Review

What monozygotic twins discordant for phenotype illustrate about mechanisms influencing genetic forms of neurodegeneration


As monozygotic (MZ) twins are believed to be genetically identical, discordance for disease phenotype between MZ twins has been used in genetic research to understand the contribution of genetic vs environmental factors in disease development. However, recent studies show that MZ twins can differ both genetically and epigenetically. Screening MZ twins for genetic and/or epigenetic differences could be a useful and novel approach to identify modifying factors influencing phenotypic expression of disease. MZ twins that are phenotypically discordant for monogenic diseases are of special interest. Such occurrences have been described for Huntington’s disease, spinocerebellar ataxias, as well as for familial forms of Alzheimer’s disease. By comparing MZ twins that are phenotypically discordant, crucial factors influencing the phenotypic expression of the disease could be identified, which may be of relevance for understanding disease pathogenesis and variability in disease phenotype. Overall, understanding the crucial factors in development of a neurodegenerative disorder will have relevance for predictive testing, preventive treatment and could help to identify novel therapeutic targets.

Conflict of interest

The authors report no conflict of interest.

Although familial aggregation of phenotype is a prominent feature of most neurodegenerative disorders and has been of great value in the discovery of genetic contribution, numerous examples of lack of expected correlation between genotype and phenotype are apparent (1, 2). The direct translation of a mutated gene into a disease phenotype is challenged by phenomena such as reduced penetrance, high variability of age of onset (AO) between individuals carrying the same mutation, or by a broad spectrum of disease symptoms despite involvement of the same gene. Factors such as genetic background of the individual, (epi)genetic mutations that occur during life, and environmental factors could be relevant (1–3).

The complexity of genotype–phenotype correlations in neurological disorders is further exemplified by the existence of monozygotic (MZ) twins that are discordant for disease phenotype (4), despite having the same (or a very similar, see later) genotype. Discordance is defined here as ‘differing significantly’ in either presence or presentation of the disease (5), e.g. a twin is considered discordant if just one of the pair has the disease diagnosis, if a significant difference in AO of the disease is present, or if they significantly differ in clinical symptoms.

In monogenic disorders such as Huntington’s disease (HD), spinocerebellar ataxias (SCAs), or dominant forms of Alzheimer’s disease (AD), discordance in MZ
twins – although relatively rare (4, 6, 7) – may provide understanding of relevant factors that are involved in disease development in addition to the causative gene. Although MZ twins are historically considered as genetically identical, several recent studies suggest that MZ twins may differ in the epigenetic status of their genome (8, 9), and could even show (small) genetic variations (10, 11). This extends the classical view of environmental factors alone causing discordance, where (epi)genetic factors could be relevant in addition or could form mechanistic explanations for environmental impact on phenotype.

Here we will illustrate which new aspects of discordant MZ twins could be studied in order to increase the understanding of mechanisms contributing to genetic forms of neurodegenerative disorders, using the autosomal dominant neurodegenerative disorders HD and SCAs, as well as familial AD as an example.

A novel role for MZ twins in medical genetic research

‘Very similar, but not identical’, that is what every mother will tell you about her MZ twin. Indeed, phenotypic differences in molecularly confirmed MZ twins can vary from as small as the location of the crown to as striking as the discordance in disease development. Historically, the rate of phenotypic discordance within a twin pair formed the basis for dissecting the amount of genetic vs environmental influences on phenotype expression, especially used in complex disorders. The degree of similarity in disease expression between MZ pairs (concordance) over that of dizygotic (DZ) pairs is herein taken as evidence for genetic contribution in the etiology (12).

The central assumption of these studies is that MZ twins are genetically identical, whereas DZ twins are genetically no more similar than siblings born from separate pregnancies. In MZ twins, phenotypic differences have been assumed to merely result from a variable environment. The validity of this assumption, however, is challenged by recent studies that describe the existence of genetic and epigenetic differences in MZ twins (8–11). Therefore, one could hypothesize that – in addition to environmental factors – differences in genetic and epigenetic makeup could form alternative explanations for phenotypic discordance. Indeed, discordance of MZ twins in Beckwith–Wiedeman syndrome has been explained through a difference in epigenetic imprinting of the genes involved (13). The existence of copy number variations (CNVs) in MZ twins, indicating genetic differences, has also been identified (10).

