UNIVERSITY OF BATH

Link to publication

University of Bath

Using Openly Accessible Resources to Strengthen Causal Inference in Epigenetic Epidemiology of Neurodevelopment and Mental Health

Esther Walton [1,2,3]*, Caroline L Relton [1,2] and Doretta Caramaschi [1,2]

1   Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK;
    caroline.relton@bristol.ac.uk (C.L.R.); d.caramaschi@bristol.ac.uk (D.C.)
2   Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
3   Department of Psychology, University of Bath, Bath, UK
*   Correspondence: esther.walton@bristol.ac.uk

Abstract:

The recent focus on the role of epigenetic mechanisms in mental health has led to several studies examining the association of epigenetic processes with psychiatric conditions and neurodevelopmental traits. Some studies suggest that epigenetic changes might be causal in the development of the psychiatric condition under investigation. However, other scenarios are possible, e.g. statistical confounding or reverse causation, making it particularly challenging to derive conclusions on causality. In the present review, we examine the evidence from human population studies for a possible role of epigenetic mechanisms in neurodevelopment and mental health and discuss methodological approaches on how to strengthen causal inference including the need for replication, (quasi-)experimental approaches and Mendelian randomization. We signpost openly accessible resources (e.g. MR-Base, EWAS catalog, tissue-specific methylation and gene expression databases) to aid the application of these approaches.

Keywords: DNA methylation, epigenetics, mental health, neurodevelopment, causal inference, Mendelian randomization.

1. Epidemiological evidence linking epigenetics and mental health

Mental health and neurodevelopmental disorders are under the influence of both genetic and environmental factors. Epigenetic mechanisms regulate gene expression and are potential mediators of both these genetic and environmental effects on mental traits and disorders. Of the known epigenetic processes involved in gene regulation, DNA methylation, which consists of the covalent addition of a methyl group to a cytosine base at CpG dinucleotides, is the most widely studied. The main reason for its popularity is the availability of cost-effective, high throughput laboratory assays that utilise DNA extracted using standard protocols. To date, most epigenetic studies of mental health have measured DNA methylation at the genome-wide level using Illumina Infinium 450K or EPIC arrays in peripheral blood or saliva samples, since these tissues are most commonly available in large studies.

Epigenetic studies that have investigated the association of DNA methylation with mental health traits and conditions in peripheral blood or saliva using the Illumina 450K arrays were identified in a semi-systemic manner by searching within PubMed. The characteristics of the studies are summarised in Table 1. While this search is not meant as a systematic review, it provides examples of studies that investigated the link between DNA methylation and brain-related processes in peripheral tissues. Associations of DNA methylation variation measured in peripheral blood, in

relation to schizophrenia are among the most widespread published so far. In the largest study to date, a comparison of 689 men affected by the disease and 645 controls reported over 900 methylation variable sites across the genome. Although the authors applied a more relaxed threshold (FDR p < 0.2) in the discovery sample, many methylation sites replicated in an independent sample with effects consistent in size and direction [1]. Other associations have been reported linking methylation variable loci with suicidal behaviour within individuals with bipolar disorder [2], for depressive symptoms within the elderly [3], self-reported wellbeing [4] and panic disorder in adulthood [5]. However, in some instances conflicting evidence can be found [6] or only very weak evidence is provided, as seen in a study on post-traumatic stress and major depressive disorder [7]. With respect to neurodevelopment, DNA methylation differences were reported in relation to educational attainment and cognitive abilities measured in adulthood [8], attention-deficit hyperactivity disorder [9], oppositional defiant disorder [10], multiple risk behaviours [11], substance abuse [12], early onset conduct disorder [13] and childhood physical aggression [14], with weaker evidence for an association with violent aggression and diagnosed autism spectrum disorders [15-17]. Neurological conditions that showed differences in blood-based DNA methylation when compared to controls include mesial temporal lobe epilepsy [18], narcolepsy [19] and Parkinson's disease [20].

Table 1. Epigenome-wide association studies of mental health traits and diseases conducted in peripheral blood. A semi-systematic PubMed search was undertaken (https://www.ncbi.nlm.nih.gov/pubmed/, access date 21/11/2018) using the terms "DNA methylation", "methylome-wide", "epigenome-wide", "psychiatry", "psychiatric", "behaviour" and "human". See text for references.

Trait/disease	Study design	Tissue	Sample size	DNA methylation differences	Significance threshold	Authors, year
Wellbeing	Population study	Blood	N=2456	2 CpGs	Bonferroni p < 0.05	Baselmans et al., 2015
Schizophrenia	Case-control	Blood	N=1339 (discovery); N=497 (replication)	923 CpGs	FDR p < 0.2	Montano et al., 2016
Substance abuse	Population study	Cord blood	N=244	65 CpGs	FDR q < 0.05	Cecil et al., 2016
Suicidal behaviour	Case-control	White blood cells	N=123	None below threshold	Not specified	Bani-Fatemi et al., 2017
Post-traumatic stress disorder	Clinical study (trauma patients)	Blood	N=473	None below threshold	FDR p < 0.05	Kuan et al., 2017
Major depressive disorder	Case-control	Blood	N=473	None below threshold	FDR p < 0.05	Kuan et al., 2017
Panic disorder	Case-control	Blood	N=96	40 CpGs	FDR p < 0.05	Shimada-Sugimoto et al., 2017
Educational attainment	Population study	Blood	N=10767	9 CpGs	p < 1x10-7	Karlsson Linner et al., 2017
Mesial temporal lobe epilepsy	Case-control	Blood	N=60	216 CpGs	p < 1.03x10-7	Long et al., 2017
Parkinson's disease	Case-control	Peripheral blood mononuclear cells	N=38	2 CpGs (identified via multiple methods)	methylation difference > 15% and validation with other methods	Kaut et al., 2017
Attention-deficit hyperactivity disorder	Population study	Cord blood	N=828	13 CpGs	FDR q < 0.05	Walton et al., 2017
Oppositional defiant disorder	Population study	Cord blood	N=671	30 CpGs	FDR q < 0.05	Barker et al., 2018

Depression	Case-control	Blood	N=200	6 DMRs	Sidak corrected p < 0.05	Crawford et al., 2018
Cognitive abilities	Population study	Blood	N=2557-6809	2 CpGs	p < 0.05/(420000 CpG x 7 traits)	Marioni et al., 2018
Depressive symptoms	Case-control	Blood	N=47	None below threshold		Shimada et al., 2018
Depressive symptoms	Population study	Blood	N=7948 (discovery); N=3308 (replication)	3 CpGs	p < 1.03x10-7	Story Jovanova et al., 2018
Narcolepsy	Case-control	Blood	N=46	14 CpGs	FDR p < 0.05	Shimada et al., 2018
Violent aggression	Clinical study (schizophrenia patients)	Peripheral blood mononuclear cells	N=134 (discovery)	Weak differences	p < 1x10-6	Mitjans et al., 2018
Physical aggression	Population study	Buccal (discovery); Peripheral T cells (replication)	N=119 (discovery); N=38 (replication)	4 CpGs; 2 DMRs	FDR q < 0.05	Cecil et al., 2018
Early-onset conduct disorder	Case-control	Cord blood	N=260	7 CpGs	FDR q < 0.05	Cecil et al., 2018
Multiple risk behaviours	Population study	Blood	N=227-575	2 CpGs	FDR q < 0.10	de Vocht et al., 2018
Autism spectrum disorder	Case-control	Blood	N=1311	None below threshold	p < 1.12x10-7	Andrews et al., 2018
Autism spectrum disorder	Case-control	Cord blood	N=1263	None below threshold	p < 1x10-7	Hannon et al., 2018

2. Challenges to assess causality

Although there are indications that peripheral DNA methylation could be a plausible mechanism that leads to certain brain-related conditions, causality is often difficult to establish in epigenetic epidemiology. Many studies based on epigenome-wide associations are observational and do not allow for a direct assessment of whether the observed DNA methylation differences are a cause, consequence or confounder for the disease of interest.

Firstly, evidence is often based on studies with small sample sizes without replication. Even if the effects are replicated across studies, they might arise due to similar confounding structures in the data sets, such as the distribution of tobacco smoking behaviours. Even after adjusting for self-reported smoking, residual confounding could still be present due to reporting bias. For example, the association study of DNA methylation on educational attainment has revealed that all sites linked with education have previously been associated with smoking behaviour. Since smoking is often negatively correlated with years of education, this suggests that the observed association between DNA methylation and education is largely due to confounding, rather than describing a causal relationship [21].

Another possible scenario where DNA methylation changes are not causal for a disease arises when the disease manifestation itself causes changes in DNA methylation, also referred to as reverse causation. This could arise in cross-sectional studies where the samples for DNA methylation analysis are obtained at the same time point as the administration of a questionnaire to assess the outcome of interest or where the methylation measurement was taken after the diagnosis of a disease was made. For instance, in the large EWAS on major depressive disorder DNA methylation was measured after the diagnosis was made. Hence, based on the association study alone it is impossible to disentangle whether epigenetic changes are cause or consequence of the disease [3].

88   Most human epigenetic studies of mental health are based on peripheral samples. Although in
89 some cases methylation changes occur in CpG sites linked to genes that have relevant brain functions,
90 it is often challenging to relate changes in peripheral methylation to the development of a condition
91 that affects the central nervous system (CNS). This problem is of relevance mainly because DNA
92 methylation in the brain of living individuals cannot be quantified. Post-mortem samples, while rare,
93 only allow the assessment of DNA methylation changes after the disease has manifested [22], as for
94 instance in an EWAS of autism spectrum disorder conducted across several brain regions [23]. In this
95 case, epigenetic changes could be confounded by treatment effects, as DNA methylation changes
96 have been reported for instance in relation to antipsychotic treatment [24].
97   The 'gold standard' experimental approach used to seek causal evidence is the randomised
98 controlled trial (RCT). However, this is not a feasible option for DNA methylation research, as it is
99 not yet possible or ethical to undertake an RCT with DNA methylation as the primary controlled
100 exposure. Some studies have taken advantage of RCTs set up with other primary exposures and
101 subsequently measured DNA methylation as a surrogate or intermediate but these have tended to be
102 serendipitous, relying on RCTs that have collected DNA samples for other purposes (see below for
103 further discussion of this issue).
104   Animal studies, particularly in the laboratory, have the advantage of allowing for controlled
105 experimental conditions and access to specific tissues other than peripheral blood, therefore avoiding
106 the issue of confounding and the otherwise limited inferences that can be made with respect to tissue
107 specificity. In mouse studies, DNA methylation can for example be manipulated by deleting the
108 genes coding for DNA methyltransferases (Dnmt1/Dnmt3a/Dnmt3b), the enzymes that catalyse the
109 transfer of a methyl group to a cytosine nucleotide. A study by Hutnick et al. [25] showed that the
110 deletion of Dnmt1, even when restricted to the forebrain, caused widespread hypomethylation,
111 neuronal degeneration and behavioural impairment in learning and memory. This is in line with
112 other mouse studies, where Dnmt1 deletion seemed to cause increases in anxiety-like behaviour and
113 deleting both Dnmt1 and Dnmt3 led to synaptic abnormalities with functional consequences for
114 hippocampal plasticity [26,27]. These studies indicate a causal link between overall DNA methylation
115 and brain-related traits, however they do not allow for the identification of specific methylation loci
116 within the genome at which the changes in DNA methylation might be exerting their influence.
117 Recently, with the technology of the CRISPR-Cas9 system applied *in vivo* to laboratory mice it has
118 become possible to demonstrate that DNA methylation at the *FMR1* gene causes the molecular and
119 physiological phenotype of fragile-X syndrome [28]. While fragile-X syndrome has a specific and
120 detectable molecular phenotype (lack of FMR1 protein), the limitation of most animal studies is that
121 many human psychiatric diseases are defined by behavioural traits that can only partially observed
122 in other species. Most animal models are based on resemblance of the behavioural symptoms and
123 therefore mostly correspond to a sub-set of symptoms and traits of the modelled human psychiatric
124 diseases rather than the full disease. Similarly, the pathological mechanisms leading to the human
125 psychiatric conditions might not necessarily correspond to the changes observed in the animal
126 models that only partially mimic the human condition.

