European Journal of Human Genetics (2001) 9, 829–835

www.nature.com/ejhg

ARTICLE

Polymorphism of trinucleotide repeats in loci DM, DRPLA and SCA1 in East European populations

Svetlana N Popova*[,1], PA Slominsky[1], EA Pocheshnova[2], EV Balanovskaya[2], LA Tarskaya[3], NA Bebyakova[4], LV Bets[5], VP Ivanov[6], LA Livshits[7], EK Khusnutdinova[8,9], VA Spitcyn[2] and SA Limborska[1]

[1]Institute of Molecular Genetics RAS, Kurchatov sq.2, Moscow, Russia; [2]Medical and Genetic Scientific Center, Moscow, Russia; [3]Institute of Health, Yakutsk, Russia; [4]Archangelsk State Medical Academy, Archangelsk, Russia; [5]Moscow State University, Moscow, Russia; [6]Kursk State Medical Academy, Kursk, Russia; [7]Institute of Molecular Biology, Kiev, Ukraine; [8]Institute of Ethnography, Minsk, Republic of Belarus; [9]Biochemistry and Genetics, Ufa, Russia

A normal polymorphism at three triplet repeat loci (myotonic dystrophy (DM), dentatorubral-pallidoluysian atrophy (DRPLA) and spinocerebellar ataxia type 1 (SCA1)) were examined in healthy unrelated individuals from the Siberian Yakut (Mongoloid) population, the Adygei (Caucasian) population and nine East European populations: populations from Russia (Holmogory, Oshevensk, Kursk, Novgorod, Udmurts, Bashkir), two Ukrainian populations (Lviv and Alchevsk) and one Belarussian. The distribution of alleles for DRPLA and SCA1 were similar for all East-European populations. For the DM locus, East European populations had typical allele distribution profiles with two modes, $(CTG)_5$ and $(CTG)_{11-14}$, but some differences were found for the Bashkir population where alleles containing 11–14 CTG repeats had relatively higher frequency. The Yakut population had different allele spectra for all types of repeats studied. Higher heterozygosity levels and insignificant differences between expected and observed heterozygosity were found for all tested loci. The latter led us to suggest that the trinucleotide repeat loci analysed are not influenced by selection factors and could be useful for genetic relationship investigations in different populations. *European Journal of Human Genetics* (2001) **9**, 829–835.

Keywords: short tandem repeat polymorphism; population genetics; DM; DRPLA; SCA1 loci

Introduction

The human genome contains a large amount of highly repetitive DNA sequences including the short tandem repeats. Several characteristics make them useful in comparing human population genetics with classical polymorphism markers, such as blood group or protein markers. However, of special concern within the huge number of short tandem repeat loci are those gene sequences in which pathologically expanded trinucleotide repeats cause hereditary diseases.[1]

***Correspondence: SN Popova, Institute of Molecular Genetics RAS, Kurchatov sq.2, Moscow, Russia. Tel: +7 095 1960210; Fax: +7 095 1960221; E-mail: popova@img.ras.ru**
Received 11 May 2001; revised 12 July 2001; accepted 3 August 2001

The expansion of trinucleotide repeats in humans has been associated with several human genetic diseases, such as myotonic dystrophy, dentatorubral pallidoluysian atrophy, fragile X syndromes, and several types of spinocerebellar ataxias.[2–5] All these triplet repeat loci are highly polymorphic and their length can vary significantly in healthy populations. Myotonic dystrophy (DM) is associated with the expansion of a CTG repeat in the 3′ UTR of the myotonin protein kinase gene. In the case of healthy individuals the copy number varies from 5 to 37 CTG repeat, while in DM patient chromosomes the copy number is more than 50 and usually as high as several thousand repeats in severely affected individuals.[6] The myotonic dystrophy frequency in the Global population can vary from 1 in 475 in certain regions of Canada to 1 in 8000 in Western Europe.[7]