In addition, the existence of MZ twins that are discordant for highly penetrant disorders also reveals that other factors than the causative gene could play a role. Although rare, examples of discordant MZ twins have been described for autosomal dominant neurodegenerative disorders, including HD (14–23) and SCAs (24). In contrast to non-familial forms of AD, twin discordance for familial AD is also considered as rare, and few cases have been described (4). Since environmental impact is thought of as minor in such penetrant forms of diseases, epigenetic or other genetic alterations could be particularly relevant. As an example, MZ twins with discordant features for autosomal dominant neurofibromatosis type I (NF1) were screened for methylation differences in the NF1 region (25). The authors reported the existence of intra-pair differences in methylation, arguing the likelihood of epigenetic differences in the NF1 as a possible modifier for NF1 phenotypic variability.

Overall, such discordant cases could form an excellent resource to discover crucial factors that determine the phenotypic expression of disease, in addition to the main disease gene, thereby possibly increasing our understanding of phenomena such as reduced penetrance or high AO variability in the disease. A note of caution, however, is the expected low sample size because of the rare occurrence of discordance, especially in case of HD and SCAs (14–24), or because of difficulties in documentation of discordance. The latter is especially relevant for cases of late onset where one co-twin could have passed away for other reasons than due to the disease (4, 5).

Genetics of neurodegenerative disorders

As illustrated in Fig. 1, neurodegenerative disorders can be depicted as a continuous spectrum, ranging from monogenic, Mendelian-inherited disorders to multicausal, complex disorders that are influenced by several genes and modifying factors, each of small effect size.

Huntington’s disease

HD is a monogenic neurodegenerative disorder inherited as an autosomal dominant trait, and usually presents itself in adulthood (7). Prevalence rates are estimated between 1 and 7 persons per 100,000, dependent on the specific population (7, 26). Neuropathologically,
Mechanisms influencing genetic forms of neurodegeneration

HD is characterized by progressive neurodegeneration of mainly the striatum and cerebral cortex, producing motor, cognitive and psychiatric symptoms (7). The responsible mutation in HD is a CAG repeat elongation in the HTT gene, which is located on chromosome 4. It normally contains a stretch of <35 repeats, while it is expanded to >36 CAG repeats in patients. In the affected range, the individual CAG tract length is a major determinant for AO, where longer tracts cause earlier onset (27, 28). However, some studies indicate that only 70% of the variance in AO can be explained by CAG tract size (7, 28), leaving ∼30% to be explained by modifying factors other than the causative gene (7). It has been estimated that 10–20% of the difference in AO might be explained by ‘heritable factors, including non-CAG genetic modifiers’ and the remaining 10% might be environmental (7, 28).

MZ twins that are discordant for HD phenotype, although rare, have been described (14–23) and may be useful to determine crucial modifiers in the development of HD. Despite being genetically (almost) identical, some MZ twins have been reported to extensively differ in the nature of HD symptoms, or show a significant difference in AO. In older case reports, motor symptoms and onset of behavioral symptoms were discordant (15, 16, 23). More recent case reports have shown discordance in motor symptoms, behavioral changes and cognitive differences (14, 18–20). MZ twins significantly differing for HD AO (>6 years) have also been described (17, 22) (Table 1).

Intra-twin differences in epigenetic status influencing transcription, or acquired genetic differences in cis/trans factors modifying HD, as well as the occurrence of somatic instability of the CAG tract could explain disease discordance (9, 21, 29, 30). Moreover, it is useful to include MZ twins concordant for HD, since that would enable one to filter acquired pathogenic differences from non-pathogenic ones. Although reports on HD twins are sparse, four examples of MZ twins concordant for HD have recently been described (31), and several older case reports can be found (32–35).

