Epigenetics as the Underlying Mechanism for Monozygotic Twin Discordance

Tara L. Hogenson

Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minn., USA

Key Words

Monozygotic twins · Phenotypic discordance · Environmental factors

Abstract

Monozygotic twins share an identical DNA sequence but typically display some level of phenotypic discordance. The cause of this discordance is often unknown. Two known contributing factors to phenotype are genetics and environment. While the mechanism for the genetic effect is defined through DNA sequence, the mechanism for expression of the environmental effect is less defined. With the emergence of the field of epigenetics, researchers have begun to consider it an important contributing factor to phenotype. Exposure to various environmental factors has been shown to have an effect on an individual’s epigenetic marks and may be the primary mechanism for how the environment induces a change in phenotype. Current research indicates that epigenetic differences between monozygotic twins may be an important contributing factor to phenotypic discordance. Monozygotic twins are an ideal resource for the study of epigenetics since many confounding factors found in the general population, such as variation in DNA sequence, can be eliminated. Uncovering these epigenetic factors will increase our understanding of the molecular mechanisms that may contribute to phenotypic discordance in monozygotic twins and may lead to new theoretical and experimental opportunities in health and disease.

The Twin Method

Twins account for 1 in 80 pregnancies worldwide with one third of these being identical [1]. This amounts to approximately 1 in 250 live births [2]. Twinning rates have steadily increased in Western countries since the 1970s with the advent of assisted reproductive tech-
There are two main distinctions between twins, monozygotic (MZ) or dizygotic (DZ). MZ, or identical, twins result from a single ovum fertilized by a single sperm. DZ, or fraternal, twins result from two different ova fertilized by two different sperm. MZ twins originate from the same zygote and are genetically identical while DZ twins share on average half of their genome (the same as non-twin siblings) but have a shared prenatal environment.

There are four characterized types of MZ twins, defined by the stage at which the egg divides (fig. 1). The exact trigger of zygote splitting is unknown. MZ twinning has traditionally been viewed as a random process [5]. However, several studies have found that MZ twinning can be a familial non-random event, indicating that twinning may in part be inherited [6–9]. Recent data have even implicated epigenetic mechanisms or the possibility of a twinning gene [10].
MZ twins share a genetic code and a similar prenatal and postnatal environment. Due to their shared genetic code, differences between identical twins are typically attributed to environment. This assumption has allowed researchers to distinguish between traits that are primarily inherited versus those that are influenced by environmental factors. Twin studies help researchers determine the relative contribution of environment and genetics to complex traits and disease. Traits with a low identical twin concordance rate implicate environment modulating the phenotype.

The classic twin method involves the comparison of MZ and DZ twin concordance rates to determine the heritability of several complex traits. Heritability ($H^2$) is defined as the proportion of phenotypic variance that can be attributed to genetic variance. The heritability of a specific trait is estimated as twice the difference between MZ and DZ twin concordance rates (i.e. Falconer’s formula) [11]. In this formula, $r(MZ)$ is the MZ twin correlation rate and $r(DZ)$ is the DZ twin correlation rate.

$$H^2 = 2[r(MZ) - r(DZ)]$$

Falconer’s formula is based on the assumption that DZ twins share half of their DNA. Therefore, the contribution of genetic factors should be twice the difference in concordance between MZ and DZ twins. This equation does not take into account the effect of environment on variation since shared environment is assumed for both the MZ and DZ twin pairs. If MZ twins share a trait more often than DZ twins, this indicates the importance of genetics in phenotypic variation amongst the general population for that trait. Table 1 shows the concordance and heritability rates of several complex traits and diseases as determined by twin studies.

While MZ twins share identical genotypes, they are often discordant for disease. Discordant identical twins have given researchers important insights into the role of environment in complex disease, but identifying the mechanisms responsible for the observed discordance has been a challenge. With the emergence of the field of epigenetics, evidence suggests that epigenetic mechanisms may be a potential mediator between the environment and phenotypic expression. Epigenetics is defined as any heritable genomic mechanism influencing the phenotype of an organism that is not related to changes in DNA sequence. Epigenetic mechanisms are an important regulator of gene expression.

