



Population data on Powerplex 2.1 (FGA, vWA, TPOX, THO1, Penta E, D18S51, D21S11, D3S1358, D8S1179) and Gammastar (D16S539, D7S820, D13S317, D5S818) in a sample of Caucasian-Mestizos from Colombia

Juan J. Yunis^{a,b,*}, Oscar García^c, Sandra Moreno^a, Cielo Pineda^a,
Claudia Rodríguez^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 3 ethnic groups [5]. The Caucasian-Mestizo population represents the majority of the population, composed mainly 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), the Colombians of African origin are the main 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

* Corresponding author. Servicios Médicos Yunis Turbay y Cía. Ave. 22 #42-24, of 102, Bogotá, Colombia.
E-mail addresses: jyunis@hotmail.com (J.J. Yunis), gobies01@sarenet.es (O. García).

Table 1
Observed allele frequency for STR loci

	VWA (N=267)	TPOX (N=262)	TH01 (N=250)	FGA (N=265)	Penta E (N=247)	D18S51 (N=254)	D21S11 (N=267)	D3S1358 (N=267)	D8S1179 (N=267)	D16S539 (N=491)	D7S820 (N=491)	D13S317 (N=491)	D5S818 (N=491)
5					0.0324								
6		0.0095	0.3120										
7		0.0038	0.2560		0.0789						0.0112	0.0010	0.0407
8		0.4695	0.0720		0.0162				0.0094	0.0143	0.1029	0.0947	0.0112
9		0.0553	0.1540		0.0101				0.0056	0.1477	0.0784	0.1690	0.0774
9.3			0.1880										
10		0.0458	0.0180		0.0567	0.0079			0.0655	0.1466	0.2811	0.0662	0.0662
11		0.2901			0.0911	0.0118			0.0918	0.2637	0.2984	0.2281	0.3859
12		0.1202			0.1923	0.1161			0.1086	0.2678	0.1823	0.2576	0.2912
13	0.0037	0.0057			0.1073	0.0965		0.0037	0.3221	0.1487	0.0407	0.1191	0.1171
13.2						0.0039							
14	0.0880				0.0870	0.1634		0.0936	0.2397	0.0092	0.0051	0.0601	0.0092
15	0.1086				0.1053	0.1732		0.3708	0.1292	0.0020		0.0041	0.0010
16	0.3296				0.0526	0.1299		0.2659	0.0281				
17	0.2566			0.0019	0.0263	0.1220		0.1498					
18	0.1685			0.0038	0.0364	0.0689		0.1067					
19	0.0393			0.0585	0.0223	0.0394		0.0094					
20	0.0056			0.0868	0.0405	0.0315							
21				0.1340	0.0263	0.0098							
22				0.1528	0.0121	0.0118							
22.2				0.0038									
23				0.1377	0.0040	0.0079							

24				0.1604	0.0020	0.0059								
25				0.1377										
26				0.0887						0.0037				
27				0.0245						0.0187				
28				0.0038						0.0880				
29										0.1873				
29.2										0.0075				
30				0.0038						0.2959				
30.2										0.0281				
31										0.0749				
31.2										0.0918				
32										0.0243				
32.2										0.1124				
33										0.0056				
33.2										0.0506				
34										0.0037				
34.2										0.0019				
35										0.0037				
35.2										0.0019				
45.2				0.0019										
Min. freq	0.0106	0.0106	0.0112	0.0132	0.0131	0.0122	0.0110	0.0106	0.0109	0.0061	0.0060	0.0061	0.0059	
χ^2 test	0.2845	0.7845	0.6790	0.0560	0.2380	0.6790	0.2170	0.7360	0.9730	0.4220	0.2045	0.4005	0.2455	
Exact test	0.1030	0.7875	0.5140	0.1365	0.0880	0.2245	0.0675	0.7170	0.9750	0.4035	0.3035	0.5905	0.5300	
$H_{\text{obs.}}$	0.7341	0.7023	0.7200	0.9434	0.8907	0.8504	0.7903	0.7416	0.7790	0.7841	0.7810	0.8227	0.7410	
$H_{\text{exp.}}$	0.7775	0.6770	0.7741	0.8772	0.9077	0.8826	0.8401	0.7505	0.7982	0.7938	0.7719	0.7862	0.7373	
MEC	0.4829	0.4318	0.4599	0.8846	0.7764	0.6956	0.5811	0.4955	0.5607	0.5699	0.5480	0.5736	0.4882	
DP	0.9176	0.8501	0.9135	0.9653	0.9829	0.9714	0.9557	0.9009	0.9338	0.9241	0.9179	0.9459	0.8981	

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 Central Andean region of Colombia for Powerplex 2.1 (FGA, TPOX, D8S1179, vWA, Penta E, D18S51, D21S11, D3S1358) ($N=247$ –267) and for GammaStar loci (D16S539, D7S820, D13S317 and D5S818) ($N=491$) (Promega, Madison, WI).

2. Materials and methods

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

PCR amplification was performed according to the manufacturer's recommendations using the Powerplex 2.1 and the GammaStar systems (Geneprint Systems, Promega). The PCR products were resolved in 5% Long Ranger denaturing gels for Powerplex 2.1 and in 4% acrylamide-*bis*-acrylamide denaturing gels from GammaStar and detected in a Hitachi FMBIO II scanner.

Allele designations were made according to recommendations of the DNA Commission of the International Society for Forensic Haemogenetics [14] 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 [15]. 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.

3. Results and discussion

The observed allele frequencies and the results of the different test procedures for testing the correspondence of the genotype frequencies with their HWE proportions for the 13 STR loci are shown in Table 1.

The genotype frequency distributions for most of the loci do not deviate from HWE expectations based on the χ^2 test and the exact test (in all cases, the data were shuffled 2000 times). Minimum allele frequencies for PCR-based loci based on statistical and population genetics theory were determined [16,17].

Table 1 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 analysed (data not shown) 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.

We have previously reported the allele frequencies and other parameters of forensic importance for vWA, TPOX, and THO1 [10] and D16S539, D7S820, D13S317 [12]. The results obtained in the present study do not show any statistically significant difference with those obtained previously for the same loci. In addition, the results obtained for FGA and D7S820 by using the STR Multiplex II kit (Lifecodes), do not show any statistically significant difference from those found using the Powerplex 2.1 system (Promega).

In conclusion, a Colombian population database has been established for the analysed systems. The combined power of exclusion is estimated as 99.9995% and the combined power of discrimination is >99.99999%. 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. Censo 1993, Resumen Nacional, Población ajustada y proyecciones, Santa Fé de Bogotá, República de Colombia, 1996.
- [6] Departamento Administrativo Nacional de Estadística, DANE. Los grupos étnicos de Colombia en el censo de 1993, Memorias, Santa Fé de Bogotá, República de Colombia, 1998.
- [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] J.J. Yunis, H. Ossa, M. Salazar, M.B. Delgado, R. Deulofeut, A. de la Hoz, D.H. Bing, O. Ramos, E.J. Yunis, 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. U. N. 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 Africans 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] 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.
- [15] 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.
- [16] 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.
- [17] B.S. Weir, Independence of VNTR alleles defined by fixed bins, *Genetics* 130 (1992) 873–887.