The continuing value of twin studies in the omics era

Jenny van Dongen1, P. Eline Slagboom2, Harmen H. M. Draisma1, Nicholas G. Martin3 and Dorret I. Boomsma1

Abstract | The classical twin study has been a powerful heuristic in biomedical, psychiatric and behavioural research for decades. Twin registries worldwide have collected biological material and longitudinal phenotypic data on tens of thousands of twins, providing a valuable resource for studying complex phenotypes and their underlying biology. In this Review, we consider the continuing value of twin studies in the current era of molecular genetic studies. We conclude that classical twin methods combined with novel technologies represent a powerful approach towards identifying and understanding the molecular pathways that underlie complex traits.

Classical twin design

The classical twin study has been used for decades to estimate the importance of genetic and environmental influences on complex trait variation. Its results have contributed to the awareness that variation in almost every conceivable facet of the human condition is influenced by genetic variation (BOX 1). Traits include intrinsic physical, medical and biochemical characteristics, life-outcome variables, such as income, divorce and mortality, and behavioural traits, including apparently trivial ones such as television watching and Internet use. In fact, for many human phenotypes, heritability estimates derived from twin studies initially encouraged the search for the responsible genetic variation. Through their collaboration in genome-wide association study (GWAS) consortia, large twin registries (TABLE 1; Supplementary information S1 (table)) are nowadays also making an important contribution towards identifying the genetic variation that underlies complex traits and disorders.

Twins offer unique opportunities to genetic research that extend beyond the analysis of phenotypic heritability (BOX 2). Twin designs can provide insight into the genetic aetiology of disease development over time and can aid in the detection of biomarker profiles for medical conditions. For heritable traits, the comparison of discordant monozygotic twins (discordant MZ twins) represents a powerful improvement over the traditional case–control study to search for disease-associated biological marks. The power of this design is demonstrated in a recent study that compared the DNA methylation patterns of MZ twins who were discordant for systemic lupus erythematosus (SLE), and this study identified several genomic regions in which DNA methylation changes were associated with the disease1. Novel applications of the classical twin design can provide fundamental insights into the biological mechanisms underlying complex traits. For example, gene expression studies in MZ and dizygotic (DZ) twins have highlighted that variation in genome-wide expression between individuals is due to both genetic and environmental influences and that the importance of these influences may vary across genes and tissues2,3.

This Review addresses the continuing value of twin studies. We describe various twin study designs with examples of traditional applications, and we describe how twin approaches are now used for tracing disease-causing mutations and for studying various other newly emerging phenotypes (for example, the epigenome, transcriptome, metabolome, proteome and microbiome). We address using discordant MZ twins for the identification of biological mechanisms that are associated with complex traits, for the inference of causality and for the genome-wide analysis of genotype-by-environment (G×E) interaction at variability genes. We also discuss various questions that can be addressed by contrasting data from MZ and DZ twins to establish the heritability of biological marks and to unravel the shared aetiology of associated traits. A range of twin studies is presented, focusing on the initial level of the DNA sequence, down to its expression and intermediate phenotypes, such as metabolites, and ultimately to the clinical endpoints of interest.
Heritability
The proportion of variation in a trait that is due to heritable differences between individuals in a population: that is, the proportion of variation due to additive genetic effects (that is, narrow-sense heritability) or the proportion of variation due to all genetic effects (that is, broad-sense heritability).

Discordant monozygotic twins
(Discordant MZ twins). Twins who derive from a single fertilized egg cell but who are dissimilar for a certain characteristic or disease. By contrast, concordant MZ twins are phenotypically similar.

Case–control study
The comparison of individuals with a trait or disease of interest (cases) to controls to identify genes or other aspects associated with the trait. Cases and controls can be unrelated or can be relatives (within-family case–control design).

Epigenome
The entire collection of epigenetic marks, including DNA methylation and histone modifications, that regulate the expression of the genome. In contrast to the genome, the epigenome is specific to each cell.

The continuing importance of twin study designs
Quantitative analysis of genetic and environmental influences.
The classical twin design has traditionally been used to study the heritability of disease-related phenotypes and clinical endpoints (Table 2). This design has also been widely applied to estimating the extent to which different traits are influenced by the same or different genetic and environmental factors. Multivariate twin models of, for example, symptoms of anxiety and depression provided evidence that comorbidity of these disorders is due to genetic influences that affect the vulnerability to both disorders but that different environments determine whether a vulnerable person develops major depression or generalized anxiety disorder. Longitudinal data can be analysed in a similar way: genetic variation in intelligence quotient from ages 1 to 16 is largely attributable to the same genetic influences, and the increase in heritability is due to amplification of genetic effects with age. The classical twin design can be extended to model polygenic G × E interactions by testing whether the heritability of a trait varies across different levels of environmental exposures. The heritability of body mass index (BMI) is moderated by physical activity: the higher the level of physical activity, the smaller the genetic influence on BMI.

Extending twin models with data from other relatives (such as their parents, siblings, spouses or offspring) enhances statistical power and allows for testing of a much wider range of hypotheses about the causes of human variation, including the role of cultural transmission, social interactions among relatives, genetic non-additivity and various mechanisms of assortative mating. The ‘offspring-of-twins’ design is a powerful tool for studying intergenerational associations between environmental variables and outcomes in children. Also, comparing the phenotypic similarity of children of female MZ twins (who socially are cousins but genetically are half-siblings) to the similarity of children of male MZ twins gives insight into the differential importance of paternal and maternal effects; if paternal and maternal effects are equally important, children of male twins and female twins are expected to be equally similar. For birth weight, larger correlations have been observed in children of female twins compared with children of male twins, highlighting the importance of maternal effects for this trait.

Classical twin methods continue to be a valuable addition to genetic association studies to establish, for example, the proportion of the heritability that can be explained by newly identified SNPs from GWASes. The current discussion about ‘missing heritability’ largely stems from the (often great) disparity between estimates of total heritability from twin studies and the proportion of variance accounted for by SNPs from GWASes, for which many explanations have been proposed, including implications that heritability estimates from twin studies may be too high. In our later section on testing classical assumptions, we discuss the relevance of recent molecular findings in twins in the light of the current discussion on ‘missing heritability.’
The comparison of discordant MZ twins offers an alternative to the traditional case–control study. Here, the primary interest is not to infer causality but to identify factors associated with a trait of interest that differ between cases and controls who are perfectly matched for age, sex and genetic background, and who are partly matched for early environmental influences. Molecular phenotypes and the causes of quantitative trait variation. Technological advances allow an assessment of the extent to which twins resemble each other at the level of molecular processes that contribute to their phenotypic similarity. Thereby, the comparison of discordant MZ twins can lead us into novel pathways associated with disease. A unique advantage of the MZ twin design is the ability to study biological discordance

The total set of RNA transcripts that are produced in a cell or tissue by transcription of DNA.

