

Syddansk Universitet

## Twin methodology in epigenetic studies

Tan, Qihua; Christiansen, Lene; von Bornemann Hjelmberg, Jacob; Christensen, Kaare

*Published in:*  
Journal of Experimental Biology

*DOI:*  
[10.1242/jeb.107151](https://doi.org/10.1242/jeb.107151)

*Publication date:*  
2015

*Document version*  
Publisher's PDF, also known as Version of record

*Citation for published version (APA):*  
Tan, Q., Christiansen, L., von Bornemann Hjelmberg, J., & Christensen, K. (2015). Twin methodology in epigenetic studies. *Journal of Experimental Biology*, 218(1), 134-139. <https://doi.org/10.1242/jeb.107151>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## REVIEW

## Twin methodology in epigenetic studies

Qihua Tan<sup>1,2,\*</sup>, Lene Christiansen<sup>1,2</sup>, Jacob von Bornemann Hjelmberg<sup>1</sup> and Kaare Christensen<sup>1,2</sup>**ABSTRACT**

Since the final decades of the last century, twin studies have made a remarkable contribution to the genetics of human complex traits and diseases. With the recent rapid development in modern biotechnology of high-throughput genetic and genomic analyses, twin modelling is expanding from analysis of diseases to molecular phenotypes in functional genomics especially in epigenetics, a thriving field of research that concerns the environmental regulation of gene expression through DNA methylation, histone modification, microRNA and long non-coding RNA expression, etc. The application of the twin method to molecular phenotypes offers new opportunities to study the genetic (nature) and environmental (nurture) contributions to epigenetic regulation of gene activity during developmental, ageing and disease processes. Besides the classical twin model, the case co-twin design using identical twins discordant for a trait or disease is becoming a popular and powerful design for epigenome-wide association study in linking environmental exposure to differential epigenetic regulation and to disease status while controlling for individual genetic make-up. It can be expected that novel uses of twin methods in epigenetic studies are going to help with efficiently unravelling the genetic and environmental basis of epigenomics in human complex diseases.

**KEY WORDS:** Twins, Epigenomics, Complex phenotypes**Introduction**

Human twins are produced when one pregnancy results in two offspring. A pair of identical twins is developed when one oocyte is fertilized by a single sperm (monozygotic, MZ) and the embryo splits into two, resulting in two genetically identical offspring. In contrast, a pair of fraternal twins is produced when two oocytes (dizygotic, DZ) are released at a single ovulation and fertilized by two different sperm at the same time, resulting in the same type of genetic relationship as siblings. The sharing of intrauterine and rearing environments and genetic similarity and dissimilarity in twins make them unique subjects for genetic studies. For example, by modelling phenotypic covariance in identical and fraternal twins and assuming equal sharing of the rearing environment, the classical twin design is able to decompose observed phenotype variation in twins into genetic (additive, dominant) and environmental (common and unique) components, enabling estimation of the genetic contribution to phenotype development or diseases (heritability) without knowledge of the individual genotypes. Although there can be critical flaws in the assumptions (Rijsdijk and Sham, 2002), the classical twin method has made a remarkable contribution to the literature on human genetics in the past century, especially in the last few decades.

<sup>1</sup>Epidemiology, Biostatistics and Bio-demography, Institute of Public Health, University of Southern Denmark, DK-5000, Odense C, Denmark. <sup>2</sup>Human Genetics, Institute of Clinical Research, University of Southern Denmark, DK-5000, Odense C, Denmark.

\*Author for correspondence (qtan@health.sdu.dk)

A significant change in 21st century genetics will be the shift from structural genetic variations, with genes regarded as a static concept, to functional genomics, where the dynamic patterns of gene activity are analysed jointly from gene interaction to gene regulation and functional genomics analysis (Peltonen and McKusick, 2001). Epigenetics is the study of changes in the regulation of gene activity and expression that are not dependent on gene sequence. In a broad sense, the epigenetic control over gene activity involves multiple molecular mechanisms including binding of small molecules to specific sites in DNA or chromatin, non-coding RNAs (ncRNAs), microRNAs, etc., all of which serve as ‘volume controls’ that tune up or down a gene’s expression. Of the different mechanisms, DNA methylation is the most robust form of epigenetic modification readily measurable using high-throughput techniques. Although current techniques allow epigenetic profiling at the genome scale, identifying the epigenetic patterns in disease conditions and under given environmental exposures imposes new challenges both in experimental design and in methodological issues.

Like any complex phenotype, the epigenetic control of gene expression activity can be influenced by both genetic and environmental factors. In fact, recent research showed that the impact of the environment can be acquired via the epigenome (Fraga et al., 2005; Wong et al., 2005; Poulsen et al., 2007; Szyf et al., 2008; Ling and Groop, 2009), a hot area in complex disease studies including cancer that draws active research. With the rapid development in epigenomic analysis using next-generation sequencing or array-based technologies, the newly emerging epigenetic epidemiology is serving as a bridge linking gene activity with environmental conditions. In this regard, twins are useful study subjects that can help with assessing the genetic and environmental influences in epigenetic regulations (Tan et al., 2013).

