Molecular Biology, Vol. 30, No. 6, Part 1, pp. 763–769, 1996

Allele Polymorphism of Two Tetranucleotide Tandem Repeats in Two City Populations of Russia

D. A. Chistyakov,[1] I. A. Efremov,[1] O. N. Odinokova,[2] and V. V. Nosikov[1]

UDC 575.1:577.1

Translated from *Molekulyarnaya Biologiya*, Vol. 30, No. 6, pp. 1274–1283, November-December, 1996.

Original article submitted February 15, 1996.

PCR analysis of the allele polymorphism of two tetranucleotide tandem repeats (STR) HUMvWFII and D6S366 among unrelated Russians living in Moscow and Tomsk (Siberia) revealed seven HUMvWFII alleles of 154–174 bp and eight D6S366 alleles of 142–170 bp. For both loci the experimental genotype frequencies corresponded to the Hardy–Weinberg equilibrium in both population samples. No significant differences in allele frequency distribution were found between the two city populations; comparisons were made with foreign samples. Family analysis (14 cases) demonstrated Mendelian inheritance of the alleles of these loci. The polymorphism information content of the loci was calculated. These STR could be used for personal identification and paternity tests.

Key words: polymerase chain reaction; allele polymorphism; Russian population; microsatellites

Short tandem repeats (STR), or microsatellites, are the most widespread type of DNA repeats; they usually have simple motifs up to 10 nt in the repeat unit [1]. Some microsatellites of 100–300 bp exhibit high allele polymorphism. Highly polymorphic tetranucleotide tandem repeats include microsatellites D6S366 and HUMvWFII.

Intron 40 of the von Willebrand factor gene (vWF, 12p13.3–p13.2 [4, 5]) was shown to contain a polymorphic region with multiple copies of an ATCT tandem repeat. PCR [6] revealed three distanced polymorphic microsatellites in this locus. Since mutations in the vWF gene may result in von Willebrand's disease, the study of these polymorphisms may be useful for molecular diagnosis.

The microsatellite HUMvWFI located in the 5′-proximal part of intron 40 was found to have 11 alleles of 95–131 bp in European populations [7–10]. HUMvWFII in the 3′-proximal part of the polymorphic region (positions 2215–2380 in intron 40) has seven alleles of 154–178 bp [8, 11]. Amplification of the polymorphic region (744–820 bp) including HUMvWFI and HUMvWFII permits simultaneous analysis of these loci after cleaving the PCR products with *Alu*I [12].

For microsatellite HUMvWFIII (positions 1640–1794 in intron 40), seven alleles of 138–162 bp were found upon examination of 100 unrelated Britons [13]. This polymorphic STR, designated HUMvWFA31/A, is now used by the Forensic Science Service (Birmingham, UK) in a panel of seven STR for personal identification; analysis of different ethnic groups revealed 12 HUMvWFIII alleles [14].

In various population samples, 10 alleles of 138–174 bp were found for microsatellite D6S366 (6q21-qter) [15]. This microsatellite is included into a panel of 13 highly polymorphic nonlinked STR proposed for personal identification and paternity tests in the USA [15].

The aims of the present work were to analyze the allele polymorphism of microsatellites HUMvWFII and D6S366 in different samples of Russian urban populations; to compare the allele frequency distribution with analogous data for other populations; and to assess the applicability of these polymorphisms in personal identification and other work on human genetics.

MATERIALS AND METHODS

Taq^R DNA polymerase was from NPK Biotekh (Moscow). Oligonucleotide primers were synthesized in the Institute of Bioorganic Chemistry (Moscow).

[1]National Research Center GosNIIgenetika, Moscow, 113545.
[2]Institute of Medical Genetics, Russian Academy of Medical Sciences, Tomsk, 634007.