Dentatorubral pallidoluysian atrophy (DRPLA) and spinocerebellar ataxia type 1 (SCA1) are caused by an amplified CAG repeat within the coding sequence of the gene, resulting in the lengthening of a polyglutamine tract in the corresponding protein.[8,9] These disease related trinucleotide repeats are now widely used in the investigations of the history and diversity of humans both in the Global population and in individual population groups. Higher trinucleotide repeat polymorphism degree was found for populations with different origins and ethnicity. Recent data show that such polymorphisms can be useful for genetic population studies and genome mapping, especially for determination of population origin. Moreover, the allele-size distribution can be used for interpopulation distance calculation.[2,7,10,11]

Taking the above into account, we approached the investigation of population relationships in Russia and allele polymorphisms in disorders with a triplet repeat expansion. In this study, three loci with different triplet repeat polymorphisms were investigated (CTG-repeat in myotonic dystrophy locus and CAG-repeats in SCA1 and DRPLA loci) in normal individuals of 11 human populations from separate ethnic groups (four Russian, two Ukrainian, one Belorussian, one Bashkir, one Udmurt, one Adygei and one Yakut) living in the Russian Federation and ex-USSR territory.

Materials and methods

DNA samples for study were obtained from anonymous random donors of 11 populations from different ethnic groups with different levels of Mongoloid and Caucasoid components (Figure 1).

Two population groups originate from the Turkic linguistic family: these are the Siberian Yakut and Bashkir populations. The Yakut population is smaller in number compared with original Siberian peoples. In classical anthropology they belong to the Central-Asian anthropologic type. From a historical and ethnographic point of view, three groups are isolated in Yakut ethnicity: Central, Viluy and Yakut-nomads. In this study, samples from the Central group of Yakut people were examined. Based on historical events we consider the Baskhir population as one with a combination of Caucasoid and Mongoloid components. Also in this study, samples from different regions of Bashkiria were examined. The East Slavonic linguistic family were represented by the Kursk, Novgorod, Oshevensk and Holmogory of Russia, Lviv and Alchevsk of Ukraine and mixed Belarussian populations. The Oshevensk and Holmogory are isolated populations from the Archangelsk region of North Russia where external influences are very low, the Novgorod are a North-Western Russian population, and the Kursk a South Russian population. The Lviv and Alchevsk are two populations originating from West Ukraine (Lviv) and East Ukraine near the Russian border (Alchevsk). The Finno-Ugric linguistic family was represented by the Udmurtes from Izhevsk. The Adygei population came from Adyg-Shabsug and represented North Caucasian linguistic family. Blood samples were collected after individual's informed consent according to following criteria: all individuals belong to the native population of the regions studied (with at least three female generations living in the region); all individuals are non-related and healthy. DNA was extracted from peripheral leukocytes as described previously.[12] Analyses were made based on 50 to 250 people for each population.

For determination of allele size we performed polymerase chain reaction (PCR) amplification of each target sequence and denaturing PAAG electrophoresis for the separating of PCR products. The PCR amplification was performed using a Perkin Elmer Cetus thermal cycler. The reaction mixture contained 0.1–0.2 mg of genomic DNA, 2 pmol of each primer, 0.2 mM each dNTPs, 0.5 u Taq-polymerase (Sileks, Russia) and polymerase buffer (50 Tris-HCl, pH 8.8, 15 mM $(NH_4)_2SO_4$, 5 mM $MgC1_2$, 0.01% gelatin) in 20 μl of total

