Improvement in the Molecular Diagnosis of Machado-Joseph Disease

Patricia Maciel, PhD; Maria do Carmo Costa, BSc; Anabela Ferro, BSc; Marylene Rousseau, BSc; Cláudia Sofia Santos, BSc; Claudia Gaspar, PhD; José Barros, MD; Guy A. Rouleau, MD, PhD; Paula Coutinho, MD, PhD; Jorge Sequeiros, MD, PhD

Background: Direct detection of the gene mutation allows for the confirmation of the clinical diagnosis of Machado-Joseph disease (MJD), the most frequent cause of autosomal dominant spinocerebellar ataxia worldwide.

Objective: To address the main difficulties in our national MJD predictive testing program. The first was the emergence of intermediate alleles, for which it is not yet possible to determine whether they will cause disease. The second was the issue of homoallelism, ie, homozygosity for 2 normal alleles with exactly the same (CAG)n length, which occurs in about 10% of all test results.

Methods: A large pedigree with 1 affected patient carrying a 71 and a 51 CAG repeat and 2 asymptomatic relatives carrying the 51 CAG repeat and normal-size alleles underwent clinical and molecular studies. Intragenic haplotypes for these alleles were determined. A representative sample of the healthy population in the region was obtained to assess the distribution of the normal (CAG)n length. We established the genotype for 4 intragenic polymorphisms in the gene for MJD (MJD1) in 21 homoallelic individuals, to distinguish their 2 normal chromosomes. In addition, we developed a new Southern blot method to completely exclude cases of nonamplification of expanded alleles in the homoallelic individuals.

Results: The study of the family in which the 51 CAG repeat was found suggests that the allele is apparently not associated with disease. These intermediate alleles were not present in a large sample of the healthy population from the same region. Intragenic polymorphisms allowed distinction of the 2 different normal alleles in all cases of homoallelism. The absence of an expanded allele was also confirmed by Southern blot.

Conclusions: We propose an improved protocol for molecular testing for MJD. These strategies, developed to overcome the practical difficulties mostly in the presymptomatic and prenatal diagnosis of MJD, should prove useful for other polyglutamine-related disorders.

Arch Neurol. 2001;58:1821-1827

From the UnIGENE, Instituto de Biologia Molecular e Celular (Drs Maciel, Gaspar, and Sequeiros and Ms Costa, Ferro, and Santos), and the Departamento Estudos de Populações, Instituto de Ciências Biomédicas de Abel Salazar (Dr Sequeiros), Universidade do Porto, and Serviço de Neurologia, Hospital de St António (Dr Barros), Porto, Portugal; Instituto Superior de Ciências da Saúde-Norte, Paredes, Portugal (Dr Maciel); Centre for Research in Neuroscience, McGill University and the Montreal General Hospital Research Institute, Montreal, Quebec (Ms Rousseau and Drs Gaspar and Rouleau); and Serviço de Neurologia, Hospital de St Sebastião, Santa Maria da Feira, Portugal (Dr Coutinho).
SUBJECTS, MATERIALS, AND METHODS

SUBJECTS

A family from the Tagus River Valley (family MJD75) was identified during an ongoing survey of inherited ataxias in Portugal. Results of molecular testing showed the proband and his 45-year-old affected son to carry an expanded allele at the MJD1 gene; however, the son also carried another allele of intermediate size, never previously described in the affected or the healthy population (Figure 1). To clarify the possible association of this allele with the presence or absence of disease, an important issue for genetic counseling of the family (including the 2 asymptomatic siblings), we examined additional members of the family, including the proband’s spouse and 12 of her relatives (pedigree MJD75b) (Figure 2).

To determine the (CAG)_n size in normal chromosomes and look for intermediate alleles, we tested anonymous Guthrie cards (from the national program for phenylketonuria screening); 20 samples were collected at random from 16 villages in the region, totaling 320 chromosomes tested. During a 2-year period, the molecular diagnosis of MJD was obtained at our laboratory for 149 patients referred by a neurologist or another physician, and for 55 at-risk family members after indication by a medical geneticist. Of these 204 subjects, 21 were shown to be homoallelic and underwent studies for intragenic polymorphisms and Southern blot analysis as described herein.

