

Translational Epigenetics

# Twin and Family Studies of Epigenetics

Volume 27

Series Editor

**Trygve Tollefsbol**

Professor of Biology, University of Alabama at Birmingham, and Senior Scientist, Comprehensive Cancer Center, Comprehensive Center for Healthy Aging, Comprehensive Diabetes Center and Nutrition Obesity Research Center  
Director, Cell Senescence Culture Facility, Birmingham, AL, United States

Edited by

**Shuai Li**

Centre for Epidemiology and Biostatistics, The University of Melbourne,  
Parkville, VIC, Australia

**John L. Hopper**

Centre for Epidemiology and Biostatistics, The University of Melbourne,  
Parkville, VIC, Australia



**ACADEMIC PRESS**

An imprint of Elsevier

# Combining twin-family designs with measured genetic variants to study the causes of epigenetic variation

**Camelia C. Minică<sup>a,b,c</sup>, Michael C. Neale<sup>d</sup>, Dorret I. Boomsma<sup>a</sup>, and Jenny van Dongen<sup>a</sup>**

<sup>a</sup>*Department of Biological Psychology, Netherlands Twin Register, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*

<sup>b</sup>*Stanley Center for Psychiatric Disease, Broad Institute of MIT and Harvard, Cambridge, MA, United States* <sup>c</sup>*Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, United States* <sup>d</sup>*Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, United States*

## 1 Introduction

In recent years, epigenome-wide association studies (EWASs) have identified thousands of loci where DNA methylation is associated with complex human traits, diseases, and environmental exposures. The majority of these studies have focused on DNA methylation,<sup>1</sup> which is one of the best-understood epigenetic mechanisms. The role of DNA methylation in development and environmental responses is well established; however, many more epigenetic mechanisms exist that are at present largely under-explored in EWASs.<sup>2,3</sup>

The results from EWA studies are valuable for understanding the underlying pathways of traits, and for the development of biomarkers of diseases and environmental exposures. Yet, the interpretation of epigenetic associations is not easy. Because DNA methylation is dynamic, DNA methylation signatures associated with a disease might be causal to the disease or may have arisen secondary to disease development. Another option is that the association is caused by a third factor that influences DNA methylation and the disease independently (for example, lifestyle, prenatal exposures, or genetic variants). Although the epigenome is often discussed in the context of environmental explanations for diseases, the epigenome is also shaped by genetic influences. In fact, the epigenome is an important mediator of genetic variation in regulatory regions of the genome on complex traits and disease.<sup>4</sup> For instance, disease-associated SNPs are often associated with expression levels of transcription factors, which in turn drive variation in DNA methylation level of distal binding sites.<sup>5</sup>

EWA studies are contributing to the creation of catalogs of disease-, trait-, and exposure-associated epigenetic variation.<sup>6</sup> At the same time, large-scale methylation Quantitative Trait Loci (mQTL) analyses are performed. In mQTL studies the goal is to map associations between genetic variants (typically, Single Nucleotide Polymorphisms, SNPs) and DNA methylation levels across the genome.<sup>7</sup> First, the results of such mQTL projects have led to catalogs that are valuable for providing insights into the

downstream effects of genetic variants identified in genome-wide association studies (GWAS), because SNPs identified as genome-wide significant can be then be characterized in their roles as mQTLs. Second, catalogs deliver “instruments” for so-called Mendelian Randomization (MR) analysis. MR has become a feasible and popular approach to examine causal relationships between DNA methylation and a trait of interest.<sup>8,9</sup> The basic MR idea was proposed by Katan<sup>10,11</sup>: if low cholesterol has a causal effect on the risk to develop cancer, a genotype that reduces cholesterol level should increase the risk of cancer.<sup>10</sup> The rationale for using genotypes is to rule out effects of confounding factors: genotypes are randomly assigned at conception.

In this chapter, we discuss how twin-family studies combined with measured genetic information (e.g., SNPs) and epigenetic information (e.g., DNA methylation) can make a unique complementary contribution to these ongoing efforts in the field of human epigenetics. We discuss two distinct applications. We first describe the classical twin design and its extensions with measured genetic information that can provide information on the causes of epigenetic variation, including estimating the total heritability of epigenetic marks, estimating the contribution of measured SNPs to the total heritability, and examining age- and sex-specific genetic and environmental influences on epigenetic variation. Such studies contribute to catalogs of between-individual variation in DNA methylation.<sup>12</sup> Second, we describe an integration of MR with the Classical Twin Design.<sup>13</sup> This model can contribute to elucidating the nature of the relationship between, for example, DNA methylation level at a CpG site and a complex trait or exposure, allowing to test the role of genetic pleiotropy and the role of causation and its direction.

---

## 2 Combining twin-family designs with measured genetic variants to study the causes of epigenetic variation

In the first part of this chapter, we discuss models that focus on the question: what are the causes of variation in epigenetic marks between individuals, and more specifically what part of variation is due to genetics and what part is due nongenetic causes? We will illustrate the application of these models by discussing a study conducted by ourselves and colleagues<sup>12</sup> in the Netherlands Twin Register.<sup>14</sup> In this study, we analyzed data from a large cohort of twins and family members for whom DNA methylation was measured across the genome with the Illumina 450k array in whole blood. The data were used to estimate the overall contribution of genetic and environmental effects to variation in DNA methylation level at individual sites, the variance explained by genome-wide SNPs, and to examine interaction of genetic and environmental influences with age and sex. The results of this study are described in a catalog that is available at <http://bbmri.researchlumc.nl/atlas/>.

### 2.1 SNP heritability

Classical twin and family designs allow to estimate the proportion of variation in a trait that is explained by all genetic variants in a population (i.e., the heritability;  $h^2$ ), because the genetic relations among biological relatives are known: parents and offspring share 50% of their genes identical by descent (IBD), siblings and DZ twins 50% on average, and MZ twins share all their genetic material. When resemblance for observed traits reflects the differential genetic resemblance, this supports genetic influences on these traits, which can be estimated for example by genetic structural equation modeling.<sup>15,16</sup>

Methods that incorporate measured genotypes allow to estimate the proportion of variation in a trait that is explained by a given set of genetic variants, usually a large set of SNPs that are commonly

genotyped or imputed and which have a minor allele frequency above a certain threshold. Any two individuals will differ in the extent to which they share a copy of a genetic variant and measured genetic relationships among “unrelated” individuals form the basis to estimate the “SNP heritability” of traits ( $h_{SNPs}^2$ ), that is, the variance in a trait that is explained by the additive effects of all measured or imputed genome-wide SNPs together.<sup>17</sup> This estimate of  $h_{SNPs}^2$  is not expected to equal  $h^2$  of the same trait because  $h_{SNPs}^2$  only includes the effects variants that are tagged by genotyped or imputed SNPs. On the other hand,  $h^2$  captures the effects of all variants, common and rare. The difference between  $h_{SNPs}^2$  and  $h^2$  can provide insight into the contribution of common variants to a trait of interest.

The method consists of two main steps. In the first step, a so-called genetic relationship matrix (GRM) is constructed based on genome-wide SNP data. The GRM summarizes the genetic similarity between all individuals in the dataset. The GRM of (presumed) unrelated people contains the very small genetic relationships between all pairs of individuals. Each value in the matrix represents (a standardized measure of) the proportion of alleles that is shared identity-by-state (IBS) between a pair of individuals. The approach is based on the idea that all people are distantly related to each other because we all share common ancestors in the distant past. In the second step, the phenotypic resemblance of individuals is modeled as a function of their genetic relatedness in order to estimate the contribution of the set of input SNPs that was used to construct the GRM to the variation in the phenotype of interest. When genetically more similar individuals are also phenotypically more similar, the SNP heritability will be greater than zero. The principle is similar to classical twin/family designs, in which the heritability ( $h^2$ ) is estimated by modeling (observed) phenotypic similarity as a function of the known genetic similarity of twins (or other relatives).

One frequently used software program for performing this type of analysis is Genome-wide Complex Trait Analysis (GCTA).<sup>18</sup> GCTA can calculate GRMs based on genome-wide SNP data from a group of individuals, and can estimate in a subsequent step the heritability with a method called genome-based restricted maximum likelihood (GREML).<sup>18,19</sup>

The approach was first introduced to the human genetics community (it has been in used in animal breeding for longer) in a paper by Jian Yang and colleagues in 2010, where it was applied to estimate the proportion of variation in human height that is explained by all genome-wide common SNPs<sup>20</sup> and later by multiple other papers that refined the method. The results illustrated that 45% of the variance in height is explained by SNPs. Two important observations followed from this: (1) the SNP heritability of height was much larger than the variance explained by (genome-wide significant) SNPs that had been identified in 2010 by GWAS (~5%). This finding revealed that height is influenced by many more common SNPs than the ones that had already been identified, but that the effects of those SNPs are too small to individually reach significance in GWAS and (2) the SNP heritability of height was smaller than the total heritability of height (which is about 80%), which suggested that height is also influenced by genetic variants that are not or imperfectly tagged by the SNPs tested in GWAS (common SNPs).