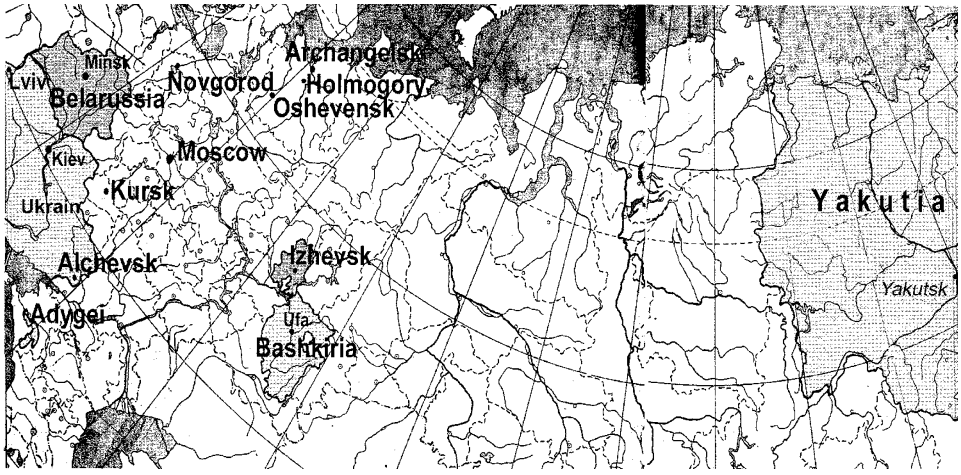

Figure 1 Geographical location of populations tested.

volume. One of the primers was labelled by ^{33}P-γATP before PCR. Labelled reaction was performed in 20 μl buffer (2 μl 10 × kinase buffer, 50 pmol of appropriate primer, 50 μCi ^{33}P-γATP, 2.5 u of T4 polynucleotide kinase (Sileks, Russia)). The cycling conditions included a denaturation at 95°C for 5 min, followed by 30 cycles of appropriate programme for each locus and the procedure was completed by a final extension step at 72°C for 10 min. The programme conditions and primers used were the same as those described previously.[6–8]

The mixture, containing 2 μl of amplification product and 2 μl loading buffer (0.01% bromphenol blue, 0.01% xylene cyanol, 98% formamide), was denatured at 95°C for 5 min before electrophoresis. The exact size for trinucleotide repeats was determined using 6% denaturing PAAG electrophoresis with coelectrophoresis of an M13 sequencing ladder.

Results

The polymorphic CAG and CTG repeat were examined by PCR in 11 human populations. The number of chromosomes typed in each population, as well as expected and observed heterozygosity are shown in Table 1. Each population was typed for three loci. For technical reasons, not all individuals were typed in each population for all three loci, because some DNA samples didn't amplify sufficiently. The modal number of repeats and the number of different alleles were used as a comparison between populations. Triplet repeat length analysis showed some differences in allele spectra between Caucasoid and Mongoloid populations. For all investigated populations, higher heterozygosity level was observed (from 65 to 84%).

Twenty-six allelic variants with repeat numbers from 5 to 32 CTG were found in DM locus studies (Table 2). For all populations tested, allele spectra had bimodal character with modes at the $(CTG)_5$ and $(CTG)_{11-14}$ triplet repeats, coinciding with previously mentioned data from other investigations for this locus.[11,13,14] Comparison of allele spectra for populations of different ethnic origin shows a decrease of $(CTG)_5$ allele frequencies (from 42% in Kursk to 7% in Yakut) and an increase of $(CTG)_{11-14}$ allele frequency (from 45% in Adygei to 85% in Yakut). The Bashkir population has an intermediate position: the $(CTG)_5$ allele frequency is 28.7% and the $(CTG)_{11-14}$ is 57.6%. A marked reduction of the 5-repeat allele frequency was found for the Yakut population (7%). In this population the majority (89.4%) of alleles are $(CTG)_{11-15}$. When comparing the Yakut with the Slavonic (Kursk) population the difference in frequency of the $(CTG)_5$ and $(CTG)_{11-14}$ alleles was highly significant ($P<0.001$). Comparison for allele distribution carried out with RxC test[15] showed statistically significant differences between the Kursk population (East Slavonic group) and non-Slavonic populations (Yakut and Bashkir) (Table 3).

