Methylome-wide association study of antidepressant use in Generation Scotland and the Netherlands Twin Register implicates the innate immune system

Barbu MC, PhD[1], Huider F, MSc[2], Campbell A, MA[3], Amador, C, PhD[4], Adams, MJ, PhD[1], Lynall, ME, PhD[5], Howard, DM, PhD[1,6], Walker, RM, PhD[3], Morris SW, MSc[3], Van Dongen J[2], Porteous, DJ, PhD[3], Evans, KL, PhD[3], Bullmore, E, PhD[5], Willemsen, G[2], Boomsma DI[2], Whalley, HC, PhD[1], McIntosh, AM, MD[1]

[1]Division of Psychiatry, The University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK

[2]Faculty of Behavioural and Movement Sciences, Biological Psychology, Vrije Universiteit Amsterdam, Netherlands

[3]Centre for Genomic and Experimental Medicine, The Institute of Genetics and Cancer, The University of Edinburgh, UK

[4]MRC Human Genetics Unit, The Institute of Genetics and Cancer, The University of Edinburgh, UK

[5]Department of Psychiatry, University of Cambridge, Cambridge, UK

[6]Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

Correspondence can be sent to:

Dr Miruna C Barbu, mbarbu@ed.ac.uk, 0131 537 6691

Division of Psychiatry, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh EH10 5HF

Running title: Methylome-wide association study of self-reported antidepressant use

Keywords: DNA methylation, methylome-wide association study, methylation-based score, innate immune system, major depressive disorder, antidepressant use

Abstract

Antidepressants are an effective treatment for major depressive disorder (MDD), although individual response is unpredictable and highly variable. Whilst the mode of action of antidepressants is incompletely understood, many medications are associated with changes in DNA methylation in genes that are plausibly linked to their mechanisms. Studies of DNA methylation may therefore reveal the biological processes underpinning the efficacy and side effects of antidepressants.

We performed a methylome-wide association study (MWAS) of self-reported antidepressant use accounting for lifestyle factors and MDD in Generation Scotland (GS:SFHS, N=6,428, EPIC array) and the Netherlands Twin Register (NTR, N=2,449, 450K array) and ran a meta-analysis of antidepressant use across these two cohorts.

We found 10 CpG sites significantly associated with self-reported antidepressant use in GS:SFHS, with the top CpG located within a gene previously associated with mental health disorders, *ATP6V1B2* (β=-0.055, $p_{corrected}$=0.005). Other top loci were annotated to genes including *CASP10, TMBIM1, MAPKAPK3,* and *HEBP2*, which have previously been implicated in the innate immune response. Next, using penalised regression, we trained a methylation-based score of self-reported antidepressant use in a subset of 3,799 GS:SFHS individuals that predicted antidepressant use in a second subset of GS:SFHS (N=3,360, β=0.377, $p=3.12\times10^{-11}$, R^2=2.12%). In an MWAS analysis of prescribed selective serotonin reuptake inhibitors, we showed convergent findings with those based on self-report. In NTR, we did not find any CpGs significantly associated with antidepressant use. The meta-analysis identified the two CpGs of the ten above that were common to the two arrays used as being significantly associated with antidepressant use, although the effect was in the opposite direction for one of them.

Antidepressants were associated with epigenetic alterations in loci previously associated with mental health disorders and the innate immune system. These changes predicted self-reported antidepressant use in a subset of GS:SFHS and identified processes that may be relevant to our mechanistic understanding of clinically relevant antidepressant drug actions and side effects.

65

66

67 **Introduction**

68 Major Depressive Disorder (MDD) is a leading cause of disability worldwide (1) and is
69 caused by a combination of environmental and genetic factors (2). MDD has a number of
70 effective treatments, with antidepressant drugs amongst the most commonly prescribed
71 evidence-based therapies worldwide (3,4). Response to antidepressants is nevertheless
72 unpredictable and highly variable, with only approximately 50% of individuals achieving
73 remission after completing two treatments (5). The unmet needs of many individuals with
74 MDD indicate the urgency of understanding the mechanisms of effective antidepressant
75 action (6) so that their efficacy can be improved and their delivery targeted to those most
76 likely to benefit.

77 The mode of action of antidepressants was originally assumed to be through the
78 inhibition of monoamine re-uptake. However, since synaptic monoamine concentration
79 poorly mirrors the trajectory of symptomatic improvement, their mechanism of action
80 remains uncertain (7). Studies of antidepressants in both animal and human studies have
81 implicated a number of possible modes of action. These include evidence that
82 antidepressants lead to changes in DNA methylation (DNAm) and gene expression in a
83 number of potentially relevant pathways (8). Francois et al. (2015) (9) demonstrated that
84 antidepressants reversed stress-induced changes in DNAm and expression of *5-HT1A* in an
85 animal model, while Zimmermann et al. (2012) found that antidepressants affect activity of
86 DNA methyltransferase 1 (*DNMT1*) (10). These studies demonstrate that antidepressants
87 may lead to epigenetic changes that reveal their biologically relevant mechanisms of action.

88 Epigenetic processes are associated with alterations in DNA activity without
89 alterations to the underlying genome sequence (11). DNAm is perhaps the most commonly
90 investigated epigenetic change, owing to the availability of reliable high-throughput array
91 technologies that can identify changes in DNAm at over 800K locations throughout the
92 genome. Methylome-wide association studies (MWAS) have begun to identify a number of
93 cytosine-phosphate-guanine (CpG) sites that are associated with MDD (12) and relevant
94 environmental risk factors (13). DNAm changes may therefore capture an environmental
95 archive of exogenous factors, including drug treatments, relevant to the onset and
96 maintenance of MDD.

97 Recent studies have shown the potential of a DNAm-based risk score to predict MDD
98 and its associated lifestyle and environmental factors, including antidepressant use (14,15).
99 We previously (15) trained a DNAm predictor of MDD in 3,047 individuals using a penalised
100 regression model and showed that it was significantly associated with self-reported
101 antidepressant use in an independent cohort. This suggests that CpG sites conferring risk to
102 MDD may also be linked to antidepressant use. A methylation-based predictor of self-
103 reported antidepressant use may therefore be able to predict antidepressant use and its
104 effects in other samples and may signpost the development of clinical biomarkers
105 quantifying drug action.
106 Here, we sought to identify the DNAm changes associated with self-reported
107 antidepressant use in 6,428 individuals ($N_{antidepressant\ use}$=740) from the Generation Scotland
108 study (GS:SFHS) using the Illumina Infinium MethylationEPIC array (16), capturing DNAm at
109 approximately 850K sites and in 2,449 individuals ($N_{antidepressant\ use}$=74) from the Netherlands
110 Twin Register (NTR) using the Infinium HumanMethylation450 BeadChip Kit (17), capturing
111 DNAm at 450K sites. We then conducted a meta-analysis using the two cohorts. Further, by
112 splitting GS:SFHS into two DNAm datasets where the number of participants differed
113 according to data availability, we were able to train a DNAm-based methylation score (MS)
114 of self-reported antidepressant use in one dataset (N=3,799; $N_{antidepressant\ use}$=585) and test
115 its ability to predict self-reported antidepressant use in a second (N=3,360; $N_{antidepressant}$
116 $_{use}$=317). We also addressed confounding by indication and by smoking in planned sensitivity
117 analyses (18). Using linked National Health Service (NHS) Scotland medication prescribing
118 data, we were able to identify whether the MWAS findings from self-report were also found
119 when using contemporaneous electronic healthcare information collected around the time
120 of the blood draw.
121
122
123
124
125
126
127
128

Methods

Study populations

Generation Scotland – the Scottish Family Health Study (GS:SFHS)

GS:SFHS is a family-based population cohort designed to investigate the genetic and environmental causes of common diseases and well-being in approximately 24,000 participants aged 18-98 years in Scotland. Baseline data was collected between 2006 and 2011 (19,20) and contains detailed information on a broad range of variables. DNA is also available from blood samples taken on more than 20,000 consenting participants.

GS:SFHS received ethical approval from NHS Tayside Research Ethics Committee (REC reference number 05/S1401/89) and has Research Tissue Bank Status (reference: 20/ES/0021). Written informed consent was obtained from all participants.

Netherlands Twin Register (NTR)

NTR is a population-based cohort of over 200,000 participants, consisting of twins, their parents, spouses, and siblings. Between 2004 and 2010, biosamples were collected in the NTR-Biobank project. Full details have been reported previously (21,22). DNA was collected from peripheral whole blood.

Informed consent was obtained from all participants. The study was approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Centre, Amsterdam, an Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB00002991 under Federal-wide Assurance-FWA00017598; IRB/institute codes, NTR 03-180).

Phenotypes

GS:SFHS

Self-reported antidepressant use

For self-reported antidepressant use, a text-based questionnaire was sent to participants 1-2 weeks before attending an in-person appointment where DNAm was collected. The questionnaire recorded medication use through a “yes/no” checkbox with the

160 following accompanying question: "Are you regularly taking any of the following
161 medications?" where available options included "Antidepressants".
162 Demographic data on individuals with self-reported antidepressant use is presented
163 in Table 1. Participants with no self-reported antidepressant use were defined as those
164 individuals who answered "No" to the "Antidepressants" sub-section of the questionnaire.
165 There were 6,428 individuals with available self-reported data who also had DNAm, lifestyle,
166 and MDD status data available.
167
168 *Selective serotonin reuptake inhibitors (SSRI) prescription and data linkage*
169 Here, we selected antidepressants from BNF paragraph code 4.3.3 containing the
170 following SSRIs: Citalopram, Escitalopram, Fluoxetine, Fluvoxamine Maleate, Paroxetine,
171 and Sertraline (23) for individuals with medication data in GS:SFHS (see Supplementary
172 Materials for more information). To investigate DNAm signatures of SSRI use, we restricted
173 analyses to those individuals with SSRI dispensing records in the 12 months prior to blood
174 draw date (N=401). Individuals with DNAm data that did not have any SSRI dispensed
175 prescription records were marked as controls (N=5,688).
176 We used dates of dispensing, not prescription, when restricting antidepressant use
177 based on time interval. Demographic data on individuals included here is presented in Table
178 1. Further details with regards to SSRI prescriptions are presented in Supplementary Table 1.
179
180 *Lifestyle factors and MDD status*
181 Body mass index (BMI) was computed using height (cm) and weight (kg) as measured
182 by clinical staff at baseline recruitment. Participants reported the number of units of alcohol
183 consumed during the past week and their smoking status (never, former, current); pack
184 years was used to measure heaviness of smoking in current smokers (24).
185 MDD status was assessed using the Structured Clinical Interview of the Diagnostic
186 and Statistical Manual, version IV (SCID) (25). Participants with no MDD were defined as
187 those individuals who did not fulfil criteria for a current or previous MDD diagnosis following
188 the SCID interview (26).
189 Further details about other diagnoses and medication intake are presented in
190 Supplementary Tables 2 and 3. Participants who suffered from other psychiatric disorders
191 (e.g. bipolar disorder, psychosis) were excluded from analyses (N=11).

