Low Variability of the POLG (CAG)$_n$ Repeat in North Eurasian Populations

BORIS A. MALYARCHUK,[1] MARTA PAPUGA,[2] TOMASZ GRZYBOWSKI,[2] IGOR B. ROGOZIN,[3] MARCIN WOZNIAK,[2] MIROSLAVA V. DERENKO,[1] SERGE YU. RYCHKOV,[4] JAKUB CZARNY,[2] ILIA A. ZAKHAROV,[4] AND DANUTA MIŚCICKA-ŚLIWKA[2]

Abstract We investigated the frequency of different repeat-length alleles of the trinucleotide CAG microsatellite repeat in the coding sequence of the nuclear gene for the catalytic subunit of mitochondrial DNA polymerase gamma (POLG) in 12 ethnic groups from northern Eurasia. The population sample consisted of 1,330 individuals from 3 large geographic areas: Europe, Southwest Asia, and Siberia/East Asia. We found that the 10-repeat allele of the POLG gene is the most frequent in all analyzed populations, with a frequency of 88–96%. The heterozygosity level ranges from 22% in Europe to 13.6% in Southwest Asia with the lowest value of 7.4% in Siberia/East Asia. The present study provides evidence of clinal distribution of POLG gene heterozygosity in North Eurasian populations. In general, we found an extremely low variability of the trinucleotide CAG microsatellite repeat, suggesting that purifying selection acts against deleterious alleles, although low mutability of the repeated region cannot be ruled out.

The last decade revealed rapid progress in the investigation of population variability of the maternally inherited mitochondrial DNA (mtDNA). The accumulation of mutations (deletions, insertions, nucleotide substitutions) in human mtDNA was suggestive of a defect of nuclear genes responsible for mtDNA replication and maintenance (Copeland et al. 2003). Among these genes, mtDNA polymerase gamma (POLG; MIM 174763) has been identified as an enzyme directly involved in induction of mtDNA mutations (Longley et al. 2001; Van Goethem et al. 2001; Ponamarev et al. 2002; Del Bo et al. 2003).

[1]Institute of Biological Problems of the North, Far-East Branch of the Russian Academy of Sciences, Magadan 685000, Russia.

[2]Ludwik Rydygier Medical University, Forensic Medicine Institute, 85-094 Bydgoszcz, Poland.

[3]Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia.

[4]Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow 119991, Russia.

Human Biology, June 2005, v. 77, no. 3, pp. 355–365.

KEY WORDS: MITOCHONDRIAL DNA MUTATIONS, POLYMERASE GAMMA (POLG), HUMAN POPULATIONS, GENETIC DIVERSITY, NORTHERN EURASIAN ETHNIC GROUPS, RUSSIANS, POLES, BOSNIANS, PERSIANS, KURDS, TAJIKS, BURYATS, SHORS, ALTAIANS, TUVINIANS, KOREANS, MONGOLIANS, SIBERIANS.

Lestienne (1987) provided the first evidence of a role of POLG in the replication of human mtDNA, and Ropp and Copeland (1996) cloned the human POLG genes. Zullo et al. (1997) mapped the POLG gene to 15q24–q26, and Walker et al. (1997) mapped this gene to 15q25. Most known POLG gene mutations appear to be associated with multiple mtDNA deletions as well as with nucleotide substitutions, revealed in patients with neuromuscular diseases, such as progressive external ophthalmoplegia (PEO), mitochondrial neurogastrointestinal encephalomyopathy (MNGIE), and sensory ataxic neuropathy, dysarthria, and ophthalmoparesis (SANDO) syndrome (Chen et al. 2002; Lamantea et al. 2002; Ponamarev et al. 2002; Di Fonzo et al. 2003; Van Goethem et al. 2003). Recently, Del Bo et al. (2003) showed that patients with PEO and MNGIE characterized by mutations in the exonuclease domain of the POLG gene have the highest frequency of individually rare point mutations only in the mtDNA control region.