## 2.2 Estimating total heritability based on measured genetic relationships

The estimation of  $h^2$  with twin-family data does not require measured genetic information (i.e., genotyped DNA); however, if such information is available, one can calculate the empirical amount of genetic material that family members share. Although the proportion of sharing is fixed for parent–offspring pairs (because a parent passes exactly half of his/her chromosomes to a child), and for MZ twins (who always inherit 100% of the same genetic material from their parents, because they are derived from a single zygote), the proportion of genetic material that is shared by DZ twins and siblings varies. Because

of the random process of segregation, some sib-pairs will have inherited more than 50% of their DNA from the same parent, and other pairs less than 50%, with the population average 50%.

Measured data on genome-wide genetic variants have been used to address the assumption of the classical twin model that the proportion of genetic material that DZ twins have inherited from the same parent (IBD sharing) is on average 0.5 (50%). Using genome-wide microsatellite marker data, Peter Visscher and colleagues demonstrated that DZ twins and siblings indeed on average share about 50% IBD.<sup>21,22</sup> They also showed that most (95%) DZ twins and siblings share between 42% and 58% of their genome IBD.

In addition to verifying assumptions of the classical twin design, measured IBD information can also be integrated with the classical twin design. Rather than specifying the theoretical value for the correlation between additive genetic values (e.g., 0.5 for all DZ twins), one can plug in the actual (measured) genetic relationships of all twins in the model instead. Similar to classical twin analyses, these analyses can be used to estimate  $h^2$ , which will capture the contribution of all genetic variants (common and rare). Using measured IBD sharing of sib-pairs instead of the fixed value of 0.5, the heritability of height was estimated at 0.86, which is highly consistent with results from classical twin studies.<sup>22</sup> The approach can also be used to estimate the heritability due to all variants (common and rare) within a particular genomic region. For example, using SNP data, we calculated IBD sharing of DZ twins and sib-pairs for variants in the *IL6R* gene on chromosome 1, which indicated that additive effects of all variants segregating in this region explain 69% of the variation in soluble interleukin-6 receptor level.<sup>23</sup> This estimate was very close to the total heritability of this trait and suggests that the genetic variance of soluble interleukin-6 receptor level is almost entirely explained by genetic variation at this locus.

With genome-wide SNP data that are typically used for GWAS and are nowadays widely available in many twin cohorts, genetic relationships can be calculated with programs such as GCTA between any pair of individuals, both unrelated people and relatives.

### 2.3 Estimating total and SNP heritability simultaneously

In cohorts that include both closely and distantly related individuals, the total heritability ( $h^2$ ) and SNP heritability ( $h^2_{SNPs}$ ) can be estimated simultaneously. The approach, introduced by Noah Zaitlen and colleagues,<sup>24</sup> makes use of two GRMs: one GRM that contains the relationships between all individuals in a sample and a second GRM for the same individuals in which all genetic relationships smaller than 0.05 (distant genetic relationships) are set to zero. This makes the estimates of genetic relatedness in the second GRM equivalent to the proportion in the genome shared identity-by-descent (IBD). In essence, the covariance between individuals for a trait of interest, for instance DNA methylation level at a CpG site, is modeled as a function of the (very small) genetic covariance between unrelated individuals and the (larger) genetic covariance between relatives (Box 1). The first allows to estimate the variance explained by SNPs, and the second allows to estimate the variance explained by all other genetic variants. The sum of these two gives the total genetic variance.

### 2.4 Application to DNA methylation data from blood

We will illustrate an application to DNA methylation by discussing a study from the Netherlands Twin Register conducted by ourselves and colleagues.<sup>12</sup> In 2603 adult individuals (mean age 37.2, s.d.=13.3, 66% females), genome-wide DNA methylation was measured with the Illumina 450k array

**Box 1 Equations**

For each individual CpG (CpGi), the expected covariance was modeled as a function of two GRMs, the additive genetic variance ( $\sigma_{IBD}^2$ ), and the variance explained by genome-wide SNPs ( $\sigma_{SNPs}^2$ ) as follows:

$$\text{cov}(\text{CpGi})_{n \times n} = \text{GRM}_{n \times n}^{\text{IBS}} \otimes \sigma_{SNPs}^2 + \text{GRM}_{n \times n}^{\text{IBS} > 0.05} \otimes (\sigma_{IBD}^2 - \sigma_{SNPs}^2) + \mathbf{I}_{n \times n} \otimes \sigma_e^2,$$

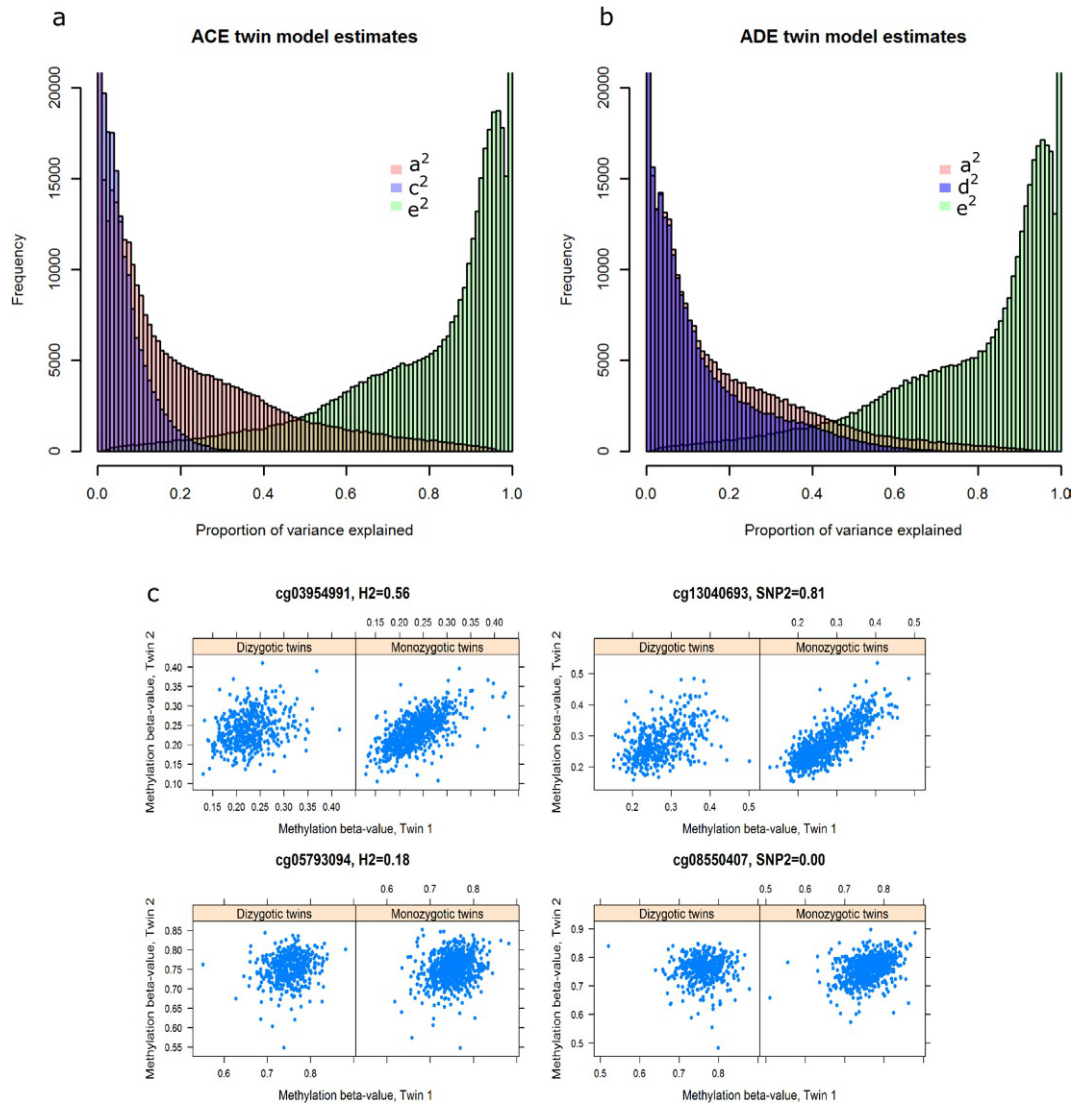
where  $\text{cov}(\text{CpGi})_{n \times n}$  is the expected covariance of DNA methylation at CpGi between individuals, adjusted for covariates,  $\text{GRM}_{n \times n}^{\text{IBS}}$  is the GRM describing the relationships between all individuals,  $\text{GRM}_{n \times n}^{\text{IBS} > 0.05}$  is the GRM in which all genetic relationships  $< 0.05$  IBS (distant genetic relationships) are set to zero making the estimates of genetic relatedness equivalent to the proportion in the genome shared identity-by-descent (IBD),  $\sigma_{SNPs}^2$  is the variance explained by all SNPs, the term  $(\sigma_{IBD}^2 - \sigma_{SNPs}^2)$  denotes the difference between the total genetic variance and the variance explained by SNPs,  $\mathbf{I}_{n \times n}$  is an identity matrix, which always have ones on the main diagonal and zeros elsewhere, and  $\sigma_e^2$  reflects the variance attributable to residual effects (“unique environment,” which may include environmental influences unique to each individual, stochastic influences and measurement error). The symbol  $\otimes$  is used to denote multiplication of matrices. The total heritability ( $h^2$ ) was calculated as:  $h^2 = \sigma_{IBD}^2 / (\sigma_{IBD}^2 + \sigma_e^2)$ . The proportion of variance explained by genome-wide SNPs was calculated as:  $h_{SNPs}^2 = \sigma_{SNPs}^2 / (\sigma_{IBD}^2 + \sigma_e^2)$  and the proportion of the heritability explained by SNPs was calculated as:  $h_{SNPs}^2 / h_{IBD}^2$ .  $n \times n$  denotes the dimension of each matrix (where  $n$  = the number of individuals).