192

193

194 ***NTR***

195 Demographic data on individuals with antidepressant use is presented in Table 2.

196 *Self-reported antidepressant use*

197 At blood draw, participants were asked about all current medication use and for all
198 medicines the dose, brand and chemical names were recorded directly from the medication
199 packaging. Following the Anatomical Therapeutic Chemical (ATC) classification system,
200 subclasses of N06A were considered antidepressants. Antidepressant medication use was
201 coded as 1 (current user) / 0 (non-user).

202

203 *Lifestyle factors and MDD status*

204 Body mass index (kg/m2) was computed based on weight and height obtained at the
205 time of blood sampling. Current and past smoking behaviour was dummy-coded as
206 current/former/never smokers and pack-years were recorded as the number of cigarettes
207 smoked per day/20 × number of years smoked.

208 Lifetime MDD was derived from the Composite International Diagnostic Interview
209 (CIDI) (27), the Lifetime Depression Assessment Self-report (LIDAS) (28), and DSM-oriented
210 Adult Self Report scales (ASR) of the Achenbach System of Empirically Based Assessment
211 (ASEBA) (29) (Supplementary Materials). Data from the CIDI, LIDAS and ASR scales were
212 combined to form one numeric variable coded as 1 (lifetime MDD case) / 0 (lifetime MDD
213 control). When data from multiple sources were available for an individual, the information
214 from CIDI was favoured over that of LIDAS and ASR, and data from LIDAS over that of ASR.

215

216 **DNA methylation**

217 ***GS:SFHS***

218 Genome-wide DNAm data profiled from whole blood samples was available for
219 9,873 individuals in GS:SFHS using the Illumina Human-MethylationEPIC BeadChip (16).
220 DNAm data for individuals was initially released in two waves (wave 1_N=5101; wave
221 2_N=4,450). Quality control (QC) and normalisation were conducted using R packages
222 ShinyMethyl (30) and watermelon (31). Details of the protocol are described in the
223 Supplementary Materials.

224 ***NTR***

225 DNAm in the NTR-Biobank study (21,32,33) was assessed with the Infinium
226 HumanMethylation450 BeadChip Kit ((17), Illumina, San Diego, CA, USA) by the Human
227 Genotyping facility (HugeF) of ErasmusMC, the Netherlands (http://www.glimdna.org/) as
228 part of the Biobank-based Integrative Omics Study (BIOS) consortium (34). DNAm
229 measurements have been described previously (33,34) (see Supplementary Materials for
230 more information).

231

232 **Statistical methods**

233 *Methylome-wide association*

234 *GS:SFHS*

235 We used linear regression models run in the "limma" package (35) in R to analyse
236 the association of each CpG site, included as an outcome variable, with self-reported
237 antidepressant use included as the predictor variable. The R code for the current analyses is
238 available in the Supplementary Materials. The following covariates were included: age, sex,
239 wave (indicating different DNAm data pre-processing waves), as well as BMI, alcohol units,
240 smoking status, and pack years to observe whether the inclusion of lifestyle factors
241 attenuates the effect of self-reported antidepressant use. MDD status was also included to
242 observe whether associations between CpGs and self-reported antidepressant use were
243 influenced by the presence of MDD and potentially confounded by the indication for
244 prescription, as some individuals not classed as having MDD have indicated antidepressant
245 use (N=320). There were 674,246 CpGs available after QC and a Bonferroni correction
246 (0.05/674,246) was used to define methylome-wide significance ($p \leq 7.42 \times 10^{-8}$).

247 Using the same model as above, we further carried out sensitivity analyses in (1)
248 non-smokers (N=3,351, $N_{antidepressant\ use}$=304), (2) individuals with (N=5,368, antidepressant
249 users=420) and without (N=5,268, antidepressant users=320) MDD, and (3) individuals with
250 SSRI dispense dates within 12 months prior to blood draw date (N=7,106, $N_{antidepressant}$
251 $_{use}$=401) (see Supplementary Materials for more information).

252 We compared findings in the main antidepressant use MWAS with summary
253 statistics from a MDD MWAS in the same sample (N=8,844, MDD cases=1,461), by looking at
254 effect sizes for significant CpGs in both analyses (Supplementary Figures 2 and 3).

255

256 *NTR*

257 The association between antidepressant medication and DNAm was tested under a
258 generalized estimation equation (GEE) model with DNAm M-values as outcome and self-
259 reported antidepressant use as a predictor. The model included the following covariates:
260 MDD status, sex, age at blood sampling, percentages of monocytes, eosinophils, and
261 neutrophils, HM450k array row, 96-wells bisulfite sample plate (dummy coding), BMI,
262 smoking, and smoking pack-years. MWAS analyses were performed with the R package
263 'gee'. The following settings were used: Gaussian link function (for continuous data), 100
264 iterations, and the 'exchangeable' option to account for the correlation structure within
265 families and within persons.

266

267 *Meta-analysis*

268 There were N=6,428 and N=2,449 participants in GS and NTR, respectively, and
269 N=318,856 overlapping CpGs between the two cohorts. Meta-analysis of these two datasets
270 was performed in METAL (36) using p-value based analysis (N=8,877). The meta-analysis
271 was based on N=318,856 total overlapping CpGs across GS:SFHS and NTR and a Bonferroni
272 correction (0.05/318,856) was used to define methylome-wide significance ($p \leq 1.57 \times 10^{-7}$).

273

274 *DNAm score analysis*

275 To create a training and testing dataset, we separated individuals in GS:SFHS by
276 DNAm pre-processing wave ($N_{wave\ 1}$=3,799; $N_{wave\ 2}$=3,360). This number differed from the
277 final number included in the MWAS (N=6,428) due to data availability for covariates
278 included in different analyses. R package "biglasso" was used to train DNAm predictors
279 (Supplementary Materials).

280 As in the MWAS, we carried out sensitivity analyses by training a score in wave 1 in
281 (1) non-smokers (N=1,952, $N_{antidepressant\ use}$=226) and (2) individuals with no MDD diagnosis
282 (N=2,791, antidepressant users=195) (see Supplementary Materials).

283 There were 76, 20, and 35 CpG sites identified for the full (MS), smoker-excluded
284 (MS-ns), and MDD-excluded (MS-control) datasets described above, respectively. The list of
285 CpG sites and their corresponding weights are presented in Supplementary Tables 4, 5, and
286 6. MS, MS-ns, and MS-control were created for each individual in wave 2 (N=3,360;
287 antidepressant users=317) by taking the sum of the product of the identified DNAm

288 residualised M-values and supplied model coefficient values. Regression models were then
289 run to identify associations between the scores and self-reported antidepressant use,
290 lifestyle factors and MDD.

291 *Pathway and regulatory element overlap analysis*

292 We used the Infinium MethylationEPIC BeadChip database to annotate significant
293 CpG sites in GS:SFHS to genes (16). The database provides information with regards to
294 genes, chromosome location, start and end sites, and other features.

295 To assess pathway enrichment for differentially methylated CpG sites while
296 correcting for biases in the representation of genes on the Infinium BeadChip, we used
297 missMethyl (37), accessed via methylGSA (38). Gene Ontology (GO) terms were accessed
298 using the msigdbr package (39). Pathways included in the analysis were all GO Biological
299 Process pathways of size 20-500 genes inclusive. CpG sites included in the analysis were
300 those significant at a threshold of $p<1x10^{-5}$ (N=144), as used in previous studies (40).

301 To assess overlap of differentially methylated CpG sites with the 15 chromatin states
302 across 127 tissues from the Consolidated Roadmap Epigenomics dataset, we used eFORGE
303 2.0 (41), accessed via the web tool at https://eforge.altiusinstitute.org and using default
304 settings.

305

306 *Identification of differentially-methylated regions (DMR)*

307 We used the dmrff function implemented in the dmrff R package (42) to identify
308 differentially methylated regions in GS:SFHS, defined as regions containing 2-30 CpG sites
309 separated by ≤ 500 bp. DMR analysis uses CpGs with unadjusted MWAS p-values ≤ 0.05 and
310 methylation changes in a consistent direction, and DMRs with a Bonferroni-corrected p-
311 value ≤ 0.05 are considered statistically significant.

312

313

314

315

316

317

318

319

320

321

322

323 **Results**

324

325 *Demographic characteristics*

326 Individuals with self-reported antidepressant use who had incomplete lifestyle (BMI,
327 alcohol consumption, and smoking) and disorder (MDD status) data were excluded. There
328 were 6,428 individuals in the final MWAS of self-reported antidepressant use in GS:SFHS
329 ($N_{antidepressant\ use}$: 740) and 2,449 ($N_{antidepressant\ use}$=74) in NTR. Descriptive and demographic
330 characteristics for these individuals, as well as individuals with prescribing data, are
331 presented in Table 1 for GS:SFHS and Table 2 for NTR.

332

333 *Methylome-wide association*

334

335 *GS:SFHS*

336 *Self-reported antidepressant use*

337 MWAS identified 10 CpG sites that were associated with self-reported
338 antidepressant use ($p \leq 7.42 \times 10^{-8}$). Nine CpG sites were significantly hypomethylated in
339 antidepressant users, while 1 site (cg26277237) was hypermethylated. Information about
340 each CpG site is shown in Table 3 and depicted as a Manhattan plot in Figure 1A.