**Twin methods for epigenetic studies****The classical twin design**

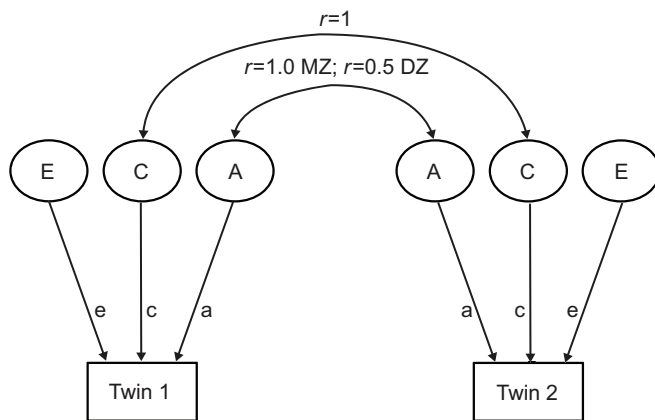
In the classical twin design, the observed resemblance in MZ twins versus DZ twins allows the decomposition of the phenotypic variation into that due to additive genetic (A), common or shared environmental (C), and unique environmental or residual (E) influences (Fig. 1), with A resulting from the sum of allelic effects across multiple genes; C from environmental influences shared by twin members, such as prenatal condition, home environment and family socioeconomic status; and E from influences that are unshared or unique, such as personal lifestyle and experiences, stochastic biological effects, as well as measurement error. Given the fact that MZ twins share all their genes, while DZ twins share on average half of their segregating genes and assuming that environmental influences are shared to the same extent by MZ and DZ twins (equal environment assumption, Fig. 1), the classical twin method is able to estimate the different variance components using phenotype data on twins. The genetic and environmental sharing in twins means that any phenotypic correlation between MZ twins ( $r_{MZ}$ ) can be attributed to their shared genetic make-up and their shared environmental influences, i.e.  $r_{MZ}=A+C$ . Likewise, any

**List of abbreviations**

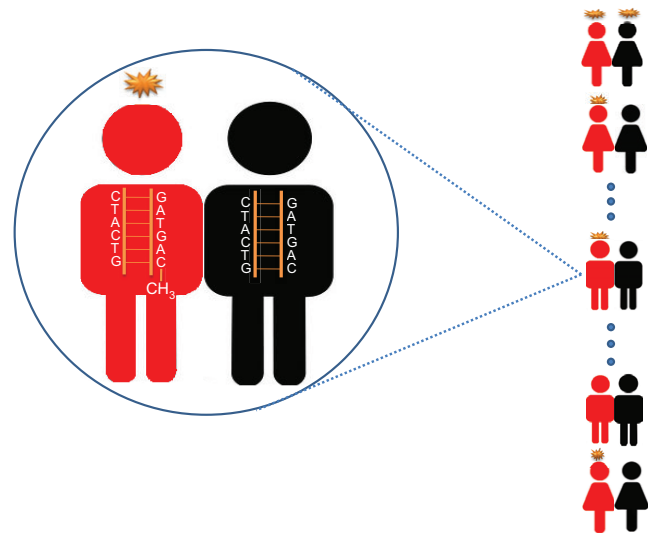
DZ	dizygotic
IBD	identical-by-descent
MZ	monozygotic
ncRNA	non-coding RNA
QTL	quantitative trait loci
$r_{DZ}$	correlation between MZ twins
$r_{MZ}$	correlation between MZ twins

observed correlation between DZ twins is due also to their shared genetic (50% on average) as well as environmental influences with  $r_{DZ}=0.5A+C$ . The variance of each twin is due to their genetic make-up, shared environmental influences and residual factors unique to each twin, i.e.  $A+C+E$ . The structural equation modelling (SEM) can be applied to estimate the A, C and E components, i.e. the ACE model, while adjusting for the effect of covariates (e.g. age), providing confidence intervals for parameter estimates and reporting performance of the model in fitting to the data. By calculating the Akaike information criterion (AIC) (Akaike, 1974), performance of the full ACE model can be compared with its nested models including the AE model (dropping the C component), the CE model (dropping the A component) and E model (dropping the A and C components). This enables selection of the best fitting and most parsimonious model for a given set of data. Instead of the popular ACE model, one can also fit an ADE model with D standing for the dominant genetic effect. The presence of dominance can be recognized when  $r_{DZ}$  is much lower than half  $r_{MZ}$ . Likewise, nested models (AE, DE and E) can be fitted and performances compared to find the best fitting model.

