

RESEARCH ARTICLE

Genetic admixture patterns in Argentinian Patagonia

María Laura Parolin^{1*}, Ulises F. Toscanini², Irina F. Velázquez¹, Cintia Llull², Gabriela L. Berardi², Alfredo Holley¹, Camila Tamburrini¹, Sergio Avena^{3,4}, Francisco R. Carnese³, José L. Lanata⁵, Noela Sánchez Carnero⁶, Lucas F. Arce¹, Néstor G. Basso¹, Rui Pereira^{7,8}, Leonor Gusmão^{7,9}

1 Instituto de Diversidad y Evolución Austral (IDEAus), CCT CONICET-CENPAT, Puerto Madryn, Argentina, **2** Primer Centro Argentino de Inmunogenética (PRICAI), Fundación Favaloro, Buenos Aires, Argentina, **3** Instituto de Ciencias Antropológicas (ICA), Facultad de Filosofía y Letras, Universidad de Buenos Aires, Buenos Aires, Argentina, **4** Centro de Estudios Biomédicos, Biotecnológicos, Ambientales y Diagnóstico (CEBBAD), Universidad Maimónides, Buenos Aires, Argentina, **5** Instituto de Investigaciones en Diversidad Cultural y Procesos de Cambio (IIDyPCa), CONICET-UNRN, San Carlos de Bariloche, Argentina, **6** Centro para el Estudio de Sistemas Marinos (CECIMAR), CCT CONICET-CENPAT, Puerto Madryn, Argentina, **7** Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal, **8** Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal, **9** DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil

* parolin@cenpat-conicet.gob.ar



OPEN ACCESS

Citation: Parolin ML, Toscanini UF, Velázquez IF, Llull C, Berardi GL, Holley A, et al. (2019) Genetic admixture patterns in Argentinian Patagonia. *PLoS ONE* 14(6): e0214830. <https://doi.org/10.1371/journal.pone.0214830>

Editor: Francesc Calafell, Universitat Pompeu Fabra, SPAIN

Received: March 15, 2019

Accepted: June 3, 2019

Published: June 17, 2019

Copyright: © 2019 Parolin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: Financial support was granted by Agencia Nacional de Promoción Científica y Tecnológica, Argentina (ANPCyT; <https://www.argentina.gob.ar/ciencia/agencia>; grants ref. MPL PICT 2013-2414, JLL PICT 2014-1558), and Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil (CNPq; <http://www.cnpq.br>; grant ref. LG 305330/2016-0).

Abstract

As in other Latin American populations, Argentinians are the result of the admixture amongst different continental groups, mainly from America and Europe, and to a lesser extent from Sub-Saharan Africa. However, it is known that the admixture processes did not occur homogeneously throughout the country. Therefore, considering the importance for anthropological, medical and forensic researches, this study aimed to investigate the population genetic structure of the Argentinian Patagonia, through the analysis of 46 ancestry informative markers, in 433 individuals from five different localities. Overall, in the Patagonian sample, the average individual ancestry was estimated as 35.8% Native American (95% CI: 32.2–39.4%), 62.1% European (58.5–65.7%) and 2.1% African (1.7–2.4%). Comparing the five localities studied, statistically significant differences were observed for the Native American and European contributions, but not for the African ancestry. The admixture results combined with the genealogical information revealed intra-regional variations that are consistent with the different geographic origin of the participants and their ancestors. As expected, a high European ancestry was observed for donors with four grandparents born in Europe (96.8%) or in the Central region of Argentina (85%). In contrast, the Native American ancestry increased when the four grandparents were born in the North (71%) or in the South (61.9%) regions of the country, or even in Chile (60.5%). In summary, our results showed that differences on continental ancestry contribution have different origins in each region in Patagonia, and even in each locality, highlighting the importance of knowing the origin of the participants and their ancestors for the correct interpretation and contextualization of the genetic information.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Argentina, like many other Latin American countries, has a large proportion of inhabitants with admixed ancestry, with a high Native American and European contribution and to a lesser extent from Africa. Soon after the Spanish conquest in the beginning of the 16th century, and until the first half of the 18th century, the Buenos Aires port received more than 25,000 African slaves, who mainly came from West Africa and Mozambique [1]. The most significant European migration occurred between 1860 and 1950. In that period, Argentina welcomed an estimated 5.5 million migrants, mainly from Spain and Italy [2,3]. According to the 1869 census, the Argentinian population had only 1.8 million inhabitants (National Institute of Statistics and Census of Argentina (INDEC), 2010), placing this migratory event as the most significant worldwide, considering the size of the recipient population. In this context, and for more than a century, the popular imaginary defined Argentina as a “white” country where most of its population was supposed to descend from European immigrants. However, the geographic distribution of the Europeans was strongly biased towards the city of Buenos Aires and the rest of the Central region of Argentina [2,3]. On the other hand, the Argentinian Patagonia was, until the last third of the 19th century, a virtually control-free area of the Republican state, a situation that allowed the native populations to preserve their autonomy for a long period [4]. This region lacked permanent settlements until the late 1800, when the so-called “Conquista del Desierto” (Desert Conquest) annexed the Patagonia to the Argentine Republic.

