Original Research Article

Analysis of Locus D9S1120 and Its Genetic Admixture Correlation in Seven Argentina Native American Ethnic Groups

MARIELA CAPUTO,[1,2] AND DANIEL CORACH[1,2]*
[1]*Servicio de Huellas Digitales Genéticas, School of Pharmacy and Biochemistry, Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*CONICET, National Scientific and Technical Research Council, Buenos Aires, Argentina*

Objective: Genetic data have complemented archaeological and linguistic investigations for understanding the peopling of the Americas. Aiming to investigate the Native South American genetic background in Argentina, seven Amerindian and one urban population were selected. The analysis focused on locus D9S1120 due to its potential anthropological information about Native American origins.

Methods: The sample set included 603 individuals belonging to nine isolated Argentinean aboriginal communities from seven tribes ($N = 296$), 100 individuals living in Buenos Aires city, and three potentially parental population references samples ($N = 207$). We computed allele and genotype frequency distributions, genetic distances, and pairwise differences among and within them. Admixture proportion was determined by means of typing 13 autosomal short tandem repeats plus D9S1120 in all populations, and comparing the data with those from three parental groups including Native American, European and Sub Saharan West African.

Results: The Native American-specific allele 9RA was found at an average frequency of 0.26 in aboriginal groups. Statistically significant differences were observed among the Native American groups when compared with the Buenos Aires urban population using analysis of molecular variance (AMOVA) ($F_{st} = 0.05669$; $P < 0.0001$). Admixture analysis denoted different results between the cohorts of Amerindian samples displaying the specific 9RA allele, compared with those lacking it. A linear correlation was established between positive 9RA and Native American ancestry.

Conclusions: Autosomal-based genetic admixture showed that the studied communities have considerable European and Native America contributions. Our results concerning D9S1120 further contribute to a better understanding of the admixture process between Sub Saharan African, Native American, and European individuals that shaped the genetic background of Argentinean extant population. Am. J. Hum. Biol. 28:57–66, 2016.

Archaeological, ethno-historical, and linguistic evidence and genetic markers provided robust approaches for tracing the migration processes that determined the demographic scenario of present day global populations (Cavalli-Sforza et al., 1994; Reich et al., 2012; Salzano and Callegari-Jacques, 1988; Ward et al., 1993). Bi-parentally transmitted genetic polymorphisms such as blood groups (Halverson and Bolnick, 2008; O'Rourke et al., 1992), immunoglobulin GM allotypes (Williams et al., 1985), minisatellites, short tandem repeats (STRs), and single nucleotide polymorphisms (SNPs) (Reich et al., 2012; Sala et al., 1999) on one hand, and markers uniparentally transmitted on the other, such as mitochondrial DNA (mtDNA) and Y-chromosome polymorphisms (Underhill et al., 1996), complement cultural and archaeological information.

These synergies lead to a better understanding about the initial peopling of the American continents. Particular genetic markers showed to be private or very high frequency in the American pre-European populations. These markers include: ABO, allele O (Halverson and Bolnick, 2008); Diego, Di a+ (Saguier Negrete, 1964); DYS199 C to T transition defining M3-Q haplogroup (Underhill et al., 1996; Zegura et al., 2004); A, B, C, D founding mitochondrial haplogroups (Wallace and Torroni, 1992), and the controversial minor X founding lineage detected in North Amerindian groups (Brown et al., 1998; Dornelles et al., 2005). More recently, at locus D9S1120, a Native American private allele called 9-repeat-allele (9RA), was identified and characterized. It has been suggested that it might have experienced an expansion during the initial colonization of the Americas (Schroeder et al., 2009; Schroeder et al., 2007; Wang et al., 2007). It was proposed that a limited number of immigrants from the eastern edge of Siberia arrived in the last continent to be inhabited by modern humans, by crossing the Bering Strait, sometime at the end of the Pleistocene, over 15,000 years before present (ybp) (Reich et al., 2012). Throughout the Americas, historical differences in the pattern of settlement and migration have resulted in a wide range of parental genetic contributions to rural or urban populations (Sans, 2000).

The 9RA in locus D9S1120 is present in Native American populations at an overall average frequency of 0.354 (0.301 and 0.471 in North and South Americas, respectively) (Schroeder et al., 2009; Wang et al., 2007; Zhivotovsky et al., 2003). Sampling areas covers from the coast of the Bering Sea to the Southern end of Argentina and Chile including: Pima, Maya, Karitiana, Surui, and Awa-Pijao-Coiyama communities. This allele is absent in all other populations outside America and Central South and East Asia (Phillips et al., 2008; Wang et al., 2007;

Contract grant sponsor: University of Buenos Aires; Contract grant number: 20020100100744 UBACyT; Contract grant sponsor: CONICET; Contract grant number: PIP 112-200801-02836.

*Correspondence to: Daniel Corach, Servicio de Huellas Digitales Genéticas, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956- 7mo piso -Ciudad Autónoma de Bs. As. -(1113), Argentina. E-mail: shdg@ffyb.uba.ar

Received 6 October 2014; Revision received 4 May 2015; Accepted 27 May 2015

DOI: 10.1002/ajhb.22755
Published online 19 June 2015 in Wiley Online Library (wileyonlinelibrary.com).

Zhivotovsky et al., 2003). Schoreder and colleagues concluded that the 9RA is identical by descent and not a result of positive selection, as the haplotypic background at 9RA was not unusually long and did not have a high frequency in the Americas. Futhermore, in Latin American Mestizo populations, the proportion of Native American ancestry at D9S1120 is not unusual compared with that observed at other genome wide microsatellites (Schroeder et al., 2009).

Prior to European contact, the territory was inhabited by aboriginal tribes whose ancestors were Asian immigrants that entered the continent along a temporal range from 25.000 to 15.000 years before present (ybp) (Cavalli-Sforza et al., 1994). It has been suggested that the extant populations immigrated following diverse corridors to modern Argentinean territory. Hence, Guaraní people arrived from the Atlantic forest by following the natural course of the Paraná and Paraguay rivers (Tarragó, 2000), while the Wichi, Toba, Pilagá, and Mocoví developed in Gran Chaco, a transitional eco-geographical zone between the Amazon tropical plains and Argentina's pampas (Demarchi and Garcia Ministro, 2008). Concerning the Mapuche, complex patterns of interaction between the native inhabitants on both sides of the Southern Andes have been recorded in numerous archaeological and ethno-historical studies (Bechis, 2010; Hajduk and Ménica Cúneo, 1997). The intensity of contact, with important eastward migrations, increased notably from the 17th century, resulted in "the araucanization of Pampa-Patagonia," representing an important cultural and biological influence of trans-Andean populations (Dana, 2006; Garcia et al., 2006; Hajduk and Ménica Cúneo, 1997). The Tehuelche population was affected by two major events, the arrival of the Araucans from Chile through the Andes in the West during the sixteenth century, and the admixture between Native Patagonian women and males from Argentinean military troops that participated in the military conquest of Patagonia during the 1870s (Ratto, 2003). The Araucanian invasion also imposed the Mapuche language on Tehuelche and other Patagonian tribes. This language called *Mapudugun* is, nowadays, a living language used by the descendants of Patagonian tribes. Finally, the Spanish language could be acquired by natives from their contact with catholic missionaries since the seventh century (Bjerg, 2007).

