Identifying the genetic components underlying the pathophysiology of movement disorders

Mario Ezquerra
Yaroslau Compta
Maria J Marti
Parkinson’s Disease and Movement Disorders Unit, Service of Neurology, Institute of Clinical Neurosciences, Hospital Clinic of Barcelona, IDIBAPS, CIBERNED, Spain

Abstract: Movement disorders are a heterogeneous group of neurological conditions, few of which have been classically described as bona fide hereditary illnesses (Huntington’s chorea, for instance). Most are considered to be either sporadic or to feature varying degrees of familial aggregation (parkinsonism and dystonia). In the late twentieth century, Mendelian monogenic mutations were found for movement disorders with a clear and consistent family history. Although important, these findings apply only to very rare forms of movement disorders. Already in the twenty-first century, and taking advantage of the modern developments in genetics and molecular biology, growing attention is being paid to the complex genetics of movement disorders. The search for risk genetic variants (polymorphisms) in large cohorts and the identification of different risk variants across different populations and ethnic groups are under way, with the most relevant findings to date corresponding to recent genome wide association studies in Parkinson’s disease. These new approaches focusing on risk variants may enable the design of screening tests for early or even preclinical disease, and the identification of likely therapeutic targets.

Keywords: genetics, movement disorders, Parkinson’s disease, parkinsonism, dystonia

Introduction

Movement disorders are a group of neurological conditions consisting of abnormal control of voluntary and automatic movements. According to whether they imply excessive or diminished speed, range and/or accuracy of movements, they can be classified as hyperkinesia or hypokinesia. Chorea and parkinsonism are, respectively, the paradigmatic syndromes of these two categories. Table 1 provides a summary of hyperkinetic and hypokinetic disorders. Movement disorders, taken together, constitute frequent neurological disorders. In particular, Parkinson’s disease and essential tremor, whose prevalence range from 1.7% and 4% among adults, are the two most frequent movement disorders, with the former being the second most frequent neurodegenerative disease, after Alzheimer’s disease. Other types of movement disorders, most importantly dystonia, are much rarer, but still clinically and socially relevant due to the disability they imply.

From an anatomo-functional point of view, movement disorders are generally viewed as a disturbance of the basal ganglia or their connections with other brain structures. Classically, a misbalance between two key basal ganglia neurotransmitters, namely dopamine and acetylcholine, has been used as a model to understand the dichotomy between hypokinesia (lack of dopamine, excess of acetylcholine) and
hyperkinesia (the opposite),2 although nowadays many other neurotransmitters, neuropeptides, and proteins have been implicated.

The last decades of the twentieth century have brought relevant progress in the knowledge of the genetic basis of movement disorders. Not surprisingly, the first movement disorders to be genetically characterized were those that had already, based on their description, been reported as being clearly familial, suggesting the presence of a Mendelian inheritance. Hence, Huntington’s disease, a dominant autosomal condition featuring a combination of chorea, psychiatric disturbances and dementia, was found in 1993 to be due to an excessive number of CAG repeats (polyglutaminopathy) within the huntingtin gene (chromosome 4).3,4 Almost simultaneously, Wilson’s disease, a recessive autosomal disease consisting of abnormal handling of copper mostly leading to familial brain and liver dysfunction which had already been mapped to chromosome 13 almost a decade before, was found to be due to a mutation of the ATP7B gene, encoding for a copper-dependent transmembrane ATPase, whose dysfunction leads to defective copper excretion.5,8

In contrast, most of the remainder of movement disorders, namely Parkinson’s disease (PD) and atypical parkinsonisms,9–11 along with dystonia,12 have been considered as sporadic disorders until recent decades, when the analysis of familial cases has enabled the characterization of a number of classic Mendelian genetic mutations, mostly in PD and primary dystonia.13,14 Notwithstanding the undeniable relevance of the identification of these monogenic forms, these account for a small proportion of cases. Thus, at present increasing attention is being paid to the complex genetics of these disorders, putting the stress on common risk genetic variants playing a partial role in these disorders, rather than on rare genetic mutations directly and univocally causing them. In this review, we describe the different strategies that have been used to establish the genetic basis of PD and dystonia, as examples of two frequent movement disorders, which has led to a better understanding of the molecular processes involved in these diseases.

Identifying the genetic components underlying the pathophysiology of movement disorders

Over the past decades, the genetics of movement disorders have been studied in several different ways, including linkage analysis in affected families, sib-pairs analysis, whole genome scanning methods and, through more simple association studies, candidate-gene strategies. Following these research methods, both causative mutations and risk gene variants have been studied. Interestingly, some of these causative or risk genes are common to different movement disorders, mostly parkinsonisms, suggesting that they share some pathophysiological mechanisms. In this section we present a general view of the main genetic research approaches available and suitable for studying movement disorders, mostly focusing on the candidate gene and large-scale methods.

