Molecular Biology, Vol. 31, No. 6, 1997, pp. 832–837. Translated from Molekulyarnaya Biologiya, Vol. 31, No. 6, 1997, pp. 978–984.

GENE MOLECULAR BIOLOGY

UDC 575.1:577.1

Allele Polymorphism of Tetranucleotide Tandem Repeat SE33 among the Udege People and Two Urban Populations of Russia

R. I. Turakulov[1], D. A. Chistyakov[1], O. N. Odinokova[2], O. K. Galaktionov[3], and V. V. Nosikov[1]

[1]*National Research Center GosNIIgenetika, Moscow, 113545 Russia*
[2]*Institute of Medical Genetics, Tomsk Research Center, Russian Academy of Medical Sciences, Tomsk, 634050 Russia*
[3]*Institute of Psychical Health, Tomsk Research Center, Russian Academy of Medical Sciences, Tomsk, 634014 Russia*

Received May 7, 1997

Abstract—The allele polymorphism of the SE33 tetranucleotide microsatellite repeat (locus ACTBP2) was analyzed in the Udege, Moscow, and Tomsk populations. PCR revealed 21 alleles of 246–326 nt. No significant differences in allele frequency distribution were found in the three populations. Genotype frequencies were shown to obey the Hardy–Weinberg equilibrium. Family analysis (two cases) demonstrated Mendelian inheritance of the alleles. The heterozygosity (0.830–0.875), the polymorphism information content (0.874–0.887), and the power of discrimination (0.991–0.992) showed that this locus can be used for personal identification and paternity tests.

Key words: polymerase chain reaction, allele polymorphism, Moscow and Tomsk populations, Udege, SE33 microsatellite

INTRODUCTION

Highly polymorphic (having more than six alleles) and highly informative microsatellite genetic markers are widely employed in genome mapping, studies of hereditary diseases, and personal identification [1, 2]. Identification panels of microsatellite markers are used in forensic medicine in the United States, United Kingdom, and Germany [3, 4]. We made an analogous panel of six microsatellite markers and studied their polymorphism in the Russian population [5–7]. For exact personal identification in the Russian population, the probability of random match (pM) for genotypes of two unrelated persons must be $6.9 \cdot 10^{-7}$. To meet this requirement, the panel must include one or two additional highly polymorphic loci, e.g., the SE33 microsatellite.

The SE33 microsatellite containing multiple copies of an AAAG tandem repeat was revealed in the intron of the β-actin pseudogene (locus ACTBP2) located on chromosome 5qter [8]. In total, about 20 β-actin pseudogenes are known [9, 10]. In Caucasians, 26 SE33 alleles of 222–322 nt were found [8, 11]. Thus, SE33 is comparable in polymorphism with the minisatellite MCT118 (locus D1S80), which has 28 alleles and suballeles [12] and is widely employed in genome fingerprinting.

The aims of this work were to analyze the allele polymorphism of the SE33 microsatellite in samples of the Moscow, Tomsk, and Udege populations and to assess its applicability in personal identification.

EXPERIMENTAL

Genomic DNA from venous blood was isolated according to [13]. DNA from saliva and blood spots and from hair roots was isolated using Chelex-100 (Bio-Rad) [14]. Before isolation, a hair was incubated in 200 μl of 10 mM Tris-HCl, pH 7.5, 0.32 M sucrose, 5 mM $MgCl_2$, 1% Triton X-100, 0.5 mg/ml proteinase K at 60°C overnight.

In Moscow and Tomsk, blood samples of unrelated healthy donors were collected at traumatological units and hemotransfusion stations. Blood samples of the