127 **3. Epidemiological approaches to investigate causality**

128 *3.1. Strength and robustness of the associations*

129   True epigenetic associations often tend to replicate in population samples with similar
130 characteristics and confounding structures, thus the associations observed could be due to real effects
131 or to other non-causal explanations. To assess the strength and robustness of the associations it is
132 recommended, where feasible, to work collaboratively across multiple studies, as true causal
133 associations ought to be reproduced across studies with different confounding structures. Such
134 collaborations can be achieved within consortia where several studies with available epigenomic data
135 can contribute to addressing the same research questions according to agreed and standardized
136 analysis plans. Selected examples of such consortia that have been used in the field of epigenetic
137 epidemiology are listed in Table 2 below.

Table 2. Selection of consortia in the field of epigenetic epidemiology.

Resource	Description	Link
Pregnancy and Childhood Epigenetics (PACE) consortium [29]	focus on the effect of early life exposures on DNA methylation in childhood	https://www.niehs.nih.gov/research/atniehs/labs/epi/pi/genetics/pace/
Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) [30]	focus on facilitating genetic and epigenetics meta-analyses and replication opportunities among cohort studies	http://www.chargeconsortium.com/
Genetics of DNA Methylation (GoDMC) consortium [31]	focus on the genetic basis of DNA methylation variation in participants of different ages and ethnicities	http://www.godmc.org.uk/information.html

For these cross-cohort analyses, it is however essential to standardize pre-processing steps, including normalisation, quality checks, and epigenome-wide association study (EWAS) analyses procedures. Data sharing is often a limiting factor in analyses of this type and harmonizing data across studies can sometimes be resource intensive. Software packages have been developed to facilitate such analyses. For example, the *meffil* R package, which was created to enable cross-cohort harmonization without data sharing, is available for download at https://github.com/perishky/meffil [32].

Where there is no opportunity for collaboration or the phenotypes of interest are not available in consortia, it is sometimes possible to access DNA methylation data and their association with the phenotype from openly available online repositories, such as Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). In the GEO repository, data can be downloaded or analysed online with the interactive GEO2R tool [33].

Replicating associations across different datasets also provides an opportunity to verify that results are not due to technical artefacts. Although replication does not necessarily increase the likelihood of associations being causal, it can be a further step in supporting the veracity of the observed association. For instance, investigating the same CpG sites-trait associations across the Illumina 450K or the more recent EPIC array or using different techniques, including pyrosequencing, bisulphite sequencing and qPCR will strengthen the inferences that can be made with respect to the confidence in true associations.

3.2. Experimental and quasi-experimental approaches

The conventional epidemiological design to investigate causality, an RCT requires the participants to be randomly assigned to groups that are similar except for the exposure of interest (here DNA methylation). Although theoretically it is possible to conduct an RCT of a demethylating agent and assess its impact on a mental health outcome, a targeted manipulation of specific methylation sites is currently not achievable with the available tools.

RCTs are however more tractable where methylation is considered as a secondary outcome to investigate the effects of an intervention. For example, an RCT design has been exploited to assess the effects of pollution [34] and folate intake [35] on DNA methylation. Linking changes in methylation, which have been identified to be a causal consequence of environmental exposures, to psychiatric disorders could be an interesting and worthwhile extension of such findings.

Natural experiments, where populations are exposed to an unplanned disaster or event, provide valuable data to reveal changes in DNA methylation that are causal for psychiatric conditions. For example, methylation changes due to prenatal exposure to the Dutch famine [36] have been shown to cause changes in mental health in adulthood [37] and suggest that DNA methylation could be a potential mediating mechanism. Similarly, prenatal maternal stress due to a significant ice storm in Quebec in 1998 affected DNA methylation [38] and autism-related traits [39].

3.3. Mendelian randomization (MR)

One widely adopted approach to strengthen causal inference is the method of Mendelian randomization (MR), a form of instrumental variable analysis. In MR, the instrument is comprised of one or more genetic variants that are robustly associated with the exposure of interest. As individuals

179 inherit alleles at random, these individuals are assigned to experience a higher-than-average dosage
180 of the exposure.
181   MR relies on the availability of genetic variants to use as instrumental variables (for a discussion
182 on additional assumptions, see [40,41]). Where genetic variants can be identified that correlate
183 strongly with DNA methylation levels, MR can be applied to study causal effects of DNA methylation
184 on mental health. Depending on the research question, the sample characteristics and data
185 availability, different MR methodologies can be applied such as one-sample, two-sample,
186 bidirectional, multivariable and two-step MR, the details of which can be found elsewhere [42,43].
187 Due to limitations in data availability and computational resources required, MR has predominantly
188 been performed to date on selected methylation loci (e.g. top hits of a robust EWAS), with a few
189 notable exceptions [44,45]. However, with the advent of more detailed data on genetic variants that
190 tag methylation variation, the approach promises to be more widely adopted.

191 3.3.1. Instruments for epigenetic MR analysis

192   Potential instruments for DNA methylation are single nucleotide polymorphisms (SNPs) that
193 are strongly associated with methylation at the CpG sites of interest – often referred to as methylation
194 quantitative trait loci (mQTL). These can be found in online databases that have performed GWAS
195 of DNA methylation (Table 3). The overwhelming majority of catalogued mQTLs have been derived
196 from populations of European ancestry and are based on peripheral blood DNA, raising the issue of
197 whether the same SNP-DNA methylation relationship is observed in other ethnicities or tissues.
198 Emerging evidence suggests that this assumption might be plausible in some instances [46].
199 However, as DNA methylation is often tissue-specific, brain tissue specific databases (Table 3) can be
200 used to identify mQTLs when the hypothesis implies a biological mechanism that acts via changes in
201 brain DNA methylation.
202   Alternatively, blood-derived mQTLs can be used in MR when an EWAS of a brain-related trait
203 has been conducted in blood and it is plausible that changes in methylation in blood cells are reflected
204 in changes in brain activity, for instance via circulating hormones that cross the blood-brain barrier
205 (see section 3.4.1 for a more detailed discussion).

206   **Table 3.** Resources that can be used to identify genetic effects on DNA methylation probes.

Resource	Description	Link
mQTL database [47]	1000 mother-child pairs across the life course; based on blood	http://mqtldb.org
BIOS QTL browser	3841 adult blood samples of varying ages	https://genenetwork.nl/biosqtlbrowser
GoDMC [48]	largest mQTL consortium to date; focus on blood tissue	http://www.godmc.org.uk/projects.html
Brain xQTL Serve [48]	411 frontal cortex brain samples of older adults	http://mostafavilab.stat.ubc.ca/xqtl
Brain Epigenomics [49]	166 foetal brain samples	https://epigenetics.essex.ac.uk/mQTL

207
208
209   Some of the resources listed in Table 3 are based on data from specific developmental periods
210 (e.g. foetal sample, cord blood) – however, our ability to use these resources in a developmentally
211 sensitive manner is still restricted and heterogeneity in ethnicity and cell type composition between
212 the target and the reference datasets limits any conclusions drawn from these analyses.
213   Most mQTLs are *cis*-associations, i.e. they are located proximal to the CpG of interest. *Cis*-SNPs
214 have large effects on the CpGs in their proximity, whereas *trans*-SNPs have smaller effects and tend
215 to act polygenically on several target loci. For these reasons *cis*-SNPs, rather than *trans*, are preferred
216 as instruments for use in MR.

217 3.3.2. Methodologies in epigenetic MR analyses

218   If mQTLs are available for the CpGs of interest, these can be used as instruments for MR. In
219 studies where genotypes, DNA methylation data and the outcome (e.g., mental health trait), are
220 available, it is possible to perform one-sample MR using the 2-stage-least-square regression (Figure
221 1, top panel). This is easily implemented with the *ivreg2* command in the STATA software or the
222 function *tsls* in the *gmm* R package (https://cran.r-project.org/web/packages/gmm/index.html) [50].
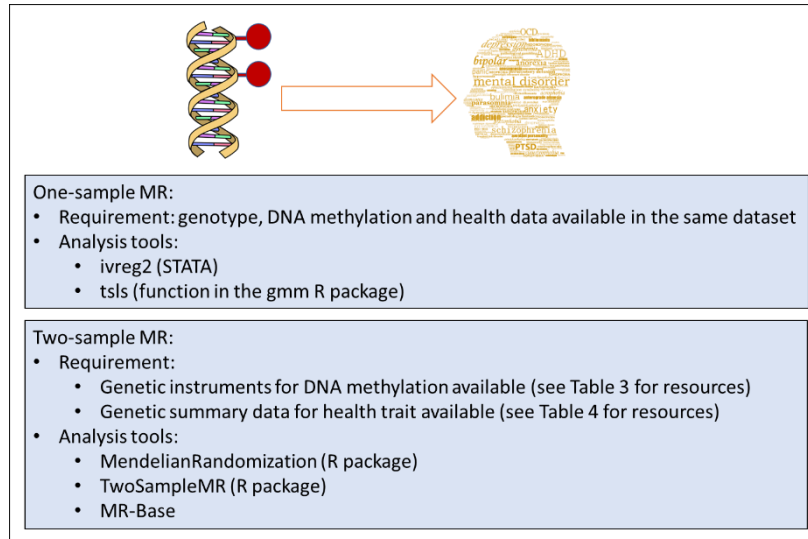
Figure 1. Overview on methodologies in epigenetic MR analyses.

225   When this data is not available, a two-sample MR approach can be used (Figure 1, bottom panel).
226 This relies on extracting the genotype-methylation (G-M) summary statistics (beta regression
227 coefficients and standard errors) from one study and the genotype-outcome (G-O) statistics from
228 another, independent study. For one SNP, the causal estimate is the ratio of the genotype-outcome
229 beta coefficient divided by the genotype-methylation beta coefficient. The standard error of the causal
230 estimate is estimated via the delta-method as described in Thomas et al. [51]. When at least three
231 genetic variants are available, the G-M/G-O ratio estimates are meta-analysed using standard meta-
232 analysis methods such as the inverse variance weighted approach with fixed or random effects
233 models. Two-sample MR can be easily performed using the MR-Base online tool
234 (http://www.mrbase.org/) and the *TwoSampleMR* R package available for download at the github
235 online repository (https://github.com/MRCIEU/TwoSampleMR) [52] . Similarly, the
236 *MendelianRandomization* R package performs two-sample MR using existing summary data on genetic
237 associations with exposure and outcome [53]. When several SNPs are available it is useful to choose
238 the MR-Egger model, which provides a test for horizontal pleiotropy and a pleiotropy-adjusted
239 causal estimate [54]. However, this method has lower power and is recommended primarily as a
240 sensitivity analysis. GWAS summary statistics for the G-O associations can be found in several online
241 databases (Table 4).

242   **Table 4.** Resources providing GWAS summary statistics for (mental health) traits.

