Genetics and Molecular Biology, 26, 2, 115-119 (2003)

www.sbg.org.br

Research Article

The *SCA1* (Spinocerebellar ataxia type 1) and *MJD* (Machado-Joseph disease) CAG repeats in normal individuals: segregation analysis and allele frequencies

Cláudia Emília Vieira Wiezel, Maria do Carmo Tomitão Canas and Aguinaldo Luiz Simões

Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brazil.

Abstract

Spinocerebellar ataxia type 1 (*SCA1*) and Machado-Joseph disease (*MJD/SCA3*) are autosomal dominant neurodegenerative diseases caused by expansions of a CAG trinucleotide repeat in the *SCA1* and *MJD* genes. These expanded sequences are unstable upon transmission, leading to an intergeneration increase in the number of repeats (dynamic mutation). The transmission of the CAG repeat was studied in normal mother-father-child trios, referred for paternity testing (*SCA1*, n = 367; *MJD*, n = 879). No segregation distortion was detected. The CAG allele frequencies were determined in 330 unrelated individuals (fathers from couples tested for paternity). The allele frequency distributions did not differ from those previously reported for European populations. The estimated values for the statistic parameters indicating diversity at the *SCA1* locus did not differ much from those reported previously for other STRs in the Brazilian population, while those for the *MJD* locus were close to or higher than the maximum values of previous reports. This shows that *SCA1* and *MJD* are highly informative loci for applications in genetic and population studies and for forensic analysis.

Key words: segregation distortion, Spinocerebellar ataxia type 1, Machado-Joseph disease.

Received: April 2, 2002; accepted: April 7, 2003.

Introduction

Spinocerebellar ataxia type 1 (*SCA1*) and Machado-Joseph disease (*MJD*/*SCA3*) are autosomal dominant neurodegenerative diseases caused by expansions of a CAG repeat in the *SCA1* and *MJD* genes, respectively. Segregation distortion favoring the transmission of mutated or normal alleles during meiosis has been reported for both genes. (Ikeuchi *et al.*, 1996; Riess *et al.*, 1997; Takiyama *et al.*, 1997; Iughetti *et al.*, 1998). Rubinsztein and Leggo (1997) added to these observations, reporting the preferential transmission of alleles with smaller CAG repeats by normal females. More recently, however, Mac Millan *et al.* (1999) did not find any evidence of segregation distortion upon transmission of CAG repeat alleles by normal individuals.

Normal variation in size of the CAG repeats of *SCA1* and *MJD* genes has been reported in a few surveys (Rubinsztein *et al.*, 1995; Watkins *et al.*, 1995; Limprasert *et al.*, 1996; Richards *et al.*, 1996; Jodice *et al.*, 1997; Limprasert *et al.*, 1997; Takano *et al.*, 1998), as well as in comparative studies carried out in affected and control groups (Dürr *et al.,* 1996; Goldfarb *et al.,* 1996; Maruyama *et al.,* 1996; Matsumura *et al*., 1996; Hsieh *et al.,*1997; Soong *et al.,* 1997; Zhou *et al.,* 1997; Lokkegaard *et al.,* 1998).

In spite of being highly polymorphic, triplet repeats in the normal range are consistently transmitted unchanged from parents to children (Limprasert *et al.*, 1994; Richards and Sutherland, 1994), and the repeat size variation in the *DRPLA* (dentatorubral-pallidoluysian atrophy) gene has recently been shown to be reliable for use in forensic tests, especially paternity tests (Pelotti *et al.*, 1998).

There are no frequency estimates of the *SCA1* and *MJD* alleles in the Brazilian population, and segregation analysis has produced conflicting results (Rubinsztein and Leggo, 1997; MacMillan *et al.,* 1999). We investigated allele segregation in mother-child and father-child pairs, and estimated the frequencies of *SCA1* and *MJD* alleles in a sample of clinically normal individuals from the northeastern region of the State of São Paulo, Brazil. We also evaluated their applications in paternity investigations, by comparisons with other markers. The absence of segrega-

Send corresponding to Aguinaldo Luiz Simões. Departamento de Genética, Faculdade de Medicina, USP, Avenida Bandeirantes 3900, 14049-900 Ribeirão Preto, SP, Brazil. E-mail: alsimoes@fmrp.usp.br.

tion distortion herein reported corroborates the data presented by MacMillan *et al.* (1999).

Material and Methods

The study sample comprised individuals seeking paternity investigation at the Clínica Civil of the Ribeirão Preto University Hospital, University of São Paulo, during the years 1996 and 1997. All individuals were white (skin color was determined visually), originating from Ribeirão Preto and nearby cities. They formed mother-child-alleged father trios, routinely submitted to the investigation of a set of six or more STRs.