0026-8933/97/3106-0832$18.00 © 1997 МАИК Наука/Interperiodica Publishing

Table 1. Frequencies of ACTBP2 alleles in the Russian and Udege populations

Allele		Moscow		Tomsk		Udege	
Repeat number	Length, nt	Case number	Frequency	Case number	Frequency	Case number	Frequency
15	238	0	0	0	0	1	0.006 ± 0.006
16	242	0	0	0	0	3	0.018 ± 0.010
17	246	5	0.023 ± 0.010	3	0.028 ± 0.016	4	0.027 ± 0.015
18	250	6	0.027 ± 0.011	7	0.066 ± 0.024	3	0.018 ± 0.010
19	254	17	0.077 ± 0.018	11	0.104 ± 0.030	9	0.055 ± 0.018
20	258	11	0.050 ± 0.015	11	0.104 ± 0.030	20	0.122 ± 0.026
21	262	7	0.032 ± 0.012	8	0.075 ± 0.026	14	0.085 ± 0.022
22	266	13	0.059 ± 0.016	7	0.066 ± 0.024	7	0.043 ± 0.016
23	270	11	0.050 ± 0.015	11	0.104 ± 0.030	14	0.085 ± 0.022
24	274	7	0.032 ± 0.012	4	0.038 ± 0.019	13	0.079 ± 0.021
25	278	6	0.027 ± 0.011	2	0.019 ± 0.013	10	0.061 ± 0.019
26	282	10	0.045 ± 0.014	2	0.019 ± 0.013	6	0.037 ± 0.015
27	286	14	0.064 ± 0.016	2	0.019 ± 0.013	6	0.037 ± 0.015
28	290	24	0.109 ± 0.021	0	0.000 ± 0.000	3	0.018 ± 0.010
29	294	15	0.068 ± 0.017	1	0.009 ± 0.009	7	0.061 ± 0.019
30	298	7	0.032 ± 0.012	3	0.028 ± 0.016	10	0.061 ± 0.019
31	302	12	0.055 ± 0.015	5	0.047 ± 0.021	7	0.043 ± 0.016
32	306	22	0.100 ± 0.020	13	0.123 ± 0.032	6	0.037 ± 0.015
33	310	19	0.086 ± 0.019	9	0.085 ± 0.027	11	0.067 ± 0.020
34	314	10	0.045 ± 0.014	4	0.038 ± 0.019	5	0.030 ± 0.013
35	318	3	0.014 ± 0.008	2	0.019 ± 0.013	2	0.012 ± 0.009
36	322	1	0.005 ± 0.005	1	0.009 ± 0.009	3	0.018 ± 0.010
Total		220	1.000 ± 0.280	106	1.000 ± 0.383	164	1.000 ± 0.338

Udege were collected in Krasnyi Yar (Pozharskii raion, Primorskii krai).

PCR primers based on published sequences [8] were synthesized by Evios-Ros (Moscow). PCR was run on a Techne PHC-2 thermal cycler in 60 μl of the reaction mixture containing 67 mM Tris-HCl (pH 6.8), 16.6 mM ammonium sulfate, 0.01% Tween-20, 1.0 mM $MgCl_2$, 0.2 mM of each dNTP, 2 units of *Taq* DNA polymerase (Biotekh), 0.1–0.2 μg genomic DNA, and 6 pmol of each primer. Amplification included 35 cycles of 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C, with first denaturation for 4 min and last synthesis for 7 min.

Alleles were identified using a marker "ladder" that was amplified in 100 μl of the reaction mixture containing 10 pmol of each primer and 1 μl of an equimolar mixture of all amplified alleles diluted 10^3 times.

Amplification products were analyzed in 8% PAG containing 7% glycerol; 10 μl of the reaction mixture per lane was applied; gels were stained with silver [15].

Genotype frequencies were checked for deviation from the Hardy–Weinberg equilibrium by the χ^2 test and G-statistics, using the Rows × Columns program based on a published algorithm [16]. This program was also used to compare the allele frequency distributions in different population samples.

The expected hetrozygosity (H_{exp}), mean exclusion chance (W), pM, and polymorphism information content (PIC) were calculated as described previously [17]; the power of discrimination (PD) was obtained as PD = 1 – pM [18].

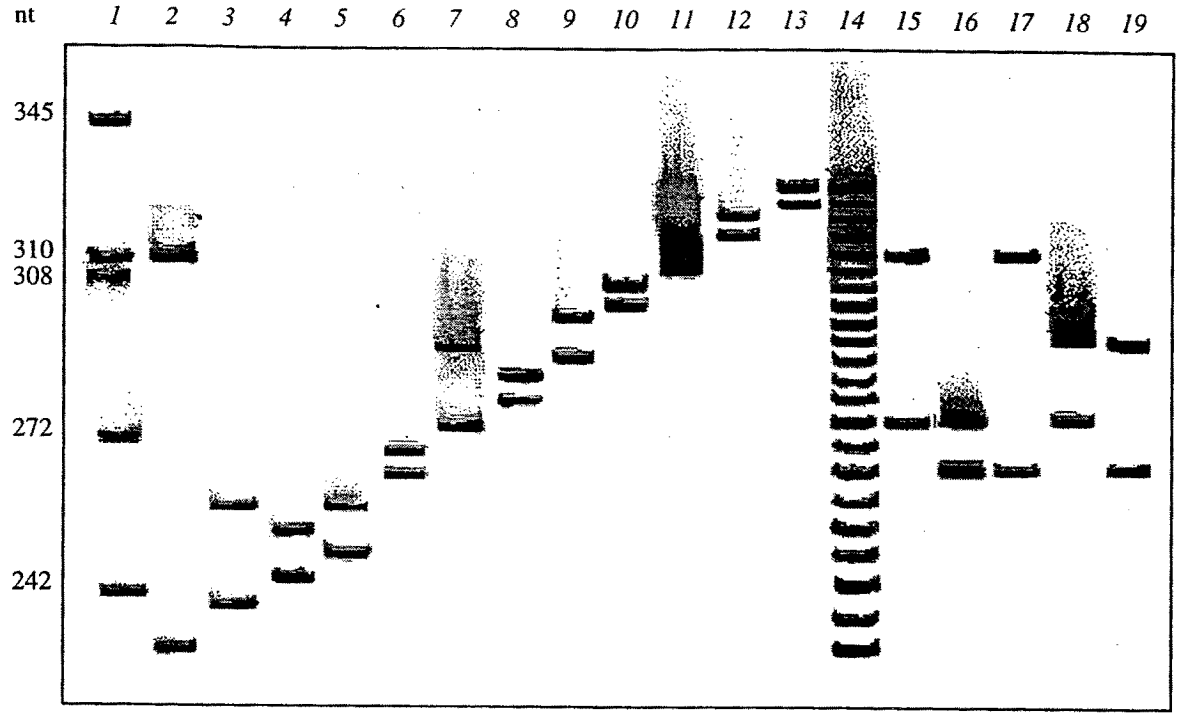

Fig. 1. Electrophoretic separation of amplified ACTBP2 alleles. (*1*) Markers (λ/*Eco*RI); (*2–13*) genotyping of 12 unrelated individuals: 15/33 (*2*), 16/20 (*3*), 17/19 (*4*), 18/20 (*5*); 21/22 (*6*), 23/27 (*7*), 24/25 (*8*), 26/28 (*9*), 29/30 (*10*), 31/32 (*11*), 33/34 (*12*), 35/36 (*13*); (*14*) allele ladder; (*15–19*) a family study: 23/32, father (*15*); 21/23, child 1 (*16*); 21/32, child 2 (*17*); 23/27, child 3 (*18*); 21/27, mother (*19*).

RESULTS AND DISCUSSION

Analysis of 110 and 53 persons in the Moscow and Tomsk populations revealed 20 alleles of the SE33 microsatellite, ranging from 246 to 322 nt. In addition, two other alleles, of 238 and 242 nt, were found in the Udege population (86 individuals). The data obtained are given in Table 1 and Fig. 1. The alleles were designated according to the repeat number within a tandem array. The shortest allele of 238 nt included 15 tetranucleotide units [11]. The most frequent were alleles 28 and 32 in Moscow, 19 and 32 in Tomsk, and 20 in the Udege population (Table 1).

Among the 210 genotypes possible, 76 were observed in Moscow and 48 in Tomsk. Heterozygotes 32/33 and homozygotes 19/19 were the most common in Moscow; five genotypes were the most frequent in Tomsk. In the Udege, 64 out of 253 possible genotypes were found; heterozygotes 20/21 and 21/23 were the most common (Table 2).

The observed genotype frequencies in all samples obeyed the Hardy–Weinberg equilibrium (Table 3), indicating the absence of intrinsic sample heterogeneity.

The SE33 allele frequency distributions proved much the same in the two rather remote Russian urban populations (Fig. 2), as observed previously with six other microsatellites [5–7]. This suggests that the allele frequencies may be extrapolated with sufficient confidence over the entire Russian population.

Comparisons were also made with the data available on other populations. German Caucasians (180 individuals) showed the highest polymorphism of the ACTBP2 locus (26 alleles) and had additional alleles 11–16 of 222–242 nt not found in the Russian populations; alleles 16, 28, and 30 occurred at a frequency higher than 0.07 [11]. US Caucasians (39) had 21 alleles with predominant 19 and 28–30 (frequency higher than 0.09) [8]. Contrary to the Moscow and Tomsk populations, they had alleles 13–16, but had no alleles 34–36 found in Russians and German Caucasians [8, 11].

R × C analysis revealed no differences in the ACTBP2 allele frequency distribution between the Moscow and Tomsk populations, German and US Caucasians, and the Udege (Mongoloids) (Fig. 2). However, significant interrace differences (between Russians, US Mongoloids, and Afro-Americans) were previously detected for microsatellites D6S366, D19S253, HUMCYAR04, HUMCD4, and HUMF13A01 [5–7]. The similarity of the ACTBP2 allele frequency distributions in Russians and the

Table 2. Frequencies of ACTBP2 genotypes in Russian and Udege populations

Genotype	Moscow		Tomsk		Udege		Genotype	Moscow		Tomsk		Udege	
	Frequency	Case number	Frequency	Case number	Frequency	Case number		Frequency	Case number	Frequency	Case number	Frequency	Case number
1	2	3	4	5	6	7	1	2	3	4	5	6	7
15/33	0	0	0	0	0.012	1	20/24	0	0	0	0	0.037	3
16/20	0	0	0	0	0.012	1	20/25	0.009	1	0	0	0.012	1
16/22	0	0	0	0	0.012	1	20/26	0	0	0	0	0.012	1
16/34	0	0	0	0	0.012	1	20/29	0.009	1	0	0	0.012	1
17/17	0	0	0.019	1	0	0	20/30	0	0	0	0	0.012	1
17/19	0	0	0.019	1	0	0	20/31	0	0	0	0	0.012	1
17/20	0.009	1	0	0	0	0	20/32	0.018	2	0.019	1	0	0
17/24	0.018	2	0	0	0	0	20/33	0	0	0.019	1	0.012	1
17/25	0	0	0	0	0.012	1	20/36	0	0	0	0	0.012	1
17/26	0.009	1	0	0	0	0	21/21	0.009	1	0.019	1	0.012	1
17/27	0	0	0	0	0.012	1	21/22	0.009	1	0.019	1	0.012	1
17/32	0	0	0	0	0.012	1	21/23	0.009	1	0	0	0.049	4
17/34	0.009	1	0	0	0	0	21/27	0.009	1	0	0	0	0
17/36	0	0	0	0	0.012	1	21/29	0	0	0	0	0.012	1
18/19	0.009	1	0	0	0	0	21/30	0	0	0	0	0.012	1
18/20	0.009	1	0.019	1	0.012	1	21/32	0	0	0.019	1	0	0
18/24	0	0	0	0	0.012	1	21/33	0.009	1	0	0	0.012	1
18/26	0	0	0.019	1	0	0	21/34	0.009	1	0	0	0	0
18/27	0.009	1	0	0	0	0	22/22	0.009	1	0	0	0	0
18/28	0.009	1	0	0	0	0	22/23	0.018	2	0	0	0.012	1
18/30	0	0	0	0	0.012	1	22/26	0.009	1	0	0	0	0
18/31	0.018	2	0.019	1	0	0	22/27	0.027	3	0	0	0	0
18/32	0	0	0.038	2	0	0	22/28	0.009	1	0	0	0	0
18/33	0	0	0.038	2	0	0	22/29	0.009	1	0	0	0	0
19/19	0.036	4	0.038	2	0	0	22/30	0	0	0	0	0.012	1
19/20	0.009	1	0.019	1	0.012	1	22/31	0	0	0.019	1	0.012	1
19/21	0	0	0.019	1	0	0	22/32	0.018	2	0.019	1	0	0
19/22	0	0	0.019	1	0	0	22/33	0	0	0.019	1	0.012	1
19/23	0.009	1	0.019	1	0	0	23/23	0	0	0.019	1	0.024	2
19/24	0	0	0	0	0.024	2	23/24	0.009	1	0	0	0	0
19/25	0	0	0.019	1	0.024	2	23/25	0	0	0	0	0.012	1
19/26	0	0	0.019	1	0.024	2	23/27	0.009	1	0.019	1	0	0
19/28	0.009	1	0.019	1	0	0	23/28	0.018	2	0	0	0	0
19/32	0.027	3	0	0	0	0	23/29	0.009	1	0.019	1	0.036	2
19/33	0.018	2	0	0	0.024	2	23/30	0	0	0	0	0.012	1
20/20	0.018	2	0.019	1	0.012	1	23/31	0	0	0	0	0.012	1
20/21	0	0	0.019	1	0.049	4	23/32	0.009	1	0	0	0	0
20/22	0	0	0.019	1	0.012	1	23/33	0.009	1	0.019	1	0	0
24/24	0	0	0	0	0.012	1	28/31	0.009	1	0	0	0	0
24/25	0	0	0.019	1	0.012	1	28/33	0.009	1	0.019	1	0	0
24/26	0	0	0	0	0.012	1	28/34	0.018	2	0	0	0	0
24/27	0	0	0	0	0.024	2	28/35	0.009	1	0	0	0	0

Table 2. (Contd.)