Ropp and Copeland (1996) found a potentially unstable CAG repeat in the second exon of the catalytic subunit of mtDNA polymerase. Rovio et al. (1999) subsequently investigated the frequency of different repeat-length alleles in populations of diseased and healthy individuals. They demonstrated that the common POLG allele of 10 CAG repeats is found in both Finnish and ethnically mixed population samples at a uniformly high frequency (approximately 88%), with homozygosity close to the equilibrium prediction. The number of repeats was found to vary between 5 and 14, with a mean value of 10 (Rovio et al. 1999, 2004). The number of CAG repeats was not associated with inherited mtDNA diseases, although a correlation between infertility in men and the absence of the common 10-CAG repeat was observed (Rovio et al. 2001). In a study of the variation in the number of CAG repeats in the POLG gene in the Danish population, Jensen et al. (2004) found that the absence of one or both 10-CAG alleles was more frequent among subfertile patients than among fertile control subjects; however, this finding was not confirmed in a study of infertile patients with different sperm counts and a control group of normospermic men of Italian origin (Michielotto et al. 2003; Krausz et al. 2004).

Therefore several studies of the variation in the number of CAG repeats in the POLG gene have shown that the most frequent alleles in the Finnish, Danish, Dutch, German, Scottish, English, Italian, and French samples of healthy males are 10-CAG and 11-CAG repeats, with a high prevalence of 10-CAG repeats ($\cong$90%) (Rovio et al. 1999, 2004; Chen et al. 2002; Jensen et al. 2004; Krausz et al. 2004; Aknin-Seifer et al. 2005). However, the frequency of the 10-CAG allele was only 0.81 in normal individuals of Italian origin (Krausz et al. 2004). This result raises a question about the distribution of POLG alleles and genotypes in different human populations. To address this question, we have studied POLG CAG variation in a broad range of North Eurasian populations.

Materials and Methods

Population Samples. We analyzed 1,330 individuals from 3 different geographic regions: 760 individuals from Europe, 166 from Southwest Asia, and

404 from Siberia/East Asia. Three ethnic groups—Russians ($n = 619$), Poles ($n = 102$), and Bosnians ($n = 39$)—represented the European sample. The sample from Southwest Asia was made up of three ethnic groups: Persians ($n = 97$) and Kurds ($n = 22$) from eastern Iran and Tajiks ($n = 47$) from Tajikistan. Siberian samples were collected mainly in the Altai-Sayan and Baikal region and included native peoples of this region: Buryats ($n = 102$), Shors ($n = 33$), Altaians ($n = 80$), and Tuvinians ($n = 42$). The population sample from East Asia included two ethnic groups living in Korea (Koreans, $n = 103$) and Mongolia (Mongolians, $n = 44$). The study of Russians ($n = 619$) from different regions of the European part of Russia was performed on the subpopulation level. Samples of Russians were collected in the southern (Stavropol, Belgorod, and Saratov), central (Kaluga, Tula, Vladimir, and Yaroslavl), and northwestern (Pskov, Velikij Novgorod, and Volot) parts of European Russia. Appropriate informed consent was obtained from all participants in this study.

CAG Variability Analysis. Genomic DNA was prepared from either blood (fresh or dried on filter paper) or hair roots by means of cell lysis in the presence of proteinase K and 1% SDS, followed by phenol/chloroform extractions. All samples were screened for CAG-repeat polymorphism in the second exon of the POLG gene through a fluorescent PCR with primers H43-FAM (5′-AGC GAC GGG CAG CGG CGG CGG CA-3′) and H42 (5′-CCC TCC GAG GAT AGC ACT TGC GGC-3′), matching positions 388–410 and 456–479, respectively (according to the POLG mRNA sequence with GenBank accession number X98093) (Davis et al. 1996).

PCR amplification was performed using 10–20 ng of genomic DNA in a 50-μl reaction volume containing 1× Promega buffer, 1.5 m*M* $MgCl_2$, 200 μ*M* of each dNTP (Promega Madison, Wisconsin), 0.5 μ*M* of each primer, and 1 U of *Taq* polymerase (Promega). After a hot start at 95°C for 2 min, samples were amplified for 30 cycles for 30 s at 94°C, 1 min at 60°C, and 30 s at 72°C (GeneAmp PCR System 9700, Applied Biosystems, Foster City, California). The PCR products were separated in a 6% denaturing polyacrylamide gel and sized on an ABI 377 DNA Sequencer (Applied Biosystems) using size standard CXR 60-400 (Promega). The fragment sizes were analyzed using GeneScan, version 3.1 (Applied Biosystems).