Resource	Description	Link
MRInstruments	R package that contains a number of data files from various sources to provide instruments in two-sample MR	https://github.com/MRCIEU/MRInstruments
Phenoscanner [55]	lists over 65 billion GWAS associations, hosted at the University of Cambridge	http://www.phenoscanner.medschl.cam.ac.uk
GWAS catalogue [56]	curated catalog in collaborative between the EMBL-EBI and NHGRI	https://www.ebi.ac.uk/gwas
Psychiatric Genomics Consortium [57]	genome-wide summary data for psychiatric disorders	https://www.med.unc.edu/pgc/results-and-downloads

ENIGMA brain structure [58]	genome-wide summary data for brain structure phenotypes	http://enigma.ini.usc.edu/research/ download-enigma-gwas-results

243

244   Following this strategy, two-sample MR has recently been applied to test for a causal effect of
245 methylation in the *DRD4* gene on physical aggression and did not support a causative link [14].

246   The direction of the association, if not known a priori, can be queried using bi-directional MR,
247 where both a causal effect of methylation on the trait and a causal effect of the trait on methylation
248 are estimated. Effectively, this procedure involves two MR analyses, requires a set of independent
249 SNPs for each analysis and can be carried out within the one-sample or the two-sample setting.

250   When the research interest is to estimate the effect of an exposure on an outcome via DNA
251 methylation, to supplement the conventional observational mediation approach, it is useful to adopt
252 an MR strategy that involves two MRs, one from exposure to methylation and one from methylation
253 to the outcome of interest. In the two-step MR approach, the SNPs used as instruments for each step
254 need to be independent. Each MR step adopts the usual assumptions for MR and is performed using
255 the same general principles and methods for MR. This implies that several independent study
256 samples are needed to obtain the summary statistics for the genotype-exposure (G-E), G-M and G-O
257 associations, which can be identified using the resources listed in Table 3 and 4. Two-step MR has
258 been applied to test the causal role of prenatal nutrients involved in the one-carbon metabolism on
259 schizophrenia via epigenetic changes [59] and to reveal DNA methylation as a mediator between the
260 exposure to prenatal vitamin B_{12} and cognitive abilities [60].

261   Other methods using genetic variants to strengthen causal inference are based on the integration
262 of genome-wide genetic and epigenetic data with the disease of interest using polygenic risk scores
263 (PRS) for the disease and co-localisation analyses. PRS are defined as the sum of trait-associated
264 alleles across many genetic loci, weighted by the GWAS effect size. Similar to the MR approach, the
265 epigenetic and phenotypic variation associated with PRS is less likely to be confounded by lifestyle
266 exposures such as smoking and environmental factors such as pollution and is less prone to reverse
267 causation. For example, EWAS studies on schizophrenia where PRS rather than diagnosis were used
268 in the analysis have identified DNA methylation differences at novel CpGs [61]. Furthermore,
269 Bayesian co-localisation analysis, where the results of a GWAS of methylation at the CpG sites and
270 the results from an independent GWAS for schizophrenia were compared, supported the hypothesis
271 that some of the genetic variants within the overlapping sites had a regulatory role in the disease via
272 influencing DNA methylation [62]. PRS for brain-related disease can be computed using summary
273 statistics from published GWAS (see Table 4 for a list of resources; to derive polygenic scores, see
274 **https://choishingwan.github.io/PRSice**, **https://www.cog-genomics.org/plink/1.9/score** and [63]).
275 Bayesian colocalization analysis can be performed using existing summary data from mQTL
276 databases and the *coloc* R package (https://cran.r-project.org/web/packages/coloc/) [64].

277 *3.4. Plausibility of biological mechanisms*

278 3.4.1. A word of caution: mechanism vs biomarker

279   The excitement of obtaining an epigenetic signal, that is strong, robust and potentially causal,
280 can be exhilarating. However, before deriving at conclusions about the "aetiological mechanism of
281 disease", it is advisable to recall the original aim of the study. Frequently, the aim is to identify causes
282 of disease, which is imperative for interventions to be successful. On the other hand, establishing
283 non-causal associations (often referred to as biomarkers, see below) can be useful in prediction.
284 However, a biomarker can be causal or non-causal. Whether the aim is to identify a causal pathway
285 and/or a biomarker (of risk or of disease) should be set out in the initial stages of the project. Caution
286 is advised with respect to the conclusions that can be drawn from the study design and data in terms
287 of biological mechanisms. The interpretation of results will differ, depending on the underlying
288 assumptions about the likelihood of system-wide effects of the exposure (i.e. genetic or
289 environmental causes of disease), the relationship between the studied tissue and the primary tissue

290 of pathophysiology. In most cases, methylation profiles would have been obtained from peripheral
291 tissues (blood or saliva), with a small proportion of studies using post-mortem brain tissue.
292   Under the assumption that the causal (but not necessarily initial, see argument below) tissue of
293 pathophysiology is the brain, at least three potential scenarios are possible describing the relationship
294 between peripheral and CNS methylation profiles: a shared common cause, periphery-mediated or
295 CNS-mediated pathways to disease (left, middle and right panel in Figure 2). Note that a scenario, in
296 which DNA methylation is a direct consequence – rather than a precursor - of disease, is an equally
297 likely possibility, but not the focus of the current discussion. A mechanistic interpretation of findings
298 based on peripheral tissue only makes sense assuming that the initial cause of pathophysiology
299 originates in the periphery (scenario b and e in Figure 2) or at the very least assuming concordance
300 of methylation patterns across tissues (top panel Figure 2, although see below for additional
301 assumptions).
302   'Concordance' in this case shall be defined as the consistency in effect of the exposure (i.e. cause
303 of disease) on DNA methylation across tissue. This is different from 'correlation' of DNA methylation
304 across tissue. For example, *relative* (but meaningful) perturbations in DNA methylation due to an
305 exposure might be comparable across tissue, while *absolute* DNA methylation levels themselves are
306 less correlated across tissues (Figure 3a). This assumes that small levels of perturbations can have
307 large effects in some but not in other tissues. Likewise, without knowing what precisely causes cross-
308 tissue correlations in DNA methylation, DNA methylation levels might be correlated across tissue,
309 but the effect of an exposure on DNA methylation in each tissue is different (Figure 3b). Therefore,
310 while correlation of DNA methylation profiles across tissues is often an important indication, it is
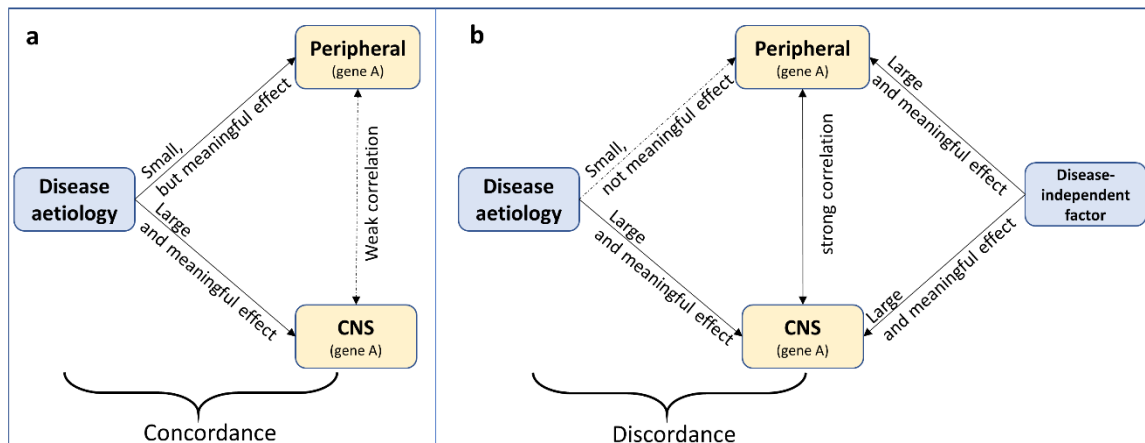311 neither necessary nor sufficient for cross-tissue concordant effects.
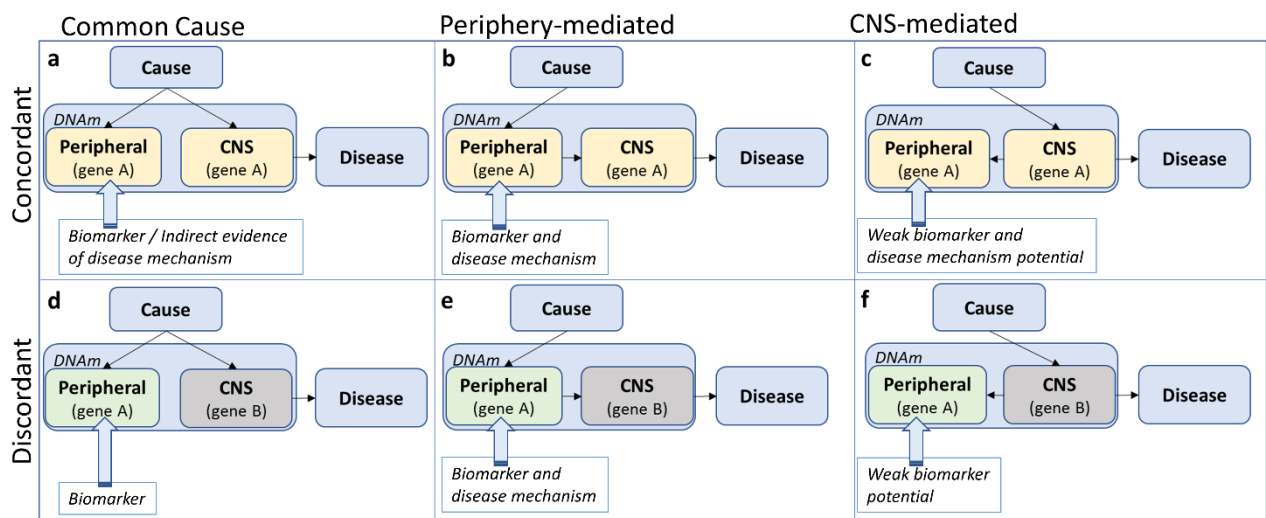
312
313   **Figure 2.** Three potential scenarios (by column) describing the relationship between peripheral and
314   central nervous system (CNS) DNA methylation (DNAm) profiles within the pathway from cause to
315   disease, assuming either consistency in effect of the risk exposure on DNA methylation across tissue
316   (i.e. concordance; top panel) or discordance (bottom panel).

317   All too often, cross-tissue concordance and correlation are implicitly assumed and findings are
318 interpreted as potentially mechanistic. However, there is evidence that cross-tissue *correlation* seems
319 to be the exception, rather than the norm [65]. *Concordance* of methylation profiles across tissues is
320 hardly ever investigated, due to the difficulty (and costs) in measuring the effect of a risk factor on
321 DNA methylation across several tissues in the same individuals. The notable exception of this is the
322 investigation of tissue-specific mQTLs. For online available resources to investigate cross-tissue
323 concordance and correlation, see sections 3.4.2 and 3.4.3.
324

Figure 3. Two scenarios demonstrating that correlation of DNA methylation profiles across tissues is neither (a) necessary nor (b) sufficient for cross-tissue concordant effects.

Even in the case of cross-tissue concordance, it is easy to overstate risk pathways to disease. In the concordant, common cause scenario (a in Figure 2), the tendency is to assume system-wide causal effects, but it might be equally likely that a disease risk factor impacts methylation of the same gene in different tissues independently. In all concordant scenarios (a-c in Figure 2), concordant gene function across tissues is presumed, although genes can have different functions in different tissues. For example, assuming that in an analysis based on data from whole blood a methylation site was identified with a potential relevance for serotonin function. In the periphery the primary function of serotonin is digestion, while in the CNS serotonin is mainly involved in sleep and mood [66]. In the "shared common cause" scenario (Figure 2, scenario a), we do not need to focus on digestion-related functions, as these are not likely involved in the disease pathophysiology. In the "periphery-mediated" scenario (Figure 2, scenario b), however, digestion should be a main pathway-of-risk while in the "CNS-mediated" scenario (Figure 2, scenario c), digestion is – if anything - a downstream pathway of disease. Any mechanistic interpretation of findings depends fundamentally on which scenario is most likely.