0026-8933/97/3006-0763$15.00

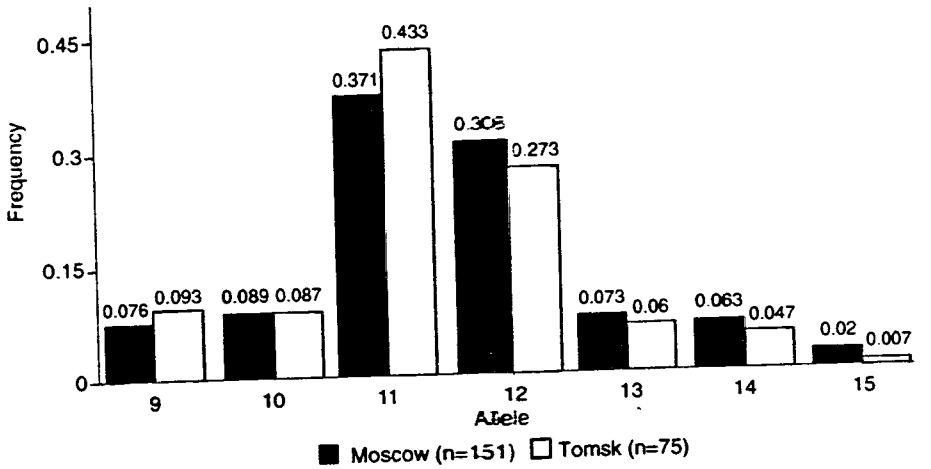

Fig. 1. Frequencies of HUMvWFII alleles in Russian urban populations: allele designation denotes the number of tandem repeats therein; n is the sample volume.

Isolation of genomic DNA from venous blood followed a published protocol [16]. DNA from saliva and blood spots was isolated using Chelex-100 (Bio-Rad) [17].

The sample of unrelated Muscovites was composed of specimens collected at traumatological units, Institute of Rheumatology, Bureau of Legal Medicine, and from the staff of GosNIIgenetika. The sample of unrelated inhabitants of Tomsk comprised specimens from maternities and healthy volunteers.

PCR was run on a Techne PHC-2 thermal cycler in 50 µl of reaction mix containing 67 mM Tris-HCl (pH 8.8), 16.6 mM ammonium sulfate, 0.01% Tween-20, 1.0 or 1.5 mM $MgCl_2$ (HUMvWFII or D6S366, respectively), 0.2 mM of each dNTP, 2.5 un. *Taq* polymerase, 10–100 ng genomic DNA or 20 µl of Chelex-100 DNA extract, and 66 ng each of primers based on published sequences ([11] for HUMvWFII and [15] for D6S366). Amplification included 30–35 cycles of 1 min at 94°C, 1 min at 55°C (HUMvWFII) or 65°C (D6S366), and 1 min at 72°C, with first denaturation for 5 min and last synthesis for 7 min.

Amplification products were analyzed in 12% PAG (16 cm long, 0.7 mm thick, with 7% glycerol), applying 10 µl of the reaction mix per lane; gels were stained with silver [18].

The observed genotype frequencies were checked for deviation from the Hardy–Weinberg equilibrium by the χ^2 and G-statistic tests, using the RowsxColumns program based on a published algorithm [19] that permits statistical assessment when the number of observations in many classes is below five and the standard χ^2 test is inapplicable. The RxC program was also used to compare the allele frequency distributions in different population samples.

This program generated 1000 random 2D data sets of the same dimensionality and with the same limit sum as the initial set of compared observations (frequencies). The χ^2 and G-statistic values were computed for each set (further results given for the initial one). The program evaluated the probability of obtaining a set (percentage of sets) with the test values not lower than the initial. Low probability (i.e., ≈0.05) indicates than rows and columns in the set are not independent, and the data are not homogeneous enough to be united; values above 0.05 indicate that the data may be observed with a probability of 5% in the case of independent rows and columns, and can be legitimately united.

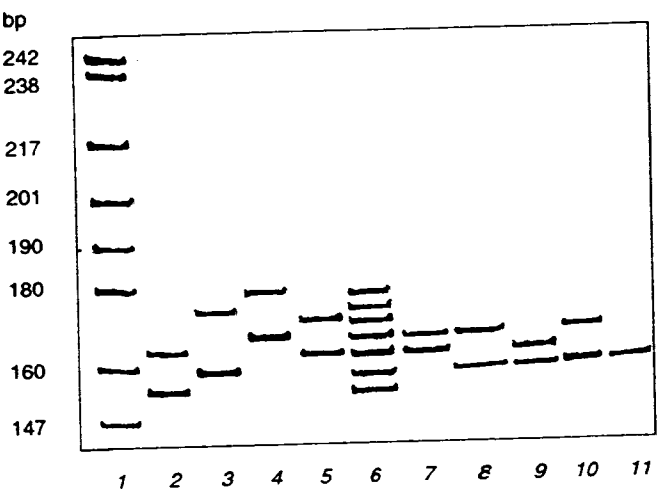

Fig. 2. Electrophoretic resolution of amplified HUMvWFII alleles. *1*) Markers (pBR322/*MspI*); *2–5*) genotyping of four unrelated individuals: 9/11 (2), 10/14 (*3*), 12/15 (*4*), 11/13 (*5*); *6*) allele ladder: *7–11*) a family study: genotypes *7*(father) 11/12, *8*(child 1) 10/12, 9(child 2) 10/11, *10*(child 3) 10/12, *11*(mother) 10/10.

