Non-Mendelian Genetic Aspects in Spinocerebellar Ataxias (SCAS): The Case of Machado-Joseph Disease (MJD)

Manuela Lima¹,², Jácome Bruges-Armas¹,³ and Conceição Bettencourt¹,²,⁴

¹Genetic and Arthritis Research Group, Institute for Molecular and Cell Biology (IBMC), University of Porto, Porto
²Center of Research in Natural Resources (CIRN) and Department of Biology, University of the Azores, Ponta Delgada
³Serviço Especializado de Epidemiologia e Biologia Molecular, Hospital de Santo Espírito de Angra do Heroísmo
⁴Laboratorio de Biología Molecular, Instituto de Enfermedades Neurológicas, Fundación Socio-Sanitaria de Castilla-La Mancha, Guadalajara
¹,²,³Portugal
⁴Spain

1. Introduction

Monogenic disorders of Mendelian nature, defined as those resulting of mutation at a single locus, and in which the observed alteration is both necessary and sufficient for phenotypic manifestation (Gropman & Adams, 2007), constituted, until recently, the main target of gene-finding studies. Mendelian or otherwise “simple” phenotypes are frequently referred in the scientific literature in opposition to the “complex” ones; the designation of “Mendelian”, therefore, should reflect the occurrence of such diseases in accordance with simple, predictable family patterns, with a single locus determining its manifestation. It has, however, become very evident that even in the case of individual causative genes, the associated phenotypes can display attributes which result in non-Mendelian patterns of the trait or disease whose expression is being considered (Gropman & Adams, 2007; Sherman, 1997). In practical terms, this implies that the number of diseases for which the respective phenotypes can be explained by the effect of mutations at a single locus is dramatically diminishing (Gropman & Adams, 2007). Several diseases, initially characterized as monogenic, are now known to be modulated by a yet undetermined number of loci. Incomplete genotype-phenotype correlations observed in a large number of diseases have, therefore, forced the widening of the monogenic model, to allow the accommodation of the remaining factors, which can potentially explain the spectrum of the phenotypic variability (Badano & Katsanis, 2002). The incompleteness of the genotype-phenotype correlations seen in such situations confirms that the product of the primary mutation is imbedded in a highly complex system, in which polymorphic variation, mutations at other loci as well as environmental variables modulate the differences observed amongst individuals (Van Heyningen & Yeyati, 2004).
Factors that produce atypical or irregular patterns of inheritance are frequently referred as “complicating factors”. The understanding of the mechanisms on the basis of such patterns is pertinent not only at a theoretical level, but also due to implications in terms of diagnosis and genetic risk assessment, conditioning, furthermore, the ability to predict the initiation and course of disease (Haines & Pericak-Vance, 2006; Van Heyningen & Yeyati, 2004). On the other hand, in the context of research related with the identification of deleterious genes, such irregularities can imply, amongst other aspects, severe obstacles in the interpretation of pedigrees, as well as difficulties in the selection of families (e.g., in the context of linkage studies). Several of such complicating factors are frequently cited, namely clinical variability, expressivity, pleiotropism, anticipation, incomplete penetrance, age-dependent penetrance and meiotic drive. Underlying these observations are mechanisms such as allelic and locus heterogeneity, presence of modifier genes, intergenerational instability, somatic instability, genomic imprinting and de novo mutations (Gilchrist et al., 2000).

Amongst the several diseases with a known causative mutation but which, nevertheless, display complicating features are those associated with triplet repeat expansions. Four classes of triplet repeat disorders are usually considered, based on the location of the repeat motif in 5’ or 3’ untranslated regions, in introns or in coding regions (revised in Bettencourt et al., 2007). Polyglutamine (Poly-Q) disorders are part of this “expansion disorders” group, being caused by a CAG repeat expansion in the coding region of the respective causative genes; enclosed within this group are several spinocerebellar ataxias (SCAs), namely SCA1, SCA2, SCA3, SCA6, SCA7, SCA17 as well as dentato-rubro-pallido-luysian atrophy (DRPLA). Poly-Q diseases exhibit atypical features, difficult to integrate in the classic mendelian expectations (Tsuji, 1997). Machado-Joseph disease (MJD/SCA3) is considered the most frequent SCA worldwide (Cagnoli et al., 2006; Paulson, 2007; Schöls et al., 2004); using MJD as a paradigm, this review aims to explore complicating, non-Mendelian aspects of Poly-Q SCAs, from a perspective that synthesizes the current knowledge concerning such complicating factors as well as their implications at several levels, namely at the level of genetic counseling (GC).

