Int J Legal Med (2000) 113:175–178

SHORT COMMUNICATION

J. J. Yunis · O. García · I. Uriarte · E. J. Yunis

Population data on 6 short tandem repeat loci in a sample of Caucasian-Mestizos from Colombia

Received: 24 September 1998 / Received in revised form: 22 December 1998 / Accepted: 11 January 1999

Abstract Blood samples from 409–452 unrelated Colombian Caucasian-Mestizo individuals were amplified and typed for six short tandem repeat (STR) markers (HUMF13A01, HUMFES/FPS, HUMVWA, HUMCSF1PO, HUMTPOX, HUMTH01). The allele frequencies, genotype frequencies, heterozygocity, mean paternity exclusion chance, polymorphism information content, discrimination power, assumption of independence within and between loci and Hardy Weinberg equilibrium were determined. The results demonstrate that all markers conform to Hardy-Weinberg equilibrium expectations. In addition, the results demonstrate the assumption of independence within and between the loci analysed. The mean exclusion chance (MEC) was 0.9851 for all six STR loci analysed and the discrimination power (DP) was 0.9999973. Therefore, this Colombian population database can be used in identity testing to estimate the frequency of a multiple PCR-based locus DNA profile in forensic cases as well as in paternity testing.

Key words Colombia · PCR · Hardy-Weinberg equilibrium · Linkage equilibrium · Paternity testing

Introduction

Short tandem repeats (STR) are widely used in forensics as well as in paternity testing [1–4]. However, a population database for the relevant population must be established for statistical evaluation of the evidence. Few STR studies exist in South America native and mestizo populations [5–7]. Extensive work may be required in this field due to different genetic admixture processes that took place in different South American countries since the discovery of America in 1492. On the other hand, extensive studies of STR profiles have been carried out for the ancestral Spanish and Portuguese populations [8–13].

The population of Colombia with nearly 41 million people is composed of three ethnic groups [14]. 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 [15]. 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 [16, 17]. Previous studies have revealed different degrees of genetic admixture in different regions of the country [18]. In the Andean region the Caucasian-Mestizo population predominates showing different degrees of admixture, mainly with Amerindians [18]. Thus, a gradient of admixture was found where the Amerindian component is stronger in the southwest section decreasing towards the north section of the Andes where the Caucasian component is stronger. In the Pacific and Caribbean coast individuals of African origin predominate where different degrees of admixture are present [18].

To our knowledge, this study represents the first report of STR allele frequencies obtained from a Caucasian/Mestizo population sample from the Andean Region of Colombia (n = 409–452) for the six STR loci. HUMTH01 (11p15.5), HUMTPOX (2p13), HUMCSF1PO (5q33.5-q34), HUMVWA (12p12-pter), HUMFES/FPS (15q25-qter) and HUMF13A01 (6p24-p25).

J. J. Yunis (✉)
Instituto de Genética, Universidad Nacional de Colombia,
Ciudad Universitaria calle 53,
Santa Fé de Bogotá, D.C. Colombia
e-mail: jyunis@hotmail.com

J. J. Yunis · E. J. Yunis
Servicios Médicos Yunis Turbay y Cia. Ave 22 # 42–24,
Santa Fé de Bogotá, D.C. Colombia

O. Garcia · I. Uriarte
Area de Laboratorio Ertzaintza, C/ Avda. Montevideo 3,
E-48002 Bilbao, Spain

Materials and methods

Sample preparation

Whole blood was obtained after informed consent from unrelated individuals requesting paternity testing studies. Genomic DNA was extracted by the Quick light DNA isolation kit (Lifecodes Corporation, Stanford, Conn.) or by the Wizard Genomic DNA isolation kit (Promega Corporation, Madison, Wis.) following the manufacturer's recommendations.

PCR amplification and typing

The HUMTH01, HUMTPOX, HUMCSF1PO (CTT) and HUMVWA, HUMF13A01, HUMFES/FPS (FFv) loci were amplified using two triplex systems. The reaction assay and the amplification conditions were performed using GenePrint STR systems (Promega Corporation) according to the manufacturer's recommendations in a PTC100 thermocycler (MJResearch, Watertown, Mass.). The PCR amplification conditions were identical to those recommended by the manufacturer for the PE480 thermal cycler. The PCR products were resolved in 4% acrylamide-bis-acrylamide denaturing gels following the manufacturers recommendations and detected by silver nitrate staining [11].

Allele designations were made according to recommendations of the DNA Commission of the International Society for Forensic Haemogenetics [19] with the aid of the allelic ladders provided by the manufacturer. A total of 410 unrelated individuals were analysed for HUMCSF1PO, 414 for HUMTPOX, 409 for HUMTH01, 448 for HUMF13A01, 450 for HUMFES/FPS and 452 for HUMVWA.