The candidate gene approach

A traditional method investigating the potential impact of genes on genetic disorders is to perform a study focusing on a candidate gene, which is supposed to be involved in a given disease. The researchers analyze directly the gene or genes that may plausibly be involved in the disease, taking into account the related molecular biology background.15 These studies may include members of affected families in order to perform linkage analysis, or alternatively, unrelated cases and controls. In a “case-control” association study, the genotype frequencies of some polymorphisms located in a particular gene are analyzed in order to address whether there is or is not an increased frequency of a single allele or genotype in patients, compared with controls. These polymorphisms most frequently are single nucleotide polymorphisms (SNPs), that is, a change of a single nucleotide in a given region of the genome. The HapMap (http://hapmap.ncbi.nlm.nih.gov/) is a database of the location and frequencies of the SNPs in the human genome in different populations,16 which are useful to analyze and design genetic associations. An example of successful application of the candidate gene approach is the finding that some polymorphisms of the microtubule associated tau protein (MAPT) gene are associated with PD.17–19 This particular gene was considered a candidate gene for PD after the evidence of its implication in another parkinsonism, progressive supranuclear palsy (PSP).

An alternative to the analysis of polymorphisms is to sequence the codifying exons of a candidate gene, in order to

---

### Table 1 Hypo- and hyperkinesias

<table>
<thead>
<tr>
<th>Hypokinesias</th>
<th>Hyperkinesias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parkinsonism (bradykinesia, akinesia)</td>
<td>Chorea, dystonia, ballism</td>
</tr>
<tr>
<td>Freezing</td>
<td>Myoclonus, myokymia, myorhythmia</td>
</tr>
<tr>
<td>Rigidity</td>
<td>Tics</td>
</tr>
<tr>
<td>Stiffness</td>
<td>Tremor</td>
</tr>
<tr>
<td>Apraxia</td>
<td>Ataxia (dysmetria)</td>
</tr>
<tr>
<td>Cataplexy</td>
<td>Akathisia</td>
</tr>
<tr>
<td>Catatonia, obsessional slowness</td>
<td>Stereotypy</td>
</tr>
</tbody>
</table>
find a causative mutation in an affected family, which could be present in the patients, and absent in healthy controls. This was successfully applied in PSP where, besides the above mentioned polymorphisms approach, a complete sequence of the MAPT gene allowed for the identification of mutations leading to a PSP lookalike clinical picture.20 Again, this was possibly due to the a priori hypothesis that a causative MAPT mutation could exist in some cases in familial PSP after the observation of the presence of tau aggregates in brain parenchyma of PSP cases and, in a related neurodegenerative disease (frontotemporal dementia with parkinsonism linked to chromosome 17), was already mapped to the MAPT gene.21–23

Large-scale approaches

Genome-wide scan linkage analysis

The candidate approach is by definition not suitable for diseases where there are no a priori candidate genes. In other instances, it is inefficient, because it is necessary to analyze candidate genes one by one. In contrast, the large-scale approach is unbiased because it does not need an a priori hypothesis about which are the candidate genes or biological pathways involved in the disease.

The basic approach consists of the analysis of the segregation of multiallelic markers, called microsatellites, which are more informative variants than SNPs, and are distributed all over the genome.24 A set of about 300 microsatellites spaced uniformly along the human genome (Genome Wide Linkage) is the one currently in use. In affected families, 1 or few alleles of microsatellites can eventually be identified to co-segregate with the disease, giving statistical significant values (LOD score).25 A subsequent analysis of all the genes near this genetic marker could enable the identification of the gene involved in the disease. This linkage approach led to the discovery of LRRK2 gene mutations as a cause of familial PD.26

Additionally, the increasing knowledge on human genome variation and the technology advances on SNP genotyping have made it possible to simultaneously examine hundreds of thousands of SNPs in a single DNA chip array (http://www.affymetrix.com/technology/design/index.affx; www.illumina.com). This kind of analysis was successfully performed on a large pedigree affected with Parkinsonian-pyramidal syndrome, which led to the discovery of a disease-associated gene.27

