

## TECHNICAL NOTE

Noelia Turret,<sup>1</sup> M.D.; Jorge López Camelo,<sup>2</sup> Ph.D.; and Lidia Vidal-Rioja,<sup>3</sup> M.D., Ph.D.

# Allele Frequencies of Six STR Loci in Argentine Populations\*

**REFERENCE:** Turret N, Camelo JL, Vidal-Rioja L. Allele frequencies of six STR loci in Argentine populations. *J Forensic Sci* 1999;44(6):1265–1269.

**ABSTRACT:** Allele frequencies of six short tandem repeat (STR) loci were determined in a Caucasian urban sample of La Plata city and three Amerindian sample populations of Argentina. Allele frequencies showed differences between urbans and Amerindians, and among Amerindians as well. The degree of genetic differentiation of subpopulations was mainly due to the Amerindian contribution. Mapuche, Mocovi, and pooled Amerindian populations showed little evidence of HW disequilibrium, and association of alleles. In the urban sample, there is no evidence of population substructuring. Forensic probabilities of exclusion and matching showed high differences between the population groups. Finally, La Plata sample did not show differences with Caucasians from other geographic regions.

**KEYWORDS:** forensic science, DNA typing, Argentina, Caucasian, Amerindian, population genetics, TH01, CSF1PO, TPOX, F13B, F13A01, FESFPS, allele frequencies, STR loci, Argentine Caucasian and Amerindian populations, forensic probabilities

The Argentine population comprises predominantly European descendants with two to five generations born in the country. This structure is particularly valid for the largest urban areas of Buenos Aires, Córdoba, Mendoza, and Santa Fe provinces. The genetic composition of Argentine population is also influenced by genetic mixture with native aborigine populations. Although the Amerindian component has not the same quantitative occurrence than the one observed for other South American countries, it constitutes a well defined section of our people composition (1). In Argentina, the city of La Plata is an urban center surrounded by important industrial settlements and a wide university area which attract immigration from other provinces. For this reason, during the last 40 years, La Plata has almost doubled its population not only due to the natural increase of its inhabitants but also to immigration. This report presents allele frequency data of the STR loci HUMTH01, HUMTPOX, HUMF13B, HUMF13A01, HUMCSF1PO, and

HUMFESFPS for an urban sample of La Plata city and three samples of the major Argentine Amerindian populations. The survey was performed by PCR amplification. The level of variation among and within all four populations and the potential usefulness of these loci for paternity and identification testing were also examined. Moreover, to test the genetic homogeneity of the current La Plata population, it was compared with Caucasian samples from other geographic populations.

## Materials and Methods

### Population Sample

Population studies were carried out on 100 unrelated donors of the La Plata city and three Argentine Amerindian populations comprising 50 Mapuche individuals, 36–50 Mocovies, and 50 Wichis. The three Amerindian samples were also analyzed as a group referred to pooled Amerindians. The urban sample includes workers undergoing blood testing for job admission, suspects and victims from forensic cases, and laboratory personnel. Mapuche sample belongs to localities of Cerro Policía and Aguada Guzmán, province of Río Negro, South Argentina. The two other Amerindian groups inhabit the North of Argentina: the Wichis, proceeding from Santa Victoria, province of Salta, and the Mocovies, from Colonia Dolores, in Santa Fe.