<table>
<thead>
<tr>
<th>Twin registry name</th>
<th>Registry characteristics</th>
<th>Age</th>
<th>Website</th>
<th>Number of twins or subjects (approximate)*</th>
<th>Number of twins or subjects with DNA available (approximate)*</th>
<th>Biospecimens (available for at least subset of the sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandim Health Project twin registry (Guinea-Bissau)</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>0–30</td>
<td><a href="http://www.bandim.org">http://www.bandim.org</a></td>
<td>2,500 (twins and singleton controls)</td>
<td>200 twin pairs</td>
<td>Whole blood, plasma</td>
</tr>
<tr>
<td>Australian Twin Registry</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>0–90</td>
<td><a href="http://www.twins.org.au">http://www.twins.org.au</a></td>
<td>66,000</td>
<td>12,000 (twins and other family members)</td>
<td>Serum, plasma, buccal cells</td>
</tr>
<tr>
<td>Chinese National Twin Registry (CNTR)</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>All</td>
<td><a href="http://cntr.bjmu.edu.cn">http://cntr.bjmu.edu.cn</a></td>
<td>35,000 twin pairs</td>
<td>3,200</td>
<td>Serum, DNA</td>
</tr>
<tr>
<td>South Korean Twin Registry (SKTR)</td>
<td>Volunteer preschoolers, cohort of school children, volunteer young adults</td>
<td>0–30</td>
<td><a href="http://www.ktrc.org">http://www.ktrc.org</a></td>
<td>10,000 twin pairs</td>
<td>800 twin pairs</td>
<td>Hair, saliva</td>
</tr>
<tr>
<td>Keio Twin Registry (Japan)</td>
<td>Adult and adolescent twins from the general population in the Tokyo area</td>
<td>14–30</td>
<td><a href="http://totcop.keio.ac.jp">http://totcop.keio.ac.jp</a>; <a href="http://kts.keio.ac.jp">http://kts.keio.ac.jp</a>; <a href="http://kotrec.keio.ac.jp">http://kotrec.keio.ac.jp</a></td>
<td>4,000 twin pairs (plus other family members)</td>
<td>600 twin pairs</td>
<td>Buccal cells, blood</td>
</tr>
<tr>
<td>Sri Lankan Twin Registry</td>
<td>Voluntary twin registry component and a population-based database with ongoing data collection</td>
<td>6–94</td>
<td><a href="http://www.ird.lk/twin%20Registry.php">http://www.ird.lk/twin%20Registry.php</a></td>
<td>35,000</td>
<td>Plans to collect DNA from 4,000</td>
<td>Buccal cells</td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Danish Twin Registry (DTR)</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>0–107</td>
<td><a href="http://www.sdu.dk/dtr">http://www.sdu.dk/dtr</a></td>
<td>170,000</td>
<td>20,000</td>
<td>Serum, plasma, buffy coat, saliva, buccal cells, urine</td>
</tr>
<tr>
<td>Finnish Twin Cohort study</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>11–100+</td>
<td><a href="http://www.twinstudy.helsinki.fi">http://www.twinstudy.helsinki.fi</a></td>
<td>45,000 (plus family members)</td>
<td>14,600 (twins and family members)</td>
<td>Whole blood, serum, plasma, saliva, urine, fat and muscle by biopsy</td>
</tr>
<tr>
<td>Netherlands Twin Register (NTR)</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>0–100</td>
<td><a href="http://www.tweelingenregister.org/en">http://www.tweelingenregister.org/en</a></td>
<td>87,500 (plus family members)</td>
<td>18,000</td>
<td>DNA, RNA, cell lines, serum, plasma, buccal cells, urine, stool</td>
</tr>
<tr>
<td>Norwegian Twin Registry (NTR)</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>18+</td>
<td><a href="http://www.fhi.no/twins">www.fhi.no/twins</a></td>
<td>40,000</td>
<td>4,800</td>
<td>Whole blood, buccal cells, plasma</td>
</tr>
<tr>
<td>Swedish Twin Register (STR)</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>5–100+</td>
<td><a href="http://ki.se/kj/isp/poldopoly.jsp?sessionid=acR0j1HrWECiIOcN/">http://ki.se/kj/isp/poldopoly.jsp?sessionid=acR0j1HrWECiIOcN/</a></td>
<td>194,000</td>
<td>44,600</td>
<td>Whole blood, serum, saliva</td>
</tr>
<tr>
<td>TwinsUK registry</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>18–90</td>
<td><a href="http://www.twinsuk.ac.uk">http://www.twinsuk.ac.uk</a></td>
<td>12,000</td>
<td>7,000</td>
<td>Whole blood, serum, plasma, buffy coat, saliva, buccal cells, urine, skin, fat, muscle</td>
</tr>
</tbody>
</table>
against an equivalent genetic background. Divergence of epigenic profiles in MZ twins depends on the locus and has been documented for both younger and older age groups. In fact, differences in DNA methylation and gene expression are already evident in newborn MZ twins. Clearly, environmental and stochastic factors start in utero and operate throughout life.

In addition to traditional organisational quantitative traits (such as height and BMI), molecular characteristics (such as gene expression levels, the methylation state of CpG sites in the DNA and the concentration of metabolites in blood and urine) may be regarded as quantitative traits. Variation in molecular traits measured in groups of MZ and DZ twins can be analysed using the classical twin method, like any other phenotype. Multivariate twin analyses address questions that are not easily resolved in any other study design, such as to what extent is the epigenetic regulation and expression of genes across genomic regions influenced by shared genetic factors and to what extent is each region influenced by unique factors? And to what degree do common genetic and environmental mechanisms underlie biological variation across different cells and tissues? The availability of genome-wide DNA marker data allows for novel approaches towards studying G×E interactions, in which MZ twins can play a vital part. By studying variation in a phenotypic trait of interest in MZ twins, it is possible to see not only whether some genotypes confer higher levels of risk for that trait but also whether some contribute to its variability; high variability in the expression of a trait from a common genetic background could explain phenotypic differences between MZ co-twins. Of interest, genetic and environmental factors may influence disease through different pathways. Twin studies can be used to identify aspects of disease that are most related to the underlying genetic liability of individuals and can thereby help to establish clinical criteria and phenotypic definitions that will facilitate the success of GWASs. Other approaches, such as the offspring-of-twins design, may provide insight into transgenerational inheritance of epigenetic regulation and the importance of maternal effects and imprinting on epigenetic marks, although such studies have not yet been published.