Genome-wide scan association analysis

These new genotyping platforms (the so-called DNA chip arrays) can also be used in case-control association studies.28–30 The genome-wide association approach provides an effective and unbiased approach to discovering risk polymorphisms for genetically complex disorders. To perform this type of study it is mandatory to analyze large numbers of cases and controls, to use alleles covering the whole genome that can be inexpensively and efficiently genotyped, and to perform accurate statistical methods for the proper interpretation of such large-scale data. In contrast to the linkage studies, the main premise when performing a genome-wide association study (GWAS) is that extensive common variation in the human genome, exhibited by disease-producing alleles with frequencies greater than 1%, are responsible for the risk of most genetically complex disorders.31,32 For instance genetic variants in tau and synuclein (SNCA) genes, have been identified as risk factors for PD by using this methodology. This methodology is suitable for settings where the risk effect of these polymorphisms is relatively low but the pathogenic allele variants are frequent in the overall population.28,30 The opposite scenario is that of multiple rare allelic variants with low frequencies where this technique is not accurate as even larger cohorts would be necessary to reach enough statistical power, for example, the mutations in the glucocerebrosidase gene, where multiple rare variants cause a high increased risk of suffering familial and nonfamilial PD in a low number of cases.33,34

The main difficulty in elucidating the genetic causes and biological processes involved in complex diseases is that probably several factors contribute to their pathophysiology. Some of these genetics factors can either have an additive or an epistatic or multiplicative effect. Epistasis occurs when the effects of one gene are modified by another gene. The presence of epistatic effects is not usually detected with the GWAS approach.35 but a number of methods have been developed to tackle epistasis, such as the Hypothesis Free Clinical Cloning (HFCC) analysis which allows for searching genome-wide epistasis in a case-control design under a fast algorithm that tests a variety of different epistatic models in multiple combinations of SNPs. With HFCC samples are split into independent groups, where the analysis is taken further to look for consistent results across the groups.35,36 This approach could be explored in future studies to detect the interaction effects of different genes on the risk of parkinsonism.

At present, a number of hypotheses try to explain the different modalities of complex genetic associations. Thus, the common variant/multiple disease (CV/MD) hypothesis postulates that common genetic factors that contribute to a given disease, under a certain combination of interacting genes and environmental factors, may contribute to the risk of another, related, disease when occurring in other genetic
or environmental backgrounds. Alternatively, the common disease/rare variant (CD/RV) hypothesis postulates that in some cases, some rare genetic changes with relatively high penetrance, contribute to genetic susceptibility to common diseases. The H1 MAPT haplotype, which is frequent in the normal population and a known risk factor for both PD and PSP (probably depending on additional risk factors as the rs242557 polymorphism, which would act by modifying the final presentation of the parkinsonism) would be an example of the former model, whereas the rare glucocerebrosidase mutations that cause Gaucher disease and are also known to increase the risk for PD would be an example of the latter.

Finally, it is noteworthy that SNPs are just one source of variation in the human genome. Other types of mutations, like small deletions and insertions of few nucleotides, or large-scale mutations, including gene duplications, chromosomal translocations, and inversions or deletions, could have implications in disease etiology. Specifically, copy-number variants (CNVs) are DNA fragments in which copy-number differences have been found in different genomes, ranging usually from a few hundred bp to several million bp in size. The importance of the CNVs in the total genetic variability and disease-susceptibility in human populations has recently been highlighted. Therefore, the GWAS SNP association studies could be supplemented with CNVs association analysis, which could be addressed by using the same SNP genotyping platforms. Other genomic analysis platforms can be used to detect specifically the CNVs, like the genomic hybridization method (CGH) or isothermic oligonucleotide-based CGH arrays.

Massive sequencing

The new systems of massive sequencing (http://www.454.com/: http://www.illumina.com/technology/sequencing_technology.ilmn) make it possible to obtain the complete genomic sequence or, alternatively, to obtain the complete sequence of all the exons of the genes of the genome in a single individual at a reasonable speed. This provides a very informative method of analyzing the human genome, and offers new possibilities in the search for genetic risk or causative factors, which could apply in the future to the study of movement disorders. This approach is particularly useful for studying rare disorders with an expected frequency of mutation <1% (Table 2), where the sequence of an affected individual can be compared to the SNPs present in healthy controls from the HapMap and dbSNP database (www.ncbi.nlm.nih.gov/projects/SNP/), allowing the common SNPs present in healthy subjects to be filtered out, thus removing the noninformative noise. In silico methods allow for analysis of other SNPs to determine the most probable candidate rare variants, depending on their location in the genome and inside the genes, and are investigated in detail to further confirm their pathogenicity.