### DNA Analysis

Genomic DNA was isolated from 20  $\mu$ L blood samples by Chelex 100 resin (BioRad, CA) according to Walsh P et al. (2). DNA samples were amplified in monoplex PCR reactions each containing a set of commercial primers from the Gene Print-STR System, (Promega Corp., Madison, WI) for HUMTH01, HUMCSF1PO, HUMTPOX, HUMF13B, HUMF13A01, and HUMFESFPS loci. Each PCR mixture included 50 ng of DNA extract; 1.5 mM  $MgCl_2$ ; 200  $\mu$ M each dATP, dGTP, dTTP; 10  $\mu$ M dCTP; 0.08  $\mu$ Ci  $\alpha$ -<sup>32</sup>P-dCTP (NEN, Dupont, Boston, MA), 0.45 units of Taq polymerase (Promega Corp.) in a final volume of 10  $\mu$ L of 1  $\times$  Taq buffer. Amplifications were carried out in a PTC-100 programmable thermocycler (MJ Research, Inc., Walertown, MA) following cycling conditions devised by Promega Corp. for each primer set. PCR products were electrophoresed through 6% polyacrylamide, 7 M urea on a Sequi-Gen cell (Bio-Rad, Hercules, CA) and the alleles detected by autoradiography. Alleles assignment was made by size comparison with appropriate allelic ladders from the Promega Corp. STR kits.

<sup>1</sup> Legal medicine specialist, Laboratorio de Identificación Genética; <sup>2</sup> Population genetics specialist, Departamento de Epidemiología Genética; <sup>3</sup> Medical geneticist, Head, Departamento de Citogenética Molecular and Laboratorio de Identificación Genética, respectively, Instituto Multidisciplinario de Biología Celular (IMBICE), La Plata, Argentina.

\* This investigation was performed with grants from CIC and CONICET of Argentina.

Received 27 Aug. 1998; and in revised form 31 Dec. 1998 and 8 Feb. 1999; accepted 8 Feb. 1999.



TABLE 1—(Continued)

Population	Allele	HUMTH01		HUMCSF1PO		HUMTPOX		HUMF13B		HUMF13A01		HUMFESFPS	
		Nc	Freq	Nc	Freq	Nc	Freq	Nc	Freq	Nc	Freq	Nc	Freq
N=50*	17	—	—	—	—	—	—	—	—	—	—	—	—
Exact test			0.010		0.262		0.142		0.721		0.573		0.061
G <sup>2</sup> (df)			16.2 <sub>(15)</sub>		16.5 <sub>(18)</sub>		6.6 <sub>(21)</sub>		2.6 <sub>(15)</sub>		17.9 <sub>(42)</sub>		17.7 <sub>(21)</sub>
Likelihood			0.367		0.559		0.998		0.999		0.999		0.668
Wichi	3.2	—	—	—	—	—	—	—	—	27	0.270	—	—
	4	—	—	—	—	—	—	—	—	9	0.090	—	—
	5	—	—	—	—	—	—	—	—	49	0.490	—	—
	6	13	0.130	—	—	2	0.020	—	—	6	0.060	—	—
	7	77	0.770	—	—	—	—	—	—	9	0.090	—	—
	8	—	—	—	—	21	0.210	1	0.010	—	—	—	—
	9	3	0.030	1	0.010	1	0.010	65	0.650	—	—	—	—
	9.3	7	0.070	—	—	—	—	34	0.340	—	—	—	—
	10	—	—	29	0.290	—	—	—	—	—	—	15	0.150
	11	—	—	18	0.180	39	0.390	—	—	—	—	65	0.650
	12	—	—	51	0.510	37	0.370	—	—	—	—	11	0.110
	13	—	—	1	0.010	—	—	—	—	—	—	9	0.090
	14	—	—	—	—	—	—	—	—	—	—	—	—
	15	—	—	—	—	—	—	—	—	—	—	—	—
	16	—	—	—	—	—	—	—	—	—	—	—	—
N=50	17	—	—	—	—	—	—	—	—	—	—	—	—
Exact test			0.732		0.730		0.576		0.800		0.191		0.352
G <sup>2</sup> (df)			6.1 <sub>(15)</sub>		6.9 <sub>(18)</sub>		6.4 <sub>(21)</sub>		2.3 <sub>(15)</sub>		16.2 <sub>(42)</sub>		10.5 <sub>(21)</sub>
Likelihood			0.978		0.991		0.999		0.999		0.999		0.971

N = sample size, \* H01 (N=50), IPO (N=49), POX, 13B, A01 (N=42), FPS (N=36).  
 Nc = chromosome number, Freq = allelic frequencies.  
 Exact test = chi square test (1 df) based observed/expected homozygosity.  
 Likelihood G<sup>2</sup> test based observed/expected genotype.

is particularly evident for allele number seven at HUMTH01 locus, which shows a frequency of 77% in the Wichi population, and 22.5% in the Caucasian urban sample of La Plata.

