European Journal of Human Genetics (1999) 7, 147–156

http://www.stockton-press.co.uk/ejhg

ARTICLE

Study of three intragenic polymorphisms in the Machado-Joseph disease gene (*MJD1*) in relation to genetic instability of the $(CAG)_n$ tract

Patrícia Maciel[1,5], Claudia Gaspar[1,5], Laura Guimarães[5], Jun Goto[3], Iscia Lopes-Cendes[1], Sean Hayes[1], Karin Arvidsson[1], Aureliano Dias[1], Jorge Sequeiros[2,5], Alda Sousa[4,5] and Guy A Rouleau[1]

[1]*Centre for Research in Neuroscience, McGill University and the Montreal General Hospital Research Institute, Montreal, Québec, Canada*
[2]*Lab. Genética Médica, ICBAS*
[3]*Department of Neurology, Tokyo University Medical School, Tokyo, Japan*
[4]*Genética Aplicada, ICBAS, Universidade do Porto, Portugal*
[5]*UnIGENe, IBMC Universidade do Porto, Portugal*

Intergenerational instability is one of the most important features of the disease-associated trinucleotide expansions, leading to variation in size of the repeat among and within families, which manifests as variable age at onset and severity, and is probably the basis for the occurrence of anticipation. Several factors are known to affect the degree of instability, namely the type of repeated sequence, its initial size, the presence or absence of interruptions in the repetitive tract and the gender of the transmitting parent. A recent study demonstrated the effect of an intragenic polymorphism ($C^{987}GG/G^{987}GG$) in the Machado-Joseph disease causative gene, immediately downstream of the CAG repeat, on the intergenerational instability of the expanded repeat. Surprisingly, there was an effect not only of the specific allele in *cis* to the disease chromosome, but also of the allele on the normal chromosome, suggesting the existence of an interaction between the normal and expanded alleles that affects the fidelity of replication of the $(CAG)_n$ tract. This effect could be a direct effect of the polymorphism studied or, alternatively, this polymorphism could be in disequilibrium with some other flanking sequence which affects the instability of the repetitive $(CAG)_n$ tract. In order to confirm the previous results in a different population and to distinguish between a direct and indirect effect of the CGG/GGG polymorphism, we typed 70 parent–progeny pairs for which the variation in the $(CAG)_n$ length in the *MJD1* gene was known, for three intragenic polymorphisms: $C^{987}GG/G^{987}GG$ and two additional, newly described ones, $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ and $\underline{A}^{669}TG/\underline{G}^{669}TG$. We also typed a control population of 125 individuals for the $\underline{A}^{669}TG/\underline{G}^{669}TG$, $\underline{C}^{987}GG/\underline{G}^{987}GG$ and $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ polymorphisms, in an attempt to identify any association between haplotype and $(CAG)_n$ length in normal chromosomes,

Correspondence: Patrícia Maciel, UnIGENe, IBMC - Universidade do Porto, R. Campo Alegre, 823, 4050 Porto, Portugal. Tel: 351 2 6074941; Fax: 351 2 6099157 or 5506519; E-mail: patricia.maciel@writeme.com
Received 24 February 1998; revised 8 July 1998; accepted 18 August 1998

suggestive of an instability-predisposing effect of the repeat-flanking sequences, which could have led to the origin of the MJD mutation in the human population. We confirmed the effect of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism on intergenerational instability when present in *trans*. Our results suggest that this effect is restricted to a small region of the gene, immediately downstream of the CAG repeat, which includes this particular nucleotide substitution and the stop codon of the *MJD1* cDNA, and is not a more widespread chromosomal effect. The lack of a significant association of any specific intragenic haplotype with larger CAG repeats in normal chromosomes, together with the absence of an effect of the intragenic haplotype in *cis* on the intergenerational instability of the expanded $(CAG)_n$ in MJD families does not indicate the existence of an instability-predisposing haplotype.

Keywords: ataxia; neurodegenerative disorder; polyglutamine; trinucleotide repeats; dynamic mutations; haplotypes; origin of mutation

Introduction

Machado-Joseph disease (MJD) is an autosomal dominant multisystem degeneration of adult onset[1] with variable clinical presentation. Its manifestations include cerebellar ataxia and progressive external ophthalmoplegia, associated in a variable degree with pyramidal signs, dystonia, rigidity, amyotrophies, and peripheral neuropathy.[1–4] The mutation associated with this disorder is the expansion of a $(CAG)_n$ tract in the coding region of the *MJD1* gene,[5] which in normal individuals contains from 12 to 41 repeat units and in expanded chromosomes from 61 to 86 repeat units.[5–7]

Although the mechanisms of trinucleotide repeat instability are still not understood, a variety of factors have been implicated as modulating intergenerational repeat instability in the diseases associated with dynamic mutations, namely the sex of transmitter,[6–18] the expanded repeat size in the transmitting parent,[8–12,18–20] the presence or absence of an interruption of the repeat,[16,21–25] and even the sex of the descendent.[26]

We and others have previously shown that in MJD families the sex of the transmitting parent has a significant effect on intergenerational instability,[6,7,17] with male meiosis associated with larger variations, both contractions and expansions, of repeat size. However, there was no association between size of CAG repeat length in the transmitting parent and change in CAG repeat length during transmission,[6] nor with the presence of the CAA interruption at the 5' end of the CAG repeat, since both control and expansion-carrying individuals have this interruption.[17,27]

A single-base substitution polymorphism ($\underline{C}^{987}GG/\underline{G}^{987}GG$) at the 3' end of the CAG repeat in the *MJD1* gene[5] was recently demonstrated to be associated with differential intergenerational instability of $(CAG)_n$ length.[17] Surprisingly, this was not observed uniquely for the allele in the disease chromosome (in *cis* of the expansion); the combination with the allele in the normal chromosome (in *trans* of the expansion) was shown to effect the degree of intergenerational instability.[28] The biological relevance of these findings remain unclear: whilst it is possible that this association results from a direct effect of the base substitution at the site examined, it may alternatively reflect the effect of a sequence or gene closely linked to this site.