Until then, the Argentinian Patagonia had been inhabited mainly by hunter-gatherer native groups of low demographic density and high residential mobility. Although these groups had a long history of sporadic contacts with foreigners, only at the end of the 19th century there was a significant demographic change in the region, due to strong migratory flows coming from inland provinces and from bordering countries, mainly attracted by new labor opportunities [5].

Our research group has previously investigated the genetic composition of urban Argentinian populations with focus on the Patagonia region by means of uniparental and autosomal markers. In five populations of Chubut and Río Negro provinces, the analysis of mitochondrial DNA and Y-chromosome markers revealed heterogeneity through autochthonous and allochthone contributions. In this sense, mitochondrial analysis showed a higher Native American maternal contribution in populations of Esquel, San Carlos de Bariloche, Comodoro Rivadavia, Puerto Madryn and Trelew [4,6–9]. For the same populations, the Native American fraction of parental lineages, analyzed by Y-STR and Y-SNPs, was relatively high in the Andean populations of Esquel and San Carlos de Bariloche, and lower in the coastal cities of Comodoro Rivadavia, Puerto Madryn and Trelew [10]. These studies showed that a biased admixture process between European males and Native women lead to a partial replacement of native paternal lineages with allochthone ones. This process has been widely observed in other regions of Argentina [11–14], as well as in other countries in South America [15].

Regarding autosomal markers, the studies carried out in urban populations of Argentinian Patagonia varied on the number of populations, the source populations and the number of samples, and the genetic markers analyzed. In this respect, Toscanini et al. [16] analyzed 46 ancestry informative markers (AIMs) in samples from Santa Cruz and Neuquén provinces, and Corach et al. [13] studied 24 AIMs in samples from Río Negro. These studies showed similar results for the Native American (32.3%-27.7%), European (68.4%-65.6%) and African (3.8%-2%) genetic contributions.

Studies carried out by our research team in Argentinian Patagonia comprised the analysis of blood systems [4,6,7], autosomal STRs [17], AIMs [18] and Alu markers [19]. Admixture results obtained showed relatively wide ranges of continental ancestry between markers: 44–

29% Native American, 67–51% European, and 4–2% African. The study based on Alu markers showed that the admixture process was not homogeneous across Patagonia region [19].

The most comprehensive ancestry studies by Seldin et al. [20], Wang et al. [21] and Muzzio et al. [22] were mainly focused on the differential distribution of Native American, European and African components in the Central and Northern regions of Argentina, while samples from the Southern region of the country were not analyzed.