An important strength of twin registries lies in the extensive longitudinal collection of data on various phenotypes. Twin studies have indicated that approximately 20–30% of the overall variation in adult lifespan is accounted for by genetic factors. Longitudinal twin studies can be used to identify biomarkers that are associated with ageing: a co-twin control analysis showed that telomere length at advanced age is predictive of survival. MZ twins with the shortest telomeres at the baseline had a threefold greater risk of death during a follow-up period of 7 years than their co-twins with the longest telomere measurements (relative risk = 2.8). The discordant MZ twin design and the classical twin design have received much interest in recent years for studying molecular biology. The following sections will provide an overview of findings from such studies.

Table 1 (cont.) | A selection of twin registries worldwide

<table>
<thead>
<tr>
<th>Twin registry name</th>
<th>Registry characteristics</th>
<th>Age</th>
<th>Website</th>
<th>Number of twins or subjects (approximate)*</th>
<th>Number of twins or subjects with DNA available (approximate)*</th>
<th>Biospecimens (available for at least subset of the sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid-Atlantic Twin Registry (MATR; United States)</td>
<td>Population-based, ascertained at birth</td>
<td>0–94</td>
<td><a href="http://www.matrycu.edu">http://www.matrycu.edu</a></td>
<td>56,000</td>
<td>1,500</td>
<td>Whole blood, serum, plasma, buffy coat, saliva, buccal cells</td>
</tr>
<tr>
<td>Minnesota Twin Family Study (MTFS; United States)</td>
<td>Ongoing population-based longitudinal study</td>
<td>11–47</td>
<td><a href="http://mctfr.psych.umn.edu">http://mctfr.psych.umn.edu</a></td>
<td>5,000 (plus family)</td>
<td>10,000 (twins and family members)</td>
<td>Blood-derived or saliva-derived DNA</td>
</tr>
<tr>
<td>Wisconsin Twin Panel (WTP; United States)</td>
<td>Population-based, longitudinal data, extensive phenotypic characterization, follow-up of selected samples</td>
<td>0–23</td>
<td><a href="http://www.waisman.wisc.edu/twinresearch">http://www.waisman.wisc.edu/twinresearch</a></td>
<td>19,638 twins (plus parents and siblings)</td>
<td>3,489 (twins, parents and siblings)</td>
<td>Saliva, buccal cells</td>
</tr>
<tr>
<td>South America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuban Twin Registry</td>
<td>Population-based with ongoing longitudinal data collection</td>
<td>All</td>
<td><a href="http://www.waisman.wisc.edu/TwinsStudy.aspx">http://www.waisman.wisc.edu/TwinsStudy.aspx</a></td>
<td>55,400 twin pairs</td>
<td>250 twin pairs</td>
<td>Blood-derived DNA</td>
</tr>
</tbody>
</table>

*Numbers refer to individual twins (rather than twin pairs) unless indicated otherwise. This table shows a selection of some of the large twin registries worldwide. For a more comprehensive table, see Supplementary information S1 (table).
Zygosity assessment
The assessment whether same-sex twins are monozygotic or dizygotic is often based on the comparison of DNA markers or alternatively on standardized questionnaires.

Multivariate twin models
Models used for the simultaneous analysis of multiple traits measured in monozygotic and dizygotic twins to estimate the importance of genetic and environmental influences shared (overlapping) between traits in explaining their clustering, comorbidity or covariance.

Tracing the origin of new mutations
Identifying sequence differences between twins. Although MZ twins originate from one zygote, there is some evidence that their somatic cells are not always identical at the DNA sequence level. A study of healthy MZ twins and singletons suggested that copy number variations (CNVs) may accumulate with ageing in a dynamic fashion. By comparing CNVs in longitudinally collected blood samples of MZ pairs, both increases and decreases in CNV content were found after 10 years (between co-twins and within individual twins). This may reflect fluctuations in the proportions of peripheral blood cells carrying aberrant DNA. By comparing copy numbers in buccal cells of twins and their parents, evidence was found for a pre-twinning de novo duplication in a healthy twin pair (that was present in both twins but not in their parents) and a post-twinning de novo deletion in one twin from a pair of twins who were discordant for attention problems. A comparison of CNVs in the blood of MZ pairs who were discordant for congenital diaphragmatic hernia and oesophageal atresia found no evidence for structural genomic differences between twins. All of these studies used microarrays, which cover a limited portion of the total content of structural variation in the genome. The application of whole-genome-sequencing techniques may unravel many more sequence differences between MZ twins, including single-nucleotide substitutions.

In 2010, the first study was published that applied whole-genome-sequencing technology in discordant MZ twins. The study entailed a combination of techniques — including whole-genome sequencing, RNA sequencing and genome-wide SNP microarrays — to measure multiple molecular marks in CD4+ cells from female twins who are discordant for multiple sclerosis. Only a small fraction of SNPs and structural variants differed within twin pairs, but no differences were replicated across methods. However, this study should be interpreted as exploratory, as only three discordant pairs were studied. Larger studies are needed to establish whether molecular differences may explain discordance for multiple sclerosis and other diseases in MZ twin pairs.

Timing the occurrence of de novo mutations
A unique advantage of studying disease-causing mutations in MZ twins is that the developmental timing of de novo mutations may be tracked if DNA from multiple cell lines is available for both twins. The timing of a mutation in the sodium channel α1 subunit gene (SCN1A) that causes Dravet’s syndrome was determined by sequencing DNA from several embryonic tissue lineages from a pair of discordant MZ twins. As the mutation was present in all analysed cell lines of the affected twin but not in those of the unaffected co-twin, it was concluded that the mutation had probably occurred at the two-cell stage in the pre-morula embryo. For any disease caused by de novo mutations, information about the timing of mutagenesis is of major importance for genetic counselling. Mutations that occur in parental gametes are associated with a negligible risk of recurrence in additional offspring. By contrast, parental germline mosaicism for the mutation is associated with a high recurrence risk because many existing parental gametes will carry the mutation.