In order to confirm the previous observations in a different population and to distinguish between a direct and indirect effect of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism, we analysed the influence of three intragenic single-base substitution polymorphisms, $\underline{C}^{987}GG/\underline{G}^{987}GG$ and two additional, newly described ones, $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ and $\underline{A}^{669}TG/\underline{G}^{669}TG$,[29] on intergenerational instability of the expanded CAG tract in the *MJD1* gene in our set of families. In addition, we studied a control population for the three intragenic polymorphisms, in an attempt to identify an association between haplotype and $(CAG)_n$ length in normal chromosomes that might give an insight into the origin of the MJD mutation in the human population.

Methods

Subjects

Our study is based on 117 individuals with a previously detected $(CAG)_n$ expansion in the *MJD1* gene, including affected individuals and asymptomatic carriers, constituting 70 parent–progeny pairs, belonging to 21 different families.

Given that the great majority of families in our MJD population was traceable to a Portuguese-Azorean origin, we used a control population of Azorean origin for the study of the association between haplotypes and $(CAG)_n$ length,

including 124 normal individuals with no known relationship to an MJD family.

Blood samples were obtained after informed consent from all individuals, and genomic DNA was extracted from lymphocytes as described elsewhere.[30]

Size of the CAG Repeat and Typing of Intragenic Polymorphisms

Amplification of CAG repeat-containing fragment of the *MJD1* gene was performed by PCR using previously described conditions,[5] and the size of the PCR products was determined by denaturing polyacrylamide gel electrophoresis, in parallel with an M13 sequence ladder.

The intragenic polymorphism $\underline{C}^{987}GG/\underline{G}^{987}GG$, situated at the 3' end of the CAG repeat, was detected by allele-specific PCR using primers ASP1 (ACTCTGTCCTGATAGGTCCCC) or ASP2 (ACTCTGTCCTGATAGGTCCCG) in combination with MJD52.[5] The stop codon polymorphism $TA\underline{A}^{1118}/TA\underline{C}^{1118}$[29] was detected by allele-specific PCR using primers ASP3 (GCAAAAATCACATGGAGCTCG) or ASP4 (GCAAAAATCACATGGAGCTCG) in combination with MJD52. In both cases, the PCR reaction was performed using the same conditions used for amplification of the CAG repeat, except for the annealing step, which was performed at 61°C for 30 s. The PCR products were analysed on a denaturing 6% polyacrylamide gel and visualised by autoradiography.

The $\underline{A}^{669}TG/\underline{G}^{669}TG$ polymorphism[29] was detected by SSCP analysis using a 0.5X MDE gel (FMC BioProducts) with 5% glycerol for electrophoresis; the PCR amplification was performed using primers MJDIVSR (TACTAGAGCTTATTTGCCAG) and MJD734R (CAGAGCCCTCTGCAAATCCT), in a solution containing 1X PCR buffer (Perkin Elmer), 0.2 mM dNTP-A, 0.03 mM dATP, 8 ng/µl each primer, 0.2 µCi ^{35}S dATP, 0.1 µg/µl BSA, 0.12 U AmpliTaq (Perkin Elmer). The amplification conditions were as follows: initial denaturation for 5 min at 94°C, followed by 1 min at 94°C, 1 min at 56°C and 1 min at 72°C for 30 cycles and a final elongation period of 5 min at 72°C. When necessary, the polymorphisms were typed for both parents and progeny, in order to determine phase.

Statistical Methods

In order to analyse the associations between intragenic polymorphisms and the size of the CAG repeats in normal chromosomes (of control individuals), we compared the frequencies of the alleles in three classes of normal chromosomes, grouped according to size of $(CAG)_n$ (up to 21, 22 to 26, and more than 27 repeat units), using Fisher's exact test $2 \times k$ and Pearson χ^2 test.

The intergenerational changes in CAG repeat length were analysed by comparison of absolute change in size of the expanded allele during transmission. Fisher's exact test (2×2) was used to determine if intergenerational change in repeat number between parent and offspring, expressed as two classes of variation (difference of more than two repeat units between parent and progeny – large variation, and difference of up to two repeat units – small variation), was associated with sex of transmitting parent.

The distribution of the variable intergenerational changes in CAG repeat length as defined is highly skewed, therefore non-parametric methods are appropriate for hypothesis testing; Mann-Whitney U test or Kruskal-Wallis ANOVA were used to assess associations between alleles or haplotypes and instability.

Linkage disequilibrium analysis was performed using Fisher's exact test.

All analyses were performed using procedures of SPSS, release 6.1.3, ©SPSS Inc. (1989–1995) and Exact $2 \times k$- exact test for a $2 \times k$ table, version 2.04, ©JH Abramson & PM Gahlinger (1993–1996).