The results of HWE tests are depicted in Table 1. The exact test of heterozygosity indicates disequilibrium at HUMTH01 ( $p = 0.010$ ) locus for Mocovies. In the pooled Amerindian sample, this test indicates disequilibrium at HUMTH01 ( $p = 0.022$ ), HUMTPOX ( $P = 0.042$ ), and HUMF13A01 ( $P = 0.038$ ) loci (not shown). The test based on the number of distinct homozygous and heterozygous genotypes detected an excess of 10–10 homozygous individuals (observed = 4, expected = 1) at HUMFESFPS locus for the Mocovi population. The  $G^2$  test, which seems to be more appropriate for small samples, detected deviations from HWE at HUMTH01 ( $P = 0.025$ ) for Mapuches, and HUMTH01 ( $P = 0.032$ ) and HUMCSF1PO ( $p = 0.049$ ) for the pool of Amerindians. Departures from HWE may be caused by factors such as population substructuring, sampling size, nonrandom mating, or disturbing forces like admixture, genetic drift, selection, endogamy, etc. (5,20,21). Our data showed 8 HWE departures out of the 4–5 possible random deviations, considering the 90 comparison and  $\alpha = 0.05$ . Deviations observed in the pooled Amerindian sample were expected and they can be explained by the Whalund’s principle (22). Other HW disequilibrium might be caused by sampling error although in Amerindian populations a biological effect could also be assumed.

Population substructuring can be detected either as deviation from Hardy-Weinberg equilibrium or as gametic phase equilibrium (21). Since the loci examined here are located in different chromosomes, we could expect no association of their alleles. In spite of this,  $Sk^2$  results show disequilibrium in the pairs CSF1PO/FESFPS ( $Sk^2 = 0.520$ , CI 95% = 0.517) for Mapuche population; CSF1PO/F13B ( $Sk^2 = 0.628$ , CI 95% = 0.478) CSF1PO/F13A01 ( $Sk^2 = 0.539$ , CI 95% = 0.479) and F13B/FESFPS ( $Sk^2 = 0.533$ ,

CI 95% = 0.523) for Mocovies; and TH01/CSF1PO ( $Sk^2 = 0.527$ , CI 95% = 0.501), CSF1PO/F13B ( $Sk^2 = 0.575$ , CI 95% = 0.513) CSF1PO/F13A01 ( $SK^2 = 0.446$ , CI 95% = 0.427), CSF1PO/FESFPS ( $SK^2 = 0.520$ , CI 95% = 0.505) and F13B/F13A01 ( $SK^2 = 0.510$ , CI 95% = 0.482) for the pool of Amerindian samples. The allelic association found in our samples might be interpreted as the consequence of heterogeneity within each of these populations. In this context, the level of non-Amerindian admixture found in Indian populations (23,24) could possibly explain these findings, though the size of the sample should not rule out. With the purpose of evaluating whether the HW and gametic phase disequilibrium found in the aborigine samples have any effect on forensic statistics, we compared the estimated genotype frequencies by applying the counting method as in Scholl S et al. 1996 (25), which can be used when no assumption of independence is presumed. The results indicated that such disequilibrium has little impact on the genotype frequencies and does not seem to be relevant for forensic purposes. On the other hand, all tests performed with the Caucasian urban sample of La Plata showed that it is homogeneous and provided no evidence of population substructuring. These results also support the use of observed allele frequencies to estimate genotype frequencies, whereas the product rule might be applied to estimate the frequencies of a multilocus profile.