Furthermore, an ongoing international scientific project, the 1000 Genomes Project (http://www.1000genomes.org/) may have important implications for genetic diseases research. This project consists of massive sequencing of the complete genome of hundreds of individuals from different populations worldwide by using massive sequencing platforms, with the final goal of providing information about the overall human genetic variability, including low frequency polymorphisms in different human populations. The combination of these genotype data with specific data from a GWAS will allow for imputing computationally the genotypes from a given sample for the additional genetic variants of the 1000 Genomes study, without need of direct genotyping. Thus, this information will help to finemap the genetic regions associated with a disease and will suggest, “suspect” SNPs likely to be involved in the studied disease.

Overview of genetics of PD and other parkinsonisms

PD

Monogenic PD

Monogenic PD is believed to account for around 15% of cases and includes both dominant and recessive forms (the so-called PARK variants) which are summarized in Table 3. Because this review is devoted to the genetics of movement disorders in general and thus precludes a detailed revision of every single monogenic form of the large variety of known movement disorders, this section focuses on the three monogenic forms of PD considered the more biologically and clinically relevant for a number of reasons. The first monogenic form of PD (PARK1) was originally detected in 3 unrelated Greek families and 1 Italian kindred with autosomal dominant inherited PD characterized by a typical clinical phenotype but with young age at onset. The mutation in those families was a single base pair change at position 209 from G to A (G209A) resulting in an Ala to Thr substitution at position 53 (A53T) of the alpha-synuclein (SNCA) gene (chromosome 4). This gene encodes for the homonymous presynaptic nerve protein, which is believed to

- Ezquerra et al
be involved in synaptic vesicle formation and exocytosis. \textsuperscript{51} The A53T mutation is predicted to disrupt the α helix and extend the β sheet structures, thought to be involved in protein misfolding and self-aggregation. Soon after the report of this mutation, and using immunohistochemical techniques, Spillantini et al found that alpha-synuclein is the main component of Lewy bodies and neurites, the histopathological hallmark of Parkinson’s disease, suggesting that SNCA and its protein product are key not only to familial but also to sporadic PD. \textsuperscript{52} In recent years a second \textit{SNCA} mutation in a Basque kindred (E46K) added to the first one, with the peculiarity that those patients developed dementia early in the disease course in contrast to the A53T cases. \textsuperscript{53} The E46 K mutation substitutes glutamic acid with lysine in a highly conserved area of the protein and is, thus, likely to result in protein dysfunction. These first observations of mutations leading to protein misfolding or dysfunction of some type, have been followed in recent years by further reports of \textit{SNCA} gene duplications and triplications (PARK4). \textsuperscript{54,55} These studies indicate that not only protein dysfunction resulting from gene mutations, but also an overexpression of the protein can enhance protein aggregation and subsequent formation of Lewy bodies. It is noteworthy that cases with gene duplication are very similar to typical PD whereas those with a gene triplication have younger age at onset and associate early dementia, suggesting a dose-response effect for the \textit{SNCA} gene. However, these \textit{SNCA} gene mutations and duplications/triplications, albeit clearly relevant for a better knowledge of PD pathophysiology, constitute an infrequent cause of PD.

The second gene mutation causing familial PD was identified in 2 Japanese families carrying homozygous or compound heterozygous exon deletions in the \textit{parkin} gene (located on chromosome 6; PARK2), which encodes for a protein similar to the ubiquitin family. \textsuperscript{56} It has been speculated that the homozygous mutation of this gene could partly account for a defective function of the proteosome and lead to multi-ubiquitin-like protein aggregates. Carriers of these

### Table 2: Overview of the strategies used in genetic research depending on the disease prevalence and assumptions about the genetic heredity

<table>
<thead>
<tr>
<th>Initial candidate gene</th>
<th>No candidate gene</th>
<th>Type of polymorphism usually analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linkage analysis</td>
<td>GWS</td>
<td>SNPs, microsatellite polymorphism</td>
</tr>
<tr>
<td>Associations studies</td>
<td>GWAS</td>
<td>SNPs, CNVs</td>
</tr>
<tr>
<td>Sequencing</td>
<td>Massive sequencing</td>
<td>SNPs, CNVs, microsatellite polymorphism</td>
</tr>
</tbody>
</table>

**Abbreviations:** GWS, genome wide scan; GWAS, genome wide association study; SNP, single nucleotide polymorphism; CNV, copy number variation.