and genome-wide SNP data were generated with the Affymetrix6 array (or extracted from whole-genome sequence data for a small subset of subjects). The study included predominantly twins (769 MZ and 424 DZ twin pairs), and a small group of their parents, siblings, and spouses.

DNA was extracted from whole blood, and DNA methylation level was measured following bisulfite treatment of the DNA at more than 450,000 sites (mostly CpG sites) across the genome.<sup>25</sup> In DNA that is derived from a mixture of cells (such as whole blood), the methylation level represents a continuous variable with values that may range between 0 and 1. For example, a methylation level of 1 means that all DNA strands had a methyl group attached at this position and a value of 0.5 means that 50% of all DNA strands had a methyl group attached at this position. The latter may occur when the location is methylated on one chromosome, as happens at many positions on one of the two copies of each X-chromosome in women, or when the position is methylated in one cell type, or even a subset of cells from the same cell type, from which the DNA was extracted. The methylation data were adjusted for several commonly used covariates, including technical batch information, age, sex, and white blood cell counts. The analyses focused on autosomal methylation sites and genotyped, common (minor allele frequency (MAF) $>0.01$ ), autosomal SNPs measured by the Affymetrix6 genotyping array.

**2.4.1 Classical twin model**

We first computed classical twin correlations and fitted models, which estimated the variance explained by additive genetic effects (A), nonadditive genetic effects (D), common environment (C), and unique environmental effects (E) at individual methylation sites (Fig. 1A and B). The pattern of the MZ and DZ twin correlations suggested that the resemblance of twins for DNA methylation level is mainly due to additive genetic influences; the average correlation in MZ twins was 0.20; approximately twice as large as the average correlation in DZ twins ( $r=0.09$ ). Full models that allowed for resemblance due to sharing of the environment showed that so-called common environmental effects (C) shared by twins explained relatively little variation across the adult peripheral blood methylome (on average 3%, s.d.=5%), with significant C effects at 185 methylation sites. Models that included genetic dominance (D: nonadditive genetics) showed that these effects explained on average 8% of variance in DNA methylation (s.d.=12%), with significant effects at 241 methylation sites. Additive genetic influences had a larger



**FIG. 1**

Genetic and environmental influences on genome-wide DNA methylation levels.

Estimates from classical twin modeling in 769 and 424 DZ twin pairs. (A) Histograms of genome-wide ACE model estimates: variance explained by additive genetic effects ( $a^2$ , red), common environmental effects ( $c^2$ , purple) and unique environmental effects ( $e^2$ , green). (B) Histograms of genome-wide ADE model estimates: variance explained by additive genetic effects ( $a^2$ , red), nonadditive genetic effects ( $d^2$ , purple) and unique environmental effects ( $e^2$ , green). Y-axes are truncated. (C) Scatterplots of DNA methylation levels in MZ and DZ twin pairs at four exemplary CpG sites. For illustrative purposes, methylation beta values (which represent methylation proportion) obtained after normalization are plotted in this figure, whereas all analyses were performed on normalized methylation M values, corrected for a number of covariates (such as age, sex, technical batch information, and white blood cell counts). Four examples of CpG sites were selected from the most variable CpG sites with high heritability ( $h^2=0.56$ ; upper left), high SNP heritability ( $h^2_{SNPs}=0.81$ ; upper right), low heritability ( $h^2=0.18$ ; lower left) and low SNP heritability ( $h^2_{SNPs}=0.00$ , lower right).

Reproduced with permission from van Dongen, J. et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. Nat. Commun. 2016;7:11115.

impact on the methylome, accounting for 20% of the variance (s.d.=21%) on average across all methylation sites. Unique environmental influences, which can include measurement error, explained most of the variation in genome-wide methylation levels, with a genome-wide average of 80% (s.d.=21%). The heritability of individual methylation sites across the genome showed large variation (Fig. 1A and B).

Fig. 1C shows an example of the DNA methylation levels in monozygotic and dizygotic twin pairs at four exemplary CpGs. Each panel shows the DNA methylation level at one CpG in twin pairs (each dot represents one twin pair), with dizygotic twins in the left panel and monozygotic twins in the right panel. The  $x$ -axis shows the methylation level in twin 1 and the  $y$ -axis shows the methylation level in his/her cotwin. A larger difference between the correlation in MZ twins and in DZ twins (stronger correlation in MZ twins) indicates a higher heritability ( $h^2$ ) of DNA methylation level at this particular location. The figure shows that for sites with a higher heritability, the values of MZ cotwins tend to be more similar to each other. The figure also illustrates that the resemblance of MZ and DZ twins does not tell how much of the variance is explained by genome-wide common SNPs. For instance, if a large proportion of the variance in DNA methylation is due to rare genetic variants that are not tagged by SNPs, MZ twins will resemble each other more than DZ twins, but the SNP heritability will be low because the SNP heritability does not include all effects of rare variants.

#### 2.4.2 Heritability based on measured genetic relationships

Next, we estimated heritability based on measured genetic relationships. We constructed a genetic relationship matrix (GRM) based on all common genotyped autosomal SNPs that are measured by the Affymetrix 6 genotype array with GCTA.<sup>18</sup> We applied the method of Zaitlen et al.<sup>24</sup> to simultaneously estimate  $h^2$  and  $h^2_{SNPs}$  in R as described in Nivard et al.<sup>26</sup> and van Dongen et al.<sup>12</sup> Nowadays, it is also possible to apply “the Zaitlen model” in GCTA. The genome-wide average heritability ( $h^2$ ) was 0.19, similar to the estimate based on the classical twin method (mean heritability 0.20), and the estimates of  $h^2$  obtained by the two methods were strongly correlated ( $r=0.99$ ). The strong correlation between these estimates was expected, because they both are estimates of total heritability.