When concordance is not assumed (see discordant scenarios d-f in Figure 2), the default position is often that - even though the epigenetic variation is not likely to be mechanistically involved - it may act as a biomarker of disease risk. However, the precise "biomarker" definition referred to is often not clear. According to the National Institute of Health Biomarkers Definition Working group, a biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathologic processes or biological responses to a therapeutic intervention" [67]. While it is beyond the scope of this review to discuss the role of DNA methylation as a biomarker of risk or disease, this term should not be used too lightly. Biomarkers should be: easily (in terms of tissue accessibility) and robustly measurable with little measurement error, reproducible across studies (e.g. it is not advised to claim biomarker potential based on a single study without replication) and have predictive power (or alternative advantages such as reducing costs). Finally, it should be clear, what exactly the established biomarker indexes (risk, disease or treatment). While it is often claimed that methylation-based biomarkers have the potential to inform intervention strategies, studies designed to explicitly demonstrate this are rarely seen [68].

It is impossible to test these scenarios (Figure 2) directly without access to longitudinal and repeated measures of both peripheral and brain tissue in living humans, but their likelihood can be assessed by using tissue-specific causal inference method such as Mendelian randomization (see section 3.3) and the increasing body of online resources as described in the following sections.

3.4.2. Biological characterisation

Characterizing the biological relevance of an identified methylation site is often part of an epigenome-wide analysis, regardless of whether a potential disease mechanism has been established.

363 While methylation sites are often primarily viewed in relation to the nearest coding gene, it can be
364 equally important to consider DNA methylation in the context of regulation of gene expression via
365 impacting chromatin accessibility and transcription factor binding. For instance, studies have
366 confirmed that DNA methylation around the transcription start site is largely associated with
367 reduced gene expression locally [48]. In a study based on brain samples, DNA methylation and
368 histone modifications were located in regulatory regions and seemed to mediate the association of
369 genetic variants with gene expression [69]. Many of those epigenomic loci were also replicated in
370 peripheral blood samples and were associated with psychiatric diseases such as schizophrenia and
371 bipolar disorder. To characterize the biological context of a methylation site, the results of an EWAS
372 can first be matched to the annotation file usually provided with the data or openly accessible online
373 (Illumina 450k and EPIC array annotation are for example available via various R packages such as
374 *meffil* [32]). This will provide CpG information on genomic location, SNPs located in or close to the
375 probe, associated genes, and location with respect to the transcription start site of these genes or CpG
376 islands. Furthermore, information is provided on low- or high-CpG density regions associated with
377 FANTOM 4 promoters [70], although the reader should keep in mind that this information was based
378 on human myeloid leukaemia cell lines and is not specific to CNS tissue. Finally, in the annotation
379 file the reader will find information on enhancer elements, DNase I Hypersensitivity Sites, open
380 chromatin regions and transcription factor binding sites (all based on the Encyclopaedia of DNA
381 Elements (ENCODE) data [71]).

382 Whenever possible, however, querying several databases (see Table 5 for selected resources) is
383 advocated to corroborate results and to summarize all findings to avoid selective reporting. Also, to
384 achieve a more meaningful interpretation of the regulatory nature of the genomic region in question,
385 investigating these regulatory characteristics in a cell-type specific manner is advisable, which can be
386 achieved using ENCODE data (**www.encodeproject.org**), usually via platforms such as
387 **genome.ucsc.edu**. For example, DNase I hypersensitivity clusters – indicative of regulatory
388 chromatin regions that are sensitive to cutting by the enzyme DNase - can be viewed for 125 cell types
389 (including cells derived from blood and brain tissue) as part of the ENCODE project. Histone marks
390 and transcription levels are available for up to nine cell lines (including blood, embryonic stem cells,
391 skeletal muscle among others). Transcription factor bindings sites are listed for 161 factors in 91 cell
392 types (for a list on cell types, see here: **https://genome.ucsc.edu/cgi-**
393 **bin/hgEncodeVocab?type=%22cell%22**). Note that information on CNS-specific cell types is not
394 always available but high (or low) correspondence across these diverse cell types could indicate
395 similarly (un-)correlated profiles in brain tissue. For cell-type specific profiles related to brain tissue,
396 a suggestion could be to investigate DNase I and histone mark data from the Roadmap Epigenetics
397 Project (**http://www.roadmapepigenomics.org/data/**) that assayed ten different brain regions
398 (including the hippocampus, cerebellum and mid frontal lobe among others). Note though that
399 DNase I data is only available for foetal brain (not region-specific) and spinal cord tissue. Also note
400 that to view Roadmap data in the UCSC genome browser, the reader will need to import these tracks
401 via the UCSC Track data hub (https://genome.ucsc.edu/cgi-bin/hgHubConnect) or via
402 http://www.roadmapepigenomics.org/data/. PsychENCODE is a comprehensive resource with
403 exceptional relevance to brain related traits [72-82]. It provides raw and derived transcriptomic,
404 epigenomic, and genomic data of post-mortem adult and developing human brains, both at the
405 single-cell and tissue level. This dataset also includes measures on (hydroxy-)methylation, is based
406 on up to 2000 individuals and incorporates resources such as GTEx, ENCODE and Roadmap
407 Epigenetics Project, discussed above and elsewhere in this article. Data and results can be
408 downloaded from The PsychENCODE knowledge portal (http://www.synapse.org/pec) and from
409 http://resource.psychencode.org/.

410   **Table 5.** Selection of resources to aid in the biological characterisation of DNA methylation findings.

Resource	Description	Link
ENCODE data [71]	tissue-specific regulatory elements across a wide range of tissues	www.encodeproject.org; or via www.ucsc.genome.edu

Roadmap Epigenetics Project [83]	tissue-specific regulatory elements specifically in brain tissue	http://www.roadmapepigenomics.org ; or via www.ucsc.genome.edu
PsychENCODE [74]	Brain-specific tissue and single-cell transcriptomic and epigenomic data	http://www.psychencode.org/
EWAS catalog	manually curated and quality controlled catalog of epigenome-wide association studies	ewascatalog.org
Imprinted genes	List of imprinted genes (by species)	http://www.geneimprint.com/site/genes-by-species

411

412   After investigating the regulatory nature of the genomic region, it can also be helpful to query
413 whether the CpG itself or the differentially methylated region (DMR) has been implicated in other
414 epigenome-wide analyses, which can be done using a manually curated EWAS catalog hosted at
415 **http://www.ewascatalog.org/**.
416   Finally, it is advised to investigate: 1) whether a CpG-of-interest is under genetic control by
417 identifying potential mQTLs, ideally in a tissue-specific manner (see section 3.3.1 and Table 3 above
418 for a list of resources); 2) whether a genomic region might show epigenetic supersimilarity, i.e. where
419 the similarity in DNA methylation between twins is greater than expected based on shared genetics,
420 as reported by Van Baak et al. [84]; and 3) whether a CpG-linked gene might be imprinted, meaning
421 that the expression of this gene depends on the parental origin. For a list of imprinted genes, see
422 **http://www.geneimprint.com/site/genes-by-species**.

423 3.4.3. Cross-tissue comparisons

424   Cross-tissue correlation (see section 3.4.1) is an important, but not essential requirement, even
425 for a mechanistic interpretation of findings (e.g. scenario e in Figure 2). In practice, correspondence
426 can be investigated using cell-type specific data on regulatory regions (see section 3.4.2 and Table 5)
427 and several other openly accessible online resources (Table 6). BECon [85]
428 (**https://redgar598.shinyapps.io/BECon/**) is based on paired blood and post-mortem brain tissue data
429 from 16 individuals. The user can enter a CpG or gene name to visualize cross-tissue correlation
430 across blood and three brain regions (BA10 (frontal), BA20 (temporal) and BA7 (parietal)). Another
431 online resource with similar functionality is available via
432 **https://epigenetics.essex.ac.uk/bloodbrain/**, based on matched blood and four post-mortem brain
433 tissues (cerebellum, entorhinal cortex, frontal cortex and superior temporal gyrus) in 74 individuals.
434 These two resources are based on the Illumina 450k array. Methylation data based on bisulphite
435 sequencing are available via MethBase [86] (**http://smithlabresearch.org/software/methbase/**) and
436 can be imported via the Track hub option (see section 3.4.2) into the UCSC genome browser. This
437 resource provides information on methylation levels at individual sites, allele-specific methylation,
438 and hypo- or hypermethylated regions. Furthermore, MethBase does not only allow for comparisons
439 across cell types (frontal cortex, neural progenitor cells, embryonic stem cells and blood tissue cells
440 in humans), but also across development (from 35 days to 64 years in the case of brain tissue data)
441 and across species (including human, mouse, chimp, dog, zebrafish and plants).
442   Alternatively, it is possible to test for a tissue-specific enrichment of EWAS probe sets, an option
443 which is currently implemented in eFORGE (**http://eforge.cs.ucl.ac.uk/**). Relying on data from
444 ENCODE and the Epigenomics Roadmap, eFORGE compares DNase I hypersensitivity site hotspot
445 overlap between an EWAS input list and background probes in a cell-type specific manner.

446   **Table 6.** Resources for cross-tissue comparisons of methylation signals. .

Resource	Description	Link
BECon [85]	cross-tissue correlations of 450k probes across paired blood and brain regions of 16 individuals	https://redgar598.shinyapps.io/BECon

Brain Epigenomics	cross-tissue correlations of 450k probes across paired blood and brain regions of 74 individuals	https://epigenetics.essex.ac.uk/bloodbrain
MethBase [86]	methylation profiles across tissues, development and species, based on bisulphite sequencing	http://smithlabresearch.org/software/methbase
eFORGE [87]	analysis of cell type-specific signals in epigenomic data	http://eforge.cs.ucl.ac.uk/

An alternative technique to investigate cross-tissue correspondence was applied in Linnér et al. [88] using data from the Epigenomic Roadmap Consortium (see section 3.4.2; although alternative resources such as PsychENCODE listed in Table 5 could also be used). There, the authors calculated average cross-tissue methylation for a selected number of CpG sites linked to educational attainment and derived deviation from this average for a range of tissues (including brain tissue). These tissue-specific measures of deviation were then correlated with EWAS test statistics (z-scores). The authors argued that a lack of correlation between EWAS z-scores of educational attainment and tissue-specific derivation (especially in brain tissue, assumed to be the target tissue of interest) indicated an absence of brain-tissue specific effects and might be suggestive of confounding. Of note, this method is based on average methylation levels across tissue and not on correlations (i.e. methylation profiles might be correlated across tissues, but at different absolute methylation levels).

Finally, there is some evidence that the effects of mQTLs on methylation can be stable across tissues [47], although large-scale investigations across a wide range of tissue types (including brain tissue) are still missing. With this in mind, investigating consistency of mQTL effects across tissues (using resources described in section 3.3.1) can be helpful to obtain some indirect evidence for or against cross-tissue concordance.

3.4.4. Tissue-specific gene and protein expression

It is generally assumed that DNA methylation influences gene expression. However, this issue is still extensively debated [89] and the absence of a functional effect of methylation of gene expression does not preclude the possibility of a meaningful, causal mechanism. Still, it can be highly informative to investigate whether a gene linked to variation in DNA methylation at a site-of-interest also shows variation in its level of expression in the tissue-of-interest. The following section and Table 7 provide an overview of online resource to assess gene expression profile by tissue and across development.