Statistical Analysis. Microsatellite allele frequencies were estimated by direct gene counting. Statistical differences between frequencies of CAG alleles and genotypes in populations were analyzed by a *t*-test, as implemented in the Statistica package (StatSoft Inc., Tulsa, Oklahoma). Gene diversity *h*, equivalent to the heterozygosity level for diploid data, was calculated using Nei's (1987) equation, as implemented in the Arlequin package (Schneider et al. 2000).

Genetic differentiation between populations was estimated by means of distance methods based on the number of different alleles of the microsatellite (F_{ST}) and the sum of the squared number of repeat differences between two alleles (R_{ST}). The measure R_{ST} differs from F_{ST} in taking explicit account of the

Table 1. Frequencies of the POLG CAG Alleles in North Eurasian Populations

	CAG Allele							
Population (2N)	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Koreans (206)	0	0	0	0	202	3	0	1
Mongolians (88)	0	0	0	1	83	4	0	0
Shors (66)	0	0	0	0	64	1	1	0
Altaians (160)	0	0	0	0	151	9	0	0
Tuvinians (84)	0	0	0	0	82	1	1	0
Buryats (204)	0	0	0	0	195	9	0	0
Tajiks (94)	0	0	0	1	87	6	0	0
Persians (194)	0	0	0	1	181	11	1	0
Kurds (44)	0	0	1	0	40	2	1	0
Bosnians (78)	0	0	1	0	74	3	0	0
Poles (204)	0	0	0	0	181	17	6	0
Russians (1,238)	1	2	8	11	1,084	112	20	0
Stavropol (118)	0	0	0	1	100	15	2	0
Belgorod (126)	1	0	1	1	110	11	2	0
Tula (132)	0	0	0	2	112	15	3	0
Kaluga (130)	0	0	1	1	115	12	1	0
Vladimir (134)	0	1	1	0	117	15	0	0
Yaroslavl (92)	0	0	1	0	81	7	3	0
Saratov (134)	0	0	2	2	114	13	3	0
Velikij Novgorod (126)	0	1	1	0	111	10	3	0
Pskov (118)	0	0	0	1	107	8	2	0
Volot (128)	0	0	1	3	117	6	1	0
Total (2,660)	1	2	10	14	2,424	178	30	1
Total (%)	0.04	0.08	0.38	0.53	91.12	6.69	1.13	0.04

mutation process at microsatellite loci, for which a generalized stepwise mutation model appears appropriate (Slatkin 1995). Calculations of between-population differentiation measures, including random permutation procedures (10,000 permutations) to assess statistical significance, were performed using the Arlequin package (Schneider et al. 2000).

Results and Discussion

The POLG gene region with the CAG microsatellite sequence was amplified from genomic DNA prepared from blood or hairs of 1,300 healthy individuals of North Eurasian origin. The results (Table 1) show that the CAG allele with 10 repeats appears to be the most common allele in different populations. This allele was found at a uniformly high frequency, varying from 88% in Russians to 98% in Koreans. The number of CAG repeats in the studied populations varied from 6 to 13.

The second most frequent allele was one codon longer than the most frequent 10-CAG allele. This 11-CAG allele is present in North Eurasians at an

average frequency of 6.7% (Table 1). Its frequency was significantly greater in Europe ($P < 0.05$, *t*-test) than in Southwest Asia (8.7%) and Siberia/East Asia (5.7% and 3.3%, respectively). The 12-CAG allele is rare in all populations studied, and its highest frequency of 3% was observed in Poles.

The shortest alleles, with 6 and 7 CAG repeats, were observed only in Russians, whereas the longest allele, with 13 CAG repeats, was found in only 1 of the 206 chromosomes sampled in Koreans. We should note that the 6-CAG allele has not been described previously. The allele with 7 CAG repeats was found earlier as a rare allele in samples from Scotland, England, Italy, and China, whereas the 13-CAG allele was present rarely in samples from Finland, Scotland, and England (Rovio et al. 1999, 2004; Krausz et al. 2004).