By treating epigenetic measurement such as the DNA methylation level at a CpG site as a quantitative phenotype, the classical twin model can be applied to estimate the genetic and environmental components in the epigenetic control of gene activity during, for example, human development and ageing, and in disease aetiology. Note that, in the ACE model, the C component reflects the common environment in their early lives from pregnancy to at least leaving home. In fact, the ACE model can be compared with its nested AE model to examine whether the shared early-life environmental conditions contribute to epigenetic modification later in life. In this case, the ACE model nicely fits the need for linking early-life environment with epigenetic status afterwards. The twin method can



**Fig. 1. Path diagram for the classical twin design.** The classical twin model decomposes phenotypic variance into genetic and environmental components under the assumption of equal environmental sharing in monozygotic (MZ) and dizygotic (DZ) twin pairs. A, additive genetic influences; C, common or shared environmental influences; E, unique environmental or residual influences.



**Fig. 2. The case co-twin design using identical twins discordant for a disease phenotype.** The similar genetic make-up and shared rearing environment make identical twin pairs perfectly matched on known (genetics, age) as well as unknown confounding factors. Left: the figure illustrates similar DNA sequence in an identical twin pair with one methylation site ( $-\text{CH}_3$  attached) occurring in the diseased twin (red) but not in the healthy twin (black), which can be associated with differential exposure indicated by a star above the diseased twin. Right: multiple pairs of identical twins can be collected with age and sex matched as well.

also be applied to epigenetic data collected longitudinally to look for dynamic patterns of genetic and early-life environmental regulation on epigenetic control over gene function during developmental and ageing processes. Besides the univariate twin model, multivariate approaches can also be applied to multiple genomic loci, e.g. CpG sites, to analyse whether they are co-regulated genetically and/or environmentally to give a deeper insight into the interactive epigenome.

**The case co-twin design**

The case co-twin design collects twin pairs discordant for a disease or trait and tries to link disease discordance with differential environmental exposure (Fig. 2). While the case co-twin design can be applied to either MZ or DZ twins, MZ pairs are preferable because they are genetically identical, leaving discordance within pairs being of environmental origin. By focusing on identical twins, the case co-twin design is especially useful in epigenetic studies as one of the main tasks in these studies is to find environmental exposures that are associated with the observed epigenetic changes linked to disease status. Here, the healthy co-twins serve as an ideal control group. MZ twin pairs share the same genetic composition, and they may also share a common rearing environment during their childhood and adolescent years. As such, MZ pairs are perfectly matched on a multitude of known (genetics, age) and unknown potential confounding factors. This means that the case co-twin design is deemed to have higher power than the ordinary case-control design used in most of the current epigenome-wide association studies.

In the literature, the case co-twin design has already been applied in recent epigenetic studies (Table 1) although statistical methods for analysing the data produced are currently demanding. Very recently, we have proposed a mixed model that models the intra-pair differential DNA methylation as a function of environmental exposure, providing adjustment for confounding factors as fixed

**Table 1. Summary of epigenetic studies using MZ twin pairs discordant for disease or trait**

Study	Year of publication	Sample size (pairs)	Coverage	Laboratory method	Statistical method	Discordant trait	Finding
Weksberg et al.	2002	10	<i>KCNQ1OT1</i> gene	Southern blotting	Descriptive	BWS	Loss of imprinting predisposes to BWS
Petronis et al.	2003	2	<i>DRD2</i> gene	Bis-seq	Descriptive	Schizophrenia	High epigenetic similarity in affected twins
Oates et al.	2006	1	<i>AXIN1</i> gene	Bis-seq	Descriptive	Caudal duplication	High methylation at promoter region of <i>AXIN1</i> gene in affected twin
Mill et al.	2006	12	<i>COMT</i> gene	Bis-seq	Descriptive	Birth weight	Epigenetic variation in MZ twins may play a key role in the etiology of psychopathology
Kuratomi et al.	2008	1	Genome-wide	MS-RDA	Descriptive	Bipolar	Altered DNA methylation of PPIEL is associated with bipolar
Kaminsky et al.	2008	1	Genome-wide	Array	Descriptive	Risk-taking behaviour	Some DNA methylation differences may have developmental and behavioural implications
Mastroeni et al.	2009	1	Genome-wide	Immunohistochemistry	<i>t</i> -test	Alzheimer's disease	Low levels of DNA methylation observed in temporal neocortex neuronal nuclei of diseased twin
Javierre et al.	2010	5	Genome-wide	Array	Paired <i>t</i> -test	Systemic lupus erythematosus	Twin discordance is associated with widespread changes in DNA methylation
Baranzini et al.	2010	3	Genome-wide	Array	Descriptive	Multiple sclerosis	Very few changes in CpG methylation between siblings
Nguyen et al.	2010	3	Genome-wide	Array	SAM	Autism	Epigenetic contributions to autism confirmed
Harder et al.	2010	8	NF1 regions	Bis-seq	Paired <i>t</i> -test	NF1	DNA methylation of normal NF1 allele responsible for modification of NF1 phenotype.
Tierling et al.	2011	1	Candidate regions	Bis-seq	Descriptive	BWS	KvDMR1 was exclusively hypomethylated in all cell types of the affected BWS twin
Souren et al.	2011	8	Candidate regions	Bis-seq	Paired <i>t</i> -test	BMI	DNA methylation variability at nine regions does not contribute to discordancy in BMI
Dempster et al.	2011	22	Genome-wide	Infinium array	Paired <i>t</i> -test	Bipolar	Further evidence for DNA methylation in the etiology of major psychoses
Gervin et al.	2012	27	Genome-wide	Infinium array	Paired <i>t</i> -test	Psoriasis	Epigenetic alteration potentially contributes to the development of psoriasis
Runyon et al.	2012	21	Selected loci	Bis-seq	Paired <i>t</i> -test	Asthma	Differential function of T cell subsets is regulated by changes in DNA methylation
Souren et al.	2013	17	Genome-wide	Infinium array	Rank test	Birth weight	No distinguishable genome-wide DNA methylation differences due to birth weight discordance
Wong et al.	2014	50	Genome-wide	Infinium array	Paired <i>t</i> -test	Autism	Numerous differentially methylated regions associated with autism