### Table 3: Summary of monogenic forms of Parkinson’s disease (PD)

<table>
<thead>
<tr>
<th>Name</th>
<th>Disease</th>
<th>Inheritance</th>
<th>Chromosome</th>
<th>Protein</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARK1</td>
<td>Familial PD</td>
<td>Autosomal dominant</td>
<td>4q21-q22</td>
<td>α-synuclein</td>
<td>Synaptic protein</td>
</tr>
<tr>
<td>PARK2</td>
<td>Young-onset PD</td>
<td>Autosomal recessive</td>
<td>6q25.2-q27</td>
<td>Parkin</td>
<td>Ubiquitin-protein ligase</td>
</tr>
<tr>
<td>PARK4</td>
<td>Familial PD</td>
<td>Autosomal dominant</td>
<td>4q region</td>
<td>α-synuclein duplication or triplications</td>
<td>Synaptic protein</td>
</tr>
<tr>
<td>PARK5</td>
<td>Familial PD</td>
<td>Autosomal recessive</td>
<td>4p15</td>
<td>Ubiquitin carboxy-terminal hydrolase</td>
<td>Hydrolase of ubiquitin</td>
</tr>
<tr>
<td>PARK6</td>
<td>Young-onset PD</td>
<td>Autosomal recessive</td>
<td>1p35-p36</td>
<td>PINK-1</td>
<td>Mitochondrial anti-stress-induced degeneration</td>
</tr>
<tr>
<td>PARK7</td>
<td>Young-onset PD</td>
<td>Autosomal recessive</td>
<td>1p36</td>
<td>DJ-1</td>
<td>Sumolation pathway</td>
</tr>
<tr>
<td>PARK8</td>
<td>Familial PD</td>
<td>Autosomal dominant</td>
<td>1p11.2-q13.1</td>
<td>LRRK2</td>
<td>Protein phosphorylation</td>
</tr>
<tr>
<td>GBA</td>
<td>Familial PD</td>
<td>Autosomal dominant</td>
<td>1q21</td>
<td>Gluco-cerebrosidase</td>
<td>Membrane lipid</td>
</tr>
</tbody>
</table>
homozygous mutations clinically present with a well-defined clinical phenotype: that of autosomal recessive juvenile parkinsonism (AR-JP), which very often starts with kinesigenic foot dystonia and is additionally characterized by a combination of levodopa-responsive parkinsonism, susceptibility to levodopa-induced dyskinesia and motor fluctuations, and a variety of psychiatric symptoms.57 This particular clinical phenotype, the juvenile onset, and the lack of Lewy bodies at autopsy, suggest that this is not properly PD and the name “parkin-gene disease” has been proposed to designate these cases.57,58 Interestingly, simple heterozygous parkin gene deletion (haploinsufficiency) has also been suggested to be enough in some cases to induce parkinsonism, perhaps acting as a susceptibility or risk factor in conjunction with other genetic or environmental risk factors, but to date this remains an open question.59

Finally, in 2004, 2 independent groups reported on a new gene causing autosomal dominant PD (PARK8): the leucine-repeate rich kinase type 2 (LRRK2) gene, initially also named dardarin after the Basque word “dardare” which means tremor (one of the first families identified carrying mutations in these gene was Basque).26,60 LRRK2 encodes for a protein member of the ROCO family and although its function is still not fully understood, it is seemingly related to cellular cycle regulation via protein phosphorylation.26 The identification of LRRK2 is generally viewed as a relevant contribution to PD genetics for several reasons. First, because LRRK2 mutant PD cases are relatively frequent in familial cases (6%) and even in cases otherwise considered as sporadic due to negative family history (between 1% and nearly 4%), and are being identified across the different continents and in different ethnic groups.61,62 Furthermore, most LRRK2 mutants feature the classic PD clinical picture not only because of their symptoms but also at onset. All this is in complete contrast to the low prevalence of SNCA mutations and to the fact that parkin-related cases may constitute an independent nosological condition. The above discussed characteristics of LRRK2-related PD make it an attractive model to study the premotor phase of PD. First, the high frequency of a LRRK2 mutations in PD and their variable age at onset make it easy to find healthy mutation carriers and, second, LRRK2 mutation-associated PD are clinically indistinguishable with idiopathic PD, probably having a similar physiopathology. Nonmotor symptoms herald, for many years in some instances, the onset of the cardinal motor signs of the disease and might constitute a very useful means of improving our knowledge of the presumably long preclinical phase of the disease, and eventually assessing candidate preventive or neuroprotective strategies.63 However, the fact that the penetrance of the LRRK2 mutations is incomplete (with figures around 60%–70% at 70–80 years), and that in some instances LRRK2 cases lack underlying Lewy-type pathology, constitute caveats to this approach.26,61,64