The expected hetrozygosity (H_{exp}), probability of random match (pM), mean exclusion chance (W), and polymorphism information content (PIC) were calculated as described previously [20].

TABLE 1. Frequencies of HUMvWFII Genotypes in Russian Urban Populations

Genotype	Frequency			
	Moscow		Tomsk	
	Observed	Expected	Observed	Expected
9-9	0.007	0.006	0.013	0.009
9-10	0.013	0.014	0.013	0.016
9-11	0.053	0.056	0.120	0.081
9-12	0.040	0.047	0.013	0.051
9-13	0.013	0.011	0.013	0.011
9-14	0.013	0.010	0	0.009
9-15	0.007	0.003	0	0.001
10-10	0.026	0.008	0	0.008
10-11	0.033	0.066	0.093	0.075
10-12	0.046	0.055	0.053	0.047
10-13	0.020	0.013	0.013	0.010
10-14	0.013	0.011	0	0.008
10-15	0	0.004	0	0.001
11-11	0.159	0.138	0.147	0.188
11-12	0.212	0.228	0.267	0.237
11-13	0.079	0.054	0.040	0.052
11-14	0.033	0.047	0.053	0.040
11-15	0.013	0.015	0	0.006
12-12	0.106	0.095	0.053	0.075
12-13	0.020	0.045	0.053	0.033
12-14	0.066	0.039	0.040	0.026
12-15	0.020	0.012	0.013	0.004
13-13	0.007	0.005	0	0.004
13-14	0	0.009	0	0.006
13-15	0	0.003	0	0.001
14-14	0	0.004	0	0.002
14-15	0	0.002	0	0.001
15-15	0	0	0	0

RESULTS AND DISCUSSION

Genotyping of 151 and 75 unrelated individuals in Moscow and Tomsk populations, respectively, revealed seven alleles of the HUMvWFII microsatellite, ranging from 154 to 178 bp (Figs. 1 and 2). The most frequent in both samples were alleles 11 and 12. Among the 28 genotypes possible, 22 were observed in Moscow and 16 in Tomsk; heterozygotes 11/12 were the most common: 21% of cases in Moscow and 27% in Tomsk (Table 1).

For D6S366, genotyping of 102 (Moscow) and 75 (Tomsk) unrelated individuals revealed eight alleles of 142–170 bp (Fig. 3 and 4); alleles 12 and 13 were the most common in both samples. Among the 36 genotypes possible, 22 were observed in Moscow and 20 in Tomsk; heterozygotes 12/13 were the most frequent: 21% of cases in Moscow and 15% in Tomsk (Table 2).

For both loci, the observed genotype frequencies in both samples showed nice correspondence to the Hardy–Weinberg equilibrium (Table 3), testifying to the absence of intrinsic sample heterogeneity.

The observed HUMvWFII allele and genotype frequency distributions proved much the same in the two Russian samples (Table 4). Comparisons were also made with the data available on European populations. Among 24 Dutchmen, alleles 11 and 12 also were the most frequent (0.35 and 0.27, respectively) [11]. A unimodal distribution was found for 44 Britons [8]: the frequency of allele 12 was 0.42, versus 0.18 for the next-frequent allele 11. The R×C data reveal no significant differences between either Russian and Dutch samples, whereas the distinctions between Russians and Britons appear to be reliable.

Likewise, comparisons were made for the D6S366 allele frequency distributions in the two Russian samples and different ethnic groups [15] (Fig. 5). The results for Moscow and Tomsk proved very close and