2. Machado-Joseph disease: general perspective

Machado-Joseph disease, also known as spinocerebellar ataxia type 3 (MJD/SCA3 - MIM 109150) is an autosomal dominant neurodegenerative disorder. Described as a disorder of adulthood, with an average age at onset rounding 40 years (Coutinho, 1992), this disease has, nevertheless, reported onset extremes of 4 (Carvalho et al., 2008) and 70 years (Coutinho, 1992). Average survival time is of 21 years (Coutinho, 1992; Kieling et al., 2007). MJD is characterized by a complex and pleiotropic phenotype, involving the cerebellar, oculomotor, pyramidal, extra-pyramidal and peripheral systems. The high clinical variability observed in this disorder has led to its systematization into three clinical types, which can occasionally be present in a single family (Coutinho & Andrade, 1978).

MJD was originally described in North American patients of Azorean ancestry, residing in the United States (Nakano et al., 1972; Rosenberg et al., 1976; Woods & Schaumburg, 1972). The history of its initial description, as three distinct clinical entities, clearly reflects the high phenotypic variability that characterizes this disorder, whose unification was dependent of the identification, in a single family, of the different clinical forms that were described in the original reports (Coutinho & Andrade, 1978). The common ancestry of the three families...
that were described between 1972 and 1976 (known as Machado, Thomas and Joseph), has largely conditioned the interpretations about the origin of the disease, initially considered as Azorean, and designated as “Azorean disease of the nervous system” (Romanul et al., 1977). The molecular screening of MJD’s causative mutation allowed, afterwards, the differential diagnosis, which has led to an epidemiological profile clearly distinct from the one obtained on the basis of clinical criteria alone (Lopes-Cendes et al., 1996). In Portugal, MJD represents about 58% of the families with dominant ataxias (Vale et al., 2009). In the Azores, more precisely in the small island of Flores, the disease achieves the highest values of prevalence reported worldwide (Bettencourt et al., 2008a; Lima et al., 1997).

Two main studies have addressed the issue of the worldwide origin of the MJD mutation. Gaspar et al. (2001), using flanking and intragenic markers, have identified two main haplotypes in 94% of the families studied. In the Azores, these two haplotypes were present, and were related with the islands of higher prevalence (Flores and São Miguel), indicating that two mutational events were responsible for the presence of MJD in the families of Azorean origin, a result previously disclosed by genealogical analysis (Lima et al., 1998). Aiming to determine the origin, age and dispersion of these two main mutational events, Martins et al. (2007) have conducted a more extensive haplotype analysis, which revealed that the TTACAC lineage, the most frequently found in the expanded alleles of patients from all over the world, achieved its highest variability in Asia (specifically in the Japanese population). A “Short Tandem Repeat” (STR) based haplotype was identified in this population and an approximate age of 5774±1116 years was suggested.