Results

No Apparent Association between the Alleles of Intragenic Polymorphisms *in* cis *of the* $(CAG)_n$ *Expansion and Amplitude of Intergenerational Variation of Expanded* $(CAG)_n$ *Size*

We detected no significant difference in the distribution of absolute variation of $(CAG)_n$ length during transmission between individuals carrying either the $\underline{C}$ or $\underline{G}$ alleles of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism in *cis* with the expanded allele (Mann-Whitney U test, $P = 0.150$). The comparison of absolute variation in repeat size between individuals carrying either the $\underline{A}$ or the $\underline{C}$ variants of the $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ polymorphism in *cis* with the expanded CAG also revealed absence of a significant difference in the median of the absolute variation in repeat size between the two groups (Mann-Whitney U test, $P = 0.450$). The same was observed for the $\underline{A}^{669}TG/\underline{G}^{669}TG$ polymorphism (Mann-Whitney U test, $P = 0.700$).

No association was detected between the presence of any particular (complete or partial) intragenic haplotype in *cis* (defined by the three intragenic polymorphisms mentioned above) with increased instability (Kruskal-Wallis test, $P = 0.378$).

A borderline significant effect of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism in *cis* with the expanded allele was detected only when the stronger effect of the allele in *trans* was taken into account (see below).

Influence of the Genotype of Intragenic Polymorphisms on the Intergenerational Instability of the Expanded $(CAG)_n$

When the transmission events were divided into groups according to genotype of the transmitting parent for the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism, taking both the alleles in *cis* and in *trans* into account (Table 1), there was a significant difference in the median of the absolute variation in repeat length during transmission between the groups, with the parental (CAG)exp-$\underline{C}$GG/(CAG) normal-$\underline{G}$GG genotype associated with increased instability (Kruskal-Wallis test, $P = 0.028$).

When the genotype of the stop codon polymorphism was considered, the (CAG)exp-TA A/(CAG)normal-TA C polymorphism was associated with larger intergenerational variation, but the difference was not statistically significant (Kruskal-Wallis test, $P = 0.080$).

The parental genotype for the A^{669}TG/G^{669}TG polymorphisms did not have a significant effect on instability (Kruskal-Wallis test, $P = 0.632$).

Effect of the Alleles of Intragenic Polymorphisms Present in the Normal Chromosome on the Amplitude of Intergenerational Variation of Expanded $(CAG)_n$ Size in the MJD Chromosome

A highly significant association was observed (Table 2) between the presence of the G variant of the C^{987}GG/G^{987}GG polymorphism on the normal chromosome of the transmitting parent and an increased amplitude of intergenerational variation in expanded repeat size (Mann-Whitney U test, $P = 0.019$).

A similar analysis for both the TA A^{1118}/TA C^{1118} and the A^{669}TG/G^{669}TG polymorphisms revealed an association of the C variant of the stop codon polymorphism present in the normal chromosome of the transmitter with an increased intergenerational instability of the expanded repeat (Mann-Whitney U test, $P = 0.024$), but no such association was detected for the A^{669}TG/G^{669}TG polymorphism (Mann-Whitney U test, $P = 0.291$).

Given these results, we also analysed the effect of the polymorphisms present in *cis* of the expanded allele on its instability controlling for the effect of the polymorphism present in *trans*. Considering only the individuals that had a GGG variant in *trans* (which is associated with increased intergenerational instability), when we compared the degree of instability between GGG and CGG carriers in *cis* ($n = 20$ and $n = 33$, respectively), the difference in instability was of borderline significance, in spite of a small sample size (Mann-Whitney U test, $P = 0.055$), indicating an increased instability in carriers of the C variant in *cis*. For the stop codon polymorphism, however, considering the pairs where the progenitor had the C variant in the normal chromosome, the difference between C ($n = 23$) and G ($n = 16$) carriers was not significant (Mann-Whitney U test, $P = 0.19$).

We did not detect a significant association of any particular intragenic haplotype (defined by the three intragenic polymorphisms mentioned above) in the normal chromosome of the transmitter with increased intergenerational instability of the expanded repeat (Kruskal-Wallis, $P = 0.321$). However, when partial haplotypes, including only two of the intragenic polymorphisms, were analysed, the haplotype defined by the C^{987}GG/G^{987}GG and TA A^{1118}/TA C^{1118} polymorphisms seemed to have an effect on instability, with the GC haplotype in *trans* associated with larger variations of the expanded repeat size (Kruskal-Wallis, $P = 0.042$). The partial haplotypes defined by the A^{669}TG/G^{669}TG and C^{987}GG/G^{987}GG polymorphisms, and by the A^{669}TG/G^{669}TG and TA A^{1118}/TA C^{1118} polymorphisms were not associated with differential instability (Kruskal-Wallis, $P = 0.270$ and $P = 0.197$, respectively).

Table 1 Association of genotypes (alleles in *cis*+*trans* of the expanded repeat) of the three intragenic polymorphisms in the *MJD1* gene with increased intergenerational instability of the repeat (Kruskal-Wallis)

polymorphism	*all transmissions*	*male transmissions*	*female transmissions*
A^{669}TG/G^{669}TG	p=0.632	p=0.914	p=0.304
C^{987}GG/G^{987}GG	p=0.028 (C+G)	p=0.623	p=0.014 (C+G)
TA A^{1118}/TA C^{1118}	p=0.080	p=0.685	p=0.027 (C+C)

Table 2 Association of alleles of the three intragenic polymorphisms in the *MJD1* gene in *trans* of the expanded $(CAG)_n$ with increased intergenerational instability of the repeat (Mann-Whitney test)

polymorphism	*all transmissons*	*male transmissons*	*female transmissons*
A^{669}TG/G^{669}TG	p=0.291	p=0.942	p=0.092
C^{987}GG/G^{987}GG	p=0.019 (G)	p=0.301	p=0.040 (G)
TA A^{1118}/TA C^{1118}	p=0.024 (C)	p=0.650	p=0.003 (C)

Combined Analysis of the Effects of Intragenic Polymorphisms and Gender of the Transmitting Parent

Knowing that gender of transmitting parent is an important factor determining the degree of intergenerational instability of the expanded $(CAG)_n$ in MJD, we analysed the influence of intragenic polymorphisms on instability separately for maternal and paternal transmissions.