The present study aimed to analyze a panel of 46 ancestry informative insertion-deletion markers (AIM-Indels), described by Pereira et al. [23], in a relatively large number of individuals ($n = 433$), to contribute to a better understanding of the genetic structure and different admixture processes that occurred in urban Argentinian Patagonia populations. In the original publication [23], the authors showed that the 46 AIM-Indel multiplex is suitable to distinguish admixture proportions from Asian, European, African and Native American origin. Furthermore, in a collaborative study by different laboratories, apart from validating the method for forensic samples and mixtures analysis, the authors showed that it also performs well in assigning ancestry to samples of unknown geographic origin [24].

For a more in-depth interpretation, with focus in the regional and local differences, the ancestry results were analyzed taking into account genealogical and historical data. This information allowed for investigating possible associations between genetic ancestries inferred by the 46 AIMs panel and the geographic origin of the participants.

The knowledge of the genetic composition of Argentinian Patagonia, in terms of individual and population ancestry composition will provide important information for medical, forensic and anthropological studies in the region.

Materials and methods

Ethics statement

This study was approved by the Ethics Committees of Puerto Madryn Zonal Hospital (Resol.009/2015) and San Carlos de Bariloche Zonal Hospital (Resol. 1510/2015). Moreover, it has been endorsed by Hospitals' Teaching and Research Committees of Trelew, Comodoro Rivadavia and Esquel. Biological samples involved in the study were coded and anonymized, as per the Argentina National Law 25.326 of Protection of Personal Data. All samples were collected under written informed consent to participate in this study.

Biological samples and DNA extraction

A volume of 5 ml of blood was EDTA collected from 433 healthy non-related voluntary blood donors, randomly selected from hemotherapy services of public and private hospitals of five Argentinian Patagonia cities; two from the Andes region, namely San Carlos de Bariloche and Esquel, and three located in the coastal area: Puerto Madryn, Trelew, and Comodoro Rivadavia (Fig 1).

Participants were thoroughly informed of the scopes of the study and answered a questionnaire to obtain information about their birthplace and that of the two preceding generations (parents and grandparents). This information allowed us to evaluate a possible association between genetic ancestry and geographic origin. DNA extraction was performed as described in Parolin et al. [19].

Genotyping of the AIM-Indels

A panel of 46 AIM-Indels, developed by Pereira et al. [23] was genotyped in a single multiplex PCR followed by capillary electrophoresis. PCR thermocycling conditions were: an initial step