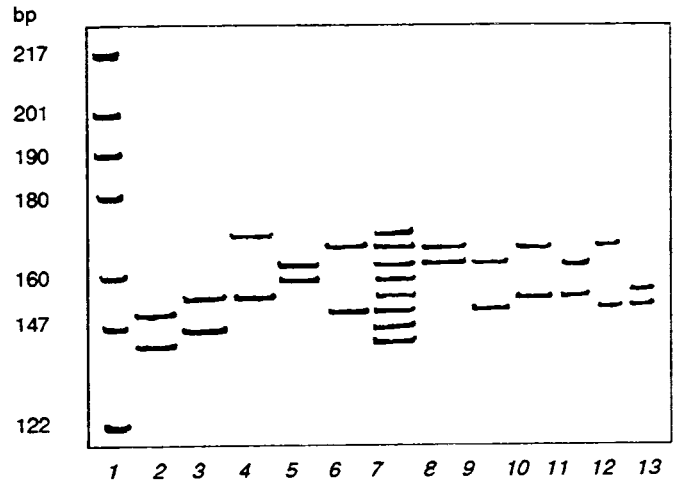

Fig. 3. Electrophoretic resolution of amplified D6S366 alleles. *1*) Markers (pBR322/*Msp*I); *2–6*) genotyping of five unrelated individuals: 10/12 (*2*), 11/13 (*3*), 13/17 (*4*), 14/15 (*5*), 12/16 (*6*); *7*) allele ladder; *8–13*) a family study: genotypes *8*(father) 15/16, *9*(child 1) 12/15, *10*(child 2) 13/16, *11*(child 3) 13/15, *12*-child 4) 12/16, *13*(mother) 12/13.

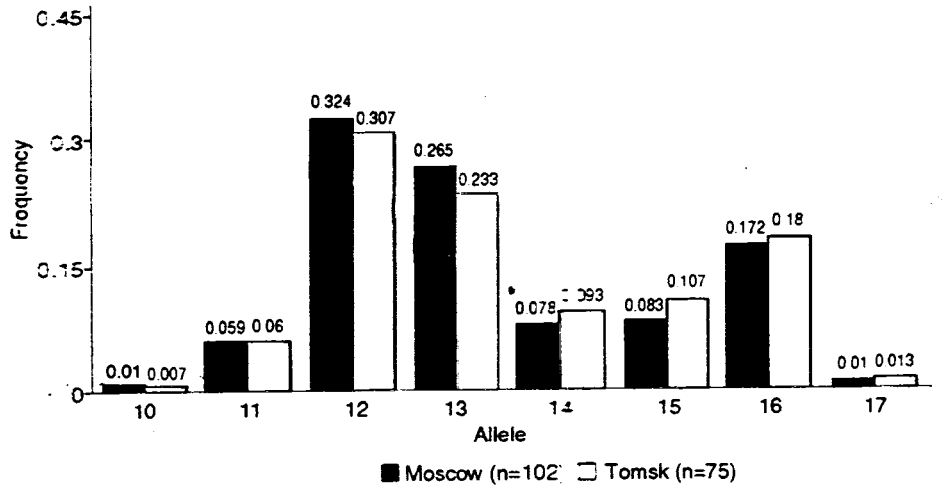

Fig. 4. Frequencies of D6S366 alleles in Russian urban populations; allele numeration denotes the number of tandem repeats therein; n is the sample volume.

rather similar to those for US Caucasians (170 individuals); however, the latter did not show such a marked prevalence of alleles 12 and 13 (respective frequencies 0.265 and 0.200), and had an additional very rare allele 18 of 174 bp (frequency 0.003). US Hispanics (164) had alleles 9 and 18 not found in the Russian population, and showed prevalence of allele 13. Afro-Americans (178) had alleles 9–17 with predominant 13 and 14. Among Mongoloids (68), alleles 12 and 13 were prevalent as among Caucasoids, but only six alleles were found: 9 and 11–15. As follows from the data in Fig. 5, differences in the D6S366 allele distribution from the Russian population were statistically reliable except for US Caucasians.

Mendelian inheritance of HUMvWFII and D6S366 alleles was observed in analysis of 14 families with up to four children (Figs. 2 and 3); no mutant alleles were detected.

The results of allele frequency comparisons suggest that for sufficiently remote and isolated Caucasian populations there are real differences, commensurate with interethnic ones: thus the numeric values of RxC probabilities are practically the same for pairs Russians–Britons (HUMvWFII) and Russians–Mexican Hispanics (D6S366). This idea is supported by the literature data on real interpopulation differences in allele frequencies for micro- and minisatellites [21, 22].

