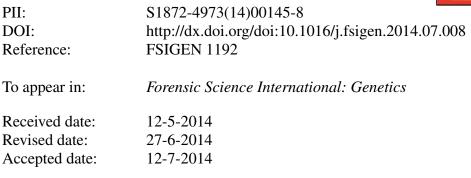
Accepted Manuscript

Title: Forensic population data for 20 STR loci in Argentina

Author: Alicia Borosky Ulises Toscanini Andrea Gómez María Laura Parolín Néstor Basso Carlos Vullo



Please cite this article as: A. Borosky, U. Toscanini, A. Gómez, M.L. Parolín, N. Basso, C. Vullo, Forensic population data for 20 STR loci in Argentina., *Forensic Science International: Genetics* (2014), http://dx.doi.org/10.1016/j.fsigen.2014.07.008

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Forensic population data for 20 STR loci in Argentina. 1 2 3 4 5 6 7 8 9 Alicia Borosky^a, Ulises Toscanini^c, Andrea Gómez^c, María Laura Parolín^b, Néstor Basso^b, Carlos Vullo^{a,*} ^a LIDMO- Córdoba- Argentina. ^bCENPAT – Puerto Madryn – Argentina. ^cPRICAI- Fundación Favaloro- CABA- Argentina. 10 11 * Corresponding author at: LIDMO- Independencia 644- 4to A- 5000- Córdoba- Argentina 12 Tel.: 54 351 4240434 13 14 E-mail address: cvullo@fibercorp.com.ar 15 16

16 Dear editor,

17 During the last years, new commercial STRs kits have become available to the forensic genetic community, including the set of markers