MZ, monozygotic; BWS, Beckwith–Wiedemann syndrome; NF1, neurofibromatosis type 1.

effects and taking into account random effects such as batch effect, arrangement of samples on the array, etc. In brief, the model takes the following form:

$$\log \left[ \frac{\text{Me}(+)}{\text{Me}(-)} \right] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + 1 | y_1 + 1 | y_2 + \dots + 1 | y_m . \quad (1)$$

Here, Me(+) and Me(−) are DNA methylation levels measured in the affected and unaffected twins of the same pair;  $\beta$  is the slope for fixed effect variables ( $x_1$  to  $x_n$ ), and  $y_1$  to  $y_m$  are the random effect variables including batch effects. In this model, the intercept  $\alpha$  is an important parameter because it stands for the mean fold-change in DNA methylation level between affected and unaffected twins. By testing the null hypothesis of  $\alpha=0$ , we are able to test whether the disease is associated with hyper- ( $\alpha>0$ ) or hypo-methylation ( $\alpha<0$ )

at the CpG site being tested. Except for the twin pair-specific variables, e.g. age and sex, the rest of the fixed effects are for differential exposure of specific environments. If the  $\beta$  of a differential exposure variable is significantly different from zero, the corresponding environment causes hyper- ( $\beta > 0$ ) or hypomethylation ( $\beta < 0$ ) at the CpG site in the affected twins, establishing the link from environmental exposure to epigenetic regulation and to disease. Note, one could argue that the effects of age and sex are matched out in the co-twin design. In epigenetic studies, however, this argument does not hold because within-twin pair difference in epigenetic measurement can be higher in old than in young twin pairs, and for some specific sites within-twin pair epigenetic difference can differ according to sex. In our regression-based mixed model, these can be easily included as pair-specific covariates with fixed effects and their effects adjusted.

#### Quantitative trait loci (QTL) mapping using dizygotic twins

Current array-based high throughput genotyping techniques, such as Illumina and Affymetrix and, more recently, next-generation sequencing, are revolutionizing the way we design and conduct genetic epidemiological studies. High resolution genomic analysis enabled by the high density and informative SNP markers offers researchers efficient tools to map QTL linked to human complex phenotypes including molecular phenotypes such as gene expression activity and DNA methylation levels.

DZ twin pairs are in fact ordinary siblings. By estimating phenotype-dependent allele-sharing identical-by-descent (IBD) within a pair, the non-parametric linkage analysis can be performed using variance component analysis (Kruglyak and Lander, 1996). Similar to disease phenotypes, the measured molecular phenotypes (e.g. DNA methylation levels) can be regarded as a quantitative trait on which non-parametric linkage analysis can be conducted. We emphasize that linkage scans using DZ twins take advantage of the fact that DZ twins are matched for pre-natal and shared rearing environmental factors and perfectly matched for their ages. These factors could affect the measured epigenetic status and thus need to be adjusted in the linkage analysis. The twin-based QTL linkage mapping gains further advantage from the low probability of non-paternity. Non-paternity can result in a biased estimation of IBD probability, leading to incorrect linkage results in non-parametric linkage analysis (MacGregor et al., 2000). In fact, DZ twins have been used in linkage mapping of gene expression QTLs (De Moor et al., 2007; Livshits et al., 2007; Perola et al., 2007; O'Connor et al., 2008). Similarly, methods using sib-pairs (i.e. a pair of siblings) for association mapping such as S-TDT (Spielman and Ewens, 1998) and QTDT (Abecasis et al., 2000; Ewens et al., 2008) apply to molecular data on DZ twins as well.

In fact, both linkage (Schadt et al., 2003; Morley et al., 2004; Monks et al., 2004) and association (Cheung et al., 2005) mapping have been used to look for genomic regions that are under epigenetic control of transcriptional activities through *cis* (local) or *trans* (distant) acting regulation. Although gene expression can be correlated with epigenetic status, direct application of linkage and association analysis to epigenetic phenotypes (DNA methylation, histone acetylation, ncRNA expression) should produce more efficient and specific results as the measured mRNA levels can be affected by multiple epigenetic mechanisms mentioned above.