The Human Protein Atlas (https://www.proteinatlas.org/humanproteome) is an excellent resource to investigate in which tissues a gene-of-interest is expressed in absolute terms, and also whether the expression of such a gene is elevated in the target tissue relative to average expression levels in all tissues. Lists on whole groups of genes that are preferentially expressed in certain tissues (e.g. n=1460 genes are listed to show elevated expression profiles in brain tissue relative to all other tissues) can be used to test for enrichment of brain-expressed genes in EWAS results.

The Genotype-Tissue Expression project (GTEx, https://gtexportal.org/home/) provides similar options, listing information on tissue-specific gene expression, regulation and eQTL information. Importantly, the eQTL function allows users to investigate tissue-specific eQTL effects (for example of SNPs that have already been identified to be mQTLs).

To gain insight into gene expression profiles across development, the reader is encouraged to consult the EMBL-EBI expression atlas (https://www.ebi.ac.uk/gxa/home), which displays data from a range of resources (including NIH Epigenomics Roadmap, ENCODE and GTEx).

Three resources are of particular relevance to brain tissue-specific gene expression: the Allen Brain Map and Brain Cloud. The Allen Brain Map portal (**http://portal.brain-map.org/**) provides a range of useful data, including the Human Brain Atlas and the Developing Human Brain resources. The former is a unique multimodal atlas of the human brain, integrating highly detailed anatomic and genomic information. The user can search for a gene-of-interest and visualize its expression profile in different brain regions using high-resolution, MRI-based 3-D histology scans.

The BrainSpan Atlas of the Developing Human Brain (**http://www.brainspan.org**) provides information on the human transcriptome (RNA sequencing and exon microarray data) across different brain regions and development. The BrainCloud application informs on genome-wide gene expression and their genetic control in the dorsolateral prefrontal cortex of normal subjects across the lifespan (**http://braincloud.jhmi.edu**).

The PsychENCODE project combines data from several resources (including GTEx and BrainSpan) to characterize a large spectrum of genomic elements with the human brain, including gene expression as well as multi-QTL maps (for expression, chromatin, transcript expression and cell fraction), enhancers, splice variants and co-expression modules, often specific to cell type, brain region or developmental period. For a more detailed discussion on brain-based resources, see Keil et al [90].

Table 7. Resources to investigate tissue-specific gene expression.

Resource	Description	Link
Human Protein Atlas	expression profiles for all protein-coding genes in 44 tissues and organs in the human body	https://www.proteinatlas.org/humanproteome
Genotype-Tissue Expression project (GTEx)	information on tissue-specific gene expression, regulation and eQTL information based on 53 non-diseased tissues across 714 individuals	https://gtexportal.org/home/
BIOS QTL Browser [48]	methylation QTL data based on up to 3,841 whole-blood samples	https://genenetwork.nl/biosqtlbrowser/
EMBL-EPI expression atlas	gene expression profiles across development based on a range of resources	https://www.ebi.ac.uk/gxa/home
Human Brain Atlas [91]	multimodal atlas of the human brain, integrating highly detailed anatomic and genomic information based on six adult brains	http://human.brain-map.org/
Developing Human [92] Brain	human transcriptome in up to 16 brain regions from 4 weeks post conception to over 40 years	http://www.brainspan.org/
BrainCloud [93]	gene expression and their genetic control in the dorsolateral prefrontal cortex of normal subjects across the lifespan	http://braincloud.jhmi.edu/
PsychENCODE [74]	Integration of expression and other regulatory elements across different brain cell types, regions and developmental periods	http://www.psychencode.org/

Finally, it is important to note that gene expression levels (either in absolute terms or relative to average levels across tissues) can be misinterpreted. For example, *DRD4* (coding for the dopamine D4 receptor) does not appear to be preferentially expressed in brain tissue, but it would be misleading to come to the conclusion therefore that *DRD4* has no role in psychopathology, as numerous studies have demonstrated DRD4 functioning to be involved in emotion and complex behaviours such as novelty seeking [94-96]. Furthermore, there is a renewed interest in dopamine D4 receptor-based pharmacological treatments for substance use and Parkinson disease [97]. As highlighted throughout this review, molecular phenotypes including DNA methylation and gene expression vary over time and across tissues meaning that any measure will be specific to the temporal context at which the sample was taken and thus limiting the inferences that can be made with respect to cause.

514 3.4.5. Gene ontology analysis

515 At last, it can be of interest to carry out an ontology analysis (or, relatedly, pathway or gene
516 property analyses) to investigate whether the most associated CpG probes cluster within distinct
517 biological functions. A plethora of online resources are available for ontology analyses and the reader
518 is referred to excellent reviews on the topic [98,99]. In general, analysis tools with the option to carry
519 out tissue-specific analyses are recommended. For example, FUMA (**http://fuma.ctglab.nl**, [100])
520 tests the relationships between tissue-specific gene expression and disease-gene associations, using
521 gene expression data from GTEx and the BrainSpan project, among others. As this resource was
522 primarily designed for genetic data, the user needs to map CpGs first to a gene before carrying out
523 the analysis using the GENE2FUNC option. With this functionality, Linnér et al. [88] reported that
524 genes closest to CpG probes linked to educational attainment were not preferentially expressed in
525 brain tissue, suggesting that findings might have been driven by confounding factors.

4. Strengths and limitations

527 Epigenetic epidemiological studies of mental health and related phenotypes continue to be the
528 focus of much interest with the hope of enhancing understanding of the biological mechanisms
529 underlying the aetiology and progression of psychiatric diseases. However, they still present
530 challenges and limitations.
531 The platforms to generate data that have been most widely employed sample only a very small
532 portion of CpG sites in the genome. Studies using sequencing-based approaches such as a recent
533 methylome-wide association study of major depressive disorder that measured DNA methylation in
534 28 million CpGs promise to unlock more information on epigenetic variation and will unravel more
535 insights into the role of methylation in mental health [101]. Moreover, while the majority of the
536 current studies focus on CpG methylation, DNA methylation is also present at non-CpG sites,
537 particularly in brain tissue, suggesting a potential role in neurodevelopment and mental health [102].
538 Methylome sequencing only recently allowed the characterisation of non-CpG methylation in brain
539 tissue [103], but could provide an additional avenue to discover novel effects in relation to
540 neuropsychiatric traits.
541 Mendelian randomization is proving to be a useful tool to strengthen causal inference and
542 explore molecular mediation by DNA methylation. It does however have recognized limitations and
543 is unlikely to provide definitive evidence of causal pathways without triangulation using
544 complementary approaches in epidemiology and other disciplines.
545 Epidemiological studies of methylation and brain-related processes using peripheral tissue
546 alone may not be able to unravel true biological mechanisms, but the associations found can be
547 translated in useful biomarkers (whether causal or not) for diseases or their progression and therefore
548 are worthwhile investigating. They can also be used to establish how substantial the contribution of
549 genetic factors to variance in methylation is. Also, it is often of interest to know whether a CpG
550 impacts gene expression (or vice versa), even if not causally linked to disease. Finally, these
551 approaches are useful to explain the correlation between peripheral DNA methylation to brain-based
552 processes, even if these processes index (non-causal) disease correlates. Even with the limitation of
553 not necessarily addressing the issue of causal correlates of psychiatric diseases that could be
554 translated into intervention, peripheral epigenetic associations can answer biological questions that
555 ultimately help understanding mental health.

5. Future perspectives and conclusions

557 In conclusion, recently developed openly accessible resources allow epigenetic epidemiological
558 studies of mental health and offer multiple opportunities to understand the aetiology and
559 progression of psychiatric conditions. Future advances in software development specific for
560 epigenetics and statistical methodologies for causal inference as well as large biobanks in multiple
561 complementary populations will increase substantially our understanding of mental health and lead
562 to the generation of reproducible results to inform prevention and intervention strategies.

568 **References**

569

570

571 1. Montano, C.; Taub, M.A.; Jaffe, A.; Briem, E.; Feinberg, J.I.; Trygvadottir, R.; Idrizi, A.; Runarsson, A.;
572 Berndsen, B.; Gur, R.C., et al. Association of DNA Methylation Differences With Schizophrenia in an
573 Epigenome-Wide Association Study. *JAMA Psychiatry* **2016**, *73*, 506-514,
574 doi:10.1001/jamapsychiatry.2016.0144.

575 2. Bani-Fatemi, A.; Jeremian, R.; Wang, K.Z.; Silveira, J.; Zai, C.; Kolla, N.J.; Graff, A.; Gerretsen, P.;
576 Strauss, J.; De Luca, V. Epigenome-wide association study of suicide attempt in schizophrenia. *J
577 Psychiatr Res* **2018**, *104*, 192-197, doi:10.1016/j.jpsychires.2018.07.011.

578 3. Story Jovanova, O.; Nedeljkovic, I.; Spieler, D.; Walker, R.M.; Liu, C.; Luciano, M.; Bressler, J.; Brody,
579 J.; Drake, A.J.; Evans, K.L., et al. DNA Methylation Signatures of Depressive Symptoms in Middle-
580 aged and Elderly Persons: Meta-analysis of Multiethnic Epigenome-wide Studies. *JAMA Psychiatry*
581 **2018**, *75*, 949-959, doi:10.1001/jamapsychiatry.2018.1725.

582 4. Baselmans, B.M.; van Dongen, J.; Nivard, M.G.; Lin, B.D.; Consortium, B.; Zilhao, N.R.; Boomsma,
583 D.I.; Bartels, M. Epigenome-Wide Association Study of Wellbeing. *Twin Res Hum Genet* **2015**, *18*, 710-
584 719, doi:10.1017/thg.2015.85.

585 5. Shimada-Sugimoto, M.; Otowa, T.; Miyagawa, T.; Umekage, T.; Kawamura, Y.; Bundo, M.; Iwamoto,
586 K.; Tochigi, M.; Kasai, K.; Kaiya, H., et al. Epigenome-wide association study of DNA methylation in
587 panic disorder. *Clin Epigenetics* **2017**, *9*, 6, doi:10.1186/s13148-016-0307-1.

588 6. Shimada, M.; Otowa, T.; Miyagawa, T.; Umekage, T.; Kawamura, Y.; Bundo, M.; Iwamoto, K.;
589 Ikegame, T.; Tochigi, M.; Kasai, K., et al. An epigenome-wide methylation study of healthy
590 individuals with or without depressive symptoms. *J Hum Genet* **2018**, *63*, 319-326, doi:10.1038/s10038-
591 017-0382-y.

592 7. Kuan, P.F.; Waszczuk, M.A.; Kotov, R.; Marsit, C.J.; Guffanti, G.; Gonzalez, A.; Yang, X.; Koenen, K.;
593 Bromet, E.; Luft, B.J. An epigenome-wide DNA methylation study of PTSD and depression in World
594 Trade Center responders. *Transl Psychiatry* **2017**, *7*, e1158, doi:10.1038/tp.2017.130.

595 8. Marioni, R.E.; McRae, A.F.; Bressler, J.; Colicino, E.; Hannon, E.; Li, S.; Prada, D.; Smith, J.A.; Trevisi,
596 L.; Tsai, P.C., et al. Meta-analysis of epigenome-wide association studies of cognitive abilities. *Mol
597 Psychiatry* **2018**, 10.1038/s41380-017-0008-y, doi:10.1038/s41380-017-0008-y.

598 9. Walton, E.; Pingault, J.B.; Cecil, C.A.; Gaunt, T.R.; Relton, C.L.; Mill, J.; Barker, E.D. Epigenetic
599 profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. *Mol Psychiatry* **2017**,
600 *22*, 250-256, doi:10.1038/mp.2016.85.