The MJD locus was mapped to 14q24.3-q32 in 1993 (Takiyama et al., 1993). In the following year, Kawaguchi and colleagues (Kawaguchi et al., 1994) isolated and characterized a cDNA clone, designated as MJD1a, identifying MJD’s causative mutation as an expansion of a CAG motif in the coding region of the ATXN3 gene. Initially described as containing 11 exons, ATXN3 spans a genomic region of around 48 kb, with the CAG tract located in exon 10, at 5’ (Ichikawa et al., 2001). Two novel exons were identified recently, in a study that used information from cDNA clones of Azorean MJD patients and controls (Bettencourt et al., 2010a). In the MJD locus, normal chromosomes have from 12 to 44 CAG repeats, whereas in patients, usually heterozygous, the number of repeats in the mutated allele consensually ranges between 61 and 87 (Maciel et al., 2001). Intermediate alleles are rare and, as a result, their behavior is poorly understood. For example, in a family described by Maciel and colleagues (Maciel et al., 2001), an allele with 51 repeats apparently was not associated with the disease. On the other hand, alleles of intermediate size have been associated with the MJD phenotype (e.g., Padiath et al., 2005; Van Alfen et al., 2001); in the study of Van Alfen et al. (2001), four symptomatic family members presented CAG tract sizes between 53 and 54 repeats. Although rare, the cases of intermediate alleles imply that the distinction, initially very clear, between normal and mutated alleles has become much more difficult to establish. Alleles with around 50 repeats seem, in certain cases, to act as fully penetrant, a scenario that remains, nevertheless, rare (Paulson, 2007). In Portugal, and despite the high prevalence of the disease, a study of nearly 2000 chromosomes from the general population, representing all Portuguese districts, failed to detect the presence of intermediate alleles (Lima et al., 2005). Normal and pathogenic repeat size ranges are not definitive, and gathering of up-dated information should be a concern of laboratories offering molecular testing.
The ATXN3 gene is ubiquitously expressed in neuronal and non-neuronal tissues; it encodes for ataxin-3, a protein with an approximate molecular weight of 42kD, in its native form. In the neurons, ataxin-3 is found essentially in the cytoplasm (Paulson et al., 1997). Five products of the ATXN3 gene, referring to transcripts of different sizes, were described by Ichikawa and colleagues (Ichikawa et al., 2001). More recently, the sequence of 56 distinct transcripts, generated by alternative splicing, was described, and the high transcriptional variability of MJD’s causative gene has been demonstrated (Bettencourt et al., 2010a).

Ataxin-3 belongs to the family of cysteine-proteases; structurally it is composed by 339 aa, to which a variable number of glutamines is added (Poly-Q tract) (Kawaguchi et al., 1994). This protein is composed by a Josephine domain (JD), located in its N-terminal portion, containing, at its C-terminal, two or three ubiquitin-interacting motifs (UIMs) and the Poly-Q tract. It has been proposed that the native form of the protein acts as a deubiquitinating enzyme in the ubiquitin-proteosome pathway (revised in Bettencourt & Lima, 2011). Therefore, evidence concerning the proprieties of ataxin-3 suggests that this protein participates in cellular pathways related to quality control of proteins (see, amongst others, Schmitt et al., 2007). The involvement of the normal form of the protein in the regulation of transcription has also been suggested (Chou et al., 2008).

In the MJD locus, the presence of an expanded allele leads to a protein enriched in glutamines. The Poly-Q tract expansion should lead to a gain of neurotoxic function of the corresponding protein and, as a consequence, to neuronal death, by a process which remains, nevertheless, incompletely understood. Models of pathogenesis include the formation of toxic oligomers of ataxin-3, as well as aberrant protein-protein interactions, which disrupt normal cellular functions; revisions on the aspects of MJD’s pathogenesis can be found, amongst others, in Paulson (2007) and Katsuno et al. (2008).

Notwithstanding the fact that MJD constitutes a relatively well defined clinical entity, its clinical diagnosis can, in many situations, be difficult to establish. Such is the case, when the disease is at its initial stage and minor, but more specific signs are absent. Moreover, in apparently isolated cases, or in cases that appear associated to a less common geographic distribution, a clinical diagnosis can also be hard to establish with certainty (Lopes-Cendes et al., 1996). Therefore, in the differential diagnosis of MJD, molecular testing, available after the identification of the causative mutation, has become of major importance (Maciel et al., 2001). Furthermore, predictive testing (PT), as well as prenatal diagnosis (PND) became available (Sequeiros et al., 1998). More recently, pre-implantation genetic diagnosis (PGD) is also feasible (Drüsedau et al., 2004). The program of PT and GC, available for MJD in Portugal since the end of 1995, was based on the previous experience with Huntington’s disease (HD) (Sequeiros, 1996). This program is ongoing in the Azores since 1996, and its impact in patients and families is periodically revised (Gonzalez et al., 2004; Lima et al., 2001).