The extant population of Argentina has been genetically influenced by three major differentiated continental contributors (Avena et al., 2006; Avena et al., 2012; Corach et al., 2010). They were Native American, European, and Sub Saharan West African.

As there is a lack of information about Argentinean Amerindian tribes concerning genetic admixture, we selected seven Native American communities and an admixed urban population from Argentina to evaluate, on one hand, if the presence of 9RA could predict Native American ancestry. On the other hand, aiming to estimate the proportions of genetic admixture, we included in the comparison a reference sample set previously typed with 13 autosomal STRs and obtained from the Control Panel of the Centre des Etudes de Polymorphisms Humane (CEPH). It comprised Native American, European, and Sub Saharan West African populations that represent the parental populations, which contributed to the make-up of the contemporary urban Argentine population, roughly representing the extant Argentinean gene pool.

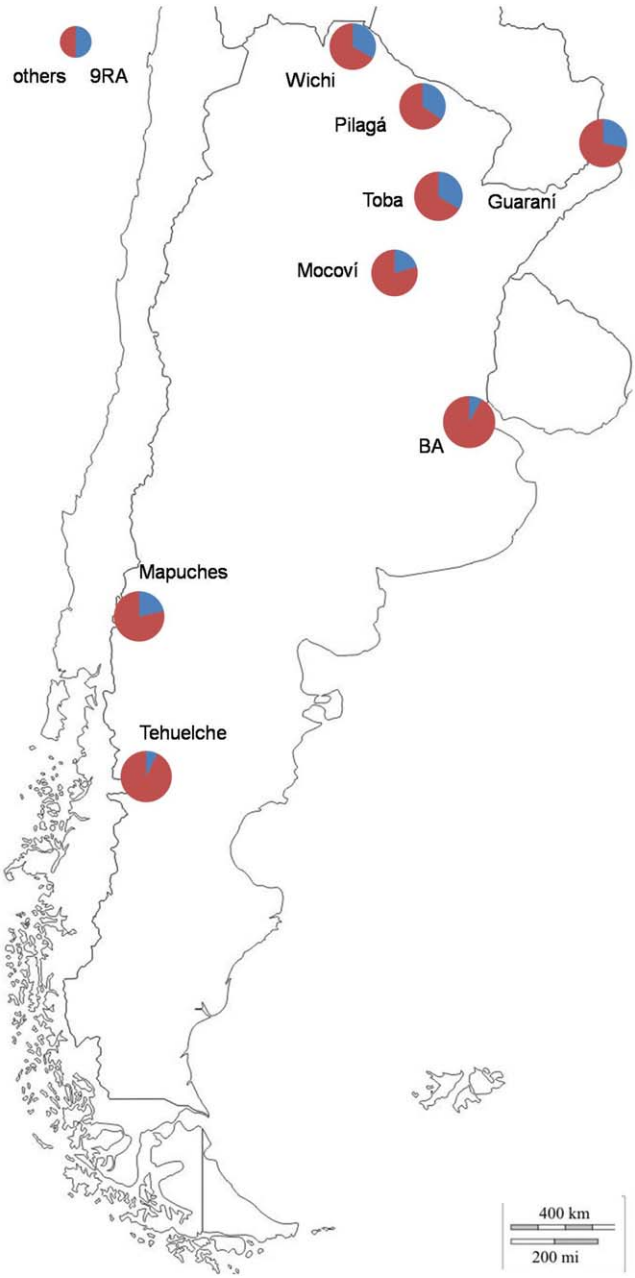

Fig. 1. Geographic locations of the studied populations are indicated within the map of Argentina. In each location, the 9RA frequency is plotted in blue and the other alleles frequencies are in red.

MATERIALS AND METHODS

Population studied

We selected a set of 296 participants from nine Argentinean Native American populations for this study. The tribes included belong to three major linguistic groups (Greenberg and Ruhlen, 2007): Ge-Pano-Carib; Macro-Panoan sub-branch *Mataco-Guaycurú* including Wichi (formerly called Mataco), Pilagá, Toba, and Mocoví communities (*Guaycurú* speakers); Equatorial-Tucanoan;

TABLE 1. D9S1120 frequencies in Native and an urban Argentinean population

	Guaraní	Wichi	Toba	Pilagá	Mocoví	Mapuche	Tehuelche	BA
9RA	0.2821	0.3333	0.3375	0.3462	0.2051	0.2167	0.0714	0.0800
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0167	0.0000	0.0100
13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0150
14	0.0000	0.0093	0.0125	0.0000	0.0000	0.0000	0.0179	0.0150
15	0.0385	0.3889	0.1438	0.1154	0.2949	0.1833	0.2679	0.1650
16	0.3718	0.1759	0.4250	0.4038	0.2949	0.3500	0.4643	0.4850
17	0.2436	0.0093	0.0625	0.0769	0.1923	0.1833	0.1607	0.1500
18	0.0641	0.0833	0.0125	0.0577	0.0000	0.0500	0.0179	0.0750
19	0.0000	0.0000	0.0063	0.0000	0.0128	0.0000	0.0000	0.0050
Ho	0.667	0.574	0.70	0.692	0.739	0.667	0.643	0.710
Different alleles	5	6	7	5	5	6	6	9
Total alleles	78	108	80	52	78	60	56	200

TABLE 2. F_{st} *comparison among the regional populations:* F_{st} *values are below the diagonal and the* P *values (1,023 permutations) above the diagonal*

	Guaraní	Wichi	Toba	Pilagá	Mocoví	Mapuche	Tehuelche	Buenos Aires
Guaraní	0.00000	**0.00000**	**0.00000**	0.27928	**0.00901**	0.19820	**0.00000**	**0.00000**
Wichi	0.12456	0.00000	**0.00000**	**0.00000**	**0.00901**	**0.00000**	**0.00000**	**0.00000**
Toba	0.02738	0.07921	0.00000	0.86486	**0.00000**	0.09009	**0.00000**	**0.00000**
Pilagá	0.01088	0.07348	−0.01011	0.00000	**0.02703**	0.29730	**0.00000**	**0.00901**
Mocoví	0.04227	0.04219	0.04128	0.03731	0.00000	0.51351	0.06306	**0.00000**
Mapuche	0.00521	0.06351	0.01530	0.00620	−0.00255	0.00000	0.18018	0.09910
Tehuelche	0.06124	0.11066	0.05561	0.05740	0.01795	0.01215	0.00000	0.34234
Buenos Aires	0.04458	0.13481	0.05071	0.04655	0.04246	0.01627	−0.00089	0.00000

$P < 0.05$ were considered as statistically significant indicated in bold.

sub-branch *Tupí-Guaraní* including MByá Guaraní and Andean Southern, including Mapuche and Tehuelche (*Mapudugun* speakers). The aboriginal samples included 39 Mbyá Guaraní from FortínMBororé, Puerto Iguazú (S 25° 32′ W 54° 35′) in Northwest Misiones Province;54 Wichi from PozoYacaré- Ing. Juarez (S 23° 32′ W 62° 14′) in Formosa Province; 26 Toba from San Martin, Chaco Province (S 26° 35′ W 59° 20′); 40 Toba from Rosario (S 32° 57′ W 60° 42′) in Santa Fe Province, and 14 Toba from Las Lomas (S 31° 34′ W 60° 43′) in Northwest of Santa Fe City in Santa Fe Province, 26 Pilagá from Las Lomitas, Formosa Province (S 24° 42′ W 60° 36′); 39 Mocoví from Tostado (S 29° 14′ W 61° 46′) in Northeast Santa Fe Province; 30 Mapuche inhabitants of Cerro Policía, Loncovaca Arriba (S 39° 37′ W 68° 36′) in Western Río Negro Province and 28 Tehuelche sampled from the communities of Chalia (S 45° 28′ W 70° 39′) in Southwestern Chubut Province. The sampling locations are shown in Figure 1. Most of the samples were collected by one of the authors, DC. Mapuche and Tehuelche samples were kindly provided by Dr. Raul Carnese and were investigated previously (Ginther et al., 1993; Sala and Corach, 2014; Sala et al., 1999).

