



Analysis of the *DMPK* gene CTG repeat in healthy Brazilians

Silvia Regina Sampaio Freitas¹, Pedro H. Cabello¹, Rodrigo Soares Moura-Neto² and Luiz Alves Duro³

¹*Departamento de Genética, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil.*

²*Departamento de Genética, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.*

³*Instituto de Neurologia Deolindo Couto, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.*

Abstract

Myotonic dystrophy (DM) is a neuromuscular disorder caused by the expansion of the cytosine-thymine-guanine (CTG) repeat of the myotonic dystrophy protein kinase gene (*DMPK*). This repeat is highly polymorphic in healthy individuals [(CTG)₅₋₃₇], and it has been proposed that expanded CTG alleles originated from larger sized normal alleles [(CTG)₁₉₋₃₇]. According to this hypothesis, a positive correlation should be expected between the frequency of these large-sized normal alleles and the prevalence of the disorder in a population. We determined the distribution of CTG alleles of the *DMPK* gene in 156 healthy Brazilians from Rio de Janeiro city. Our analyses of 312 chromosomes detected 20 different alleles ranging in size from 5 to 27 CTG repeats, with 24 alleles having more than 18 repeats (7.69%). This frequency of (CTG)₃₋₁₉ alleles observed in our population suggests that the prevalence of DM in Rio de Janeiro should not be different from the prevalence in European populations.

Key words: myotonic dystrophy, dynamic mutation, trinucleotide repeat, CTG repeat, *DMPK* gene.

Received: December 14, 2005; Accepted: July 21, 2006.

Myotonic Dystrophy (DM) is a multi-systemic neuromuscular disorder caused by a dynamic mutation of cytosine-thymine-guanine (CTG) trinucleotide repeats in the 3' untranslated region of the myotonic dystrophy protein kinase gene (*DMPK*) (Brook *et al.*, 1992). The number of CTG repeats observed in normal chromosomes varies from 5 to 37, while DM chromosomes have more than 3000 repeats (Harley *et al.*, 1992). The high variability in the size of expanded CTG alleles is associated with disparity in the ages of DM onset and severity of symptoms observed both within and between families (Brook *et al.*, 1992).

The allelic distribution of the CTG polymorphism and the incidence of DM are also highly variable among various ethnic groups. European populations exhibit an elevated frequency of large-sized normal alleles [(CTG)₁₉₋₃₇] (8%) (Zerylnick *et al.*, 1995) and a DM incidence of 2.2×10^{-5} to 5.5×10^{-5} (Harper, 1989), while in Africans the frequency of (CTG)₁₉₋₃₇ is 1% (Harley *et al.*, 1992) and the occurrence of the disease is apparently zero (Ashizawa and Epstein, 1991). The correlation between the disease incidence and large-sized normal alleles suggests that (CTG)₁₉₋₃₇ repeats predispose towards DM.

Investigations of the origins of expanded DM alleles [(CTG)₃ 50] showed that these large expansions of CTG repeats are in almost complete linkage disequilibrium with several intragenic and extragenic polymorphisms (Goldman *et al.*, 1996), among which the *Alu* insertion/deletion polymorphism, located in intron 8 of the *DMPK* gene, is one of the most studied (Harley *et al.*, 1992; Yamagata *et al.*, 1996). The (CTG)₅ and (CTG)₃ 19 alleles were found to be in complete association with *Alu*(+), while (CTG)₁₁₋₁₃ alleles were in nearly complete association with the *Alu*(-) allele (Deka *et al.*, 1996; Tishkoff *et al.*, 1998). Moreover, these large-sized normal alleles exhibit a mutation rate which increases with repeat length, suggesting a predisposition toward expansion (Harley *et al.*, 1992). These results indicate that DM alleles can arise from (CTG)₃ 19 alleles. Several authors (Deka *et al.*, 1996; Tishkoff *et al.*, 1998) suggest that the frequency of *DMPK* alleles with more than 18 CTG repeats could serve as an estimator of the DM prevalence in a population.

In Brazil, ethnic variability exists due to migrations from European and African populations, this migration having occurred mainly during the post-colonization period. According to Callegari-Jaques and Salzano (1999), 58% of the immigrants who arrived in Brazil between 1500 and 1972 were Europeans, 40% Africans and 2% Asians. An analysis of the *DMPK* gene alleles in the population of

the Brazilian city of Rio de Janeiro (RJ) would be interesting because the degree of ethnic heterogeneity may be reflected in the allelic profile of CTG repeats. During the work described in this paper we determined the allelic distribution of the CTG polymorphism in 156 healthy Brazilians from Rio de Janeiro city and correlated the frequency of large-sized normal CTG alleles with the prevalence of DM.

Genomic DNA from each individual was extracted by a standard salting out method. The polymerase chain reaction (PCR) with the DM-A and DM-B primers was employed to amplify the DMPK gene region containing CTG repeats, as previously described by Ashizawa *et al.* (1993). Denatured PCR products were loaded onto 8% polyacrylamide sequencing gel, submitted to electrophoreses at 1700V for 2 hours and visualized by silver staining. With all tested samples we also ran three samples of known allele sizes [(CTG)₅, (CTG)₁₅, and (CTG)₂₄] as standards. Allele sizing was performed using the pBR322-HaeIII molecular weight standard and the Sigma Gel® program version 1.0 (Sigma, USA). The GDA statistical program, version 1.19 (Lewis and Zaykin, 2001), was used to calculate allelic frequencies, expected heterozygosity, and to evaluate whether the DMPK locus allele distribution was in Hardy-Weinberg equilibrium. The chi-squared test (χ^2) test was used to compare the allele frequencies of our sample to those reported for African, European and Amerindian populations.

In our sample of 326 chromosomes we detected 20 different alleles with sizes ranging from 5 to 27 CTG repeats (Figure 1), although alleles with 6, 20 and 25 CTG repeats were not found. The distribution of allelic frequencies was tri-modal, with peaks at 5 (~5% of chromosomes analyzed), 11 to 15 and 21 to 24 repeats. The sample distribution was in Hardy-Weinberg equilibrium for this locus ($p = 0.666$), and values of observed (0.789) and expected (0.819) heterozygosities were statistically similar ($p > 0.05$). Comparative analysis of the allele distributions in RJ-Brazilians and Europeans indicated that they shared the same CTG repeats profile ($\chi^2 = 0.664$; $p = 0.705$) (Table 1). Alleles with more than 18 repeats were found equally frequently in RJ-Brazilians (7.69%) and other European populations (7.63%) (Zerylnick *et al.*, 1995; Tishkoff *et al.*,

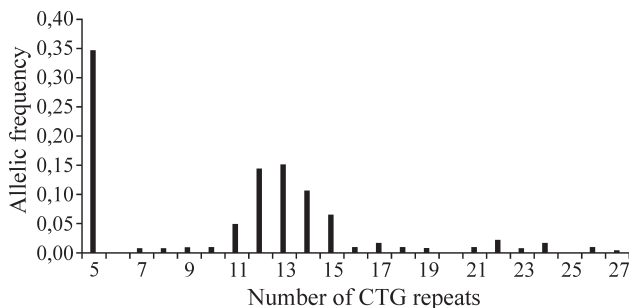


Figure 1 - Distribution of allelic frequency at DMPK locus in a sample (n = 156) of the healthy Rio de Janeiro population.

Table 1 - Frequencies of (CTG)_n classes in Africans, Europeans, Amerindians and the Brazilians from the present study.

Classes	Ethnic Groups (Number of chromosomes assessed)			
	Africans ^{1,4} (1044)	Europeans ^{1,2} (550)	Amerindians ^{1,2} (738)	Rio de Janeiro (312) ⁵
(CTG) ₅	0.2624	0.3745	0.0392	0.3462
(CTG) ₆₋₁₈	0.7337	0.5490	0.9363	0.5769
(CTG) _{≥ 19}	0.0038	0.0763	0.0243	0.0769

¹Zerylnick *et al.*, 1995; ²Deka *et al.*, 1996; ³Goldman *et al.*, 1996; ⁴Gennarelli *et al.*, 1999; ⁵Our present study.