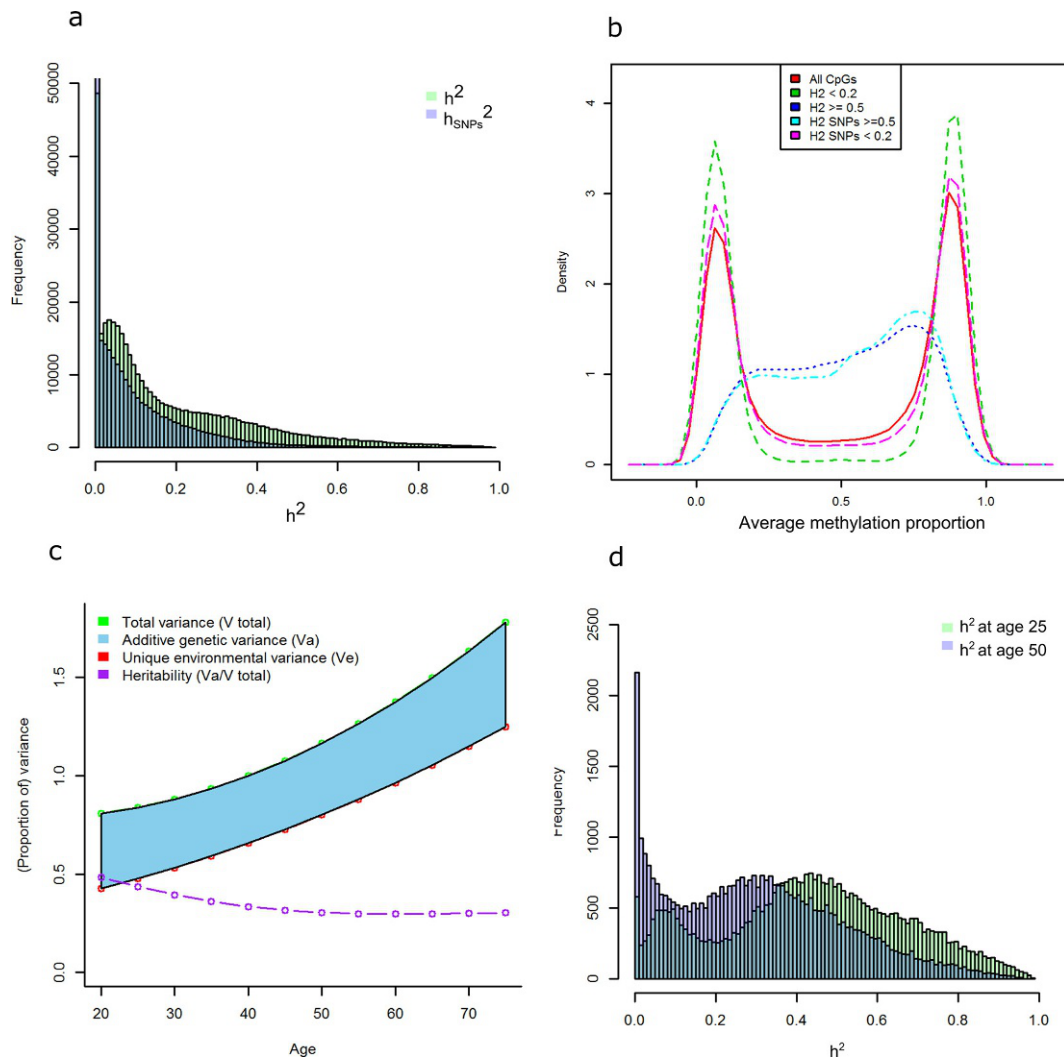
Across all sites, on average 7% (s.d.=12%) of the variance of DNA methylation was explained by common genetic variants in the genome ( $h^2_{SNPs}$ ; Fig. 2A). Although it was already known from mQTL studies that SNPs affect DNA methylation, it was unknown how much of the variance of DNA methylation levels in total is explained by common SNPs. On average across the genome, the proportion of the total heritability that can be explained by SNPs  $\left(\frac{h^2_{SNPs}}{h^2}\right)$  was 0.37, again, with large variation across individual methylation sites (s.d.=0.40). At many sites, a relatively large proportion of total estimated heritable variation in DNA methylation was explained by common genetic variants: at almost a fifth (18%) of all methylation sites, almost the entire total heritability can be explained by common SNPs. There are also many sites where DNA methylation shows heritable variation, but common genetic variants explain little of this genetic variance. For example, SNPs explained less than 1% of the total heritability at 39% of genome-wide DNA methylation sites.

While the overall distribution of DNA methylation level across the genome is bimodal, methylation sites with a large heritability are characterized by an intermediate methylation level (Fig. 2B).

### 2.5 Interaction with sex and age

EWA studies have detected many loci where methylation differences between males and females exist, which are not limited to the X-chromosomes. Similarly, EWA studies of age have detected widespread changes in DNA methylation level with age. The relationship between DNA methylation and age is so



**FIG. 2**

Contribution of SNPs to the variation in DNA methylation level and age-related changes.

(A) Histogram of total additive heritability ( $h^2$ ) of DNA methylation in blood (green) and variance explained by genome-wide SNPs ( $h^2_{SNPs}$ , purple) for genome-wide CpGs. The y-axis is truncated. (B) Density plot showing the distribution of the average methylation level in blood at methylation sites with a high vs low heritability. It shows that methylation sites that are to a large extent explained by genetic variants tend to have an intermediate methylation level. (C) Total variance of DNA methylation, additive genetic variance, unique environmental variance, and heritability plotted against age, based on estimates obtained in interaction models on data from 2603 individuals. (D) Histogram of heritability at age 25 (green) and at age 50 (purple) for 39,455 CpGs with significant interaction between age and genetic variance or between age and unique environmental variance. Dark blue denotes the overlap of green and purple bars.

Adapted with permission from van Dongen, J. et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat. Commun.* 2016;7:11115.

strong at certain loci, that this information has been used to build DNA methylation-based predictors of chronological and biological age, with so-called epigenetic clock algorithms.<sup>27</sup> In 2005 Fraga and colleagues reported for the first time that the epigenetic profiles of older MZ twins are less similar than those of young MZ twins based on a comparison of global and locus-specific DNA methylation and histone acetylation.<sup>28</sup>

While the majority of EWA studies of age have focused on changes in the mean methylation level with age, some have tested changes in the variance of DNA methylation with age. For example, with genome-wide methylation data from blood, Roderick Sliker and colleagues detected many loci where the variance of DNA methylation level correlates positively with age.<sup>29</sup> This means that the differences between individuals in a population become larger as a function of age; older individuals show more variation in DNA methylation level at these loci. The cause of age-related changes in variance can be genetic (for instance, as people age, changes in the expression of genes may cause an increase in genetic variance), environmental (as people become older, effects of the environment on DNA methylation might accumulate, leading to greater variation in older people), or stochastic (possibly, stochastic effects are amplified at a higher age due to a loss of proper maintenance of the epigenome as people age).<sup>30</sup> With twin-family data, the causes of age-related changes in variance (genetic and environmental or stochastic) can be examined. This can be done by adding a moderator variable to the classical twin model<sup>31</sup> to model the interaction between age and the genetic and environmental effects. Similarly, moderation by sex can be modeled to examine which methylation sites show sex differences in the genetic or environmental variance. Such interaction models give age-specific and sex-specific estimates of heritability, respectively.

In the DNA methylation dataset from the NTR, we applied models that included age- or sex-interaction terms. In our study, the interaction term was added to the model based on measured genetic relationships between relatives, according to the methods described by Nivard et al.,<sup>26</sup> but we note that these interaction terms can also be added to classical twin models.<sup>31</sup>

### **2.5.1 Sex interaction results**

If we look at all genome-wide methylation sites, the average heritability of DNA methylation is almost the same in males (mean  $h^2=0.199$ ) and females (mean  $h^2=0.198$ ); however, 2667 methylation sites (0.7% of all measured sites) showed a significant interaction effect of sex and genetic or environmental effects on DNA methylation level. At 59% of these sites, the heritability was lower in women. Importantly, if sex differences in heritability are observed, this can be due to a difference in the environmental variance or due to a difference in the genetic variance between the sexes. Both can also occur simultaneously. We found that at 76% of all sites with significant sex interaction, the unique environmental variance (rather than the additive genetic variance) differed between the sexes. At sites with a lower heritability in females, it was usually the case that the variance of DNA methylation due to environmental influences was larger in females. These sites are important candidates for sex-specific variation in epigenetic regulation, and occurred for example at loci connected to metabolism. Several CpGs whose methylation levels correlate with serum levels of the metabolites tryptophan, mannose, and 5-oxoproline displayed a significant sex difference in the environmental variance of DNA methylation level.

### **2.5.2 Age interaction results**

Age interaction effects were highly common: 39,455 sites (10.4% of all measured sites) showed a significant interaction effect of age and genetic or environmental effects on DNA methylation level. At 82% of these sites, the unique environmental variance changed with age. These sites typically showed

an increase in the unique environmental variance and total variance with age, and a decreasing heritability (Fig. 2C and D): At 90% of sites with significant age interaction, the heritability was lower at age 50 than at age 25, although the difference in heritability between younger and older people was usually modest. For example, a decrease in heritability with age was found at a CpG site located in the *TNIP1* gene (cg22178392). The methylation level of this CpG correlates with serum LDL cholesterol. The heritability of DNA methylation level at this CpG site in blood decreases from 54% at age 25 to 39% at age 50 due to an increase in the unique environmental variance with increasing age.

The results from this large cross-sectional study are consistent with the observations from small longitudinal studies of MZ twin pairs that the epigenetic profiles of MZ twins diverge with age.<sup>28,32</sup>

### 2.5.3 Integration with findings from EWASs of diseases, traits, and exposures

Since the publication of this twin-family study in 2016, a large number of EWASs of diseases, traits, and exposures have been conducted. The findings of these studies now allow us to better interpret the implications of methylation sites with age- and sex-specific variance and the possible causes of differences in environmental variance. The EWAS atlas (<http://bigd.big.ac.cn/ewas/index><sup>6</sup>) includes all published associations between DNA methylation (Illumina array) identified in EWASs of traits, diseases, and exposures. We used the EWAS atlas to examine the overlap of methylation sites detected in 333 EWASs of diseases and traits on January 14, 2019 and our list of 2034 methylation sites with a sex difference in environmental variance and 32,234 methylation sites with age-dependent environmental variance.