Additionally, a set of 100 voluntary donors from the urban area of Buenos Aires city provided their samples for comparison; these samples were obtained at Hospital Italiano, a private hospital at Buenos Aires. In all cases, Native and urban sample donors read and signed written consent statements approved by the bioethical committee, School of Pharmacy and Biochemistry, Buenos Aires University.

DNA extraction

DNA was obtained from saliva (Quinque et al., 2006) or blood samples spotted onto filter paper extracted by conventional Proteinase K-SDS incubation followed by organic extraction and ethanol precipitation (Sambrook et al., 1989).

Native American patrilineage and matrilineage determination

Native American sample donors were selected based on their having aboriginal uniparental lineages. DNA samples were previously analyzed for identifying the M3-Q Y-haplogroup and any of the major founding Native American mitochondrial haplogroups (A, B, C, or D). This goal was achieved by means of real time polymerase chain reaction (PCR) assays, followed by high-resolution melting analysis. The procedure followed the approach proposed by Zucarelli et al. (2011). For this purpose, syto9 intercalating dye (Lifetechnologies) and real-time PCR Rotor-Gene equipment (Corbet Inc., Sidney, Australia) were used.

PCR amplification

D9S1120 locus was typed using primers and amplification conditions according to Phillips et al. (2008). Amplicons were separated and visualised on an Applied Biosystems 3100 Avant Genetic Analyzer, and analysed with the expert software Genemapper IDX version 1.0 (Applied Biosystems, Foster City, CA). In order to ensure that the identical allele designation was the same as Phillips and colleagues, three homozygous individuals for allele 9RA and other three homozygous for allele 16 were sequenced using BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) confirming the number of repeats and STR structure. Sequence analysis was carried out with the help of Sequencher® v.4.8 software (Gene Codes Corporation). The allele ladder was constructed by mixing PCR products from samples containing 9, 14, 15, 16, and 17 repeats.

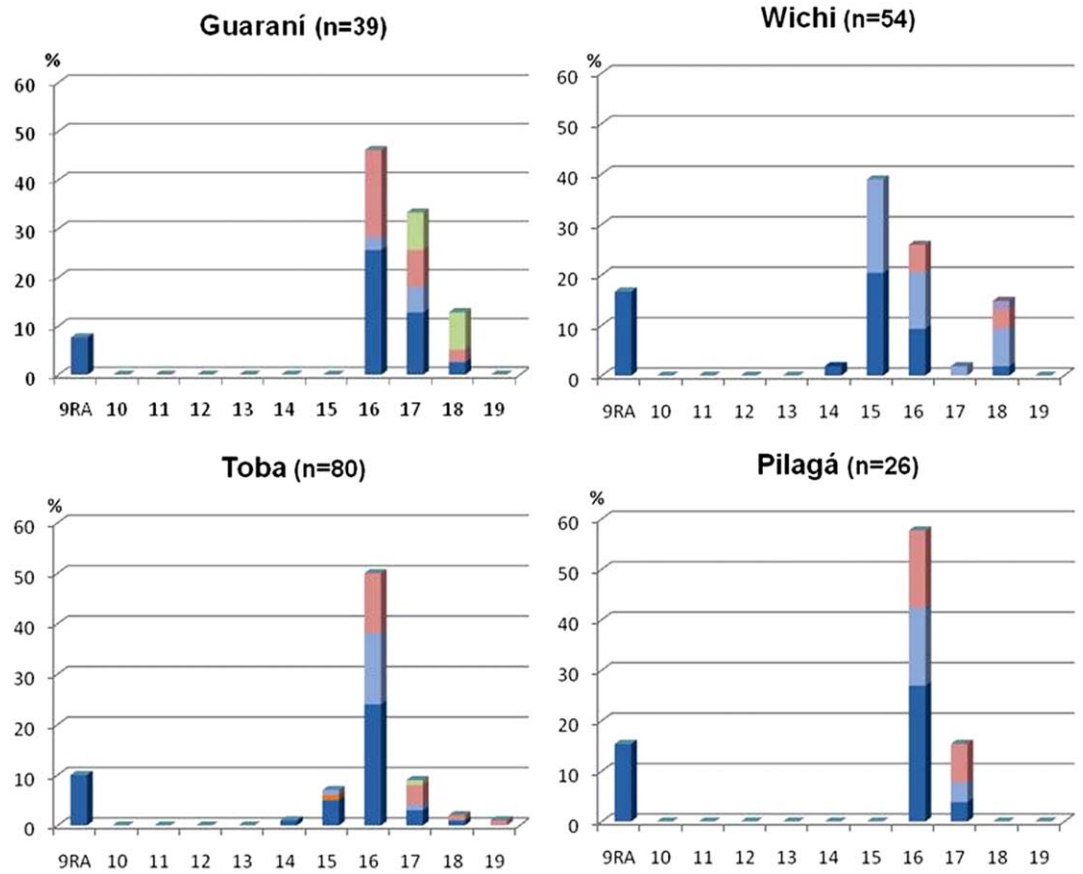

Fig. 2. Locus D9S1120 genotype frequency distribution detected in the seven Native American Argentinean groups and the sample representing the urban populations of Buenos Aires city. Combinations of the color codes identifying the number of repeats are displayed in the lower right of the figure and the values of number of repeats for the second allele are indicated on the *x*-axis. Combinations of color and axis values define each genotype. The *y*-axis indicates the percentage of individuals exhibiting each genotype.

Data analyses

Allele frequency distribution was computed using GenAlEx 6.5software (Peakall and Smouse, 2006, 2012). For each population sample, the number of alleles, heterozygosity, and Hardy Weinberg expectation were determined and verified. analysis of molecular variance (AMOVA) (Excoffier et al., 1992; Weir, 1996; Weir and Cockerham, 1984) and differentiation exact test of sample differentiation based on genotype frequencies (Raymond and Rousset, 1995) were evaluated by using Arlequinv3.5 software. Pairwise F_{st}s and Nei's average number of pairwise differences between population (Nei and Li, 1979) were determined by Arlequinv3.5 software and plotted with R-lequin (Excoffier et al., 2007).