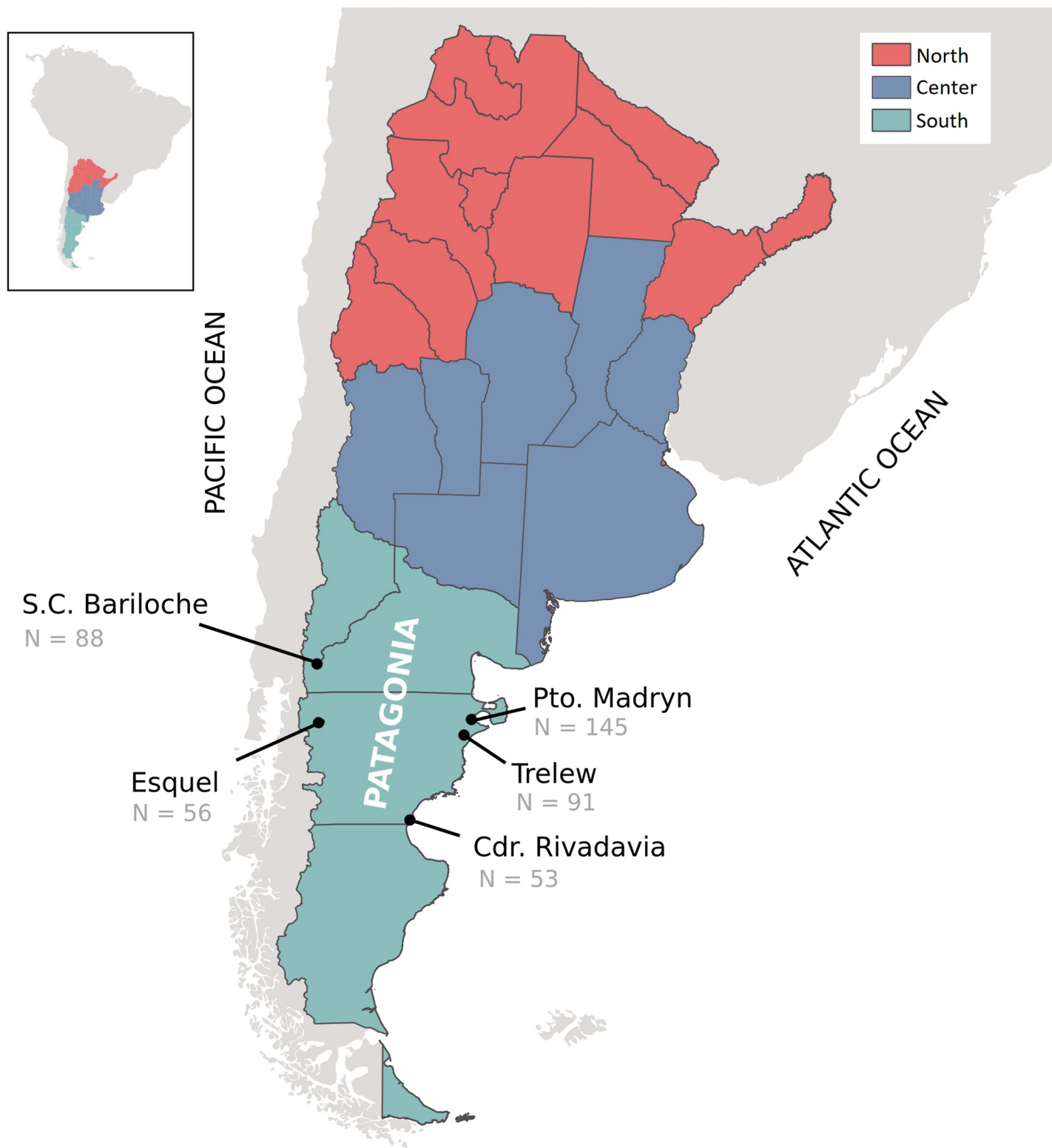


Fig 1. Map of Argentina showing the three main regions of the country, and the five studied populations.

<https://doi.org/10.1371/journal.pone.0214830.g001>

of 15 min at 95°C; followed by 27 cycles at 94°C for 30 s, 60°C for 1.5 min, 72°C for 45 s; and a final extension at 72°C for 60 min. Dye-labeled amplified fragments were separated and detected using an ABI3130 Genetic Analyzer (Life Technologies), and automated allele calls were obtained with GeneMapper v.3.2 (Life Technologies). This panel showed to be highly informative to analyze genetic admixture in different Latin-American populations, including Argentina [16], Brazil [25], Bolivia [26] and Colombia [27,28].