As is shown in Table 1, the significance of the effect of genotype of intragenic polymorphisms on intergenerational instability was increased when we considered maternal transmissions only: for paternal transmissions, no significant effect of the genotype was observed. Similarly, when the analysis of the effect of the polymorphisms in *trans* of the expanded allele was

performed separately for paternal and maternal transmissions, we observed a significant effect only in maternal transmissions (Table 2).

The haplotype analysis confirmed these observations: the GC haplotype defined by the $C^{987}GG/G^{987}GG$ and TAA^{1118}/TAC^{1118} polymorphisms in *trans* was associated with increased instability in maternal transmissions only (Kruskal-Wallis, $P = 0.029$).

The direct comparison between haplotypes GA and GC defined by the $C^{987}GG/G^{987}GG$ and TAA^{1118}/TAC^{1118} polymorphisms in maternal transmissions revealed a borderline significant difference in the degree of instability among carriers of these haplotypes ($P = 0.047$, Mann-Whitney U test), which was not detected when the two genders were analysed together ($P = 0.656$, Mann-Whitney U test). The comparison of the CC and GC haplotypes revealed no significant difference in the degree of instability ($P = 0.555$, Mann-Whitney U test).

Larger Sizes of the $(CAG)_n$ in MJD1 are not Associated with Specific Alleles of the Intragenic Polymorphisms

The distribution of CAG repeat sizes in the normal chromosomes of our Portuguese-Azorean control population is grossly trimodal, with peaks at 15, 24 and 28 repeats. To test for association between intragenic polymorphisms and $(CAG)_n$ length in normal chromosomes, three groups of normal alleles (defined by $(CAG)_n$ lengths from 15 to 21, 22 to 26 and 27 to 36) were compared. The frequencies of the C and G alleles of the $C^{987}GG/G^{987}GG$ polymorphism were significantly different between groups (Fisher's exact test, $P = 0.0021$). However, the detailed analysis of the distribution of these alleles revealed that, although there is an increased relative frequency of the C allele among the chromosomes with large CAG repeats when compared with small and medium CAG repeats, this allele is not exclusively associated with the larger repeat lengths or even the predominant allele in this group of chromosomes (Figure 1a).

The same analysis was performed for the TAA^{1118}/TAC^{1118} polymorphism in 231 normal chromosomes, and we observed a difference in distribution ($P < 0.0001$). But although there was an enrichment of allele A among the large CAG repeat-containing normal chromosomes when compared with the small or medium size repeats, no predominance or exclusive association of this allele with larger repeats was observed (Figure 1b). Similar results were obtained for the $A^{669}TG/G^{669}TG$ polymorphism ($n = 154$, $P = 0.0051$ – Figure 1c). A summary of these results is presented in Table 3.

Haplotypes could be determined for 133 individuals, and their distribution according to $(CAG)_n$ sizes revealed that four different haplotypes are present in

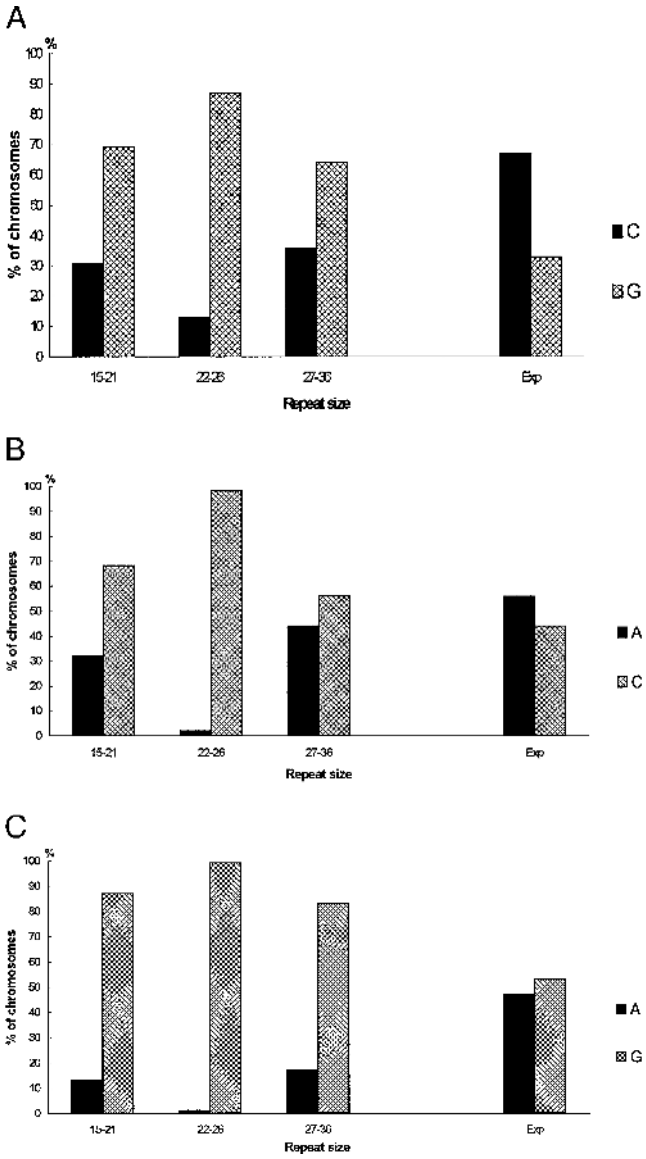

Figure 1a *Distribution of the alleles of the $C^{987}GG/G^{987}GG$ polymorphism according to $(CAG)_n$ size, in a control population* (n = *232) of Portuguese-Azorean origin and in the MJD families of Portuguese descent.* **1b** *Distribution of the alleles of the TAA^{1118}/TAC^{1118} polymorphism according to $(CAG)_n$ size, in a control population* (n = *231) of Portuguese-Azorean origin and in the MJD families of Portuguese descent.* **1c** *Distribution of the alleles of the $A^{669}TG/G^{669}TG$ polymorphism according to $(CAG)_n$ size, in a control population* (n = *154) of Portuguese-Azorean origin and in the MJD families of Portuguese descent*