Possible (epi)genetic differences derived from a study toward factors underlying disease discordance in MZ twins could be relevant to explain low penetrance and high AO variability in HD singletons.

Spinocerebellar ataxias

SCAs are a group of autosomal dominant-inherited neurodegenerative disorders, characterized by atrophy in cerebellar and spinal tissue, resulting in motor symptoms such as unsteady gait, clumsiness and dysarthria, and by neurological symptoms including pyramidal signs, ophthalmoplegia, and cognitive impairment (6, 36). It has been estimated that approximately 1–3 per 100,000 persons exhibit SCA (6, 36, 37). SCAs can be phenotypically heterogeneous, depending on the specific gene and mutation involved. However, even between patients carrying the same dominant mutation in the same gene, a wide range of phenotypic expression is seen, as illustrated by different symptoms or high AO variability (6, 36), similar to HD. Genetically, they can be grouped in coding CAG-expansion forms, non-coding expansion SCAs or conventional-mutation SCAs, but all showing autosomal dominant inheritance.

| Table 1. Examples of MZ twins with discordant phenotype for HD, SCA2, and familial AD* |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **Huntington’s disease**        |                  |                  |                  |                  |                  |
| Motor symptoms                  | Cognitive symptoms | Behavioral symptoms | Age of onset | References     |
| x                               |                  |                  |                  | (15)            |
| x                               |                  |                  |                  | (23)            |
| x                               |                  |                  |                  | (18)            |
| x                               |                  |                  |                  | (15)            |
| x                               |                  |                  | (Δ 3 years)      | (20)            |
| x                               |                  |                  | (Δ >7 years)     | (17)            |
| x                               |                  |                  | (Δ >7 years)     | (19)            |
| x                               |                  |                  | (Δ 6 years)      | (22)            |

| **Spinocerebellar ataxia-2**    |                  |                  |                  |                  |
| Motor symptoms                  |                  |                  |                  |                  |
| x                               |                  |                  |                  | (Δ >3 years)     |

| **Alzheimer’s disease**         |                  |                  |                  |                  |
| Type of AD                      | Motor symptoms    | Behavioral symptoms | Age of onset | Lewy body-related pathology | References |
| FAD-APOE-ε4 I                   |                  |                  | (Δ 18 years)    |                  | (4)        |
| FAD-APOE-ε4 II                  |                  |                  | (Δ 7 years)     |                  | (4)        |
| EOFAD-PSEN1                     |                  |                  | (Δ 4 years)     |                  | (4)        |

AD, Alzheimer’s disease; EOFAD-PSEN1, early onset familial Alzheimer’s disease with PSEN1 mutation; FAD-APOE-ε4, familial Alzheimer’s disease with APOE-ε4 mutation; HD, Huntington’s disease; MZ, monozygotic; SCA, spinocerebellar ataxia.

*Cross-indicated are the discordant features that the authors described for the twin.
CAG-expansion SCAs are more frequent than the other SCAs, with SCA1-3 being the most frequent CAG-expansion forms (6, 36).

CAG-expansion SCAs share the same mutation type with HD; with clinical manifestation above a certain threshold of repeats. In non-coding expansion SCAs, untranslated expansions cause disease, probably through a gain of function of RNA (6), while in conventional-mutation SCAs a deletion, insertion or other conventional mutation in a coding gene is causing SCA (6, 36).

Several studies suggest that CAG-expansion SCAs also exhibit a negative correlation between CAG tract size and AO (6, 36). Nevertheless, a high variability in AO, and the presence of a reduced penetrance range in several SCAs (6), indicates the involvement of modifiers in disease development – as in HD. However, only about 50% of the AO variance for SCAs is explained by CAG size (6, 37) compared to ∼70% in HD (7), suggesting a larger influence of other factors than the mutated gene in case of SCAs.

Applying an MZ twin approach could shed light on relevant modifying factors in SCAs, and be relevant for CAG-expansion disorders in general, as applies for HD, since similar modifiers could be involved in the etiology of CAG-expansion disorders as a whole (36).