Sites with a sex difference in environmental variance showed significant enrichment (an overlap that is greater than expected by chance) for 15 traits. The top enriched traits are allergy- and immune system-related. This could mean that males and females experience differences in the frequency of exposure to environmental influences that shape epigenetic regulation related to allergy and the immune system, or it could be the case that males and females are equally exposed to such influences but that their epigenomes do not respond to this exposure in the same way. Methylation sites with sex-specific environmental variance identified by our twin-family study could be interesting candidates for the development of sex-specific biomarkers for immune system-related conditions.

Sites with age-dependent environmental variance showed significant enrichment for 58 traits. The top enriched traits are related to aging, rare genetic disorders, cancer, and smoking. The enrichment for smoking-associated sites illustrates that accumulating effects of exposure to cigarette smoke might be responsible for increases in DNA methylation variance with age. The overlap with cancer-associated sites suggests that methylation changes induced by environmental influences throughout the lifespan or amplification of stochastic epigenetic variation with age might be linked to the increased risk of cancer as people age.

## 2.6 Conclusions

The findings of this twin-family study can be summarized as follows:

- The estimates of heritability of DNA methylation level obtained in the classical twin model are highly similar to estimates of heritability obtained in models based on measured genetic relationships of twins.
- The heritability of DNA methylation, which had previously only been estimated based on classical twin models, can be explained partly by SNPs. This confirms that SNPs contribute to the heritability of DNA methylation.

- Averaged over all genome-wide methylation sites, common SNPs explain 37% of the total heritability of DNA methylation. The remaining heritability might be explained by other genetic variants that were not tagged by the genotyping array.
- The proportion of the heritability of DNA methylation level that can be explained by common SNPs varies widely across genome-wide methylation sites.
- The SNP heritability gives an estimate of the variation in DNA methylation at a CpG site that can ultimately be explained by mQTL SNPs once all mQTLs for this site have been identified.
- On the genome-wide level, environmental or stochastic influences are a more important determinant of sex-specific and age-specific methylation variation between individuals than genetic influences.
- Twin-family studies of epigenetic marks can provide valuable information on locations in the genome where epigenetic variation between people may reflect disease-relevant environmental exposures or genetic variation, and whether this variation is sex- or age-specific.

## 2.7 Further applications

We have discussed an application to whole blood data from adult twins with DNA methylation measured by the Illumina 450k array. Future studies on other tissues, data from other age groups including children, and higher resolution DNA methylation data (such as the Illumina EPIC array and methylation sequencing data) or data on other epigenetic marks will allow to examine the contribution of genome-wide SNPs to DNA methylation variation in other tissues, to examine the causes of variation in children, and to examine the causes of variation in DNA methylation and other epigenetic marks that are not captured by the 450k array.

In addition to creating a GRM based on all genome-wide SNPs, it is also possible to estimate the variance that is explained by particular subsets of genome-wide SNPs by creating multiple GRMs; for instance, separate GRMs based on SNPs on each chromosome,<sup>23,33</sup> SNPs in *cis* vs SNPs in *trans*,<sup>34</sup> or genome-wide significant SNPs identified in GWAS (or mQTL analysis, in the case of DNA methylation data) vs all other SNPs that remain to be detected.<sup>35</sup>

It is also possible to estimate the variance explained by all SNPs based on whole-genome sequencing data.<sup>36</sup> This can provide insight into the contribution of rare variants to the heritability of DNA methylation levels. Application to height and BMI showed that the entire heritability of height and BMI can be captured by all variants detected with whole-genome sequencing data.<sup>37</sup> For DNA methylation, it remains to be established whether differences between the total heritability and SNP heritability are explained by rare variants.

---

## 3 Combining twin-family designs with measured genetic variants to study cause and effect

Often findings from EWAS are interpreted as reflecting causal effects of methylation on a trait of interest (e.g., methylation causes a disease) or of an exposure of interest on methylation (e.g., DNA methylation is a consequence of smoking behavior). However, methylation differences observed in, for instance, an EWAS of smoking, may also reflect effects of other exposures that correlate with smoking (e.g., cannabis or alcohol use), or they may reflect the genetic predisposition to smoking. For instance,

one of the most strongly associated genetic variants for nicotine dependence is located in the DNA methyltransferase gene *DNMT3B*.<sup>38</sup> This strongly implies a role for DNA methylation in nicotine addiction but might also cause genome-wide differences in DNA methylation between people with and people without a genetic predisposition for nicotine dependence, regardless of their smoking behavior. Alternatively, it is important to note that while methylation levels at some loci revert to those observed in nonsmokers in the first year after smoking cessation,<sup>39</sup> others remain differentially methylated up to ~35 years after quitting smoking.<sup>40</sup> Intriguingly, and in line with these observations, the risk of relapse to smoking is greatest in the first year (above 50%), and decreases as a function of time to a low of 10% after 30 years of abstinence.<sup>41</sup>

Understanding the nature of these associations might provide insights crucial for efficient prediction, prevention, and treatment of common complex diseases, and requires one to distinguish between these alternative hypotheses. Toward this aim, one is invariably confronted with questions about the causal mechanisms that generated the observed association.

### 3.1 Mendelian Randomization

In studying causality one typically relies on randomized controlled trials, but these may not be feasible, or fail to adhere to ethical standards. For example, ethical constraints preclude assigning individuals to a “heavy smoking experimental condition” (the exposure variable) in order to test the causal effects of heavy smoking on DNA methylation (the outcome variable). To resolve such questions one must rely on observational studies. Mendelian Randomization (MR) represents an alternative to the randomized controlled trials.

MR uses genetic variants as “instrumental variables” to detect the causal effect of the exposure on disease/outcome in nonexperimental settings. An “instrument” is a variable that (a) associates directly with the exposure of interest, (b) is independent of confounders (other factors affecting the outcome), and (c) does not directly affect the outcome (Fig. 3). MR is a popular method for establishing causality in observational studies, as demonstrated by the large number of substantive applications over the last years. In the field of epigenetics, MR has been applied to DNA methylation data to examine, for instance, the causal relationship between DNA methylation and lipid levels,<sup>8</sup> and between DNA methylation and BMI.<sup>42</sup>

MR exploits the fact that genetic variants are randomly assigned at meiosis, are largely distributed in the population independent of environmental confounding, and are not modified by disease outcomes. Correspondingly, MR might be seen as “Nature’s randomized controlled trial.” For example, analogous to a randomized controlled trial testing the efficacy of statin dose levels in reducing LDL-Cholesterol and the risk of cardiovascular disease, in “Nature’s randomized controlled trial” (MR) “dose levels” in the form of loss of function variants in the *PCSK9* gene are allocated randomly at conception. Like statins, these mutations decrease LDL-cholesterol level and the risk of cardiovascular disease.<sup>44–47</sup>

In regard to the assumption that causality flows unidirectionally (Fig. 3) one might obviously argue that bidirectional relationships between complex traits are also plausible, and there are MR approaches to test such bidirectional hypotheses. Ideally, an accurate description of bidirectional relationships involves estimating simultaneously reciprocal causal paths. In the literature, “bidirectional MR” typically involves two unidirectional tests employed sequentially: Model 1 representing hypothesis 1: X causes Y (parameter  $g_1$  estimated), and Model 2 representing hypothesis 2: Y causes X (parameter  $g_2$  estimated). In Fig. 3, bidirectionality involves reciprocal causal paths between exposure and outcome

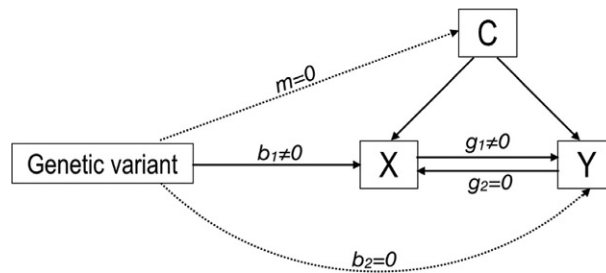


FIG. 3

Diagrammatic representation of the Mendelian randomization design. Variable X—the exposure variable—is associated with variable Y—the outcome variable. In the context of epigenetic studies, X may represent, for example, smoking behavior, and Y may represent DNA methylation at a particular CpG site. The association between X and Y might arise due to: (i) confounders C (risk factors affecting both X and Y), (ii) a direct causal effect of X on Y (parameter  $g_1 \neq 0$ ), (iii) a causal effect of Y on X (parameter  $g_2 \neq 0$ ). One may probe these hypotheses using an instrumental variable. An instrumental variable associates robustly with the exposure and is independent of the outcome and possible confounders (conditional on exposure). Mendelian Randomization employs genetic variants as instrumental variables to test causal hypotheses. The MR design yields information about causality under the assumptions that the genetic variant employed as an instrumental variable is: (i) associated with the exposure, i.e.,  $b_1 \neq 0$ , (ii) independent of confounders, i.e.,  $m = 0$ , (iii) its effect on the outcome is fully mediated by the exposure variable, i.e., there is no direct effect of the instrumental variable on the outcome or *there is no pleiotropy*, parameter  $b_2 = 0$ , and (iv) causation flows unidirectionally from exposure to outcome, parameter  $g_2 = 0$  (henceforth the MR approach based on these standard assumptions is called the *standard MR*).