Presently there is no effective pharmacological approach for SCAs. Specific drugs have been prescribed to minimize some of the symptoms, such as ataxia or dystonia (reviewed in Bettencourt & Lima, 2011). Cell and animal models have also been fundamental in the understanding of the pathogenesis and gene therapy search (e.g., Gould, 2005); the use of interference RNA and the administration of antisense oligonucleotides showed promising results according to Alves et al. (2008; 2010) and Hu et al. (2009), respectively.
3. Complicating factors in MJD

Several aspects of non-compliance with the Mendelian expectations can be readily recognized for MJD. Variation in the age at onset, variability in clinical presentation, presence of anticipation as well as repeat instability (somatic and germ line), have been described as the main complicating factors in MJD (Tsuji, 1997). Other factors, which will be referred, are also noteworthy.

3.1 Clinical variability

As previously referred, MJD is characterized by a complex phenotype, which is highly variable amongst patients. The recognition of its high degree of clinical variability has led to the proposal of Coutinho & Andrade (1978), in which patients could be classified into three clinical types. According to Coutinho (1992), type 1 has an early onset (average of 24 years) and a more rapid progression of symptoms, being characterized by pyramidal signs and dystonia. Type 2 is the most frequent, and occurs around 40 years of age, being dominated by ataxia and ophthaloplegia, with or without pyramidal signs. Type 3 has a latter onset (average of 47 years) and progresses slowly, with amyothrophies. The three clinical types can, occasionally, be present in the same family. A fourth type, a rare presentation that associates parkinsonism to cerebellar signs (Suite et al., 1986), and a fifth type, associated with spastic paraplegia (Sakai & Kawakami, 1996), have also been proposed.

Notwithstanding the fact that some clinical features, if present, can help in the differential diagnosis of MJD (e.g., ophthalmoplegia, bulging eyes or face and tongue fasciculations), phenotypic overlapping with other SCAs has implications for GC and PT. Therefore, at-risk individuals entering the PT program must have an affected close relative with a definitive molecular diagnosis - “mutation-positive” familial control (Sequeiros et al., 2010).

Variation in age at onset, evidenced by its reported extremes (4 and 70 years) described by Carvalho et al. (2008) and Coutinho (1992), constitutes a particular aspect of the clinical variability of this disorder. A significant, but partial, negative correlation between the size of the expanded allele (and thus, the extension of the Poly-Q tract) and the age of appearance of symptoms explains between nearly 50 to 75% of the variation in the age at onset, depending on the analyzed series of patients (e.g., Maciel et al., 1995; Maruyama et al., 1995). The size of the mutated allele also correlates with the frequency of particular clinical signs, such as pyramidal signs, which are more frequent in patients with larger repeats (Takiyama et al., 1995). In the Azorean series of patients, for example, the number of CAG repeats, determined in genomic DNA and in mRNA, explains 68% and 67% of variation in the age at onset, respectively (Bettencourt et al., 2010b). The incompleteness of the correlation observed between the size of the CAG tract and the age at onset implies that such information cannot be used for counseling purposes; whether allele sizes should be disclosed in the molecular test report is still being debated (Sequeiros et al., 2010).

The reported incompleteness of the genotype-phenotype correlation, observed in MJD as well as in other SCAs, confirms the involvement of non-CAG factors in the explanation of the phenotype, that can either be genetic or environmental (van de Warrenburg et al., 2005). For MJD, the hypothesis that an important fraction of the residual of the disease onset (after accounting for the CAG repeat size) is of genetic nature has been reinforced (DeStefano et al., 1996; van de Warrenburg et al., 2005). The few studies that attempted to
identify MJD modifiers have used candidate gene approaches; Jardim et al. (2003), in a study that considers the effect of the CAG tract at several expansion loci (namely SCA2, SCA6 and DRPLA), only found a correlation between the severity of fasciculations and the size of the CAG tract at the SCA2 locus. Recently, Bettencourt and colleagues (2011) found a significant association between the presence of the APOE ε2 allele and an earlier onset in MJD.