Tables 2 and 3 show gene diversity values  $h$  for the different loci, the level of variation among and within the groups, and the  $F_{ST}$  values for each polymorphic marker. The expected heterozygosity values for each locus varied in all four populations and no evidence of locus characteristic profile was detected. Accordingly, for La Plata sample, HUMTH01 ( $h = 0.796$ ) was the most variable locus while HUMTPOX was the less variable one ( $h = 0.635$ ). Conversely, for the Wichi population, the lowest value corresponded to the HUMTH01 ( $h = 0.384$ ) locus and the highest one to HUMTPOX ( $h = 0.678$ ). The genetic diversity ( $F_{st}$ ) of

TABLE 2—Gene diversity analysis of six STR loci in four Argentine population samples.

Locus	Gene Diversity (h)				Source of Variation in the Four Populations			Source of Variation in Amerindian		
	Urban	Mapuche	Mocovi	Wichi	Within	Among	F <sub>ST</sub>	Within	Among	F <sub>ST</sub>
TH01	0.796 ± 0.040	0.618 ± 0.069	0.652 ± 0.067	0.384 ± 0.069	0.649	0.073	0.101	0.551	0.074	0.118
CSF1PO	0.732 ± 0.044	0.746 ± 0.062	0.725 ± 0.064	0.624 ± 0.069	0.712	0.011	0.015	0.698	0.014	0.019
TPOX	0.635 ± 0.048	0.664 ± 0.067	0.715 ± 0.070	0.678 ± 0.066	0.661	0.037	0.053	0.679	0.028	0.039
F13B	0.724 ± 0.045	0.584 ± 0.069	0.640 ± 0.074	0.462 ± 0.074	0.626	0.035	0.053	0.558	0.031	0.052
F13A01	0.789 ± 0.041	0.776 ± 0.059	0.753 ± 0.067	0.668 ± 0.067	0.754	0.043	0.054	0.810	0.031	0.038
FESFPS	0.673 ± 0.047	0.632 ± 0.068	0.675 ± 0.078	0.534 ± 0.071	0.635	0.009	0.014	0.607	0.011	0.018
Mean	0.725 ± 0.058	0.670 ± 0.069	0.693 ± 0.041	0.558 ± 0.109			0.048			0.047

TABLE 3—Gene diversity analysis of the six STR loci in Caucasian population samples from different sources.

Locus	Gene Diversity (h) in Caucasian Samples				Source of Variation		
	1*	2	3	4	Within	Among	F <sub>ST</sub>
TH01	0.796 ± 0.040	0.783 ± 0.029†	0.773 ± 0.052	0.776 ± 0.030**	0.780	0.002	0.003
CSF1PO	0.732 ± 0.044	0.734 ± 0.031†	0.722 ± 0.036	0.722 ± 0.045††	0.727	0.001	0.001
TPOX	0.635 ± 0.048	0.618 ± 0.035†	0.585 ± 0.057	0.587 ± 0.049††	0.609	0.001	0.002
F13B	0.724 ± 0.045	0.715 ± 0.031‡	0.735 ± 0.031¶	0.711 ± 0.084‡‡	0.708	0.002	0.002
F13A01	0.789 ± 0.041	0.730 ± 0.034§	0.726 ± 0.035	0.768 ± 0.027**	0.751	0.004	0.005
FESFPS	0.673 ± 0.047	0.700 ± 0.034§	0.647 ± 0.041	0.690 ± 0.034**	0.688	0.006	0.010

\* Data from La Plata sample.

† Data from U.S. Caucasian sample (Budowle B et al. 1997).

‡ Data from U.S. Caucasian sample (Promega Manual, 1997).

§ Data from U.S. Caucasian sample (Hammond H et al. 1994).

|| Data from German Caucasians (Puers et al. (1993).