For calculating the *SCA1* and *MJD* allele frequencies, we considered mothers and alleged fathers of 165 trios (Sample I) consecutively submitted to the same set of tests for paternity determination, totaling 330 genetically unrelated individuals, or 660 chromosomes. For estimating segregation distortion, 118 trios with confirmed paternity were selected from this sample. In addition, other 249 trios were examined for the *SCA1* locus, and 761 trios for the *MJD* locus. The last two groups (Sample II) were neither consecutively collected, nor submitted to the same set of paternity tests, but the individuals were ethnically similar to those in Sample I, and paternity was confirmed with a probability of 99,99%. The difference in sample size between the two groups reflects nothing but the longer time of use of the *MJD* locus in the laboratorial routine.

Genomic DNA was extracted from 300 µL of whole blood, as described by Higuchi (1989). The fragments containing the CAG repeats of the *SCA1* and *MJD* genes were amplified by PCR, using the primers and conditions described by Orr *et al.* (1993) and Kawaguchi *et al.* (1994), respectively. The amplified products were electrophoresed on a 12% denaturing polyacrylamide gel, followed by silver nitrate staining. The PCR products of about 150 samples were initially genotyped after electrophoresis; the procedure was repeated with samples of apparently equal mobility placed side by side, thus allowing the visualization of an allele ladder and the selection of samples for sequencing, in order to determine the exact number of CAG repeats. These sequenced samples were used as reference.

Exact tests were performed using the GENEPOP program (Raymond and Rousset, 1995). For estimating segregation distortion, we prepared a program that allows to determine the inherited allele by direct genotype analysis of the trios, and records which one of the alleles, the larger or the smaller, was transmitted. The proportions were compared using a χ^2 test, considering the expected ratio of 50%.

This study was approved by the Research Ethics Committee of the Ribeirão Preto University Hospital, Ribeirão Preto School of Medicine, University of São Paulo (HCRP nr. 5158/98).

Results

The alleles were designated according to the number of CAG repeats. Nineteen different *SCA1* alleles were identified, with 19 to 39 repeats. At the *MJD* locus, 21 alleles were found, with 14 to 40 repeats. *SCA1* alleles 20, 24, 38 and 39, as well as the *MJD* allele 40, were observed only in Sample II, and were not considered in the allele frequency analysis. In the segregation distortion analysis, 367 trios were analyzed for the *SCA1* locus, including 734 meioses (Table 1), and 879 trios for the *MJD* locus, totaling 1758 meioses (Table 2). No segregation distortion for the *SCA1* alleles was revealed by the analysis of informative meioses ($p = 0.8516$ and $p = 0.3604$, for 257 maternal and 269 paternal meioses, respectively; Table 1) no distortion was observed for the *MJD* locus either (Table 2), analyzed in 745 maternal and 695 paternal informative meioses ($p = 0.6339$ and $p = 0.2713$, respectively).

Table 1 - Transmission of larger and smaller *SCA1* alleles by normal individuals.

Meiosis	Transmitted allele				χ^2
	Smaller	Larger	Non-informative	Total	
Maternal	130	127	110	367	0.035 (p = 0.8516)
Paternal	142	127	98	367	0.8364 (p = 0.3604)
Total	272	254	208	734	0.6158 (p = 0.4326)

Table 2 - Transmission of larger and smaller *MJD* alleles by normal individuals.

Meiosis	Transmitted alleles				χ^2
	Smaller	Larger	Non-informative	Total	
Maternal	366	379	134	879	0.2268 (p = 0.6339)
Paternal	362	333	184	879	1.21 (p = 0.2713)
Total	728	712	318	1758	0.1776 (p = 0.6734)

The allele frequencies for the *SCA1* and *MJD* loci were determined in 330 unrelated individuals (660 chromosomes) of Sample I: 44 genotypes for *SCA1* and 75 for *MJD* were detected; 76 individuals were homozygous at the *SCA1* locus and 50, at the *MJD* locus.

The allele frequencies found were similar to those reported in the literature for European populations (Figures 1 and 2). At both loci, the most frequent alleles (29 and 30 for *SCA1* and 14 and 23 for *MJD*) were less frequent in African populations.

The parameters of forensic interest were calculated for both loci: Power of Discrimination (PD): *SCA1* = 0.9087 and *MJD* = 0.9572; Power of Exclusion (PE): SCA1 = 0.4583 and *MJD* = 0.50; Polymorphism Information Content (PIC): *SCA1* = 0.7259 and *MJD* = 0.8218; Heterozygocity (H): *SCA1* = 0.7697 and *MJD* = 0.8485.