601 10. Barker, E.D.; Walton, E.; Cecil, C.A.M.; Rowe, R.; Jaffee, S.R.; Maughan, B.; O'Connor, T.G.; Stringaris,
602 A.; Meehan, A.J.; McArdle, W., et al. A Methylome-Wide Association Study of Trajectories of
603 Oppositional Defiant Behaviors and Biological Overlap With Attention Deficit Hyperactivity
604 Disorder. *Child Dev* **2018**, *89*, 1839-1855, doi:10.1111/cdev.12957.

605   11.   de Vocht, F.; Suderman, M.; Tilling, K.; Heron, J.; Howe, L.D.; Campbell, R.; Hickman, M.; Relton, C.
606   DNA methylation from birth to late adolescence and development of multiple-risk behaviours. *J*
607   *Affect Disord* **2018**, *227*, 588-594, doi:10.1016/j.jad.2017.11.055.

608   12.   Cecil, C.A.; Walton, E.; Smith, R.G.; Viding, E.; McCrory, E.J.; Relton, C.L.; Suderman, M.; Pingault,
609   J.B.; McArdle, W.; Gaunt, T.R., et al. DNA methylation and substance-use risk: a prospective, genome-
610   wide study spanning gestation to adolescence. *Transl Psychiatry* **2016**, *6*, e976, doi:10.1038/tp.2016.247.

611   13.   Cecil, C.A.M.; Walton, E.; Jaffee, S.R.; O'Connor, T.; Maughan, B.; Relton, C.L.; Smith, R.G.; McArdle,
612   W.; Gaunt, T.R.; Ouellet-Morin, I., et al. Neonatal DNA methylation and early-onset conduct
613   problems: A genome-wide, prospective study. *Dev Psychopathol* **2018**, *30*, 383-397,
614   doi:10.1017/S095457941700092X.

615   14.   Cecil, C.A.M.; Walton, E.; Pingault, J.B.; Provencal, N.; Pappa, I.; Vitaro, F.; Cote, S.; Szyf, M.;
616   Tremblay, R.E.; Tiemeier, H., et al. DRD4 methylation as a potential biomarker for physical
617   aggression: An epigenome-wide, cross-tissue investigation. *Am J Med Genet B Neuropsychiatr Genet*
618   **2018**, 10.1002/ajmg.b.32689, doi:10.1002/ajmg.b.32689.

619   15.   Mitjans, M.; Seidel, J.; Begemann, M.; Bockhop, F.; Moya-Higueras, J.; Bansal, V.; Wesolowski, J.;
620   Seelbach, A.; Ibanez, M.I.; Kovacevic, F., et al. Violent aggression predicted by multiple pre-adult
621   environmental hits. *Mol Psychiatry* **2018**, 10.1038/s41380-018-0043-3, doi:10.1038/s41380-018-0043-3.

622   16.   Andrews, S.V.; Sheppard, B.; Windham, G.C.; Schieve, L.A.; Schendel, D.E.; Croen, L.A.; Chopra, P.;
623   Alisch, R.S.; Newschaffer, C.J.; Warren, S.T., et al. Case-control meta-analysis of blood DNA
624   methylation and autism spectrum disorder. *Mol Autism* **2018**, *9*, 40, doi:10.1186/s13229-018-0224-6.

625   17.   Hannon, E.; Schendel, D.; Ladd-Acosta, C.; Grove, J.; i, P.-B.A.S.D.G.; Hansen, C.S.; Andrews, S.V.;
626   Hougaard, D.M.; Bresnahan, M.; Mors, O., et al. Elevated polygenic burden for autism is associated
627   with differential DNA methylation at birth. *Genome Med* **2018**, *10*, 19, doi:10.1186/s13073-018-0527-4.

628   18.   Long, H.Y.; Feng, L.; Kang, J.; Luo, Z.H.; Xiao, W.B.; Long, L.L.; Yan, X.X.; Zhou, L.; Xiao, B. Blood
629   DNA methylation pattern is altered in mesial temporal lobe epilepsy. *Sci Rep* **2017**, *7*, 43810,
630   doi:10.1038/srep43810.

631   19.   Shimada, M.; Miyagawa, T.; Toyoda, H.; Tokunaga, K.; Honda, M. Epigenome-wide association study
632   of DNA methylation in narcolepsy: an integrated genetic and epigenetic approach. *Sleep* **2018**, *41*,
633   doi:10.1093/sleep/zsy019.

634   20.   Kaut, O.; Schmitt, I.; Tost, J.; Busato, F.; Liu, Y.; Hofmann, P.; Witt, S.H.; Rietschel, M.; Frohlich, H.;
635   Wullner, U. Epigenome-wide DNA methylation analysis in siblings and monozygotic twins
636   discordant for sporadic Parkinson's disease revealed different epigenetic patterns in peripheral blood
637   mononuclear cells. *Neurogenetics* **2017**, *18*, 7-22, doi:10.1007/s10048-016-0497-x.

638   21.   Karlsson Linner, R.; Marioni, R.E.; Rietveld, C.A.; Simpkin, A.J.; Davies, N.M.; Watanabe, K.;
639   Armstrong, N.J.; Auro, K.; Baumbach, C.; Bonder, M.J., et al. An epigenome-wide association study
640   meta-analysis of educational attainment. *Mol Psychiatry* **2017**, *22*, 1680-1690, doi:10.1038/mp.2017.210.

641   22.   Bakulski, K.M.; Halladay, A.; Hu, V.W.; Mill, J.; Fallin, M.D. Epigenetic Research in Neuropsychiatric
642   Disorders: the "Tissue Issue". *Curr Behav Neurosci Rep* **2016**, *3*, 264-274, doi:10.1007/s40473-016-0083-4.

643   23.   Ladd-Acosta, C.; Hansen, K.D.; Briem, E.; Fallin, M.D.; Kaufmann, W.E.; Feinberg, A.P. Common
644   DNA methylation alterations in multiple brain regions in autism. *Mol Psychiatry* **2014**, *19*, 862-871,
645   doi:10.1038/mp.2013.114.

646   24.   Kinoshita, M.; Numata, S.; Tajima, A.; Yamamori, H.; Yasuda, Y.; Fujimoto, M.; Watanabe, S.;
647   Umehara, H.; Shimodera, S.; Nakazawa, T., et al. Effect of Clozapine on DNA Methylation in

648   Peripheral Leukocytes from Patients with Treatment-Resistant Schizophrenia. *Int J Mol Sci* **2017**, *18*,
649   doi:10.3390/ijms18030632.

650 25. Hutnick, L.K.; Golshani, P.; Namihira, M.; Xue, Z.; Matynia, A.; Yang, X.W.; Silva, A.J.; Schweizer,
651   F.E.; Fan, G. DNA hypomethylation restricted to the murine forebrain induces cortical degeneration
652   and impairs postnatal neuronal maturation. *Hum Mol Genet* **2009**, *18*, 2875-2888,
653   doi:10.1093/hmg/ddp222.

654 26. Feng, J.; Zhou, Y.; Campbell, S.L.; Le, T.; Li, E.; Sweatt, J.D.; Silva, A.J.; Fan, G. Dnmt1 and Dnmt3a
655   maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci*
656   **2010**, *13*, 423-430, doi:10.1038/nn.2514.

657 27. Morris, M.J.; Na, E.S.; Autry, A.E.; Monteggia, L.M. Impact of DNMT1 and DNMT3a forebrain
658   knockout on depressive- and anxiety like behavior in mice. *Neurobiol Learn Mem* **2016**, *135*, 139-145,
659   doi:10.1016/j.nlm.2016.08.012.

660 28. Liu, X.S.; Wu, H.; Krzisch, M.; Wu, X.; Graef, J.; Muffat, J.; Hnisz, D.; Li, C.H.; Yuan, B.; Xu, C., et al.
661   Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell* **2018**,
662   *172*, 979-992 e976, doi:10.1016/j.cell.2018.01.012.

663 29. Felix, J.F.; Joubert, B.R.; Baccarelli, A.A.; Sharp, G.C.; Almqvist, C.; Annesi-Maesano, I.; Arshad, H.;
664   Baiz, N.; Bakermans-Kranenburg, M.J.; Bakulski, K.M., et al. Cohort Profile: Pregnancy And
665   Childhood Epigenetics (PACE) Consortium. *Int J Epidemiol* **2017**, 10.1093/ije/dyx190,
666   doi:10.1093/ije/dyx190.

667 30. Psaty, B.M.; O'Donnell, C.J.; Gudnason, V.; Lunetta, K.L.; Folsom, A.R.; Rotter, J.I.; Uitterlinden, A.G.;
668   Harris, T.B.; Witteman, J.C.; Boerwinkle, E., et al. Cohorts for Heart and Aging Research in Genomic
669   Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide
670   association studies from 5 cohorts. *Circ Cardiovasc Genet* **2009**, *2*, 73-80,
671   doi:10.1161/CIRCGENETICS.108.829747.

672 31. Ligthart, S.; Marzi, C.; Aslibekyan, S.; Mendelson, M.M.; Conneely, K.N.; Tanaka, T.; Colicino, E.;
673   Waite, L.L.; Joehanes, R.; Guan, W., et al. DNA methylation signatures of chronic low-grade
674   inflammation are associated with complex diseases. *Genome Biol* **2016**, *17*, 255, doi:10.1186/s13059-016-
675   1119-5.

676 32. Min, J.L.; Hemani, G.; Davey Smith, G.; Relton, C.; Suderman, M. Meffil: efficient normalization and
677   analysis of very large DNA methylation datasets. *Bioinformatics* **2018**, 10.1093/bioinformatics/bty476,
678   doi:10.1093/bioinformatics/bty476.

679 33. Barrett, T.; Troup, D.B.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.;
680   Marshall, K.A.; Phillippy, K.H.; Sherman, P.M., et al. NCBI GEO: archive for functional genomics data
681   sets--10 years on. *Nucleic Acids Res* **2011**, *39*, D1005-1010, doi:10.1093/nar/gkq1184.

682 34. Li, H.; Chen, R.; Cai, J.; Cui, X.; Huang, N.; Kan, H. Short-term exposure to fine particulate air
683   pollution and genome-wide DNA methylation: A randomized, double-blind, crossover trial. *Environ*
684   *Int* **2018**, *120*, 130-136, doi:10.1016/j.envint.2018.07.041.

685 35. Richmond, R.C.; Sharp, G.C.; Herbert, G.; Atkinson, C.; Taylor, C.; Bhattacharya, S.; Campbell, D.;
686   Hall, M.; Kazmi, N.; Gaunt, T., et al. The long-term impact of folic acid in pregnancy on offspring
687   DNA methylation: follow-up of the Aberdeen Folic Acid Supplementation Trial (AFAST). *Int J*
688   *Epidemiol* **2018**, 10.1093/ije/dyy032, doi:10.1093/ije/dyy032.

689 36. Tobi, E.W.; Goeman, J.J.; Monajemi, R.; Gu, H.; Putter, H.; Zhang, Y.; Slieker, R.C.; Stok, A.P.; Thijssen,
690 P.E.; Muller, F., et al. DNA methylation signatures link prenatal famine exposure to growth and
691 metabolism. *Nat Commun* **2014**, *5*, 5592, doi:10.1038/ncomms6592.

692 37. van den Broek, T.; Fleischmann, M. Prenatal famine exposure and mental health in later midlife.
693 *Aging Ment Health* **2017**, 10.1080/13607863.2017.1402293, 1-5, doi:10.1080/13607863.2017.1402293.