¶ Data from Nishimura D and Murray J, 1992.

\*\* Data from French Caucasians (Pftzeinger H et al. 1994).

†† Data from Swiss Caucasians (Hochmeister M et al. 1994).

‡‡ Data from Austrian Caucasians (Klitschar M and Crevenna R 1996).

TABLE 4—Statistical parameters of medical/legal interest.

Locus	Statistic	Urban	Mapuche	Mocovi	Wichi	Amerindian
TH01	Pex	0.595	0.361	0.404	0.209	0.369
	PM	0.088	0.200	0.184	0.418	0.202
CSF1PO	Pex	0.486	0.514	0.477	0.351	0.423
	PM	0.130	0.128	0.128	0.222	0.134
TPOX	Pex	0.394	0.417	0.466	0.387	0.449
	PM	0.191	0.162	0.128	0.186	0.132
F13B	Pex	0.479	0.304	0.361	0.187	0.303
	PM	0.123	0.268	0.203	0.398	0.252
F13A01	Pex	0.592	0.566	0.528	0.428	0.554
	PM	0.073	0.096	0.112	0.174	0.094
FESFPS	Pex	0.431	0.375	0.441	0.318	0.384
	PM	0.157	0.212	0.170	0.244	0.196

Pex (Power of exclusion), PM (Power of matching).

subpopulations ranged from 0.014 for HUMFESFPS locus to 0.101 for HUMTH01 locus, with an average of 0.048 (Table 2). For Amerindian  $F_{ST}$  values, the average is 0.047. Thus, the degree of genetic differentiation for the four samples analyzed is mainly due to the Amerindian contribution. On the other hand, the results of gene diversity ( $h$ ) analysis in Caucasian samples showed homogeneity among the populations with  $F_{ST}$  values ranging 0.001–0.010 (Table 3). These findings suggest that although Argentine population is mixed (26), the level of admixture in the urban sample of La Plata is not enough to anticipate differences in

DNA profile frequency when compared with other Caucasian populations.

Statistic data of medical/legal interest are presented in Table 4. Combined probability of exclusion calculation is 0.985 for La Plata sample, 0.966 for Mapuches, 0.972 for Mocovies, 0.90 for Wichis, and 0.962 for the pool of Amerindians. The combined matching probability ranged from  $\cong 1/3400$  individuals for Wichis to  $1/325,000$  for Urbans,  $\cong 1/44,000$  for Mapuches,  $1/86,000$  for Mocovies, and  $1/60,000$  for the pooled Amerindians. The high interpopulation variability registered in this study justifies the need for

local populations studies that may corroborate the relative efficiency of a genetic marker for these populations.

### Acknowledgments

We thank Dr. F. R. Carnese from the Instituto de Ciencias Antropológicas of the Buenos Aires University and Dr. P. Zukas from the Centro de Tipificación No. 10 of Santa Fe province for providing the Amerindian samples, and the anonymous donors who provided the Caucasian samples. We also appreciate the advice of Dr. Nestor O. Bianchi during this investigation, and Beatriz Tosti for the English correction of this manuscript.