### Complex genetics of PD

Most of PD cases are still sporadic, though. This, along with compelling evidence that there must also be environmental factors in PD pathogenesis, has suggested that, as in many other human illnesses, complex genetics may be important in PD.65 Under a complex genetic model certain genetic variants not causative of disease per se (the so-called gene polymorphisms) might cause minimal dysfunction which might lead to neurodegeneration when interacting with other gene variants, exposure to the aforementioned environmental risk factors, and/or lack of exposure to possible environmental protective factors.65 Such might be the case, though still speculative, of the above discussed heterozygote parkin mutations.59 Another widely investigated and discussed example would be that of the heterozygous mutations of the glicocerebrosidase gene, responsible for Gaucher’s disease, the most frequent lysosomal storage disorder when occurring in homozygosity.66 But even before the identification of SNCA mutations, variants of the MAPT (the microtubule associated tau protein) gene, and particularly its H1H1 haplotype, had been consistently shown as a significant risk factor for PD.17,67 This has opened the debate about how this gene is involved in a neurodegenerative disease that does not belong to tauopathies (neurodegenerative diseases where the pathological hallmark is the deposition of tau protein aggregates within the brain). One of the proposed explanations is that the H1H1 haplotype of the MAPT gene might enhance neuronal loss in PD, whereas other MAPT genetic changes observed in the tauopathies would account for an over-expression of the tau protein itself favoring its hyperphosphorilation and abnormal deposition.41

With this premise, and taking advantage of the recent developments in genetics, most remarkably GWAS, several attempts have been made to identify different genetic risk variants in PD. The two largest GWAS published to date, of European and Japanese ancestries respectively have been consistent in: a) the finding that, despite being a rare cause of monogenic PD, the SNCA gene also constitutes a significant risk factor for sporadic PD; b) a trend towards more representation of a LRRK2 polymorphism among PD patients than in the control populations; c) the confirmation of a new locus implicated in PD: PARK16, located in chromosome
1q32. In contrast they diverge in terms of MAPT association being confirmed as a genetic risk factor for PD in the European study only, whereas the new locus BST1 (located in chromosome 4p15) was identified in the Japanese study solely.69,68 However, this latter locus has recently been replicated in both French and Dutch GWAS studies.69,70 Other gene variants such as the cyclin G-associated kinase (GAK) are surfacing as relevant PD risk factors in other GWAS association studies.70,71

Atypical parkinsonisms

Monogenic forms of atypical parkinsonisms

PSP and multisystem atrophy (MSA) are the most frequent forms of the so-called atypical parkinsonisms, which constitute a difficult and important differential diagnosis with PD.10,11

Although no clear monogenic forms of MSA have been identified, an autopsy confirmed case from a German MSA family with involvement of 2 successive generations has been reported, and increased frequency of parkinsonism in first degree relatives has been found among MSA patients in a recent case-control study.72,73 Regarding PSP, mutations of the MAPT gene have rarely been associated with a clinical picture of PSP.30,74

Complex genetics of atypical parkinsonisms

As mentioned above, linkage association techniques and the study of SNPs have also helped established the MAPT H1 haplotype and more specifically the H1c subhaplotype (probably more related to the expression of the gene and subsequently with its brain protein-dosage) as a risk factor for PSP, to an even greater extent than PD.58,75,76

A SNP association study has found a risk association between the SNCA gene and MSA — again not surprisingly, as MSA is a synucleinopathy in the same way as PD, although the two diseases differ in that in the former the alpha-synuclein aggregates are intraglial (glial citoplasmatic inclusions), and in the latter they are intraneuronal (Lewy bodies and neurites).77

Ongoing and future GWAS studies in PSP and MSA are expected to add more information to the complex genetics of both diseases.

Overview of genetics of dystonia

Dystonia is a movement disorder characterized by involuntary sustained muscle contractions that lead to twisting and repetitive movements and abnormal postures.1 According to the etiology, dystonia can be classified as primary torsion dystonia (PTD) and nonprimary dystonia. PTD is a syndrome where only dystonia is present (except for tremor) and there is no evidence of acquired cause or neurodegeneration. Nonprimary dystonia includes the “dystonia plus” (syndromes which are inherited disorders with additional neurological signs besides dystonia, but without neuropathologically detectable lesions), dystonia in heredodegenerative diseases, and dystonia of exogenous causes.

PTD is the most frequent cause of dystonia. Seven genes have been mapped for PTD including DYT1, 2, 4, 6, 13, and 17, but only 2, DYT1 (TOR1A) and DYT6 (THAP1), have been identified.14 Most are associated with an early onset generalized phenotype, except for DYT7 and 13, which have been associated with adolescent or adult-onset focal or segmental dystonia. In this section, following the premise of the PD and parkinsonisms section, only DYT1 and 6 are reviewed as the 2 most relevant examples of monogenic forms of dystonia. No information is available yet on complex genetics of dystonia and cooperative national and multinational initiatives are warranted.