Phenotypic impact of epigenetic variation
DNA methylation and disease. In addition to de novo mutations in the DNA, epigenetic variation may be another important source of phenotypic variation and discordance in MZ twins. The following example demonstrates this point. In 1997, a pair of MZ girls was born; one of them was healthy, but the other had a severe spinal malformation in which the spinal cord was duplicated. This defect resembled a condition in mice with a mutation in the Axin1 gene, but no mutation was found in this gene in the twins. However, increased methylation of CpG sites at the AXIN1 promoter was found in the affected twin as compared with the unaffected twin, and this may have suppressed gene expression and caused the malformation.

Although epigenetic variation has not yet been investigated in large twin studies, several small studies

Box 2 | The classical twin design
In the classical twin design, the extent to which phenotypic variation in a trait (V_P) is due to genetic (V_G) and environmental (V_E) influences is estimated as \( V_P = V_G + V_E \). Genetic variance can be further decomposed into additive genetic variance (V_A) and variance due to non-additive genetic effects (dominance variance (V_D)); \( V_G = V_A + V_D \). Most twin studies, unless they are very large, consider the narrow-sense heritability (h²), which refers to the proportion of variation that is due to additive genetic variance: \( h² = V_A / V_P \).

Environmental influences (V_E) comprise those that are shared by family members (‘the common environment’ (V_C)); \( V_E = V_C + V_U \). These unobserved variance components can be estimated from the observed resemblance (that is, the phenotypic covariance) in monozygotic (MZ) and dizygotic (DZ) twin pairs. MZ twins are derived from a single fertilized egg cell and share (nearly) 100% of their segregating genes, whereas DZ twins are derived from two distinct zygotes and share on average 50% of their segregating genes. Twins of both types share 100% of the common environment and 0% of the unique environment. Therefore, the phenotypic covariance of MZ twins is expected to equal \( V_C + V_U + V_A \), and the phenotypic covariance of DZ twins is expected to equal 0.5 \( V_C + 0.25 V_U + V_A \). These expectations are the input (that is, the structural equations) for genetic structural equation modelling (GSEM), a technique by which maximum likelihood estimates of variance components are obtained from twin data. GSEM obtains the expected MZ and DZ covariances given the equations above and compares the outcome to the covariances observed in the data. The maximum likelihood estimates of \( V_C, V_U, V_A \), and \( V_E \) are those estimates that predict covariances that are most consistent with the observed data. MZ and DZ data, \( V_C, V_U, \) and \( V_A \) cannot be estimated simultaneously. \( V_A \) is estimated if there is stronger evidence for non-additive effects (if the MZ correlation is more than twice as large as the DZ correlation), and \( V_U \) is estimated if there is stronger evidence for common environmental effects (if the MZ correlation is less than twice as large as the DZ correlation). In extended-twin-family designs, the information from additional types of family relations together with the information from twins allows for estimating \( V_C, V_U, V_A \), and \( V_E \) simultaneously.

In multivariate twin models, extending the set of equations for the expected covariances allows the modelling of the cross-twin—cross-trait covariance — that is, the covariance of trait one in one twin with trait two in the co-twin. To estimate the degree to which the clustering of different traits or comorbidity of disorders is explained by genetic and environmental influences, the same principles apply as for the expected covariances of twins (for example, MZ twins are expected to share 100% of genetic influences that overlap between traits, whereas DZ twins are expected to share 50%, resulting in a larger cross-twin—cross-trait covariance for MZ twins if the association between traits has a genetic basis).
Genetic non-additivity

Refers to genetic effects that contribute to the phenotypic variance in a non-additive manner. These include the effects of interacting alleles at a single locus (dominance) and interactions between different loci (epistasis).

Assortative mating

Refers to the situation whereby a trait is correlated in spouses because it influences partner choice (phenotypic assortment) or because it correlates with certain environments that influence partner choice (social homogamy). It is also called nonrandom mating.

Maternal effects

Effects that are transmitted from mother to offspring including genetic effects. The phenotype in offspring can be influenced by: the maternal allele, mitochondrial inheritance, the effects of the prenatal environment (for example, nutrient supply in utero) or the maternal supply of RNA or proteins to the egg cell.

Co-twin control method

A method of examining the associations between traits using discordant twins. If monozygotic twins who are discordant for trait 1 are also discordant for trait 2, the association between these traits is unlikely to be confounded by underlying shared genetic or early environmental influences.

Transgenerational inheritance

The transmission of a trait across generations (genetic or cultural inheritance). Epigenetic variation may also be transmitted across generations.

Imprinting

The mechanism that can occur at some loci to silence the expression of one of the two alleles, depending on the parent-of-origin of the allele.

Copy number variations (CNVs). These refer to large DNA segments (>1 kb) of which the number of copies is variable (for example, between individuals or between cells within an individual) — for example insertions, deletions and duplications.

demonstrate the promise of the discordant twin design for epigenetics, including studies of Alzheimer’s disease, autism, bipolar disorder, birth weight, cancer and SLE. In MZ twins who are discordant for the autoimmune disorders SLE, rheumatoid arthritis or dermatomyositis, a global decrease in DNA methylation (hypomethylation) was identified in SLE-affected twins, as were regional DNA methylation changes at 49 genes that were enriched for immune function. Many of the genes that were hypomethylated in SLE-affected twins also showed increased expression compared with the healthy co-twin. Integrated studies of DNA methylation and gene expression in discordant twins are particularly valuable for identifying loci at which epigenetic regulation may be associated with disease. Importantly, the dynamic nature of epigenetic variation makes results of epigenetic studies more difficult to interpret compared with genetic studies. Alternatively to being the cause of disease discordance, epigenetic differences may also reflect the effects of disease or the effect of an event occurring in one twin that independently triggered both the disease and the epigenetic changes. Some twin registries have collected longitudinal biological samples, and this allows for identifying epigenetic differences between twins that were already present before the onset of discordance for some diseases. Functional studies will ultimately be required to verify the effect of epigenetic variation.