341 The EWAS catalogue (http://www.ewascatalog.org/) includes CpG sites represented
342 on the Illumina 450K methylation array that have shown associations with traits at $p \leq 1 \times 10^{-4}$
343 in genome-wide analyses. The catalogue was used to cross-reference all CpG sites
344 associated with self-reported antidepressant use with the wider literature. cg05603985 was
345 previously found to be associated with smoking ($p=1.8 \times 10^{-43}$; (18)) and alcohol consumption
346 ($p=7.9 \times 10^{-13}$ in African ancestry and $p=1.9 \times 10^{-8}$ in European ancestry; (43); cg27589594 was
347 found to be associated with gestational age ($p=3.5 \times 10^{-16}$; (44)). Eight of the CpG sites were
348 represented only on the Infinium MethylationEPIC BeadChip (containing approximately
349 850K CpG sites) and were therefore not included in the catalogue. Searches conducted on
350 other databases, including EWASdb (45) indicate that these 8 CpG sites have not been
351 previously associated with any trait.

352 The Pearson's correlation between the MWAS effect sizes from the full-sample
353 MWAS and MWAS in non-smokers across all CpGs was *r*=0.655 (95% C.I.=0.636-0.674).
354 When restricting CpGs to those that reached a p-value≤$5x10^{-5}$, the correlation increased
355 further to *r*=0.877 (95% C.I.=0.869-0.884), and Bonferroni-corrected CpGs (N=10) in the full
356 sample and the same CpGs in the non-smoker sample had a high correlation (*r*=0.980, 95%
357 C.I.=0.979-0.981). Supplementary Table 7 indicates the effect size and nominal p-value for
358 all CpGs found to be associated with self-reported antidepressant use in the full sample and
359 in the non-smoker sample.
360 The Pearson's correlation between effect sizes obtained from the MWAS of
361 antidepressant use in individuals with MDD and MWAS of antidepressant use in those
362 without MDD across all CpGs was *r*=0.094 (95% C.I.=0.069-0.118). Restricting CpGs to a
363 significance cut-off of p≤$5x10^{-5}$ increased the correlation in effect sizes to *r*=0.961 (95%
364 C.I.=0.959-0.963). Supplementary Table 8 indicates the effect size and nominal p-value for
365 all CpGs found to be associated with self-reported antidepressant use in the full sample and
366 in the sensitivity analysis sample.
367
368 *SSRI*
369 MWAS conducted using individuals with dispensing data identified 12 CpG sites, all
370 hypomethylated, associated with SSRI use (the most commonly prescribed treatment for
371 depression) within 12 months prior to blood draw date, which are shown in Supplementary
372 Table 9 and Supplementary Figure 1.
373 Additionally, all 10 CpG sites that were associated with self-reported antidepressant
374 use (p≤$7.42x10^{-8}$) were found to be nominally significant (p<0.05) in the MWAS including
375 SSRI dispensing records as a phenotype. The direction of effect was consistent for all CpG
376 sites identified in the self-report MWAS: 9 were hypomethylated and 1 was
377 hypermethylated, presented in Supplementary Figure 1 and Supplementary Table 10 (β and
378 p-value for smallest and largest associations included cg09511513 (β=-0.026, p=0.028) and
379 cg26277237 (β=0.027, p=$3.09x10^{-7}$)). The majority (10/12) of CpG sites identified in the
380 MWAS of SSRI also were nominally significant in the MWAS where the self-report variable
381 was input as the predictor (Supplementary Table 11).
382
383 *NTR*

384 There were no methylome-wide significant findings in NTR. The top ten CpG sites
385 were annotated to genes implicated in psychiatric- and cognition-related phenotypes and in
386 disorders of the immune system and are presented Supplementary Table 12. Figure 1B
387 represents a Manhattan plot of NTR results.
388
389 *Meta-analysis*
390 Meta-analysis of antidepressant use MWAS across GS:SFHS and NTR revealed two
391 Bonferroni-corrected CpG sites, which were on the 450K array; both were identified in the
392 GS:SFHS MWAS; cg05603985 (β=-0.021, p=3.01x10^{-8}, q=0.009) had an opposite direction of
393 effect across cohorts (negative in GS:SFHS and positive in NTR), and cg27589594 (β=-0.024,
394 p=6.46x10^{-8}, q=0.021) had the same direction of effect (negative) in both. Methylome-wide
395 results are depicted in Figure 1C.
396 There were N=318,856 overlapping CpGs between GS:SFHS and NTR, and effect sizes
397 were not strongly correlated (Pearson's correlation *r*=0.081, 95% C.I.=0.06-0.102). When
398 restricting CpGs to those that reached a p-value≤0.05 across both MWAS (N=1,714), the
399 correlation between effect sizes increased (*r*=0.32, 95% C.I.=0.301-0.339). Further
400 information with regards to concordant signals across the two cohorts is in the
401 Supplementary Materials.
402
403 *DNAm score analysis*
404 LASSO regression selected 76, 20, and 35 CpGs to calculate MS, MS-ns, and MS-
405 control for self-reported antidepressant use, respectively, in GS:SFHS wave 2.
406 Briefly, the MS was associated with self-reported antidepressant use in a model
407 adjusted for age, sex, and 10 genetic PCs (β=0.377, p=3.12x10^{-11}, R^2=2.12%); After
408 adjustment for BMI, alcohol units, smoking status, pack years, and MDD, this association
409 remained significant, although the variance explained decreased (β=0.213, p=0.0035,
410 R^2=0.56%). The association between MS-ns and self-reported antidepressant use was in the
411 same direction but became non-significant, both in the covariate-free model (β=0.106,
412 p=0.075, R^2=0.16%) and in the model with the above covariates included (β= 0.04, p=0.565,
413 R^2=0.02%). Finally, MS-control was significantly associated with self-reported antidepressant
414 use in the covariate-free model (β=0.254, p=1.15x10^{-5}, R^2=0.93%), but not when including

415 covariates (β=0.064, p=0.381, R^2=0.05%). Table 4 indicates associations with lifestyle traits
416 and MDD for all scores.
417
418 *Pathway and regulatory element overlap analysis*
419 Following FDR adjustment for multiple comparisons, the only over-represented
420 pathways were related to regulation of myeloid cell differentiation (GO Negative Regulation
421 of Myeloid Cell Differentiation and GO Negative Regulation of Myeloid Leucocyte
422 Differentiation, $p_{adjusted}$=0.04 for both). Enrichment of these pathways was driven by
423 hypomethylation in the antidepressant use group in the body or 5'UTR of the following
424 genes: *GAT2, HOXA7, INPP5D, MEIS1, RAR* and *UBASH3B*.
425 Testing for overlap of differentially methylated CpG sites at $p<1x10^{-5}$ with chromatin
426 states in 127 tissues from the Roadmap Epigenomics Consortium, the greatest enrichment
427 was seen in blood cell tissues. Of the peripheral blood cell subsets tested, the cell type
428 showing the greatest enrichment was monocytes (Figure 2), with significant enrichment for
429 monocyte enhancers ($q=4x10^{-11}$) and regions flanking active transcription start sites in
430 monocytes ($q=1x10^{-9}$).
431 We repeated the tests above using the results from the MEWAS of antidepressant
432 prescriptions (N CpGs at $p<1x10^{-5}$=484). We did not find
433 enrichment of any specific GO biological processes but replicated the finding from the self-
434 reported antidepressant use MWAS that monocytes were the peripheral blood cell subset
435 most enriched in the MWAS signal (enhancers in primary monocytes from peripheral
436 blood were enriched at $q=4x10^{-94}$) and also found strong enrichment for enhancers in
437 peripheral blood neutrophils ($q=4x10^{-79}$).
438
439 *DMR analysis*
440 We identified 159 DMRs, ranging from 2-11 CpGs in length, to be associated with
441 self-reported antidepressant use using the DMR adjusted p-values (Supplementary Excel File
442 1). The region with the largest significance (p.adjust=7.54×10^{-12}) spanned 4 CpGs annotated
443 to *GATA2* on chromosome 3, a gene involved in stem cell maintenance and in which
444 mutations are associated with wide-ranging immunological disorders (46,47). The largest
445 DMR (p.adjust=2.04×10^{-8}) spanned 11 CpGs located on chromosome 2, where most were
446 annotated to *OTX1*, a gene previously associated with risk-taking behaviour (48). The gene

447 may also play a role in the brain and development of sensory organs (49). Only one DMR
448 spanning 2 CpGs (p.adjust=2.55×10^{-7}) contained a CpG that was also associated with the
449 trait at the level of individual probes (cg05186879, p=4.91×10^{-9}).

450 **Discussion**

451 Self-reported antidepressant use is associated with differences in DNAm at 10 CpG
452 sites in GS:SFHS located in genes that have previously been associated with psychiatric
453 disorders and the innate immune system. There were no methylome-wide findings in NTR.
454 We found that all 10 CpG sites in GS:SFHS were also nominally associated with
455 antidepressant prescription data obtained from the electronic health record, indicating
456 agreement between self-reported and record linkage data. Sensitivity analyses in GS:SFHS
457 indicated that results were not significantly confounded by smoking. As shown in
458 Supplementary Table 7, the effect for all 10 associated CpGs are in the same direction
459 across sensitivity analyses, and all p-values in the sensitivity analyses are nominally
460 significant. The effect sizes of the top 10 CpGs were also highly correlated in the full sample
461 and in the non-smoker sample (r=0.980). A DNAm score trained on one sub-sample was
462 associated with self-reported antidepressant use, four lifestyle factors, and MDD in a second
463 unrelated sample from the same study. Pathway analysis revealed an enrichment of
464 differentially methylated genes in myeloid cell differentiation. However, the effect size
465 correlation between the 10 associated CpGs and the same CpGs in a MDD MWAS (r=0.674)
466 may indicate that DNAm signatures of antidepressants act as a proxy for MDD, although
467 further analyses with much larger samples will be needed to disentangle the effects
468 between antidepressant use and MDD on DNAm. Lastly, meta-analysis of MWAS in GS:SFHS
469 and NTR identified two CpGs that were also found in GS:SFHS.