Ancestry analysis

The sample set investigated (excluding the Mocoví group) was additionally typed for 13 autosomal STRs widely used in human identification studies including: CSF1PO, FGA, THO1, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11 to complement ancestry analysis. Parental samples, including 70 Sub Saharan Africans, 53 European, and 84 Native American, were chosen from CEPH panel typing results that were kindly supplied by Dr. Peter de Knijff from Leiden University. Genotypic results at D9S1120 for parental samples were obtained from the supplementary materials of Wang et al (2007). Admixture proportions were established by using STRUCTURE v 2.3.4 software (Pritchard et al., 2000). Five iterations for parental populations (k) from two to four were set and 40.000 burning followed by 60.000 Markov Chain Monte Carlo simulations were performed for each round. Boolean flag and start at pop info were set for parental populations. Admixture model and correlated allelic frequencies were chosen. A value of $k = 3$ best fit our dataset and was selected for all further analysis. Data analysis was refined using *CLUMPP* software (Jakobsson and Rosenberg, 2007) and bar plot was performed with the help of *DISRUPT* software (Rosenberg, 2004).

RESULTS

The D9S1120 genotype distribution was in Hardy Weinberg equilibrium in all the analyzed populations. The frequency distribution observed for D9S1120 is summarized in Table 1. As there were no significant differences in the allele distribution for the three Toba communities, the results were pooled together for further analysis. We found nine alleles including: 9, 12, 13, 14, 15, 16, 17, 18,

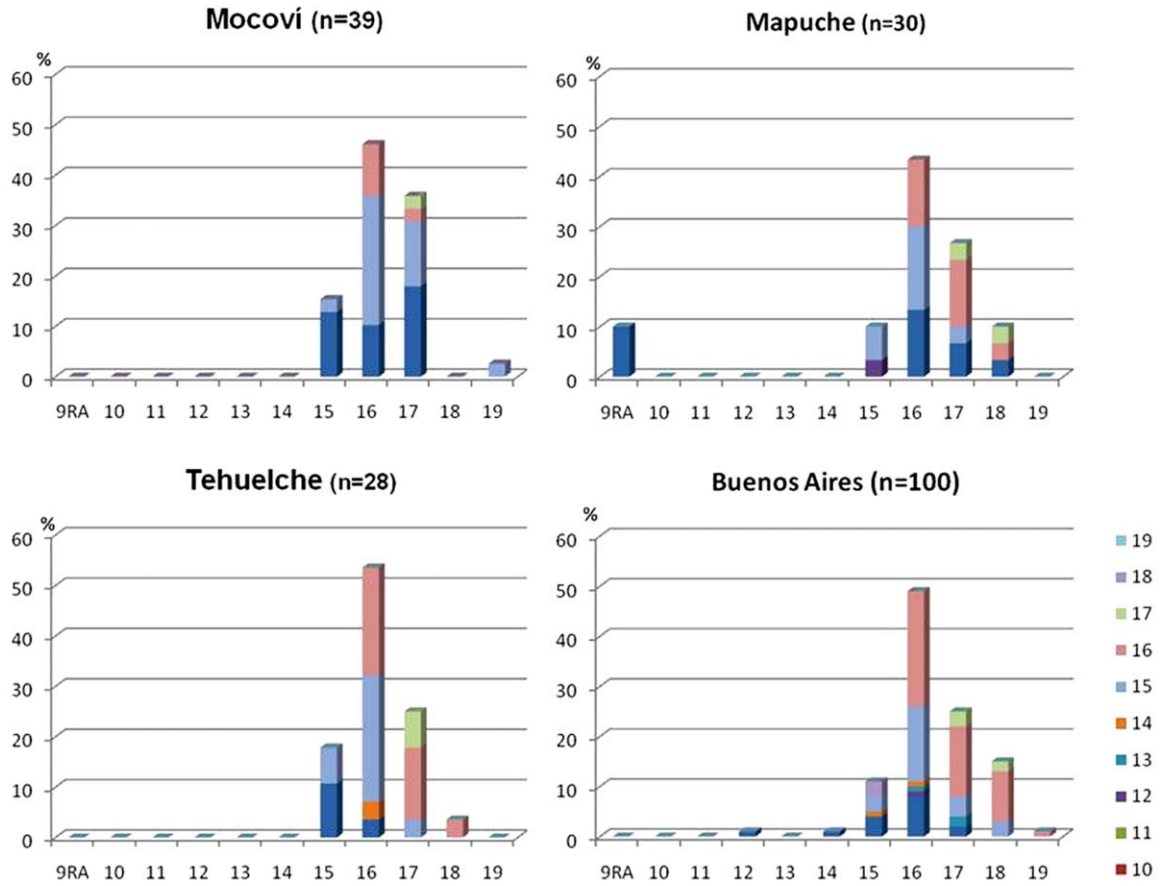

Fig. 2. Continued

and 19 repeats in the urban population. Native American communities showed a reduced number of alleles for this locus. Pilagá, Guaraní, and Mocoví only showed five; Wichi, Tehuelche, and Mapuche denoted six and seven alleles were present in the Toba. The observed heteregozity (Ho) ranged between 0.57 and 0.74. The Wichi and Mocoví groups showed the lowest and the highest Ho, respectively (Table 1), as can be seen in Figure 2. The 10, 11, nor 13 alleles were found in Argentinean Native American samples while the 13 repeat was only found in the urban population. The 9RA allele was found at an average frequency of 0.26 in the seven Native American populations. In the urban population, the 9RA frequency was 0.08, with increased allele frequency for 15, 16, and 17, reflecting the European contribution to the extant urban admixed population (Table 1 and Fig. 1).

Mataco-Guaycurú speakers showed a variable allele 9RA frequency distribution. Within this linguistically defined group, the highest frequency was detected in the Pilagá (0.35), followed by the Toba (0.34), Wichi (0.33), and Mocoví (0.21). An intermediate frequency was detected in MByá Guaraní (0.28), which was slightly lower in Mapuche (0.22). The Tehuelche population showed the lowest frequency for 9RA (0.07), even lower than the most admixed urban populations (0.08). Regarding allele 16, the highest value was found in the urban population and Tehuelche (0.49 and 0.46, respectively), being the lowest value in Wichi (0.18) (Table 1).

Concerning genotype distributions, Wichi and Pilagá communities showed the highest frequency of 9RA homozygous genotypes, whereas Buenos Aires, Mocoví, and Tehuelche groups did not have 9RA homozygous genotypes. A similar genotype distribution was observed between Tehuelche and Buenos Aires urban population (Fig. 2).

F_{at} was estimated (Table 2) and Nei's average number of pairwise differences were plotted in Figure 3. The largest genetic distance was observed between the Wichi and the Buenos Aires urban population followed by Wichi versus Guaraní (Table 2 and Fig. 3) ($P < 0.0001$). Based on these results, the Wichi group seemed to be the most isolated community (Fig. 3), furthermore, all the genetic distances for this population were statistically significant (Table 2). The lowest genetic distance was observed between Tehuelche and Buenos Aires urban population, as was expected from their similar genotype distribution. The genetic distance estimated by Nei or F_{st} between Mapuche versus Mocoví and Toba versus Pilagá seemed to be the lowest ($P < 0.05$) (Table 2 and Fig. 3). No statistically significant difference in terms of genetic distance was observed among the *Andean* Amerindian speakers group (Mapuche and Tehuelche) (Table 2). Finally, the Toba and Tehuelche showed minor differences whereas the Mapuche and Mocoví showed the largest within population difference (Fig. 3).

