International Congress Series 1239 (2003) 201–205

Population data on D7S820, FGA, D1S533 and D9S304 in a sample of Caucasian-Mestizos from Colombia

Juan J. Yunis [a,b,*], Oscar García [c], Sandra Moreno [a], Ion Uriarte [d], Emilio J. Yunis [a]

[a] *Servicios Médicos Yunis Turbay y Cía, Ave 22, #42-24, of 102, Bogotá, Colombia*
[b] *Departamento de Patología, Facultad de Medicina e Instituto de Genética, Universidad Nacional de Colombia Ciudad Universitaria, Calle 53, Bogotá, Colombia*
[c] *Area de Laboratorio Ertzaintza, Avda. Montevideo 3, Bilbao, Spain*
[d] *Unidad de Policía Científica, Ertzaintza, Bilbao, Spain*

1. Introduction

Short tandem repeats (STR) are widely used in forensics as well as in paternity testing [1–4]. However, before a new marker system can be introduced into forensic casework, a population database for the relevant population must be established for statistical evaluation of the evidence.

The population of Colombia with nearly 41 million people is composed of three ethnic groups [5]. The Caucasian-Mestizo population represents the majority of the population, composed mainly of inhabitants of Spanish descent and, in a minor degree, of other European, Arab and Jewish populations. However, in certain regions (Pacific coast, Caribbean coast and islands), Colombians of African origin are the predominant group [6]. The third ethnic group, the Amerindians, are located mainly in the plains, the Amazonian jungle, in some regions of the Colombian Andes (southwest) and in the northeast section of the country [7,8]. Previous studies have revealed different degrees of genetic admixture in different regions of the country [9]. In the Andean region, the Caucasian-Mestizo population predominates, showing different degrees of admixture, mainly with Amerindians. In the Pacific and Caribbean, coast individuals of African origin predominate where different degrees of admixture are present.

We have previously reported the population frequencies for several STR loci [10–13]. This report presents allele frequency data in a sample of Caucasian-Mestizos from the

* Corresponding author. Servicios Médicos Yunis Turbay y Cia, Ave 22 # 42-24, Bogotá, Colombia.
E-mail addresses: jyunis@hotmail.com (J.J. Yunis), gobies01@sarenet.es (O. García).

0531-5131/03
PII: S0531-5131(02)00562-9

Central Andean region of Colombia ($n = 919-921$) for D7S820, FGA, D1S533 and D9S304 loci.

2. Materials and methods

Whole blood was obtained after informed consent from unrelated individuals requesting paternity testing studies. Genomic DNA was extracted by the Wizard Genomic DNA isolation kit (Promega, Madison, WI) following the manufacturer's recommendations.

The FGA, D7S820, D1S533 and D9S304 loci were amplified using a quadruplex system. The reaction assay and the amplification conditions were performed using the STR Multiplex II kit (Lifecodes, Stamford, CT) according to the manufacturer's recommendations in a PTC100 thermocycler (MJResearch, Watertown, MA). The PCR products were resolved in 4% acrylamide-*bis*-acrylamide denaturing gels following the manufacturers recommendations and detected by silver nitrate staining [14]. A total of 921 unrelated individuals were analyzed for FGA, D7S820 and D1S533 and 919 unrelated individuals for D9S304.

Allele designations were made according to recommendations of the DNA Commission of the International Society for Forensic Haemogenetics [15] with the aid of the allelic ladders provided by the manufacturer.

Statistical evaluations were performed using the computer program Genetic Data Analysis (GDA) as previously described [16]. Analyses included the possible divergence from Hardy–Weinberg expectations and other parameters of forensic importance: observed and expected heterozygosities, mean exclusion chance (MEC), polymorphic information content (PIC), discrimination power (DP) and the possible associations between loci. Results demonstrate the assumption of independence within and between the loci analysed.

3. Results and discussion

The allele frequencies for FGA, D7S820, D1S533 and D9S304 are shown in Table 1. Minimum allele frequencies for PCR-based loci, based on statistical and population genetics theory [17,18], were also determined (Table 1). Thus, a greater confidence of the DNA profile frequency estimates can be attained with current size databases.

Results of the different test procedures for testing the correspondence of the genotype frequencies with their HWE proportions are shown in Table 2. The genotype frequency distributions for most of the loci do not deviate from HWE expectations based on χ^2 test and the exact test (in all cases, the data were shuffled 2000 times).

Table 2 shows several statistical parameters of forensic importance, such as expected and observed heterozygosities, mean exclusion chance (MEC), mean paternity exclusion, polymorphic information content (PIC) and discrimination power (DP).

An interclass correlation test analysis demonstrated that there is no evidence for correlation between the alleles at any of the pairs of loci (Table 3) and support the view that the use of the product rule would provide a good approximation of the estimate of the rarity of a multiple locus profile.

Table 1
Observed allele frequency for STR loci

	D7S820	FGA	D1S533	D9S304
	$N=921$	$N=921$	$N=921$	$N=919$
4				0.0560
5				0.0011
6				0.0185
7	0.0119		0.0114	0.0076
8	0.1064		0.0706	0.3455
9	0.0803		0.0852	0.0881
10	0.2752		0.0445	0.0354
11	0.2980		0.0820	0.1104
12	0.1895		0.1971	0.2133
13	0.0364		0.2682	0.1001
14	0.0022		0.1580	0.0201
15			0.0738	0.0033
16		0.0011	0.0092	0.0005
17		0.0043		
18		0.0103		
19		0.0700		
20		0.1026		
21		0.1238		
22		0.1281		
22.2		0.0005		
23		0.1205		
24		0.1623		
24.2		0.0022		
25		0.1607		
26		0.0901		
27		0.0201		
28		0.0027		
29		0.0005		
Min. freq.	0.0033	0.0035	0.0035	0.0034

We have previously reported the allele frequencies and other parameters of forensic importance for D7S820 [12]. The results obtained in the present study do not show any statistically significant difference with those obtained previously for the same locus. In addition, the results obtained for FGA by using the STR Multiplex II kit (Lifecodes), do

Table 2
HWE tests on the analysed loci and statistical parameters of forensic importance

	D7S820	FGA	D1S533	D9S304
χ^2 test	0.6075	0.6170	0.1905	0.2310
Exact test	0.4755	0.4520	0.1630	0.1335
$H_{obs.}$	0.7655	0.8512	0.8404	0.8074
$H_{exp.}$	0.7807	0.8779	0.8381	0.8004
MEC	0.5366	0.6973	0.6760	0.6128
PIC	0.7471	0.8647	0.8188	0.7766
DP	0.9184	0.9723	0.9537	0.9338

Table 3
Two-loci interclass correlation test for the analysed loci (p values)

	D7S820	FGA	D1S533	D9S304
D7S820	–			
FGA	0.8315	–		
D1S533	0.8445	0.7960	–	
D9S304	0.6875	0.9275	0.7700	–

not show any statistically significant difference from those found using the Powerplex 2.1 system (Promega) obtained by us (see accompanying paper in this volume).

In conclusion, a Colombian population database has been established for the D7S820, FGA, D1S533 and D9S304 loci. The combined power of exclusion is estimated as 98.24% and the combined power of discrimination is 99.9993%. These four STR systems have been shown to be a useful tool for personal identification. The allele frequency data can be used for deriving estimates of multiple locus profile frequencies for identity testing purposes using the product rule, as well as for paternity testing calculations.

References

[1] A. Edwards, A. Civitello, H.A. Hammond, C.T. Caskey, DNA typing and genetic mapping with trimeric and tetrameric tandem repeats, Am. J. Hum. Genet. 49 (1991) 746–756.
[2] A. Edwards, H.A. Hammond, L. Jin, C.T. Caskey, R. Chakraborty, Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups, Genomics 12 (1992) 241–253.
[3] C. Puers, H.A. Hammond, L. Jin, C.T. Caskey, J.W. Schumm, Identification of repeat sequence heterogeneity at the polymorphic STR locus HUMTH01[AATG]n and reassignment of alleles in population analysis using a locus-specific allelic ladder, Am. J. Hum. Genet. 53 (1993) 953–958.
[4] H.A. Hammond, L. Jin, Y. Zhong, C.T. Caskey, R. Chakraborty, Evaluation of 13 short tandem repeat loci for use in personal identification applications, Am. J. Hum. Genet. 55 (1994) 175–189.
[5] Departamento Administrativo Nacional de Estadística, DANE (1996). Censo 1993, Resumen Nacional. Población ajustada y proyecciones, Santa Fé de Bogotá, República de Colombia.
[6] Departamento Administrativo Nacional de Estadística, DANE (1998). Los grupos étnicos de Colombia en el censo de 1993, Memorias. Santa Fé de Bogotá, República de Colombia.
[7] R. Arango, E. Sanchez, Los pueblos indígenas de Colombia (Guía etnográfica de Colombia para actividades de planeación), Programa N.U.D, y D.N.P. Departamento Nacional de Planeación, Bogotá, Colombia, 1988.
[8] E.J. Yunis, H. Ossa, M. Salazar, M.B. Delgado, R. Deulofeut, A. de la Hoz, D.H. Bing, O. Ramos, E.J. Yunis, Major histocompatibility complex class II alleles and haplotypes and blood groups of four Amerindian tribes of northern Colombia, Hum. Immunol. 41 (1994) 248–258.
[9] C. Sandoval, A. De la Hoz, E. Yunis, Estructura genética de la población Colombiana. Análisis de Mestizaje, Rev. Fac. Med. UN Col. 41 (1993) 3–14.
[10] J.J. Yunis, O. Garcia, I. Uriarte, E.J. Yunis, Population data on 6 short tandem repeat loci in a sample of Caucasian-Mestizos from Colombia, Int. J. Legal Med. 113 (2000) 175–178.
[11] J.J. Yunis, O. Garcia, A. Baena, G. Arboleda, I. Uriarte, E.J. Yunis, Population frequency of short tandem repeats loci D18S849, D3S1744, and D12S1090 in Caucasian-Mestizo and African descent populations of Colombia, J. Forensic Sci. 45 (2000) 429–431.
[12] J.J. Yunis, O. Garcia, I. Uriarte, E.J. Yunis, Population data on D16S539, D7S820, D13S317, LPL, F13B and D1S80 loci in a sample of Caucasian-Mestizos from Colombia, Forensic Sci. Int. 115 (2001) 117–118.
[13] J.J. Yunis, A. Baena, O. Garcia, I. Uriarte, E.J. Yunis, Population data of F13AO1, FES/FPS, VWA, CSF1PO, TPOX and THO1 short tandem repeat loci in a sample of African descent individuals of Colombia, Forensic Sci. Int. 117 (2001) 235–236.

[14] P. Martin, A. Alonso, B. Budowle, C. Albarrán, O. Garcia, M. Sancho, Spanish population data on 7 tetrameric short tandem repeat loci, Int. J. Legal Med. 108 (1995) 145–149.
[15] DNA Recommendations, Report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems, Int. J. Legal Med. 107 (1994) 159–160.
[16] O. García, P. Martín, B. Budowle, C. Albarrán, A. Alonso, Basque Country autochthonous population data on 7 short tandem repeat loci, Int. J. Legal Med. 111 (1998) 162–164.
[17] B. Budowle, K.L. Monson, R. Chakraborty, Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci, Int. J. Legal Med. 108 (1996) 173–176.
[18] B.S. Weir, Independence of VNTR alleles defined by fixed bins, Genetics 130 (1992) 873–887.