#### Twins for epigenetic study of human diseases

Changes in gene expression resulting from global epigenetic regulation triggered by genetic, environmental and stochastic effects accumulated over time (Fraga et al., 2005) can contribute to the

development of individual susceptibility to complex diseases (Poulsen et al., 2007; Gilbert, 2009; Gröniger et al., 2010; Holliday, 2010; Bocklandt et al., 2011), adding molecular evidence that nature and nurture are inextricably linked. Analysing the complex interactive relationship between gene and environment in the development of complex phenotypes poses a new challenge.

As mentioned above, identical twins are useful study subjects in epigenetic studies of complex phenotypes because of their perfectly matched genetic background (Fig. 2) (Petronis, 2006; Poulsen et al., 2007). Table 1 lists the published articles on epigenetic studies using identical twins discordant for diseases or traits. As shown in Table 1, the scope of diseases covered in the application is still very limited. Current patterns of diseases in developed and developing countries is dominated by non-communicable diseases such as cardiovascular diseases, diabetes and cancer, for which environmental factors play an important role in disease development and pathology. Given the complex nature of common diseases, the use of twin design in epigenetic studies of complex diseases should be highly encouraged. The powerful case co-twin design in combination with recent high-throughput techniques for epigenomic analysis should help to identify important environmental risk factors that are behind the observed epigenetic alteration responsible for disease or health end points.

#### Twins for epigenetic studies of human ageing and development

Ageing is an inevitable process for everyone but we undergo it at our own time and pace. Though genetics could initially determine how a person ages, over time, environment could eventually play a higher role during the ageing process.

Studies have shown that the early-life environmental conditions are important modifiers of the ageing process (Szyf, 2012; Szyf, 2009; Murgatroyd and Spengler, 2011). As the genetic regulation of the ageing process is not directly controllable, studying the environmental factors that retard or accelerate it can be of more practical impact. As the first step towards this, it is necessary to explore the relative importance of gene and environment, a typical topic in twin modelling. In ageing studies, the classical twin model has been applied to various ageing-related phenotypes including physical performance (Frederiksen et al., 2002) and cognitive ability (Greenwood et al., 2011; Reynolds et al., 2005) in older subjects. Tan et al. (Tan et al., 2005) applied the twin method to gene expression data from older Danish twins and estimated heritability for the expression levels of the most active genes. Subsequently, Kaminsky et al. (Kaminsky et al., 2009) studied global DNA methylation profiles in MZ and DZ twins and reported highly significant epigenetic differences, suggesting both genetic and environmental influences on epigenetic regulation of gene expression. As a matter of fact, the impact of early-life milieu (including intrauterine life) on the epigenome has been explored in both animal and human studies with results linked to the development of stress (Murgatroyd and Spengler, 2011), obesity (Lillycrop and Burdge, 2011), diabetes (Fradin and Bougnères, 2011) and late-onset mental illness (Szyf et al., 2007; McGowan and Szyf, 2010). In a multicellular organism, many of the effects of differential gene expression in determining the structural and functional differentiation of cells arise during development. Epigenetic modifications triggered by early-life events represent a plausible mechanism by which constant exposure to a certain early-life environment could be integrated into the epigenome to programme adult hormonal and behavioural responses, facilitating adaptation to changing environmental conditions through

modification or alteration in gene activity. In this regard, identical twins discordant for birth weight are ideal subjects for studying the impact of very early life adverse environment on epigenetic regulation of health at adult ages as the genetic part of epigenetic control is cancelled out, enabling pure association between environmental exposure and epigenetic status. Using a small sample of 17 pairs of MZ twins discordant for birth weight, Souren et al. (Souren et al., 2013) reported no significant intra-pair differential DNA methylation pattern in a genome-wide analysis. We are currently working on whole-genome DNA methylation data on a relatively large sample of 150 pairs of identical twins discordant for birth weight. The observed epigenetic patterns in the discordant twins will be linked to metabolic phenotypes to look for epigenetic modification induced by birth weight discordance and its association with changes in metabolic profiles. Our results will also serve as reliable replicates to the findings of Souren et al. (Souren et al., 2013).

The ageing process is accompanied by various altered phenotypes that might be linked to epigenetic change. In the literature, a growing number of reports have demonstrated that the DNA methylation is an active process in post-mitotic cells and that the epigenome matures continuously throughout life with the effects of ageing emerging gradually over time as a result of exposure to a variety of environmental factors. For instance, Tan et al. (Tan et al., 2008) reported age-dependent patterns in gene expression in the Utah CEPH (Centre d'Etude du Polymorphisme Humain) families. Although the observed age-associated change in gene expression can be induced by both epigenetic regulation and genetic alterations accumulated from DNA damage, it can be postulated that a considerable portion of the changes can be explained by epigenetic mechanisms. In another example, Fraga et al. (Fraga et al., 2005) reported more epigenetic similarity in younger compared with older MZ twins, which suggested epigenetic regulation induced through accumulated stimulation by internal and external factors including environmental exposure. By treating gene expression or DNA methylation levels as molecular phenotypes, the classical twin design can be applied at different ages to explore the age-dependent patterns in the genetic and environmental contribution to epigenetic modification of gene activity, which can be linked to ageing-related phenotypes (e.g. physical and cognitive decline) and diseases.