In general, our study demonstrates that the frequency of alleles longer than 10 CAG repeats (7.86%) is markedly higher than that of shorter alleles (1.03%): 3.71% versus 0.12% in Siberian/East Asian populations, 6.33% versus 0.9% in Southwest Asian populations, and 10.39% versus 1.51% in European populations. These data are completely consistent with previous studies of CAG variation in humans, mostly of European origin [11.4% versus 2.2%, according to Rovio et al. (2004) data].

Out of 14 observed POLG genotypes, only 6 were in common among regional groups of the studied populations (Table 2). Five genotypes (*POLG*6/*10, *7/*10, *9/*9, *9/*12,* and **8/*11*) were private (seen only in a single ethnic group) in Russians; genotypes **10/*13* and **9/*11* were observed as private in Koreans and Mongolians, respectively, and genotype **11/*12* was revealed only in Russians and Poles. As a measure of genetic variation within a population, we have used gene diversity *h*, which is equivalent to the expected heterozygosity for diploid data. As can be seen in Table 2, the heterozygosity values range from 3.8% in Koreans to 26.8% in several Russian populations. In Siberian/East Asian populations, the heterozygosity varies from 3.8% to 11%, with an average value of 7.4% (Table 3). In Southwest Asian populations the heterozygosity level ranges from 12.7% to 17.4% (13.6% on average). In European populations its values vary from 10% to 26.8% (21.6% on average). The highest values of heterozygosity were found in Russian subpopulations, and the lowest values (16–17%) were found in northwestern Russians from Pskov and Volot (Table 2). In general, the present study provides evidence for a clinal distribution of POLG gene heterozygosity in North Eurasian populations. The average heterozygosity ranges from more than 20% in Europe and more than 13% in Southwest Asia to its lowest value of 7.4% in populations of Siberia/East Asia (Table 3).

Gene diversity analysis indicates that all the study populations are characterized by a high level of homozygosity for CAG alleles because of the presence of mainly one high-frequency allele, CAG-10 (see Tables 2 and 3). This suggests that either balancing selection or the presence of an advantageous allele can explain the observed pattern of variation; however, the basis of selection acting at the POLG microsatellite in humans remains unclear (Rovio et al. 2004). One should note, however, that a prevalent length allele is also present in African

Table 2. Frequencies of the POLG Genotypes in North Eurasian Populations

Population (*N*)	*POLG Genotype*														*Gene Diversity* ($h \pm$ SE)
	**10/*10*	**10/*11*	**10/*12*	**11/*11*	**9/*10*	**8/*10*	**11*/12*	**7/*10*	**6/*10*	**9/*9*	**9/*12*	**10/*13*	**9/*11*	**8/*11*	
Koreans (103)	99	3	0	0	0	0	0	0	0	0	0	1	0	0	0.038 ± 0.02
Mongolians (44)	40	3	0	0	0	0	0	0	0	0	0	0	1	0	0.110 ± 0.05
Shors (33)	31	1	1	0	0	0	0	0	0	0	0	0	0	0	0.060 ± 0.04
Altaians (80)	72	7	0	1	0	0	0	0	0	0	0	0	0	0	0.107 ± 0.03
Tuvinians (42)	40	1	1	0	0	0	0	0	0	0	0	0	0	0	0.047 ± 0.03
Buryats (102)	93	9	0	0	0	0	0	0	0	0	0	0	0	0	0.085 ± 0.03
Tajiks (47)	40	6	0	0	1	0	0	0	0	0	0	0	0	0	0.141 ± 0.05
Persians (97)	84	11	1	0	1	0	0	0	0	0	0	0	0	0	0.127 ± 0.03
Kurds (22)	18	2	1	0	0	1	0	0	0	0	0	0	0	0	0.174 ± 0.08
Bosnians (39)	35	3	0	0	0	1	0	0	0	0	0	0	0	0	0.100 ± 0.05
Poles (102)	82	13	4	1	0	0	2	0	0	0	0	0	0	0	0.206 ± 0.04
Russians (619)	480	91	15	8	8	7	4	2	1	1	1	0	0	1	0.225 ± 0.05
Stavropol (59)	42	13	2	1	1	0	0	0	0	0	0	0	0	0	0.268 ± 0.05
Belgorod (63)	49	9	1	1	0	1	0	0	1	0	1	0	0	0	0.232 ± 0.05
Tula (66)	48	11	3	2	2	0	0	0	0	0	0	0	0	0	0.268 ± 0.05
Kaluga (65)	51	10	1	1	1	1	0	0	0	0	0	0	0	0	0.210 ± 0.05
Vladimir (67)	51	13	0	1	0	1	0	1	0	0	0	0	0	0	0.227 ± 0.05
Yaroslavl (46)	37	4	2	1	0	1	1	0	0	0	0	0	0	0	0.220 ± 0.06
Saratov (67)	50	11	1	0	0	2	2	0	0	1	0	0	0	0	0.268 ± 0.05
Velikij Novgorod (63)	48	10	3	0	0	1	0	1	0	0	0	0	0	0	0.219 ± 0.05
Pskov (59)	49	6	2	1	1	0	0	0	0	0	0	0	0	0	0.174 ± 0.05
Volot (64)	55	4	0	0	3	0	1	0	0	0	0	0	0	1	0.163 ± 0.04
Total (1,330)	1,114	150	23	10	10	9	6	2	1	1	1	1	1	1	0.118 ± 0.04