### References

- Martinez Sarasola C. Nuestros Paisanos Los indios, 3rd edition. Emecé Ed. S.A. Buenos Aires, Argentina 1992;15–20.
- Walsh P, Metzger D, Russell H. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* 1992;10(4):506–13.
- Li C. First course in population genetics. Boxwood, Pacific Grove; CA 1976.
- Deka R, Chakraborty R, DeCuro S, Rothhammer F, Barton S, and Ferrell R. Characteristics of polymorphism at a VNTR locus 3' to the apolipoprotein B gene in five human populations. *Am J Hum Genet* 1992;51:1325–33.
- Chakraborty R, Fornage M, Gueguen R, Boerwinkle E. Population genetics of hypervariable loci: analysis of PCR based VNTR polymorphism within a population. In: Burke T, Dolf G, Jeffreys A, Wolff R, editors. *DNA fingerprinting: approaches and applications*. Birkhauser Verlag, Basel/Switzerland 1991;127–43.
- Sokal R, Rohlf F. *Biometry*, 2d Edition. WH Freeman and Company, San Francisco 1981.
- Gill P, Evett I. Population genetics of short tandem repeat (STR) loci. *Genetica* 1995;96:69–87.
- Chakraborty R. Detection of nonrandom association of alleles from the distribution of the number of heterozygous loci in a sample. *Genetics* 1984;108:731–919.
- Nei M. *Molecular evolutionary genetics*. Columbia University press. New York 1987;512.
- Budowle B, Smerick J, Key K, Moretti T. United States populations data on the multiplex short tandem repeat loci-HUMTH01, TPOX, and CSF1PO and the variable number tandem repeat locus D1S80. *J Forensic Sci* 1997;42(5):846–9.
- GenePrint STR System, Promega Technical Manual, 1997.
- Hammond H, Jin L, Zhong Y, Caskey C, Chakraborty R. Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am J Human Genet* 1994;55:175–89.
- Nishimura D, Murray J. A tetranucleotide repeat for the F13B locus. *Nucleic Acids Res* 1992;20–5:1167.
- Puers CH, Lins A, Sprecher C, Brinkmann B, Schumm J. Analysis of polymorphic short tandem repeat loci using well-characterized allelic ladders. Fourth International Symposium on Human Identification. Promega Corp. 1993;161–75.
- Pfitzinger H, Rousselet F, Mangin P. French experience in STR system: Caucasian population data bases, automated fluorescent quadruplex typing and forensic applications for HUMFESFPS, HUMTH01, HUMVWA31/A and HUMF13A01 loci. Fifth International Symposium on Human Identification. Promega Corp. 1994;85–94.
- Hochmeister M, Borer U, Dirnhofer Y. Swiss Caucasian population data and casework-application using PCR amplification kits. Fifth Symposium on Human Identification. Promega Corp. 1994;51–60.
- Klitschar M, Crevenna R. Validation of the STR system FXIIB for forensic purposes in an Austrian population sample. *Forensic Sci Int* 1996;81:35–42.
- Smouse P, Chakraborty R. The use of restriction fragment length polymorphisms in paternity analysis. *Am J Hum Genet* 1986;38:918–39.
- Jones DA. Blood samples: probability of discrimination. *J Forensic Sci Soc* 1972;12:355–9.
- Cavalli-Sforza L, Bodmer W. *Poblaciones Mendelianas*. In: *Genética de las poblaciones Humanas*. Omega Ed. Barcelona 1981;42–73.
- Edwards A, Hammond H, Jin L, Caskey T, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992;12:241–53.
- Wahlund S. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas* 1928;11:65–106.
- Salzano F, Callegari-Jacques S. Discontinuous genetic variability: description. In: *Evolution*. Clarendon Press. Oxford 1988.
- Bianchi N, Bailliet G, Bravi C, Carnese F, Rothhammer F, Martínez-Marignac V, et al. Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers. *Am J Physical Anthropology* 1997;102:79–89.
- Scholl S, Budowle B, Radecki K, Salvo M. Navajo, Pueblo, and Sioux population data on the loci HLA DQA1, LDLR, GYPA, HBGG, D7S8, Gc, and D1S80. *J Forensic Sci* 1996;41(1):47–51.
- Lopez-Camelo JS, Cabello PH, Dutra MG. A simple model for the estimation of congenital malformation frequency in racially mixed populations. *Brazilian J Genet* 1996;19(4):659–63.

Additional information and reprint requests:  
Lidia Vidal-Rioja, M.D.  
IMBICE  
C.C. 403  
1900 La Plata-Argentina  
Fax: (54-21) 253320  
E-mail: cimbic@netverk.com.ar