694 38. Cao-Lei, L.; Elgbeili, G.; Massart, R.; Laplante, D.P.; Szyf, M.; King, S. Pregnant women's cognitive
695 appraisal of a natural disaster affects DNA methylation in their children 13 years later: Project Ice
696 Storm. *Transl Psychiatry* **2015**, *5*, e515, doi:10.1038/tp.2015.13.

697 39. Walder, D.J.; Laplante, D.P.; Sousa-Pires, A.; Veru, F.; Brunet, A.; King, S. Prenatal maternal stress
698 predicts autism traits in 6(1/2) year-old children: Project Ice Storm. *Psychiatry Res* **2014**, *219*, 353-360,
699 doi:10.1016/j.psychres.2014.04.034.

700 40. Davey Smith, G.; Hemani, G. Mendelian randomization: genetic anchors for causal inference in
701 epidemiological studies. *Hum Mol Genet* **2014**, *23*, R89-98, doi:10.1093/hmg/ddu328.

702 41. Zheng, J.; Baird, D.; Borges, M.C.; Bowden, J.; Hemani, G.; Haycock, P.; Evans, D.M.; Smith, G.D.
703 Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep* **2017**, *4*, 330-345,
704 doi:10.1007/s40471-017-0128-6.

705 42. Pingault, J.B.; O'Reilly, P.F.; Schoeler, T.; Ploubidis, G.B.; Rijsdijk, F.; Dudbridge, F. Using genetic data
706 to strengthen causal inference in observational research. *Nat Rev Genet* **2018**, *19*, 566-580,
707 doi:10.1038/s41576-018-0020-3.

708 43. Teumer, A. Common Methods for Performing Mendelian Randomization. *Front Cardiovasc Med* **2018**,
709 *5*, 51, doi:10.3389/fcvm.2018.00051.

710 44. Richardson, T.G.; Haycock, P.C.; Zheng, J.; Timpson, N.J.; Gaunt, T.R.; Smith, G.D.; Relton, C.L.;
711 Hemani, G. Systematic Mendelian randomization framework elucidates hundreds of genetic loci
712 which may influence disease through changes in DNA methylation levels. *bioRxiv* **2017**,
713 10.1101/189076, 189076, doi:10.1101/189076.

714 45. Walton, E.; Hemani, G.; Dehghan, A.; Relton, C.; Smith, G.D. Systematic evaluation of the causal
715 relationship between DNA methylation and C-reactive protein. *bioRxiv* **2018**, 10.1101/397836, 397836,
716 doi:10.1101/397836.

717 46. Lin, D.; Chen, J.; Perrone-Bizzozero, N.; Bustillo, J.R.; Du, Y.; Calhoun, V.D.; Liu, J. Characterization of
718 cross-tissue genetic-epigenetic effects and their patterns in schizophrenia. *Genome Med* **2018**, *10*, 13,
719 doi:10.1186/s13073-018-0519-4.

720 47. Gaunt, T.R.; Shihab, H.A.; Hemani, G.; Min, J.L.; Woodward, G.; Lyttleton, O.; Zheng, J.; Duggirala,
721 A.; McArdle, W.L.; Ho, K., et al. Systematic identification of genetic influences on methylation across
722 the human life course. *Genome Biol* **2016**, *17*, 61, doi:10.1186/s13059-016-0926-z.

723 48. Bonder, M.J.; Luijk, R.; Zhernakova, D.V.; Moed, M.; Deelen, P.; Vermaat, M.; van Iterson, M.; van
724 Dijk, F.; van Galen, M.; Bot, J., et al. Disease variants alter transcription factor levels and methylation
725 of their binding sites. *Nat Genet* **2017**, *49*, 131-138, doi:10.1038/ng.3721.

726 49. Hannon, E.; Spiers, H.; Viana, J.; Pidsley, R.; Burrage, J.; Murphy, T.M.; Troakes, C.; Turecki, G.;
727 O'Donovan, M.C.; Schalkwyk, L.C., et al. Methylation QTLs in the developing brain and their
728 enrichment in schizophrenia risk loci. *Nat Neurosci* **2016**, *19*, 48-54, doi:10.1038/nn.4182.

729 50. Chausse, P. Computing Generalized Method of Moments and Generalized Empirical Likelihood with
730 R. *Journal of Statistical Software* **2010**, *34*, 1-35.

731   51.   Thomas, D.C.; Lawlor, D.A.; Thompson, J.R. Re: Estimation of bias in nongenetic observational
732   studies using "Mendelian triangulation" by Bautista et al. *Ann Epidemiol* **2007**, *17*, 511-513,
733   doi:10.1016/j.annepidem.2006.12.005.

734   52.   Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.;
735   Bowden, J.; Langdon, R., et al. The MR-base platform enables systematic causal inference across the
736   phenome. *eLife* **2018**.

737   53.   Yavorska, O.O.; Burgess, S. MendelianRandomization: an R package for performing Mendelian
738   randomization analyses using summarized data. *Int J Epidemiol* **2017**, *46*, 1734-1739,
739   doi:10.1093/ije/dyx034.

740   54.   Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian randomization with invalid instruments: effect
741   estimation and bias detection through Egger regression. *Int J Epidemiol* **2015**, *44*, 512-525,
742   doi:10.1093/ije/dyv080.

743   55.   Staley, J.R.; Blackshaw, J.; Kamat, M.A.; Ellis, S.; Surendran, P.; Sun, B.B.; Paul, D.S.; Freitag, D.;
744   Burgess, S.; Danesh, J., et al. PhenoScanner: a database of human genotype-phenotype associations.
745   *Bioinformatics* **2016**, *32*, 3207-3209, doi:10.1093/bioinformatics/btw373.

746   56.   MacArthur, J.; Bowler, E.; Cerezo, M.; Gil, L.; Hall, P.; Hastings, E.; Junkins, H.; McMahon, A.; Milano,
747   A.; Morales, J., et al. The new NHGRI-EBI Catalog of published genome-wide association studies
748   (GWAS Catalog). *Nucleic Acids Research* **2017**, *45*, D896-D901, doi:10.1093/nar/gkw1133.

749   57.   Sullivan, P.F.; Agrawal, A.; Bulik, C.M.; Andreassen, O.A.; Borglum, A.D.; Breen, G.; Cichon, S.;
750   Edenberg, H.J.; Faraone, S.V.; Gelernter, J., et al. Psychiatric Genomics: An Update and an Agenda.
751   *Am J Psychiatry* **2018**, *175*, 15-27, doi:10.1176/appi.ajp.2017.17030283.

752   58.   Thompson, P.M.; Andreassen, O.A.; Arias-Vasquez, A.; Bearden, C.E.; Boedhoe, P.S.; Brouwer, R.M.;
753   Buckner, R.L.; Buitelaar, J.K.; Bulayeva, K.B.; Cannon, D.M., et al. ENIGMA and the individual:
754   Predicting factors that affect the brain in 35 countries worldwide. *Neuroimage* **2017**, *145*, 389-408,
755   doi:10.1016/j.neuroimage.2015.11.057.

756   59.   Kirkbride, J.B.; Susser, E.; Kundakovic, M.; Kresovich, J.K.; Davey Smith, G.; Relton, C.L. Prenatal
757   nutrition, epigenetics and schizophrenia risk: can we test causal effects? *Epigenomics* **2012**, *4*, 303-315,
758   doi:10.2217/epi.12.20.

759   60.   Caramaschi, D.; Sharp, G.C.; Nohr, E.A.; Berryman, K.; Lewis, S.J.; Davey Smith, G.; Relton, C.L.
760   Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during
761   pregnancy and child's IQ at age 8, cognitive performance and educational attainment: a two-step
762   Mendelian randomization study. *Hum Mol Genet* **2017**, *26*, 3001-3013, doi:10.1093/hmg/ddx164.

763   61.   Viana, J.; Hannon, E.; Dempster, E.; Pidsley, R.; Macdonald, R.; Knox, O.; Spiers, H.; Troakes, C.; Al-
764   Saraj, S.; Turecki, G., et al. Schizophrenia-associated methylomic variation: molecular signatures of
765   disease and polygenic risk burden across multiple brain regions. *Hum Mol Genet* **2017**, *26*, 210-225,
766   doi:10.1093/hmg/ddw373.

767   62.   Hannon, E.; Dempster, E.; Viana, J.; Burrage, J.; Smith, A.R.; Macdonald, R.; St Clair, D.; Mustard, C.;
768   Breen, G.; Therman, S., et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for
769   co-localization of genetic associations and differential DNA methylation. *Genome Biol* **2016**, *17*, 176,
770   doi:10.1186/s13059-016-1041-x.

771   63.   Choi, S.W.; Mak, T.S.H.; O'Reilly, P. A guide to performing Polygenic Risk Score analyses. *bioRxiv*
772   **2018**, 10.1101/416545, 416545, doi:10.1101/416545.

773 64. Giambartolomei, C.; Vukcevic, D.; Schadt, E.E.; Franke, L.; Hingorani, A.D.; Wallace, C.; Plagnol, V.
774 Bayesian test for colocalisation between pairs of genetic association studies using summary statistics.
775 *PLoS Genet* **2014**, *10*, e1004383, doi:10.1371/journal.pgen.1004383.

776 65. Walton, E.; Hass, J.; Liu, J.; Roffman, J.L.; Bernardoni, F.; Roessner, V.; Kirsch, M.; Schackert, G.;
777 Calhoun, V.; Ehrlich, S. Correspondence of DNA Methylation Between Blood and Brain Tissue and Its
778 Application to Schizophrenia Research. *Schizophr Bull* **2016**, *42*, 406-414, doi:10.1093/schbul/sbv074.

779 66. Mohammad-Zadeh, L.F.; Moses, L.; Gwaltney-Brant, S.M. Serotonin: a review. *J Vet Pharmacol Ther*
780 **2008**, *31*, 187-199, doi:10.1111/j.1365-2885.2008.00944.x.

781 67. Biomarkers Definitions Working, G. Biomarkers and surrogate endpoints: preferred definitions and
782 conceptual framework. *Clin Pharmacol Ther* **2001**, *69*, 89-95, doi:10.1067/mcp.2001.113989.

783 68. Mitchell, R.E.; Paternoster, L.; Davey Smith, G. Mendelian Randomization in Case Only Studies: A
784 Promising Approach to be Applied With Caution. *Am J Cardiol* **2018**, 10.1016/j.amjcard.2018.09.035,
785 doi:10.1016/j.amjcard.2018.09.035.

786 69. Ng, B.; White, C.C.; Klein, H.U.; Sieberts, S.K.; McCabe, C.; Patrick, E.; Xu, J.; Yu, L.; Gaiteri, C.;
787 Bennett, D.A., et al. An xQTL map integrates the genetic architecture of the human brain's
788 transcriptome and epigenome. *Nat Neurosci* **2017**, *20*, 1418-1426, doi:10.1038/nn.4632.

789 70. Severin, J.; Waterhouse, A.M.; Kawaji, H.; Lassmann, T.; van Nimwegen, E.; Balwierz, P.J.; de Hoon,
790 M.J.; Hume, D.A.; Carninci, P.; Hayashizaki, Y., et al. FANTOM4 EdgeExpressDB: an integrated
791 database of promoters, genes, microRNAs, expression dynamics and regulatory interactions. *Genome*
792 *Biol* **2009**, *10*, R39, doi:10.1186/gb-2009-10-4-r39.

793 71. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome.
794 *Nature* **2012**, *489*, 57-74, doi:10.1038/nature11247.