Table 3. Gene Diversity and Genetic Differentiation of North Eurasian Regional Groups

Regional Group	*Gene Diversity (h)*	*Europe*	*West Asia*	*Siberia/East Asia*
Europe	0.216 ± 0.014	**0.00148** **−0.00084**	0.00218	0.00895[a]
West Asia	0.136 ± 0.025	0.0061[a]	**−0.01019** **−0.00804**	−0.0001
Siberia/East Asia	0.074 ± 0.013	0.0272[a]	0.00737[a]	**−0.00558** **0.00214**

The boldface diagonal shows R_{ST} (top value) and F_{ST} (bottom value).
a. Value of genetic differentiation was significant ($P < 0.05$).

great apes, but a different length variant appears to have been selected in chimpanzees than in humans and gorillas (Rovio et al. 2004). This finding suggests that a selective sweep for an advantageous allele of POLG or a nearby gene could underlie the phenomenon (Rovio et al. 2004). The low degree of length diversity of the POLG microsatellite region in Siberians/East Asians compared to Europeans could reflect either that purifying selection had a more pronounced effect in Siberia/East Asia or that evolution of the POLG gene in Europeans was accelerated (new CAG alleles were accumulated) because of relaxation of selective constraints on this gene. However, the present data do not allow resolution of these alternatives.

The values of two measures of genetic differentiation, F_{ST} and R_{ST}, estimated for the three regional groups of populations are shown in Table 3. Both measures yield a low level of between-population differentiation for the combined sample, but it is noteworthy that the F_{ST} estimate is substantially higher than the R_{ST} estimate (1.96%, $P = 0$, versus 0.65%, $P = 0.002$, respectively). The same result follows from the pairwise comparisons when regional groups of populations are compared: F_{ST} values differ significantly ($P < 0.05$) from 0 in all between-region comparisons, whereas a single significant R_{ST} value is present only in a comparison of European and Siberian/East Asian regional groups (Table 3).

This discordance between the two measures of genetic differentiation may be due to distinctive features of microsatellite variation. The R_{ST} measure takes into account differences in microsatellite allele size, but the F_{ST} measure is based on the number of different alleles of microsatellite loci in populations (Slatkin 1995). Therefore the low R_{ST} values found between regional groups of populations may suggest that demographic processes (such as human migrations and genetic drift), and not mutation, are the dominant force creating population differentiation for the microsatellite sequence of the POLG gene. Meanwhile, the low level of between-population differentiation (reflected by both R_{ST} and F_{ST} values) suggests that balancing or purifying selection may favor the maintenance of the low allelic diversity seen in North Eurasian populations. Only two alleles,