Statistical analysis

Statistical evaluations were performed using a HWE-analysis software package (HWE-Analysis, Version 3.3. Christoph Puers, Institute of Legal Medicine, University of Münster). Analyses included the possible divergence from Hardy-Weinberg expectations and other parameters of forensic importance such as observed and expected heterozygosities [20], mean exclusion chance (MEC) [21], mean paternity exclusion probability (MEP) [22], polymorphic information content (PIC) [23] and discrimination power (DP) [24]. The possible associations between loci were tested using the computer program GDA (Lewis PO, Zaykin D 1996. Genetic Data Analysis. Software for the analysis of discrete genetic data. Computer program distributed by the authors. http//chee.unm.edu/gda/).

Results and discussion

The observed allele frequencies for the six STR loci in the Colombian Caucasian/mestizo population sample are shown in Table 1. The results of the different procedures for testing the correspondence of the genotype frequencies with the HWE proportions are shown in Table 2. The genotype frequency distributions for most of the loci do not deviate from HWE expectations based on the χ^2-test, the logarithmic likelihood ratio (G) test and the exact test [25] (in all cases, the data were shuffled 2000 times). HUMFES/FPS and HUMTPOX show differences with the χ^2-test, (marked with an asterisk) but this is not as meaningful test for HWE as the exact test. In addition, the test based on the number of distinct genotypes observed in the sample population shows that the observed numbers of distinct heterozygote and homozygote genotypes [26] are in accordance with their respective HWE predictions (data not shown).

Minimum allele frequencies for PCR-based loci, based on statistical and population genetics theory [27–29], were determined (Table 1). Thus, a greater confidence of the DNA profile frequency estimates can be attained with current size databases.

Table 1 Observed allele frequencies for six STR loci n Caucasian-Mestizos from Colombia

Allele	TH01 (n = 409)	TPOX (n = 414)	CSF1PO (n = 410)	VWA (n = 452)	FES/FPS (n = 450)	F13A01 (n = 448)
3.2						0.21987
4						0.09040
5	0.00367					0.19978
6	0.37286	0.00725				0.19085
7	0.23227	0.00121	0.00854		0.00444	0.26339
8	0.09046	0.49034	0.01341		0.00889	0.00335
9	0.14059	0.06280	0.02805		0.00222	
9.3	0.14059					
10	0.01956	0.04348	0.23659		0.20000	
11		0.29469	0.25488		0.52000	
12		0.09058	0.38415		0.18111	0.00335
13		0.00966	0.06341		0.07667	0.00223
14			0.00854	0.06858	0.00667	0.00558
15			0.00244	0.06969		0.01116
16				0.32854		0.01004
17				0.31195		
18				0.16925		
19				0.04204		
20				0.00996		
Minimum frequency	0.00731	0.00660	0.00690	0.00650	0.00600	0.00670

n = sample size

Table 2 HWE tests on the loci analysed

	TH01	TPOX	CSF1PO	VWA	FES/FPS	F13A01
χ^2 test	0.3180	0.0130*	0.0755	0.4370	0.0335*	0.2575
G test	0.3490	0.2795	0.4065	0.4380	0.0755	0.2575
Exact test	0.4025	0.1115	0.2040	0.5490	0.0785	0.2490

* Number of random shuffles performed for all tests: 2000

Table 3 Statistical parameters of forensic importance

	H_{obs}	H_{exp}[a]	MEC	MEP	PIC	DP
TH01	0.7897	0.7602	0.5450	0.5273	0.7244	0.9018
TPOX	0.6667	0.6593	0.4157	0.3682	0.6074	0.8203
CSF1PO	0.7073	0.7272	0.4887	0.4716	0.6806	0.8771
VWA	0.7611	0.7555	0.5361	0.5192	0.7165	0.8995
FES/FPS	0.6400	0.6515	0.4111	0.3573	0.6046	0.8299
F13A01	0.7879	0.7984	0.5978	0.5960	0.7666	0.9277

[a] Expected heterozygosity is an unbiased estimate

Table 4 Two-loci interclass correlation test for the analyzed loci (*p* values)

TH01	–					
TPOX	0.2300	–				
CSF1PO	0.1310	0.1335	–			
VWA	0.2590	0.0330*	0.3315	–		
FES/FPS	0.3250	0.1205	0.1405	0.5615	–	
F13A01	0.7815	0.0930	0.3810	0.4735	0.0820	–
	TH01	TPOX	CSF1PO	VWA	FES/FPS	F13A01

* Number of random shuffles performed for all tests: 2000

Table 3 shows several statistical parameters of forensic importance, such as expected and observed heterozygosities, mean exclusion chance (MEC), mean paternity exclusion, mean paternity exclusion probability (MEP), polymorphic information content (PIC) and discrimination power (DP). The mean exclusion chance (MEC) for the six STR loci analysed was 0.9851 and the discrimination power (DP) was 0.9999973.

An analysis was performed to determine whether there were any detectable associations between any pairs of the loci (Table 4). There was one example of departure out of a total of 15 interclass correlation tests, which is approximately 6.6% of the comparisons and near to the expected. The departure from expectations (marked with an asterisk) was between the vWA/TPOX loci ($p = 0.033$). After applying a Bonferroni correction (used for correcting when multiple test are performed on a population sample) [30], the data suggest that overall there is little evidence for departures from independence for the sample population.