DYT1 dystonia

The DYT1-PTD is the most common and severe form of early-onset dystonia. It is a childhood onset disease with a mean age at onset of 12 years. The onset of the disease is often focal with subsequent generalization (65% of cases) and typically affects legs, trunk and arms, and much more rarely cranial structures. DYT1 dystonia is inherited as an autosomal dominant trait with a reduced penetrance of ∼30%.78 The disease is caused by an in-frame GAG deletion in the TOR1 gene, in the long arm of the chromosome 9 (9q34), which removes 1 single glutamic acid residue in the carboxy terminal region of the encoded protein.79 The TOR1 gene is composed of 5 exons and its expressions is regulated by transcription factors of the Ets family. These bind to 2 Ets binding cores in the upstream region of the gene.80

TOR1 encoded torsin A, a 332-amino acid protein, is a member of the AAA+ family of ATPase, which has been associated with different functions, including protein processing and degradation, intracellular trafficking, vesicle recycling and citoeskeletal dynamics.81 Torsin A is widely expressed in human tissues. Both torsin A protein and mRNA have been found in dopaminergic neurons of pars compacta of the substantia nigra, and neurons within the cortex, striatum, thalamus, hippocampus, cerebellum, midbrain, pons, and spinal cord.82–84 A single pathological study carried out in 4 DYT1 dystonia brains found perinuclear inclusion bodies...
in midbrain reticular formation and periaqueductal gray matter, but so far this finding has not been replicated, with the remainder of pathological studies not having detected any specific abnormalities.

Different studies have indicated that torsin A functions at different subcellular locations including the nuclear envelope, the endoplasmic reticulum, and secretory and synaptic vesicle compartments. Mutant torsin A has been observed to interfere with the linkage between the nuclear envelope/endoplasmic reticulum membranes and the cytoskeleton, which may affect neurite extension during brain development. Other studies have suggested an abnormal postnatal maturation affecting the neurons (mainly at the synaptic level), but also the glia.

Based on its similarities with the AAA+ATPases and the head shock proteins (HSP), torsin A has also been suggested as likely to act as a chaperone. Torsin A has been found to co-localize with HSPs, ubiquitin and alpha-synuclein in Lewy bodies in patients with Parkinson’s disease and its over-expression prevents the formation of alpha-synuclein aggregates in a cell culture model and in polyglutamine repeat-induced protein aggregates in *Caenorhabditis elegans.*

Several studies carried out in different mouse models generated to investigate the role of torsin A in dystonia suggest that the abnormal brain networks in *DYT1* dystonia are complex, involving multiple neurotransmitters, and probably areas. In these models, altered levels of dopamine metabolites, attenuation of release of dopamine after amphetamine stimulation, and lack of the normal inhibition of cholinergic interneurons activity after D2 dopamine receptor stimulation, have been detected, indicating that possible impairment between striatal dopaminergic and cholinergic signaling. However, increased levels of serotonin have also been observed. In agreement with the aforementioned pathological study in human *DYT1* dystonia brains, some of these transgenic animal models have also showed perinuclear inclusions and aggregates of ubiquitin and Torsin A. To sum up, and despite the numerous findings from all these studies, it can be concluded that the precise function of torsin A in dystonia remains elusive.

**DYT6 dystonia**

*DYT6* dystonia is, like *DYT1* dystonia, a dominantly inherited dystonia, but in contrast with the latter condition, at presentation it involves the cranio-cervical musculature and the arms and only rarely the legs. Furthermore, the degree of progression to other regions varies remarkably in *DYT6* dystonia. The gene is found in chromosome 8p21-q22 and the overall prevalence of the mutation ranges between 1.0% and 2.5% in PTD cohorts variably selected on the basis of family history, early onset, generalization and involvement of laryngeal or cervical regions.

In 2009, Fuchs and coworkers initially reported 2 mutations in the *THAP1* gene (THAP domain containing, apoptosis associated protein 1) causing *DYT6* dystonia in 4 ancestrally related Amish-Mennonite families and in an unrelated family of German ancestry. Subsequently, large genetic screening studies have identified more than 45 additional mutations in PTD patients of European Caucasian ancestry, but also in Brazilian and Chinese patients.