In CAG-expansion disorders, such as SCA1-3 and SCA6, the length of the non-mutated allele has been shown to influence AO, and longer normal CAG lengths of non-causal genes can also contribute to AO variability (6, 37), showing that cis and trans acting genetic factors need to be considered in CAG-expansion disorders, and could form one region of interest to investigate in discordant MZ twins. Reports on MZ twins concordant and discordant for SCA phenotype are sparse. However, one MZ twin pair discordant for SCA2 has been described, showing discordance in AO and significant differences in ocular movement, motor coordination and posture stability (Table 1) (24). The existence of somatic mutations was proposed as possibly contributing to disease development.

Alzheimer's disease

AD is the most common neurodegenerative disorder causing dementia in the elderly, 43% of persons above 85 years suffer from AD, and AD comprises between 60% and 80% of all prevalent dementia cases (38, 39). AD is characterized by global cognitive decline as well as the accumulation of beta-amyloid (Aβ) deposits and neurofibrillary tangles in the brain (3).

The majority of AD cases are sporadic, characterized by late onset (LOAD, >60 years, ∼95% of AD cases), where polygenetic and environmental factors influence the disease. However, rare and highly penetrant early onset familial AD (EOFAD, <60 years, <5% of AD cases) forms exist, which are transmitted in an autosomal dominant fashion (3, 38). To date, dominant mutations in three single genes have been reported to cause EOFAD. These include the Aβ precursor protein (APP) on chromosome 21, presenilin 1 (PSEN1) on chromosome 14 and presenilin 2 (PSEN2) on chromosome 1. The most frequently mutated gene, PSEN1, accounts for the majority of AD cases with onset prior to age 50. These AD-causing mutations share a common biochemical pathway, i.e. they are involved in altered production of Aβ, eventually causing neuronal cell death and dementia (3, 38).

LOAD, by contrast, can currently only be predicted by a range of risk factors of different nature. The heritability for LOAD is estimated between 58% and 79% (4, 40), leaving ‘other’ (i.e. non-inherited environmental, epigenetic) factors as possible modifiers. To date, one genetic factor has been established as significant modifier of LOAD; the ε4 allele of the apolipoprotein E gene (APOE-ε4), which is strongly enriched in rare familial forms of LOAD (40). APOE-ε4 is neither necessary, nor sufficient to cause LOAD, but instead operates as a genetic risk modifier. An individual who is homozygous for APOE-ε4 has an estimated 65% risk to develop AD at an age of 85 years, compared to a risk of only 20% with one APOE-ε4 allele or 10% when not carrying an APOE-ε4 variant, suggesting a dose-dependent effect of the APOE-ε4 mutation (41). This is in contrast to APP, PSEN1 and PSEN2 that are dominant causative mutations of AD. It has been suggested that APOE-ε4 could be involved in the same Aβ pathway as the EOFAD proteins (3). Found modifiers in a twin study might therefore also be of relevance for understanding both AD forms.

Several lines of evidence suggest that numerous additional LOAD modifiers – and probably also EOFAD loci – remain to be identified, since the four known genes account for probably less than 50% of the genetic variance in all AD forms together (3, 38).

As in HD and SCAs, MZ twins discordant for AD phenotype could be valuable in finding crucial modifiers among a range of potential candidates. In case of LOAD, one study estimates total MZ discordance rates of ∼39% (38). The MZ discordance rate is likely to be significantly lower in strongly inherited forms of AD like EOFAD, where genetics dominate. Indeed, only three MZ twin pairs discordant for a strongly familial form of AD have been recently reported, of which one is an EOFAD case (Table 1) (4). However, especially these cases constitute an interesting resource to discover crucial factors influencing the pathogenesis of AD. Of the described twins discordant for familial AD, two twin pairs showed discordance in AO and clinical symptoms in familial APOE-ε4 involved AD, and one pair was discordant for AD phenotype of the otherwise dominant mutation PSEN1 (4). Relevant to compare to are MZ twins concordant for familial AD, which have, for example, been described for APOE-ε4 involved AD (42, 43), but also large-scale studies toward genetic contribution in AD include MZ twins concordant for familial AD (44, 45).