Discussion

This study comprises the largest sample of individuals analyzed for *SCA1* and *MJD* loci reported in the literature, corresponding, respectively, to 24 and 28% of the total. It revealed the largest number of different alleles in a single sample hereto described, though no new allele was outside the range already reported in the literature.

We did not observe preferential transmission of normal *SCA1* and *MJD* alleles: 526 informative meioses (257 maternal and 269 paternal) for the *SCA1* locus and 1,440 informative meioses (745 maternal and 695 paternal) for the *MJD* locus. These results are in accordance with those of Mac Millan *et al.* (1999), who did not find evidence of segregation distortion in *MJD*, *SCA1* and *DRPLA* loci in their study of 377 pairs of twins and their normal parents. However, the preferential transmission of the smaller alleles of the *MJD* locus has been reported by Rubinsztein and Leggo (1997) in normal women whose smaller allele was transmitted in 166 out of the 290 meioses analyzed, while the men transmitted the smaller allele in 126 out of the 269 meioses.

In the present study, the distributions of the allele frequencies at both loci showed a great similarity with the distributions reported in the literature for European populations (Figures 1 and 2). The most prevalent alleles (29 and 30 in the *SCA1* locus, and 14 and 23 in the *MJD* locus, respectively) are less frequent in African populations, which gives them a greater heterozygozity. For the *MJD* locus, we observed the three modes present in European populations, which correspond to alleles 14, 23, and 27, while, in African populations, up to five modes occur, corresponding to alleles 14, 22, 28, 30, and 33. For the *SCA1* locus, our sample showed a mode around allele 30, similar to European populations, while two modes occur in the African populations, corresponding to alleles 26 and 30. The greater heterozygozity in African populations and the profile of frequencies with multiple modes may be either casual or the result of peculiarities of the places of origin of the few samples hereto analyzed.

Since our study is restricted to a sample from a small region of the State of São Paulo, investigations of other series of individuals from different geographical areas of the country are desirable.

For genes with trinucleotide repeats that expand as dynamic mutations, the alleles in the normal range are easily amplified by PCR, but there is a possibility that larger alleles are not amplified, as is the case in myotonic dystrophy (Gennarelli *et al.*, 1998), resulting in an apparently homozygous genotype. Nevertheless, the CAG expanded repeat has been easily diagnosed by PCR (Brice, 1998). The amplification products corresponding to the normal and expanded alleles can be visualized, and the number of repeats accurately determined after denaturing polyacrylamide gel electrophoresis and silver staining (Vuillaume *et al.*, 1998; Maruyama *et al.*, 1996). We did not detect expanded or intermediate alleles in this study, but, considering the sensitivity of the method used to measure the alleles, the frequency of homozygotes (23.3% for *SCA1* and 15.15% for *MJD*) seems likely to represent the real frequencies in the population.

The estimated values for the statistical parameters of forensic interest (Heterozygozity, Polymorphism Information Content, Power of Exclusion and Power of Discrimina-

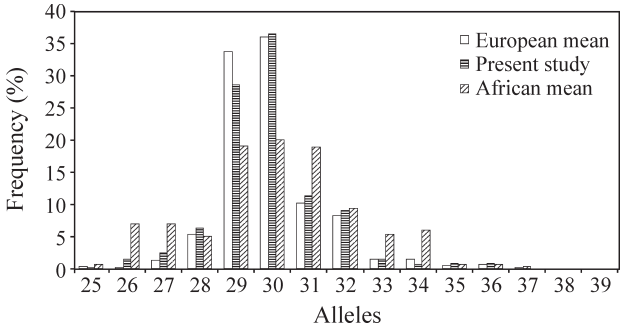

Figure 1 - Allele frequencies of the *SCA1* locus in three population samples. The European and African means were calculated from Watkins *et al*. (1995); Jodice *et al.* (1997); Limprasert *et al*. (1997); Takano *et al*. (1998).

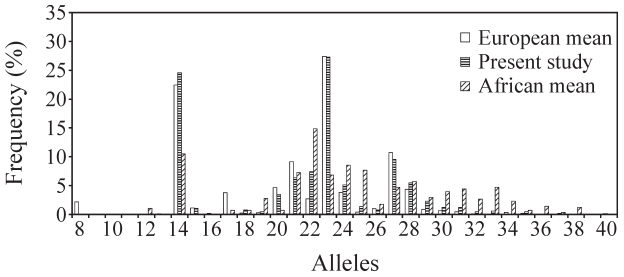

Figure 2 - Allele frequencies of the *MJD* locus in three population samples. The European and African means were calculated from Rubinsztein *et al.* (1995); Limprasert *et al*. (1996); Richards *et al*. (1996); Jodice *et al.* (1997); Takano *et al*. (1998).

tion) for the *SCA1* locus did not differ much from the values reported previously for other loci in the Brazilian population (Pagotto *et al.*, 1999; Leboute, 2000), while those for the *MJD* locus were close to or higher than the maximum values reported. This demonstrates that these loci are highly informative markers for general applications in genetics and population studies, as well as for paternity tests in forensic investigations, much in the same way as the *DRLPA* locus (Pelotti *et al.*, 1998).