(i.e., parameters  $g_1$  and  $g_2$ ). These causal paths are mutually confounded in MR, allowing only one to be estimated (the bidirectional model is not identified in cross-sectional data; in other words, cross-sectional data does not provide sufficient information to yield unique estimates of these parameters). To circumvent this identification problem, it is routinely assumed that either  $g_1$  or  $g_2$  equal zero.

The interest in MR has increased recently particularly due to the DNA genotyping costs that have rapidly tumbled down over the last decade, making available large samples of genotyped individuals including twins who participate in twin registers worldwide. In addition, large Genome-Wide Association Studies (GWASs) have identified an increasing number of genetic variants robustly associated with exposures, yielding strong instrumental variables to be employed in MR studies. Furthermore, confounding in observational association is a prominent problem in the social sciences, which can be addressed using MR. Notwithstanding MR's potential, the problem remains that the standard procedure involves strong assumptions (see Fig. 3) which are not testable currently; these assumptions are unlikely to hold for complex traits. Following we describe how one of these strong assumptions—the “no pleiotropy” assumption—can be tested and relaxed by combining twin-family designs with measured variants to study cause and effect.

### 3.2 Mendelian Randomization meets the classical twin design

Behavioral genetics models for twin resemblance have been with us for decades, but little attention has been given to the tremendous potential of integrating MR with the classical twin design. Twins share

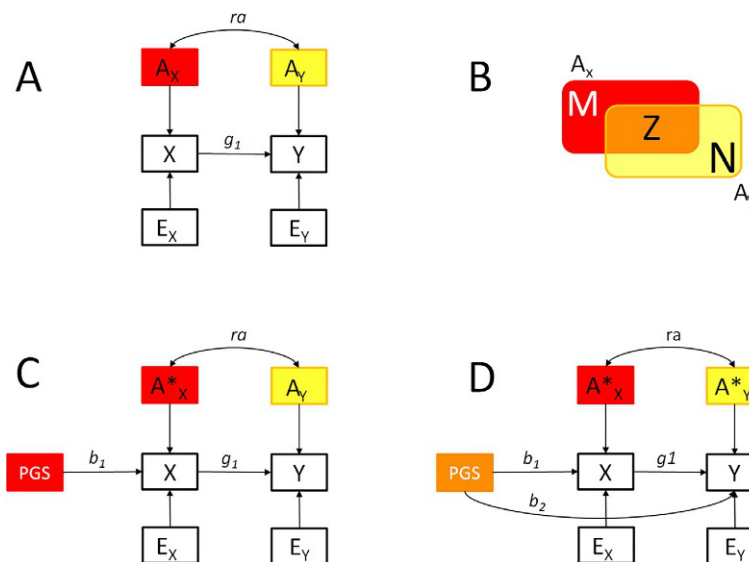


FIG. 4

A closer look at the “no pleiotropy” assumption.

(A) X directly influences Y (parameter  $g_1$ ). In addition, the additive genetic variables  $A_X$  and  $A_Y$  are correlated (parameter  $ra$ ). (B) The set of loci  $M$  underlies the variance of  $A_X$ , but does not contribute directly to the variance of  $A_Y$ , i.e., in set theory notation,  $M = A_X \setminus A_Y$  (yet note that if X causally influences Y separate powered GWASs of the two traits will associate variants from set  $M$  with both X and Y<sup>43</sup>). Likewise,  $N = A_Y \setminus A_X$ , i.e., the set of loci  $N$  contributing to the variance of  $A_Y$  but not to the variance of  $A_X$ .  $Z$  represents the intersection of  $A_X$  and  $A_Y$ , that is, the set of loci  $Z$  underlies both  $A_X$  and  $A_Y$ , i.e.,  $Z = A_Y \cap A_X$ . Note that the set  $Z$  may contain pleiotropic loci, where the pleiotropy is due to direct effects or due to linkage disequilibrium; (C) The MR model with a polygenic instrument (polygenic score; PGS) and “no pleiotropy.” PGS is associated with X (parameter  $b_1$ ), but is assumed to have no direct influence on the outcome Y. This model holds only if the instrument PGS is constructed on the basis of a subset of variants from set  $M$ . In the presence of PGS,  $A^*_X$  is a residual (in the regression of X on PGS). (D) MR with pleiotropic genetic instrument. In this model, the PGS is constructed on the basis of a sample of genetic variants taken from set  $Z$ . The parameter  $b_2$  accommodates the fact that the set of variants used to construct PGS underlies the variance of both  $A_X$  and  $A_Y$ . The “no pleiotropy” assumption implies  $b_2 = 0$ .

Adapted with permission from Minică, C.C., Dolan, C.V., Boomsma, D.I., et al. Extending causality tests with genetic instruments: an integration of mendelian randomization with the classical twin design. *Behav. Genet.* 2018. doi:10.1007/s10519-018-9904-4.

both genetic and environmental factors, therefore they provide the information needed to estimate the causal effect and the contribution of genetic and environmental factors to the exposure–outcome relationship. We embedded MR with twin information to develop the means to test causal hypotheses when the genetic variants used as instrumental variables have pleiotropic effects<sup>13</sup> (i.e., when assumption (iii) in Fig. 3 is violated), meaning that the genetic variant(s) used as instrument does not only affect the exposure, but also has a direct effect on the outcome (Box 2).

### Box 2 The “no pleiotropy” assumption

The evidence so far suggests that the effects of individual genetic variants on complex human traits are small.<sup>47</sup> This raises concerns about the causal effect obtained in finite samples using the genetic variants as instrumental variables, as this might be biased—the so-called weak instrument issue. Specifically, when the association between the instrumental variable (IV) and the exposure variable is weak (i.e., the IV explains a small percentage of variance in the exposure), the IV causal effect estimate will be biased toward the observed association. Therefore using single genetic variants as instruments might render MR liable to weak instrument bias.<sup>48</sup> Adding the weak effects of multiple genetic variants into a polygenic score (PGS) is a way to increase the power of the genetic instrument.<sup>49–52</sup> However, the “no pleiotropy” assumption is likely violated, particularly when the instrument is a polygenic score (see Fig. 4 for a closer look at the “no pleiotropy” assumption—reproduced with permission from Behavior Genetics, Extending causality tests with genetic instruments: An integration of Mendelian Randomization with the Classical Twin Design).

The absence of pleiotropy assumption is one of the standard MR assumptions unlikely to be satisfied due to the pervasive pleiotropy across the human genome.<sup>43,53,54</sup> Indeed, there are multiple examples of pleiotropic genetic variants in the GWAS literature (i.e., genetic variants that associate with multiple complex traits/risk factors). For example, *FTO* gene variants were associated with Body Mass Index,<sup>55–59</sup> obesity,<sup>56,60,61</sup> age at menarche,<sup>62,63</sup> breast cancer,<sup>64,65</sup> HDL-cholesterol,<sup>66</sup> Type-2-Diabetes.<sup>67–72</sup>

A genetic variant associated directly with both the exposure and the outcome variables of interest (i.e., pleiotropic) is an invalid instrumental variable. Standard MR assumes pleiotropy to be absent (see Fig. 3). Violation of this assumption may bias the causal effect estimate or even result in spurious causal effects.

Only when the “no pleiotropy” assumption holds does standard MR yield a correct estimate of the causal effect. In other words, when the instrument has pleiotropic effects (i.e., one of the IV assumptions is violated, see Fig. 3), standard MR biases the causal effect estimate, or even detects causal effects when there are none. In an effort to avoid such spurious results, we developed the MR-Direction of causation model (MR-DoC, Fig. 5) by incorporating data collected from relatives, such as twins.