The classical twin design provides information about the importance of genetic influences on epigenetic variation: comparison of the level of DNA methylation at the imprinted IGF2–H19 locus in MZ and DZ twins showed that variation in DNA methylation at this locus is mainly determined by heritable factors before middle age. High heritability of epigenetic variation has also been observed for some other loci, although the average heritability across all loci seems to be low.

Differential miRNA expression and disease. The role of non-coding RNAs such as microRNAs (miRNAs) is fairly unexplored. In a sample of MZ twins and sibling pairs who were discordant for autism, miRNA expression in lymphoblastoid cell lines was measured, and differential regulation of a number of miRNA transcripts was observed. For two differentially expressed brain-specific miRNAs, the putative target genes inhibitor of DNA binding 3 (ID3) and polo-like kinase 2 (PLK2), which have been implicated in circadian rhythm signalling and modulation of synapses, were validated by experiments involving knockdown or overexpression of these miRNAs. By combining miRNA data and mRNA expression data, dysregulation of miRNA expression was found to contribute to alterations in target gene expression, which in turn may contribute to disease pathology of autism. The expression of miRNAs was measured in MZ twins who were discordant for lupus nephritis, and differential expression of several miRNAs was observed. Primarily among the gene targets of the most important miRNAs were genes that have a role in interferon signalling (IFN signalling). Together, these studies indicate that the discordant MZ twin design will be a valuable approach towards exploring the role of miRNA expression in complex disease.

Gene expression: causes and disease links

There is wide variation in the heritability of transcript expression across the genome. To identify expression quantitative trait loci (eQTLs), variation in expression across tissues of healthy female twin pairs was investigated in a ‘matched co-twin analysis’. In the initial stage, SNP associations were tested in one twin of each pair. Although this method of eQTL identification does not require twins, the co-twins in this study served to replicate and validate the identified eQTLs, thus providing extra confidence in the findings.

A frequent use of twin studies is to identify gene expression alterations (on a shared genetic background) that are associated with various disease states; such genes may provide mechanistic insight into disease pathogenesis. A study of gene expression in subcutaneous fat of obesity-discordant MZ twins detected differential expression of a range of genes. Differentially expressed genes included those that are involved in inflammatory pathways (which were upregulated in obese twins) and in mitochondrial branched-chain amino acid (BCAA) catabolism (which were downregulated in obese twins). Interestingly, the largest increase in expression in obese twins was reported for the gene encoding the inflammatory cytokine osteopontin (SPP1), which has previously been associated with obesity and insulin resistance in mice. Other diseases for which gene expression changes have been identified in discordant MZ twins include rheumatoid arthritis, bipolar disorder, schizophrenia and type 1 diabetes. A comparison of the skeletal muscle transcriptomes in MZ twins who are discordant for postmenopausal osteoarthritis highlights the insights that may be obtained from MZ twins who are discordant for drug treatment, regarding the long-term effects of drug therapies. Several pathways were differentially regulated in twins who received hormonal treatment, and expression differences correlated significantly with differences in muscle performance between the twins. Large twin studies estimating the heritability of expression of individual transcripts have not yet been published.

Metabolomics

Metabolites may serve as biomarkers of health and disease and can be quantified in body fluids and tissue samples by approaches such as mass spectrometry and ¹H NMR spectroscopy. The first metabolomics study based on ¹H NMR spectroscopic analysis of urine and blood plasma from MZ and DZ twin pairs showed that familial factors (such as genetic influences and family environment) explain on average 42% of the variation in metabolite peak heights in plasma and 30% of the variation in urine. In two GWASs of metabolite profiles, data from twins allowed the proportion of variance in metabolite levels explained by significantly associated SNPs to be compared with the proportion explained by the total genetic or familial variance. Heritability estimates of metabolic measures based on data from 221 MZ twin pairs and...
340 DZ twin pairs ranged between 23% and 55% for amino acids and other small-molecule metabolites. Estimates were higher for lipids (48–62%) and lipoproteins (50–76%). Although for most direct metabolite measures the total variance explained by significantly associated SNPs was 10% at most, higher estimates of explained variance were observed for certain metabolite ratios. The highest explained variance (25%) was observed for the ratio of linoleic acid to other polyunsaturated fatty acids. The twin-based heritability for this ratio was 62%, implying that 40% of the total heritability can be ascribed to SNPs, which is high compared with most other (clinical) phenotypes.

Whereas traditional enzymatic methods usually provide composite measures of metabolites, 1H NMR gives more detailed insight into the behaviour of individual metabolites in pathways. In a direct comparison, similar estimates of heritability were found for most composite lipid measures on the basis of either enzymatic methods or 1H NMR. This supports the notion that high-resolution metabolomics techniques are reliable.

Similarly to differentially expressed genes, differential levels of other molecules can be linked to disease pathogenesis. After detecting differences in serum and fat tissue lipid profiles in MZ twins discordant for obesity, a simulation of lipid bilayer dynamics was carried out using lipidomics and gene expression data from the twins, providing novel functional insights into the biological pathways that underlie adipocyte expansion. This study shows how findings from discordant twin studies may encourage and guide further functional or bioinformatic approaches to obtain in-depth mechanistic insights into the pathological mechanisms that underlie complex traits and disease.

To date, there have been few proteomic studies in twins. A twin study of serum protein levels, as measured by antibody arrays, found that only a small proportion of the variation was attributable to familial factors; however, experimental variation in this study was fairly large.