Overall, cases discordant for strongly inherited forms of AD show that also in familial AD other factors than the causative gene may play a significant role in disease development.
Epigenetics

MZ twins can differ in the epigenetic status of their genome. Several studies have identified epigenetic differences either at selected genes of MZ twins (9) or in the overall epigenome (8). In general, the term epigenetics refers to ‘modifications of gene expression that are controlled by heritable factors other than DNA sequence, and are potentially reversible’ (12, 46). Epigenetic mechanisms are dynamic processes that are influenced by developmental stage, tissue type, environmental and stochastic factors. Typical examples of epigenetic mechanisms include DNA methylation and histone modifications, but also comprise other mechanisms including ATP-based chromatin remodeling, non-coding RNA-mediated gene silencing and transcription factor-binding mechanisms (46, 47). Histone modifications and DNA methylation particularly have been linked to phenotypic outcome, through altering transcription and genomic stability (2, 47). Epigenetic modifications have been found relevant for human disease, including many neurodegenerative disorders such as Parkinson disease (PD), AD and HD (2). In AD, hypermethylation of the major Aβ-degrading enzyme nphrinls and genome-wide altered histone acetylation patterns have been suggested as contributing to AD pathogenesis. In HD patients, elevated levels of histone methylation have been reported (2). Both differences in DNA methylation and histone modification have been shown to arise during the lifetime of MZ twins and have occasionally been linked to disease discordance (8, 9, 13, 25). In Beckwith–Wiedeman syndrome and NF1, for example, intra-twin differences in DNA methylation were correlated to differences in phenotypic expression of the disease (13, 25).

In human, DNA methylation occurs especially to the cytosine of CpG dinucleotides (2). The methylation status of regulatory regions can thereby influence gene expression or genomic stability, the latter of specific relevance for trinucleotide disorders (48).

For histone modification, acetylation and methylation have been linked to disease via influencing the formation status of DNA (2, 49). Neither histone acetylation nor histone methylation can be translated directly to gene expression profiles, but instead form a ‘code’ consisting of several methylation and/or acetylation events. Nevertheless, several studies have shown typical patterns of histone modification in regions such as promoters, enhancers, insulators, and coding regions that relate to gene expression, revealing high levels of histone acetylation and H3K4 methylation in promoter regions of active genes, but elevated levels of H3K27 methylation correlating with gene repression (2, 49). Evidently, investigating differences in histone-modification patterns as possible source for twin discordance is complex. One might therefore not assess both methylation and acetylation at the same time, but rather focus on subsets of histones. Indeed, detection of intra-twin differences in histone modification has been shown, by assessing H3 and H4 acetylation specifically (8). Interestingly, significantly different genome-wide acetylation was herein associated with gene overexpression.

Overall, acquired differences in epigenetic modifications can influence the phenotypic expression of disease, and as such could form an explanation for a variable disease phenotype in MZ twins, relevant also to singletons.

Genetics

Although MZ twins initially share all of their genotype and confirmation of monozygosity is even based on determining identical genotypes by assessing a range of single-nucleotide polymorphisms (SNPs) that have a high frequency in the general population, genetic differences may arise during embryonic development or later in life. Postzygotic mitotic recombination and other somatic events, such as deletions, single-nucleotide mutations, gene conversion, and repeat expansions have been proposed as possible genetic mechanisms causing MZ discordance (50, 51).