Statistical analysis

Allele frequency estimations, Hardy-Weinberg exact tests, and population differentiation tests were performed using the Arlequin software v.3.5.2 [29]. Population genetic analyses were carried out comparing our five populations with an additional six Argentinian populations [16] and six South American populations from Brazil [25], Bolivia [26] and Colombia [27,28]. Genetic distances (F_{ST}) between populations across 46 AIM-Indel markers were visualized in a multi-dimensional scaling (MDS) plot performed with the R v3.1.1 FactoMineR [30] statistical package. The proportion of genetic continental ancestry was estimated using the STRUCTURE v2.3.3 software [31]. In agreement with the historical and demographic settlement of Argentina we assumed a tri-hybrid model with three parental groups: Native American (NAM), European (EU) and African (AF) (i.e. $K = 3$). The allele frequencies of the Native American, European and African parental populations, from the diversity panel of human cell line samples HGDP-CEPH (Human Genome Diversity Project), were taken from Pereira et al. [23]. STRUCTURE runs consisted of 100,000 burning steps followed by 100,000 Markov Chain Monte Carlo (MCMC) iterations. For the Admixture model, the option "Use population Information to test for migrants" was used. Allele frequencies were correlated and updated only with individuals with POPFLAG = 1. Population admixture structure was represented in a ternary plot, including individual genetic ancestry estimated in the samples studied. The ternary plot was performed by the ggtern package of R v3.1.1 [32]. Average individual ancestry, with a 95% confidence interval (CI), was estimated based on: i) parametric t distribution, if the sample size was higher than 29, and ii) non-parametric bootstrapping method, if the sample size was smaller than 30. Differences in the mean Native American, European and African ancestries, between the five studied populations and between the samples with four grandparents born in different geographic places, were analyzed using the non-parametric one-way ANOVA Kruskal-Wallis test [33]. This non-parametric approach was selected because the distribution of the residuals of the parametric ANOVA deviated from normality. When the Kruskal-Wallis test was significant ($p < 0.05$), population pairwise differences were tested by the Dunn post-hoc analysis [34]. A Chi-squared test [35] was performed to analyze the independence between the mean genetic ancestry obtained (Native American/European/African) and the regional origin (North/Center/South Argentina). Statistical analyses were performed by standard R v3.1.1 functions [32] implementing *ad-hoc* scripts. Geographic representations of the birthplace and ancestry proportions of the donors were obtained by ordinary kriging using the Geostatistical Analyst module of ArcMap 9.3 (ESRI). The covariances of the studied parameters (percentage of ancestry) were modeled in terms of the distance between sampling units [36]. The maps were created with QGIS (<http://qgis.osgeo.org>) and ArcGIS Desktop 10.6.1 software using spatial data obtained from GADM webpage (<https://gadm.org/license.html>), which are freely available for academic purpose.

Results and discussion

The genotyping results obtained in the five urban populations from Argentinian Patagonia are shown in S1 Table. Allele frequencies and expected heterozygosities were estimated for the 46

AIM-Indels are presented in S2 and S3 Tables, respectively. Exact test showed no significant deviations to the Hardy-Weinberg equilibrium expectations for all loci in the five populations ($p \geq 0.0038$; for a significant level of 0.001, when applying the Bonferroni's correction for 46 tests performed) [37].

Genetic ancestry analysis

The estimation of continental admixture in 433 Argentine Patagonian individuals varied from 0.4 to 99.4% Native American, 0.3 to 99.2% European, and 0.2 to 23.4% African. Individual genetic ancestry obtained was represented in a ternary plot (Fig 2).

The ternary plot showed a linear distribution of the Patagonia samples between European and Native American parental populations, with a tendency of more density towards the European component.

The average individual ancestry estimated in the total Patagonia sample was 35.8% (95% CI: 32.2–39.4%) Native American, 62.1% (95% CI: 58.5–65.7%) European, and 2.1% (95% CI: 1.7–2.4%) African. Comparing the five studied localities (Fig 3), a prevalence of the Native American ancestry was observed in Esquel and in Comodoro Rivadavia, located in western and southeastern Chubut, respectively. Meanwhile Puerto Madryn, from northeastern Chubut, was the Patagonian population with higher European contribution. San Carlos de

