf8 signaling to regulate multiple processes
of zebrafish development. Dev. Dyn. 236, 1891–1904 (of zebrafish development. *Dev. Dyn.* **236**, 1891–1904 (2007).
Mansur, R. B. *et al.* BDNF and BMI effects on brain structures of bipolar offspring:
results from the global mood and brain science initiative. *Acta Psychiat* of zebrafish development. *Dev. Dyn.* 236, 1891–1904 (2007).
Mansur, R. B. *et al.* BDNF and BMI effects on brain structuresults from the global mood and brain science initiative. Ac
607–614 (2017).
Visone, R. *et al.* DNA
- SB. Mansur, R. B. et al. BBNF and BMF effects on brain structures or bipolar orisping.

results from the global mood and brain science initiative. Acta Psychiatr. Scand. 136,

607–614 (2017).

S9. Visone, R. et al. DNA me results from the global mood and brain science initiative. Acta Psychiatr. Scand. 136,
607–614 (2017).
Visone, R. *et al.* DNA methylation of shelf, shore and open sea CpG positions
distinguish high microsatellite instabil Visone, R. *et al*
distinguish high
colon cancer ster
Hendrickson, D.,
binding by chrom SS. Visone, R. et al. DNA methylation of shelf, shore and open sea Cpd positions
distinguish high microsatellite instability from low or stable microsatellite status
colon cancer stem cells. *Epigenomics* 11, 587–604 (2019
- distinguish high microsatellite instability from low or stable microsatellite status
colon cancer stem cells. *Epigenomics* 11, 587–604 (2019).
Hendrickson, D., Kelley, D. R., Tenen, D., Bernstein, B. & Rinn, J. L. Widespr binding by chromatin-associated proteins. Genome Biol. 17, 28 (2016).
-
- 61. Xu, Y., Xiao, G., Liu, L. & Lang, M. Zinc transporters in Alzheimer's disease. Molecular
61. Xu, Y., Xiao, G., Liu, L. & Lang, M. Zinc transporters in Alzheimer's disease. Molecular
62. Noh, H., Paik, H. Y., Kim, J. & Binding by chromatin-associated proteins. Genome Biol. 17, 28 (2010).
Xu, Y., Xiao, G., Liu, L. & Lang, M. Zinc transporters in Alzheimer's dise
Brain 12, (2019).
Noh, H., Paik, H. Y., Kim, J. & Chung, J. The alteration of 61. Xu, Y., Xiao, G., Liu, L. & Lang, M. Zinc transporters in Alzheimer's disease. Molecular
Brain 12, (2019).
62. Noh, H., Paik, H. Y., Kim, J. & Chung, J. The alteration of zinc transporter gene
expression is associated Bram 12, (2019).
Noh, H., Paik, H
expression is ass
Elem. Res. **158**, 1
Starnawska, A.
symptomatology expression is associated with inflammatory markers in obese women. *Biol. Trace*
Elem. Res. 158, 1–8 (2014).
63. Starnawska, A. *et al.* Epigenome-wide association study of depression
symptomatology in elderly monozygotic
- Elem. Res. 158, 1–8 (2014).
Starnawska, A. et al.
Symptomatology in elderly i
Zhong, R. et al. MAD1L1 Ar.
increase the risk of colorect
Pariante C. M. & Lightman
- Elem. Res. 158, 1–8 (2014).
Elem. Res. 158, 1–8 (2014).
Starnawska, A. et al. Epigenome-wide association study of depression
symptomatology in elderly monozygotic twins. Transl. Psychiatry 9, 214–214 (2019).
Zhong, R. et a Starnawska, A. et al. Epigenome-wide association study of depression
symptomatology in elderly monozygotic twins. Transl. Psychiatry 9, 214–214 (2019).
64. Zhong, R. et al. MAD1L1 Arg558His and MAD2L1 Leu84Met interaction
- Zhong, R. *et al.* MAD1L1 Arg558His and MAD2L1 Leu84Met interaction with smoking increase the risk of colorectal cancer. *Sci. Rep.* **5**, 1–9 (2015).
Pariante, C. M. & Lightman, S. L. The HPA axis in major depression: clas
- Enong, R. et al. MAD1L1 Arg558His and MAD2L1 Leub4Wet interaction with smoking
increase the risk of colorectal cancer. Sci. Rep. 5, 1–9 (2015).
65. Pariante, C. M. & Lightman, S. L. The HPA axis in major depression: classi and new developments. *Trends Neurosci*. **31**, 464–468 (2008).
Mooney, M. A. *et al.* Large epigenome-wide association studentifies peripheral DNA methylation associated with dise
burden. *Transl. Psychiatry* **10**, 1–12 (2 66. Mooney, M. A. *et al.* Large epigenome-wide association study of childhood ADHD
identifies peripheral DNA methylation associated with disease and polygenic risk
burden. Transl. Psychiatry 10, 1–12 (2020).
67. Braun, P. and new developments. *Trends Neurosci.* 31, 464–466 (2006).
Mooney, M. A. *et al.* Large epigenome-wide association studentifies peripheral DNA methylation associated with dise-
burden. *Transl. Psychiatry* 10, 1–12 (2020 66. Mooney, M. A. et al. Large epigenome wide association stady of childhood ADHD
identifies peripheral DNA methylation associated with disease and polygenic risk
burden. Transl. Psychiatry 10, 1–12 (2020).
67. Braun, P.
- burden. Transl. Psychiatry 10, 1–12 (2020).
Braun, P. R. *et al.* Genome-wide DNA methylation comparison between live human
brain and peripheral tissues within individuals. Transl. Psychiatry 9, 1–10 (2019). burden. *Transl. Psychiatry* 10, 1–12 (2020).
Braun, P. R. *et al.* Genome-wide DNA met
brain and peripheral tissues within individua braun, P. R. et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. Transl. Psychiatry 9, 1–10 (2019). brain and peripheral tissues within individuals. Transl. Psychiatry 9, 1–10 (2019).