470 The CpG site with the largest effect size, both in GS:SFHS and the meta-analysis (β=-
471 0.055) was cg09511513, located in *ATP6V1B2*. This CpG site has not been associated with
472 any other traits previously, to the best of our knowledge. *ATP6V1B2* encodes a component
473 of vacuolar ATPase, which is a multisubunit enzyme that mediates the acidification of
474 eukaryotic intracellular organelles, including endosomes and lysosomes, and may be
475 involved in neurotransmission (50). Importantly, a single nucleotide polymorphism (SNP)
476 within *ATP6V1B2*, rs1106634, has exhibited a suggestive p-value in a meta-analysis in
477 schizophrenia and bipolar disorder (p=$3.97\text{x}10^{-6}$) and has been associated with lifetime risk
478 of depression (p<0.001) in a recent study (51–53).

479 In NTR, there were no methylome-wide significant findings. The 10 most
480 significant CpGs (Supplementary Table 12) were annotated to genes implicated in similar
481 processes as those found in GS:SFHS, including the innate immune and psychiatric diseases.
482 Although there were no significant findings, the pattern of results in NTR is similar to
483 significant findings in GS:SFHS.The two CpGs identified in the meta-analysis that were
484 present on both array types were also associated in GS:SFHS and were annotated to genes
485 previously implicated in psychiatric disease and the innate immune system (*SKI, SLC5A10*)
486 (see Table 3). These were not nominally significant in NTR (Supplementary Table 13).
487 cg05603985 was hypomethylated in GS:SFHS and hypermethylated in NTR, and cg27589594
488 was hypomethylated in both cohorts. Although the signal from one CpG was the same
489 across the two cohorts, further large-scale analyses on DNAm collected on the same array
490 will be needed to uncover the role of antidepressants in relation to psychiatric and immune
491 processes in DNAm.

492 We were able to show that a MS for self-reported antidepressant use was associated
493 with self-reported antidepressant use, MDD status, and 4 lifestyle factors (BMI, alcohol
494 units, smoking status, and pack years) in a subset of GS:SFHS. In addition, when excluding
495 MDD-associated signals in a second DNAm score (MS-control), we found that the predictor
496 was still associated with antidepressant use, although with reduced variance explained (MS
497 R^2=2.12%; MS-control R^2=0.93 %). This suggests that, although antidepressant use may in
498 part be a marker of MDD effects on DNAm, antidepressant prescription associations may be
499 partly independent from the condition for which they are prescribed. However, when
500 including lifestyle factors in the association between MS-control and antidepressant use, the
501 variance explained further decreased (R^2=0.05%). This may indicate that lifestyle risk factors
502 for MDD may play an important role in this association, even in the absence of an MDD
503 diagnosis.

504 The decreased variance explained may be in part due to differences in the training
505 sample sizes for the scores (MS_N=3,799, antidepressant users=585; MS-$control_N$=2,791,
506 antidepressant users=195; MS-ns_N=1,952, antidepressant users=226), although variation
507 explained in lifestyle factors is comparable between the MS and MS-control (MS R^2:
508 BMI=0.75%; alcohol units=1.93%; smoking status=6.11%; pack years=9.14%; MS-control R^2:
509 BMI=0.53%; alcohol units=1.56%; smoking status=5.08%; pack years=8.09%).

510 When smokers were excluded from the training sample, MS-ns associations with all
511 variables remained in the same direction but became non-significant. This suggests that
512 smoking may partially confound the associations between antidepressant use and DNAm.
513 However, when excluding smokers in our MWAS analyses, correlations between the effect
514 sizes in the full-sample and non-smoker MWAS indicated that results were not likely
515 confounded by smoking (*r*=0.980). Further, effects for all significant CpGs were in the same
516 direction in the reduced sample as in the full sample, and were all nominally significant in
517 the non-smoker sample (Supplementary Table 7). Future studies would benefit from
518 training DNAm predictors of self-reported antidepressant use in larger samples of lifelong
519 non-smokers.Only one CpG overlaps between the CpG sites in our MS and CpG sites
520 comprising an MDD DNAm risk score (MS) in Barbu et al. (2020) (15). This CpG (cg09935388)
521 has been associated with smoking in previous studies (54,55), and may therefore capture
522 smoking-associated effects related to both MDD and antidepressant use. Further, in both
523 studies, MDD MS and antidepressant use MS were both more strongly associated with
524 lifestyle factors than with the trait of interest (MDD MS R^2 in Barbu et al. (2020):
525 MDD=1.75%, BMI=0.097%, alcohol units=0.7%, smoking status=3.2%, pack years=6.5%;
526 Antidepressant use MS R^2: antidepressant use=2.12%, BMI=0.75%, alcohol units=1.93%,
527 smoking status=6.11%, pack years=9.14%). The stronger associations for both MS with
528 lifestyle factors may indicate that the scores capture an archive of relevant current and past
529 lifestyle factors that are also associated with depression and the prescription of
530 antidepressants. Previous large-scale studies have found profound effects of each of the
531 lifestyle factors on the methylome (56–58). However, the MS was still associated with
532 antidepressant use after accounting for the lifestyle factors mentioned above (R^2=0.56%),
533 indicating a possible independent effect of antidepressants. Due to the complex
534 associations between DNAm, MDD, and lifestyle factors, future large-scale studies, including
535 longitudinal studies, are warranted to investigate independent effects of each variable in
536 relation to MDD.
537 To investigate potential relationships between DNAm links to antidepressant use
538 and MDD, we conducted a MDD MWAS in GS (N=8,844, cases=1,641). While none of the top
539 10 CpGs in the antidepressant use MWAS were significantly associated with MDD, the effect
540 sizes showed a moderate correlation (*r*=0.674; See Supplementary Figures 2 and 3). This

541 suggests that DNAm signatures of antidepressant use may partially reflect confounding by
542 indication and much larger samples will be needed to disentangle this relationship.

543 Using missMethyl, we identified altered DNAm nearby several genes involved in
544 myeloid cell differentiation. Myeloid cells play an important role in the innate immune
545 response, and their activation and differentiation depend in part on epigenetic mechanisms
546 (59). Several of the genes annotated to CpGs identified at genome-wide significance here
547 were associated with the innate immune response in previous GWAS; this may complement
548 the finding that genes annotated to the 144 CpG sites differentially methylated at $p<1x10^{-5}$
549 are enriched for myeloid cell differentiation.

550 Further, the differentially-methylated regions were annotated to genes that play a
551 role in cell maintenance, as well as brain and sensory organ development (46,49).
552 Importantly, mutations in *GATA2* have previously been related to wide-ranging
553 immunological disorders, providing further evidence linking antidepressant use and the
554 immune system (47).

555 The role of the innate immune system has previously been investigated in relation to
556 psychiatric disorders. In MDD, studies have found a number of pro-inflammatory cytokines,
557 such as TNF-α and IL-6, to have higher concentrations in depressed individuals, as compared
558 to healthy controls (60). Importantly, antidepressants, particularly SSRIs, have been shown
559 to exert effects on the immune system by causing a reduction in inflammatory markers.
560 Generally, antidepressants have been found to reduce levels of pro-inflammatory cytokine,
561 such as TNF-α, IL-6, and IL-1β, although their effects are complex and incompletely
562 understood (60). The current study suggests novel links between DNAm, antidepressants,
563 and processes involved in the innate immune system. Studies that investigate links between
564 antidepressants and DNAm alterations at CpG sites that are known to be associated with
565 the expression of innate immune system, myeloid cell differentiation, and MDD-specific
566 genes may be of interest.

567 There are a number of potential limitations to the current study. Firstly, although
568 self-reported antidepressant use could not be restricted to specific dates prior to the blood
569 draw date in GS:SFHS, the variable is considered contemporaneous to blood draw date as
570 self-report questionnaires were sent 1-2 weeks before the clinic appointment. However, the
571 questionnaire did not record types of antidepressants, which may explain the different top
572 CpG sites based on the predictor input in MWAS. The different CpG sites identified in each

573 MWAS may also be due in part to the proportion of overlapping individuals used in each
574 analysis ($N_{overlap\ in\ antidepressant\ users}$=289; $N_{overlap\ in\ controls}$=4,811). Nonetheless, as indicated in
575 Supplementary Table 11, most of the CpG sites significantly associated with SSRI use within
576 12 months prior to blood draw date were nominally significant in the MWAS where self-
577 reported antidepressant use was fit as the predictor. This pattern was also observed when
578 investigating the top 10 CpG sites associated with the self-reported measure in the SSRI use
579 MWAS (Supplementary Table 10). The findings above indicate there is agreement between
580 different forms of data collection, specifically self-reported medication use and record
581 linkage to PIS. The variables have previously been investigated in a study by Hafferty et al.
582 (2018) (61), who showed that self-reported antidepressant use in GS:SFHS showed very
583 good agreement with record linkage data at 3- and 6-month fixed time windows (61).

584 Secondly, since we chose to restrict PIS data linkage to 12 months before the blood
585 draw date, it seems likely that the DNAm changes reported may occur in response to
586 antidepressant treatment, although this requires further replication. Nevertheless, the
587 DNAm data provided here is cross-sectional, and it is not possible to show how the
588 trajectory of DNAm at each CpG site evolved over time. Future studies should therefore
589 collect DNAm at multiple time points and further address the possibility of confounding by
590 antidepressant indication.

591 It should also be noted that SSRI prescribing data is unable to confirm whether
592 individuals took the medication as prescribed (62). In addition, the current study is based
593 exclusively on European ancestry cohorts, and the generalisability to diverse ancestries is
594 unknown.

595 Finally, correlations of effect sizes between the two cohorts are of low magnitude
596 and confidence intervals are large. However, the top 10 CpGs identified in NTR, although
597 not significant, were implicated in similar processes as those in GS:SFHS, and the meta-
598 analysis uncovered two CpGs that were also uncovered in GS:SFHS. These novel findings
599 should serve as a stimulus for larger, international collaborations and meta-analyses in
600 future.