For the SCA1 locus, 19 alleles were found (from 8 to 14 in different populations), with each allele containing from 20 to 42 CAG trinucleotide repeats (Table 4). For the majority of samples investigated, the $(CAG)_{30}$ allele was the most common, with frequency varying from 21% to 39% in the

Table 1 Number of chromosomes analysed, expected and observed heterozygosity and allele spectra parameters for investigated populations in three loci

							Russians			*Ukrainians*	
Locus	*Yakut*	*Adygei*	*Belarussian*	*Udmurt*	*Bashkir*	*Holmogory*	*Oshevensk*	*Kursk*	*Novgorod*	*Lviv*	*Alchevsk*
DRPLA											
No. of chromosomes	180	560	632	90	280	98	154	462	228	110	112
Heterozygosity:											
expected	0.82	0.82	0.77	0.85	0.8	0.71	0.65	0.77	0.75	0.78	0.76
observed	0.77	0.84	0.81	0.76	0.84	0.65	0.7	0.73	0.76	0.78	0.66
Mode(s)	8, 13, 17	8, 12, 15	8, 15	10, 15	8, 10, 15	10, 15	10, 15	10, 15	8, 10, 15	8, 15	8, 10, 15
No. of alleles	12	13	17	11	14	11	11	14	15	10	11
SCA1											
No. of chromosomes	180	538	622	92	196	96	148	464	220	100	112
Heterozygosity:											
expected	0.72	0.72	0.76	0.78	0.77	0.83	0.78	0.78	0.77	0.78	0.77
observed	0.64	0.76	0.75	0.67	0.75	0.79	0.8	0.77	0.83	0.72	0.7
Mode(s)	28, 30	30	30	30	30	30	30	30, 32	30, 32	29, 32	30
No. of alleles	9	13	14	10	14	14	9	14	12	14	8
DM											
No. of chromosomes	170	558	674	92	308	98	156	448	228	136	112
Heterozygosity:											
expected	0.73	0.8	0.77	0.83	0.83	0.81	0.75	0.76	0.8	0.78	0.82
observed	0.72	0.83	0.75	0.8	0.8	0.78	0.68	0.79	0.87	0.68	0.82
Mode(s)	5, 13	5, 11, 13	5, 11, 13	5, 12, 14	5, 13	5, 11, 13	5, 11, 13	5, 13	5, 13	5, 12	5, 13
No. of alleles	10	21	20	10	15	11	12	22	16	12	15

Table 2 Allele variants frequencies for (CTG)n repeat in miotonin proteinkinase gene

(CTG)n	*5*	*6*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*
Yakut	7					0.6	14.7	39	31	4.1								1	1		1	0.6					
Adygei	37	0.2	0.6	0.2	0.2	15.6	10	13.7	5.7	1.7	1	0.5	0.3	0.9	4.7	3.3	0.9	1.2	0.5		1.4		0.2			0.2	
Belarussian	38.5			0.4		0.5	13.4	11.5	18.5	9.5		0.9			0.75	1.2	2	1	0.75	0.25	0.6		0.2		0.1		
Udmut	28					13	19	18	14			1			2	2					2						1
Bashkir	28.7		0.2	0.9		11.7	15	18	12.6	1.8	0.4	0.4		0.2	2	3.6	1.8	1		0.9	0.2	0.2	0.4				
Holmogory	32				1	18.4	12.2	21.4	8		1.3			2	1	2	1	1									
Oshevensk	41.7				1.3	11.5	9	22.4	8.3					0.6	1.3	1.3	1.3										
Kursk	41.9	0.3	0.5	0.2	0.3	10.1	13.6	15.9	6.8	2	1		0.2	0.6	1.3	2.3	0.3	0.7	0.4	0.2	0.9		0.3	0.2			
Novgorod	32			0.4			21.5	12	17.5	10	0.4	0.4			0.4	1.3	2	0.4		0.4	0.4			0.4		0.4	
Lviv	36			1			16	21	13	6		1.3		1	1.25	1.25						1	1.25				
Alchevsk	28			1.8		1.8	13	19	22.3	6					0.9	0.9	0.9	1.8		0.9	0.9	0.9				0.9	

Table 3 Comparison non-Slavonic populations with Kursk using RxC programme

Population	*χ-squared* χ^2	P	*G-statistic* *G*	P
DRPLA				
Yakut	255.8994	0.0000	276.9733	0.0000
Bashkir	43.4669	0.0000	45.8159	0.0000
SCA1				
Yakut	141.0113	0.0000	138.2583	0.0000
Bashkir	45.5670	0.0000	52.7617	0.0000
DM				
Yakut	160.1182	0.0000	180.3309	0.0000
Bashkir	49.7899	0.0000	59.2382	0.0000

Lviv and Adygei populations, respectively. For this locus, more than 80% of all Caucasoid populations had allele $(CAG)_{29}$ or alleles with more copies of the repeat with modes in $(CAG)_{30}$ and $(CAG)_{32}$. In the Yakut population, alleles with 28 and 30 repeats had the highest frequency and 86% of allele spectrum were shifted by 2–3 alleles to the left from spectra of other populations. Statistical analysis showed significant differences in allele distribution between the groups of different origin.

Twenty-one distinct alleles with a size range from 6 to 27 triplet repeats, were found for the DRPLA locus (Table 5). Allele spectra had bimodal character with modes in 8–10 and 13–16 CAG repeats for Caucasoid populations, while the Yakut population demonstrated trimodal spectrum with modes at 8, 13 and 17 CAG repeats: the frequency of the $(CAG)_8$, $(CAG)_{13}$ and $(CAG)_{17}$ alleles were under-represented, whereas 10 and 15 were over-represented in the Yakut compared with other groups. For the Yakut population, allele sizes were shifted to a shorter range compared with allele distribution of Caucasoid populations: 66% of Yakut alleles contained from 6 to 13 CAG repeats, whereas in other populations 40–60% of allele spectra consisted of $(CAG)_{13-15}$ alleles.

Genetic diversity for all populations was estimated with F-statistic using the jackknifing method.[16] These results are presented in Table 6. The Fst for different populations varies from 0.016 to 0.041 for the DRPLA locus, from 0.005 to 0.21 for the SCA1 locus and from 0.005 to 0.024 for the MD locus. Fst quantity was lowest for the Yakut population in comparison with all populations in all loci analysed.

For population relationship analysis we have constructed a genetic tree (using GDA computer program) based on three loci for the populations investigated. The dendrogram for genetic relations is presented in Figure 2. One can see that the Yakut population was disposed some way from the others, which formed a cluster of Caucasoid populations having their own hierarchy. Thus, two smaller clusters are obtained: the first formed by the Lviv and Adygei populations, and the second only with East Slavonic populations.

Table 4 Allele variants frequencies for (CAG)n repeat in SCA1 gene

(CAG)n	*20*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*37*	*40*	*42*
Yakut						1.7	0.6	43	19.5	23	1	8	1			1.7		0.5	
Adygei						2	0.7	5.7	33.7	39.2	9	5.4	3	0.7	0.4			0	0.2
Belarussian	0.2		0.1			0.9	2.9	4.4	29	36.3	10.6	8.8	3.7	0.4	1.8	0.9			
Udmurt						3	1	7	26	35	16	8	2			1	1		
Bashkir					0.4	2.2	0.3	15	26.2	38.6	10.2	3.9	2.6		0.3	0.3			
Holmogory			1	1		1	1	13	18.8	30	14	10	4.2	2	2	1	1		
Oshevensk					0.7		0.7	13.5	29	30.4	7.6	14	3.4			0.7			
Kursk		0.2		0.2	0.2		2.2	7.5	30	31	9.6	13	3	1	1.6	0.2			
Novgorod					2.3	1	2.3	4	27	35	10	13	0.5	2.3	1.4	1			
Lviv			3			1	1	4	39	21.4	12.3	13	2	1	1	1			
Alchevsk							1	11	25	38	11	7	4		3				