MOLECULAR CHARACTERIZATION OF THE CAG REPEITIVE TRACT AND INTRAGENIC POLYMORPHISMS IN THE MJD1 GENE

Blood samples were obtained after informed consent from all individuals in the familial study, and genomic DNA was extracted from lymphocytes, as described elsewhere. The DNA was extracted from the filter paper blots using the resin Chelex (Bio-Rad, Hercules, Calif) in a final concentration of 1.8 M NaCl. Amplification of the CAG repeat–containing fragment in the MJD1 gene was performed by means of polymerase chain reaction (PCR) analysis, using previously described conditions, and the size of the PCR products was determined by denaturing 6% polyacrylamide gel electrophoresis in parallel with an M13 sequence ladder, visualized by means of autoradiography. In every reaction, we used as a positive control genomic DNA from a patient with an expanded allele containing 86 CAG repeats, the largest we have ever amplified by means of PCR.

The intragenic polymorphisms C^669GG/G^669GG and TA^1118/TA^1118 (single nucleotide substitutions at positions indicated in each codon) were detected by means of allele-specific PCR. The polymorphism A^1118TG/G^1118TG was detected by means of single-strand conformational polymorphism analysis. The intragenic polymorphism C^1178/G^1178 (single nucleotide substitution at position 1178 in the 3’ noncoding region of the gene) was detected by means of allele-specific PCR, using primers MJD8 (5’-GATTACAAATTTACTTAAGAG-3’) or MJD9 (5’-GATTAACAAATTTACTTAAGAT-3’) combined with MJD6 (5’-GACAGATTTACATCCATGTG-3’). The allele-specific PCR was performed in the same conditions used for amplification of the CAG repeat, except for the annealing step, which was performed at 52°C for 30 seconds. The PCR products were analyzed in a denaturing 6% polyacrylamide gel and visualized by means of autoradiography. The DNA sequencing was performed using the primer MJD9a (5’-CCAGTTGTACCTTTGAGT-3’) to determine the genotype for the allele-specific PCR in a few individuals used as controls in the allele-specific PCR. Sequencing with primer MJD6 was used to determine genotype for polymorphisms TA^1118/TA^1118 and C^1178/G^1178 for the same purpose. Reactions were performed using 5 µL of DNA and a cycle-sequencing kit (Thermo Sequenase; USB, Cleveland, Ohio) following the manufacturer’s instructions.

Another approach was also used to confirm homoallelism. A 186–base pair (bp) fragment was generated by means of PCR, with primers MJD1 (5’-TGCCCAT-GATAGTTATTTGTTGA-3’) and MJD2 (5’-GGAAAATACATGTGTTCAGAATC-3’). This fragment was purified from agarose gel using an extraction kit (QIAEX II; QIAGEN, Valencia, Calif) and labeled with phosphorus 32 (32P) using a labeling kit (Prime-It II Random Primer; Stratagene, Cedar Creek, Tex), to be used as a probe in a Southern blot. The genomic DNA (40 µg) was digested with the restriction enzyme AluI, and the fragments were separated in a 1% agarose gel (Nusieve 3:1 agarose; FMC BioProducts, Rockland, Me) and blotted in a nylon membrane (Hybond H-N+; Amersham Pharmacia Biotech, Buckinghamshire, England) that was hybridized with the probe at 65°C. This results in bands of approximately 600 bp for normal alleles and bands larger than 733 bp for expanded alleles. These were visualized by means of autoradiography.

Figure 3 represents the location of all primers and restriction sites used in the molecular diagnostic procedure.

Development of new therapies. Therefore, predictive testing and genetic counseling are still the only means to diminish the impact of the disease in affected families.

Direct detection of the MJD mutation allows the confirmation of the clinical diagnosis, which can be useful given the potential clinical overlap among the different forms of spinocerebellar ataxia. Another immediate application of the molecular test was the possibility of presymptomatic diagnosis, in the context of genetic counseling programs. This is of utmost importance, particularly in areas of high prevalence of disease, such as the Azorean islands (1:3700 in São Miguel; 1:120 in Flores), where the disease is considered a public health problem, and some areas of mainland Portugal (1:1000 in a small area of the Tagus River Valley). In this study, we address the main difficulties encountered in the context of the Portuguese Predictive Testing and Genetic Counseling Program for Machado-Joseph Disease. The first was the emergence of alleles of a size between the known normal and pathogenic ranges (intermediate alleles), for which it is not possible at this point to determine whether they will cause the disease. The second was the issue of
homoallelism, i.e., the homozygosity for 2 normal alleles with exactly the same (CAG)\textsubscript{n} length. The strategies developed to overcome these difficulties are presented herein and may prove useful to other polyglutamine-related disorders.

RESULTS

INTERMEDIATE ALLELES IN A FAMILY WITH MJD

A family with the MJD mutation was identified, where the index case was a 69-year-old man (IV:7 in Figure 2). Age at onset was 45 years with gait imbalance, followed some years later by diplopia, dysarthria, and dysphagia. He used a wheelchair since 54 years of age, due to severe cerebellar ataxia. He had a complete limitation of upward movements of his eyes and partial limitation of lateral movements, without lid retraction or nystagmus. A bilateral facial palsy with atrophy and fasciculations was evident, as well as a moderate, generalized muscle weakness and atrophy of the limbs with areflexia. Except for a brisk jaw jerk, no corticospinal signs were present, and no dystonia. The clinical picture corresponded to a classic subtype 3.

The proband's oldest son (V:4, born in 1952) was also affected, with difficulty walking since 32 years of age, followed 2 years later by dysarthria and diplopia. At 45 years of age, with 13 years of evolution of the disease, he still had an independent gait but was unable to work. He showed a moderate cerebellar ataxia, more marked in the gait and in the lower limbs. A coarse nystagmus was present. Vertical eye movements were limited, and some contraction fasciculations of the face were observed. Muscular strength was normal. All deep-tendon reflexes were exaggerated, but plantar reflexes were normal, and there was no spasticity. This clinical picture corresponded to a subtype 2. The other son (V:5) and the daughter (V:6) of the proband had normal results of neurologic examinations at 35 and 32 years of age, respectively; the proband’s wife (IV:8) had normal results of neurologic examination at 67 years of age.

The proband (IV:7) carried an expanded allele with 68 CAG repeats (Figures 1 and 2). His oldest son (V:4) carried an expanded allele of 71 repeats, and, in addition, an allele with 51 repeats (Figures 1 and 2). The unaffected 35-year-old son (V:5) carried the paternal normal allele with 21 repeats and the intermediate allele with 51 repeats, whereas the 32-year-old unaffected daughter (V:6) carried 2 normal alleles (21 and 23 CAGs). Their unaffected mother (IV:8) had a normal allele (23 repeats) and the intermediate allele of 51 repeats.

The length of the (CAG)\textsubscript{n} in 12 other members on the maternal side of the family is shown in Figure 2; all were thought, owing to history, not to have MJD. On examination, all but 4 were considered healthy. Two sisters had juvenile parkinsonism (IV:1 and IV:2), and 1 male member was bedridden due to spinal cord injury (III:12); all 3 had 2 normal alleles. One female member (III:11), however, who was bedridden after a stroke,
intermediate alleles were found. (n=320) obtained from the region of origin of family MJD75. No with anterior horn signs could have been developing the peripheral neuropathy), cerebellar ataxia combined (probably associated with the occurrence of stroke and deterioration, advanced age, and her history of diabetes previous stroke). In conclusion, along with mental weakness only on the right side (probably due to her generalized fasciculations, contrasting with muscular was a moderate atrophy of the hands and legs, with Table 1: Position of primers and restriction sites used in the molecular diagnostic procedures. MJD indicates Machado-Joseph disease (MJD); ASP, allele-specific polymerase chain reaction analysis. had an expanded allele with 63 CAG repeats. She was aged 85 years at the time of examination and had been in good health until 78 years of age, when she had a first time at the time of observation. She presented with speaking worsened. Dysphagia was noticed for the other symptoms will begin. Below this limit, there are re- individual, although it is not possible to establish at what son lives long enough, the disease will develop in this If someone carries alleles with more than 60 CAG repeats, one can predict with some certainty that if the per- repeat known to date in patients. The intragenic markers A\textsuperscript{669}TG/G\textsuperscript{669}TG, C\textsuperscript{987}GG/ G\textsuperscript{987}GG, TAA\textsuperscript{1118}/TAC\textsuperscript{1118}, and C\textsuperscript{178}/A\textsuperscript{178} were used to de- fine the haplotypes associated with the intermediate and expanded alleles in the family under study. The 51 CAG repeat coming from the unaffected mother (IV:8) was associated with the same haplotype as the expanded allele coming from the affected father (IV:7), GGCA. The expanded allele present in the 85-year-old relative (III:11) was also associated with haplotype GGCA.