601 In conclusion, we conducted an MWAS of self-reported antidepressant use across
602 two large, population-based cohorts and identified 10 novel CpG sites in GS:SFHS located in
603 genes previously associated with mental health disorders, implicating the innate immune
604 system. Further, MS predictors were associated with self-reported antidepressant use,

605 MDD, and a number of lifestyle factors in a second sample from the same study. SSRI use
606 derived from linkage health records showed convergent effects indicating agreement with
607 the findings derived using only self-report. Finally, a meta-analysis revealed similar findings
608 to the ones in GS:SFHS. Our findings highlight biological processes that may be relevant to
609 furthering our understanding of antidepressant actions and their side effects.
610
611
612 **References**

613 1. World Health Organization. Depression and Other Common Mental Disorders. Who.
614 2017;24.
615 2. Wray NR, Gottesman II. Using summary data from the Danish National Registers to
616 estimate heritabilities for schizophrenia, bipolar disorder, and major depressive
617 disorder. Front Genet. 2012;3(JUL).
618 3. I K. Antidepressants and the Placebo Effect. Z Psychol. 2014;222(3).
619 4. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, et al.
620 Comparative efficacy and acceptability of 21 antidepressant drugs for the acute
621 treatment of adults with major depressive disorder: a systematic review and network
622 meta-analysis. Lancet [Internet]. 2018;391(10128):1357–66. Available from:
623 http://dx.doi.org/10.1016/S0140-6736(17)32802-7
624 5. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al.
625 Acute and longer-term outcomes in depressed outpatients requiring one or several
626 treatment steps: A STAR*D report. Am J Psychiatry [Internet]. 2006 [cited 2020 Nov
627 19];163(11):1905–17. Available from: https://pubmed.ncbi.nlm.nih.gov/17074942/
628 6. Keers R, Uher R. Gene-environment interaction in major depression and
629 antidepressant treatment response. Vol. 14, Current Psychiatry Reports. Curr
630 Psychiatry Rep; 2012. p. 129–37.
631 7. Berton O, Nestler EJ. New approaches to antidepressant drug discovery: Beyond
632 monoamines. Vol. 7, Nature Reviews Neuroscience. Nat Rev Neurosci; 2006. p. 137–
633 51.
634 8. Menke A, Binder EB. Epigenetic alterations in depression and antidepressant
635 treatment. Dialogues Clin Neurosci. 2014;16(3):395–404.
636 9. Le François B, Soo J, Millar AM, Daigle M, Le Guisquet AM, Leman S, et al. Chronic

637 mild stress and antidepressant treatment alter 5-HT1A receptor expression by
638 modifying DNA methylation of a conserved Sp4 site. Neurobiol Dis. 2015 Oct
639 1;82:332–41.
640 10. Zimmermann N, Zschocke J, Perisic T, Yu S, Holsboer F, Rein T. Antidepressants inhibit
641 DNA methyltransferase 1 through reducing G9a levels. Biochem J. 2012 Nov
642 15;448(1):93–102.
643 11. Bird A. Perceptions of epigenetics. Vol. 447, Nature. Nature Publishing Group; 2007.
644 p. 396–8.
645 12. Jovanova OS, Nedeljkovic I, Spieler D, Walker RM, Liu C, Luciano M, et al. DNA
646 methylation signatures of depressive symptoms in middle-aged and elderly persons:
647 Meta-analysis of multiethnic epigenome-wide studies. JAMA Psychiatry.
648 2018;75(9):949–59.
649 13. McCartney DL, Hillary RF, Stevenson AJ, Ritchie SJ, Walker RM, Zhang Q, et al.
650 Epigenetic prediction of complex traits and death. Genome Biol. 2018;19(1):136.
651 14. Clark SL, Hattab MW, Chan RF, Shabalin AA, Han LKM, Zhao M, et al. A methylation
652 study of long-term depression risk. Mol Psychiatry. 2019 Jun 1;25(6):1334–43.
653 15. Barbu MC, Shen X, Walker RM, Howard DM, Evans KL, Whalley HC, et al. Epigenetic
654 prediction of major depressive disorder. Mol Psychiatry [Internet]. 2020 Jun 10 [cited
655 2020 Jun 11];1–12. Available from: http://www.nature.com/articles/s41380-020-
656 0808-3
657 16. Hansen K. IlluminaHumanMethylationEPICanno.ilm10b2.hg19: Annotation for
658 Illumina's EPIC methylation arrays. R Packag version 060 [Internet]. 2016; Available
659 from: https://bitbucket.com/kasperdanielhansen/Illumina_EPIC
660 17. Hansen K. IlluminaHumanMethylation450kanno. ilmn12. hg19: annotation for
661 Illumina's 450k methylation arrays. 2016;
662 18. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al.
663 Epigenetic Signatures of Cigarette Smoking. Circ Cardiovasc Genet. 2016;436–47.
664 19. Smith BH, Campbell H, Blackwood D, Connell J, Connor M, Deary IJ, et al. Generation
665 Scotland□: the Scottish Family Health Study□; a new resource for researching genes
666 and heritability. BMC Med Genomics. 2006;9:1–9.
667 20. Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, et al. Cohort
668 Profile□: Generation Scotland□: Scottish Family Health Study (GS□: SFHS). The study ,

669 its participants and their potential for genetic research on health and illness. Int J
670 Epidemiol. 2013;(July 2012):689–700.
671 21. Willemsen G, Geus EJC de, Bartels M, Beijsterveldt CEMT van, Brooks AI, Burk GFE, et
672 al. The Netherlands Twin Register Biobank: A Resource for Genetic Epidemiological
673 Studies. Twin Res Hum Genet [Internet]. 2010 Jun [cited 2021 Sep 7];13(3):231–45.
674 Available from: https://www.cambridge.org/core/journals/twin-research-and-
675 human-genetics/article/netherlands-twin-register-biobank-a-resource-for-genetic-
676 epidemiological-studies/94AA0448D70BFEA4780D0CD01EB7DF05
677 22. Ligthart L, Beijsterveldt CEM van, Kevenaar ST, Zeeuw E de, Bergen E van, Bruins S, et
678 al. The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-
679 Family Designs. Twin Res Hum Genet [Internet]. 2019 Dec 1 [cited 2021 Sep
680 7];22(6):623–36. Available from: https://www.cambridge.org/core/journals/twin-
681 research-and-human-genetics/article/netherlands-twin-register-longitudinal-
682 research-based-on-twin-and-twinfamily-
683 designs/C68488FDD92716CB9ED7C040B9776E2F
684 23. Britain RPS of G. British National Formulary 58. 2009.
685 24. Karen Leffondré, Michal Abrahamowicz, Jack Siemiatycki BR. Modeling Smoking
686 History: A Comparison of Different Approaches. Am J Epidemiol. 2002;156(9):813–23.
687 25. First MB, Spitzer RL, Gibbon M, Williams JBW, Davies M, Borus J, et al. The structured
688 clinical interview for DSM-III-R personality disorders (SCID-II). Part II: Multi-site test-
689 retest reliablity study. J Pers Disord. 1995 Feb 8;9(2):92–104.
690 26. Fernandez-Pujals AM, Adams MJ, Thomson P, McKechanie AG, Blackwood DHR, Smith
691 BH, et al. Epidemiology and heritability of major depressive disorder, stratified by age
692 of onset, sex, and illness course in generation Scotland: Scottish family health study
693 (GS: SFHS). PLoS One. 2015 Nov 1;10(11).
694 27. Kessler RC, Wittchen H-U, Abelson JM, Mcgonagle K, Schwarz N, Kendler KS, et al.
695 Methodological studies of the Composite International Diagnostic Interview (CIDI) in
696 the US national comorbidity survey (NCS). Int J Methods Psychiatr Res [Internet].
697 1998 Feb 1 [cited 2021 Sep 7];7(1):33–55. Available from:
698 https://onlinelibrary.wiley.com/doi/full/10.1002/mpr.33
699 28. Bot M, Middeldorp CM, Geus EJC de, Lau HM, Sinke M, Nieuwenhuizen B van, et al.
700 Validity of LIDAS (LIfetime Depression Assessment Self-report): a self-report online

701 assessment of lifetime major depressive disorder. Psychol Med [Internet]. 2017 Jan 1
702 [cited 2021 Sep 7];47(2):279–89. Available from:
703 https://www.cambridge.org/core/journals/psychological-medicine/article/validity-of-
704 lidas-lifetime-depression-assessment-selfreport-a-selfreport-online-assessment-of-
705 lifetime-major-depressive-disorder/97602A258E2DABA0745F8039C0F20907
706 29. The Achenbach System of Empirically Based Assessment (ASEBA) for Ages 18 to 90
707 Years. - PsycNET [Internet]. [cited 2021 Sep 7]. Available from:
708 https://psycnet.apa.org/record/2004-14941-004
709 30. Fortin J-P, Fertig E, Hansen K. shinyMethyl: interactive quality control of Illumina 450k
710 DNA methylation arrays in R. F1000Research [Internet]. 2014 Jul 30 [cited 2020 Nov
711 9];3:175. Available from: https://doi.org/10.12688/f1000research.4680.1
712 31. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven
713 approach to preprocessing Illumina 450K methylation array data. BMC Genomics
714 2013 141 [Internet]. 2013 May 1 [cited 2021 Sep 30];14(1):1–10. Available from:
715 https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-293
716 32. Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, Sondervan D, et al.
717 Genome-wide association of major depression: description of samples for the GAIN
718 Major Depressive Disorder Study: NTR and NESDA biobank projects. Eur J Hum Genet
719 2008 163 [Internet]. 2008 Jan 16 [cited 2021 Sep 7];16(3):335–42. Available from:
720 https://www.nature.com/articles/5201979
721 33. van Dongen J, Nivard MG, Willemsen G, Hottenga J-J, Helmer Q, Dolan C V., et al.
722 Genetic and environmental influences interact with age and sex in shaping the human
723 methylome. Nat Commun 2016 71 [Internet]. 2016 Apr 7 [cited 2021 Sep 7];7(1):1–
724 13. Available from: https://www.nature.com/articles/ncomms11115
725 34. Bonder MJ, Luijk R, Zhernakova D V., Moed M, Deelen P, Vermaat M, et al. Disease
726 variants alter transcription factor levels and methylation of their binding sites. Nat
727 Genet. 2017 Jan 1;49(1):131–8.
728 35. Matthew E Ritchie, Belinda Phipson, Di Wu, Yifang Hu, Charity W Law, Wei Shi GKS.
729 limma powers differential expression analyses for RNA-sequencing and microarray
730 studies. Nucleic Acids Res. 2015;43(7):47.
731 36. Willer CJ, Li Y, Abecasis GR. METAL: Fast and efficient meta-analysis of genomewide
732 association scans. Bioinformatics [Internet]. 2010 Jul 8 [cited 2020 Nov