Acknowledgments

We thank Ana Lúcia Pimentel for laboratory technical assistance, Dr. Lewis Joel Greene for comments on earlier drafts of this article, to CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and to FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) for financial support.

References

Brice A (1998) Unstable mutations and neurodegenerative disorders. J Neurol 245:505-510.

Dürr A, Stevanin G, Cancel G, Duyckaerts C, Abbas N, Didierjean O, Chneiweiss H, Benomar A, Lyon-Caen O, Julien J, Serdaru M, Penet C, Agid Y and Brice A (1996) Spinocerebellar Ataxia 3 and Machado-Joseph Disease: Clinical, molecular and neuropathological features. Ann Neurol 39:490-499.

Gennarelli M, Pavoni M, Amicucci P, Novelli G and Dallapiccola B (1998) A single Polymerase Chain Reaction-bases protocol for Detecting normal and expanded alleles in Myotonic Dystrophy. Diag Mol Pathol 7 (3):135-137.

Goldfarb LG, Vasconcelos O, Platonov FO, Lunkes A, Kipnis V, Kononova S, Chabrashvili T, Vladimirtsev VA, Alexeev VP and Gajdusek DC (1996) Unstable triplet repeat and phenotypic variabily of Spinocerebellar Ataxia Type 1. Ann Neurol 39:500-506.

Higuchi R (1989) Simple and rapid preparation of samples for PCR. In: Erlich, HA (ed) PCR technology – principles and aplications for DNA amplification. New York, Stockton Press, pp 36.

Hsieh M, Tsai HF, Lu TM, Yang CY, Wu HM and Li SY (1997) Studies of the CAG repeat in the Machado-Joseph disease gene in Taiwan. Hum Genet 100:152-162.

Ikeuchi T, Igarashi S, Takiyama Y, Onodera O, Oyake M, Takano H, Koide R, Tanaka H and Tsuji S (1996) Non- Mendelian transmission in Dentatorubral-Pallidoluysian Atrophy and Machado-Joseph Disease: the mutant allele is preferentially transmitted in male meiosis. Am J Hum Genet 58:730-733.

Iughetti P, Otto P, Zatz M, Bueno MRP and Marie SK (1998) Different behavior in the paternally *vs.* maternaly inherited mutated allele in Brazilian Machado-Joseph (MJD1) families. Am J Med Genet 77:246-248.

Jodice C, Giovannone B, Calabresi V, Bellocchi M and Terrenato L (1997) Population variation analysis at nine loci containing expressed trinucleotide repeats. Ann Hum Genet 61:425-438.

Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiguchi I, Kimura J, Narumiya S and Kakizuka A (1994) CAG expansions in a novel gene from Machado-Joseph disease at chromosome 14q32.1. Nat Genet 8:221-228.

Leboute APM (2000) Polimorfismos genéticos em três amostras da população urbana brasileira. PhD Thesis, Universidade de São Paulo, Ribeirão Preto.

Limprasert P, Nouri N and Keats BJB (1994) Meiotic stability and polymorphism of CAG repeat in normal chromosome at SCA1 locus. Am J Hum Genet (supl): A55:193.

Limprasert P, Nouri N, Heyman RA, Nopparatana C, Kamonsilp M, Deininger PL and Keats BJB (1996) Analysis 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 5(2):207-213.

Limprasert P, Nouri N, Nopparatana C, Deininger PL and Keats BJB (1997) Comparative studies of the CAG repeats in the Spinocerebellar Ataxia Type 1 (SCA1) gene. Am J Med Genet 74:488-493.

Lokkegaard T, Nielsen JE, Hasholt L, Fenger K, Werdelin L, Tranebjaerg L, Lauritzen M, Colding-Jorgensen E, Gronbech-Jensen M, Henriksen AO and Sorensen SA (1998) Machado-Joseph disease in three Scandinavian families. J Neurol Sci 156:152-157.

Mac Millan JC, Voisey J, Healey SC and Martin NG (1999) Mendelian segregation of normal CAG trinucleotide repeat alleles at normal CAG trinucleotide repeat alleles at three autosomal loci. J Med Genet 36:258-259.