- 18 in the widely used PowerPlex® 16 HS System (Promega Corp, USA) and AmpFISTR® Identifiler (Applied Biosystems, USA) kits, and
- 19 different numbers of 'new' STRs. In order to provide useful information of forensic interest for the new markers included in the
- 20 PowerPlex® 21 System (Promega Corp, USA), namely D1S1656, D6S1043 and D12S391, a population study was conducted in a
- 21 sample of 907 unrelated healthy individuals from Argentina. Samples were randomly chosen from routine paternity testing. Blood
- 22 samples or buccal swabs were collected after informed consent, taken from individuals of different urban populations from 7 provinces of
- 23 Argentina: 464 individuals from a Region 1, including the provinces of Catamarca (N=27), Córdoba (N=67), Entre Ríos (N=24) and
- Buenos Aires (N=346), and 443 individuals form a Region 2 including the provinces of Neuquén (N=134), Chubut (N=223) and Santa
 Cruz (N=86).
- 26 Argentina's population ancestry shows variable degree of admixture among Native American populations with Europeans mainly from
- 27 Spain and Italy [1-4]. Provinces in the Region1 comprise the 51 % of the total populations of the country, whereas the Region2 has the
- 28 3% of the total population of Argentina [1]. Besides, Region 2 exhibits a higher proportion of Native American contribution with
- approximately 8% of Native American component against 1.8% of Native American component in the Region 1 [1].
- 30 DNA was extracted either from peripheral blood or buccal swabs using standard procedures and DNA typing was performed following
- 31 manufacturer instructions in an ABI PRISM 310 Genetic Analyzer. Analysis of raw data and allele calls were done using Genemapper
- 32 v3.2 Analysis Software (Applied Biosystems). Forensic parameters (i.e. Typical Paternity Index (TPI), Discrimination Power (PD) and
- 33 Power of Exclusion (PE) were calculated using an *in home* modified version of the PowerstatsVersion 1.2 software package (Promega,
- 34 Madison, WI, USA) [5]. Exact tests for Hardy–Weinberg equilibrium were performed with Arlequin Version 3.5 [6], as well as for
- 35 population comparisons. Additionally, frequency distribution for the three 'new' markers included in the PowerPlex® 21 System (i.e.
- 36 D1S1656, D6S1043 and D12S391) were also compared with those in other populations previously studied [7-28] (Supplementary Table
- 37 1). A significance level of .05 was used for all the statistical tests. Bonferroni's correction was used for Hardy–Weinberg equilibrium and
- linkage disequilibrium tests between pairs of loci, which assume that a 0.05 significance level obtained for 20 tests (one per locus) yields
 an actual significance threshold of 0.0025 [29].
- 40 Allele frequencies for the 20 STR loci analyzed and statistical parameters of forensic interest are shown in Supplementary Table 2a to
- 41 2i.Since we are aware that the sample size for some populations is low for forensic purposes (e.g. Catamarca and Entre Ríos) we carried
- 42 out an Exact Test of Sample Differentiation Based on Haplotype Frequencies as implemented in Arlequin V3.5, between any possible
- 43 pair of samples in order to test if some geographical grouping could be possible. Thus, for these comparisons we also considered two sets
- 44 of pooled data: (i) Buenos Aires, Córdoba, Entre Ríos and Catamarca referred as Region 1-; and (ii) Neuquén, Chubut and Santa Cruz –
- 45 referred as Region 2–. Supplementary Table 3 shows the non-differentiation exact *p*-values for the global test, and also the significant *p*-
- 46 values observed for different pairs of populations for each locus. Although *p*-values for the global test were always above the significant
- 47 level for every locus, significance between some pairs of different Provinces were observed at particular loci. However, no significant
- 48 differences were observed between pooled data for Region 1 and the Provinces involved in this group, and between pooled data for the
- 49 Region 2 and the Provinces of this group. Therefore, frequency distributions for the two pooled datasets could be reliably used for the
- 50 interpretation of caseworks in some of the populations that may be underrepresented in this study.
- 51 Supplementary Table 1a, 1b and 1c show the results of single locus comparisons for the three 'new' markers included in the PowerPlex®
- 52 21 System with published data from other populations of America, Europe and Asia.
- 53 Exact test method showed no departures from the Hardy–Weinberg equilibrium (HWE) in the populations studied. The combined Power
- 54 of Exclusion (PE) and Combined Power of Discrimination (PD) for the PowerPlex® 21 System were 0.99993 and
- 56 Argentina, respectively. Similar values were obtained for these parameters in the original datasets included in each geographical group.
- 57 The use of these 20 STR markers is very useful in the analysis of incomplete paternity cases, paternity testing with genetic
- 58 inconsistencies, to increase the resolution in database searching and for the correct statistical interpretation, given that population
- 59 substructure across Argentina has been demonstrated [30,31]. The results of this study add complementary information to that existing
- 60 for the common 17 STR set included in the PowerPlex® 16 System [30] and AmpFISTR® Identifiler [31], and provide new valuable data
- 61 of forensic interest for D1S1656, D6S1043 and D12S391 loci in different urban populations of Argentina. Our database might also be
- 62 useful for other population genetics and diversity studies.