733 9];26(17):2190–1. Available from: https://pubmed.ncbi.nlm.nih.gov/20616382/
734 37. Phipson B, Maksimovic J, Oshlack A. MissMethyl: An R package for analyzing data
735 from Illumina's HumanMethylation450 platform. Bioinformatics. 2016;32(2):286–8.
736 38. Ren X, Kuan PF. methylGSA: a Bioconductor package and Shiny app for DNA
737 methylation data length bias adjustment in gene set testing. Bioinformatics.
738 2019;35(11):1958–9.
739 39. Dolgalev I. msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format.
740 R package version 7.1.1. [Internet]. 2020 [cited 2020 Jul 29]. Available from:
741 https://cran.r-project.org/web/packages/msigdbr/index.html
742 40. Chu AY, Tin A, Schlosser P, Ko YA, Qiu C, Yao C, et al. Epigenome-wide association
743 studies identify DNA methylation associated with kidney function. Nat Commun
744 [Internet]. 2017 Dec 1;8(1):1–12. Available from:
745 www.nature.com/naturecommunications
746 41. Breeze CE, Reynolds AP, Van Dongen J, Dunham I, Lazar J, Neph S, et al. EFORGE v2.0:
747 Updated analysis of cell type-specific signal in epigenomic data. Bioinformatics.
748 2019;35(22):4767–9.
749 42. Suderman M, Staley J, French R, Arathimos R, Simpkin A, Tilling K. dmrff: identifying
750 differentially methylated regions efficiently with power and control. bioRxiv
751 [Internet]. 2018 Dec 31 [cited 2021 May 28];508556. Available from:
752 https://doi.org/10.1101/508556
753 43. Liu C, Marioni RE, Hedman ÅK, Pfeiffer L, Tsai P, Reynolds LM, et al. A DNA
754 methylation biomarker of alcohol consumption. Mol Psychiatry. 2018;(February
755 2016):422–33.
756 44. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, et al. Prediction of
757 gestational age based on genome-wide differentially methylated regions. Genome
758 Biol. 2016 Oct 7;17(1).
759 45. Liu D, Zhao L, Wang Z, Zhou X, Fan X, Li Y, Xu J, Hu S, Niu M, Song X LY. EWASdb:
760 epigenome-wide association study database. Nucleic Acids Res. 2019;47:989–93.
761 46. Li Y, Qi X, Liu B, Huang H. The STAT5–GATA2 Pathway Is Critical in Basophil and Mast
762 Cell Differentiation and Maintenance. J Immunol [Internet]. 2015 May 1 [cited 2021
763 May 31];194(9):4328–38. Available from:
764 http://www.jimmunol.org/content/194/9/4328http://www.jimmunol.org/content/1

765 94/9/4328.full#ref-list-1
766 47. Collin M, Dickinson R, Bigley V. Haematopoietic and immune defects associated with
767 GATA2 mutation [Internet]. Vol. 169, British Journal of Haematology. Blackwell
768 Publishing Ltd; 2015 [cited 2021 May 31]. p. 173–87. Available from:
769 https://onlinelibrary.wiley.com/doi/full/10.1111/bjh.13317
770 48. Karlsson Linnér R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al.
771 Genome-wide association analyses of risk tolerance and risky behaviors in over 1
772 million individuals identify hundreds of loci and shared genetic influences. Nat Genet
773 [Internet]. 2019 Feb 1 [cited 2020 Nov 9];51(2):245–57. Available from:
774 /pmc/articles/PMC6713272/?report=abstract
775 49. Frantz GD, Weimann JM, Levin ME, McConnell SK. Otx1 and Otx2 define layers and
776 regions in developing cerebral cortex and cerebellum. J Neurosci [Internet]. 1994 Oct
777 1 [cited 2021 May 31];14(10):5725–40. Available from:
778 https://www.jneurosci.org/content/14/10/5725
779 50. Petschner P, Gonda X, Baksa D, Eszlari N, Trivaks M, Juhasz G, et al. Genes Linking
780 Mitochondrial Function, Cognitive Impairment and Depression are Associated with
781 Endophenotypes Serving Precision Medicine. Vol. 370, Neuroscience. Elsevier Ltd;
782 2018. p. 207–17.
783 51. Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci
784 associated with schizophrenia and bipolar disorder. Schizophr Res. 2010 Dec;124(1–
785 3):192–9.
786 52. Shyn SI, Shi J, Kraft JB, Potash JB, Knowles JA, Weissman MM, et al. Novel loci for
787 major depression identified by genome-wide association study of Sequenced
788 Treatment Alternatives to Relieve Depression and meta-analysis of three studies. Mol
789 Psychiatry. 2011 Feb;16(2):202–15.
790 53. Gonda X, Eszlari N, Anderson IM, Deakin JFW, Bagdy G, Juhasz G. Association of
791 ATP6V1B2 rs1106634 with lifetime risk of depression and hippocampal
792 neurocognitive deficits: possible novel mechanisms in the etiopathology of
793 depression. Transl Psychiatry. 2016 Nov 8;6(11):e945.
794 54. Joubert BR, Håberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K
795 epigenome-wide scan identifies differential DNA methylation in newborns related to
796 maternal smoking during pregnancy. Environ Health Perspect. 2012

797 Oct;120(10):1425–31.
798 55. Küpers LK, Xu X, Jankipersadsing SA, Vaez A, La Bastide-van Gemert S, Scholtens S,
799 Nolte IM, Richmond RC, Relton CL, Felix JF DL. DNA methylation mediates the effect
800 of maternal smoking during pregnancy on birthweight of the offspring. Int J
801 Epidemiol. 2015;44(4):1224–37.
802 56. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide
803 association study of body mass index, and the adverse outcomes of adiposity. Nature
804 [Internet]. 2017 Jan 5 [cited 2021 Mar 29];541(7635):81–6. Available from:
805 https://www.nature.com/articles/nature20784
806 57. Liu C, Marioni RE, Hedman AK, Pfeiffer L, Tsai PC, Reynolds LM, et al. A DNA
807 methylation biomarker of alcohol consumption. Mol Psychiatry. 2018 Feb
808 1;23(2):422–33.
809 58. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al.
810 Epigenetic Signatures of Cigarette Smoking. Circ Cardiovasc Genet [Internet]. 2016
811 Oct 1 [cited 2021 Mar 29];9(5):436–47. Available from:
812 http://circgenetics.ahajournals.org/lookup/suppl/doi:10.1161/CIRCGENETICS.116.00
813 1506/-/DC1.
814 59. Álvarez-Errico D, Vento-Tormo R, Sieweke M, Ballestar E. Epigenetic control of
815 myeloid cell differentiation, identity and function [Internet]. Vol. 15, Nature Reviews
816 Immunology. Nature Publishing Group; 2015 [cited 2020 Jul 29]. p. 7–17. Available
817 from: https://pubmed.ncbi.nlm.nih.gov/25534619/
818 60. Eyre HA, Lavretsky H, Kartika J, Qassim A, Baune BT. Modulatory Effects of
819 Antidepressant Classes on the Innate and Adaptive Immune System in Depression.
820 Vol. 49, Pharmacopsychiatry. Georg Thieme Verlag; 2016. p. 85–96.
821 61. Hafferty JD, Campbell AI, Navrady LB, Adams MJ, MacIntyre D, Lawrie SM, et al. Self-
822 reported medication use validated through record linkage to national prescribing
823 data. J Clin Epidemiol. 2018 Feb 1;94:132–42.
824 62. Arnet I, Abraham I, Messerli M, Hersberger KE. A method for calculating adherence to
825 polypharmacy from dispensing data records. Int J Clin Pharm. 2014 Nov 29;36(1):192–
826 201.
827 63. Battram T, Richmond RC, Baglietto L, Haycock PC, Perduca V, Bojesen SE, et al.
828 Appraising the causal relevance of DNA methylation for risk of lung cancer. Int J

829 Epidemiol [Internet]. 2019 Oct 1 [cited 2020 Nov 10];48(5):1493–504. Available from:
830 https://academic.oup.com/ije/article/48/5/1493/5573019
831 64. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to
832 1.2 million individuals yield new insights into the genetic etiology of tobacco and
833 alcohol use [Internet]. Vol. 51, Nature Genetics. Nature Publishing Group; 2019 [cited
834 2020 Nov 9]. p. 237–44. Available from: /pmc/articles/PMC6358542/?report=abstract
835 65. Chen MH, Raffield LM, Mousas A, Sakaue S, Huffman JE, Moscati A, et al. Trans-ethnic
836 and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global
837 Populations. Cell [Internet]. 2020 Sep 3 [cited 2020 Nov 10];182(5):1198-1213.e14.
838 Available from: https://pubmed.ncbi.nlm.nih.gov/32888493/
839 66. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic
840 Functional Enrichment to Improve GWAS Power. Am J Hum Genet [Internet]. 2019
841 Jan 3 [cited 2020 Nov 9];104(1):65–75. Available from:
842 https://pubmed.ncbi.nlm.nih.gov/30595370/
843 67. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of
844 Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell
845 [Internet]. 2016 Nov 17 [cited 2020 Nov 19];167(5):1415-1429.e19. Available from:
846 https://pubmed.ncbi.nlm.nih.gov/27863252/
847 68. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe
848 interactions have shaped the genetic architecture of inflammatory bowel disease.
849 Nature [Internet]. 2012 Nov 1 [cited 2020 Nov 19];491(7422):119–24. Available from:
850 https://pubmed.ncbi.nlm.nih.gov/23128233/
851 69. Liu JZ, Van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association
852 analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight
853 shared genetic risk across populations. Nat Genet [Internet]. 2015 Aug 27 [cited 2020
854 Nov 19];47(9):979–86. Available from: https://pubmed.ncbi.nlm.nih.gov/26192919/
855 70. Din L, Sheikh M, Kosaraju N, Smedby KE, Bernatsky S, Berndt SI, et al. Genetic overlap
856 between autoimmune diseases and non-Hodgkin lymphoma subtypes. Genet
857 Epidemiol [Internet]. 2019 Oct 1 [cited 2020 Nov 19];43(7):844–63. Available from:
858 https://pubmed.ncbi.nlm.nih.gov/31407831/
859 71. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and
860 polygenic prediction from a genome-wide association study of educational