Maruyama H, Kawakami H and Nakamura S (1996) Reevaluation of the exatc CAG repeat length in hereditary cerebellar ataxias using highly denaturing conditions and long PCR Hum Genet 97:591-595.

Matsumura R, Takayanagi T, Murata R, Futamura N, Hirano M and Ueno S (1996) Relationship of (CAG)nC configuration to repeat instability of the Macahdo-Joseph disease gene. Hum. Genet 98:643-645.

Orr HT, Chung M, Banfi S, Kwiatkowski Jr TJ, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LPW and Zoghbi HY (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nat Genet 4:221-225.

Pagotto RC, Canas MCT, Brito ROAA and Simões AL (1999) Allele frequencies of the human von Willebrand factor gene (vWF) in a Brasilian population sample. Int J Legal Med 112:326-328.

Pelotti S, Mantovani V, Esposti PD, D'Apote L, Bragliani M, Maiolini E, Abbondanza A and Pappalardo G (1998) The DRPLA CAG repeats in aan Italian population sample: evaluation of the polymorphism for forensic applications. J Forensic Sci 43(2):410-412.

Raymond M and Rousset F (1995) GENEPOP (version 1.2): population genetics software for exat tests and ecumeniscism. J Hered 86:248-249.

Richards RI and Sutherland GR (1994) Simple repeat DNA is not replicated simply. Nat.Genet 6:114-116.

Richards RI, Crawford J, Narahara K, Mangelsdorf M, Friend K, Staples A, Denton M, Easteal S, Hori TA, Kongo I, Jenkins T, Goldman A, Panich V, Ferakova E and Sutherland GR (1996) Dynamic mutation loci: alleles distribuitions in different populations. Ann Hum Genet 60:391-400.

Riess O, Epplen JT, Amoiridis G, Przuntek H and Schöls L (1997) Transmission distortion of the mutant alleles in spinocerebellar ataxia. Hum Genet 99:282-284.

Rubinsztein DC, Leggo J, Coetzee GA, Irvine R, Bucley M and Ferguson-Smith MA (1995) Sequence variation and size ranges of CAG repeats in the Machado-Joseph disease, spinocerebellar ataxia type 1 e androgen receptor genes. Hum Mol Genet 4(9):1585-1590.

Rubinsztein DC and Leggo J (1997) Non-Mendelian transmission at the Machado-Joseph disease locus in normal females: preferential transmission of the alleles with smaller CAG repeats. J Med Genet 34:234-236.

Soong B, Cheng C, Liu R and Shan D (1997) Machado-Joseph disease: Clinical, molecular, and metabolic characterization in Chinese Kindreds. Ann Neurol 41:446-452.

Takano H, Cancel G, Ikeuchi T, Lorenzetti D, Mawad R, Stevanin G, Didierjean O, Dürr A Oyake M, Shimohata T, Sasaki R, Koide R, Igarashi S, Hayashi S, Takiyama Y, Nishizawa M, Tanaka H, Zoghbi H, Brice A and Tsuji S (1998) Close associations between prevalences of dominantly inherited spinocerebellar ataxias with CAG-repeat expansions and frequencies of large normal CAG alleles in japanese and caucasian populations. Am J Hum Genet 63:1060-1066.

Takiyama Y, Sakoe K, Soutome M, Namekawa M, Ogawa T, Nakano I, Igarashi S, Oyake M, Tanaka H, Tsuji S and Nishizawa M (1997) Single sperm analysis of the CAG repeats in the gene for Machado-Joseph disease (MJD1): evidence for non-Mendelian transmission of the MJD1 gene and for the effect of the intragenic CGG/GGG polymorphism on the intergenerational instability. Hum Mol Genet 6(7):1063-1068.

Vuillaume I, Schraen S, Rousseaux J and Sablonnière B (1998) Simple nonisotopic assays for detection of (CAG)n repeats expansions associated with seven neurodegenerative disorders. Diagn Mol Pathol 7(3):174-179.

Watkins WS, Bamshad M and Jorde LB (1995) Population genetics of trinucleotide repeat polymorphisms. Hum Mol Genet 4(9):1485-1491.

Zhou YX, Takiyama Y, Igarashi S, Li YF, Zhou BY, Gui DC, Endo K, Tanaka H, Chen ZH, Zhou LS, Fan MZ, Yang BX, Weissenbach J, Wang GX and Tsuji S (1997) Machado-Joseph disease in four Chinese pedigrees: Molecular analysis of 15 patients including two juvenile cases and clinical correlations. Neurology 48:482-485.

Editor: Francisco Mauro Salzano