### LENGTH OF THE (CAG\textsubscript{n}) TRACT IN THE MJD1 GENE IN A CONTROL POPULATION

The distribution of (CAG\textsubscript{n}) tract length in 302 chromosomes of control individuals born in the same district where the individual with the 51 CAG repeat originated is shown in **Figure 4**. The largest allele found was 37 CAG repeats. No intermediate alleles were found in this sample.

**APPARENT HOMOALLELISM OF THE MJD1 GENE IN MEMBERS OF MJD FAMILIES**

We have performed a total of 204 diagnostic and predictive tests. According to our results, 107 family members carried an expansion, 79 had 2 normal alleles of different size, and 21 (10.3%) were apparently homoallelic, ie, had 2 normal alleles with the same (CAG\textsubscript{n}) size. When these 21 individuals underwent typing for 3 intragenic polymorphisms of the MJD1 gene to try to distinguish the 2 normal chromosomes (**Figure 5**), in 2 cases (9.5%) the distinction was possible using polymorphism C\textsuperscript{987}GG/G\textsuperscript{987}GG; in 4 (19.0%), using A\textsuperscript{178}/C\textsuperscript{178}; and in 18 (85.7%), using TAA\textsuperscript{1118}/TAC\textsuperscript{1118}. In combination, these 3 intragenic polymorphisms allowed for the distinction of both normal alleles in all 21 cases of homoallelism.

Results of Southern blot analysis also confirmed that none of these individuals carried an expanded allele (an example is shown in **Figure 6**). The molecular diagnosis of MJD is currently based on the determination of the (CAG\textsubscript{n}) length in the MJD1 gene. If someone carries alleles with more than 60 CAG repeats, one can predict with some certainty that if the person lives long enough, the disease will develop in this individual, although it is not possible to establish at what age symptoms will begin. Below this limit, there are reports of 3 patients with alleles containing 56, 55, and 54 CAG repeats,\textsuperscript{32-34} making the last one the shortest CAG repeat known to date in patients. Precision of size determination has its limitations; although we can apply more accurate methods using densitometry to select the peak with the largest area.
and using a cloned, sequenced allele as control, one still has to consider that somatic mosaicism exists, with differences in (CAG)_n length between lymphocytes (where length is usually measured) and central nervous system cells, as well as among subpopulations of lymphocytes. An error of ±1 CAG repeat is considered acceptable. In addition, (CAG)_n size on its own is not useful as a more precise predictor of outcome. Although the variability in clinical presentation is partly explained by the length of the expanded allele, this inverse correlation is incomplete and not applicable for prediction of age at onset or clinical presentation.3,25,26 For that reason, the sizes of the alleles are not usually communicated to the consultant. We believe it is important, however, to determine and to keep an accurate record of the sizes of normal and expanded alleles, since the molecular diagnosis of MJD is still in a research phase. In the future, knowledge of other factors affecting disease onset and progression, eg, environmental agents or modifier genes, may contribute to a more accurate prediction.

To illustrate this point, we describe herein an MJD1 allele with a repeat length not previously encountered. Although studies in several different populations had shown a wide gap between the normal (12-44 CAG repeats) and the disease (61-87 CAG repeats) range, these limits are expected to change with the increasing size of our sample. The present identification of a formerly undescribed “intermediate” allele containing 51 CAG repeats was the source of a potential problem for genetic counseling, since we are not able to predict whether or not the disease will develop in an individual carrying this allele.

The study of the nuclear family in which this allele with 51 CAG repeats was found suggested, however, that this allele might not be pathogenic. First, the 67-year-old transmitting mother was still unaffected. Second, the individual carrying this allele in addition to a full expansion or juvenile onset, as could possibly be expected in a homozygous patient.35-37 Third, this allele was stable on at least 2 transmissions. The reduced number of cases with the allele does not, however, allow us to establish this conclusion with certainty. When we studied other living members of this family, no other individuals carrying the allele with 51 CAG repeats were found. We did, however, find an expanded allele with 63 CAG repeats in an 85-year-old woman (Figure 2) in whom the MJD phenotype had not been detected previously, possibly only because of masking by the sequelae of a previous stroke and a diabetic neuropathy.