Nonetheless, the distributions for the two loci proved virtually indistinguishable in samples from two remote Russian urban populations. In probability calculations for personal identification, the allele frequencies determined for a single representative sample may

TABLE 2. Frequencies of D6S366 Genotypes in Russian Urban Populations

Genotype	Frequency			
	Moscow		Tomsk	
	Observed	Expected	Observed	Expected
10-10	0	0	0	0
10-11	0.010	0.001	0	0.001
10-12	0.010	0.006	0	0.004
10-13	0	0.005	0	0.003
10-14	0	0.002	0.013	0.001
10-15	0	0.002	0	0.001
10-16	0	0.003	0	0.002
10-17	0	0	0	0
11-11	0.020	0.003	0	0.004
11-12	0.049	0.038	0.053	0.037
11-13	0.020	0.031	0.040	0.028
11-14	0	0.009	0	0.011
11-15	0	0.010	0	0.013
11-16	0	0.020	0.027	0.022
11-17	0	0.001	0	0.002
12-12	0.078	0.105	0.107	0.094
12-13	0.206	0.171	0.147	0.143
12-14	0.020	0.051	0.053	0.057
12-15	0.049	0.054	0.027	0.065
12-16	0.147	0.111	0.107	0.110
12-17	0.010	0.006	0.013	0.008
13-13	0.039	0.070	0.053	0.054
13-14	0.088	0.042	0.040	0.044
13-15	0.049	0.044	0.067	0.050
13-16	0.078	0.091	0.067	0.084
13-17	0.010	0.005	0	0.006
14-14	0.010	0.006	0	0.009
14-15	0.010	0.013	0.027	0.020
14-16	0.020	0.027	0.053	0.034
14-17	0	0.002	0	0.002
15-15	0.010	0.007	0.013	0.011
15-16	0.039	0.029	0.053	0.038
15-17	0	0.002	0.013	0.003
16-16	0.029	0.029	0.027	0.032
16-17	0	0.003	0	0.005
17-17	0	0	0	0

be extrapolated with sufficient confidence over the entire Russian population, without risk of overestimating the individualization potential.

Comparison of the information parameters for HUMvWFII and D6S366 (Table 3) shows that the latter is substantially more informative, with heterozygosity of about 80% for the Russian population. On the whole, the values of the information parameters for these loci in the Moscow sample are comparable to those of the six minisatellite loci we have studied earlier [20, 23], which vouches for the usefulness of these microsatellites in personal identification among Russians.

As compared with the VNTR alleles, the chance that microsatellite copies would be preserved in heavily degraded DNA is much higher since the STR alleles are shorter (100–350 bp). Therefore, in addition to the existing panel of six minisatellite loci [20, 23], it would be expedient to make an analogous panel of 6–7 microsatellites for reliable personal identification, especially when there is shortage of biological material.

The authors are indebted to G. Carmody (Carleton University, Ottawa) for the RxC program and to O.

TABLE 3. Information Parameters and Goodness-of-Fit Tests for HUMvWFII and D6S366

	UMvWFII		D6S366	
	Moscow	Tomsk	Moscow	Tomsk
H_{obs}	0.695	0.787	0.814	0.800
H_{exp}	0.747	0.720	0.793	0.801
pM	0.103	0.121	0.081	0.071
W	0.678	0.632	0.762	0.794
PIC	0.587	0.543	0.626	0.659
χ^2	12.5265	8.4883	14.3408	7.4906
Probability	0.9900	0.9930	0.9530	1.0000
(±SE)	(0.0031)	(0.0026)	(0.0041)	(0)
G-Statistic	14.3040	10.1836	17.5434	9.4731
Probability	0.9920	0.9960	0.9530	1.0000
(±SE)	(0.0028)	(0.0020)	(0.0041)	(0)

TABLE 4. Comparison of HUMvWFII Allele Frequency Distribution in the Moscow Population with Those in Other Caucasians

	Russians (Tomsk)	Dutchmen [9]	Britons [18]
Sample volume	75	24	44
R×C set	7×2	7×2	7×2
χ^2	3.6024	8.2531	15.6530
Probability ± SE	0.7230±0.0142	0.1970±0.0126	0.0130±0.0036
G-statistic	3.7821	8.1846	16.1609
Probability ± SE	0.7030±0.0144	0.2460±0.0136	0.0180±0.0042