795 72. Li, M.; Santpere, G.; Imamura Kawasawa, Y.; Evgrafov, O.V.; Gulden, F.O.; Pochareddy, S.; Sunkin,
796 S.M.; Li, Z.; Shin, Y.; Zhu, Y., et al. Integrative functional genomic analysis of human brain
797 development and neuropsychiatric risks. *Science* **2018**, *362*, eaat7615, doi:10.1126/science.aat7615

798 73. Gandal, M.J.; Zhang, P.; Hadjimichael, E.; Walker, R.L.; Chen, C.; Liu, S.; Won, H.; van Bakel, H.;
799 Varghese, M.; Wang, Y., et al. Transcriptome-wide isoform-level dysregulation in ASD,
800 schizophrenia, and bipolar disorder. *Science* **2018**, *362*, eaat8127, doi:10.1126/science.aat8127

801 74. Wang, D.; Liu, S.; Warrell, J.; Won, H.; Shi, X.; Navarro, F.C.P.; Clarke, D.; Gu, M.; Emani, P.; Yang,
802 Y.T., et al. Comprehensive functional genomic resource and integrative model for the human brain.
803 *Science* **2018**, *362*, eaat8464, doi:10.1126/science.aat8464.

804 75. Zhu, Y.; Sousa, A.M.M.; Gao, T.; Skarica, M.; Li, M.; Santpere, G.; Esteller-Cucala, P.; Juan, D.;
805 Ferrández-Peral, L.; Gulden, F.O., et al. Spatiotemporal transcriptomic divergence across human and
806 macaque brain development. *Science* **2018**, *362*, eaat8077, doi:10.1126/science.aat8077.

807 76. Amiri, A.; Coppola, G.; Scuderi, S.; Wu, F.; Roychowdhury, T.; Liu, F.; Pochareddy, S.; Shin, Y.; Safi,
808 A.; Song, L., et al. Transcriptome and epigenome landscape of human cortical development modeled
809 in organoids. *Science* **2018**, *362*, eaat6720, doi:10.1126/science.aat6720.

810 77. Rajarajan, P.; Borrman, T.; Liao, W.; Schrode, N.; Flaherty, E.; Casiño, C.; Powell, S.; Yashaswini, C.;
811 LaMarca, E.A.; Kassim, B., et al. Neuron-specific signatures in the chromosomal connectome
812 associated with schizophrenia risk. *Science* **2018**, *362*, eaat4311, doi:10.1126/science.aat4311.

813 78. An, J.-Y.; Lin, K.; Zhu, L.; Werling, D.M.; Dong, S.; Brand, H.; Wang, H.Z.; Zhao, X.; Schwartz, G.B.;
814 Collins, R.L., et al. Genome-wide de novo risk score implicates promoter variation in autism spectrum
815 disorder. *Science* **2018**, *362*, eaat6576, doi:10.1126/science.aat6576

816   79.   Rhie, S.K.; Schreiner, S.; Witt, H.; Armoskus, C.; Lay, F.D.; Camarena, A.; Spitsyna, V.N.; Guo, Y.;
817   Berman, B.P.; Evgrafov, O.V., et al. Using 3D epigenomic maps of primary olfactory neuronal cells
818   from living individuals to understand gene regulation. *Science Advances* **2018**, *4*, eaav8550,
819   doi:10.1126/sciadv.aav8550

820   80.   Kozlenkov, A.; Li, J.; Apontes, P.; Hurd, Y.L.; Byne, W.M.; Koonin, E.V.; Wegner, M.; Mukamel, E.A.;
821   Dracheva, S. A unique role for DNA (hydroxy)methylation in epigenetic regulation of human
822   inhibitory neurons. *Science Advances* **2018**, *4*, eaau6190, doi:10.1126/sciadv.aau6190.

823   81.   Chen, C.; Meng, Q.; Xia, Y.; Ding, C.; Wang, L.; Dai, R.; Cheng, L.; Gunaratne, P.; Gibbs, R.A.; Min, S.,
824   et al. The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from
825   patients with psychiatric disorders. *Science Translational Medicine* **2018**, *10*, eaat8178,
826   doi:10.1126/scitranslmed.aat8178

827   82.   Meng, Q.; Wang, K.; Brunetti, T.; Xia, Y.; Jiao, C.; Dai, R.; Fitzgerald, D.; Thomas, A.; Jay, L.; Eckart,
828   H., et al. The DGCR5 long noncoding RNA may regulate expression of several
829   schizophrenia-related genes. *Science Translational Medicine* **2018**, *10*, eaat6912,
830   doi:10.1126/scitranslmed.aat6912

831   83.   Roadmap Epigenomics, C.; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-
832   Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J., et al. Integrative analysis of 111 reference human
833   epigenomes. *Nature* **2015**, *518*, 317-330, doi:10.1038/nature14248.

834   84.   Van Baak, T.E.; Coarfa, C.; Dugue, P.A.; Fiorito, G.; Laritsky, E.; Baker, M.S.; Kessler, N.J.; Dong, J.;
835   Duryea, J.D.; Silver, M.J., et al. Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol*
836   **2018**, *19*, 2, doi:10.1186/s13059-017-1374-0.

837   85.   Edgar, R.D.; Jones, M.J.; Meaney, M.J.; Turecki, G.; Kobor, M.S. BECon: a tool for interpreting DNA
838   methylation findings from blood in the context of brain. *Transl Psychiatry* **2017**, *7*, e1187,
839   doi:10.1038/tp.2017.171.

840   86.   Song, Q.; Decato, B.; Hong, E.E.; Zhou, M.; Fang, F.; Qu, J.; Garvin, T.; Kessler, M.; Zhou, J.; Smith,
841   A.D. A reference methylome database and analysis pipeline to facilitate integrative and comparative
842   epigenomics. *PLoS One* **2013**, *8*, e81148, doi:10.1371/journal.pone.0081148.

843   87.   Breeze, C.E.; Paul, D.S.; van Dongen, J.; Butcher, L.M.; Ambrose, J.C.; Barrett, J.E.; Lowe, R.; Rakyan,
844   V.K.; Iotchkova, V.; Frontini, M., et al. eFORGE: A Tool for Identifying Cell Type-Specific Signal in
845   Epigenomic Data. *Cell Rep* **2016**, *17*, 2137-2150, doi:10.1016/j.celrep.2016.10.059.

846   88.   Karlsson Linnér, R.; Marioni, R.E.; Rietveld, C.A.; Simpkin, A.J.; Davies, N.M.; Watanabe, K.;
847   Armstrong, N.J.; Auro, K.; Baumbach, C.; Bonder, M.J., et al. An epigenome-wide association study
848   meta-analysis of educational attainment. *Molecular Psychiatry* **2017**, *22*, 1680-1690,
849   doi:10.1038/mp.2017.210.

850   89.   Schlosberg, C.E.; VanderKraats, N.D.; Edwards, J.R. Modeling complex patterns of differential DNA
851   methylation that associate with gene expression changes. *Nucleic Acids Res* **2017**, *45*, 5100-5111,
852   doi:10.1093/nar/gkx078.

853   90.   Keil, J.M.; Qalieh, A.; Kwan, K.Y. Brain transcriptome databases: a user's guide. *J Neurosci* **2018**,
854   10.1523/JNEUROSCI.1930-17.2018, doi:10.1523/JNEUROSCI.1930-17.2018.

855   91.   Hawrylycz, M.J.; Lein, E.S.; Guillozet-Bongaarts, A.L.; Shen, E.H.; Ng, L.; Miller, J.A.; van de
856   Lagemaat, L.N.; Smith, K.A.; Ebbert, A.; Riley, Z.L., et al. An anatomically comprehensive atlas of the
857   adult human brain transcriptome. *Nature* **2012**, *489*, 391-399, doi:10.1038/nature11405.

858 92. Miller, J.A.; Ding, S.L.; Sunkin, S.M.; Smith, K.A.; Ng, L.; Szafer, A.; Ebbert, A.; Riley, Z.L.; Royall, J.J.;
859 Aiona, K., et al. Transcriptional landscape of the prenatal human brain. *Nature* **2014**, *508*, 199-206,
860 doi:10.1038/nature13185.

861 93. Colantuoni, C.; Lipska, B.K.; Ye, T.; Hyde, T.M.; Tao, R.; Leek, J.T.; Colantuoni, E.A.; Elkahloun, A.G.;
862 Herman, M.M.; Weinberger, D.R., et al. Temporal dynamics and genetic control of transcription in the
863 human prefrontal cortex. *Nature* **2011**, *478*, 519-523, doi:10.1038/nature10524.

864 94. Munafò, M.R.; Yalcin, B.; Willis-Owen, S.A.; Flint, J. Association of the dopamine D4 receptor (DRD4)
865 gene and approach-related personality traits: meta-analysis and new data. *Biol. Psychiatry* **2008**, *63*,
866 197-206, doi:10.1016/j.biopsych.2007.04.006.

867 95. Ptácek, R.; Kuzelová, H.; Stefano, G.B. Dopamine D4 receptor gene DRD4 and its association with
868 psychiatric disorders. *Med. Sci. Monit.* **2011**, *17*, RA215-220.

869 96. Di Ciano, P.; Grandy, D.K.; Le Foll, B. Dopamine D4 receptors in psychostimulant addiction. *Adv.*
870 *Pharmacol.* **2014**, *69*, 301-321, doi:10.1016/B978-0-12-420118-7.00008-1.

871 97. Lindsley, C.W.; Hopkins, C.R. Return of D4 Dopamine Receptor Antagonists in Drug Discovery. *J.*
872 *Med. Chem.* **2017**, *60*, 7233-7243, doi:10.1021/acs.jmedchem.7b00151.

873 98. Cirillo, E.; Parnell, L.D.; Evelo, C.T. A Review of Pathway-Based Analysis Tools That Visualize
874 Genetic Variants. *Front Genet* **2017**, *8*, doi:10.3389/fgene.2017.00174.

875 99. Kao, P.Y.P.; Leung, K.H.; Chan, L.W.C.; Yip, S.P.; Yap, M.K.H. Pathway analysis of complex diseases
876 for GWAS, extending to consider rare variants, multi-omics and interactions. *Biochimica et Biophysica*
877 *Acta (BBA) - General Subjects* **2017**, *1861*, 335-353, doi:10.1016/j.bbagen.2016.11.030.

878 100. Watanabe, K.; Taskesen, E.; van Bochoven, A.; Posthuma, D. Functional mapping and annotation of
879 genetic associations with FUMA. *Nat Commun* **2017**, *8*, 1826, doi:10.1038/s41467-017-01261-5.

880 101. Aberg, K.A.; Dean, B.; Shabalin, A.A.; Chan, R.F.; Han, L.K.M.; Zhao, M.; van Grootheest, G.; Xie, L.Y.;
881 Milaneschi, Y.; Clark, S.L., et al. Methylome-wide association findings for major depressive disorder
882 overlap in blood and brain and replicate in independent brain samples. *Mol Psychiatry* **2018**,
883 10.1038/s41380-018-0247-6, doi:10.1038/s41380-018-0247-6.

884 102. Jang, H.S.; Shin, W.J.; Lee, J.E.; Do, J.T. CpG and Non-CpG Methylation in Epigenetic Gene
885 Regulation and Brain Function. *Genes (Basel)* **2017**, *8*, doi:10.3390/genes8060148.

886 103. Rizzardi, L.F.; Hickey, P.F.; Rodriguez DiBlasi, V.; Tryggvadottir, R.; Callahan, C.M.; Idrizi, A.;
887 Hansen, K.D.; Feinberg, A.P. Neuronal brain-region-specific DNA methylation and chromatin
888 accessibility are associated with neuropsychiatric trait heritability. *Nat Neurosci* **2019**, *22*, 307-316,
889 doi:10.1038/s41593-018-0297-8.

890

891