### Tissue specificity of molecular variation

In concordance with most molecular and genetic epidemiological studies, most twin studies have been based on peripheral blood. But how well does a molecular profile in blood cells reflect epigenetic and gene expression changes that are associated with different phenotypes and diseases in relevant tissues? Epigenetic changes that arise at later stages of development and throughout life are more likely to be limited to specific tissues or even cells. DNA methylation profiles of MZ twins who are discordant for major psychosis suggest that epigenetic changes related to psychosis may be reflected in peripheral blood. In this study, the most significant methylation change in psychosis-affected twins — that is, hypomethylation at the promoter of the ST6GALNAC1 gene — was also evident in post-mortem brain tissues of some psychosis patients. However, large studies are warranted to establish how well molecular profiles in blood reflect those occurring in...
**Table 2 (cont.)** Heritability estimates from twin studies

<table>
<thead>
<tr>
<th>Trait</th>
<th>Heritability</th>
<th>Number of twin pairs (or study type for multiple data sets)</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cancer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>0.42</td>
<td>21,000</td>
<td>114</td>
</tr>
<tr>
<td>Breast cancer (in females)</td>
<td>0.27</td>
<td>23,788</td>
<td>114</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>0.35</td>
<td>44,788</td>
<td>114</td>
</tr>
<tr>
<td><strong>Ageing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td>0.25</td>
<td>Review</td>
<td>34</td>
</tr>
<tr>
<td>Telomere length</td>
<td>0.56</td>
<td>175</td>
<td>35</td>
</tr>
<tr>
<td><strong>Lifestyle and life events</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise participation</td>
<td>0.48–0.71†</td>
<td>37,051</td>
<td>89</td>
</tr>
<tr>
<td>Dietary patterns</td>
<td>0.41–0.48</td>
<td>3,262†</td>
<td>90</td>
</tr>
<tr>
<td>Smoking initiation</td>
<td>M: 0.37; F: 0.55</td>
<td>Meta-analysis</td>
<td>147</td>
</tr>
<tr>
<td>Smoking persistence</td>
<td>M: 0.59; F: 0.46</td>
<td>Meta-analysis</td>
<td>147</td>
</tr>
<tr>
<td>Alcohol abuse or dependence</td>
<td>0.50–0.70</td>
<td>Review</td>
<td>148</td>
</tr>
<tr>
<td>Stressful life events</td>
<td>0.28</td>
<td>Meta-analysis</td>
<td>92</td>
</tr>
</tbody>
</table>

*Note that numbers refer to twin pairs unless stated otherwise, and most heritability estimates refer to the narrow-sense heritability (h²; BOX 2). †Range of heritabilities from different countries or study samples. ‡Female twin pairs with child (offspring-of-twin design). §Only females. ††The original paper reports estimates for various age categories from 3–71 years, separately for males and females, F, females; M, males.

### Congenital diaphragmatic hernia
A birth defect that is characterized by malformation of the diaphragm, lung hypoplasia and pulmonary hypertension.

### Oesophageal atresia
A congenital malformation of the oesophagus in which the oesophagus does not form an open passage to the stomach and may be connected to the trachea.

### Maximum likelihood
Maximum-likelihood estimation obtains estimates of population parameters from a data set by computing the probability (likelihood) of obtaining the observed data for a range of different parameter values and evaluating for which values the probability of observing the data is highest.

### Dravet’s syndrome
A childhood-onset epileptic encephalopathy that is also called severe myoclonic epilepsy of infancy.

### Host genetic influences on the microbiome
Studies of the human gut microbiome have revealed considerable variation in the composition of microbial communities between individuals. It remains to be established to what degree this variation is controlled by host genetics, but greater similarity has been observed in family members compared to unrelated individuals. A few studies have explored the role of host genetics by comparing various measures of the microbiome in small groups of MZ and DZ twins, but findings have so far been inconclusive, with some studies suggesting that the microbiota are slightly more similar in MZ twins compared with DZ twins and other studies observing comparable levels of similarity of the faecal microbiome of MZ and DZ twins. An important factor in the comparison of similarity of individuals is the level that is compared: the overlap between relatives may be small at the organismal level but might be larger at relevant functional levels (for example, the degree to which microbial genes and metabolic pathways are shared).

A few studies in twins searched for microbial signatures that are associated with disease. A comparison of the faecal microbial communities in (discordant) obese and lean MZ twins showed that obesity is associated with various changes, including reduced bacterial diversity and differences in the representation of specific bacterial genes and metabolic pathways. In MZ twins who are discordant for inflammatory bowel diseases, certain gastrointestinal bacterial populations differed in abundance among individuals with different clinical phenotypes of Crohn’s disease, which is relevant to our understanding of the pathogenesis behind inflammatory bowel diseases. MZ twins who are discordant for ulcerative colitis differed in the composition of the microbiota and in the expression of human RNA transcripts that are related to oxidative and immune responses in the mucosal epithelium. In affected twins, fewer RNA transcripts correlated with bacterial genera than in unaffected twins, suggesting that ulcerative colitis may be associated with a loss of interaction between the mucosal transcriptional profile and the colonic microbiota.

### The interplay of genes and environment
Genetic and environmental influences in many cases do not act independently. Gene–environment correlation (G×E) refers to the situation in which exposure to certain environments is under genetic control. For instance, twin and adoption studies have found that lifestyle factors (for example, exercise participation and diet), life events (for example, divorce) and life circumstances (for example, family environment and social support) are moderately heritable. Thus, influences that are usually considered as measures of ‘environment’ might often be better described as external factors that are partly under genetic control. By contrast, G×E interaction refers to the scenario in which different genotypes have different reactions to the same environmental exposure. By comparing differences in serum lipid levels in MZ twins across pairs with different genotypes, it was found that the Kidd (JK) blood group locus is associated with variability in the total cholesterol level.
**Box 3 | The value of twins in neuroimaging genetics**

A similar approach was used to test whether an interaction between the length polymorphism in the SLC6A4 serotonin transporter gene and environmental stress is associated with MZ discordance for depression; no evidence was found for this hypothesis.

**Testing classical assumptions**

MZ twins share all of their segregating genes, whereas DZ twins share on average 50%. The assumptions of the classical twin method and the interpretation of results have always been a subject of debate (for a...
Table 3 | MZ and DZ twin concordance for complex disease

<table>
<thead>
<tr>
<th>Trait</th>
<th>MZ twins</th>
<th>DZ twins</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1 diabetes</td>
<td>42.9</td>
<td>7.4</td>
<td>129</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>34</td>
<td>16</td>
<td>130</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>25.3</td>
<td>5.4</td>
<td>149</td>
</tr>
<tr>
<td>Crohn’s disease</td>
<td>38</td>
<td>2</td>
<td>150</td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>15</td>
<td>8</td>
<td>150</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>32.2</td>
<td>8.7</td>
<td>134</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>15.5</td>
<td>11.1</td>
<td>151</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>40.8</td>
<td>5.3</td>
<td>152</td>
</tr>
<tr>
<td>Major depression</td>
<td>31.1(^a) or 47.6(^b)</td>
<td>25.1(^a) or 42.6(^b)</td>
<td>153</td>
</tr>
<tr>
<td>Attention-deficit hyperactivity disorder</td>
<td>82.4</td>
<td>37.9</td>
<td>154</td>
</tr>
<tr>
<td>Autism spectrum disorders</td>
<td>93.7</td>
<td>46.7</td>
<td>155</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>11</td>
<td>5</td>
<td>114</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>13(^a)</td>
<td>9(^a)</td>
<td>114</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>18</td>
<td>3</td>
<td>114</td>
</tr>
</tbody>
</table>

*Defined as \(2C / (2C + D)\), where \(C\) is the number of concordant affected twin pairs, and \(D\) is the number of discordant twin pairs. \(^a\)Concordance in male twin pairs. \(^b\)Concordance in female twin pairs. DZ, dizygotic; MZ, monozygotic.