Indeed, there is accumulating recent evidence that MZ twins can differ genetically (10, 11). Although such studies are the first of their kind, these suggest that genetic differences could be one source of variation in MZ twins discordant for phenotype. The occurrence of \textit{de novo} CNVs in MZ twins has been reported (10, 11). The frequency of \textit{de novo} occurrence of these larger alterations has been estimated as \(~5\%\) per person per generation, indicating their relevance (10, 52). Furthermore, the occurrence of differing SNPs within sets of MZ twins has been documented (11), and the frequency of \textit{de novo} point mutations has been estimated on average as \(2.5 \times 10^{-8}\) per nucleotide per generation (or 175 mutations per diploid genome per generation) (53). On the basis of the idea that MZ twins can exhibit genetic differences, one could hypothesize that disease \textit{discordance} in MZ twins can also derive from acquired genetic differences, alternatively to epigenetic explanations. The detection of genetic differences relies on the comparison with a reference genome. In case of an MZ twin approach, the co-twins form each other’s reference. Since the rest of the genome of an MZ twin is identical, mismatches found could theoretically be selected for validation as possible mutations underlying disease discordance. That is, discordance for the autosomal dominant disease NF1 in an MZ twin pair has been explained by the presence of an \textit{NF1} mutation; the affected twin carried the \textit{de novo} \textit{NF1} mutation in all investigated cells, while the unaffected twin was mosaic (54).

Genetic discordance in strongly inherited disorders could also be an opportunity to determine crucial \textit{novel genetic modifiers}. The benefit of such an approach has recently been illustrated by a study wherein a nonsense mutation in the interferon regulator factor 6 was linked to development of the autosomal dominant Van der Woude syndrome, based on occurrence of that mutation in only the affected twin of a pair discordant for the disease (55). Furthermore, recently a set of genes likely involved in schizophrenia has been
Ketelaar et al.

identified, based on genomic differences within MZ twins discordant for the disorder (11). However, one should keep in mind that any genetic difference cannot always be linked to disease, since non-pathogenic, or false-positive differences can be present (29, 56).

Additionally to de novo SNPs and CNVs, somatic instability of a mutation could also explain disease variability, especially applicable to trinucleotide disorders. In HD models, for example, it has been shown that the CAG repeat size can highly differ within and between tissues, the extent of instability thereby influencing disease progression (57). In humans, such genetic variation within and between brain tissues has recently been associated with AO of HD (58). In the clinic, estimates for AO of HD are mainly based on CAG tract sizes determined in one tissue (mainly blood), and a significant proportion of variability in AO remains unexplained (7). However, by determining the CAG tract length of several types of tissues in addition to blood cells, and by using a sensitive detection method, one could possibly get an indication of the presence of somatic instability. Indeed, disease discordance in an MZ HD twin pair could be explained by a different abundance of the mutation, with the affected twin carrying the mutation in all cells assessed, while the yet unaffected twin was mosaic (21). But the inclusion of more than a single-cell type is also relevant for detection of somatic mosaicism of other type (i.e. conventional mutations).

The expected abundance of an acquired mutation is hard to estimate, since the percentage of cells involved is dependent on several factors, including type and timing of the mutational event (i.e. the earlier in life, the more cells could be affected), and tissue type involved, whereby disease-relevant tissues have occasionally been found to exhibit a higher abundance of the de novo mutation (57, 58). But in example, a de novo deletion present in only 10–20% of blood cells has recently been found in one twin of a pair discordant for disease phenotype (10).

In addition to epigenetic factors, MZ twins could thus form an interesting resource to determine crucial genetic factors underlying phenotypic variability in inherited disorders, paradoxically just because of their similarity in genetic background.

Methods to assess genetic differences: how to find the needle in the haystack?

Two major considerations when choosing the methods in a twins design toward detecting genetic differences are the abundance and type of the mutation.

Abundance influences the required sensitivity threshold of a method determining the detection of the mutation, but also influences the cell/tissue type to include.

The type (size) of mutation determines the required resolution of coverage of a screening method, in addition to the sensitivity. SNP genotyping arrays, for example, could also be used to determine CNVs, while genetic hybridization arrays are designed specifically for larger than SNP mutations. Next generation sequencing (NGS) offers a high coverage for detecting a broad range of genetic alterations, since it exhibits single-nucleotide resolution (59).