3.2 Gene dosage effect

The reduced number of homozygous patients described for MJD makes any generalization, concerning the impact of gene dosage on clinical presentation, hard to perform. The few cases described in the literature, nevertheless, reinforce the fact that the phenotype is more severe and the onset is earlier in homozygous carriers of the mutated allele (e.g., Carvalho et al., 2008; Lerer et al., 1996). This indicates that gene dosage is an important determinant of the onset. The increased severity observed in homozygous is common in dominant diseases, which do not follow the Mendelian expectation of phenotypic overlapping between homo- and heterozygous. In specific populations, known to have a particularly high prevalence of MJD, such as the Azorean island of Flores, the possibility of mating between carriers of the MJD mutation must be taken into consideration when planning GC sessions.

3.3 Incomplete and age-dependent pattern of penetrance

In MJD pedigrees, skipped generations are rarely observed. Coutinho (1992) refers that the majority of such cases can be explained by a premature death of the obligate carrier, in relation to the average onset of the disease. Other factors, such as migration, can further prevent the confirmation of the disease status in the obligate carrier. With an estimated value of 98%, the MJD gene penetrance presents an age-dependent pattern, which implies that the a posteriori, or residual risk, differs considerably depending on the age of the individual. Residual risk tables constitute, in this context, important tools for GC (Bettencourt et al., 2008a), since they allow the geneticist to estimate the probability that an asymptomatic at-risk individual has to develop the disease at a certain age.

3.4 Intergenerational instability of the CAG tract

Within the Poly-Q disorders, anticipation is more marked for DRPLA, SCA2, SCA7 and HD (Paulson, 2007; Takiyama et al., 1999). In MJD, however, the aggravation of symptoms, and the decrease in the age at onset in successive generations, is also observed. On the basis of anticipation in MJD is the instability of the ATXN3 gene region containing the expanded CAG tract, which, during cellular division, can lead to alterations in the size of the repeat tract, resulting in expansions or, more rarely, in contractions. Although the decrease in the age at onset is highly related with the increase in the size of the CAG tract, families with a higher degree of anticipation than it would be expected for each repeat unit increase were identified (Takiyama et al., 1998).

Germline mosaicism has been consistently described for MJD (e.g., Cancel et al., 1995). The tendency for the increase of the repeat size is more marked in male than in female meiosis; processes specific to sperm or oocyte development should be involved in such differences.
Maciel et al. (1995) reported that the size of the repeat tract varies in 55% of transmissions; from these, 78% correspond to expansions and 22% to contractions.

Several factors have been implicated as modulating the intergenerational instability in MJD, such as the sex of the transmitting parent as well as the intragenic environment (see, amongst others, Igarashi et al., 1996; Maciel et al., 1999). The results obtained by Takiyama et al. (1997) suggests the presence of an inter-allelic association involved in the instability of the CAG tract, trough yet poorly known mechanisms. Other inter-allelic and \textit{cis} factors have also been studied by Martins et al. (2008); these authors have concluded that distinct origins of the mutation (established on the basis of haplotypes constructed using intragenic SNPs) present different behaviors on what concerns repeat instability. Evidences gathered so far support the presence of a mechanism associated to DNA repair, rather than associated with replication, on the basis of meiotic instability observed in this locus (Martins et al., 2008).

Little is known about the mutational process that leads to the emergence of repeats within the pathological size. It has been postulated that a mutational bias, in favor of expansions, exists in trinucleotide repeat loci, suggesting that the upper end of the normal allele distribution would provide a “reservoir” from which expanded alleles would be generated (Rubinsztein et al., 1994). The hypothesis that normal alleles of larger size could constitute such a reservoir was not supported by a study of nearly 2000 chromosomes of a representative sample of the general Portuguese population (Lima et al., 2005). On the contrary, the report from Lima and colleagues (2005), shows that the allelic distribution was skewed towards smaller size alleles. In a subsequent study, Martins et al. (2006) have integrated not only the analysis of the CAG repeat, but also information on haplotypes built using intragenic and flanking markers; the conclusions indicate that a multistep mechanism is on the basis of the evolution of the CAG repeats in MJD, originated either by gene conversion or DNA slippage.

### 3.5 Somatic mosaicism

The instability of the expanded polyglutamine-coding (CAG)\textsubscript{n} tracts during mitotic cell division can lead to changes in repeat length, either contractions or more frequently expansions, resulting in certain populations of cells carrying different repeat sizes. For MJD, the somatic mosaicism has been described by several authors (see, amongst others, Cancel et al., 1995; Lopes-Cendes et al., 1996; Maciel et al., 1997).