the normal chromosomes with large CAG repeats (Figure 2), and that, although we could observe an increased relative frequency of the ACA haplotype among these chromosomes when compared with the chromosomes with medium and small $(CAG)_n$ tracts, there was no exclusive or even predominant association of this haplotype among normal chromosomes carrying large CAG repeats. The difference in distribution of haplotypes by $(CAG)_n$ size was not statistically significant ($P = 0.30$).

Analysis of Linkage Disequilibrium Between the Intragenic Polymorphisms in the MJD1 *Gene, in Normal Chromosomes*

In order to clarify the meaning of the results obtained with the partial haplotypes defined by the $\underline{C}^{987}GG/\underline{G}^{987}GG$ and $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ polymorphisms, in *trans* of the MJD mutation, we tested our control population for the existence of linkage disequilibrium between these two intragenic polymorphisms in the *MJD1* gene. Linkage disequilibrium was detected between alleles C and A of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ and $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ polymorphisms ($P < 0.0001$).

Table 3 Comparison of the distribution of alleles of the three intragenic polymorphisms in the *MJD1* gene according to $(CAG)_n$ length in normal chromosomes (Fisher's exact test, considering three size groups)

polymorphism	*allele*	n	*significance*
$\underline{A}^{669}TG/\underline{G}^{669}TG$	a	12	0.0051
	g	142	
$\underline{C}^{987}GG/\underline{G}^{987}GG$	c	56	0.0021
	g	176	
$TA\underline{A}^{1118}/TA\underline{C}^{1118}$	a	49	0.0000
	c	182	

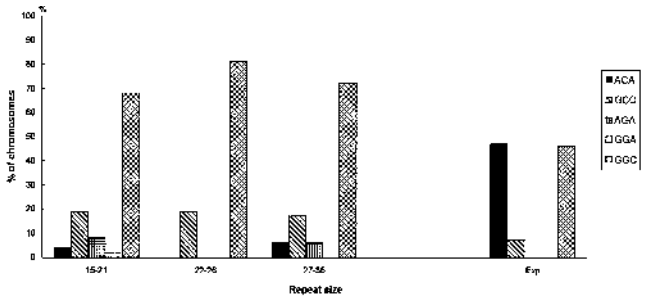

Figure 2 *Distribution of the different intragenic haplotypes, defined by the* $\underline{A}^{669}TG/\underline{G}^{669}TG$, $\underline{C}^{987}GG/\underline{G}^{987}GG$ *and* $TA\underline{A}^{1118}/TA\underline{C}^{1118}$ *polymorphisms, according to* $(CAG)_n$ *size in a control population (*n = *133) of Portuguese-Azorean origin and in the MJD families of Portuguese descent*

Discussion

Intergenerational instability is one of the most important features of the disease-associated trinucleotide expansions, leading to variation in size of the repeat among and within families. This phenomenon partly explains the variable age of onset and severity of the disease, and is the likely basis for the occurrence of anticipation.

The identification of factors such as the length of repetitive tract, presence or absence of interruptions, gender of transmitting parent, chromosomal environment and others that contribute to genetic instability in human disorders caused by expanded trinucleotide repeats provides important clues for the understanding of the biological processes which lead to this instability, and any proposed models for mutation in the trinucleotide repeats should account for these observations. Furthermore, the identification of factors that clearly determine trinucleotide repeat instability in human diseases could possibly allow for prediction of the behaviour of the expanded repeat during transmission in affected families, which would be of manifest relevance for genetic counselling.

Previous studies of the segregation of the expanded $(CAG)_n$ in Machado-Joseph disease have demonstrated the influence of the gender of the transmitting parent[6,7,17] and of an intragenic single-base substitution polymorphism immediately 3' to the $(CAG)_n$ tract on intergenerational instability.[17] In a study by Igarashi and colleagues, the C variant of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism was shown to be preferentially associated both with larger repeats in normal chromosomes and with the expanded repeats in MJD chromosomes, raising the possibility that this substitution confers an increased instability to the $(CAG)_n$ tract. This hypothesis would be in agreement with the fact that most MJD families, even from highly diverse ethnic origins, share a common intragenic haplotype.[31] Surprisingly, in the same study, the analysis of the effect of this polymorphism on intergenerational instability of the expanded repeat in MJD pedigrees demonstrated that the allele carried by the transmitting parent in the normal chromosome strongly affected intergenerational instability, with the (CAG)exp-CGG/(CAG)normal-GGG genotype associated with increased instability.

The results of our analysis of the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism and two additional polymorphisms in a different population confirm some but not all of the previous observations. When the association between $(CAG)_n$ length and the $\underline{C}^{987}GG/\underline{G}^{987}GG$ polymorphism

was assessed in controls, an increased proportion of the C allele was observed among larger normal alleles, in agreement with previous studies.[17,32] Similarly, increased frequencies of allele A of the A^{669}TG/G^{669}TG polymorphism and allele A of the TAA^{1118}/TAC^{1118} polymorphism were detected among longer repeats. It is interesting, however, that in our study as well as the two previous studies[17,32] there was no exclusive association of the larger alleles with any allele of the polymorphisms in the population of Portuguese origin, in contrast with the observations in populations of other ethnic origins (French and Japanese, but not North African). In terms of the intragenic haplotypes defined by these polymorphisms, four different haplotypes were found among the group of larger normal alleles in our population. Although the ACA haplotype, previously shown to be present in approximately 70% of the MJD chromosomes, in a large population of MJD patients of various ethnic origins,[31] was found with increased relative frequency among the larger normal alleles, the difference in distribution of sizes between haplotypes was not statistically significant.

Also, the sequence variations surrounding the expanded $(CAG)_n$ do not seem to have a strong destabilising effect during transmission of this repeat in MJD pedigrees: the analysis of the effect of the C^{987}GG/G^{987}GG polymorphism on the mutation rate of the expanded $(CAG)_n$ in our population of expansion carriers revealed that on its own the allele in *cis* of the expansion has no significant effect on intergenerational instability. This is in agreement with results obtained by Igarashi and colleagues, who could not demonstrate a significant effect of the C^{987}GG/G^{987}GG polymorphism in *cis* of the expansion on instability.

Our results are not compatible with the suggestion that the expanded alleles in MJD chromosomes might have originated from a pool of expansion-prone normal alleles. On the one hand, although instability-predisposing haplotypes have been demonstrated in other trinucleotide repeat-related human disorders, with specific haplotypes in linkage disequilibrium with repeat copy numbers at the high end of the normal range and with the expanded alleles,[18,23,28,33–43] this does not seem to be the case in MJD, at least not in all populations. On the other hand, if there was a instability-predisposing haplotype, its effect should also be detectable on intergenerational instability of the expansion-carrying chromosomes, therefore there should be a noticeable effect of the haplotype on instability in *cis*, which we do not detect. In this sense, it will be interesting to determine whether the recently identified intermediate size CAG repeats in the *MJD1* gene[44,45] are preferentially associated with a particular intragenic haplotype. In one instance, a 51 CAG repeat-containing allele was associated with the GGC haplotype (Maciel, 1997 unpublished).

Both our results and the results of the previous study[17] did, nevertheless, demonstrate the effect of the genotype (combination of the alleles in *cis* and in *trans*) of the C^{987}GG/G^{987}GG polymorphism on intergenerational instability. Furthermore, in our study, the analysis of this polymorphism in *trans* of the expanded $(CAG)_n$ separately revealed a highly significant effect, suggesting that the effect observed when the genotype is considered could be mainly contributed by the allele of the C^{987}GG/G^{987}GG polymorphism in *trans* of the expansion. Although in the study by Igarashi and colleagues there is no mention of the analysis of the allele in *trans* only, it is possible that this is the case in their population as well. The allele in *cis* is also relevant: when the effect of the G allele in *trans* was taken into account, there was a borderline significant difference between the alleles of the polymorphism in *cis*, with allele C associated with increased instability; this reinforces the idea of an allelic interaction being relevant for the mechanism of genetic instability.

Additionally, our results with two other intragenic polymorphisms in *trans* of the expanded MJD allele suggest that this effect of the normal chromosome could be due specifically to the C/G substitution, and not to a more widespread chromosomal effect. The A^{669}TG/G^{669}TG polymorphism does not show an effect on instability, whereas the TAA^{1118}/TAC^{1118} polymorphism does. This could be due to the fact that this polymorphism is in linkage disequilibrium with the C^{987}GG/G^{987}GG polymorphism in the normal population. Unfortunately, the size of our sample does not allow us to study interactions between haplotype and gender and to develop a model considering the contribution of multiple factors to intergenerational instability of the expanded $(CAG)_n$ in the *MJD1* gene. However, when we analysed the effect of the polymorphisms on the normal chromosome and instability separately in maternal and paternal transmissions, we found that this effect was significant only for maternal transmissions, both when we consider each polymorphism separately or when we analyse the effect of the haplotype defined by the C^{987}GG/G^{987}GG and TAA^{1118}/TAC^{1118} polymorphisms. This is in contrast with the observations by Igarashi and colleagues, who

observe a more significant effect during paternal transmissions, and was an unexpected result, considering the previously demonstrated higher instability of the expanded repeat during paternal transmissions. A simple explanation for the differences observed would be the existence of sampling biases, aggravated by relatively small sample sizes in both studies. It is also conceivable that different mechanisms might be involved in instability during maternal and paternal transmissions.

The biological mechanisms by which the contribution of the normal chromosome to instability might occur are not clear. One possibility is the occurrence of a direct interaction between chromosomes, such as in gene conversion events.[17] Gene conversion has been shown to occur in the Huntington disease gene, with variation in repeat size but no alteration of flanking markers.[46,47] Since the variation in size of the expanded allele in MJD pedigrees occurs with no corresponding change in normal allele, the mechanism of unequal crossing-over would have to involve complex interactions, such as hydrogen bonding between hairpins on the homologous alleles during replication, as suggested by McMurray and colleagues.[48] Besides, any changes in repeat number must occur on the 3' side of the CAG tract, because the CAA interruption seems to be in the same position in most chromosomes, or else absent.[27] This polarity of variation was previously demonstrated in gene conversion events leading to mutation of human minisatellites.[49,50] In this perspective, it is noteworthy that the intragenic polymorphisms which seem to affect the genetic instability of the CAG repeat would be localised in the 3' flanking region of this repeat. To conclude, our findings, although provocative, remain mostly unexplained at the molecular level, stressing the need for further studies of the mechanisms by which the normal chromosome might contribute to instability of the expanded CAG tract in the MJD causative gene.