Twins provide additional within-family data that can be used to examine whether the genetic variant employed as IV has a direct (pleiotropic) effect on the outcome variable (parameter  $b_2$ , Fig. 5). By explicitly modeling this pleiotropy path, MR-DoC correctly detects causation even when standard MR’s “no pleiotropy” assumption is violated (see table 2 in Ref. <sup>13</sup> for results of a simulation study). That is, twins provide the critical advantage of allowing one to estimate the causal parameter (parameter  $g_1$  in Fig. 5) and pleiotropic effects (parameters  $b_1$  and  $b_2$  in Fig. 5), conferring crucial power advantages (as we previously showed is also the case when twins are included in the context of GWAS<sup>73,74</sup>). Specifically, power analyses showed that MR-DoC has greater statistical power than a standard MR analysis, especially if the unshared environmental effects on the exposure  $X$  and the outcome  $Y$  are uncorrelated (i.e., parameter  $re$  in Fig. 5 is zero). The power of MR-DoC increases with: (1) decreasing residual  $X$ – $Y$  correlation ( $re$ ), and (2) with decreasing additive genetic variance of the exposure variable (see fig. 6 in Ref. <sup>13</sup>).

### 3.3 Overall prospects and potential limitations of the MR methods

Currently there is great interest in developing MR methods that are robust to assumption violation. While the first two standard MR assumptions in Fig. 3 (i.e., robust association with the exposure and independence of confounders) are likely to hold (as genetic variants are randomly assigned at meiosis, are largely distributed in the population independent of environmental confounding, and the association of many genetic variants with exposures/risk factors has been demonstrated and replicated in large



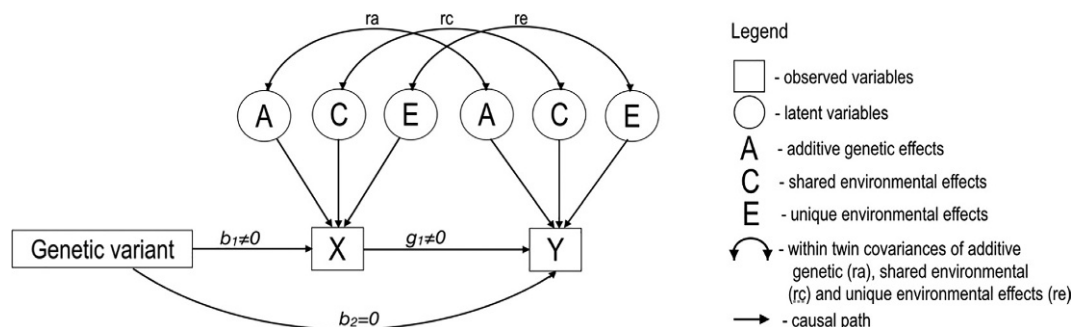


FIG. 5

The Mendelian randomization-direction of causation model (MR-DoC; to avoid clutter we show the model in twin 1 only, but see fig. 4 in Ref. <sup>13</sup> for details).

Twin data allow one to estimate the causal effect of the exposure variable  $X$  on the outcome variable  $Y$  (parameter  $g_1$ ), the pleiotropic effects of the genetic variant employed as an instrumental variable (parameters  $b_1$  and  $b_2$ ), and the contribution of genetic (A), shared (C), and unique (E) environmental factors to the  $X$ – $Y$  relationship.

GWASs), a particular attention has been given to the “absence of pleiotropy” assumption discussed above (see e.g., Refs. <sup>13,75–77</sup>). This assumption is unlikely to hold on the account of the ubiquity of pleiotropy across the human genome. More recently, the focus has shifted also to other important assumptions specific to studies that employ genetic variants as instrumental variables, like e.g., dynastic effects (these effects occur because the parents transmit to their children not only genotypes, i.e., affecting the offspring phenotype directly, but the parental genotypes also have an indirect effect on the children’s phenotype via the environments they provide) and nonrandom mating (occurs when husband–wife pairs are selecting each other based on phenotypic similarity, and this results in parental genotypes becoming correlated<sup>78–82</sup>). It is important to note that MR models that utilize family data (like MR-DoC) generally yield robust results when such assumptions are violated.

A potential concern related to the method described before surrounds the fact that the MR-DoC model requires access to individual-level data (anonymized genetic and phenotypic data collected in individuals, data more difficult to share than summary statistics, due to ethical constraints). Yet, current approaches based on summary statistics cannot resolve the “no pleiotropy” and the “non-random mating” assumptions simultaneously.<sup>78</sup> These assumptions can be tested, and possibly relaxed in the MR-direction of Causation approach we described here,<sup>13</sup> or in the approach proposed by Hartwig and colleagues<sup>78</sup> that also exploits data collected in families (trios). Indeed, family-based studies are due for a renaissance, as individual-level data collected from relatives combined with MR provide additional within-family data that partially resolve these strong assumptions and hence, strengthen causal inferences.<sup>80</sup>

Another possible concern may relate to the adoption of these twin-based MR methods, given the requirement for twin-family datasets. However, there are large-scale twin data registries (e.g., Netherlands, Scandinavia, UK, Australia, USA; see also<sup>83</sup>) and new twin registries are being added.

These large twin registries are often at the forefront of method development and empirical genome-wide and epigenome-wide association studies. Their data are typically representative of the general population, which guarantees generality of the results using twin and family data. That is, if one establishes a causal mechanism using twin data (e.g., salt intake leads to hypertension) others do not need to repeat this. Moreover, these MR models that use twin data (MR-DoC based) can be adapted and applied to sibling data, and importantly, to unrelated individuals. One of our future aims is to facilitate the use of our MR methods also in cohorts of unrelated individuals, by incorporating them in models for Genetic Relationship Matrices (GRMs).

### 3.4 Conclusion

We have shown how combining within-family designs with measured genetic variants might provide a means to test and relax one of the standard MR assumptions, and might strengthen causal inference in random (nonexperimental) samples. This combined approach can test and, if necessary, relax MR's strong "no pleiotropy" assumption, which the standard MR approaches cannot test empirically. The MR-DoC twin model is appealing as it allows one to use, as instrumental variables, both single genetic variants and powerful polygenic scores, and estimates correctly the causal effect in the presence of pleiotropy (i.e., the violation of the standard MR's "no pleiotropy" assumption). This and other robust MR methods that utilize within-family information to test cause and effect will greatly increase the value of the wealth of existing data collected at twin registries and the rich epidemiological samples of unrelated individuals, as they correctly detect causation and estimate effect sizes even when assumptions like "absence of pleiotropy" or "random mating" may be untenable. It should be noted that inferring cause–effect from nonexperimental samples is very challenging and has been subjected to multiple methodological developments. It is important to emphasize that robust causal inference will require evidence obtained with multiple approaches, i.e., methodological triangulation,<sup>81</sup> coupled with "sound theoretical thinking, constant vigilance, and a thorough understanding of the potential limitations of the methods being used."<sup>82</sup>

---

## 4 Overall conclusion

In this chapter, we have described two types of designs that integrate twin data with measured genetic information. In the first part we showed how integrating genetic relationship matrices (GRMs) based on measured SNPs with twin data allows for more in-depth questions concerning the causes of genome-wide DNA methylation levels. In the second part we described an integration of MR with twin data (the MR-DoC model) to examine causal relationships in observational settings. When writing this chapter, MR-DoC has not yet been applied to epigenetic data, but we believe epigenetic data could become a valuable application of the model. MR-DoC could be applied to dissect the nature of the associations between DNA methylation and smoking behaviors, and between DNA methylation and disease outcomes. Many worldwide twin registers have collected both epigenomic data (such as genome-wide DNA methylation arrays) and genomic information (genome-wide SNP arrays) for the same participants. We believe that the integration of such data with twin/family designs will make a valuable contribution to ongoing EWAS and mQTL mapping efforts.

---

## Acknowledgments

We acknowledge funding from the National Institute on Drug Abuse grant DA049867, the Netherlands Organization for Scientific Research (NWO): Biobanking and Biomolecular Research Infrastructure (BBMRI-NL, NWO 184.033.111) and the BBRMI-NL-financed BIOS Consortium (NWO 184.021.007), the “Aggression in Children: Unraveling gene-environment interplay to inform Treatment and Intervention strategies” project (ACTION). ACTION received funding from the European Union Seventh Framework Program (FP7/2007-2013) under grant agreement no 602768; NWO Large Scale infrastructures, X-Omics (184.034.019), and the Royal Netherlands Academy of Science Professor Award (PAH/6635).