Lipid bilayer dynamics

The dynamic properties of lipid bilayer membranes, such as thickness, fluidity and permeability, that influence the physiological properties of a cell.

Lipidomics

The comprehensive study of the entire set of lipids in biological systems, such as cells, tissues and organs, using metabolomics techniques.

Chimerism

The situation in which an individual carries some of the genetic material originating from another individual (for example, originating from the co-twin or originating from the mother).

Microbiota

The collection of all microorganisms living in a certain environment (for example, the human gut).

Identity-by-descent sharing (IBD sharing)

Refers to the proportion of alleles in two individuals that are derived identically by descent from a common ancestor.

Detailed discussion of the difficulties related to the concept of heritability, see REF. 98. A first assumption is that MZ twins are genetically identical, for which it has now been proved that there are exceptions to the rule. Still, the difficulty of various whole-genome-sequencing efforts to find any replicable differences between MZ twins\(^{98,41}\) suggests that DNA sequence differences between MZ twins are not large, although an exact estimation of somatic sequence variation (given the nontrivial error rate in sequencing itself) has not been reported.

The availability of genome-wide marker data also allows us to address the assumption that DZ twins share on average 50% of their segregating genetic material by estimating the true amount of genetic material that DZ twins have inherited from the same parent (that is, identity-by-descent sharing (IBD sharing)). Using genome-wide microsatellite marker data, it was demonstrated that the proportion of IBD sharing in most (95%) DZ twins and siblings lies within the range of 42–58%, with an average close to 50%\(^{46}\). Using the empirical IBD measure instead of assumptions about genetic sharing, the heritability of height was estimated at 0.86, which is highly consistent with results from traditional twin studies, providing perhaps the most pertinent evidence to support the estimates of narrow-sense heritability from twin studies.

MZ twins share environmental influence to the same degree as DZ twins. Now that the classical twin design is being used to study epigenetic variation, it is becoming evident that novel attention has to be paid to the assumption that MZ twins share environmental influences to the same degree as DZ twins. Because MZ twins are derived from a single zygote, they may start out with more similar epigenomes than DZ twins, who originate from two zygotes with unique epigenetic profiles. DZ twins may thus start with more epigenetic differences than MZ twins owing to a cause that is not necessarily related to genetic differences. Although this hypothesis remains to be tested, an important observation in this light has been provided by a comparison of small groups of MZ twins that were either monochorionic or dichorionic. The DNA methylation profiles of buccal epithelial cells were more similar in dichorionic MZ twins than in monochorionic MZ twins\(^{80}\), and this may be related to the timing of splitting of the zygote. Thus, differences in epigenetic resemblance of monochorionic and dichorionic twins may be due to epigenetic divergence of embryonic cells that takes place after the blastomeric stage. Although this issue requires further study in larger samples, it shows that prenatal developmental processes related to twinning may influence the epigenetic resemblance of twins. Importantly, if MZ twins are epigenetically more similar than DZ twins owing to non-genetic causes, the heritability of phenotypes that are epigenetically regulated may be overestimated.

Twin concordance and disease liability

Relationship between heritability and discordance rates in MZ twins. A high concordance of MZ twins on its own does not imply a high heritability, as demonstrated by concordance for measles. Before immunization was introduced, concordance was close to 100% in both MZ and DZ twins\(^{100}\). This indicates that, despite the high concordance in MZ twins, genetic differences between individuals actually contribute little to differences in the vulnerability to this infectious disease. Likewise, a high rate of disease discordance in MZ twins does not rule out the importance of genetic influences. Although MZ twins are usually remarkably similar in appearance, MZ twins who are discordant for disease are often observed (TABLE 3). It is generally assumed that liability to disease is continuous, and disease becomes evident after a threshold has been passed. The probability of observing discordant MZ twins thus depends on the heritability of the underlying liability and on the level of the threshold\(^{101}\). Especially for rare disorders (for which the threshold is high), many affected MZ twins are discordant even if the heritability is high (for example, schizophrenia, attention-deficit hyperactivity disorder, autism, multiple sclerosis or type 1 diabetes). From the dimensional view of disease liability, it also follows that despite striking differences in clinical appearance, discordant MZ twins can be quite similar in terms of underlying disease liability (FIGURE 1).

Trait concordance in MZ twins, penetrance and disease risk prediction. The presence of disease-discordant twins indicates that genomes cannot completely predict the disease outcome of individuals, even if most variation in disease outcome between individuals is caused by genetic differences. For example, for schizophrenia, despite the high heritability of 80%, the probandwise concordance between MZ co-twins is only 40–50%.
The fact that MZ twin concordance for common disorders is not always high has important implications for genomic risk prediction and the ethical concerns that have been raised in this light. Even if we knew all of the genetic variants that contribute to differences in disease risk between individuals, we would still not be able to predict with certainty the disease risk of all individuals on the basis of their DNA sequence.

**Conclusions**

Insights that can be obtained from twin studies extend far beyond the classical estimates of heritability. Traditional comparisons of the phenotypic resemblance of twins have been extended to studies of molecular variation across biological samples, providing functional insights into the underlying biology of heritable traits. The study of discordant MZ twins is a powerful method to identify DNA sequence variants, epigenetic variation and metabolites that are associated with disease.

One might feel that there are few aspects of the human condition that have not been investigated in twins; however, new aspects emerge all the time. We have emphasized the value of twin studies in refining phenotypic and clinical definitions and to evaluate biomarkers for disease, but the use of twins can go even further. In recent years, political scientists, sociologists and even economists have become engaged in twin studies. A study of MZ twins who were infected with HIV through blood transfusion at birth but who had strikingly different clinical outcomes used the identical genetic background of twins as a model to study the evolutionary processes and population dynamics that shape viral diversity.