Acknowledgements

We are very grateful to the members of the MJD families for their cooperation in the molecular studies. We would also like to thank Drs Anita de Stefano and Lindsay Farrer (Boston University School of Medicine), and Dr Ridha Joober (McGill University) for their suggestions and the insightful discussion of the statistical analysis.

This project was supported by an NIH grant (NS 31687), the Fonds de Recherche en Santé du Québec, the Association Canadienne de l'Ataxie de Friedreich, the Medical Research Council of Canada, the Network of Centres of Excellence (Canadian Genetic Diseases Network) and a PRAXIS XXI grant. Patrícia Maciel and Claudia Gaspar are the recipients of a scholarship from Programa PRAXIS XXI, Ministério da Ciência e Tecnologia, Portugal.

References

1 Coutinho P, Andrade C: Autosomal dominant system degeneration in Portuguese families of the Azores Islands. *Neurology* 1978; **28**: 703–709.
2 Rosenberg RN, Nyhan WL, Bay C, Shore P: Autosomal dominant striatonigral degeneration: a clinical, pathological, and biochemical study of a new genetic disorder. *Neurology* 1976; **26**: 703–714.
3 Coutinho P, Sequeiros J: Aspects cliniques, génétiques et pathologiques de la maladie de Machado-Joseph. *J Génét Hum* 1981; **29**: 203–209.
4 Barbeau A, Roy M, Cunha L *et al*: The natural history of Machado-Joseph disease. *Can J Neurol Sci* 1984; **11**: 510–525.
5 Kawaguchi Y, Okamoto T, Taniwaki M *et al*: CAG expansions in a novel gene from Machado-Joseph disease at chromosome 14q32.1. *Nat Genet* 1994; **8**: 221–227.
6 Maciel P, Gaspar C, DeStefano AL *et al*: Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *Am J Hum Genet* 1995; **57**: 54–61.
7 Maruyama H, Nakamura S, Matsuyama Z *et al*: Molecular features of the CAG repeats and clinical manifestation of Machado-Joseph disease. *Hum Mol Genet* 1995; **4**: 807–812.
8 Fu Y-H, Kuhi DPA, Pizzuti A *et al*: Variation of CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991; **67**: 1047–1058.
9 Snow K, Doud LK, Hagerman R, Pergolizi RG, Erster SH, Thibodeau SN: Analysis of a CGG sequence at the FMR-1 locus in fragile X families and in the general population. *Am J Hum Genet* 1993; **53**: 1217–1228.
10 Fisch GS, Snow K, Thibodeau SN *et al*: The fragile X premutation in carriers and its effect on mutation size in offspring. *Am J Hum Genet* 1995; **56**: 1147–1155.
11 Lavedan C, Radvanyi HH, Shelbourne P *et al*: Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am J Hum Genet* 1993; **52**: 875–883.
12 Duyao M, Ambrose C, Myers R *et al*: Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat Genet* 1993; **4**: 387–392.
13 Goldberg YP, Kremer B, Andrew SE *et al*: Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. *Nat Genet* 1993; **5**: 174–179.
14 De Rooij KE, De Kooning Gans PAM, Skraastad MI *et al*: Dynamic mutation in Dutch Huntington's disease patients: increased paternal repeat instability extending to within the normal size range. *J Med Genet* 1993; **30**: 996–1002.