Table 5 Allele variants frequencies for (CAG)n repeat in DRPLA gene

(CAG)n	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*27*
Yakut			13.6		14.2		4		3	35.2	19	4	3	1.7	1.7				0.6		
Adygei		0.2	13.6		4.7		8	2	6.7	33.4	16.4	8.3	4.3	2	0.2	0.2					
Belarussian		0.3	10.2		14.2	0.3	2.9	1	2	39.8	21	3.3	1	1.7	1.4	0.2	0.2	0.54			
Udmurt	2.2		12.2		9	1		17	22.2	24.4	10		1					1			
Bashkir		0.3	7.8	0.3	26.6		3	0.3	2.4	30.7	13.4	7.5	4.4	1.5	1.2		0.3				0.27
Holmogory			5.6		17		1	1		47	19.4	1	4	2	1		1				
Oshevensk			6		17		2.6	0.6	1.3	55	12.4	0.6	0.6	2.6	1.3						
Kursk			13.6		14.2		4		3	35.2	19	4	3	1.7	1.7				0.6		
Novgorod			15	0.4	9.6		2.2	0.9	2.6	42.1	19	3.9	1.3	0.9	0.4		0.4	0.9		0.4	
Lviv			15		14		0.9	0.9	4.5	36	22	2	2.7	2							
Alchevsk			12.5		12.5		3	1	3	41	20	2	2	2			1				

Table 6 Fst amount for DRPLA, SCA1 and DM loci for investigated populations

Population	Fst DRPLA	Fst SCA1	Fst DM
Yakut	0.015988	0.005679	0.004840
Adygei	0.039045	0.020365	0.023070
Udmurt	0.032792	0.019784	0.020724
Bashkir	0.034046	0.020883	0.022592
Russian			
Holmogory	0.034700	0.019535	0.020972
Oshevensk	0.032553	0.020188	0.021026
Kursk	0.040651	0.022662	0.021988
Novgorod	0.037014	0.020487	0.021294
Ukrainians			
Lviv	0.035877	0.018410	0.020610
Alchevsk	0.035825	0.020127	0.020719

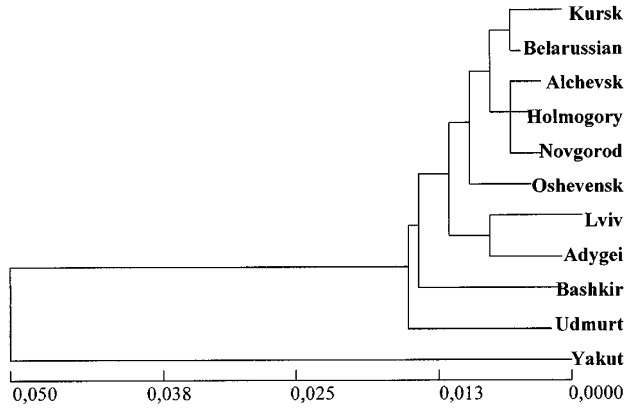

Figure 2 Dendrogram for genetic relationship for investigated populations constructed using GDA-program.

Discussion

We tested populations with different levels of Mongoloid and Caucasoid components, belonging to different linguistic families – East Slavonic, Turkic, Finno-Ugric, and North Caucasian – and of normal triplet polymorphisms of DM, DRPLA and SCA1 loci. As discovered recently,[1,10] these loci are highly polymorphic in healthy individuals and significant diversity exists for allele spectra in groups with different organization. As expected, similar results were observed in our research. For all three loci, the most remarkable differences in allele distribution were found for the Yakut population, which derives from the Turkic linguistic group and has Mongoloid ancestry. Significant divergence by the Yakut population from other population groups was also shown by the genetic diversity test. Fst index for all loci investigated in this population was 2 – 3 times lower than in other populations. The data suggest more homogeneity of this group and close relationship between individuals in it. Genetic analysis carried out with the GDA program[16] demonstrates that similar polymorphic loci can be used for genetic relationship investigation both in the Global population and small intrapopulation groups. In the dendrogram constructed, almost all East Slavonic populations form a single cluster, with only the Lviv population lying closer to some non-Slavonic groups. This population originates from West Ukraine and it is likely that some Central European influences (Polish, Hungarian, Slovac) might have changed its gene pool. The Kursk and Belorussian populations, situated in the South and West regions analysed, take closest positions in the tree and are possibly most similar in genetic variability. Disposition of the Alchevsk population between two North Russian populations (Holmogory and Novgorod) might be evidence of more complexity in East Slavonic population history. This Ukrainian population derives from East Ukraine and lies near the Russian border. Future studies of Russian and Ukrainian genetic variability could provide new insight into the origination of and relationships between these East Slavonic groups.