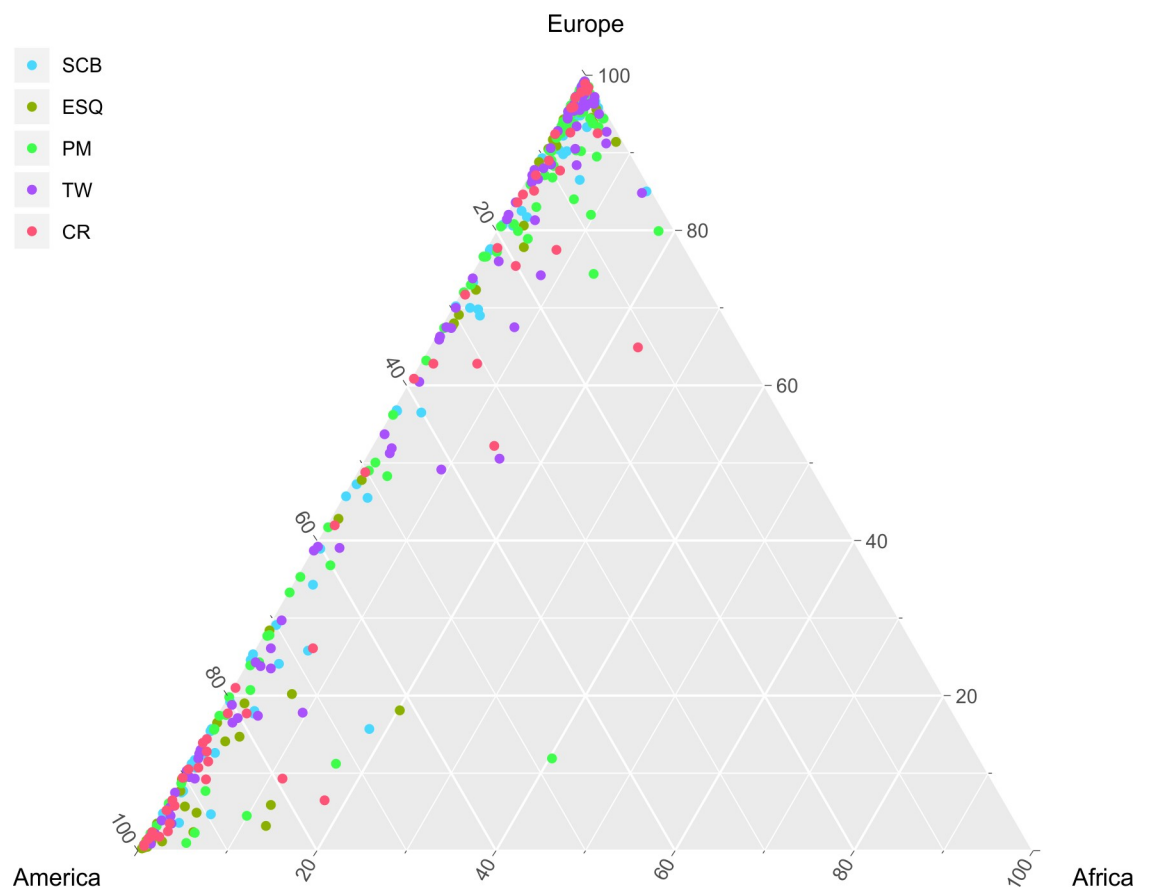


Fig 2. Ternary plot showing the continental ancestry distribution of the Patagonian samples. SCB: San Carlos de Bariloche, ESQ: Esquel, PM: Puerto Madryn, TW: Trelew, CR: Comodoro Rivadavia.

