REVIEW ARTICLE

The use of the twin model to investigate the genetics and epigenetics of skin diseases with genomic, transcriptomic and methylation data

V. Bataille,†‡*, M. Lens,† T.D. Spector‡

†Twin Research and Genetic Epidemiology Unit, Kings College, London SE1 7EH, UK
‡Dermatology Department, West Herts NHS Trust, Herts, UK
*Correspondence: V. Bataille. E-mail: bataille@doctors.org.uk

Abstract

Twins have always fascinated medical research even before the discovery of DNA and the understanding of the differences between identical and non-identical twins. Dermatology with the benefit of being able to visualize phenotypes was one of the first specialties reporting on the fascinating concordance in identical (MZ) twins in the 1920s. Over the last 20 years, the heritability of skin diseases using twins has been clearly demonstrated, across a wide variety of traits including melanoma, polymorphic light eruption, psoriasis, eczema and acne. Other rarer diseases have also been shown to have a significant genetic basis such as lupus, sarcoidosis and lichen sclerosus. Following evidence of heritability for many skin disease the next step was Genome-Wide Association Studies (GWAS) which are uncovering new genes in large twin cohorts. The twin model is also ideal for the new field of epigenetics, investigating subtle differences in DNA methylation within discordant MZ pairs for a disease, as well as differences in CNVs. Twins are also valuable for examining differences in gene function via RNA expression in twins discordant for a skin trait or disease.

Received: 22 July 2011; Accepted: 16 December 2011

Conflict of interest

None declared.

The twin model

The twin study model has always been based on the comparison of monozygotic (MZ) twins who share almost 100% of their DNA sequences with dizygotic (DZ) twins who share on average 50% of their genetic polymorphisms. Although it is now apparent that MZ twins do not always have completely identical DNA sequences. These minor differences between identical twins are not really affecting the heritability estimates in the sense that one can estimate the relative influence of genetic factors on a disease. Twin studies allow estimating heritability by comparing concordance of a disease or trait between MZ and DZ pairs. Path analyses in the 1990s allowed to dissect several components of the variance for a trait or disease to estimate the relative contribution of genetic, shared and unique environmental factors on the variance. Many complex traits or diseases have a heritability of roughly 50–80%, meaning that 50–80% of the variance in the liability to the trait or diseases can be explained by genetic factors. The rest of the variance is explained by shared family environment and unique environmental factors. However, even for very heritable traits, MZ twins are often discordant for diseases. Progress made in high throughput genotyping, transcriptomics, methylation studies and RNA sequencing are now showing that some of the unexplained genetic effect may be explained by epigenetic processes that do not alter DNA sequence, but alter transcription or expression of proteins. Again, twins have proven to be invaluable to examine these epigenetic effects, as MZ twins with the same DNA sequence can be compared for epigenetics differences explaining trait or disease discordance.

Twin studies in dermatology

Dermatologists have an important place in twin research as they were the first to report on the similarities in many skin traits including naevi in MZ twins. Herman Werner Siemens, a German dermatologist, was one of the inventors of the twin study design. In 1924, he published ‘Zwillingspathologie’ (Twin Pathology) and introduced the widely used ‘classical twin method’. Siemens reported higher concordance in naevus counts in MZ twins compared with siblings amongst many other dermatological traits. Already in the 1920s, he commented ‘twins of one egg origin will be of immense value for geneticists and psychologists’. Siemens’ work was very instrumental in highlighting the scientific value of MZ and DZ twins with very detailed collection of various skin phenotypes, well before DNA was discovered. Twin experts recognize the contribution of Siemens in discovering the twin model.
that we still use today. In the following 50 years from the discovery of the twin model, studies were mainly focused on behavioural traits and psychiatric diseases. Not until the 1980s, twin studies restarted to look at non-behavioural traits and common diseases. Several teams have collected very large and well documented twin datasets including the UK Twin Registry (TwinsUK cohort), the Netherlands Twin Registry (NTR), the Queensland Twin Registry as well as national twin registries in Sweden, Denmark and Finland (Table 1).

### Naevi

Siemens was the first one who tried to determine the role of genes over environment in for the appearance of naevi. In 1991, almost seventy years after the reports by Siemens, a UK research group published the results of the study examining the concordance of naevus counts in 23 MZ and 22 DZ twin pairs. This study demonstrated a higher concordance in the number of common naevi in MZ twins compared with DZ twins, which suggested that total naevus count has a strong inherited basis. In 1999, the Queensland Twin Registry reported a significantly higher concordance for naevus counts in 153 MZ adolescent twins (0.94) compared with the concordance observed in 199 DZ twins (0.60). Analysis of linkage to a highly polymorphic marker D9S942, located adjacent to CDKN2A gene, a detected quantitative-trait-loci (QTL) effects accounting for 27% of variance in total naevus count. Longer alleles at D9S942 were associated with flat naevi, but not with raised naevi. In 2000, the UK Twin Registry reported on the heritability of naevus counts based on 127 MZ twin pairs and 323 DZ twin pairs; the intraclass correlation for total body naevus counts was 0.83 in MZ twins compared with 0.51 in DZ twins (Fig. 1). Using quantitative genetic analyses, this study showed the importance of age. The contribution of additive genetic factors on naevus count increased significantly with age, with more than 80% of the variance on naevus attributed to genes in twins aged 45 years or over compared with only 36% in those aged less than 45 years. Freckle counts were also highly heritable in the same study (Fig. 2). In 2005, another UK twin study examining gene–environment interactions for nevi was published included 103 MZ and 118 DZ twin pairs from Yorkshire and Surrey using the same protocol as in the large Australian study conducted by Zhu et al. Correlations in nevus density were higher in MZ pairs (0.94) than in DZ pairs (0.61). The reported heritability of 65% for total body naevus counts in the UK adolescent twins was comparable to the 68% heritability reported for the Australian adolescent twins.

The number of naevi decrease with age which is one of the most important confounders in naevi studies. Why naevi undergo a senescence process remains unknown, but studies investigating

### Table 1 Heritability estimates in skin

<table>
<thead>
<tr>
<th></th>
<th>Heritability (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freckle counts</td>
<td>91</td>
<td>9</td>
</tr>
<tr>
<td>Varicose veins</td>
<td>86</td>
<td>45</td>
</tr>
<tr>
<td>Polymorphic light eruption</td>
<td>84</td>
<td>43</td>
</tr>
<tr>
<td>Acne</td>
<td>81</td>
<td>31</td>
</tr>
<tr>
<td>Androgenic alopecia (female)</td>
<td>80</td>
<td>44</td>
</tr>
<tr>
<td>Naevus counts below age 45 (UK)</td>
<td>36</td>
<td>9</td>
</tr>
<tr>
<td>Naevus count above 45 years (UK)</td>
<td>84</td>
<td>9</td>
</tr>
<tr>
<td>Naevus counts in adolescents (UK)</td>
<td>65</td>
<td>10</td>
</tr>
<tr>
<td>Naevus counts in adolescents (Australia)</td>
<td>68</td>
<td>8</td>
</tr>
<tr>
<td>Eczema (adults)</td>
<td>61</td>
<td>23</td>
</tr>
<tr>
<td>Eczema (children)</td>
<td>90</td>
<td>28</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>66</td>
<td>36</td>
</tr>
</tbody>
</table>

![Figure 1](image1.png) Similar naevus types and distribution in a pair of MZ twins.

![Figure 2](image2.png) Freckle counts and photoageing in MZ twins.
senescence in melanocytes have shown that patients with an excess of naevi have a delayed senescence in vitro. This finding suggests that different genes may have an effect on the appearance of naevi in early adulthood compared with the senescence process which occurs from the age of 40 years onwards. It has been observed clinically that the senescence of naevi appears to be delayed in patients with multiple atypical naevi, especially in those belonging to susceptible melanoma families. The individuals with multiple atypical naevi also appear to have delayed photoageing with less solar elastosis and a lower prevalence of solar keratoses (VB personal observation). The UK Twins Registry already looking at the genetics of ageing took the opportunity to investigate whether an excess of naevi may be a marker of delayed ageing. A study of 1897 twins on whom white cell telomere length data were available was performed to establish the correlation between telomere length and naevus counts. Telomere length is regarded as a surrogate biological clock, as it slowly decreases with age until the cell undergoes apoptosis. The speed at which this process occurs is in part genetic. Naevus counts were found to be positively correlated with telomere length. Longer telomeres were observed in twins with an excess of naevi compared with those with few naevi. This equates to roughly 6 years of chronological ageing between twins with more than 100 naevi compared with those having less than 25 naevi. These findings suggest that the delayed senescence observed in melanocytes in patients with a large number of naevi may not be melanocyte-specific and may also affect other cell types. This may also explain why patients with an excess of naevi have delayed photoageing. Another study from the same group reported that twins with an excess of naevi have higher Bone Mineral Density (BMD) and are relatively protected against osteoporosis than those with few naevi. The results from this study support the hypothesis that the reduced ageing in these individuals affects other systems.

The increased heritability of an ageing phenotype with increasing age is supported by studies of longevity using twin models. These studies show that the genetic influence on longevity increases with age. The chance of reaching old age is becoming more concordant in MZ pairs with age (vB15). Genetic influences on lifespan are minimal prior to age 60, but increase thereafter (vB15). The Department of Twin Research & Genetic Epidemiology at Kings College in London, UK, will use telomere length assays and transcriptome sequencing integrated with genome-wide association studies in a systems genetics framework to infer causal interactions between telomere length, skin ageing and naevus counts. These datasets will provide novel insights to ageing process.

In contrast to naevi, cutaneous melanoma, like many other cancers, is not highly heritable. Part of this difference may be methodological, because the naevus as an intermediate phenotype is a continuous trait, whereas cancer is a dichotomous trait and rarer, and so, more difficult to study in very large numbers. The Finish Twin Cohort, as a population-based study, suggested that environmental and not hereditary effects are most important in the development of cutaneous malignancies. In 2009, the Queensland Twins Registry published results of a twin study determining the proportion of genetic and environmental factors influencing variation in liability to melanoma. Melanoma concordance was higher in MZ twins (4 of 27 pairs) than in DZ twins (3 of 98 pairs). MZ twins with a co-twin with melanoma were 9.8 more likely to have melanoma than a DZ pairs, where the risk was only 1.8. This study estimated that 55% of the variation in liability to melanoma was due to genetic influences.

With the advances in high throughput genotyping, twin data have now led to valuable genome-wide analyses due to large and well phenotyped cohorts. The UK Twin Registry published the first genome-wide linkage analysis on naevi counts in 2006 using 700 markers along the whole genome. This study identified regions linked to total body naevus counts on chromosome 9p21 and 5q31. In 2007, the Queensland Twin Registry reported linkage to regions on chromosome 2, 9, 8 and 17 for flat naevi using 796 microsatellite markers across the genome. In 2009, the UK Twin Registry published a genome-wide association study on a much denser marker set of over 300 000 markers using the Illumina Hap 300K duo Chip. A genome-wide association study looks at possible associations between half a million genetic markers across the whole genome and a disease and a trait. This led to the discovery of two new naevus genes, MTAP on chromosome 9p21 close to the CDKN2A locus and PLA2G6 on chromosome 22q13. These two genes were also found to be melanoma genes and were replicated in two case-control melanoma studies in the UK and Australia, respectively, as well as in melanoma families from the Genomel Melanoma Consortium. MTAP has already been implicated in other cancers such as pancreatic cancer, which is found to cluster in some melanoma families. The fact that MTAP lies close to the CDKN2A raises the question of possible interactions between these two genes. They appear to often be deleted together in many cancers, but MTAP has also been reported to act as an independent tumour suppressor gene. PLA2G6 has been shown to be involved in cell growth and apoptosis in various cancers.

**Eczema**

Eczema has also been investigated using the twin model. A population-based study conducted in Denmark involving 812 twin pairs showed that the concordance rate of atopic dermatitis was 0.72 in MZ and 0.23 in DZ twin pairs. The atopic phenotype is associated with increased IgE levels. In a study including 340 MZ and 533 DZ British adult female twins, Strachan et al. found that IgE levels were under significant heritable control with 60% of the variance explained by genetic factors. However, MZ twins are often discordant for atopic disorders, and so gene–environment interactions or geneti/epigenetic differences may explain discordant MZ pairs. In a study of 1480 Swedish twin pairs (7–9 year old), Lichtenstein and Svartengren showed that 39–76% of the
liability to eczema was due to genetic factors. In 2005, a Norwegian study based on 3334 twins aged 18–35 years reported a concordance of 54% for eczema and itch.35 In a population-based study, Lerbaek et al.36 showed that 41% of the variance in hand eczema in more than 4000 twins can be explained by genetic factors. Thomsen et al.27 published the results from a large twin study based on more than 11 515 twins in Denmark showing a heritability of 82% for atopic dermatitis with a sevenfold increased risk of atopic dermatitis for a co-twin of an MZ pair compared with a threefold increase for a co-twin of a DZ pair. The highest heritability estimate for eczema (approximately 90%) was reported in a study enrolling 8633 twins in Holland.29 A very high contribution of genetic factors in the pathogenesis of eczema found in this study may be explained by the fact that the twins were young children, and in children, it is possible that the environmental differences on the expression of the diseases may be less significant. Multicentre genome-wide association studies (GWAS) on eczema are currently under way.

Acne
Acne was first studied using the twin model in 1984 in a study including 930 twin pairs in the USA.29 A twin study was conducted in the UK measured sebum excretion and acne scores in an adolescent twins and found a higher concordance of sebum excretion in MZ twins compared with DZ twins.30 In a large twin study enrolling 458 pairs of MZ and 1099 pairs of DZ twins, Bataille et al.31 showed that acne is one the most heritable skin phenotypes with 81% of the variance in acne liability explained by additive genetic factors. Hormonal factors, reproductive factors, BMI, glucose levels and lipids were not found to be associated with acne. Acne is also a symptom of the Polycystic Ovary Syndrome (PCOS). The Netherlands Twin Register study including 1332 MZ and 1873 DZ twins showed that this syndrome is also influenced by genetic factors with a concordance for PCOS of 71% in MZ twins compared with 38% in the DZ twins.32 The UK Twin Registry has conducted a GWAS in twins with acne, but so far, no obvious hits have been reliably identified. The power of the study was an issue, as only several hundred cases were available compared with several thousand controls. A larger acne case-control–GWAS study based in the UK is underway, and the results are awaited.

Psoriasis
Many anecdotal case reports on the higher concordance for psoriasis and psoriatic arthritis in MZ twins were published in the literature since the 1950s. In 1978, a study based on MZ twins in Denmark not only showed that 18 of 32 MZ pairs were concordant for psoriasis which supports a genetic influence but also highlighted the issue of discordance of psoriasis in MZ pairs.33 The Queensland Twin Registry reported a concordance of 35% for psoriasis in MZ twins compared with 12% in DZ pairs.34 A study of psoriatic arthritis in 228 affected twins showed a higher concordance for the disease in MZ twins compared with DZ twins.35 Grijovski et al.36 found that the heritability for psoriasis was 66% in more than 8000 twins in Norway. Data from this study showed that zygosity and gender did not affect the disease prevalence reported at 4.2%. Zhang et al.37 conducted the first large GWAS study to identify susceptibility variants for psoriasis. Using a two-stage case-control design 1139 cases and 1132 controls were included with a replication set of 5721 cases and 7340 controls. They reported strong replication for two known susceptibility loci, MHC and IL12B, and identified a new susceptibility locus within the LCE gene cluster on 1q21. The most significant finding of another GWAS, carried out in the USA, with 1359 cases and 1400 controls, was SNP rs12191877, which is in tight linkage disequilibrium with HLA-Cw*0602, the consensus risk allele for psoriasis.38

Other skin diseases
Many skin diseases including very rare ones have been reported to be more correlated within MZ twins compared with DZ twins, and these include sarcoidosis, lupus, lichen sclerosus and atrophicus, albinism, vitiligo, Netherton syndrome amongst many.39–42 Polymorphic light eruption has also been shown to be under significant genetic influence with heritability of 84%.43 The UK Twin registry in collaboration with other groups identified new genes for male pattern baldness and female hair thinning.44 The same research group also found that the development of varicose veins is highly genetic with heritability of 86%.45

The future of the twin model
Missing heritability
Genetic studies of many common diseases have benefited from the twin model with heritability studies and pathways analyses. More recently, GWAS of well phenotyped twin cohorts have uncovered new genes and new pathways. Whilst it is clear that GWAS do not need to rely on the twin model, the collection of large twin data-sets with varied and detailed phenotypes has led to very valuable twin cohorts with genotypes. The hundreds of novel variants published over the last few years prove that GWAS do deliver. However, the GWAS studies published so far do not provide clues to the source of the vast majority of the genetic variance for these common disorders.46 The question in the last 2 years in genetics has therefore been ’where is the missing heritability?’.47 More than 500 GWAS have been carried out to date reporting on hundreds of SNPs associations, but cannot explain more than a small percentage of the genetic effect. For example, genes for naevi and melanoma have been discovered with the recent GWAS, but they only explain a small proportion of the variance in naevi and melanoma susceptibility. With a few exceptions such as macular degeneration, few diseases have explained more than 10% of the variance despite large studies including thousands of patients. There is still a large part of the heritability missing. This may be explained in part by common variants that are not detected by the current GWAS, as these may be in even greater number with even
lower effects, therefore needing higher number of cases and controls to produce some hits. More recent GWAS for height and weight have exceeded 150,000 patients, and these studies still explain less than 14% of variance for these two traits with hundreds of genes. Alternatively, there may be many rarer variants with moderate effects that will not be picked up by the current GWAS. Sequencing is a likely next step, which can accurately detect variants in the range of 0.1–10%, which, although currently expensive (over $2,000 per person), is reducing in price rapidly. Several twin cohorts, including the TwinsUK, will have whole genome sequence within the next year.

### Epigenetics

Epigenetics looking at methylation and histone modification is now a rapidly moving area of research in the genetics of complex traits. These epigenetic changes affect transcription and hence disease expression. The twin model has an invaluable role in epigenetic research that is reviewed in details elsewhere.47 Another DNA variation which is structural so by definition not epigenetics is the presence of CNVs. CNVs are stretches of DNA, tens or hundreds of base pair long, that are either deleted or duplicated and vary between individuals. CNVs could possibly explain part of the missing heritability, as they go undetected in GWAS, because they do not alter DNA sequences. Techniques such as comparative genomic hybridization can identify CNVs by hybridizing DNA to microarrays. The small CNVs from 2 to 50 m base pairs will always be very difficult to pick up and may also alter disease expression. However, recent studies suggest that common large CNVs are, with the exception of neuropsychiatric diseases, less important in complex traits than previously thought.48,49 CNVs could be responsible for the discordance observed in MZ twins for some diseases including cancer, but this is still only supported by anecdotal case reports only.50,51 A very recent study looking in detail at three discordant identical twin pairs for multiple sclerosis has shown that DNA sequence, CNVs, RNA or methylation status did not differ significantly within the three pairs.52 The only difference was in allelic imbalance that can only be detected with allelic specific expression. Recent studies show that many different techniques will be needed to fully elucidate the discordance for many diseases in MZ pairs and help to unravel the complexities of gene expression variations.

Twins are ideal to look at differences in DNA methylation profiles between MZ and DZ pairs. Kaminsky et al.53 conducted DNA methylation analysis in white cells, buccal and gut epithelial cells in 114 MZ and 80 DZ twins. It appears that DNA methylation differences within MZ and DZ pairs are the greatest at regions without clearly defined regulatory function compared with important regulatory DNA regions, suggesting that there is a functional stratification of the epigenome. It is now thought that the difference seen in concordance in MZ twins is more due to these epigenetic differences rather than due to environmental differences per se. Javierre et al.54 suggested that lupus discordance in MZ twins associates with changes in the pattern of DNA methylation. However, these epigenetic factors can themselves be influenced by environmental factors, and so it is difficult to fully dissociate. As expected, the studies of DZ twins show more difference within pairs for the rate of DNA methylation compared with MZ, and so, the amount of methylation is partly heritable. The methylation profiles are also tissue specific and, for example, were more pronounced in the buccal mucosa than white cells.53 So, it appears that MZ can be more similar not only because of DNA sequence but also because they originate from one zygote with similar epigenomic profile, whereas DZ originate from two zygotes with different epigenomic profiles. In the future, methylation studies will have to be based on tissues rather than blood, as epigenetic events appear to be tissue specific, and skin would be ideal as it is more readily accessible. The Illumina 450K Infinium Methylation BeadChip is promising in allowing examination of more than 450 K of methylation sites, and MZ pairs discordant for skin diseases or traits would be ideal for these genome-wide methylation studies.

Another possible explanation for some missing heritability in GWAS studies may be the definition of a trait or disease. By assuming that all diseases are the same, for example, atopic dermatitis, and analysing them together in a GWAS looking for variants, there may be some clinical subtypes with many different genetic pathways, and this will lead to GWAS missing many variants because of inaccurate definition of the disease or the phenotype.

### Transcriptomics

Twins are also very useful to look at differences in expression profiles in different tissues by comparing these differences within MZ and DZ pairs. Human regulatory variation is a very complex area of research as gene expression is an important regulatory process that determines proportion of the phenotypic difference. The twin model helps to investigate MZ twins with identical DNA sequence and assess expression profiles in various tissues. This is achieved using microarrays and second generation RNA sequencing (RNA-seq). The TwinsUK registry has already collected genome-wide RNA data on more than 800 skin biopsies using the Illumina whole genome expression array Human HT-12 version and the data are currently being analysed (the MuTHER study, http://www.muther.ac.uk). The skin expression data will be compared with data on fat and white cells collected in the same patients. Around 15,000 expressed genes were detected in the skin samples, and at first glance, age is an important variable in the expression profiles in skin samples.55 These studies are particularly helpful to further investigate the results of GWAS studies. Some phenotypes may only manifest themselves in one tissue especially in skin diseases, and therefore, studying the effect of genes on various tissues in terms of significance and effect size is important. RNA sequencing, as the next step, is now taking this further by identifying new putative coding exons, and allows us to investigate the variation in transcription, splicing and allele specific expression between individuals.56 The TwinsUK twin registry as part of the
EU funded EuroBATS Consortium has just started RNA sequencing in skin fat and white cells, and these data will be correlated to the extensive genotypic and phenotypic database already available.

Conclusions
Twins have always fascinated scientists, and to this day, with recent progresses in high throughput genotyping, twin models are still providing an invaluable resource to discover new genes and understand how they work. Skin research has advanced with the help of twins, as this has allowed us to realize the importance of genetics on all common skin diseases, some of which may have been ignored. Epigenetics is a new and fast moving area of genetic research, and identical discordant twins may be more informative than concordant twins, as they will shed more light on aetiology. Skin being a readily accessible tissue has proven to be very useful for transcriptomics, RNA sequencing and methylation studies that are currently under way. These studies are not only likely to provide significant insights into skin diseases but also will serve as a window to other less accessible diseases and tissues.

References