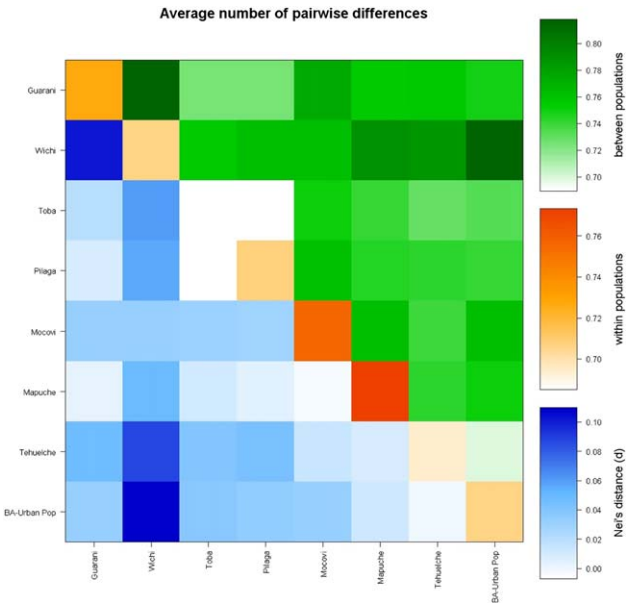

Fig. 3. R-lequin population average pairwise differences and distance plot. The graphical representation depicts: in green between populations differences (plotted above the diagonal); within population differences are indicated in the diagonal in orange, and finally, below the diagonal is coded Nei's distance in blue. On the right side of the plot are indicated the gradient colors associated with the values of differences or distances between or differences within or between populations.

Statistically significant differences were observed among the Native American groups versus Buenos Aires urban population when an AMOVA analysis was carried out ($F_{st} = 0.05669$; $P < 0.0001$) and differentiation exact test conducted ($P < 0.0001$). No statistical differences were observed among linguistic groups (*Mataco-Guaycurú*, *Tupí-Guaraní*, and *Mapudugun* speakers). A slight statistically significant difference was observed when the Native American groups were divided by geographic localization. As Northern populations, we included: Guaraní, Toba, Wichi, Pilagá, and Mocoví and within the Southern group, Tehuelche, and Mapuche. AMOVA values were $F_{st} = 0.06078$ ($P < 0.05$).

Admixture analysis using 13 autosomal STRs plus D9S1120 was carried out. This set of markers provide an acceptable approximation of genetic substructure (Barnholtz-Sloan et al., 2005; Salazar-Flores et al., 2014) and allows determination of the ancestral contributions of the parental populations to the population under study. Our results agree with those previously reported for the population of Argentina (Corach et al., 2010), which used 24 ancestry informative markers and uniparental transmitted genotypes. Genetic structure analysis using 14 autosomal STR was able to detect the three major continental contributors even though the percentage of each one was slightly different because the markers used herein were not strictly ancestry informative. Genetic structure analysis was performed defining the two groups: individuals with at least one 9RA and individuals without any 9RA. Ancestry for the 9RA positive cohort showed: 5% African, 74.5% Native American, and 20.4% European. In contrast, 9RA negative denoted a different proportion: 24.1% African, 55.9% Native American, and 20% European (Fig. 4a). Moreover, when grouping by aboriginal community, similar results were observed (Fig. 4b). The fivefold excess of African contribution in those individuals lacking 9RA represents an interpretation challenge and might be explained by selective forces acting on 9RA negative individuals but this hypothesis has not yet been tested.

A linear correlation between the tribal groups with increasing Native American ancestry percentage and frequency of 9RA was observed ($r^2 = 0.59$) (Fig. 5). As mentioned above, the urban population and the Tehuelche had similar allele and genotype frequency distributions. This situation may reflect a major degree of gene flow from the urban admixed population to the Tehuelche, which is supposed to be an isolated group. In order to reduce this bias, Tehuelche was omitted, and the new regression coefficient was increased to $r^2 = 0.98$.

DISCUSSION

In this study, we characterized the D9S1120 locus in seven Argentinean tribes and in the urban Buenos Aires population. The average frequency of 9RA in the studied tribes (0.26) was in line with that reported by others (Wang et al., 2007; Phillips et al., 2008; Schroeder et al., 2009; Schroeder et al., 2007).

Comparing our results with those already published for *Tupí-Guaraní* speakers, the Guaraní group studied by Wang and colleagues showed a higher 9RA frequency than the group of inhabitants Fortin MBororé analyzed (0.45 vs. 0.28) (Wang et al., 2007). This difference could be explained by the different levels of Native American autosomal ancestry present in our group. Although previous studies in the Guaraní group showed 100% of Native American mtDNA haplogroups (A2: 40.5%; D1: 36.0%; B2: 17.5%; C1: 6.0%) and 88.8% for Native American Y- haplogroup Q1a3a (Sala et al., 2010), analysis of genetic admixture using autosomal markers showed a Native American ancestry of 57.3% and an African contribution of 9%. In contrast, Brazilian Guaraní show clear differences both at mtDNA where haplogroup B is absent and autosomal markers that denote a 97% Native American, 3% European, and 3% African contributions (Marrero et al., 2007). The different biparentally transmitted markers used in each study might explain the differences observed.

Concerning *Mataco-Guaycurú* speakers, in particular the Mocoví community, we observed statistically significant genetic distances between other Gran Chaco inhabitants such as the Toba, Pilagá and Wichi. These results are in concordance with the results at X, Y, and autosomal chromosomes previously analyzed (Catanesi et al., 2007; Demarchi and Garcia Ministro, 2008; Demarchi and Mitchell, 2004; Glesmann et al., 2013). In contrast, a very small distance was obtained between the Mocoví and Tehuelche, probably due to the similar genotype frequencies. The Mocoví showed an intrapopulation diversity and the highest Ho which may be due to the genetic variation present in this group. Unfortunately, no autosomal markers were analyzed in the Mocoví to assess their Native American ancestry.