Figure 2. Manhattan plot for the discovery methylome-wide association study (MWAS) for PRS of pT at 5×10^{-8} in Generation Scotland (GS). Each dot represents a CpG probe. X-axis represents the relative position of the probes in the genome. Y-axis represents -log10-transformed p values. The red dashed line represents the significance threshold for Bonferroni correction.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.30.21259731; this version posted July 5, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/tunder, who has granted med The copyright holder for this prepriation that is a selection posted July 6, 2021. The copyright holder for this preprint the copyright of this preprint. who has granted by perfact in the authorycle and has granted medRxix a license to display the prepriation in perpetuity. . [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) It is made available under a

It is made available under a [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) . **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2021.06.30.21259731;](https://doi.org/10.1101/2021.06.30.21259731) this version posted July 5, 2021. The copyright holder for this preprint

Figure 3. Replication MWAS in Lothian Birth Cohort (LBC) 1921, LBC1936 and Avon
Longitudinal Study of Parents and Children (ALSPAC) adults. (a) Manhattan plot for the Longitudinal Study of Parents and Children (ALSPAC) adults. (a) Manhattan plot for the replication MWAS for PRS of pT at 5×10°. Each dot represents a CpG probe. X-axis
محمد represents the relative position of the probes in the genome. Y-axis represents -log10-
transformed p values. The red dashed line represents the significance threshold for Bonferroni correction. (b) Number of CpG probes significant associated with polygenic risk scores (PRS) at nine different pTs for replication analysis. X-axis represents the pTs used for scores (PRS) at nine different pTs for replication analysis. A-axis represents the pTs used for generating PRS. Y-axis shows the number of probes significant associated with the given PRS. The four different lines represent four types of methods to define significance. (c) Scatter plot showing the correlation of standardised regression coefficients between the discovery (GS) and replication (LCB1921+LBC1936+ALSPAC adults) analysis. Each dot represents a CpG probe. Probes shown in the figure are those associated with depression-PRS of pT at 5×10⁻⁸
8. A figure - Annual Annual Annual Colorador - Annual Colorador - Annual Colorador - Annual Colorador - Annual in the discovery MWAS (in GS). Dots in green represent probes locate in the major histocompatibility complex (MHC) regions and those in red represent other probes that locate outside of the MHC region.

Figure 4. Heatmap showing the SNP-CpG mapping. Each row and column represent a CpG probe and depression risk locus, respectively. Those tests that are not significant after Bonferroni-correction are left blank. For those significant associations, a darker cell represents a higher -log10-transformed p-value. All CpG probes and depression risk loci were categorised based on which chromosome (CHR) they locate. Within each chromosome, probes and SNPs are aligned from left to right or from bottom to top based on their genomic position.

Figure 5. Mendelian randomisation (MR) analysis on DNA methylation and depression using data from the Genetics of DNA Methylation Consortium (GoDMC). (a) Discovery MR testing causal effect of DNA methylation to depression using GoDMC data. (b) Replication MR testing effect of DNA methylation using GS data to depression. (c) MR of reversed directionality testing the causal effect from depression to DNA methylation. X-axes represent p-values for MR analyses. Y-axes represent the individual tests conducted.

Table 1. Results for gene ontology (GO) analysis for the MWAS on PRS at pT 5×10^{-8} . Analyses were conducted separately for including and excluding the MHC region. Top ten GO terms are listed in the table. BP=biological process, CC=cellular component and MF=molecular function.

Table 2. Results for pathway analysis for the MWAS on PRS at $pT 5\times10^{-8}$. Analyses were conducted separately for including and excluding the MHC region. Top ten KEGG pathways are listed in the table.