In the central nervous system (CNS), the pattern of mosaicism for mutated alleles is similar in the several structures, with the exception of the cerebellar cortex, which presents slightly reduced tracts (Cancel et al., 1998; Hashida et al., 1997). In non-neuronal tissues, the instability is lower in muscle (Tanaka et al., 1999). The studies conducted have consistently failed to demonstrate a correlation between the degree of mosaicism and the selective neuronal vulnerability (Cancel et al., 1998; Ito et al., 1998). The pattern of mosaicism in genomic DNA is maintained in mRNA, and the variation in the size of the CAG repeat in mRNA is also not relatable with the severity of the pathological involvement of the several tissues (Ito et al., 1998).

Somatic mosaicism contributes to the limitations in the precision of sizing the MJD repeat motif, since it originates differences in (CAG)\textsubscript{n} length among subpopulations of lymphocytes as well as between lymphocytes (where length is usually measured) and CNS
cells (revised in Lima et al., 2006). Thus an error of ±1 CAG repeat is considered as acceptable (Maciel et al., 2001; Sequeiros et al., 2010).

3.6 Segregation distortion

Alterations to the Mendelian proportions in the segregation of the ATXN3 alleles were firstly highlighted by Ikeuchi and colleagues, in 1996. These authors suggested the occurrence of “meiotic drive” to justify the observation of an excess of affected descendents of MJD patients, a fact hardly explainable by the Mendelian principle of random segregation of alleles (Ikeuchi et al., 1996). Their results pointed to the existence of segregation distortion, in the male meiosis. This issue, however, is far from consensual. A single-sperm typing performed in Japanese patients, by Takiyama et al. (1997) indicated a preferential transmission of mutated alleles. On the other hand, a similar study by Grewal et al. (1999), which used sperm samples from patients of French origin, failed to report the presence of segregation distortion. Using a methodology based on the analysis of pedigrees of patients belonging to Azorean MJD families, Bettencourt et al. (2008a) also investigated the presence of segregation distortion in the transmission of mutated ATXN3 alleles. According to that study, segregation is done in agreement with the expected Mendelian proportions.

The behavior postulated for the wild-type and the mutated ATXN3 alleles is not necessarily comparable. Nevertheless, studies with normal individuals have also been conducted, aiming to contribute to the understanding of this issue (Bettencourt et al., 2008b; MacMillan et al., 1999; Rubinsztein & Leggo, 1997). Rubinsztein and Leggo (1997), in a segregation study of MJD alleles in normal heterozygous individuals, reported the presence of segregation distortion only when the transmitting parent was a female, with the smaller allele being preferably transmitted. Results from another study, by Bettencourt et al. (2008b), followed the same trend, with a preference for the transmission of the smaller allele. These last authors also reinforced the importance of the genotypic constitution of the sample being analyzed, which may act as a confounding factor in the detection of segregation distortion.

4. Conclusion

Poly-Q diseases occur as a result of a mutation at the respective causative genes, representing, from that perspective, simple, monogenic diseases. Several aspects of complexity are, nevertheless, present in this group of disorders. Many of the complicating factors present are displayed by MJD and were approached here; the majority of them have implications for patients management and, therefore, its understanding is of major importance.

5. Acknowledgements

C.B. is a postdoctoral fellow of “Fundação para a Ciência e a Tecnologia” – FCT [SFRH/BPD/63121/2009].

6. References


Igarashi, S.; Takiyama, Y.; Cancel, G.; et al. (1996). Intergenerational instability of the CAG repeat of the gene for Machado-Joseph disease (MJD1) is affected by the genotype of the normal chromosome: implications for the molecular mechanisms of the instability of the CAG repeat. Human Molecular Genetics, 5(7): 923-32.


(MJD), dentatorubral pallidoluysian atrophy (DRPLA), and spinal and bulbar muscular atrophy (SBMA). *Neurochemical Research*, 23(1): 25-32.


The purpose of this book has been to depict as many biochemical, genetic and molecular advances as possible, in the vast field of the spinocerebellar ataxias.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: