Acta Neurol Scand 2005: 111: 274–277 DOI: 10.1111/j.1600-0404.2005.00402.x

ACTA NEUROLOGICA
SCANDINAVICA

Haplotype analysis of the DM1 locus in the Serbian population

Krndija D, Savić D, Mladenović J, Rakočević-Stojanović V, Apostolski S, Todorović S, Romac S. Haplotype analysis of the DM1 locus in Serbian population.
Acta Neurol Scand 2005: 111: 274–277. Blackwell Munksgaard 2005.

Objectives – Analysis of the CTG-repeat number and three biallelic markers, *Alu*(+/−), *Hin*fI(+/−), and *Taq*I(+/−), in the *DMPK* gene in healthy and myotonic dystrophy type 1 (DM1) Serbian individuals. Also, the consideration of haplotypes in the light of the proposed models of CTG-repeat evolution and origin of the DM1 mutation. *Materials and methods* – Markers were analyzed by PCR and haplotypes were obtained on 203 unrelated normal chromosomes and 24 unrelated DM1 chromosomes. *Results* – A strong linkage disequilibrium was detected between the three biallelic markers alone ($P < 0.0001$) and between distinct CTG-repeat size classes and reconstructed haplotypes. Greater than 98% of normal chromosomes contain (+ + +) and (− − −) haplotypes. The (+ + +) haplotype is the most common, while the $(CTG)_{9-17}$ are the most frequent alleles. We found a complete association of (+ + +) haplotype with $(CTG)_{\geq 18}$ and mutated alleles. *Conclusions* – $(CTG)_{9-17}$/(+ + +) haplotype is the ancestral haplotype and DM1 mutation occurred on $(CTG)_{18-35}$/ + + + chromosome.

D. Krndija[1], D. Savić[1], J. Mladenović[2], V. Rakočević-Stojanović[3], S. Apostolski[3], S. Todorović[2], S. Romac[1]

[1]Faculty of Biology, [2]Clinic for Pediatrics Neurology and Psychiatry and [3]Institute of Neurology, School of Medicine, University of Belgrade, Belgrade, Serbia and Montenegro

Key words: myotonic dystrophy type 1; DM1; CTG repeats; haplotype analyses; linkage disequilibrium

Prof. Stanka Romac PhD, Faculty of Biology, University of Belgrade, Studentski trg 16, PO Box 52, 11000 Belgrade, Serbia and Montenegro
Tel./Fax: +381 11 639100
e-mail: stanka@bf.bio.bg.ac.yu

Accepted for publication December 16, 2004

Myotonic dystrophy type 1 (DM1) is an autosomal dominant disorder and is present in global populations, including Serbian. DM1 is caused by an expansion of a $(CTG)_n$ repeats in the 3′-UTR of the myotonin protein kinase gene (*DMPK*) (1–3). In a healthy population, the number of CTG repeats is highly polymorphic and varies from 5 to 35, whereas in patients, this number varies from 50 to several thousands (4). Most human populations have a bimodal distribution of allele sizes, with peaks at 5 and between 9 and 17 repeats (mid-sized alleles), but most of the non-African populations have a third group of large-sized alleles, ranged from 18 to 35 repeats (5). There is a positive correlation between the incidence of the disease and the frequency of large-sized alleles in different ethnic groups (5–9). Therefore, it was postulated that $(CTG)_{\geq 18}$ alleles constitute a reservoir for recurrent mutations, which compensate for loss of the mutated alleles due to the negative effect of anticipation, and maintain the disease incidence at a constant rate (10).

Haplotype studies were conducted in order to determine the mechanism that leads to disorder and its high frequency. Polymorphic markers in and around the DM1 locus were analyzed to reconstruct the haplotype background of the various DM1 alleles. Also, the haplotype analyses established different patterns of linkage disequilibrium among the markers and provided various plausible models for the evolution of the CTG repeats at the DM1 locus. The most important among those is the model given by Tishkoff et al. (5). They investigated two intragenic polymorphisms, *Alu*(+/−) deletion polymorphism (11) and *Hin*fI(+/−) restriction-site polymorphism (RSP, 11), and one extragenic *Taq*I(+/−) RSP (D19S463, 12). The relative locations of these polymorphisms are shown in Fig. 1. Tishkoff et al. (5) concluded that the mid-sized alleles are the most ancient and that they arose on *Alu*(+)–*Hin*fI(+)–*Taq*I(+) haplotype background. Other types of chromosomes with higher and lower numbers of CTG repeats arose from this ancestral chromosome. Haplotype analyses of DM1-causing chromosomes

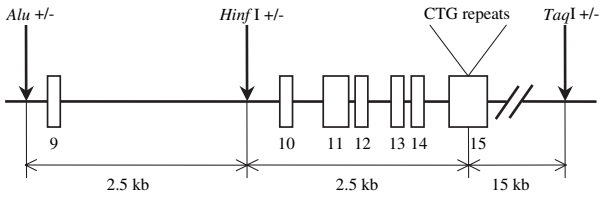

Figure 1. Relative positions of *Alu*(+/−), *Hin*fI(+/−), *Taq*I(+/−) and CTG-repeat polymorphism located in and around *DMPK* gene. Open boxes represent the exons of *DMPK* gene and black lines represent the introns and genomic region 3′ from the *DMPK* gene. *Alu*(+/−) deletion polymorphism is located in intron 8, *Hin*fI(+/−) RSP is located in intron 9, while CTG repeats are located in 3′-UTR of the *DMPK* gene. Extragenic *Taq*I(+/−) RSP is located 15 kb centromeric to the CTG repeats.

found mutated alleles only in association with (+++) haplotype in DM1 individuals of a European origin (12–14). Accordingly, a founder event established $(CTG)_{18-35}$ alleles on a (+++) haplotype, and subsequent expansions produced DM1-causing alleles from one or more of these large-sized alleles (5).

The objectives of this paper were: (i) to provide haplotype data on $(CTG)_n$-*Alu*-*Hin*fI-*Taq*I loci on chromosomes independently drawn from the Serbian healthy population and DM1 individuals, and make conclusions about genomic structure of the Serbian population; (ii) to identify haplotypes that may be specific to Serbians; and (iii) to examine this data in the light of the proposed models for the evolution of $(CTG)_n$ repeats and origin of the DM1 mutation.

Materials and methods

Population sample

A total of 36 healthy and 25 DM1 Serbian families were studied in order to reconstruct haplotypes. In 30 healthy and 24 DM1 families we managed to unambiguously reconstruct haplotypes. Haplotypes were obtained on 203 unrelated normal chromosomes drawn from healthy and DM1 families, and 24 unrelated DM1 chromosomes.

DNA analyses

Genomic DNA was isolated from peripheral blood, with consent of subjects included in the study. Analysis of $(CTG)_n$ repeats on normal chromosomes was performed as described (15). The *Alu* polymorphism was typed by the published three-primer protocol (11). The *Hin*fI and the *Taq*I polymorphisms were typed as described (8, 12). DM1 mutation was characterized by small-pool, long-range PCR-based Southern blot analysis (16).

Statistical analyses

Allele frequencies were calculated by gene counting. Heterozygosity value was calculated as $1 - \sum r^2$, where r is the individual allele frequency. Statistical significance was determined using the chi-squared test. Linkage disequilibrium was calculated and a modified chi-squared test was used to assess the significance of linkage disequilibrium for each haplotype (17).

Results

In the sample of 203 normal chromosomes drawn from the Serbian population, the number of CTG repeats ranged from 5 to 28. We observed a trimodal distribution of allele sizes, with peaks at 5, between 9 and 17 and at 20 repeats (Table 1). The $(CTG)_{9-17}$ size class is the most frequent (52.7%), whereas, the $(CTG)_5$ allele is the most common single allele (35.5%). $(CTG)_n$ locus was found to be highly polymorphic with the average heterozygosity value 0.805.

The population was found to be in Hardy–Weinberg equilibrium for the studied DM1 polymorphisms (genotype frequencies for the three polymorphisms are close to predicted Hardy–Weinberg expectations; $P > 0.99$).

Table 1 The frequency of $(CTG)_n$ alleles and associated haplotypes

$(CTG)_n$ alleles	Number of alleles	Frequency (%)	(+++)	(+ − −)	(− − +)	(− − −)
5	72	35.47	72			
8	1	0.49				1
9	1	0.49				1
10	5	2.46	5			
11	17	8.37		1		16
12	28	13.79	3		1	24
13	38	18.72	1		1	36
14	14	6.89	9			5
15	1	0.49				1
16	3	1.48	3			
19	3	1.48	3			
20	7	3.45	7			
21	4	1.97	4			
22	5	2.46	5			
23	1	0.49	1			
25	1	0.49	1			
26	1	0.49	1			
28	1	0.49	1			
Total	203		116	1	2	84

Table 2 Frequency of (*Alu*–*Hin*fI–*Taq*I) haplotypes and association with $(CTG)_n$ alleles

Alu–*Hin*fI–*Taq*I haplotype	Frequency (%)	Frequency of the$(CTG)_n$ associated with a distinct *Alu*–*Hin*fI–*Taq*I haplotype (%)			
		$(CTG)_5$	$(CTG)_{6-8}$	$(CTG)_{9-17}$	$(CTG)_{\geq 18}$
(+++)	57.14	100	0	19.63	100
(+ − −)	0.49	0	0	0.93	0
(− − +)	0.98	0	0	1.87	0
(− − −)	41.38	0	100	77.57	0
n	203	72	1	107	23

n, number of chromosomes.

We detected four out of eight possible haplotypes defined by the three biallelic polymorphisms [*Alu*(+/−), *Hin*fI(+/−), *Taq*I(+/−)] (Table 1). Haplotype frequencies were calculated from three biallelic markers alone, as well as from data on all four polymorphic sites. However, the (+ + +) and (− − −) haplotypes account for >98% of all normal chromosomes studied. Minor (+ − −) and (− − +) haplotypes are present in <2%. On DM1 chromosomes only (+ + +) haplotype was found. Pair-wise linkage disequilibrium (D) between the *Alu*, *Hin*fI, and *Taq*I polymorphisms is highly significant ($P < 0.0001$).

A striking linkage disequilibrium was observed between the CTG-repeat alleles and certain biallelic markers. The (+ + +) haplotype is completely associated with $(CTG)_5$ alleles, and the (− − −) haplotype is strongly associated with $(CTG)_{9-17}$ alleles (78% of all $(CTG)_{9-17}$ alleles) (Table 2). The $(CTG)_{18-35}$ group of alleles was found only in association with the (+ + +) haplotype. Minor (+ − −) haplotype was found only in association with $(CTG)_{11}$ allele, whereas (− − +) haplotype was found in conjunction with $(CTG)_{12}$ and $(CTG)_{13}$ alleles.

Discussion

Observed frequencies of $(CTG)_n$ alleles grouped in size classes in the Serbian population are similar to those published for a mixed European population (5).

The most common DM1 haplotypes (98.52%) in Serbian population are (+ + +) and (− − −). Genetic drift is probably responsible for prevalence and almost equal frequencies of (+ + +) and (− − −) haplotypes in European and Middle Eastern populations. We found that (+ + +) is the most frequent haplotype in our population (57.14%), which is in concordance with data on mixed European population and supports the hypothesis of Tishkoff et al. (5) that (+ + +) is the ancestral haplotype. This hypothesis was drawn from the facts that (+ + +) is the only haplotype found in Primates, and it is the most abundant haplotype in sub-Saharan African populations.

The (+ + +) haplotype is completely associated with $(CTG)_5$ alleles, and the (− − −) haplotype was strongly associated with $(CTG)_{9-17}$ alleles, which is in agreement with previous studies based on *Alu* or *Alu* and *Hin*fI polymorphisms in conjunction with the $(CTG)_n$ polymorphism in Caucasians (10, 12). In this study $(CTG)_{\geq 18}$ alleles were found to be in strong linkage disequilibrium with (+ + +) chromosomes, speaks in favor of the hypothesis that $(CTG)_{\geq 18}$ alleles form a pool of unstable alleles that may further expand (10).

$(CTG)_{9-17}$ alleles are the most frequent in our population as in many other populations and they occur on four different haplotypes (in sub-Saharan Africans, mid-sized alleles are associated with many different haplotypes), speaks in favor of hypothesis that the ancestral (+ + +) haplotype in humans contained a mid-sized repeat (5).

We found DM1-causing alleles only in association with (+ + +) haplotype and our results confirm findings from previous haplotype analyses of DM1 individuals of European descent, which demonstrated complete association between DM1 alleles and (+ + +) haplotype (12, 14). These results indicate a common origin for DM1 mutation in European populations. Occurrence of DM1 mutation on a haplotype other than (+ + +) is very rare, with only a few reported cases (18).