861 attainment in 1.1 million individuals. Nat Genet. 2018;50(8):1112–21.
862 72. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, et al. Shared genetic and
863 experimental links between obesity-related traits and asthma subtypes in UK
864 Biobank. J Allergy Clin Immunol [Internet]. 2020 Feb 1 [cited 2020 Nov
865 19];145(2):537–49. Available from: https://pubmed.ncbi.nlm.nih.gov/31669095/
866 73. Lam M, Chen CY, Li Z, Martin AR, Bryois J, Ma X, et al. Comparative genetic
867 architectures of schizophrenia in East Asian and European populations. Nat Genet
868 [Internet]. 2019 Dec 1 [cited 2020 Nov 19];51(12):1670–8. Available from:
869 https://pubmed.ncbi.nlm.nih.gov/31740837/
870 74. Edwards TL, Giri A, Hellwege JN, Hartmann KE, Stewart EA, Jeff JM, et al. A trans-
871 ethnic genome-wide association study of uterine fibroids. Front Genet [Internet].
872 2019 [cited 2020 Nov 19];10(JUN). Available from:
873 https://pubmed.ncbi.nlm.nih.gov/31249589/
874 75. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic
875 analysis of over 1 million people identifies 535 new loci associated with blood
876 pressure traits. Nat Genet [Internet]. 2018 Oct 1 [cited 2020 Nov 10];50(10):1412–25.
877 Available from: https://pubmed.ncbi.nlm.nih.gov/30224653/
878 76. Wang H, Yang J, Schneider JA, De Jager PL, Bennett DA, Zhang HY. Genome-wide
879 interaction analysis of pathological hallmarks in Alzheimer's disease. Neurobiol Aging
880 [Internet]. 2020 Sep 1 [cited 2020 Nov 10];93:61–8. Available from:
881 https://pubmed.ncbi.nlm.nih.gov/32450446/
882 77. Gallois A, Mefford J, Ko A, Vaysse A, Julienne H, Ala-Korpela M, et al. A
883 comprehensive study of metabolite genetics reveals strong pleiotropy and
884 heterogeneity across time and context. Nat Commun [Internet]. 2019 Dec 1 [cited
885 2020 Nov 19];10(1). Available from: https://pubmed.ncbi.nlm.nih.gov/31636271/
886 78. Inouye M, Ripatti S, Kettunen J, Lyytikäinen LP, Oksala N, Laurila PP, et al. Novel Loci
887 for Metabolic Networks and Multi-Tissue Expression Studies Reveal Genes for
888 Atherosclerosis. PLoS Genet [Internet]. 2012 Aug [cited 2020 Nov 19];8(8). Available
889 from: https://pubmed.ncbi.nlm.nih.gov/22916037/
890 79. Li M, Maruthur NM, Loomis SJ, Pietzner M, North KE, Mei H, et al. Genome-wide
891 association study of 1,5-anhydroglucitol identifies novel genetic loci linked to glucose
892 metabolism. Sci Rep [Internet]. 2017 Dec 1 [cited 2020 Nov 19];7(1). Available from:

https://pubmed.ncbi.nlm.nih.gov/28588231/

80. Li Q, Wineinger NE, Fu DJ, Libiger O, Alphs L, Savitz A, et al. Genome-wide association study of paliperidone efficacy. Pharmacogenet Genomics [Internet]. 2017 Jan 1 [cited 2020 Nov 19];27(1):7–18. Available from: https://pubmed.ncbi.nlm.nih.gov/27846195/

Acknowledgements

Generation Scotland is currently supported by the Wellcome Trust Investigator Award in Science 01/06/2021 to 31/05/26 'Exploiting genomic approaches to identify the environmental basis of depression'. (Reference: 220857/Z/20/Z) to McIntosh AM (PI). The DNA 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: McIntosh AM) and through funding from NARSAD (Ref: 27404; PI: Dr DM Howard and Ref: 21956; PI: Dr Kathryn Evans) and the Royal College of Physicians of Edinburgh (Sim Fellowship; PI: Dr HC Whalley). Dr CA was supported by an MRC University Unit award to the MRC Human Genetics Unit, University of Edinburgh, MC_UU_00007/10. 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 of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Dr DMH is supported by a Sir Henry Wellcome Postdoctoral Fellowship (Reference 213674/Z/18/Z). Dr MCB is supported by a Guarantors of Brain Non-clinical Post-Doctoral Fellowship. Dr EB is an NIHR Senior Investigator and is a member of the scientific advisory board of Sosei Heptares. Dr ML is supported by a fellowship from the Medical Research Council (MR/S006257/1).

Financial disclosures

AMM has previously received grant support from Pfizer, Lilly and Janssen. These studies are not connected to the current investigation. Remaining authors report no conflicts of interest.

957
958
959
960
961
962
963
964
965

	Self-reported AD use sample		SSRI sample from NHS Scotland records	
Demographic characteristics	**AD use** (N=740)	**No AD use** (N=5,688)	**SSRI use in 12-month interval** (N=401)	**No SSRI use** (N=6,705)
Age				
Mean (SD), range	51.33 (11.07), 18 - 87	49.5 (13.73), 18 - 87	49.16 (12.35), 18 – 85	50.21 (13.65), 18 - 94
Sex (%)				
Female	564 (76%)	3,336 (59%)	364 (75%)	3,631 (54%)
Male	176 (24%)	2,352 (41%)	123 (25%)	3,074 (46%)
Wave (%)				
1	480 (65%)	2,918 (51%)	260 (65%)	3,389 (51%)
2	260 (35%)	2,770 (49%)	141 (35%)	3,316 (49%)
BMI				
Mean (SD), range	28.17 (5.58), 16.11 – 51.29	26.60 (4.95), 14.78 – 67.62	28.37 (5.97), 17.50 – 51.29	26.68 (4.96), 15.93 – 67.62
Alcohol units				
Mean (SD), range	8.99 (11.75), 0 - 105	10.56 (11.36), 0 – 146	9.46 (10.59), 0 – 72	10.91 (12.24), 0 - 326
Smoking status (%)				
Current smoker	192 (26%)	861 (15%)	104 (26%)	1,006 (15%)
Former smokers (quit < 1 year ago)	15 (2%)	146 (2%)	2 (1%)	154 (2%)
Former smokers (quit > 1 year ago)	229 (31%)	1,634 (29%)	116 (30%)	1,907 (29%)
Never smoked tobacco	304 (41%)	3,047 (54%)	179 (43%)	3,638 (54%)
Pack years				
Mean (SD), range	11.03 (16.03), 0 - 116	7.01 (13.51), 0 - 133	9.59 (14.58), 0 – 72.85	7.27 (14.07), 0 - 133
MDD status (%)				
Cases	420 (57%)	740 (13%)	201 (50%)	706 (11%)
Controls	320 (43%)	4,948 (87%)	200 (50%)	5,999 (89%)

966 **Table 1**. Demographic characteristics for GS individuals with self-reported antidepressant
967 use and SSRI prescribing data in 12 months prior to blood draw date included in MWAS,
968 including lifestyle variables and MDD; AD=antidepressant.

	Self-reported AD use sample	
Demographic characteristics	**AD use** (N=74)	**No AD use** (N=2,375)
Age		
Mean (SD), range	37.74 (12.20), 18 – 72	36.76 (3.97), 18 - 80
Sex (%)		
Female	52 (70%)	1,618 (68%)
Male	22 (30%)	757 (32%)
BMI		
Mean (SD), range	25.06 (4.25), 18.30 – 34.60	24.24 (3.97), 14.50 – 50.70
Smoking status (%)		
Current smokers	13 (17%)	434 (18%)
Former smokers	19 (26%)	506 (21%)
Never smoked tobacco	42 (57%)	1,435 (61%)
Pack years		
Mean (SD), range	3.56 (5.83), 0 - 25	3.95 (8.69), 0 – 105
MDD status (%)		
Cases	10 (14%)	395 (17%)
Controls	64 (86%)	1,980 (83%)

Table 2. Demographic characteristics for NTR individuals with self-reported antidepressant use data within a week of blood draw included in MWAS, including lifestyle variables and MDD; AD=antidepressant.

CpG site	Gene	Chrom	β	P-value	P-corr	CpG site information	Gene information
cg05603985	*SKI*	1	-0.022	3.92×10^{-10}	0.0002	Smoking (63), alcohol consumption (64)	Platelet count (65), Red cell distribution width (65), body height (66), lung function (66), self-reported educational attainment (66)
cg05273171	*PNKD; TMBIM1*	2	-0.013	2.35×10^{-8}	0.017	-	PNKD: red cell distribution width (65), body height (66), monocyte, reticulocyte, leukocute, granulocyte, neutrophil, and eosinophil count (67), waist-hip ratio (66) TMBIM1: red cell distribution width (65), body height (66), inflammatory bowel syndrome (68), Crohn's disease (69), cardiovascular disease (66) Related pathways: innate immune system
cg03864397	*CASP10*	2	-0.030	1.05×10^{-8}	0.007	-	Mean corpuscular haemoglobin (65), chronic lymphocytic leukaemia or rheumatoid arthritis (70)
cg05186879	*MAPKAPK3; CISH*	3	-0.023	3.3×10^{-9}	0.002	-	MAPKAPK3: self-reported educational attainment (71), BMI (72), schizophrenia (73) CISH: eosinophil count (66)
cg16315329	*CISH; MAPKAPK3*	3	-0.039	2.52×10^{-8}	0.018	-	MAPKAPK3: self-reported educational attainment (71), BMI (72), schizophrenia (73) CISH: eosinophil count (66)
cg25753411	*HEBP2*	6	-0.036	2.89×10^{-10}	0.0002	-	Related pathways: innate immune system
cg09511513	*ATP6V1B2*	8	-0.055	6.73×10^{-9}	0.005	-	Depression, mental or behavioural disorder (53)

cg26277237	*KANK1*	9	0.024	2.87×10^{-8}	0.020	-	Uterine fibroid (74), body height (66), systolic blood pressure (75), PHF-tau measurement (76)
cg20494891	*MYO1E*	15	-0.026	2.32×10^{-8}	0.017	-	Serum metabolite measurement (77), pulse pressure (75), coronary artery calcification (78), BMI (66), diffuse plaque measurement (76)
cg27589594	*SLC5A10*	17	-0.025	4.91×10^{-9}	0.004	Gestational age (44)	1.5 anhydroglucitol measurement (79), response to paliperidone (schizophrenia) (80)