In the coming years, longitudinal phenotypic information coupled with biological material collected by worldwide twin registries (TABLE 1; Supplementary information S1 (table)) will be an important resource for large-scale molecular studies. To make optimal use of genetic data collected within twin registries, methods for family-based association analysis are being explored further. In recent years, political scientists, sociologists and even economists have become engaged in twin studies. With the increasing interest in rare genetic variants, there may be renewed interest in linkage studies, in which DZ twins can have an important role. Linkage analysis in DZ twins, contrary to the analysis of non-twin siblings, is not affected by age differences within pairs and is less likely to suffer from non-paternity. Next-generation sequencing across multiple tissues and cell types will facilitate the detection of genome-wide SNPs, CNVs and epigenetic variation in discordant twins at an unprecedented scale, suggesting that twins will continue to provide valuable insights to human genetics.

**Figure 1** Liability threshold model and disease discordance in monozygotic twins. The liability threshold model assumes that multifactorial diseases result from an underlying continuous character (that is, liability) that is normally distributed in the population [11]. If the combined effects of genetic and environmental influences push an individual’s liability across a certain threshold level, the individual is affected. In the population, the proportion of individuals with a liability above the threshold is reflected in the disease prevalence. In discordant monozygotic (MZ) twin pairs, only one twin has a liability above the threshold, although the liability of the unaffected twin may also be high. The black arrow displays the potential range of liabilities of affected twins from discordant MZ twin pairs, and the grey arrow displays the potential range of liabilities of unaffected twins. A comparison of MZ twins who were discordant for congenital diaphragmatic hernia and oesophageal atresia found no differences in genomic structural variation between co-twins [15]. However, structural events in relevant genomic regions that may have contributed to the genetic predisposition of both twins were detected in several pairs; these events were rarely observed in individuals from a healthy control population.

A metabolomic study of MZ twins who were discordant for schizophrenia found that, relative to healthy individuals in concordant pairs, the unaffected twins from discordant female pairs showed similar (although smaller) metabolic changes than the affected co-twins [13]. These examples demonstrate that the liability of unaffected twins from discordant pairs may also be elevated. However, this feature does not argue against the value of studying discordant MZ twin pairs to search for the molecular events that caused the affected twin to pass the threshold or events that protected the unaffected twin. Of interest, a study of neurofibromatosis type 1 (NF1) in MZ twins with the same causal mutation in the NF1 gene but highly variable disease phenotypes revealed considerable variation between twins in DNA methylation at the NF1 gene [19].

8. Haworth, C. M. et al. The heritability of general cognitive ability increases linearly from childhood to
This paper describes a study based on a large sample of twins from six twin cohorts, showing that the heritability of general cognitive ability increases significantly from childhood to young adulthood.


This paper describes the implementation of GxE interaction tests in variance component analysis, by estimating unmeasured genetic effects as a linear function of one or more measures of the environment or moderators.


This review provides an overview of studies of MZ twins who are discordant for chromosomal abnormalities, Mendelian disorders and other genetic disorders.


This paper describes a study of a female MZ twin who are discordant for obsessive-compulsive, anxious and depressive symptoms in a population-based sample.


This paper describes a study on how epigenetic factors related to the lifetime of monozygotic twins. Hum. Mol. Genet. 21, 416–418 (2012)

This is the first study indicating that epigenetic changes are responsible for MZ twin discordance. A study on the Swedish Twin Registry.


This paper describes an association study on the Swedish Twin Registry.


This paper describes a study of monozygotic twins discordant for childhood-onset type 1 diabetes.


This paper describes a study on the Swedish Twin Registry.


This paper describes a study of the human neonatal epigenome.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.


This paper describes a study of how epigenetic differences arise during fetal growth and development.
In investigations into the influence of host genetics on the development of inflammatory bowel disease phenotypes, microbiome composition and activity play a pivotal role. The interplay between the gut microbiota and host genomic structures is critical for the development of these conditions. For example, studies from the Turnbaugh lab (Turnbaugh, P.J., et al., Nature, 354, 360–366, 2005) have demonstrated the importance of gut microbial diversity in the etiology of inflammatory bowel disease. Moreover, the role of host genetics in determining the expression of these microbial contributions is becoming increasingly evident. For instance, Pietiläinen et al. (J. Med. Genet., 47, 561–566, 2010) found that genetic factors influence the expression of inflammatory bowel disease susceptibility markers.

Additionally, the study of host-genetic influences on the gut microbiome is revealing fascinating insights into the mechanisms underlying the emergence of metabolic disorders. The work of Pietiläinen et al. (J. Med. Genet., 47, 561–566, 2010), for example, indicated that genetic factors influence the expression of inflammatory bowel disease susceptibility markers. These findings emphasize the importance of understanding the genetic basis of these complex traits.

Furthermore, the analysis of genetic factors in the gut microbiome is crucial for understanding the etiology and pathogenesis of metabolic disorders. The work of Pietiläinen et al. (J. Med. Genet., 47, 561–566, 2010) indicated that genetic factors influence the expression of inflammatory bowel disease susceptibility markers. These findings emphasize the importance of understanding the genetic basis of these complex traits.

In conclusion, the study of host-genetic influences on the gut microbiome is revealing fascinating insights into the mechanisms underlying the emergence of metabolic disorders. The analysis of genetic factors in the gut microbiome is crucial for understanding the etiology and pathogenesis of metabolic disorders. The work of Pietiläinen et al. (J. Med. Genet., 47, 561–566, 2010) indicated that genetic factors influence the expression of inflammatory bowel disease susceptibility markers. These findings emphasize the importance of understanding the genetic basis of these complex traits.


Acknowledgements

This work was supported by the European Research Council (ERC 230374) and the Institute for Health and Care Research (EMGO +).

Competing interests statement

The authors declare no competing financial interests.

FURTHER INFORMATION

Dorret I. Boomsma’s homepage: [http://www.tweelingen-register.org](http://www.tweelingen-register.org)


Netherlands Consortium for Healthy Ageing: [http://www.healthy-ageing.nl](http://www.healthy-ageing.nl)


SUPPLEMENTARY INFORMATION

See online article: S1 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF