



Forensic considerations on STR databases in Argentina

U. Toscanini^a, G. Berardi^a, A. Amorim^{b,c}, A. Carracedo^d,
A. Salas^d, L. Gusmão^b, E. Raimondi^{a,*}

^a *PRICAI-FUNDACIÓN FAVALORO, Buenos Aires, Argentina*

^b *IPATIMUP, Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Portugal*

^c *Faculdade de Ciências, Universidade do Porto, Portugal*

^d *Instituto de Medicina Legal, Facultad de Medicina, Universidad de Santiago de Compostela, Galicia, Spain*

Abstract. A genetic comparison study was conducted between populations from different regions of Argentina in order to determine if a pooled population STR database could be used for general forensic purposes. Samples were from urban populations of six geographically distant provinces of Argentina (Tucumán, San Luis, La Pampa, Buenos Aires, Neuquén and Santa Cruz), and two Native American populations from the North and Northwest region of the country (Tobas and Collas). Fifteen autosomal markers were analyzed. Exact tests did not show deviations from Hardy–Weinberg equilibrium. Low F_{st} values were observed for population pairwise comparisons. However, significant differences were found when comparing Buenos Aires with Neuquén, San Luis and Santa Cruz, and Neuquén with La Pampa. Concerning Native American populations, F_{st} P -values were statistically significant when comparing Tobas and Collas with all urban populations and between them. Single locus comparisons showed significant differences when comparing Neuquén and Buenos Aires at D5S818, FGA and Penta E loci, between Buenos Aires and Santa Cruz at PENTA E locus, and between Neuquén and Santa Cruz at PENTA E locus. However, no significant differences were found between the urban samples from San Luis, Santa Cruz, Tucumán and La Pampa. When comparing urban populations with Amerindians and European populations, significant P -values were observed at 12–15 loci. The four non-differentiated urban populations were pooled and P -values showed no deviation for Hardy–Weinberg equilibrium. A combined matching probability of 2.0×10^{-18} and an a priori chance of exclusion of 0.9999995 were obtained. These results suggest that it would be possible to use a combined database for Tucumán, San Luis, La Pampa and Santa Cruz, as no significant differences were found between them. Caution should be taken concerning small isolates where Native American component could be much more relevant. When comparing

* Corresponding author. Av. Belgrano 1782, 1° SS-(1093) Capital Federal, Argentina. Tel.: +54 11 4378 1205; fax: +54 11 4383 1197.

E-mail address: eraimondi@ffavaloro.org (E. Raimondi).

Argentina urban population with two Iberian samples, some significant differences were found. Therefore, an Iberian database might not adequately represent the Argentinean genetic makeup. © 2006 Elsevier B.V. All rights reserved.

Keywords: STR; Argentina; Database; Forensic; Powerplex 16

1. Introduction

Argentina's current population is about 36 million inhabitants, with more than 14 million residing in the province of Buenos Aires and in the autonomous district. Argentinians are a fusion of different native groups with descendants of European immigrants, mainly from Italy and Spain, that arrived in Argentina during the late 19th and early 20th centuries. Several Native American groups exist nowadays in Argentina, mainly concentrated in the provinces of the Northwest, Central and South regions of the country, with a current population estimated at about 200,000. Given that the Native American and the immigrant contribution may be different across the country, we conducted a genetic comparison study for 15 autosomal STR loci in 6 urban and 2 Native American populations from Argentina, in order to determine if significant differences exist among them that could be of forensic interest.

2. Materials and methods

DNA was extracted from blood samples taken from non-related individuals living in urban areas of 6 provinces, namely, Tucumán ($N=51$), San Luis ($N=42$), La Pampa ($N=147$), Buenos Aires ($N=879$), Neuquén ($N=355$) and Santa Cruz ($N=82$), and in two native communities, namely, Tobas ($N=129$) and Collas ($N=43$), of Argentina. The 15 autosomal STR loci included in the Powerplex 16 kit were analyzed according to manufacturer's instructions (PowerPlex®16 System Technical Manual; Promega Corp.) Detection of amplified fragments were done using either an ABI PRISM 377 sequencer or an ABI PRISM 310 genetic analyzer (Applied Biosystems). Population comparisons were performed by means of Slatkin's distance (Rst) between every possible pair of populations, using the Arlequin software version 2.000 [1]. Matching probability and a priori chance of exclusion as well as frequency estimates were calculated using the Powerstats v1.2 software (Promega Corp.).

3. Results and discussion

Exact test P -values did not show deviations from Hardy–Weinberg equilibrium for both, urban and Native American populations. Low F_{st} values were observed for the population pairwise comparisons of urban populations; however, significant differences were found when comparing Buenos Aires with Neuquén, San Luis and Santa Cruz and Neuquén with La Pampa (Table 1). Concerning Native American populations, F_{st} P -values were statistically significant when comparing Tobas and Collas with every urban population and between themselves (Table 1). Single locus comparisons showed some significant differences when comparing Neuquén and Buenos Aires, namely, at D5S818, FGA and Penta E ($P=0.000$), between Buenos Aires and Santa Cruz at PENTA E ($P=0.000$) locus, and between Neuquén and Santa Cruz at PENTA E ($P=0.002$) locus. No significant differences were found at the single locus comparisons between the urban

Table 1

Fst (Slatkin's distance) (below diagonal) and Fst *P*-values (above diagonal) between pairs of populations

	BA	NQ	TU	SL	SC	LP	TO	CO
BA		0.000*	0.051	0.017*	0.001*	0.146	0.000*	0.000*
NQ	0.003		0.290	0.538	0.392	0.006*	0.000*	0.000*
TU	0.002	0.000		0.766	0.811	0.068	0.000*	0.000*
SL	0.003	0.000	0.000		0.618	0.238	0.000*	0.000*
SC	0.002	0.000	0.000	0.000		0.093	0.000*	0.000*
LP	0.000	0.001	0.002	0.000	0.001		0.000*	0.000*
TO	0.047	0.034	0.029	0.039	0.036	0.045		0.000*
CO	0.031	0.024	0.018	0.025	0.021	0.031	0.022	

BA: Buenos Aires; NQ: Neuquen; TU: Tucumán; SL: San Luis; SC: Santa Cruz; LP: La Pampa; TO: Tobas; CO: Collas; Pooled: Pooled data from San Luis, Santa Cruz, La Pampa and Tucumán; *significant *P*-values.

samples from San Luis, Santa Cruz, Tucumán and La Pampa. When comparing urban populations with Amerindians and European populations, significant *P*-values were observed at 12–15 loci ($0.000 < P < 0.0498$). The four non-differentiated urban populations studied were pooled in a single population database ($N=322$). *P*-values showed no deviation from Hardy–Weinberg equilibrium in this pooled sample. The combined matching probability and a priori chance of exclusion for pooled data were 2.0×10^{-18} and 0.9999995, respectively. These analyses show that it would be possible to use a combined database for Tucumán, San Luis, La Pampa and Santa Cruz, since no significant differences were found between any of these populations. Nevertheless, caution should be taken concerning small isolates where Native American component could be much more relevant. In addition, it is remarkable that when comparing Argentina urban population with two Iberian samples (a major population source of the country European stock), some significant differences were found. Therefore, an Iberian database might not adequately represent the Argentinean genetic makeup. The real impact in forensic casework would require further investigation.

Acknowledgements

This work was partially supported by POCTI, Programa Operacional Ciência, from Fundação para a Ciência e Tecnologia.

Reference

- [1] S. Schneider, D. Roessli, L. Excoffier, Arlequin version 2000, A Software for Population Genetic Data Analysis, University of Geneva, 2000.