Haplotype studies in a large number of populations suggested that $(CTG)_{18-35}$/+ + + is the plausible ancestral chromosome, on which low-frequent mutations gradually lead to full DM1 mutation (5, 12). A complete association of (+ + +) haplotype with the DM1 mutation, as well as a strong association of (+ + +) haplotype with $(CTG)_{\geq 18}$ alleles in the Serbian population, confirms this hypothesis. An explanation for the existence of one haplotype associated with DM1 mutation could be that there was not enough time for mutations and recombinations to occur, which could cause variations in the DM1 mutation's haplotype (5, 10, 12, 14). Another explanation is that it could represent the rare haplotype which may predispose CTG repeats to expand. According to the ori-shift model for expansion of trinucleotide repeats (which considers a change in relative orientation or distance between trinucleotide repeats and its replication origin as the primary event that puts a repeat into an expansion-prone position) (19). *Alu* insertion might change the distance between the CTG repeats and its replication origin, therefore expanding.

Acknowledgements

This work was supported by the Serbian Ministry of Science (grant no. 1521).

References

1. Brook JD, McCurrach ME, Harley HG et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein family member. Cell 1992;**68**:799–808.
2. Fu Y-H, Pizzuti A, Fenwick RG et al. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 1992;**255**:1256–8.
3. Mahadevan M, Tsilfidis C, Sabourin L et al. Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. Science 1992;**255**:1253–5.
4. The International Myotonic Dystrophy Consortium (IDMC). New nomenclature and DNA testing guidelines for myotonic dystrophy type 1 (DM1). Neurology 2000;**54**: 1218–21.
5. Tishkoff SA, Goldman A, Calafell F et al. 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 1998;**62**:1389–402.
6. Yamagata H, Miki T, Ogihara T et al. Expansion of unstable DNA region in Japanese myotonic dystrophy patients. Lancet 1992;**339**:692.
7. Zerylnick C, Torroni A, Sherman SL, Warren ST. Normal variation at the myotonic dystrophy locus in global human populations. Am J Hum Genet 1995;**56**:123–30.
8. Goldman A, Ramsay M, Jenkins T. New founder haplotypes at the myotonic dystrophy locus in Southern Africa. Am J Hum Genet 1995;**56**:1373–8.
9. Deka R, Majumder PP, Shriver MD et al. Distribution and evolution of CTG repeats in the myotonin protein kinase gene in human populations. Genome Res 1996;**6**:142–54.
10. Imbert G, Kretz C, Johnson K, Mandel J-L. Origin of the expansion mutation in myotonic dystrophy. Nat Genet 1993;**4**:72–6.
11. Mahadevan MS, Foitzik MA, Surh LC, Korneluk RG. Characterization and polymerase chain reaction (PCR) detection of an Alu deletion polymorphism in total linkage disequilibrium with myotonic dystrophy. Genomics 1993;**15**:446–8.
12. Neville CE, Mahadevan MS, Barceló JM, Korneluk RG. High resolution genetic analysis suggests one predisposing haplotype for the origin of the myotonic dystrophy mutation. Hum Mol Genet 1994;**3**:45–51.
13. Harley HG, Brook JD, Rundle SA et al. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 1992;**355**:545–6.
14. Goldman A, Krause A, Ramsay M, Jenkins T. Founder effect and the prevalence of myotonic dystrophy in South Africans: molecular studies. Am J Hum Genet 1996;**59**:445–52.
15. Čuljković B, Stojković O, Vukosavić S et al. CTG repeat polymorphism in DMPK gene in healthy Yugoslav population. Acta Neurol Scand 2002;**105**:55–8.
16. Savić D, Rakočević-Stojanović V, Keckarević D et al. 250 CTG repeats in DMPK is a threshold for correlation of expansion size and age at onset of juvenile-adult DM1. Hum Mutat 2002;**19**:131–9.
17. Weir BS. Genetic data analysis II. Sunderland, MA: Sinauer, 1996.
18. Basu P, Gangopadhaya PK, Mukherjee SC, Das SK, Sinha KK, Bhattacharyya NP. Molecular anatomy of CTG repeat expansion in myotonin protein kinase gene among myotonic dystrophy patients from eastern India. Hum Mutat 2000;**16**:372.
19. Cleary JD, Nichol K, Wang YH, Pearson CE. Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nat Genet 2002;**31**:37–46.