There are several well-known epigenetic mechanisms, including DNA methylation, chromatin remodeling, and microRNA expression (fig. 2). These epigenetic modifications affect the genome by either inducing or suppressing gene expression, thus resulting in a change in phenotype. Epigenetic functions are responsible for controlling gene expression during

### Table 1. Heritability of several complex traits and diseases from twin studies

<table>
<thead>
<tr>
<th>Trait</th>
<th>MZ concordance rate, %</th>
<th>DZ concordance rate, %</th>
<th>Heritability, %</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>83</td>
<td>46</td>
<td>74</td>
<td>12</td>
</tr>
<tr>
<td>Type 1 diabetes</td>
<td>53</td>
<td>11</td>
<td>72</td>
<td>13</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>50</td>
<td>37</td>
<td>26</td>
<td>14</td>
</tr>
<tr>
<td>SZ</td>
<td>41–65</td>
<td>0–28</td>
<td>80–85</td>
<td>15</td>
</tr>
<tr>
<td>Obesity</td>
<td>74</td>
<td>32</td>
<td>50–90</td>
<td>16</td>
</tr>
<tr>
<td>Autistic disorders</td>
<td>92</td>
<td>10</td>
<td>90</td>
<td>17</td>
</tr>
<tr>
<td>Celiac disorders</td>
<td>83</td>
<td>17</td>
<td>57–87</td>
<td>18</td>
</tr>
</tbody>
</table>

MZ probandwise concordance.

MZ probandwise concordance.
normal cell development and physiological responses, but dysregulation of epigenetics can be associated with diseased phenotypes. Differences in epigenetic marks, such as DNA methylation or histone code acetylation, may lead to MZ twin discordance in complex traits and disease. When no genetic variant is identified between discordant MZ twins, analysis of the epigenome may reveal changes in gene expression responsible for discordance.

**Epigenetic Mechanisms Contributing to Twin Discordance**

Environment has been shown to play an important role in phenotype. What is less clear is the mechanism by which environmental factors induce a change in phenotype. Research using the twin model and isogenic animal models indicates that environment may modulate
the phenotype through the epigenome. Differential exposure to various environmental factors may cause phenotypic discordance between MZ twins through changes in the epigenome. Several environmental factors have demonstrated an influence on epigenetic modifications, including chemical exposure, dietary, behavioral and social interactions [19, 20].

Endocrine disruptors are environmental chemicals that interfere with the function of the endocrine system by blocking or mimicking the actions of hormones [19]. Bisphenol A (BPA) is an endocrine disruptor found in polycarbonate plastic used in many common products including food and beverage containers and baby bottles. Exposure to BPA has been associated with high body weight, increased risk for breast and prostate cancer, and altered reproductive function [21]. In a study by Dolinoy et al. [21], researchers studied the effect of maternal exposure to BPA on the fetal epigenome using the Agouti mouse model.

The Agouti mouse is an example of an animal model that has contributed to our understanding of how environment modulates the epigenome. The Agouti mouse is a unique biosensor that allows measurement of gene expression through a metastable epiallele. A metastable epiallele is an allele that is variably expressed in genetically identical individuals due to differential epigenetic marks between the individuals [22]. These epigenetic marks are of specific interest to researchers since they have been shown to be altered by exposure to different environmental agents. The Agouti mouse genome contains the metastable epiallele, \( A^{vy} \). No metastable epialleles have currently been identified in humans.

The \( A^{vy} \) allele was created by a retrotransposon insertion of a murine intracisternal A particle (IAP). Long terminal repeats flanking the IAP contain promoters that initiate IAP transcription along with adjacent sequences [22]. The \( Agouti \) gene contains an IAP insertion in the 5′ end, which includes a cryptic promoter that promotes constitutive \( Agouti \) expression [21]. CpG methylation of the IAP is established during embryonic development. Methylation levels of the promoter vary amongst isogenic mice, causing phenotypes to range from a yellow to black coat, with complete methylation in the black mice and reduced methylation associated with a yellow coat [5].

In the experiment by Dolinoy et al. [21], maternal dietary exposure to BPA during prenatal development was shown to decrease CpG methylation at the IAP upstream of the \( Agouti \) gene as shown by a statistically significant shift in offspring coat color with increased exposure. CpG methylation was also decreased at another metastable epiallele locus, \( Cabp^{IAP} \), which encodes for the CDK5 activator-binding protein involved in kinase inhibition. This shows multiple murine loci with DNA hypomethylation after exposure to maternal dietary BPA. The mechanism for how BPA lowers the CpG methylation status of promoters is currently unknown. Decreased DNA methylation can lead to key transcriptional changes in genes resulting in an alteration of the fetal phenotype. If the hypomethylation occurs in the promoter region of an oncogene, this may lead to increased expression and the development of cancer in humans.

Exposure to certain environmental factors may cause imprinting defects in the germline or during prenatal development. In a study by Wu et al. [23], researchers studied the effect of exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) during fetal growth in mice. TCDD is a dioxin-like compound and a persistent environmental contaminant. Based on DNA methylation analysis of the \( H19/IGF2 \) imprinting control region, the authors found increased DNA methylation and increased methyltransferase activity of mice exposed to TCDD leading to decreased expression of both a tumor suppressor gene, \( H19 \), and insulin growth-like factor 2 (\( IGF2 \)). The fetuses exposed to TCDD also had a lower birth weight.

The authors proposed that the mechanism by which TCDD increases methylation status of the imprinted genes is through modulation of the methyltransferases responsible for the transfer of methyl groups to the \( H19 \) and \( IGF2 \) promoters [23]. Two distinct methyltransferases involved in the maintenance and establishment of methylation of the genome are
DNMT1 and DNMT3. It has been reported that transcription factor Sp1 controls expression of these DNA methyltransferases [24–26]. The effects of TCDD are mediated by the aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator, which has been reported to act synergistically with Sp1 [27, 28]. Exposure to TCDD was shown to enhance Sp1 activity, indicating that TCDD may increase DNA methyltransferase transcription by increasing Sp1 activity. However, further research needs to be performed to identify the exact methyltransferases involved in the methylation of H19 and IGF2 and to determine the mechanism for control of methylation.

Dietary intake may also play a significant role in disease development. Alzheimer’s disease (AD) is influenced by environmental factors and has been shown to be discordant in MZ twins. Risk factors for AD include health conditions, dietary habits, genetics, age, and lifestyle [29]. Diets high in saturated fat and cholesterol have been shown to increase the risk of AD onset while a low-fat diet and increased physical activity have been related to lowering the risk.

In a study by Mastroeni et al. [30] looking at cortical neurons of twins discordant for AD, the authors noted that there was a significant decrease in DNA methylation in the temporal neocortex neuronal nuclei of the affected twin through immunoreactivity for a DNA methylation marker. The specific hypomethylated promoter regions were not identified. Many dietary factors have been shown to govern DNA methylation including folate, riboflavin, cobalamin, pyridoxine, methionine, and choline [31]. As an example, folate or folic acid deficiency has been implicated in playing a role in AD progression [29, 30].

Folic acid is a water-soluble vitamin essential for numerous body functions. One primary role of folate during DNA synthesis is to provide the methyl group required for DNA methylation. Folate deficiency is rare in Western diets due to the fortification of bread and cereals with folic acid. Intake of folic acid has been significantly associated with a decreased risk of AD, with one study showing a 55% reduced risk [29, 32]. Lowered consumption of folic acid throughout life may lead to hypomethylation across the genome in the folate-deficient AD patient. This could lead to increased expression of genes associated with the AD phenotype.

Dietary deficiencies during prenatal development have a lasting effect on the epigenome of the fetus. Studies have shown that parental diet can impact DNA methylation levels and the phenotype of the next generation in humans [33]. Heijmans et al. [34] demonstrated that individuals who were prenatally exposed to famine during the Dutch Hunger Winter had less DNA methylation in the IGF2 gene than their siblings who were not exposed six decades thereafter. The Dutch Hunger Winter was a distinct period of famine in the western part of the Netherlands in the winter of 1944–1945. Hypomethylation of IGF2 may be due to a deficiency of methyl donors in the diet during gestation, leading to a persistent change in the epigenome.

Exposure to environmental agents is not the only way environment can modulate the epigenome. Social interaction has been shown to have a clear effect on epigenetic modifications. Adoption studies have demonstrated an increased risk of psychiatric disorders for individuals raised in a stressful family environment with no biological risk [35]. One study showed that chronic psychosocial stress (social defeat) in mice led to lasting downregulation of brain-derived neurotrophic factor (BDNF) transcripts [36]. The researchers identified several lasting epigenetic modifications in the hippocampus leading to lowered expression of BDNF. This included histone dimethylation of H3K27 in the BDNF promoter leading to decreased expression. Lowered BDNF transcription has been demonstrated in schizophrenic patients when compared to normal control subjects, and is seen as an early event in the onset of psychosis [37].