When investigating discordant MZ twins, a highly sensitive method with high resolution and genome-wide coverage should ideally be applied, such as NGS-based screens (60), as only a small number of intra-twin genetic differences are expected (10, 53). Moreover, a high resolution approach would allow detection of genetic differences of several sizes, varying from SNPs to CNVs and repeat expansions. Furthermore, a genome-wide coverage allows for a non-biased approach, not restricted to certain a priori selected regions. Loci-restricted genetic-screening approaches, such as Sanger sequencing or small pool polymerase chain reaction (61, 62), could be applied to validate suggested modifiers derived from genome-wide screens or other sources.

Epigenetic differences: how to map the epigenome?

Because of the dynamic nature of epigenetic modifications, they potentially form an excellent source for variation underlying disease discordance. In addition to stochastic factors, environmental factors can give rise to differences in the epigenome acquired over lifetime within MZ twin pairs (2). Thus, epigenetic modifications can form the molecular link between the historically assessed environmental differences and twin discordance.

Because often any a priori hypothesis is lacking regarding where in the genome twins can differ, a genome-wide epigenetic screen would be the preferred method. Most current genome-wide epigenetic approaches use NGS techniques (63), and thus high resolution and sensitivity can be accomplished genome-wide, allowing identification of novel disease-relevant regions.

On the other hand, site-specific epigenetic methods (63, 64) could be used when one has a clear candidate region in mind, to reduce costs, or to study one locus more extensively for different epigenetic modifications. That is, bisulfate conversion followed by pyrosequencing is commonly used to investigate site-specific DNA methylation patterns, while MetlyLight can offer sensitive quantitative analysis. Site-specific assessment of histone modifications is performed by immunoprecipitation followed by local sequencing. Local epigenetic methods, however, preclude the opportunity to map novel disease-relevant regions. Such methods are therefore often used to validate candidates derived from genome-wide epigenetic screens (64).

Environmental factors

Environmental factors remain a source for phenotypic variability. Beside stochastic factors, environment could affect phenotypic expression via induction of (epi)genetic alterations (65). Furthermore, environmental factors could influence phenotype via altering physiological and biochemical pathways. Head trauma
Mechanisms influencing genetic forms of neurodegeneration

for instance could be a risk factor in AD and PD, influencing physiology, while environmental toxins can form biochemical risk factors in AD, HD and PD (65, 66). Therefore, in addition to epigenetic or genetic screens, assessing the presence of major environmental differences will still add to the understanding of disease discordance.

Study design – some considerations

Using twin pairs to study phenotypic variability can be a powerful tool to detect crucial disease involved factors, which otherwise would be hard to filter from epidemiological studies using singletons. MZ twins are naturally matched for age, gender and genetic background, and subtle, but crucial differences in (epi)genetic status might arise and cause disease discordance. However, the success of such design depends on the comparison made, samples used, method chosen and selection criteria applied.

It is important to appropriately define discordance for the phenotype studied, on which successful comparisons and conclusions rely (5, 12, 56). A logical approach regarding the design would then be to perform at first an intra-twin comparison of the (epi)genome of discordant twins, an approach that has worked previously (9, 10). Each discordant twin set should be considered as a separate case study, as it is likely that every pair will exhibit their own type of discordance. Only for diseases showing often MZ discordance, a higher number of twins with the same type of discordance might be found to compare for shared differences. Withal, it is worthwhile to compare to twins discordant for the disease, to determine loci particularly relevant for disease discordance, while including (epi)genetic information of parents or siblings would enable to determine which variants have arisen de novo in either twin. Obtained loci could then be further selected for downstream validation (see also Fig. 2).

Blood-derived DNA is commonly used for (epi)genetic studies, but might not necessarily be representative for other (disease relevant) tissues (57). In case of neurodegenerative disorders, neuronal tissue would clearly be preferred, but since the availability may be problematic, hair-root cells might be used instead, which have developed from the same embryonic cell type (ectoderm) as neurons (21). Useful would be to first establish the true representativeness of blood-derived DNA and hair-root cells in (epi)genetic studies for neurological disorders, which requires parallel studying of such tissue compared to disease-relevant tissues in pilot cohorts of single individuals. Another point of consideration when using blood cell-derived DNA is the possible masking of intra-twin differences, due to intrauterine transfer of bone marrow-derived cells from one co-twin to the other via a shared placenta. This result in chimerism; when one twin obtains part of its blood cells from its co-twin. One third of all MZ twins are monochorionic, and chimerism is a well-known phenomenon (67). In such cases especially, more than one tissue type should be included to detect the existence and abundance of (epi)genetic differences.