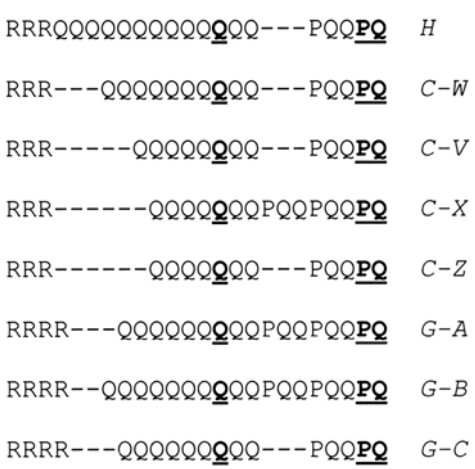

Figure 1. Major allelic variants of the POLG microsatellite region in human (H), chimpanzee (C-W, C-V, C-X, and C-Z), and gorilla (G-A, G-B, and G-C) sequences [from Rovio et al. (2004)]. Glutamines encoded by CAA and prolines encoded by CCG are underlined. Other prolines are encoded by CCT (Rovio et al. 2004).

10 CAG repeats and 11 CAG repeats, with a high prevalence of the 10-CAG allele, have been distributed in different human populations, suggesting that mutations in this microsatellite region of the POLG gene might have some functional effect on mtDNA replication.

However, low mutability of the repeated region together with drift or other nonselective factors might also explain the low allelic diversity. The only argument against such low mutability is high mutability of microsatellite sequences in human genes (Chakraborty et al. 1997; Jodice et al. 1997; Andres et al. 2002; Kondrashov and Rogozin 2004) and high variability of the $(CAG)_n$ repetitive sequence in primate POLG genes (Rovio et al. 2004). Aligned sequences of the POLG microsatellite region in human, chimpanzee, and gorilla from Rovio et al. (2004) are shown in Figure 1. There is a likely duplication of the PQQ sequence in chimpanzee and gorilla lineages (Figure 1). The $(PQQ)_2$ variant may have emerged in the last common ancestor of human, chimpanzee, and gorilla and has been maintained for 5–8 million years [for a detailed discussion of divergence times of primates, see Glazko et al. (2005)]. This would be an unusual case of ancient polymorphism because the incidence of ancestral polymorphism is low or absent on the time scale of chimpanzee-human divergence, indicating that the PQQ/$(PQQ)_2$ polymorphism is maintained by natural selection (Asthana et al. 2005).

Another possibility is that the PQQ duplication occurred independently in two lineages; however, this scenario would imply several independent insertions

or deletions of Q's, indicating that the POLG microsatellite region is highly variable on the time scale of gorilla-chimpanzee-human divergence and suggesting that there were selectional restrictions on the microsatellite length (see Figure 1). Both scenarios do not support the hypothesis that the observed distribution of POLG microsatellite alleles is due to low mutability of the repeated region combined with drift and/or other nonselective factors. In general, Figure 1 suggests that a model implying natural selection acting on the repeated region combined with high mutability of the repeated sequence is the current model of choice.

Earlier analyses of the repeated regions in the POLG gene suggested that the mutant genotype (absence of the common 10-CAG allele) is found at an elevated frequency in infertile individuals, specifically in infertile men with moderate oligozoospermia (Rovio et al. 2001). It was suggested that variant POLG alleles are somehow deleterious to sperm function or differentiation but have no phenotypic effects in other tissues (Rovio et al. 2001). We found an extremely low frequency of alleles other than the 10-CAG and 11-CAG alleles ($\cong$2.2%; see Table 1). A similar narrow distribution of $(CAG)_n$ alleles was found in the SCA2 gene (22 and 23 repeat alleles being greater than 95% of the total of human normal alleles); in another seven $(CAG)_n$-containing genes a much wider distribution of alleles was found (Andres et al. 2002). This is compatible with purifying selection acting against deleterious $(CAG)_n$ alleles. Spelbrink et al. (2000) found that deletion of the POLG $(CAG)_n$ repeat did not affect enzymatic properties but modestly up-regulated expression in cultured human cells. In addition, cells with mutant POLG enzyme containing the polyglutamine repeat deletion had a relatively low mtDNA content compared to cells containing the wild-type or exonuclease-defective POLG enzymes (Spelbrink et al. 2000).