While the Wichi seemed to be the most isolated community in showing the lowest Ho, a very small genetic distance was estimated between the Wichi and the Toba and Pilagá. No differences between these two last groups were

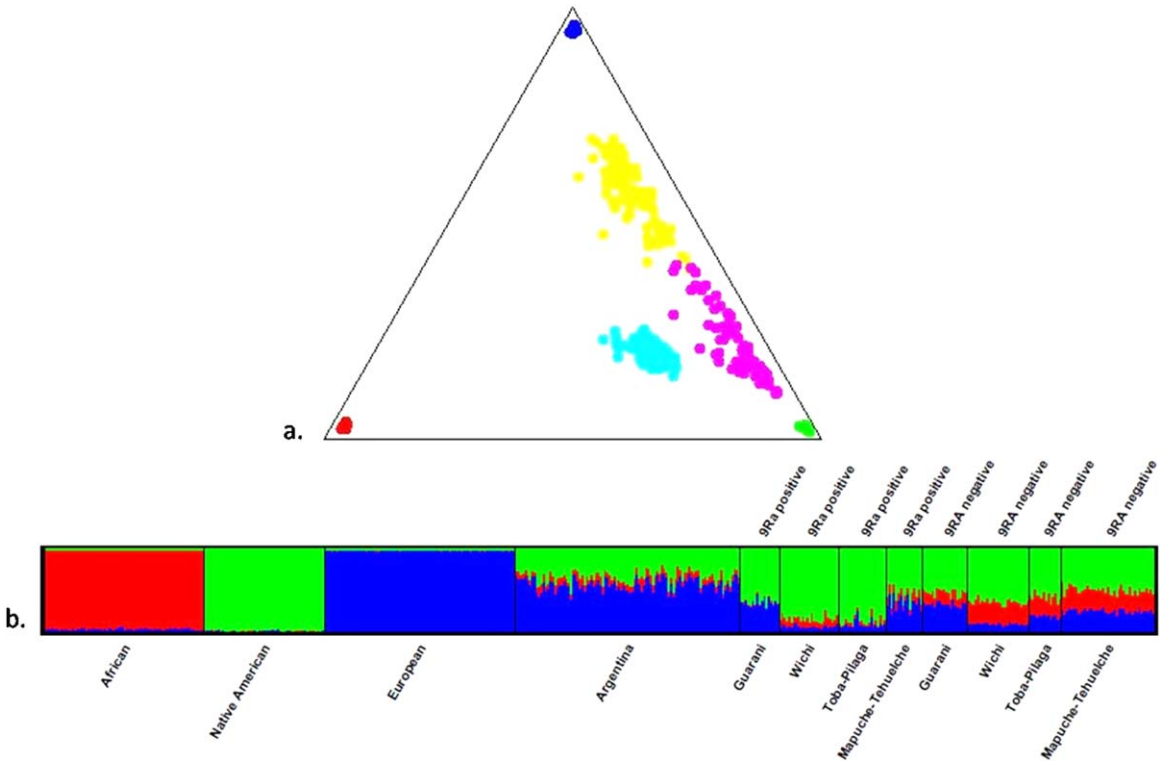

Fig. 4. (a) Genetic admixture is represented as a triangle-plot obtained after typing all samples studied with 14 autosomal STRs, 13 of which are commonly used in human identification and D9S1120. Three parental populations were included (Native American right corner indicated in green, European displayed in the upper corner colored in blue, and finally, Sub Saharan West African in the left corner colored in red). Genetic admixture analysis and plotting was performed with Structure v2.3.4 software assuming admixed model, correlated allele frequencies, pop info, and $k = 3$. A colored point represents each individual, and the distance to each corner of the triangle indicates the correspondent admixture proportions. The Buenos Aires urban population was coded in yellow, 9RA positive individuals as pink, and 9RA negative as light blue. (b) Bar plot representation: genetic structure analysis was performed defining the two groups: individuals with at least one 9RA and individuals without any 9RA.

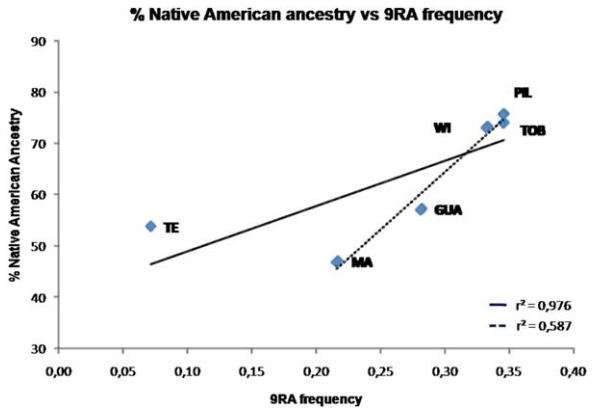

Fig. 5. Correlation between 9RA frequency and Native American Ancestry proportion. The continuous line represent the correlation, including all Native American Argentinean tribes under study ($r^2 = 0.587$). The dashed line depicts the correlation after removing the Tehuelche population from the analysis, improving the regression coefficient ($r^2 = 0.976$).

detected, which can be explained by genetic admixture between these two last communities given their linguistic affinities and their geographical proximity (Demarchi and Garcia Ministro, 2008). These three communities showed a 70–84% proportion of Native American ancestry, higher than the other studied groups at the D9S1120 locus.

Regarding the 9RA frequencies obtained for Patagonian inhabitants, the Mapuche and Tehuelche were lower than the 0.25 obtained for the Huilliche community (also *Andean* speakers) (Wang et al., 2007). Previous studies in the Mapuche and Tehuelche showed that even mtDNA haplogroups were Amerindian in all individuals (de Saint Pierre et al., 2012; Sala et al., 2014), 12.5% and 41.2%, respectively, denoting non-Native American Y-haplogroups (Sala et al., 2014). In addition, autosomal information for both groups showed 82% Native American and 17% European ancestry (Sala et al., 2014). Our results of structure analysis using autosomal markers showed that these groups had both Native American and European genetic contributions. Regarding the Tehuelche group, the low frequency of 9RA and the high frequency of allele 16 could also suggest a considerable degree of gene flow from the urban admixed population to the supposedly isolated tribal group. Moreover, the genetic relationships between Amerindian groups may have occurred as was described in blood markers and serum proteins (Goicoechea et al., 2001). The results exhibited for autosomal markers, including D9S1120, mtDNA, and Y-chromosome markers, showed reliable evidence that a

degree of non-Amerindian admixture in these populations has taken place, probably due to the movement of the Mapuche to the south at first, and, then, the introgression of both Mapuche, admixed Chileans and members of Argentinean military troops as was described in the introduction.

AMOVA showed no statistically significant differences among the, *Mataco-Guaycurú, Tupí-Guaraní*, and *Mapudugun* speakers. Nevertheless, a slight statistically significant difference was detected if the groups were divided geographically (Northern and Southern). The observed difference might be attributed to the higher European contributions in Tehueche population.

In previous reports, it was demonstrated that Argentineans carried a large fraction of European genetic heritage in their Y-chromosomal (94.1%) and autosomal (78.5%) DNA, but their mitochondrial gene pool is mostly of Native American ancestry (53.7%) with a European autosomal contribution of 78.5%, illustrating strong sex-asymmetric migrations (Avena et al., 2012; Bobillo et al., 2010; Corach et al., 2010). Our results for the 9RA frequency in the urban population from Hospital Italiano, a private hospital at Buenos Aires, showed to it to occur at a much lower frequency (0.08) than seen in the Native American groups analyzed in this study. Accordingly, Schroeder and colleagues reported for the Catamarca a frequency of 0.036, 0.21 for the Salta and 0.105 for the Tucuman population (Schroeder et al., 2009). Unfortunately, the aforementioned authors do not provide information about the source of the samples, as this information might clarify, in part, the potential degree of admixture of the population selected.