<https://doi.org/10.1371/journal.pone.0214830.g002>

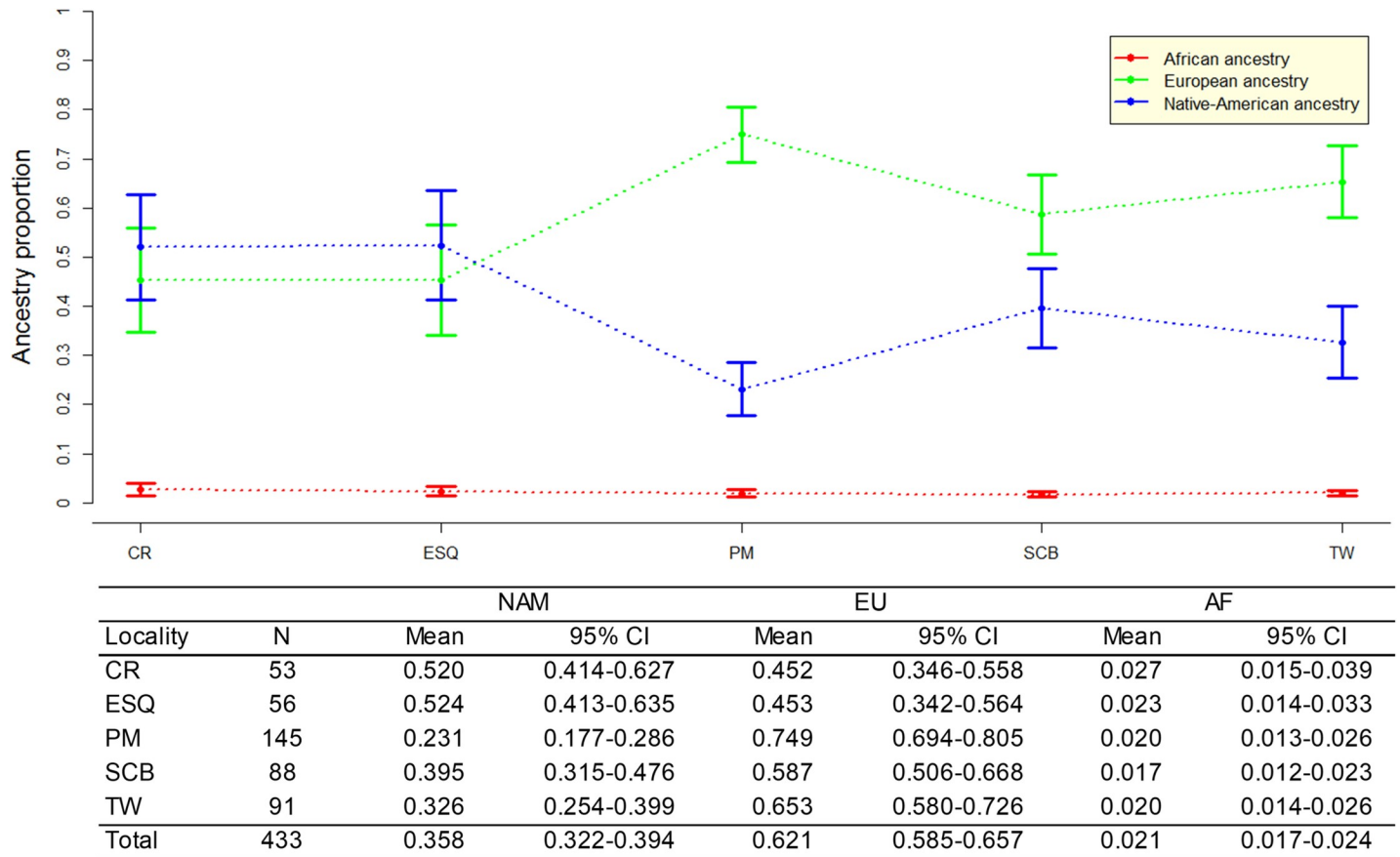


Fig 3. Average individual ancestry and the associated 95% CI for the five studied population. The blue line represents the Native American component (NAM), the green line the European component (EU), and the red line the African component (AF).

<https://doi.org/10.1371/journal.pone.0214830.g003>

Bariloche and Trelew were closer to the Native American and European averages in the global sample. The African ancestry was observed in similar low proportions in all populations.

The Kruskal-Wallis test showed that the differences observed for the Native American and European ancestry distributions among the five samples are statistically significant ($H \geq 45.285$; $p < 0.0001$), while no significant differences ($H = 4.631$; $p = 0.33$) were observed for the African ancestry. In all pairwise comparisons, significant differences ($p < 0.05$) were observed for both for Native American and European ancestry proportions, except between Comodoro Rivadavia and Esquel (NAM $p = 0.389$, EU $p = 0.439$) and between Trelew and San Carlos de Bariloche (NAM $p = 0.1124$, EU $p = 0.152$).

The overall results showed three main groups among central Patagonian localities: i) Comodoro Rivadavia and Esquel, ii) Trelew and San Carlos de Bariloche, and iii) Puerto Madryn.