63	This pape	r follows the guidelines for publication of population data requested by the journal [33].		
64	Quality control: Proficiency testing of the GHEP-ISFG WG (Spanish-Portuguese Speaking Working Group of International Society for			
65	Forensic (Forensic Genetics) (http://www.gep-isfg.org) and the proficiency testing of the Argentine Society for Forensic Genetics (SAGF)		
66	(<u>http://www.sagf.org.ar</u>) are regularly carried out.			
67				
68 60	Reference			
69 70	r11	Lections de Este Matieure Concerne INIDEC - premier des concernes		
70	[1]	Instituto de Estadísticas y Censos – INDEC, - www.indec.gov.ar. U. Toscanini, L. Gusmão, G. Berardi, A. Gómez, R. Pereira, E. Raimondi. Ancestry proportions in urban populations of		
72	[2]	Argentina. Forensic SciInt Genet SS 3 (2011) e387–e388		
73	[3]	U. Toscanini, L. Gusmão, G. Berardi, A. Amorim, A. Carracedo, Salas, E. Raimondi. Y chromosome microsatellite genetic		
74	[0]	variation in two Native American populations from Argentina: Population stratification and mutation data. Forensic SciInt		
75		Genet 2(2008)274-280		
76	[4]	Catelli L, Romanini C,Borosky A, Salado-Puerto M, Prieto L,Vullo C. Common mitochondrial DNA haplogroups observed in		
77		an Argentine population database sample. Forensic SciInt Genet SS 2 (2009) 329–330.		
78	[5]	A. Tereba, Profiles in DNA 3, Tools for Analysis of Population Statistics, Promega Corporation, 1999,		
79		http://www.promega.com/geneticidtools		
80	[6]	L. Excoffier and H.E. L. Lischer. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses		
81		under Linux and Windows. Molecular Ecology Resources10(2010)564-567.		
82	[7]	P. Hatzer-Grubwieser, B. Berger, D. Niederwieser, M. Steinlechner. Allele frequencies and concordance study of 16 STR loci		
83		- including the new European Standard Set (ESS) loci - in an Austrian population sample. Forensic SciInt Genet 6 (2012)		
84		e50-e51		
85 86	[8]	L. Fernandez-Formoso, C. Phillips, A. Rodriguez, R. Calvo, A. Barbaro, M.V. Lareu, Á. Carracedo. Allele frequencies of 20		
86 87	[0]	STRs from Northwest Spain (Galicia).Forensic SciInt Genet6 (2012) e149-e150		
88	[9]	A. Molnár, A.Zalán, G. Horváth, H. Pamjav. Allele distribution of the new European Standard Set (ESS) loci in the Hungarian population. Forensic SciIntGenet 5 (2011) 555-556		
89	[10]	A. Berti, F. Brisighelli, A. Bosetti, E. Pilli, C. Trapani, V. Tullio, C. Franchi, G. Lago, C. Capelli. Allele frequencies of the		
90	[10]	new European Standard Set (ESS) loci in the Italian population. Forensic SciInt Genet 5 (2011) 548-549		
91	[11]	I. Yurrebaso, J.A. Ajuriagerra, A. Alday, I. Lezama, J.A. Pérez, E. Romón, I. Uriarte, O. García. Allele frequencies and		
92		concordance study between the Identifiler and the PowerPlex ESX17 systems in the Basque Country population. Forensic		
93		SciInt Genet5 (2011) e79-e80		
94	[12]	M. Jedrzejczyk, R. Jacewicz, S. Szram, J. Berent. Genetic population studies on 15 NGM TM STR loci in central Poland		
95		population. Forensic SciInt Genet 6 (2012) e119-e120		
96	[13]	L.N. Riccardi, S. Ceccardi, M. Falconi, D. Luiselli, C. Bini, S. Pelotti. Genetic polymorphism of the new PowerPlex® ESI 17		
97		system in a Tibetan population from Dharamsala (India). Forensic SciInt Genet SS 3 (2011) e461-e462		
98	[14]	T. Seider, R. Fimmers, P. Betz, T. Lederer. Allele frequencies of the five miniSTR loci D1S1656, D2S441, D10S1248,		
99		D12S391 and D22S1045 in a German population sample. Forensic SciInt Genet 4 (2010) e159-e160		
100	[15]	J. Manríquez, S. Rojas, M.O. Yáñez, G. Molina. Evaluation of PowerPlex ESI 171 amplification kit in an admixed Hispano-		
101		Amerindian population sample of Valparaiso, Chile. Forensic SciInt Genet SS 3 (2011) e113-e114		
102	[16]	B. Wang, J. Kuang, Y. Ma, G. Wang, W. Zhu, Y. Hou. Genetic polymorphisms of 20 STR loci in Chinese Han population in		
103 104	[17]	Tianjin, North China. Forensic SciInt Genet SS 4 (2013) e192-e193		
104	[17]	H. Asamura, M. Ota, H. Fukushima.Population data on 10 non-CODIS STR loci in Japanese population using a newly		
105	F101	developed multiplex PCR system. Journal of Forensic and Legal Medicine. 15 (2008) 519-523 R. Rodenbusch, S. Schumacher, A.C. Mardini, F. Goulart, Lanes, Chula, C. TosiMichelon, A. ZorattoGastaldo, L.		
100	[10]	R. Rodenbusch, S. Schumacher, A.C. Mardini, F. Goulart Lanes Chula, C. TosiMichelon, A. ZorattoGastaldo, L. ParticheliMaciel, G.C.R. Hirschfeld-Campolongo, P.E. Raimann, S. Esteves de Matos Almeida, C.M. Dornelles da Silva.		
107		Allele frequencies of the five new generation forensic STR (D1S1656, D2S441, D10S1248, D12S391 and D22S1045) in the		
109		population from Rio Grande do Sul, Southern Brazil. Forensic SciIntGenet6 (2012) e55-e57		