*THAP1* is a member of the THAP (Thanatos-associated protein) family of proteins, which contain an evolutionarily conserved zinc-dependent DNA-binding domain. In addition to the THAP domain at the N-terminus, THAP1 has a low complexity-proline rich region, a coiled-coil domain and nuclear localization signal (NLS) at its C-terminus. Besides its DNA binding function, THAP1 regulates cell proliferation and can function as a nuclear proapoptotic factor, after its interaction with Par-4, a proapoptotic factor linked to prostate cancer and neurodegenerative diseases. There are no specific data on THAP1 function in the brain, though.

Most mutations described in *DYT6* are thought to eliminate the DNA-binding function of the protein. These include missense mutations within the DNA binding domain, substitution mutations that disrupt the nuclear localization signal in the C-terminus, or nonsense and frameshift mutations that truncate the protein within the DNA binding domain or before the nuclear localization signal. A smaller number of missense mutations have also been identified in the C-terminal end of the protein and near the coiled-coil domain.

The nuclear localization of THAP1 suggests a potential interaction with torsin-1A and related pathways. In this sense, 2 studies have reported an interaction between THAP1 and TOR1A. THAP1 protein binds directly to the TOR1A promoter in cell lines, primary cells and mouse brain tissue and this interaction is disrupted by pathogenic THAP1 mutations. In addition wild-type THAP1 represses the expression of TOR1, whereas dystonia-associated mutant THAP1 results in decreased repression of TOR1. These data indicate transcriptional dysregulation as a cause of dystonia and link the molecular pathways of *DYT1* and *DYT6* in dystonia.

**Identifying targets for new therapies**

Genes or proteins involved in the pathological process of a particular disease are possible targets for drug
development and subsequently for clinical trials. For instance, alpha-synuclein is an obvious target for the development of pharmacological interventions in PD and related disorders also featuring alpha-synuclein brain deposition. Therefore, elucidating the exact biological mechanisms by which alpha-synuclein causes or is associated with neurodegeneration, or the upstream molecular events that eventually lead to alpha-synuclein alteration, would have a great impact on the design of new disease-modifying drugs or strategies. However, the mechanisms by which alpha-synuclein leads to neurodegeneration still remain elusive, as it has been proposed that enhanced rates of fibrillar a-synuclein and their deposition on Lewy bodies could well be the cause of the neurotoxicity, but other studies have suggested that small prefibrillar oligomers of alpha-synuclein might be the actual toxic species. 125 Nevertheless, SNCA mutations causing PD in a dominant manner are extremely rare, and other additional genes with different or unknown functions have been implicated in familial PD, in some instances without underlying alpha-synuclein deposition or Lewy bodies (parkin and some cases of LRRK2). 26,58,64 Furthermore, some of these genes are involved in proteosome or autophagic process (parkin, LRRK2, DJ1), thus suggesting that they could be involved in the incorrect processing of misfolded or aggregated alpha-synuclein. 125 Therefore those events, likely to lay upstream of synuclein deposition, could also be targets of pharmacological intervention. Finally, mutations in parkin, LRRK2, DJ1, and PINK1 genes have been described to produce mitochondrial alterations, specifically in the complex I of the mitochondrial respiratory chain, suggesting that antioxidant treatments or modulators of mitochondria function could be candidate strategies for treating PD as well. 126–130

In addition, under the CV/MD hypothesis, other parkinsonisms that would share with PD common genetic risk factors could also have similar underlying altered biological systems. For instance, there is recent evidence that mitochondrial dysfunction, largely addressed in PD, may also be impaired in PSP and in dementia with Lewy bodies, and the same may apply for proteosome dysfunction in MSA and PSP. 131–133 However, the complete and precise picture of all the events occurring in these different diseases, and which are affecting each biological system likely to be involved, such as the mitochondria, oxidative stress, autophagy, proteosome, and final alpha-synuclein deposition, are not fully understood yet.

Conclusion

In recent years, relevant progress in the knowledge of the genetic basis of movement disorders has been achieved. Using different genetic studies different causative Mendelian gene mutations have been identified and more recently growing attention is being paid to genetic risk variants of these diseases. In complex diseases, like PD, a combination of genetic causative or risk factors, and interactions between these risk factors, could be involved in the etiology of the disease. The prevalence of each disease, together with our assumptions about the characteristics and frequencies of the genetic variants in the affected populations, will condition the research methods to be prioritized. In addition, common genetic factors are involved in different related diseases and could show common pathological biological pathways. The finding of the parkinsonism-associated pathways will allow developing animal or cell models of these conditions, in order to focus research on the related molecular mechanisms, which could eventually lead to the identification of therapeutic targets.

Disclosure

The authors declare no conflicts of interest in this work.

References