## Conclusions

Most complex traits (diseases, molecular phenotypes) are under the control of both genetic and environmental factors through the interface of epigenetics. Although current techniques already allow high-throughput epigenetic profiling at the genome scale, our understanding of genetic and environmental influences on epigenetic regulation remains limited. With proper study design and analytical approaches, twin studies can help with the identification of novel epigenetic marks and the linking of these with environmental exposures including early-life experiences for measuring their impact on epigenetic control over gene activity, using classical twin models, linkage analysis, case co-twin design, etc. It is highly expected that the valuable sample of twins is going to make new contributions in unravelling and understanding the epigenetic basis of the development of human complex diseases and traits.

## Acknowledgements

The authors thank Prof. Torben Kruse and Mads Thomassen at Odense University Hospital for useful discussions.

## Competing interests

The authors declare no competing financial interests.

## Author contributions

Q.T. drafted and revised the paper; K.C., L.C. and Q.T. formulated the conception of the paper; J.V.B.H. contributed to twin methods and statistical modelling. All authors contributed to the ideas presented.

## Funding

This work was supported jointly by the European Foundation for the Study of Diabetes (EFSD) 2013 Programme for Collaborative Research between China and Europe; the Medical and Natural Sciences Research Grant of the Novo Nordisk Foundation project no. 7493; the Integrated research on DEvelopmental determinants of Aging and Longevity (IDEAL), an EU's FP7 project no. 259679.

## References

- Abecasis, G. R., Cardon, L. R. and Cookson, W. O. (2000). A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* **66**, 279-292.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716-723.
- Baranzini, S. E., Mudge, J., van Velkinburgh, J. C., Khankhanian, P., Khrebtkova, I., Miller, N. A., Zhang, L., Farmer, A. D., Bell, C. J., Kim, R. W. et al. (2010). Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* **464**, 1351-1356.
- Bocklandt, S., Lin, W., Sehl, M. E., Sánchez, F. J., Sinsheimer, J. S., Horvath, S. and Vilain, E. (2011). Epigenetic predictor of age. *PLoS ONE* **6**, e14821.
- Cheung, V. G., Spielman, R. S., Ewens, K. G., Weber, T. M., Morley, M. and Burdick, J. T. (2005). Mapping determinants of human gene expression by regional and genome-wide association. *Nature* **437**, 1365-1369.
- De Moor, M. H., Spector, T. D., Cherkas, L. F., Falchi, M., Hottenga, J. J., Boomsma, D. I. and De Geus, E. J. (2007). Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. *Twin Res. Hum. Genet.* **10**, 812-820.
- Dempster, E. L., Pidsley, R., Schalkwyk, L. C., Owens, S., Georgiades, A., Kane, F., Kalidindi, S., Picchioni, M., Kravariti, E., Touloupoulou, T. et al. (2011). Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. *Hum. Mol. Genet.* **20**, 4786-4796.
- Ewens, W. J., Li, M. and Spielman, R. S. (2008). A review of family-based tests for linkage disequilibrium between a quantitative trait and a genetic marker. *PLoS Genet.* **4**, e1000180.
- Fradin, D. and Bougnères, P. (2011). T2DM: why epigenetics? *J. Nutr. Metab.* **2011**, 647514.
- Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J. et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. USA* **102**, 10604-10609.
- Frederiksen, H., Gaist, D., Petersen, H. C., Hjelmborg, J., McGue, M., Vaupel, J. W. and Christensen, K. (2002). Hand grip strength: a phenotype suitable for identifying genetic variants affecting mid- and late-life physical functioning. *Genet. Epidemiol.* **23**, 110-122.
- Gervin, K., Vigeland, M. D., Mattingsdal, M., Hammer, M., Nygård, H., Olsen, A. O., Brandt, I., Harris, J. R., Undlien, D. E. and Lyle, R. (2012). DNA methylation and gene expression changes in monozygotic twins discordant for psoriasis: identification of epigenetically dysregulated genes. *PLoS Genet.* **8**, e1002454.
- Gilbert, S. F. (2009). Ageing and cancer as diseases of epigenesis. *J. Biosci.* **34**, 601-604.
- Greenwood, T. A., Beer, M. S., Schmeidler, J., Valerio, D., Raventós, H., Mora-Villalobos, L., Camacho, K., Carrión-Baralt, J. R., Angelo, G., Almay, L. et al. (2011). Heritability of cognitive functions in families of successful cognitive aging probands from the Central Valley of Costa Rica. *J. Alzheimers Dis.* **27**, 897-907.
- Gröniger, E., Weber, B., Heil, O., Peters, N., Stäb, F., Wenck, H., Korn, B., Winnefeld, M. and Lyko, F. (2010). Aging and chronic sun exposure cause distinct epigenetic changes in human skin. *PLoS Genet.* **6**, e1000971.
- Harder, A., Titze, S., Herbst, L., Harder, T., Guse, K., Tinschert, S., Kaufmann, D., Rosenbaum, T., Mautner, V. F., Windt, E. et al. (2010). Monozygotic twins with neurofibromatosis type 1 (NF1) display differences in methylation of NF1 gene promoter elements, 5' untranslated region, exon and intron 1. *Twin Res. Hum. Genet.* **13**, 582-594.
- Holliday R. (2010). Perspectives in aging and epigenetics. *Epigenetics of Aging* **2010**, 447-455.
- Javierre, B. M., Fernandez, A. F., Richter, J., Al-Shahrour, F., Martin-Subero, J. I., Rodriguez-Ubreva, J., Berdasco, M., Fraga, M. F., O'Hanlon, T. P., Rider, L. G. et al. (2010). Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res.* **20**, 170-179.
- Kaminsky, Z., Petronis, A., Wang, S. C., Levine, B., Ghaffar, O., Floden, D. and Feinstein, A. (2008). Epigenetics of personality traits: an illustrative study of identical twins discordant for risk-taking behavior. *Twin Res. Hum. Genet.* **11**, 1-11.
- Kaminsky, Z. A., Tang, T., Wang, S. C., Ptak, C., Oh, G. H., Wong, A. H., Feldcamp, L. A., Virtanen, C., Halfvarson, J., Tysk, C. et al. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nat. Genet.* **41**, 240-245.
- Kruglyak, L. and Lander, E. S. (1995). Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.* **57**, 439-454.
- Kuratomi, G., Iwamoto, K., Bundo, M., Kusumi, I., Kato, N., Iwata, N., Ozaki, N. and Kato, T. (2008). Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. *Mol. Psychiatry* **13**, 429-441.