Studies performed on erythrocyte genetic systems, GM/KM allotypes (genetic markers of immunoglobulin (Ig) γ and κ chains, respectively), ancestry informative markers, maternal and paternal lineages showed that there are clear differences concerning the ethnic contribution of the selected samples if they were obtained in public or private hospitals, in economically active rural areas, in small isolated tribal groups or in an urban population of randomly selected voluntary donors (Avena et al., 2006, 2012; Martínez Marignac et al., 2004). The increased frequency of 9RA allele in the Salta province, compared with the other urban populations, could be explained by the fact that significant genetic distances do exist between the Salta Province population and other Argentinean urban populations when autosomal and Y-chromosome STRs are analyzed (Marino et al., 2006a,b; 2007), suggesting a much higher Native American contribution in the North West of the country (Avena et al., 2006, 2012; Corach et al., 2010; Marino et al., 2006a).

In the Buenos Aires urban population, D9S1120 showed the highest allele variability, with the frequencies for alleles 15, 16, and 17 similar to those described by Phillips et al. (2008) in Northwest Spain. The admixture between Native American and European populations is also detectable in D9S1120 since the 9RA frequency observed in the urban population is smaller than in the Native Americans (except for the Tehuelche) but larger than in Europeans individuals.

Regarding ancestry analysis, we were able to detect differences in genetic admixture if the individuals carried the 9RA or not. Major Native American ancestry was detected when 9RA was present and a fivefold excess African contribution was observed when 9RA was absent. These results are independent of the tribe to which the individual belonged. A positive correlation was obtained when we compared the 9RA presence versus the proportion of Native American autosomal ancestry ($r^2 = 0.59$). A greater correlation ($r^2 = 0.98$) was obtained if the Tehuelche group was not considered. Moreover, by excluding Tehuelche, the average 9RA frequency increased to 0.29, which is closer to what has been previously published for Native American populations (Wang et al., 2007).

The analysis of locus D9S1120 contributes to a better understanding of the genetic background of the Argentinean population. The dramatic demographic impact imposed by immigrants on the descendants of the Native Argentinean communities strongly reduced the original Native American genetic make-up in less than eight generations of asymmetric admixture events. Our results allow us to determine genetic admixture among aboriginal populations, providing an evaluation of European and African contributions to the supposedly ethnically isolated groups.

ACKNOWLEDGMENTS

The authors would like to thank Dr. R. F. Carnese for kindly providing the Mapuche and Tehuelche aboriginal samples during the early 1990s, to Dr. Cecilia Bobillo for providing Toba samples from Chaco province and to Msc. Consuelo Martí for providing samples of the Mocoví and Toba populations from the Santa Fe province. DC and MC are members of Carrera del Investigador Científico y Tecnológico-CONICET, Argentina. Financial support was provided by grants 20020100100744 UBACyT (University of Buenos Aires) and PIP 112–200801–02836 (CONICET) to DC.

LITERATURE CITED

Avena S, Goicoechea A, Rey J, Dugoujon JM, Dejean C, Carnese FR. 2006. Mezcla génica en una muestra poblacional de la Ciudad de Buenos Aires. Medicina (Buenos Aires) 66:113–118.

Avena S, Via M, Ziv E, Pérez-Stable E, Gignoux C, Dejean C, Huntsman S, Torres-Mejía G, Dutil J, Matta J et al. 2012. Heterogeneity in genetic admixture across different regions of Argentina. PLoS One 7:e34695.

Barnholtz-Sloan JS, Pfaff CL, Chakraborty R, Long JC. 2005. Informativeness of the CODIS STR loci for admixture analysis. J Forensic Sci 50:1322–1326.

Bechis M. 2010. Piezas de etnohistoria y de antropología histórica. Buenos Aires: Sociedad Argentina de Antropología.

Bjerg M. 2007. Vínculos mestizos. Historias de amor y parentesco en la campaña de Buenos Aires, del siglo XIX. Buenos Aires: Boletín del Instituto de Historia Argentina y Americana Dr. Emilio Ravignani.

Bobillo MC, Zimmermann B, Sala A, Huber G, Rock A, Bandelt HJ, Corach D, Parson W. 2010. Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: towards a first nationwide forensic mitochondrial DNA sequence database. Int J Legal Med 124: 263–268.

Brown MD, Hosseini SH, Torroni A, Bandelt HJ, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC. 1998. mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? Am J Hum Genet 63:1852–1861.

Catanesi CI, Martina PF, Giovambattista G, Zukas P, Vidal-Rioja L. 2007. Geographic structure in Gran Chaco Amerindians based on five X-chromosome STRs. Hum Biol 79:463–474.

Cavalli-Sforza L, Piazza A, Menozzi P. 1994. History and geography of human genes. Princeton: Princeton University Press.

Corach D, Lao O, Bobillo M, van Der Gaag K, Zuniga S, Vermeulen M, van Duijn K, Goedbloed M, Vallone PM, Parson W, de Knijff P, Kayser M, 2010. Inferring continental ancestry of argentineans from Autosomal, Y-chromosomal and mitochondrial DNA. Ann Hum Genet 74:65–76.

Dana L. 2006. Indigenous peoples in Chile. Int J Entrepreneurship and Small Business 3:779–786.

de Saint Pierre M, Bravi C, Motti J, Fuku N, Tanaka M, Llop E, Bonatto SL, Moraga M. 2012. An alternative model for the early peopling of southern South America revealed by analyses of three mitochondrial DNA haplogroups. PLoS One 7:e43486.

Demarchi D, Garcia Ministro A. 2008. Genetic structure of native populations from the Gran Chaco region, South America. Int J Hum Genetics 8:131–141.

Demarchi DA, Mitchell RJ. 2004. Genetic structure and gene flow in Gran Chaco populations of Argentina: evidence from Y-chromosome markers. Hum Biol 76:413–429.

Dornelles C, Bonatto S, De Freitas L, Salzano F. 2005. Is haplogroup X present in extant South American Indians? Am J Phys Anthropol 127: 439–448.

Excoffier L, Laval G, Schneider S. 2007. Arlequin (version 3.5): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:43–50.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Garcia F, Moraga M, Vera S, Henriquez H, Llop E, Aspillaga E, Rothhammer F. 2006. mtDNA microevolution in Southern Chile's archipelagos. Am J Phys Anthropol 129:473–481.

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, Just J, Salzano FM, MC. K. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. EXS 67:211–219.

Glesmann LA, Martina PF, Catanesi CI. 2013. Genetic variation of X-STRs in the Wichi population from Chaco Province, Argentina. Hum Biol 85:687–698.

Goicoechea AS, Carnese FR, Dejean C, Avena SA, Weimer TA, Franco MH, Callegari-Jacques SM, Estalote AC, Simoes ML, Palatnik M, Salzano FM. 2001. Genetic relationships between Amerindian populations of Argentina. Am J Phys Anthropol 115:133–143.

Greenberg J, Ruhlen M. 2007. An Amerind etymological dictionary. Stanford, CA: Stanford University Press.

Hajduk A, Ménica Cúneo E. 1997. Rescate Arqueológico en San Martin de Los Andes (Departamento Lácar, Provincia de Neuquén) y algunas reflexiones acerca de la cerámica con decoración valdiviana. Relaciones de la Sociedad Argentina de Antropología XX1I-XXII1, 1997-1998 Buenos Aires 22:319–341.