However, recent study of the clinical significance of $(CAG)_n$ alleles has failed to confirm any influence of POLG polymorphism on the efficiency of spermatogenesis (Krausz et al. 2004; Aknin-Seifer et al. 2005). More important, considering that the homozygous mutant genotype has been found in normospermic fertile men, the analysis of the CAG repeat tract of the POLG gene does not appear to have any clinical diagnostic value (Krausz et al. 2004). These conclusions were based on comparisons of normal individuals and infertile men. The frequency of *POLG*10/*10* normal individuals studied by Rovio et al. (2001) (0.81) and in our data set (0.84; see Table 1) is similar. In contrast, the frequency of **10/*10* normal individuals of Italian origin studied by Krausz et al. (2004) was only 0.66. The difference between this group and our data set is statistically significant according to Fisher's exact test ($P = 6.6 \times 10^{-8}$). A low frequency of **10/*10* normal individuals (0.72) was found in France (Aknin-Seifer et al. 2005); the difference between this group and our data set is also statistically significant ($P = 8.9 \times 10^{-3}$). In general, such a low frequency of **10/*10* individuals was not found in any other studied human population [see Table 2 and Rovio et al. (1999, 2001, 2004)] and requires further investigation. Should this observation be supported by further, more extensive studies, this would be a case of a high genetic heterogeneity of European populations.

Acknowledgments We thank A. Jakubowska, C. Dorzhu, I. Dambueva, and I. Morozova for technical help. This research was supported by the Ludwik Rydygier Medical University in Bydgoszcz, Poland, through grant BW 61/2004, the Russian Fund for Basic Research (RFBR), and the Far-East Branch of the Russian Academy of Sciences through grants 04-3-A-06-039 and 05-3-A-06-073. B. A. Malyarchuk is supported by the Józef Mianowski Fund (2005).

Literature Cited

Aknin-Seifer, I. E., R. L. Touraine, H. Lejeune et al. 2005. Is the CAG repeat of mitochondrial DNA polymerase gamma (POLG) associated with male infertility? A multi-centre French study. *Hum. Reprod.* 20:736–740.

Andres, A. M., O. Lao, M. Soldevila et al. 2002. Dynamics of CAG repeat loci revealed by the analysis of their variability. *Hum. Mutat.* 21:61–70.

Asthana, S., S. Schmidt, and S. Sunyaev. 2005. A limited role for balancing selection. *Tr. Genet*. 21:30–32.

Chakraborty, R., M. Kimmel, D. N. Stivers et al. 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc. Natl. Acad. Sci. USA* 94:1041–1046.

Chen, X., J. Goudsmit, and A. C. Van der Kuyl. 2002. Lack of correlation between length variation in the DNA polymerase gamma gene CAG repeat and lactic acidosis or neuropathy during antiretroviral treatment. *AIDS Res. Hum. Retrovir.* 18:531–534.

Copeland, W. C., M. V. Ponamarev, D. Nguyen et al. 2003. Mutations in DNA polymerase gamma cause error-prone DNA synthesis in human mitochondrial disorders. *Acta Biochim. Polon.* 50:155–167.

Davis, A. F., P. A. Ropp, D. A. Clayton et al. 1996. Mitochondrial DNA polymerase gamma is expressed and translated in the absence of mitochondrial DNA maintenance and replication. *Nucleic Acids Res.* 24:2753–2759.

Del Bo, R., A. Bordoni, M. Sciacco et al. 2003. Remarkable infidelity of polymerase gamma A associated with mutations in POLG1 exonuclease domain. *Neurology* 61:903–908.

Di Fonzo, A., A. Bordoni, M. Crimi et al. 2003. POLG mutations in sporadic mitochondrial disorders with multiple mtDNA deletions. *Hum. Mutat.* 22:498–499.

Glazko, G. V., E. V. Koonin, and I. B. Rogozin. 2005. Molecular dating: Ape bones agree with chicken entrails. *Tr. Genet.* 21:89–92.

Jensen, M., H. Leffers, J. H. Petersen et al. 2004. Frequent polymorphism of the mitochondrial DNA polymerase gamma gene (POLG) in patients with normal spermiograms and unexplained subfertility. *Hum. Reprod.* 19:65–70.