- Lillycrop, K. A. and Burdge, G. C. (2011). Epigenetic changes in early life and future risk of obesity. *Int. J. Obes.* **35**, 72-83.
- Ling, C. and Groop, L. (2009). Epigenetics: a molecular link between environmental factors and type 2 diabetes. *Diabetes* **58**, 2718-2725.
- Livshits, G., Kato, B. S., Wilson, S. G. and Spector, T. D. (2007). Linkage of genes to total lean body mass in normal women. *J. Clin. Endocrinol. Metab.* **92**, 3171-3176.
- MacGregor, A. J., Snieder, H., Schork, N. J. and Spector, T. D. (2000). Twins. Novel uses to study complex traits and genetic diseases. *Trends Genet.* **16**, 131-134.
- Mastroeni, D., McKee, A., Grover, A., Rogers, J. and Coleman, P. D. (2009). Epigenetic differences in cortical neurons from a pair of monozygotic twins discordant for Alzheimer's disease. *PLoS ONE* **4**, e6617.
- McGowan, P. O. and Szyf, M. (2010). The epigenetics of social adversity in early life: implications for mental health outcomes. *Neurobiol. Dis.* **39**, 66-72.
- Mill, J., Dempster, E., Caspi, A., Williams, B., Moffitt, T. and Craig, I. (2006). Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **141B**, 421-425.
- Monks, S. A., Leonardson, A., Zhu, H., Cundiff, P., Pietrusiak, P., Edwards, S., Phillips, J. W., Sachs, A. and Schadt, E. E. (2004). Genetic inheritance of gene expression in human cell lines. *Am. J. Hum. Genet.* **75**, 1094-1105.
- Morley, M., Molony, C. M., Weber, T. M., Devlin, J. L., Ewens, K. G., Spielman, R. S. and Cheung, V. G. (2004). Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**, 743-747.
- Murgatroyd, C. and Spengler, D. (2011). Epigenetics of early child development. *Front. Psychiatry* **2**, 16.
- Nguyen, A., Rauch, T. A., Pfeifer, G. P. and Hu, V. W. (2010). Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. *FASEB J.* **24**, 3036-3051.
- O'Connor, D. T., Zhu, G., Rao, F., Taupenot, L., Fung, M. M., Das, M., Mahata, S. K., Mahata, M., Wang, L., Zhang, K. et al. (2008). Heritability and genome-wide linkage in US and Australian twins identify novel genomic regions controlling chromogranin a: implications for secretion and blood pressure. *Circulation* **118**, 247-257.
- Oates, N. A., van Vliet, J., Duffy, D. L., Kroes, H. Y., Martin, N. G., Boomsma, D. I., Campbell, M., Coulthard, M. G., Whitelaw, E. and Chong, S. (2006). Increased DNA methylation at the AXIN1 gene in a monozygotic twin from a pair discordant for a caudal duplication anomaly. *Am. J. Hum. Genet.* **79**, 155-162.
- Peltonen, L. and McKusick, V. A. (2001). Genomics and medicine. Dissecting human disease in the postgenomic era. *Science* **291**, 1224-1229.
- Perola, M., Sarnalisto, S., Hiekkalinna, T., Martin, N. G., Visscher, P. M., Montgomery, G. W., Benyamin, B., Harris, J. R., Boomsma, D., Willemsen, G. et al.; GenomEUtwin Project (2007). Combined genome scans for body stature in 6,602 European twins: evidence for common Caucasian loci. *PLoS Genet.* **3**, e97.
- Petronis, A. (2006). Epigenetics and twins: three variations on the theme. *Trends Genet.* **22**, 347-350.
- Petronis, A., Gottesman, I. I., Kan, P., Kennedy, J. L., Basile, V. S., Paterson, A. D. and Popenklyte, V. (2003). Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance? *Schizophr. Bull.* **29**, 169-178.
- Poulsen, P., Esteller, M., Vaag, A. and Fraga, M. F. (2007). The epigenetic basis of twin discordance in age-related diseases. *Pediatr. Res.* **61**, 38R-42R.