Halverson MS, Bolnick DA. 2008. An ancient DNA test of a founder effect in Native American ABO blood group frequencies. Am J Phys Anthropol 137:342–347.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

Marino M, Corach D, Sala A. 2006a. Relevant genetic contribution of Amerindian to the extant population of Argentina. Int Congress Series 1288:397–399.

Marino M, Sala A, Corach D. 2006b. Genetic analysis of the populations from Northern and Mesopotamian provinces of Argentina by means of 15 autosomal STRs. Forensic Sci Int 160:224–230.

Marino M, Sala A, Corach D. 2007. Genetic attributes of the YHRD minimal haplotype in 10 provinces of Argentina. Forensic Sci Int Genet 1:129–133.

Marrero AR, Silva-Junior WA, Bravi CM, Hutz MH, Petzl-Erler ML, Ruiz-Linares A, Salzano FM, Bortolini MC. 2007. Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. Am J Phys Anthropol 132:301–310.

Martínez Marignac V, Bertoni B, Parra E, Bianchi N. 2004. Characterization of admixture in an urban sample from Buenos Aires, Argentina, using uniparentally and biparentally inherited genetic markers. Hum Biol 76:543–557.

Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273.

O'Rourke DH, Mobarry A, Suarez BK. 1992. Patterns of genetic variation in Native America. Hum Biol 3:417–434.

Peakall R, Smouse P. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295.

Peakall R, Smouse P. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539.

Phillips C, Rodriguez A, Mosquera-Miguel A, Fondevila M, Porras-Hurtado L, Rondon F, Salas A, Carracedo A, Lareu MV. 2008. D9S1120, a simple STR with a common Native American-specific allele: forensic optimization, locus characterization and allele frequency studies. Forensic Sci Int Genet 3:7–13.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Quinque D, Kittler R, Kayser M, Stoneking M, Nasidze I. 2006. Evaluation of saliva as a source of human DNA for population and association studies. Anal Biochem 353:272–277.

Ratto S. 2003. La frontera bonaerense (1810-1828): espacio de conflicto, negociación y convivencia. La Plata: Archivo Histórico de la Provincia de Buenos Aires "Dr. Ricardo Levene."

Raymond M, Rousset F. 1995. An exact tes for population differentiation. Evolution 49:1280–1283.

Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra MV, Rojas W, Duque C, Mesa N, García LF, Triana O, Blair S, Maestre A, Dib JC, Bravi CM, Bailliet G, Corach D, Hünemeier T, Bortolini MC, Salzano FM, Petzl-Erler ML, Acuña-Alonzo V, Aguilar-Salinas C, Canizales-Quinteros S, Tusié-Luna T, Riba L, Rodríguez-Cruz M, Lopez-Alarcón M, Coral-Vazquez R, Canto-Cetina T, Silva-Zolezzi I, Fernandez-Lopez JC, Contreras AV, Jimenez-Sanchez G, Gómez-Vázquez MJ, Molina J, Carracedo A, Salas A, Gallo C, Poletti G, Witonsky DB, Alkorta-Aranburu G, Sukernik RI, Osipova L, Fedorova SA, Vasquez R, Villena M, Moreau C, Barrantes R, Pauls D, Excoffier L, Bedoya G, Rothhammer F, Dugoujon JM, Larrouy G, Klitz W, Labuda D, Kidd J, Kidd K, Di Rienzo A, Freimer NB, Price AL, Ruiz-Linares A. 2012. Reconstructing Native American population history. Nature 488: 370–374.

Rosenberg N. 2004. Distruct: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138.

Saguier Negrete E. 1964. El sistema sanguineo Diego y el origen mongoloide de nuestros indios. Rev Med Parag 1:30–34.

Sala A, Arguelles CF, Marino M, Bobillo M, Fenocchio A, Corach D. 2010. Genetic analysis of six communities of Mbya-Guarani inhabiting north-eastern Argentina by means of nuclear and mitochondrial polymorphic markers. Hum Biol 82:433–456.

Sala A, Corach D. 2014. Analysis of admixture and genetic structure of two Native American groups of Southern Argentinean Patagonia. Mol Biol Rep 41:1533–1543.

Sala A, Penacino G, Carnese R, Corach D. 1999. Reference database of hypervariable genetic markers of Argentina: application for molecular anthropology and forensic casework. Electrophoresis 20:1733–1739.

Salazar-Flores J, Zuñiga-Chiquette F, Rubi-Castellanos R, Alvarez-Miranda JL, Zetina-Hérnandez A, Martínez-Sevilla VM, González-Andrade F, Corach D, Vullo C, Alvarez JC, Lorente JA, Sánchez-Diz P, Herrera RJ, Cerda-Flores RM, Muñoz-Valle JF, Rangel-Villalobos H. 2015. Admixture and genetic relationships of Mexican Mestizos regarding Latin American and Caribbean populations based on 13 CODIS-STRs. Homo 66:44–59.

Salzano F, Callegari-Jacques S. 1988. South American Indians: a case study in evolution. New York: Oxford University Press.

Sambrook J, Fritsch E, Maniatis T. 1989. Molecular cloning. A laboratory manual. NY: Cold Spring Harbor Laboratory Press.

Sans M. 2000. Admixture studies in Latin America: from the 20th to the 21st century. Hum Biol 77:155–177.

Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, Conrad DF, Tito RY, Osipova LP, Tarskaia LA, Zhadanov SI, Wall JD, Pritchard JK, Malhi RS, Smith DG, Rosenberg NA. 2009. Haplotypic background of a private allele at high frequency in the Americas. Mol Biol Evol 26: 995–1016.

Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. 2007. A private allele ubiquitous in the Americas. Biol Lett 3:218–223.

Tarragó M. 2000. Nueva Historia Argentina: Los pueblos originarios y la conquista. Buenos Aires: Sudamericana.

Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93:196–200.

Wallace DC, Torroni A. 1992. American Indian prehistory as written in the mitochondrial DNA: a review. Hum Biol 64:403–416.

Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K,Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A. 2007. Genetic variation and population structure in native Americans. PLoS Genet 3:e185.

Ward RH, Redd A, Valencia D, Frazier B, Paabo S. 1993. Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 90: 10663–10667.

Weir B. 1996. Genetic data analysis II: methods for discrte population genetic data. Sunderland, MA: Sinauer Association.

Weir B, Cockerham C. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

Williams RC, Steinberg AG, Gershowitz H, Bennett PH, Knowler WC, Pettitt DJ, Butler W, Baird R, Dowda-Rea L, Burch TA, et al. 1985. GM allotypes in Native Americans: evidence for three distinct migrations across the Bering land bridge. Am J Phys Anthropol 66:1–19.

Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. 2004. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Mol Biol Evol 21:164–175.

Zhivotovsky LA, Rosenberg NA, Feldman MW. 2003. Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. Am J Hum Genet 72:1171–1186.

Zuccarelli G, Alechine E, Caputo M, Bobillo C, Corach D, Sala A. 2011. Rapid screening for Native American mitochondrial and Y-chromosome haplogroups detection in routine DNA analysis. Forensic Sci Int Genet. 2011 Mar; 5(2):105–108. doi: 10.1016/j.fsigen.2010.08.018. Epub 2010 Sep 28.