---

## References

1. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet*. 2018. [https://doi.org/10.1016/S0140-6736\(18\)31268-6](https://doi.org/10.1016/S0140-6736(18)31268-6).
2. Cavalli G, Heard E. Advances in epigenetics link genetics to the environment and disease. *Nature*. 2019. <https://doi.org/10.1038/s41586-019-1411-0>.
3. Greenberg MVC, Bourc’his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. 2019. <https://doi.org/10.1038/s41580-019-0159-6>.
4. Maurano MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science (80-)*. 2012;337:1190–1195.
5. Bonder MJ, et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet*. 2017;49:131–138.
6. Li M, et al. EWAS atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res*. 2019. <https://doi.org/10.1093/nar/gky1027>.
7. Min JL, et al. Genomic and phenomic insights from an atlas of genetic effects on DNA methylation. *medRxiv*. 2020.
8. Dekkers KF, et al. Blood lipids influence DNA methylation in circulating cells. *Genome Biol*. 2016;17:138.
9. Wiklund P, et al. DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clin Epigenetics*. 2019;11.
10. Katan M, Apolipoprotein B. E isoforms, serum cholesterol, and cancer. *Lancet*. 1986;327:507–508.
11. Katan MB. Commentary: Mendelian randomization, 18 years on. *Int J Epidemiol*. 2004. <https://doi.org/10.1093/ije/dyh023>.
12. van Dongen J, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat Commun*. 2016;7:11115.
13. Minică CC, Dolan CV, Boomsma DI, de Geus E, Neale MC. Extending causality tests with genetic instruments: an integration of Mendelian randomization with the classical twin design. *Behav Genet*. 2018. <https://doi.org/10.1007/s10519-018-9904-4>.
14. Ligthart L, et al. The Netherlands twin register: longitudinal research based on twin and twin-family designs. *Twin Res Hum Genet*. 2019. <https://doi.org/10.1017/thg.2019.93>.
15. Posthuma D, et al. Theory and practice in quantitative genetics. *Twin Res*. 2003. <https://doi.org/10.1375/136905203770326367>.
16. Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet*. 2002;3:872–882.
17. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. *Nat Genet*. 2017. <https://doi.org/10.1038/ng.3941>.
18. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011. <https://doi.org/10.1016/j.ajhg.2010.11.011>.

19. Yang J, Lee SH, Goddard ME, Visscher PM. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol Biol.* 2013. <https://doi.org/10.1007/978-1-62703-447-0-9>.
20. Yang J, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010. <https://doi.org/10.1038/ng.608>.
21. Visscher PM, et al. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet.* 2006. <https://doi.org/10.1371/journal.pgen.0020041>.
22. Visscher PM, et al. Genome partitioning of genetic variation for height from 11,214 sibling pairs. *Am J Hum Genet.* 2007. <https://doi.org/10.1086/522934>.
23. Van Dongen J, et al. The contribution of the functional IL6R polymorphism rs2228145, eQTLs and other genome-wide SNPs to the heritability of plasma sIL-6R levels. *Behav Genet.* 2014;44:368–382.
24. Zaitlen N, et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet.* 2013. <https://doi.org/10.1371/journal.pgen.1003520>.
25. Sandoval J, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics.* 2011;6:692–702.
26. Nivard MG, et al. Detection of gene–environment interaction in pedigree data using genome-wide genotypes. *Eur J Hum Genet.* 2016;24:1–7.
27. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet.* 2018. <https://doi.org/10.1038/s41576-018-0004-3>.
28. Fraga MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci.* 2005;102:10604–10609.
29. Slieker RC, et al. Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms. *Genome Biol.* 2016;17.
30. Teschendorff AE, West J, Beck S. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Hum Mol Genet.* 2013. <https://doi.org/10.1093/hmg/ddt375>.
31. Purcell S. Variance components models for gene-environment interaction in twin analysis. *Twin Res.* 2002;5:554–571.
32. Talens RP, et al. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell.* 2012;11:694–703.
33. Yang J, et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet.* 2011. <https://doi.org/10.1038/ng.823>.
34. Ouwens KG, et al. A characterization of cis- and trans-heritability of RNA-Seq-based gene expression. *Eur J Hum Genet.* 2019;2019. <https://doi.org/10.1038/s41431-019-0511-5>.
35. Hagenbeek FA, et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat Commun.* 2020;11(1):1–11.
36. Yang J, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet.* 2015. <https://doi.org/10.1038/ng.3390>.
37. Wainschtein P, et al. Recovery of trait heritability from whole genome sequence data. *bioRxiv.* 2019. <https://doi.org/10.1101/588020>.
38. Hancock DB, et al. Genome-wide association study across European and African American ancestries identifies a SNP in DNMT3B contributing to nicotine dependence. *Mol Psychiatry.* 2017. <https://doi.org/10.1038/mp.2017.193>.
39. Tsai PC, et al. Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health. *Clin Epigenetics.* 2018. <https://doi.org/10.1186/s13148-018-0558-0>.
40. Guida F, et al. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Hum Mol Genet.* 2015;24:2349–2359.
41. García-Rodríguez O, et al. Probability and predictors of relapse to smoking: results of the National Epidemiologic Survey on alcohol and related conditions (NESARC). *Drug Alcohol Depend.* 2013;132:479–485.

42. Dick KJ, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 2014;383:1990–1998.
43. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet*. 2013;14:483.
44. Cohen JC, Boerwinkle E, Mosley Jr TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med*. 2006;354:1264–1272.
45. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nat Rev Drug Discov*. 2013;12:581.
46. Wu N-Q, Li J-J. PCSK9 gene mutations and low-density lipoprotein cholesterol. *Clin Chim Acta*. 2014;431:148–153.
47. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012;90:7–24.
48. Davies NM, et al. The many weak instruments problem and Mendelian randomization. *Stat Med*. 2015;34:454–468.
49. Pierce BL, Ahsan H, VanderWeele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol*. 2010;40:740–752.
50. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol*. 2013;42:1134–1144.
51. Burgess S, Thompson SG. *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation*. Chapman and Hall/CRC; 2015.
52. Palmer TM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res*. 2012;21:223–242.
53. Sivakumaran S, et al. Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet*. 2011;89:607–618.
54. Bulik-Sullivan B, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47:1236–1241.
55. Thorleifsson G, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet*. 2009;41:18–24.
56. Berndt SI, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet*. 2013;45:501–512.
57. Graff M, et al. Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. *Hum Mol Genet*. 2013;22:3597–3607.
58. Locke A, Kahali B, Berndt S, Justice A, Pers T. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518:197–206.
59. Warrington NM, et al. A genome-wide association study of body mass index across early life and childhood. *Int J Epidemiol*. 2015;44:700–712.
60. Meyre D, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat Genet*. 2009;41:157.
61. Wheeler E, et al. Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat Genet*. 2013;45:513.
62. Elks CE, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet*. 2010;42:1077.
63. Perry JRB, et al. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature*. 2014;514:92–97.
64. Garcia-Closas M, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet*. 2013;45:392–398.
65. Michailidou K, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet*. 2013;45:353.
66. Willer CJ, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013;45:1274.

67. Scott LJ, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* (80-). 2007;316:1341–1345.
68. Zeggini E, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* (80-). 2007;316:1336–1341.
69. Timpson NJ, et al. Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes*. 2009;58:505–510.
70. Voight BF, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet*. 2010;42:579.
71. Perry JRB, et al. Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. *PLoS Genet*. 2012;8.
72. Consortium DS, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet*. 2014;46:234.
73. Minica CC, Boomsma DI, Van Der Sluis S, Dolan CV. Genetic association in multivariate phenotypic data: power in five models. *Twin Res Hum Genet*. 2010;13:525–543.
74. Minică CC, Boomsma DI, Vink JM, Dolan CV. MZ twin pairs or MZ singletons in population family-based GWAS? More power in pairs. *Mol Psychiatry*. 2014;19:1154.
75. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *Int J Epidemiol*. 2015;44:512–525.
76. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet*. 2018;50:1728–1734. <https://doi.org/10.1038/s41588-018-0255-0>.
77. van Kippersluis H, Rietveld CA. Pleiotropy-robust Mendelian randomization. *Int J Epidemiol*. 2017.
78. Hartwig FP, Davies NM, Davey Smith G. Bias in Mendelian randomization due to assortative mating. *Genet Epidemiol*. 2018;42:608–620.
79. Minica CC, Boomsma DI, Dolan V C, de Geus E, Neale MC. Empirical comparisons of multiple Mendelian randomization approaches in the presence of assortative mating. *Int J Epidemiol*. 2020;49:1185–1193.
80. Brumpton B, et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat Commun*. 2020;11. <https://doi.org/10.1038/s41467-020-17117-4>, 602516.
81. Munafò MR, Smith GD. *Robust Research Needs Many Lines of Evidence*; 2018.
82. Pedhazur EJ. *Multiple Regression in Behavioral Research: Prediction and Explanation*. Rinehart, Winst: New York Holt; 1982.
83. Hur Y-M, Craig JM. Twin registries worldwide: an important resource for scientific research. *Twin Res Hum Genet*. 2013;16:1–12.