A comparison of the allele frequencies in the population under study with those of previous studies in several Spanish populations (Galicia [8, 12], Northeast Spain [9], Basque Country [10], and Central Spain [11]) revealed significant differences ($p < 0.001$) in all the loci analysed. These results are likely due to the admixture process during the last 500 years.

In conclusion, a population database of Caucasian-Mestizos from Colombia has been established for the loci HUMTH01, HUMTPOX, HUMCSF1PO, HUMVWA, HUMFES/FPS and HUMF13A01. The combined power of exclusion is estimated as 98.51% and the combined power of discrimination is 99.99973%.

Acknowledgements We would like to thank Sandra Moreno and Yolanda Gonzalez for the invaluable technical assistance.

References

1. Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756
2. Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12: 241–253
3. Puers C, Hammond HA, Jin L, Caskey CT, Schumm JW (1993) 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 53:953–958
4. Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet 55:175–189
5. Jorquera H, Budowle B (1998) Chilean population data on ten PCR-based loci. J Forensic Sci 43:171–173
6. Gene M, Fuentes M, Huguet E, Pique E, Bert F, Corella A, Perez-Perez A, Corbella J, Moreno P (1998) Quechua amerindian population characterized by HLA-DQ alpha, YNZ22, 3′ APO B, HUMTHO1, HUMVWA31A polymorphisms. J Forensic Sci 43:403–405
7. Sala A, Penacino G, Corach D (1998) Comparison of allele frequencies of eigth STR loci from Argentinian Amerindian and European populations. Hum Biol 70:937–947
8. Pestoni C, Lareu MV, Rodrigues MS, Muñoz I, Barros F, Carracedo A (1995) The use of the STRs HUMTHO1, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data for Galicia (NW Spain). Int J Legal Med 107:283–290
9. Crespillo M, Luque JA, Fernández R, Ramirez E, Garcia P, Valverde JL (1997) Allele frequency distribution of 13 PCR-based systems in a population from North-East Spain. Int J Legal Med 110:223–225
10. Garcia O, Martin P, Budowle B, Uriarte J, Albarrán C, Alonso A (1998) Basque country autochthonus population data on 7 short tandem repeat loci. Int J Legal Med 111:162–164
11. Martin P, Alonso A, Budowle B, Albarrán C, Garcia O, Sancho M (1995) Spanish population data on 7 tetrameric short tandem repeat loci. Int J Legal Med 108:145–149
12. Luis J, Liste I, Caeiro B (1994) Improved conditions for genotype diagnosis of a STR of the HTPO locus. In: Bär W, Fiori A, Rossi U (eds) Advances in Forensic Haemogenetics 5. pp 366–368
13. Gene M, Carracedo A, Huguet E, Perez-Perez A, Moreno P (1998) Population genetics of the D12S391, CSF1PO and TPOX loci in Catalonia (Northeast Spain). Int J Legal Med 111:52–54
14. Departamento Administrativo Nacional de Estadística, DANE (1996) Censo 1993, Resumen Nacional. Población ajustada y proyecciones. (in Spanish). Santa Fé de Bogotá, República de Colombia

15. Departamento Administrativo Nacional de Estadística, DANE (1998) Los grupos étnicos de Colombia en el censo de 1993 (in Spanish). Memorias. Santa Fé de Bogotá, República de Colombia
16. Arango R, Sanchez E (1988) Los pueblos indígenas de Colombia (Guía etnográfica de Colombia para actividades de planeación) (in Spanish). Programa N.U.D y D.N.P. Departamento Nacional de Planeación. Bogotá, Colombia
17. Yunis JJ, Ossa H, Salazar M, Delgado MB, Deulofeut R, de la Hoz A, Bing DH, Ramos O, Yunis EJ, Yunis EJ (1994) Major histocompatibility complex class II alleles and haplotypes and blood groups of four Amerindian tribes of northern Colombia. Hum Immunol 41:248–258
18. Sandoval C, De la Hoz A, Yunis E (1993) Estructura genética de la población colombiana. Análisis de mestizaje (in Spanish). Rev Fac Med UN Col 41:3–14
19. DNA recommendations (1994) 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 107:159–160
20. Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590
21. Wiegand P, Budowle B, Rand S, Brinkmann B (1993) Forensic validation of the STR systems SE33 and TC11. Int J Legal Med 105:315–320
22. Brenner C, Morris J (1990) Paternity index calculations in single locus hypervariable DNA probes. validation and other studies. Proceedings of the international symposium on human identification. Promega Corporation, Madison Wis., pp 21–53
23. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331
24. Jones DA (1972) Blood samples: probabilities of discriminations. J Forensic Sci Soc 12:355–359
25. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361–372
26. Chakraborty R (1993) A class of population genetic questions formulated as the generalized occupancy problem. Genetics 134:953–958
27. Budowle B, Monson KL, Chakraborty R (1996) Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. Int J Legal Med 108:173–176
28. Chakraborty R (1992) Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. Hum Biol 64:141–159
29. Weir BS (1992) Independence of VNTR alleles defined by fixed bins. Genetics 130:873–887
30. Weir BS (1990) Multiple tests. In Genetic Data Analysis. Sinauer Associates, Sunderland, Mass., pp 109–110