1	2	3	4	5	6	7	1	2	3	4	5	6	7
24/31	0.009	1	0.019	1	0	0	29/29	0	0	0.019	1	0.012	1
24/32	0.018	2	0.019	1	0.012	1	29/30	0.009	1	0	0	0	0
24/33	0.009	1	0	0	0	0	29/31	0.009	1	0	0	0	0
24/35	0	0	0.019	1	0	0	29/32	0	0	0.038	2	0	0
25/25	0	0	0	0	0.012	1	29/34	0.018	2	0	0	0	0
25/26	0.009	1	0	0	0	0	29/35	0.009	1	0	0	0	0
25/27	0.009	1	0	0	0.024	2	30/30	0	0	0	0	0.024	2
25/28	0.018	2	0	0	0	0	30/31	0.018	2	0	0	0	0
25/34	0.009	1	0	0	0	0	30/32	0.009	1	0.019	1	0	0
26/26	0	0	0	0	0.012	1	30/33	0.009	1	0	0	0	0
26/27	0.009	1	0	0	0	0	30/34	0	0	0	0	0.012	1
26/28	0.009	1	0	0	0	0	30/35	0.009	1	0	0	0	0
26/29	0.018	2	0	0	0	0	31/31	0.009	1	0	0	0	0
26/32	0.009	1	0	0	0	0	31/32	0.009	1	0.019	1	0.012	1
26/33	0	0	0	0	0	0	31/33	0.009	1	0	0	0.024	2
26/34	0.018	2	0	0	0	0	31/34	0	0	0.019	1	0.012	1
27/28	0.009	1	0.019	1	0	0	32/33	0.055	6	0.019	1	0.012	1
27/29	0.018	2	0.019	1	0	0	32/34	0	0	0.019	1	0	0
27/31	0.009	1	0	0	0	0	32/35	0	0	0	0	0.012	1
27/32	0.018	2	0	0	0	0	32/36	0.009	1	0	0	0.012	1
27/33	0	0	0	0	0.012	1	33/33	0.018	2	0	0	0	0
27/34	0	0	0.019	1	0	0	33/34	0.009	1	0	0	0	0
27/35	0	0	0	0	0	0	33/35	0	0	0	0	0.012	1
27/36	0	0	0	0	0	0	33/36	0	0	0	0	0	0
28/28	0.027	3	0.038	2	0.012	1	34/34	0	0	0	0	0.012	1
28/29	0.027	3	0.019	1	0.012	1	35/36	0	0	0.019	1	0	0
28/30	0.009	1	0	0	0	0							

Udege can be due to the assimilation of the latter and the interpopulation gene exchange.

Mendelian inheritance of ACTBP2 alleles was observed in analysis of two families with three and four children (Fig. 1).

Contrary to mini- and microsatellites D1S80 [12, 15], APOB [19], RB1 [17, 20], HUMTH01 [2–4, 11], HUMOYAR04 [3, 17], etc., the ACTBP2 allele frequency distribution showed no marked prevalence of any allele. This fact, the information parameters

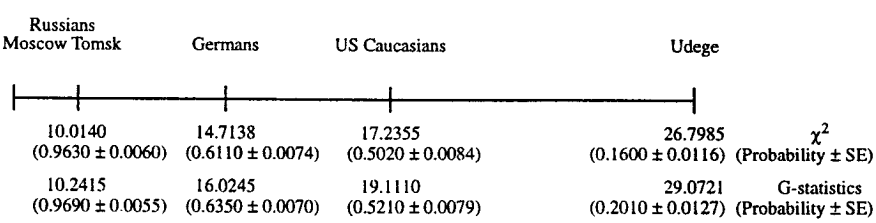

Fig. 2. Comparison of the ACTBP2 allele frequency distribution in the Moscow population with those in other populations.