Jodice, C., B. Giovannone, V. Calabresi et al. 1997. Population variation analysis at nine loci containing expressed trinucleotide repeats. *Ann. Hum. Genet.* 61:425–438.

Kondrashov, A. S., and I. B. Rogozin. 2004. Context of deletions and insertions in human coding sequences. *Hum. Mutat.* 23:177–185.

Krausz, C., E. Guarducci, L. Becherini et al. 2004. The clinical significance of the POLG gene polymorphism in male infertility. *J. Clin. Endocrinol. Metab.* 89:4292–4297.

Lamantea, E., V. Tiranti, A. Bordoni et al. 2002. Mutations of mitochondrial DNA polymerase gamma are a frequent cause of autosomal dominant and recessive progressive external ophthalmoplegia. *Ann. Neurol.* 52:211–219.

Lestienne, P. 1987. Evidence for a direct role of the DNA polymerase gamma in the replication of the human mitochondrial DNA in vitro. *Biochem. Biophys. Res. Commun.* 146:1146–1153.

Longley, M. J., D. Nguyen, T. A. Kunkel et al. 2001. The fidelity of human DNA polymerase gamma with and without exonucleolytic proofreading and the p55 accessory subunit. *J. Biol. Chem.* 276:38,555–38,562.

Michielotto, C., L. Stuppia, C. Foresta et al. 2003. The CAG-repeat polymorphism of the POLG1 gene is not associated with male oligozoospermia. *Am. J. Hum. Genet.* 73(suppl.):583.

Nei, M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Ponamarev, M. V., M. J. Longley, D. Nguyen et al. 2002. Active site mutation in DNA polymerase gamma associated with progressive external ophthalmoplegia causes error-prone DNA synthesis. *J. Biol. Chem.* 277:15,225–15,228.

Ropp, P. A., and W. C. Copeland. 1996. Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. *Genomics* 36:449–458.

Rovio, A. T., J. Abel, A. L. Ahola et al. 2004. A prevalent POLG CAG microsatellite length allele in humans and African great apes. *Mamm. Genome* 15:492–502.

Rovio, A. T., D. R. Marchington, S. Donat et al. 2001. Mutations at the mitochondrial DNA polymerase (POLG) locus associated with male infertility. *Nat. Genet.* 29:261–262.

Rovio, A., V. Tiranti, A. L. Bednarz et al. 1999. Analysis of the trinucleotide CAG repeat from the human mitochondrial DNA polymerase gene in healthy and diseased individuals. *Eur. J. Hum. Genet.* 7:140–146.

Schneider, S., D. Roessli, and L. Excoffier. 2000. *Arlequin Version 2000: A Software for Population Genetic Data Analysis.* Geneva: Genetics and Biometry Laboratory, University of Geneva.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:457–462.

Spelbrink, J. N., J. M. Toivonen, G. A. Hakkaart et al. 2000. In vivo functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cells. *J. Biol. Chem.* 275:24,818–24,828.

Van Goethem, G., B. Dermaut, A. Lofgren et al. 2001. Mutations of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nat. Genet.* 28:211–212.

Van Goethem, G., J. J. Martin, B. Dermaut et al. 2003. Recessive POLG mutations presenting with sensory and ataxic neuropathy in compound heterozygote patients with progressive external ophthalmoplegia. *Neuromusc. Disorders* 13:133–142.

Walker, R. L., P. Anziano, and P. S. Meltzer. 1997. A PAC containing the human mitochondrial DNA polymerase gamma gene (POLG) maps to chromosome 15q25. *Genomics* 40:376–378.

Zullo, S. J., L. Butler, R. J. Zahorchak et al. 1997. Localization by fluorescence in situ hybridization (FISH) of human mitochondrial polymerase gamma (POLG) to human chromosome band 15q24–q26, and of mouse mitochondrial polymerase gamma (Polg) to mouse chromosome band 7E, with confirmation by direct sequence analysis of bacterial artificial chromosomes (BACs). *Cytogenet. Cell Genet.* 78:281–284.