110	[19]	Z. Jakovski, K. Nikolova, R. Jankova-Ajanovska, B. Janeska, N. Pojskic, D. Marjanovic. Allele frequencies of the new
111		European Standard Set (ESS) loci plus SE33 locus in a population from the Republic of Macedonia. Forensic SciInt Genet 6
112		(2012) e90-e92
113	[20]	L. Albinsson, L. Norén, R. Hedell, R. Ansell. Swedish population data and concordance for the kits PowerPlex® ESX 16
114		System, PowerPlex® ESI 16 System, AmpFlSTR® NGM™, AmpFlSTR® SGM Plus™ and Investigator ESSplex. Forensic
115		SciInt Genet 5 (2011) e89-e92
116	[21]	V.C. Tucker, C. Baumgartner, G.R. Stead, A.J. Hopwood. UK population data generated with the PowerPlex® ESI 16 system.
117		Forensic SciInt Genet 6 (2012) e112–e118
118	[22]	C.R. Hill, D.L. Duewer, M.C. Kline, M.D. Coble, J.M. Butler. U.S. population data for 29 autosomal STR loci.Forensic SciInt
119		Genet 7 (2013) e82-e83
120	[23]	C.H. Shin, P. Jang, K.M. Hong, M.K. Paik. Allele frequencies of 10 STR loci in Koreans. Forensic SciInt 140 (2004) 133-135
121	[24]	D. Vanek, M. Silerova, V. Urbanova, J. Dubska, L. Saskova, E. Ehler. Allele frequencies of the new European Standard Set
122		(ESS) loci in the population of Czech Republic.Forensic SciInt Genet SS 3 (2011) e395-e396
123	[25]	M.J. Farfán, P. Sanz, M.V. Lareu, A. Carracedo. Population data on the D1S1656 and D12S391 STR loci in Andalusia (South
124		Spain) and the Maghreb (North Africa). Forensic SciInt 104 (1999) 33-36
125	[26]	S. Dognaux, M.H.D. Larmuseau, L. Jansen, T. Heylen, N. Vanderheyden, B. Bekaert, F.N. Ronny Decorte. Allele frequencies
126		for the new European Standard Set (ESS) loci and D1S1677 in the Belgian population.Forensic SciInt Genet 6 (2012) e75-e77
127	[27]	M. Aler, A. Salas, M. Gisbert, A. Carracedo. Data for nine autosomal STRs markers from Valencia (East Mediterranean coast
128		of the Iberian Peninsula).Forensic SciInt116 (2001) 37-39
129	[28]	V. Cortellini, N. Cerri, A. Verzeletti. Population data on 5 non-CODIS STR loci (D10S1248, D22S1045, D2S441, D1S1656,
130		D12S391) in a population sample from Brescia county (Northern Italy). Forensic Sci Int Genet 5 (2011) e97-e98
131	[29]	B.S.Weir, Multiple tests, in: Genetic Data Analysis II, Sinauer Associates, USA, 1990, pp. 109-110.
132	[30]	U. Toscanini, A. Salas, M. García-Magariños, L. Gusmão, E. Raimondi. Population stratification in Argentina strongly
133		influences likelihood ratio estimates in paternity testing as revealed by a simulation-based approach. Int J Legal Med 124
134		(2010) 63-9
135	[31]	U. Toscanini, L. Gusmão, G. Berardi, A. Amorim, A. Carracedo, A. Salas, E. Raimondi. Testing for genetic structure in
136		different urban Argentinian populations. Forensic Sci Int. 165(2007) 35-40
137	[32]	Borosky A, Catelli M, Vullo C, Analysis of 17 STR loci in different provinces of Argentina. Forensic SciInt Genet 3 (2009)
138		e93–e95.
139	[33]	A. Carracedo, J. Butler, L. Gusmão, et al. Update of the guidelines for the publication of genetic population data. Forensic
140		Science International: Genetics 10 (2014) A1-A2
141		