The alleles with 51 and 63 CAG repeats share the same intragenic haplotype (GGCA) and may have a common origin. We cannot determine, however, whether the 51-CAG allele resulted from the contraction of a previously expanded allele, or whether the ancestral allele was of intermediate size and expanded only in 1 branch of the family. However, the GGC haplotype (not including the polymorphism C/A1178) is known to be the most common in the healthy population in Portugal and corresponds to a small subgroup of the Portuguese families with MJD, namely those originating from the island of São Miguel and those in the Tagus River Valley.36,39

In a previous screening of a large control population from all districts of Portugal (2000 chromosomes, 100 per district) (P.M., M.do C.C., A.F., C.S.S., Laura Guimaraes, Alda Sousa, PhD, and J.S., unpublished data, August 2001), we have found no alleles larger than 36 CAG repeats, suggesting that intermediate alleles must be quite uncommon. After an additional control sample (320 chromosomes) from the very district in the Tagus River Valley where these families originated (where prevalence of MJD is 80 times that of the rest of the country) underwent screening, no intermediate alleles were found.
In Huntington disease, nonpenetration with intergenerational instability for alleles with 29 to 35 CAG repeats and low penetrance for alleles with 36 to 39 CAG repeats were described. It is possible that the same will occur for intermediate alleles in MJD, but further studies are needed to clarify this question.

It was also suggested that smaller expansions could be associated with unusual clinical presentations of MJD, such as autonomic dysfunction (present in a patient with an allele with 56 CAG repeats, combined with cerebellar ataxia), progressive proximal weakness and sensory disturbances (in a patient with an allele with 54 CAG repeats), or parkinsonian features (type 4 cases confirmed by results of molecular testing had average size repeats), or parkinsonian features (type 4 cases confirmed by results of molecular testing had average size repeats), or parkinsonian features (type 4 cases confirmed by results of molecular testing had average size repeats), or parkinsonian features (type 4 cases confirmed by results of molecular testing had average size repeats), or parkinsonian features (type 4 cases confirmed by results of molecular testing had average size repeats). Other presentations that have been suggested for MJD include pure cerebellar ataxia and spastic paraplegia phenotypes. In this context, the 2 sisters from pedigree MJD75b with juvenile parkinsonism carried 2 normal alleles; in fact, the consanguinity of their parents may suggest that another (possibly recessive) mutation might be the cause of their disease.

Another important source of ambiguity in the molecular testing of MJD was the relatively high frequency of apparent homoallelism (ie, homozygosity for exactly the same size CAG repeat). Given the highly polymorphic nature of repetitive tracts, it is not surprising that most of the individuals with 2 normal alleles at the MJD1 locus (homozygous in the classic mendelian sense) have 2 alleles with different size CAG repeats (heteroallelism). The exclusion of the disease is very clear in these cases, and the interpretation of the molecular test results does not raise major difficulties. Approximately 10% of cases of homoallelicism for the normal allele were found, however, in our diagnostic and predictive tests. Although PCR was systematically repeated in every such case, and large expansions were used as positive controls, it is impossible to completely exclude nonamplification of an expanded allele, because of either extremely large size or the presence of polymorphisms in the primer-annealing regions, leading to false-negative results such as have been described in Huntington disease. One case of homoallelicism was found in an instance of prenatal diagnosis, creating a particularly delicate situation, given time constraints and the dependence of the parents’ decision to proceed with or to terminate the pregnancy on the status of the fetus regarding the MJD mutation. In these cases, to confirm that we are in the presence of cases of true homoallelicism, our proposed approach (Figure 7) is (1) to confirm, whenever possible, that it was compatible with the parents’ genotypes, ie, that it was possible for the individual to carry 2 alleles of the same size; (2) to check the population frequency of that allele (this should be mentioned in the report); (3) to determine the intragenic haplotypes in both chromosomes (the 2 normal chromosomes carrying repetitive CAG tracts of the same length might be associated with different haplotypes); and (4) to now use, in addition, a molecular method not dependent on PCR, which has allowed us to exclude retrospectively the presence of an expansion in the MJD1 gene in all cases of homoallelicism in study, thus confirming the results obtained with the intragenic polymorphisms.

Figure 7. Flowchart for the molecular diagnosis of Machado-Joseph disease. Polymorphisms are described in the “Subjects, Materials, and Methods” section.

CONCLUSIONS

We have tried to address the major difficulties found while testing for a genetic disease largely still under study, such as MJD. We suggest specific solutions for these problems, and reinforce the need for permanent interaction between the diagnostic services and the research process.