In a study by Weaver et al. [38], researchers showed that maternal behavior by rats had an influence on the epigenome of the pups. Increased pup licking and grooming (LG) and arched-back nursing (ABN) by rat mothers during development led to differences in stress
response in offspring. As adults, offspring of high LG-ABN mothers were less fearful. In vivo and in vitro studies demonstrated increased expression of the glucocorticoid receptor (GR) with increased LG-ABN. The GR is expressed in almost every cell in the body and regulates the transcription of several genes involved in the development and stress response. The promoter region of the GR was found to be differentially methylated in the hippocampus, leading to increased GR expression in LG-ABN offspring.

McGowan et al. [39] showed a similar response in humans by examining levels of the GR promoter methylation and mRNA levels in suicide victims with a history of childhood abuse. Childhood adversity has been linked to an altered stress response in humans and is associated with decreased expression of the GR, increased risk of mental disorder, and increased suicide rates [40–42]. The researchers identified increased GR promoter methylation and decreased mRNA in hippocampal samples from suicide victims. These findings indicate the importance of parental care in gene expression in the brain and how epigenetics may mediate the long-term effects of exposures during development.

Differential exposure to these various environmental factors may lead to changes in the epigenotype that explain phenotypic discordance between MZ twins. When an epigenetic modification is introduced into the epigenome, there is one important criterion that must be met in order for it to be notable. The modification must be heritable or stable through mitosis. Epigenetic inheritance is defined as the transmission of epigenetic marks to daughter cells [43]. In the event an epimutation is present, it can be inherited across cell divisions and may exert a positive or negative effect on gene expression.

### Epigenetics-Associated Modifications in Twin Discordance

Unlike DNA sequence, which remains fairly stable throughout the lifetime of an organism, modification of the epigenome is a dynamic process. This can lead to substantial differences in gene expression in cells of organisms with identical genomes. MZ twins discordant for complex disease indicate the importance of different environmental factors in the onset of disease. Investigation into the cause for disease discordance between identical individuals may lead to the identification of nongenetic mechanisms that contribute to variation in phenotype. Several studies of twins discordant for complex disease were able to identify a potentially causal epigenetic modification underlying the phenotypic differences. In what follows is a brief overview of these studies with an emphasis on the identification of epigenetic differences that may have contributed to disease discordance.

### Autoimmune Disease

The cause of most autoimmune diseases remains unknown. It is widely believed that the disease arises due to both genetic and environmental exposures as well as defects in the immune response itself. Many autoimmune diseases do not arise until later in life, indicating that a cellular event caused the individual to lose the function of a protective element that previously prevented disease. Due to the dynamic nature of the epigenome, this hints at a role of epigenetics in disease onset.

Twin studies indicate that there is a strong genetic component to certain autoimmune diseases, such as celiac disease and type 1 diabetes (T1D), with a high concordance rate between twins [44, 45]. However, some autoimmune diseases have a small genetic component, such as rheumatoid arthritis, multiple sclerosis, systemic lupus erythematosus (SLE), and systemic sclerosis [45, 46]. Genome-wide association studies have discovered only a few
unique genetic mechanisms for the loss of immunological tolerance [47]. Epigenetic factors may be the missing link in identifying the molecular mechanisms that trigger autoimmunity.

Several studies have uncovered a correlation between DNA hypomethylation and SLE progression [48, 49]. SLE is a systemic autoimmune inflammatory disease. There are several well-established risk factors associated with SLE, including alleles in the major histocompatibility complex region [50]. However, genetic factors do not appear to be the only factor leading to disease as shown by MZ twin discordance.

In one study examining epigenetic modifications associated with SLE, Javierre et al. [50] performed sequence analysis of DNA methylation in lymphocytes of MZ twins discordant for the disease. The authors were able to identify 49 differentially methylated regions (DMRs) between the SLE and healthy twins that were potentially relevant in SLE pathogenesis. This set of 49 genes showed no significant differences in methylation when compared to normal age-matched controls, suggesting these DMRs may play a role in the development of SLE in the affected twin. Among these genes shown to have reduced methylation in SLE individuals was interferon gamma receptor 2 (IFNGR2), which has previously been associated with the pathogenesis of the disease in an SLE mouse model [51]. Using RT-PCR, the authors showed that the reduction in DNA methylation was correlated with increased expression levels in the SLE twin for 5 genes, including IFNGR2.