1998). Africans (Zerylnick *et al.*, 1995; Goldman *et al.*, 1996; Tishkoff *et al.*, 1998; Gennarelli *et al.*, 1999) ($\chi^2 = 7.037$, $p = 0.027$) and Amerindians (Zerylnick *et al.*, 1995; Tishkoff *et al.*, 1998) ($\chi^2 = 43.921$, $p = 0.000$) failed to show any similarity with the RJ-Brazilian CTG repeat profile. In addition, the Africans and Amerindians groups have less than 5% of large-sized normal alleles, which is significantly less ($p < 0.05$) than those found in the RJ-Brazilians. These results were expected because 61.69% of the Rio de Janeiro population is constituted of individuals of European ancestry (IBGE, 2002). The similar allele profile between Europeans and RJ-Brazilians, especially as regards the risk alleles (8% for the Europeans and 7% for the RJ-Brazilians), suggests a high prevalence of DM occurs in Rio de Janeiro, but until now few clinical reports of this disease have been described in this city. According to the records of the Deolindo Couto Neurology Institute (Venceslau Bras avenue, 95 - Botafogo, Rio de Janeiro/RJ, Brazil), a reference center for muscular dystrophies in Rio de Janeiro, only 67 DM cases were reported from 1961 to 2001. We speculate that the number of DM patients should be higher than those reported because: 1) only a small number of DM patients' relatives, suspected of carrying the disease, receive specialized medical care; 2) it is difficult to diagnose the disease, due to its wide phenotypic range, particularly the late form (with cataracts and baldness), which is associated to the patient's age. The discrepancy between the frequency of risk alleles and the prevalence of DM has also been observed in the city of Ribeirão Preto in the Brazilian state of São Paulo. In a sample of 492 normal chromosomes about 13% presented more than 18 CTG repeats (Camacho R., personal communication; Masters Dissertation (2001), Federal University of São Carlos, São Carlos, São Paulo state, Brazil). The elevated (CTG)₃₁₉ allele frequency in this population is similar to the frequency of 15% among Yemen Jews (Tishkoff *et al.*, 1998), but, however, while the DM prevalence in Yemen Jews was 17×10^{-5} , one of the most elevated thus far found (Tishkoff *et al.*, 1998), DM case reports were not common in Ribeirão Preto, and only ten families were clinically diagnosed. Sub-notification due to the wide phenotypic range of the disease and the absence of specialists and/or a specific molecular

diagnostic test was considered the cause of this low frequency. To improve DM identification, the association of clinical and molecular diagnoses is essential. Studies performed in Italian provinces have shown that after the implementation of DM molecular tests, the estimated prevalence increased from 4×10^{-5} to 7.24×10^{-5} (Siciliano *et al.*, 2001).

In summary, the relatively high frequency of large-sized normal CTG alleles suggests a raised prevalence of myotonic dystrophy in Rio de Janeiro. But DM molecular characterization should be implemented and extended to all relatives of patients in order to obtain precise epidemiological data.

References

- Ashizawa T and Epstein HF (1991) Ethnic distribution of myotonic dystrophy gene. *Lancet* 338:642-643.
- Ashizawa T, Dubel JR and Harati Y (1993) Somatic instability of CTG repeats in myotonic dystrophy. *Neurology* 43:2674-2678.
- Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D and Aburatani H (1992) Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3-prime end of a transcript encoding a protein kinase family member. *Cell* 68:799-808.
- Callegrì-Jacques SM and Salzano FM (1999) Brazilian Indian/non-Indian interactions and their effects. *Ciênc Cult* 51:166-174.
- Deka R, Majumder PP, Shriver MD, Stivers DN, Zhong Y and Yu LM (1996) Distribution and evolution of CTG repeats in the myotonin protein kinase gene in human populations. *Genome Res* 6:142-154.
- Goldman A, Krause A, Ramsay M and Jenkins T (1996) Founder effect and the prevalence of myotonic dystrophy in South Africans: Molecular studies. *Ann Hum Genet* 60:57-65.
- Gennarelli M, Pavoni M, Cruciani F, De Stefano G, Dallapiccola B and Novelli G (1999) CTG repeats distribution and Alu insertion polymorphism at myotonic dystrophy (DM) gene in Amhara and Oromo populations of Ethiopia. *Hum Genet* 105:165-167.
- Harley HG, Brook JD, Rundle SA, Crow S, Reardon W and Buckler AJ (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 355:545-546.
- Harper PS (1989) *Myotonic Dystrophy*. 2nd edition. W.B. Saunders Press, Philadelphia, 150 pp.
- Lewis PO and Zaykin D (2001) Genetic data analysis: Software for the analysis of discrete genetic data. Version 1.16 <http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php>.
- Siciliano G, Manca M, Gennarelli M, Angelini C, Rocchi A and Iudice A (2001) Epidemiology of myotonic dystrophy in Italy: Re-appraisal after genetic diagnosis. *Clin Genet* 59:344-349.
- Tishkoff SA, Goldman A, Calafell F, Speed WC, Deinard AS and Bonne-Tamir B (1998) A global haplotype analysis of the myotonic dystrophy locus: Implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations. *Am J Hum Genet* 62:1389-1402.
- Yamagata H, Miki T, Nakagawa M, Johnson K, Deka R and Ogihara T (1996) Association of CTG repeats and the 1-kb Alu insertion/deletion polymorphism at the myotonin protein kinase gene in the Japanese population suggests a common Eurasian origin of the myotonic dystrophy mutation. *Hum Genet* 97:145-147.
- Zerylnick C, Torroni A, Sherman SL and Warren ST (1995) Normal variation at the myotonic dystrophy locus in global human populations. *Am J Hum Genet* 56:123-130.

Associate Editor: Francisco Mauro Salzano