Table 3. Information parameters and goodness-of-fit tests for ACTBP2

Parameters	Moscow	Tomsk	Udege
H_{obs}	0.873	0.830	0.854
H_{exp}	0.935	0.932	0.937
pM	0.008	0.009	0.008
PD	0.992	0.991	0.992
W	1.411	1.458	1.447
PIC	0.887	0.874	0.885
χ^2	90.1410	41.2095	70.1047
Probability	0.9990	1.0000	0.9990
(±SE)	(0.0010)	(0)	(0.0010)
G-statistics	120.7776	54.1729	94.2545
Probability	0.9990	1.0000	0.9990
(±SE)	(0.0010)	(0)	(0.0010)

Note: SE, standard error.

(Table 3), and successful amplification of the SE33 alleles from genomic DNA of saliva, hair roots, and blood spots (data not shown) suggest that this microsatellite can be used in personal identification.

Adding the ACTBP2 locus to the existing panel of six minisatellites results in a 126 times higher resolution and pM of $6.62 \cdot 10^{-9}$. This is sufficient for reliable personal identification in a population of 161.2 million people, i.e., in the entire Russian population.

REFERENCES

1. Huang, T.H.-M., Hejtmantik, J.F., Edvards, A., Pettigrew, A.L., Herrera, C.A., Hammond, H.A., and Caskey, C.T., *Am. J. Hum. Genet.*, 1991, vol. 49, pp. 1312–1319.
2. Edwards, A., Hammond, H.A., Jin Li, Caskey, C.T., and Chakraborty, R., *Genomics*, 1992, vol. 12, pp. 241–253.
3. Hammond, H.A., Jin Li, Caskey, C.T., and Chakraborty, R., *Am. J. Hum. Genet.*, 1994, vol. 55, pp. 175–189.
4. Urquhart, A., Oldroyd, N.J., Kimpton, G.P., and Gill, P., *BioTechniques*, 1995, vol. 18, pp. 116–121.
5. Chistyakov, D.A., Efremov, I.A., Odinokova, O.I., and Nosikov, V.V., *Mol. Biol.*, 1996, vol. 30, pp. 1274–1283.
6. Chistyakov, D.A., Chelnokova, M.V., Efremov, I.A., Stupakova, M.V., Odinokova, O.I., and Nosikov, V.V., *Genetika*, 1997, vol. 33, pp. 262–268.
7. Turakulov, R.I., Chistyakov, D.A., Odinokova, O.I., and Nosikov, V.V., *Genetika*, 1997, vol. 33, pp. 829–835.
8. Polymeropoulos, M., Rath, D., Xiao, H., and Merril, C., *Nucleic Acids Res.*, 1992, vol. 20, p. 1432.
9. Warne, D., Watkins, C., Bodfish, P., Nyberg, K., and Spurr, N., *Nucleic Acids Res.*, 1991, vol. 19, p. 6980.
10. Ng, S., Gunning, P., Eddy, R., Ponte, P., Leavitt, J., Kedes, L., and Shows, L., *Cell Genet.*, 1985, vol. 40, p. 712.
11. Weigand, P., Budowle, B., Rang, S., and Brinkmann, B., *Int. J. Leg. Med.*, 1993, vol. 106, pp. 315–320.
12. Scowasch, K., Wiegand, P., and Brinkmann, B., *Int. J. Leg. Med.*, 1992, vol. 105, pp. 165–168.
13. Johns, M.B. and Paulus-Thomas, J.E., *Anal. Biochem.*, 1989, vol. 180, pp. 276–278.
14. Walish, P., Metzger, D., and Higuchi, R., *BioTechniques*, 1991, vol. 10, pp. 506–513.
15. Budowle, B., Chakraborty, R., Giusti, A.M., Eisenberg, A.J., and Allen, R.C., *Am. J. Hum. Genet.*, 1991, vol. 48, pp. 137–144.
16. Roff, D.A. and Bentzen, P., *Mol. Biol. Evol.*, 1989, vol. 6, pp. 539–545.
17. Efremov, I.A., Chistyakov, D.A., and Nosikov, V.V., *Mol. Biol.*, 1996, vol. 30, pp. 307–318.
18. Sensabaugh, C.F., *Forensic Science Handbook*, Saferstein, R., Ed., Prentice-Hall Englewood Cliffs., 1982, pp. 340–415.
19. Deka, R., Chakraborty, R., and Ferrell, R., *Genomics*, 1991, vol. 11, pp. 83–92.
20. Scharf, S.J., Bowcock, A.M., McClure, G., Klitz, W., Yandell, W., and Elrich, A.H., *Am. J. Hum. Genet.*, 1992, vol. 50, pp. 371–381.