- Reynolds, C. A., Finkel, D., McArdle, J. J., Gatz, M., Berg, S. and Pedersen, N. L. (2005). Quantitative genetic analysis of latent growth curve models of cognitive abilities in adulthood. *Dev. Psychol.* **41**, 3-16.
- Rijsdijk, F. V. and Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Brief. Bioinform.* **3**, 119-133.
- Runyon, R. S., Cachola, L. M., Rajeshuni, N., Hunter, T., Garcia, M., Ahn, R., Lurmann, F., Krasnow, R., Jack, L. M., Miller, R. L. et al. (2012). Asthma discordance in twins is linked to epigenetic modifications of T cells. *PLoS ONE* **7**, e48796.
- Schadt, E. E., Monks, S. A., Drake, T. A., Lusk, A. J., Che, N., Colino, V., Ruff, T. G., Milligan, S. B., Lamb, J. R., Cavet, G. et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**, 297-302.
- Souren, N. Y., Tierling, S., Fryns, J. P., Derom, C., Walter, J. and Zeegers, M. P. (2011). DNA methylation variability at growth-related imprints does not contribute to overweight in monozygotic twins discordant for BMI. *Obesity (Silver Spring)* **19**, 1519-1522.
- Souren, N. Y., Lutsik, P., Gasparoni, G., Tierling, S., Gries, J., Riemenschneider, M., Fryns, J. P., Derom, C., Zeegers, M. P. and Walter, J. (2013). Adult monozygotic twins discordant for intra-uterine growth have indistinguishable genome-wide DNA methylation profiles. *Genome Biol.* **14**, R44.
- Spielman, R. S. and Ewens, W. J. (1998). A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am. J. Hum. Genet.* **62**, 450-458.
- Szyf, M. (2009). The early life environment and the epigenome. *Biochim. Biophys. Acta* **1790**, 878-885.
- Szyf, M. (2012). The early-life social environment and DNA methylation. *Clin. Genet.* **81**, 341-349.
- Szyf, M., Weaver, I. and Meaney, M. (2007). Maternal care, the epigenome and phenotypic differences in behavior. *Reprod. Toxicol.* **24**, 9-19.
- Szyf, M., McGowan, P. and Meaney, M. J. (2008). The social environment and the epigenome. *Environ. Mol. Mutagen.* **49**, 46-60.
- Tan, Q., Christensen, K., Christiansen, L., Frederiksen, H., Bathum, L., Dahlgard, J. and Kruse, T. A. (2005). Genetic dissection of gene expression observed in whole blood samples of elderly Danish twins. *Hum. Genet.* **117**, 267-274.
- Tan, Q., Zhao, J., Li, S., Christiansen, L., Kruse, T. A. and Christensen, K. (2008). Differential and correlation analyses of microarray gene expression data in the CEPH Utah families. *Genomics* **92**, 94-100.
- Tan, Q., Christiansen, L., Thomassen, M., Kruse, T. A. and Christensen, K. (2013). Twins for epigenetic studies of human aging and development. *Ageing Res. Rev.* **12**, 182-187.
- Tierling, S., Souren, N. Y., Reither, S., Zang, K. D., Meng-Hentschel, J., Leitner, D., Oehl-Jaschkowitz, B. and Walter, J. (2011). DNA methylation studies on imprinted loci in a male monozygotic twin pair discordant for Beckwith-Wiedemann syndrome. *Clin. Genet.* **79**, 546-553.
- Weksberg, R., Shuman, C., Caluseriu, O., Smith, A. C., Fei, Y. L., Nishikawa, J., Stockley, T. L., Best, L., Chitayat, D., Olney, A. et al. (2002). Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* **11**, 1317-1325.
- Wong, A. H., Gottesman, I. I. and Petronis, A. (2005). Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Hum. Mol. Genet.* **14**, R11-R18.
- Wong, C. C., Meaburn, E. L., Ronald, A., Price, T. S., Jeffries, A. R., Schalkwyk, L. C., Plomin, R. and Mill, J. (2014). Methyloomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Mol. Psychiatry* **19**, 495-503.