The DM locus had the highest polymorphism and we have found the greatest number of alleles in it (26 in comparison with 21 and 19 for DPLA and SCA1 correspondingly). Moreover, this locus, in our opinion, is the most convenient for population analysis. Due to a distinctive allele spectrum with two noted modes in 5 and 11 – 14 CTG repeats, the smallest differences in population structure are easily definable and visible in the diagram.

Trinucleotide repeats show high variability among normal individuals and are therefore increasingly used to investigate genetic relationships between different ethnic groups as well as traditional genetic markers. Our study provides new data about trinucleotide repeat polymorphism in Russia and its nearest populations; this information might be useful for analysis of human history and population evolution.

Acknowledgments

We would like to thank Drs AI Mikulich, AV Stepanova, MA Solodilova and SA Kravchenko for providing the blood samples. This study was supported in part by grants from the Russian State Program 'Frontiers in Genetics', 'Human Genome', and the Russian Basic Research Foundation, grant no. 99-04-27515.

References

1 Caskey CT, Pizutty A, Fu YH, Fenwick Jr RG, Nelson DL: Triplet repeat mutation in human disease. *Science* 1992; **256**: 784 – 789.
2 Richards RI, Crawford J, Narahara K *et al*: Dynamic mutation loci: allele distribution in different populations. *Ann Hum Genet* 1996; **60**: 391 – 400.
3 Koide R, Ikeuchi T, Onodera O *et al*: Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nat Genet* 1994; **6**: 9 – 13.
4 Gusella JF, MacDonald ME: Trinucleotide instability: a repeating theme in human inherited disorders. *Annu Rev Med* 1996; **47**: 201 – 209.
5 Nelson DL: Six human genetic disorders involving mutant trinucleotide repeats. *Genome Analysis* 1993; **7**: 1 – 24.

6 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 kinase family member. *Cell* 1992; 68: 799 – 808.
7 Zerilnick 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 – 130.
8 Orr HT, Chung MY, Banfi S *et al*: Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* 1993; **4**: 221 – 226.
9 Nagafuchi S, Yanagisawa H, Sato K *et al*: Dentatorubral and palladoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nat Genet* 1994; **6**: 14 – 18.
10 Gyurus JM, Molnar J, Melegh B *et al*: Trinucleotide repeat polymorphism at five disease loci in mixed Hungarian population. *Am J Med Genet* 1999; **87**: 245 – 250.
11 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 the myotonic dystrophy mutation. *Am J Hum Genet* 1998; **62**: 1389 – 1402.
12 Miller SA, Dykes DD, Polesky HF: A sample salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acid Res* 1988; **16**: 1215.
13 Popova SN, Mikulich AI, Slominski PA, Shadrina MI, Pomasanova MA, Limborska SA: Polymorphism of the (CTG)n repeat in the myotonin protein kinase (DM) gene in Belarussian populations: analysis of interethnic heterogeneity. (In Russian). *Genetika* 1999; **35**: 994 – 997.
14 Ashizawa T, Epstein HF: Ethnic distribution of myotonic dystrophy gene. *Lancet* 1991; **338**: 642 – 643.
15 Yeh FC, Yang R-C: POPGENE version 1.21. 1997. http://www.ualberta.ca/~fyeh
16 Lewis PO, Zaykin D: Genetic Data Analysis Software (GDA) version 1d15.http://lewis.eeb.uconn.edu/lewishome/gda.html