Methodologically, it is clear that, to be able to successfully map intra-twin (epi)genetic differences in certain tissues, a sensitive method with high coverage and resolution is preferred, such as NGS-based screens. However, the main drawback is the inherent likely loss of specificity when increasing the sensitivity of NGS; a high sensitivity allows the detection of many (epi)genetic alterations, but will also increase the false-positive rate (59, 60). Therefore, when designing a genome-wide study toward intra-twin (epi)genetic differences, proper validation approaches should be considered. Possible strategies to select true intra-twin differences from false-positive (epi)genetic alterations before biological validation include bioinformatics and methodological validation. Adequate bioinformatics analysis can help filtering out methodological errors. In example, the CRISP method (68) assesses the consistency of reads between multiple pools, consistency between forward and reverse strand, and applies a likelihood-ratio model to estimate the probability of generated sequencing errors, all decreasing false-positive findings. Methodological validation is also required (62); by performing an (epi)genetic screen more than once, either with the same or a different array, repeatedly found alterations can be selected as likely true alterations. Dependent on the number of loci...
wherein the twins differ (epi)genetically, locus-specific validation of the found alteration can be performed, or the list can be narrowed down by functional and biological filter strategies (Fig. 2.) Once a selection has been obtained, the usual experimental validation procedures to link loci to phenotype can be started (56, 62, 63), and (epi)genetic mutations could be confirmed in subsequent cohorts of singletons showing disease phenotypes deviant from the expected, such as non-average AO cases.

Conclusion

The currently available (epi)genetic techniques allow for the investigation of possible differences in epigenetic or genetic status in MZ twins, in addition to environmental factors. By using MZ twins that are discordant for the genetic form of a disease, investigating small (epi)genetic differences may highlight factors significant for the phenotypic expression of the disorder, in addition to the main causative gene. Those modifying factors could be relevant for explaining phenomena such as reduced penetrance or high AO variability in the specific disorder (HD and SCAs), or for understanding the complex sporadic forms of a disease (AD). Performing (epi)genetic screening of MZ twins discordant for phenotype would increase our knowledge on the relative involvement of environmental, genetic and epigenetic factors in the expression of a disorder, and represents a novel approach to unravel disease mechanisms. Since MZ twins discordant for neurodegenerative disease are a rare phenomenon – especially in the genetic forms of a disorder – every additional single-case report will be valuable.

A better understanding of the crucial factors modifying the expression of a neurodegenerative disorder could then direct the focus of research toward relevant treatments, and would be of use for predictive testing and patient counseling.

Acknowledgements

The authors want to thank D. E. Erhnhoefer (PhD) for critically reading the manuscript. M. E. K. is supported by the UBC Graduate Student International Research Mobility Award, by the Award for Excellent Students from the University of Groningen, by the Prinses Beatrix Fund, and by the Huygens Scholarship for Talented Students from the Dutch government. R. M. W. H. is supported by grants from the Dutch Science Foundation (ZonMW) and the Maag-Lever-Darm stichting. R. M. W. H. is a University Professor at the University Medical Centre Groningen. M. R. H. is supported by grants from the Canadian Institutes of Health Research, the Huntington Society of Canada, the Huntington’s Disease Society of America, and the CHDI Foundation. M. R. H. is a Killiam University Professor and holds a Canada Research Chair in Human Genetics.

References

Mechanisms influencing genetic forms of neurodegeneration

66. Kell DB. Towards a unifying, systems biology understanding of large-scale cellular death and destruction caused by poorly liganded iron: Parkinson’s, Huntington’s, Alzheimer’s, prions, bactericides, chemical toxicology and others as examples. Arch Toxicol 2010: 84: 825–889.