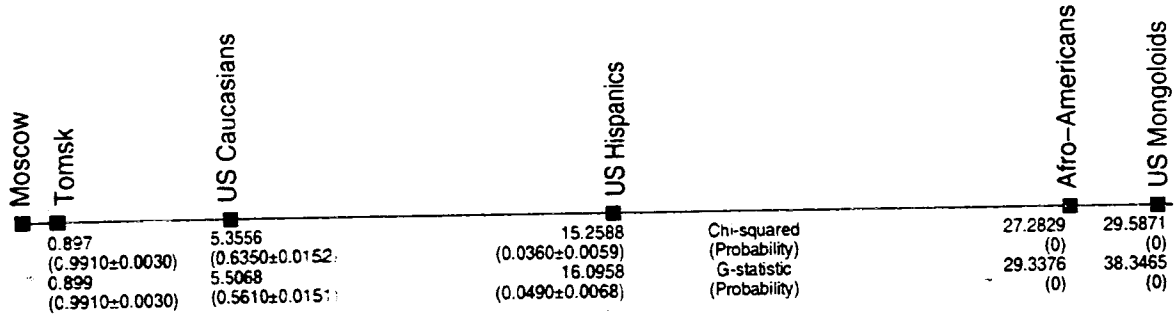

Fig. 5. Comparison of D6S366 allele frequency distribution in the Moscow population with those in other population samples.

V. Vishnichenko (Space Instrument Institute, Moscow) for designing other software.

REFERENCES

1. D. Tautz and M. Renz, *Nucl. Acids Res.*, **12**, 4127–4138 (1984).
2. A. Edwards. H. A. Hammond, Jin Li. C. T. Caskey, and R. Chakraborty, *Genomics*, **12**, 241–253 (1992).
3. I. R. Peake. M. B. Liddell, P. Moodie. G. Standen, D. J. Mancuso, E. A. Tuley, L. A. Westfiels. J. M. Sorace, J. E. Sadler, C. L. Verweij, and A. L. Bloom. *Blood*, **75**, 654–661 (1990).
4. D. J. Mancuso. E. A. Tuley, L. A. Westfield, N. K. Worall, B. B. Shelton-Inloes, J. M. Sorace, Y. G. Alevy, and J. E. Sadler, *J. Biol. Chem.*, **264**, 19514–19527 (1989).
5. D. Ginsburg, R. I. Handin, D. Bonthron, T. A. Donlon, G. A. Bruns, S. A. Latt, and S. H. Orkin. *Science*, **228**, 1401–1406 (1985).
6. R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich, *Science*, **239**, 487–491 (1988).
7. I. R. Peake, D. Bowen, P. Bignell, M. B. Liddell, J. E. Sadler, G. Standen, and A. L. Bloom, *Blood*, **76**, 555–561 (1990).

8. A. M. Cumming, J. G. Armstrong, K. Pendry, A. M. Burn, and R. T. Wensley, *Hum. Genet.*, **89**, 194–198 (1992).
9. A. V. Misyurin, V. L. Surin, and G. Ya. Solov'ev, *Genetika*, **30**, 713–717 (1994).
10. M. V. Aseev, V. N. Skakun, and V. S. Baranov, *Genetika*, **31**, 539–845 (1995).
11. H. K. Ploos van Amstel and P. H. Reitsma, *Nucl. Acids Res.*, **18**, 4957 (1990).
12. B. Mercier, C. Gaucher, and C. Mazurier, *Nucl. Acids Res.*, **19**, 4800 (1991).
13. C. Kimpton, A. Walton, and P. Gill, *Hum. Mol. Genet.*, **1**, 287 (1992).
14. A. Urquhart, N. J. Oldroyd, C. P. Kimpton, and P. Gill, *BioTechniques*, **18**, 116–121 (1995).
15. H. A. Hammond, Jin Li, Y. Zhong, C. T. Caskey, and R. Chakraborty, *Am. J. Hum. Genet.*, **55**, 175–189 (1994).
16. M. B. Johns and J. E. Paules-Thomas, *Anal. Biochem.*, **180**, 276–278 (1989).
17. P. Walsh, D. Metzger, and R. Higuchi, *BioTechniques*, **10**, 506–513 (1991).
18. B. Budowle, R. Chakraborty, A. M. Giusti, A. J. Eisenberg, and R. C. Allen, *Am. J. Hum. Genet.*, **48**, 137–144 (1991).
19. D. A. Roff and P. Bentzen, *Mol. Biol. Evol.*, **6**, 539–545 (1989).
20. I. A. Efremov, D. A. Chistyakov, and V. V. Nosikov, *Mol. Biol.*, **30**, 307–318 (1996).
21. R. Deka, R. Chakraborty, and R. Ferrell, *Genomics*, **11**, 83–92 (1991).
22. W. J. Wall, R. Williamson, M. Petrou, D. Papaioannou, and B. H. Parkin, *Hum. Mol. Genet.*, **2**, 1023–1029 (1993).
23. D. A. Chistyakov, D. K. Gavrilov, I. V. Ovchinnikov, and V. V. Nosikov, *Mol. Biol.*, **27**, 1304–1314 (1993).