987 **Table 3.** CpGs significantly associated with self-reported antidepressant use in GS (N=6,428;
988 antidepressant use=740; N CpGs=10) along with gene annotations, chromosome,
989 standardised effect size, nominal and Bonferroni-corrected p-values. Background
990 information for each CpG and gene was extracted from EWAS
991 (http://www.ewascatalog.org/; association between traits and CpGs on Illumina 450K array
992 at $p\leq1.0\times10^{-4}$); and GWAS (https://www.ebi.ac.uk/gwas/; associations between traits and
993 SNPs at $p\leq1.0\times10^{-5}$) catalogue databases. All associations included in the table from these
994 two catalogues are genome-wide significant.
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010

	MS			MS-ns			MS-control		
Outcome variable (N)	**β**	**p-value**	**R^2**	**β**	**p-value**	**R^2**	**β**	**p-value**	**R^2**
AD use (3,360)	0.377	3.12×10^{-11}	2.12%	0.106	0.075	0.16%	0.254	1.15×10^{-5}	0.93%
AD use* (3,009)	0.213	0.0035	0.56%	0.04	0.565	0.02%	0.064	0.381	0.05%
MDD (3,360)	0.28	5.78×10^{-7}	1.14%	0.158	0.007	0.35%	0.202	0.0004	0.57%
BMI (3336)	0.088	2.54×10^{-7}	0.75%	0.021	0.23	0.01%	0.075	1.33×10^{-5}	0.53%
Alcohol units (3099)	0.141	$< 2 \times 10^{-16}$	1.93%	0.007	0.665	0%	0.127	5.08×10^{-14}	1.56%
Smoking status (3,322)	0.632	$< 2 \times 10^{-16}$	6.11%	0.077	0.028	0.11%	0.567	$< 2 \times 10^{-16}$	5.08%
Pack years (3,310)	0.305	$< 2 \times 10^{-16}$	9.14%	0.061	0.0004	0.34%	0.287	$< 2 \times 10^{-16}$	8.09%

Table 4. Association between MS (methylation score), MS-ns (methylation score trained on non-smokers), and MS-control (methylation score trained on individuals with no MDD diagnosis), and self-reported antidepressant use, MDD, and 4 lifestyle factors (BMI, smoking status, pack years, alcohol units). All regression models include age, sex, and 10 genetic principal components as covariates, except for "AD use*", which also includes BMI, smoking status, pack years, alcohol units, and MDD as covariates. Effect sizes represent standardized betas. R^2 represents the variance explained in the outcome variables by each score. The number of individuals in each model varies based on different available data for each variable.

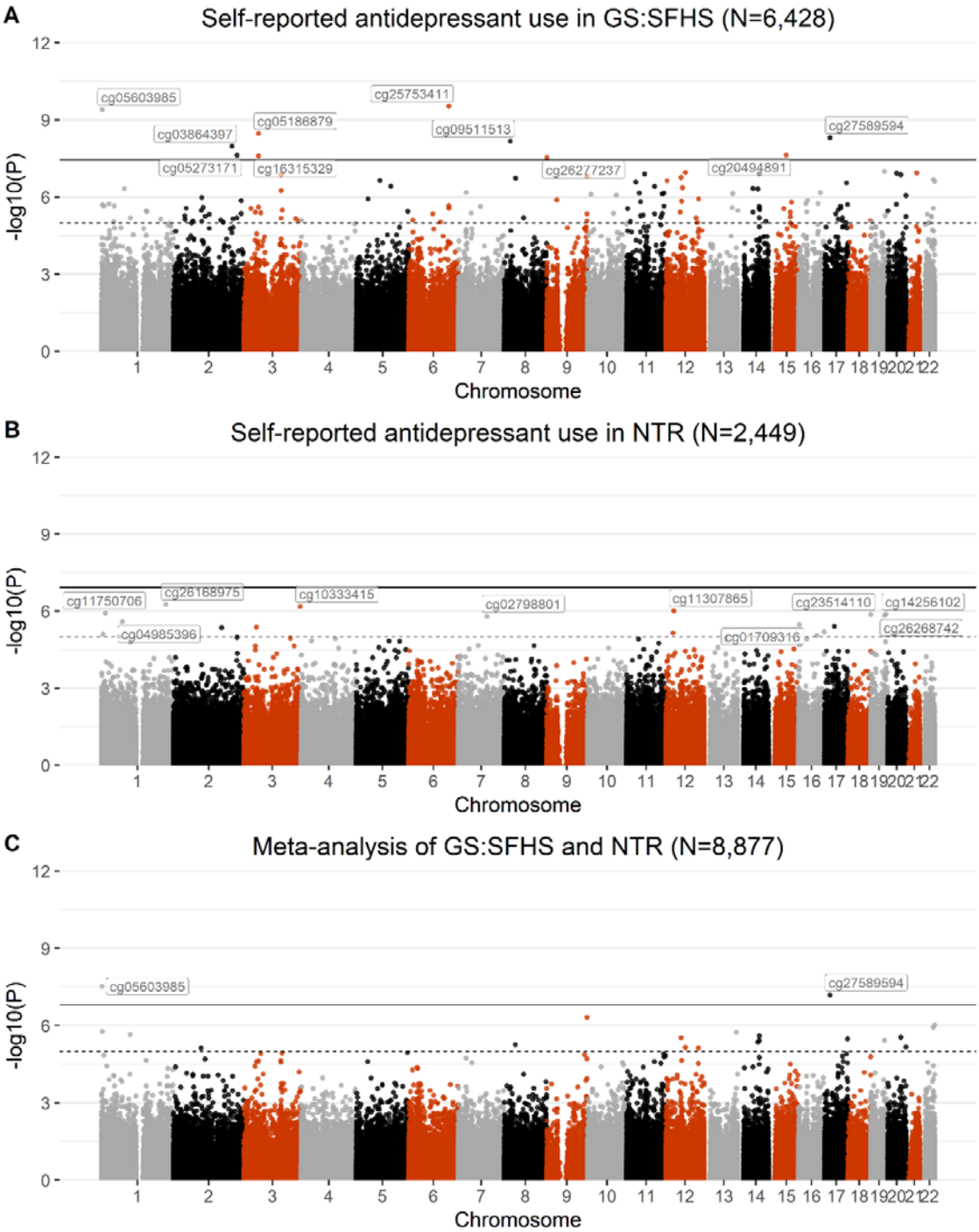
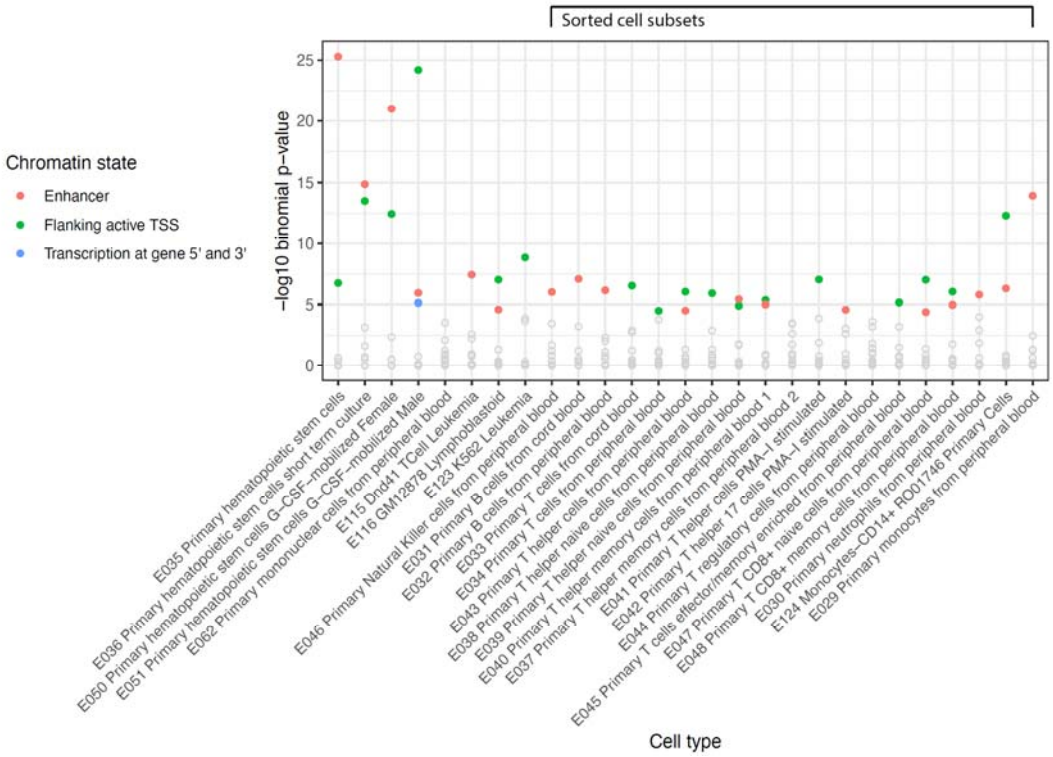


Figure 1A, 1B, 1C. Manhattan plots showing MWAS of self-reported antidepressant use in GS:SFHS (1A), MWAS of self-reported antidepressant use in NTR (1B), and meta-analysis of MWAS in GS:SFHS and NTR (1C). The black line defines methylome-wide significance for each analysis (GS:SFHS: $p \leq 3.6 \times 10^{-8}$; NTR: $p \leq 1.22 \times 10^{-7}$; meta-analysis: $p \leq 1.57 \times 10^{-8}$) and the dotted line defines $p \leq 1 \times 10^{-5}$. Methylome-wide significant hits (GS:SFHS: 10; meta-analysis: 2) and the top 10 CpGs in NTR are labelled on the graph.

1043



1044

1045 **Figure 2.** Differentially methylated CpG sites at $p<1x10^{-5}$ in GS:SFHS enriched for peripheral
1046 blood cell subsets. The x-axis represents peripheral blood cell types, while the y-axis
1047 indicates corrected p-value.