15 Ranen NG, Colin Stine O, Abbott MH *et al*: Anticipation and instability of IT-15 (CAG)n repeats in parent–offspring pairs with Huntington disease. *Am J Hum Genet* 1995; **57**: 593–602.
16 Chung M, Ranum LPW, Duvick LA, Servadio A, Zoghbi HY, Orr HT: Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nat Genet* 1993; **5**: 254–258.
17 Igarashi S, Takiyama Y, Cancel G *et al*: 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. *Hum Mol Genet* 1996; **5**: 923–932.
18 Pianese L, Cavalcanti F, De Michele G *et al*: The effect of parental gender on the GAA dynamic mutation in the *FRDA* gene. *Am J Hum Genet* 1997; **60**: 460–463.
19 Yu S, Mulley J, Loesch D, Turner G *et al*: Fragile-X syndrome: unique genetics of the heritable unstable element. *Am J Hum Genet* 1992; **50**: 968–980.
20 Hunter A, Tsilfidis C, Mettler G *et al*: The correlation of age of onset with CTG trinucleotide repeat amplification in myotonic dystrophy. *J Med Genet* 1992; **29**: 774–779.
21 Eichler E, Holden J, Popovich BW *et al*: Length of uninterrupted CGG repeats determines instability in the *FMR-1* gene. *Nat Genet* 1994; **8**: 88–94.
22 Kunst CB, Warren ST: Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. *Cell* 1994; **77**: 853–861.
23 Snow K, Tester DJ, Kruckeberg KE, Thibodeau SN: Sequence analysis of the fragile X trinucleotide repeat: correlations with stability and haplotype, and implications for the origin of fragile X alleles. *Am J Hum Genet* 1994; **55**: A243.
24 Nancarrow JK, Holman K, Mangelsdorf M *et al*: Molecular basis of p(CGG)$_n$ repeat instability at the FRA16A fragile site locus. *Hum Mol Genet* 1995; **4**: 367–372.
25 Leeflang EP, Arnheim N: A novel repeat structure at the myotonic dystrophy locus in a 37 repeat allele with unexpectedly high stability. *Hum Mol Genet* 1995; **4**: 135–136.
26 Loesch D, Huggins R, Petrovic V, Slater H: Expansion of the CGG repeat in fragile x in the FMR1 gene depends on the sex of the offspring. *Am J Hum Genet* 1995; **57**.
27 Limprasert P, Nouri N, Heyman R *et al*: Analysis of CAG repeat of the Machado-Joseph gene in human, chimpanzee and monkey populations: a variant nucleotide is associated with the number of CAG repeats. *Hum Mol Genet* 1996; **5**: 207–213.
28 Igarashi S, Takiyama Y, Takano HKS *et al*: Single base polymorphism at the 3' end of the CAG repeat in *MJD1* gene is associated with the size of CAG repeat and the intergenerational instability. *Am J Hum Genet* 1996; **5**: 923–932.
29 Goto J, Watanabe M, Ichikawa Y *et al*: Machado-Joseph disease gene products carrying different carboxyl termini. *Neurosci Res* 1997; **28**: 373–377.
30 Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
31 Gaspar C, Goto J, Lopes-Cendes I *et al*: Different origins of mutation in Machado-Joseph disease patients. Ist International Symposium on Inherited Ataxias. Montreal, 1997.
32 Stevanin G, Lebre A-S, Mathieux C *et al*: Linkage disequilibrium between the spinocerebellar ataxia 3/Machado-Joseph disease mutation and two intragenic polymorphisms, one of which, (X359Y), affects the stop codon. *Am J Hum Genet* 1997; **60**: 1548–1552.
33 Richards RI, Holman K, Friend K *et al*: Evidence of founder chromosomes in fragile X syndrome. *Nat Genet* 1992; **1**: 257–260.
34 MacPherson JN, Bullman H, Youings SA, Jacobs PA: Insert size and flanking haplotype in fragile X and normal populations: possible multiple origins for the fragile X mutation. *Hum Mol Genet* 1994; **3**: 399–405.
35 Imbert G, Kretz C, Johnson K, Mandel J-L: Origin of the expansion mutation in myotonic dystrophy. *Nat Genet* 1993; **4**: 72–76.
36 Neville C, Mahadevan MS, Barceló JM, Korneluk R: High resolution genetic analysis suggests one ancestral predisposing haplotype for the origin of myotonic dystrophy mutation. *Hum Mol Genet* 1994; **3**: 45–51.
37 Rubinsztein DC, Leggo J, Amos W, Barton DE, Ferguson-Smith MA: Myotonic dystrophy CTG repeats and the associated insertion/deletion polymorphisms in human and primate populations. *Hum Mol Genet* 1994; **3**: 2031–2035.
38 Squitieri F, Andrew SE, Goldberg YP *et al*: DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. *Hum Mol Genet* 1994; **3**: 2103–2114.
39 Rubinsztein DC, Leggo J, Goodburn S, Barton DE, Ferguson-Smith MA: Haplotype analysis of the delta-2642 and (CAG)$_n$ polymorphisms in the Huntington's disease (*HD*) gene provides an explanation for the apparent 'founder' effect. *Hum Mol Genet* 1995; **4**: 203–206.
40 Almqvist E, Chong S, Telenius H *et al*: Contribution of DNA sequence and (CAG)$_n$ size to mutational frequencies of intermediate alleles (IAs) in Huntington disease (HD): evidence from single sperm analyses. *Am J Hum Genet* 1996; **59**: A48.
41 Almqvist E, Spence N, Nichol K *et al*: Ancestral differences in the distribution of the delta-2642 glutamic acid polymorphism is associated with varying CAG repeat lengths in normal chromosomes: insights into the genetic evolution of Huntington disease. *Hum Mol Genet* 1995; **4**: 207–214.
42 Murray A, MacPherson JN, Pound MC *et al*: The role of the size, sequence and haplotype in the stability of FRAXA and FRAXE alleles during transmission. *Hum Mol Genet* 1997; **6**: 173–184.
43 Yanagisawa H, Fujii K, Nagafuchi S *et al*: A unique origin and multistep process for the generation of expanded DRPLA triplet repeats. *Hum Mol Genet* 1996; **5**: 373–379.
44 Takiyama Y, Sakoe K, Nakano I, Nishizawa M: Machado-Joseph disease: Cerebellar ataxia and autonomic dysfunction in a patient with the shortest known expanded allele (56 CAG repeat units) of the *MJD1* gene. *Neurology* 1997; **49**: 604–606.

45 Maciel P, Rousseau M, Gaspar C, Coutinho P, Sequeiros J, Rouleau GA: An intermediate allele of 51 repeats found at the Machado-Joseph disease (MJD) locus: implications for predictive testing and the study of the origin of the MJD mutation. *Am J Hum Genet* 1997; **61**: A314.
46 Warner JP, Barron LH, Fitzpatrick DR, Brock DJH: A gene conversion event at the Huntington's CAG repeat. *Am J Hum Genet* 1996; **59**: 1697.
47 Nicol K, Goldberg YP, Almqvist E, Sillence D, Richards F, MRH: Interallelic conversion events associated with trinucleotide instability in the Huntington disease gene. *Am J Hum Genet* 1996; **59**: 1590.
48 McMurray CT: Mechanisms of DNA expansion. *Chromosoma* 1995; **104**: 2–13.
49 Jeffreys AJ, Tamaki K, MacLeod A, Monckton DG, Neil DL, Armour JAL: Complex gene conversion events in germline mutation at human minisatellites. *Nat Genet* 1994; **6**: 136–145.
50 Monckton DG, Neumann R, Guram T *et al*: Minisatellite mutation rate variation associated with a flanking DNA sequence polymorphism. *Nat Genet* 1994; **8**: 162–170.