T1D is a complex autoimmune disease that is common worldwide. MZ twins are often discordant for the disease with only an approximate 50% concordance rate between twins [52]. This is also highly dependent on the age the affected twin is diagnosed. Due to this low rate of concordance, it is believed that nongenetic factors play an important role in the onset of the disease, including epigenetics. However, a specific epigenetic variant for T1D has yet to be identified between discordant MZ twins.

In a study by Rakyan et al. [53], researchers performed DNA methylation analysis through an epigenome-wide association study in CD14+ monocytes of T1D-discordant MZ twins. They identified 132 intrapair differentially methylated CpG sites of genes known to be associated with T1D or the immune response. The methylation differences significantly correlated with the diabetic state. This included the HLA class II gene, HLA-DQB1. This allele is one of the known highest genetic risk factors for T1D. They confirmed their findings in a ‘longitudinal’ study, looking at individual methylation levels in singletons before and immediately after the clinical onset of T1D. This suggests that T1D-associated methylation arises early in the development of the disease. However, the authors did not confirm the functional outcome of the differentially methylated T1D CpG sites, which could have been determined by measuring the gene expression of the affected genes.

**Autism Spectrum Disorder**

Autistic disorders are a neurological disorder characterized by difficulties with social interaction, communication, and repetitive behavior [54]. Autism spectrum disorder (ASD) is divided into three distinct disorders: autism, Asperger syndrome, and pervasive developmental disorder not otherwise specified differing in severity [55]. The rate of occurrence of ASD in the general population is 0.6% [56]. Clinical signs of ASD are usually present by age 3 but research indicates that abnormalities in the development of communication and social behavior can be detected as early as 14 months [57].

Based on the marked delay or disruption of communication in patients with ASD, researchers hypothesized that ASD may be caused by defects in the neuroanatomic structures or neurodevelopmental pathways during early brain development [58]. Neuroimaging of individuals with ASD has identified alterations in neuronal organization, neurotransmitter
pathways and brain growth [58]. These data support the notion that disruption of neurodevelopment in utero or shortly after birth is involved in the development of ASD.

Genetic factors play a primary role in the development of the disease with a concordance rate of >90% among MZ twins. Mutations in genes controlling neurodevelopment have been associated with the development of ASD. However, data from linkage studies are rarely replicated making the isolation of genes responsible for the phenotype difficult to identify [59]. Males are four times more likely to develop autism than females [60]. Despite this marked difference, research assessing linkage to the X chromosome has found no correlation [55].

Although twins are typically concordant for ASD, they are commonly found to be on different ends of the autism spectrum. This hints at a strong environmental component in disease presentation [55, 61]. The prevailing theory for the etiology of ASD is that exposure to potential environmental pollutants or neurotoxins during prenatal development along with genetic susceptibility play a role in the development of the disease. This indicates that a shared prenatal environment along with a shared genome may explain high MZ twin ASD concordance rates. Exposure to teratogenic medications in utero, such as valproate, thalidomide and misoprostol, has been associated with ASD [62]. Valproate has been shown to play a role in the etiology of ASD through epigenetic mechanisms, including interference with folate metabolism and inhibition of histone deacetylases (HDAC) [59].

Aberrant methylation or loss of methylation during critical phases of imprinting can lead to a loss of critical gene expression or overexpression [22]. Thus, environmental agents that alter imprinting may lead to severe developmental disorders. Badcock and Crespi [54] proposed in 2006 that autism may be caused by extreme imprinting in the brain. The brain tissue undergoes major genomic imprinting, which may indicate a strong component of epigenetics in the presentation of the disorder. Several autistic syndromes can be caused by loss of imprinting at a single locus, including Angelman and Prader-Willi syndromes, Rett syndrome, and Turner syndrome.

In a study by Hu et al. [63], the authors studied differences in gene expression in genetically identical individuals with different degrees of autism. They used microarray analysis to observe lymphoblastoid cell lines derived from the blood of MZ twins for biomarkers of autism with the hypothesis that expression differences between affected and unaffected individuals might be present in tissues besides the brain [63]. Based on their results, they were able to show that MZ twins discordant for ASD exhibited differential expression for genes critical in the nervous system. One gene identified to be differentially expressed in lymphoblastoid cell line cells was the serotonin transporter, which has been strongly implicated in ASD. A decrease in the expression of the serotonin transporter was associated with increased severity of the disease between the MZ twins. Based on these data, it is possible that epigenetic mechanisms may play a role in the differential expression of genes, leading to discordance of autistic phenotype between the MZ twins. However, this study did not compare the epigenome of the twins.

In a separate study by Nguyen et al. [64], the authors investigated the epigenetic contribution to ASD by observing the DNA methylation profile of lymphoblastoid cell line cells in MZ twins with a discordant diagnosis for ASD. The authors discovered distinct methylation differences in two ASD candidate genes, BCL-2 and RORA, between the discordant MZ twins. Both genes exhibited increased methylation of specific CpG sites in the ASD twins, leading to lowered expression. BCL-2 is an anti-apoptotic protein, which has been shown to be significantly reduced in patients with ASD [65]. RORA is a nuclear hormone receptor involved in the transcriptional regulation of many genes including survival of Purkinje cells [66]. Deficiency of Purkinje cells is a common feature in the postmortem brain of patients with ASD [67]. These findings not only identified key candidate genes for ASD, but also epigenetic modifications that were relevant to the ASD phenotype.
Schizophrenia and Bipolar Disorder

Schizophrenia (SZ) is a chronic and disabling brain disorder that includes symptoms such as hallucinations, delusions, and disorganized communication [68]. Onset of SZ usually occurs in early adulthood. The incidence of the disorder is fairly low with an occurrence of 0.5–1% in the general population [69]. A combination of genetics and environment appears to play a role in the development of the disease with a concordance rate between 41 and 65% for MZ twins [70]. Molecular studies have been conducted to identify DNA mutations within the affected twin, but no significant differences between MZ twins were detected [70].

Several epidemiology studies have revealed a long list of environmental factors that may contribute to disease onset including stress, mental and physical abuse, and drug use [71]. However, it is difficult to determine if certain environmental factors, such as drug use, may be due to self-medication for the presence of psychotic symptoms before SZ diagnosis. Due to the clear effect of environment on the etiology of the disease, researchers have begun to use the twin model to identify epigenetic differences in MZ twins discordant for SZ.

Dempster et al. [72] performed an epigenome-wide association study of DNA methylation in blood samples of 22 pairs of MZ twins discordant for SZ and bipolar disorder (BD). BD has an incidence of approximately 0.5–1% in both males and females [73]. BD is characterized by states of mania alternated with symptoms of depression. Genetic factors appear to contribute substantially to the onset of the disease while environmental factors have also been implicated. The concordance rate of BD is much higher (80%) in MZ twins than that of SZ [74]. SZ and BD have been shown to clearly aggregate in families indicating a strong genetic component.

Dempster et al. [72] found numerous DMRs associated with psychosis between the discordant twins, including \( SOX_1 \), \( SMUG_1 \), and \( GGN \). One gene of interest was the promoter region of \( ST6GALNAC1 \), located at 17q25.1, which was found to be hypomethylated in the affected twins. \( ST6GALNAC1 \) is a member of the sialyltransferase family and plays a role in protein glycosylation. It is expressed primarily in the brain and other CNS tissue. Increased \( ST6GALNAC1 \) expression is associated with SZ. Hypomethylation at this locus leads to increased gene expression, which may be an important contributing factor to the pathogenesis of the disease.

Human and animal model studies have suggested that \( GPR24 \), encoding G protein-coupled receptor 24, is associated with susceptibility to BD. One of the DMRs identified by the researchers was located at only 21 base pairs from the \( GPR24 \) allele, implying an epigenetic link between \( GPR24 \) and BD. These findings suggest that individuals may be affected with this disease in part due to a modification of the epigenome. However, it is difficult to draw conclusions about causality since DNA samples were not taken from the twins until after the onset of the disease.

Cancer

Cancer is an uncontrolled growth of abnormal cells in the body. Cancer is caused by a loss of function of genes required for control of cell growth (tumor suppressor genes), increased expression of genes involved in cell proliferation (oncogenes), or loss of function of genes required for DNA repair (mismatch repair genes). Genetics and environment have both been shown to play a role in the development of the disease with a strong link between cancer and some environmental toxins also known as carcinogens. Carcinogens are DNA mutagens associated with chemicals found in cigarettes, high-fat diet, and exposure to ultraviolet light.
While some environmental factors have been shown to induce DNA mutations, changes in epigenetic modifications have also been associated with cancer [75]. Alteration in the methylation patterns on DNA is one of the most common epigenetic changes identified in cancer [76]. Hypermethylation of CpG islands in the promoter region of tumor suppressor genes leads to decreased expression and increased cell proliferation. Studies have shown promoter hypermethylation of the tumor suppressor gene p16 in patients with squamous cell lung carcinoma and pancreatic cancer [77, 78].

Global hypomethylation has also been associated with an increased risk of cancer [76]. DNA hypomethylation leads to chromosomal instability, reactivation of transposable elements, activation of oncogenes, and loss of imprinting [79]. Other epigenetic modifications associated with cancer include histone modifications and modulation of expression of microRNAs. Hypoacetylated and hypermethylated histones H2 and H4 can lead to silencing of tumor suppressor genes [80]. Downregulation of subgroups of microRNAs is often seen among tumor tissues [81]. DNA methylation in the microRNA 5′ regulatory region could be responsible for decreased expression of microRNA in tumors [79].

Due to the large role of environment in disease development, cancer is often discordant between MZ twins. Genetic susceptibility accounts for approximately 30% of the heritability of breast cancer, the most common female neoplasm [82]. BRCA1 and BRCA2 are the primary genes associated with heritable breast cancer. BRCA1 and BRCA2 code for DNA repair enzymes. In nonfamilial cancers, mutations are rarely identified in these genes. However, increased methylation of BRCA1 has been observed [83].

In a study by Galetzka et al. [84], researchers analyzed the methylation patterns of BRCA1 promoter regions in primary fibroblasts between MZ twins discordant for childhood leukemia. The affected twin had increased methylation of the BRCA1 promoter region (12%) compared to the healthy twin (3%). Since this was a childhood cancer, the epimutation in the affected twin may have occurred during an early postzygotic event. Hypermethylation of the BRCA1 allele leads to a decreased expression of the gene, which may have contributed to the development of cancer in the affected twin.

In another study, Heyn et al. [82] examined DNA methylation patterns in identical twins discordant for breast cancer. The researchers examined whole-blood specimens of 15 twin pairs discordant for breast cancer and identified 403 DMRs. An additional study of a cohort of 21 twin pairs discovered an alternative gene product of the docking protein DOK7 as a candidate for blood-based cancer diagnosis. Increased methylation of DOK7 was also shown to be an important biomarker, which is detectable years before cancer diagnosis. These data have not yet been confirmed in a larger clinical dataset.

An understanding of the effect of epigenetics on diseases such as cancer is an important tool for treating disease. Comparison of epigenetic profiles of twins discordant for cancer could help identify important epigenetic modifications in the disease. Identifying specific epigenetic profiles for various types of cancer could help diagnose individuals more accurately. This information could also aid in disease prevention and drug therapy since some epigenetic modifications may be reversible with treatment.

**Future Directions**

Despite the advantages of the twin model for epigenetic studies, there are several issues within the field of epigenetic research that need to be addressed. For instance, the epigenome varies between tissues unlike genomic information, which is generally uniform among all cells. This makes tissue specificity extremely important when studying the epigenome, which can be difficult if the tissue is not easily sampled. Also, there are currently no proper tools to
differentiate causal epigenetic changes from epigenetic ‘noise’. Often times, epigenetic modifications may be caused by the disease or treatment and may not play a direct role in pathogenesis. For example, epigenetic modifications associated with disease may actually be due to a mutation in a gene that codes for a protein involved in maintaining or modifying the epigenome. Mutations in genes that code for DNA methyltransferases and other epigenetic modifiers have been shown to cause disease [85].

Unlike the genome, the epigenome is believed to undergo constant fluctuations, which makes it difficult to ascertain when a mutation first appears. This can be especially difficult given the fact that epigenetic modifications diverge over time. Longitudinal studies identifying epigenetic differences between discordant MZ twins before and after onset of disease will play an important role in establishing causal epimutations. Data from these studies may potentially aid researchers in the growing fields of personalized medicine and epigenetic therapy.

Worldwide use of drug therapy has revealed a substantial difference in response between individuals. While some therapies may be helpful to one individual, they may have an adverse effect on other patients. Often, there is a distinct molecular mechanism underlying an individual’s response to treatment [86]. Pharmacogenetics, commonly referred to as personalized medicine, involves identifying genetic variants between individuals that determine treatment response. It correlates phenotypic biomarkers of disease with genetic variants through association studies and twin studies in patients [86].

Many interindividual discrepancies in response to drugs cannot be explained by DNA sequence variation alone. Epigenetics may offer an alternative explanation for these discrepancies [87]. Pharmacoepigenetics investigates individual variation in drug response due to epigenetic variation. In the coming years, next-generation sequencing of the epigenome across multiple tissue and cell types may facilitate detection of epigenetic variants that may explain variation in drug response [88, 89].

Historically, discordant MZ twins have been particularly helpful in molecular genetic studies for identifying the locus responsible for a given phenotype. This is due to the elimination of various confounding factors, including difference in genetics and environment found in general population studies. These same advantages may be applied to epigenetic studies. An epigenome-wide association study observing DMRs, histone modifications, or variation in microRNA expression between discordant MZ twins will aid in the identification of epigenetic modifications associated with disease and treatment response.

An advantage of treatment for disease-related epigenetic modifications is that these modifications can be manipulated, unlike DNA sequence. Several drugs are currently being developed that target epigenetic modifications, leading to the growing field of pharmacoepigenomics. Epigenetic therapies include DNA methylation inhibitors and HDAC inhibitors [75]. DNA methylation inhibitors are small molecules designed to reactivate aberrantly silenced genes by interfering with DNMT activity [75]. HDAC inhibitor compounds interfere with HDAC, leading to restoration of gene function.

Several inhibitors of DNA methylation and histone deacetylation have been approved for treatment of cancer by the US Food and Drug Administration and have been used clinically [90]. DNMT inhibitors 5-azacytidine (Vidaza) and 5-aza-2'-deoxycytidine (Decitabine) are both approved by the Food and Drug Administration for treatment of myelodysplastic syndromes [91]. HDAC inhibitors suberoylanilide hydroxamic acid (Vorinostat) and depsipeptide (Romidepsin) are approved for treatment of T-cell lymphoma. Some drugs that are already widely used for noncancer diseases have only recently been discovered to have an effect on the epigenome. The HDAC inhibitor valproic acid, typically used for the treatment of epilepsy, has been approved for treatment of leukemia, breast cancer and ovarian cancer [91]. There are several new DNMT and HDAC inhibitors that are currently in clinical trials [92].
Besides targeting histone-modifying enzymes, some epigenetic therapies target proteins that have been shown to interact with epigenetic machinery. MeCP2 is an ideal target for epigenetic modification due to its role in the regulation of transcription by binding methyl domains. Zebularine is a nucleoside analog of cytidine that was originally developed as a cytidine deaminase inhibitor but has also been shown to inhibit DNA methylation and tumor growth [93]. Recent studies have shown that zebularine inhibits MeCP2 expression making it a possible candidate for epigenetic treatment [94].

Despite the advances in novel epigenetic therapeutics, there are several problems with epigenetic modifiers, including their lack of specificity. A possible solution to the pitfalls of current epigenetic drugs would be to explore other therapeutic options, such as drugs that bind directly to the epigenetic locus [75]. As the field of epigenetic therapy continues to grow, research into epigenetic modifications associated with disease using the discordant twin model may play an important role in the identification of additional therapeutic targets for the treatment of disease.

Conclusions

While there are hundreds of known genetic variants associated with disease, variants in the epigenome have been vastly understudied. As demonstrated in the previous sections, differences in epigenetic modifications between MZ twins have been shown to contribute to disease discordance. The use of the discordant MZ twin model provides a unique opportunity for exploration of epigenetic mechanisms and their role in disease discordance. Analysis of the epigenome of MZ twins discordant for disease may lead to the identification of causal epigenetic modifications. Identification of these epigenetic marks may reveal the biological mechanisms responsible for the phenotypes of affected individuals. These discoveries may lead to the identification of new therapeutic targets, and, with the isolation of definitive causal factors, may also play a role in disease prevention.

Acknowledgements

Thanks to Michael A. King for his help with the illustrations and to Martin E. Fernandez-Zapico for critically reading the manuscript.

References


Hogenson: Epigenetics as the Underlying Mechanism for Monozygotic Twin Discordance