Aiming to compare the ancestry distribution between the three main regions of the country (North/Center/South), our data was analyzed together with those obtained by Toscanini et al. [16], for the same set of 46 AIM-Indels. As also found by Toscanini et al. [16], the Chi-square test showed that the ancestry varies between regions ($p = 0.003$). In this respect, the European contribution appeared to be markedly high in the Central region of Argentina (81.4%), while the Native American component increases toward the North (38.8%) and the South (37.8%) of the country (S1 Fig).

Differences on the ancestry of Argentinian regions were also found in other studies, based on 24 AIMs [13] and 99 AIMs [18] (S4 Table). For the Southern region, a high Native American contribution was registered in all studies. However, in Corach et al. [13], the European component appears to be overestimated in the South and in the North regions. This could be due to differences in the markers analyzed and/or in the populations sampled. In fact, as noted above, a large variation exists among South region populations, and it is possible that the same occurs in the North region. It is also worth noting that sampling in both, Corach et al. [13] and Toscanini et al. [16], was based on the place of residence of the individuals and no information on the birthplaces of the participants or their relatives is provided. The higher Native American contribution found in the South by Avena et al. [18], when compared to other studies, can also be explained by the heterogeneity of the populations studied from this region. Indeed, their study included samples from the two localities, Esquel and Comodoro Rivadavia, showing the highest Native American contributions, in this study and in previous studies using Alu autosomal markers (42.3% and 40.5%, respectively) [19], and mtDNA (78% and 73%, respectively) [4,6].

At regional level, the relatively high preservation of the Native American ancestry obtained in current urban Patagonia populations might be related to the late incorporation of this region to the National State (Law 1584, 1884), which allowed the Native populations to keep their autonomy for longer. Also, it might be related to the high migratory flow that the Patagonia region received from the inland provinces and from bordering countries, mainly Chile, with high Native American genetic background [5,38–40].

Likewise, at the local level, the admixture results obtained seem to be related to the birthplace of the participants and their ancestors, since a high variation was observed among grandparents' birthplaces, for the five Patagonian samples. Comparing to the remaining samples, Esquel showed the lowest immigration from other regions, allowing for a better preservation of the original genetic background (Fig 4). Indeed, Esquel showed the highest average of Native American ancestry (52.4%; Fig 3), and the highest frequency of grandparents born in the same place (25.7%) (Fig 4).

This could be related with the gradual and diversified growth of the population, and a better preservation of the autochthonous origins [4,41]. Since before its foundation (1906) to the present day, Esquel has maintained a strong trade with Chile, being a zone of transport and gateway of migrants in both sides of the Andes (19.3% of the grandparents born in Chile). San Carlos de Bariloche, the most populated southern Andean city (512 inhabitants/km²), registered 39.5% of Native American ancestry (Fig 3), and 10.3% of the grandparents born in the same locality. However, 27.9% of the grandparents came from the Central region, mainly attracted by the tourism industry. As Esquel, San Carlos de Bariloche also has a strong migratory flow from Chile [5], as in the sample analyzed 25.5% of grandparents born in this neighboring country. Comodoro Rivadavia city was the second population with highest Native American ancestry (52%; Fig 3). It showed the highest proportion of grandparents from Chile (38.2%) and North West Argentina (18.3%) (Fig 4). Both immigrant groups, which preserve a high Native American ancestry [18,21,22,38,41], were mainly attracted to Comodoro Rivadavia by the exponential growth of the oil industry which started in 1907 and is still ongoing [6]. Meanwhile, Puerto Madryn showed the highest European ancestry (74.9%; Fig 3), and the highest proportion of grandparents from Europe (23.7%) and Central region (32.4%) (Fig 4), which presents the highest European ancestry of the country [13,16,18,22]. As mentioned in the introduction, between later 19th century and the end of the Second World War, a large number of Europeans arrived at the Central region of Argentina, who mainly settled in Buenos Aires province, and later spread to Pampean and littoral areas, attracted by the agriculture opportunities [2,3]. In Puerto Madryn, the larger proportion of immigrants from the Central

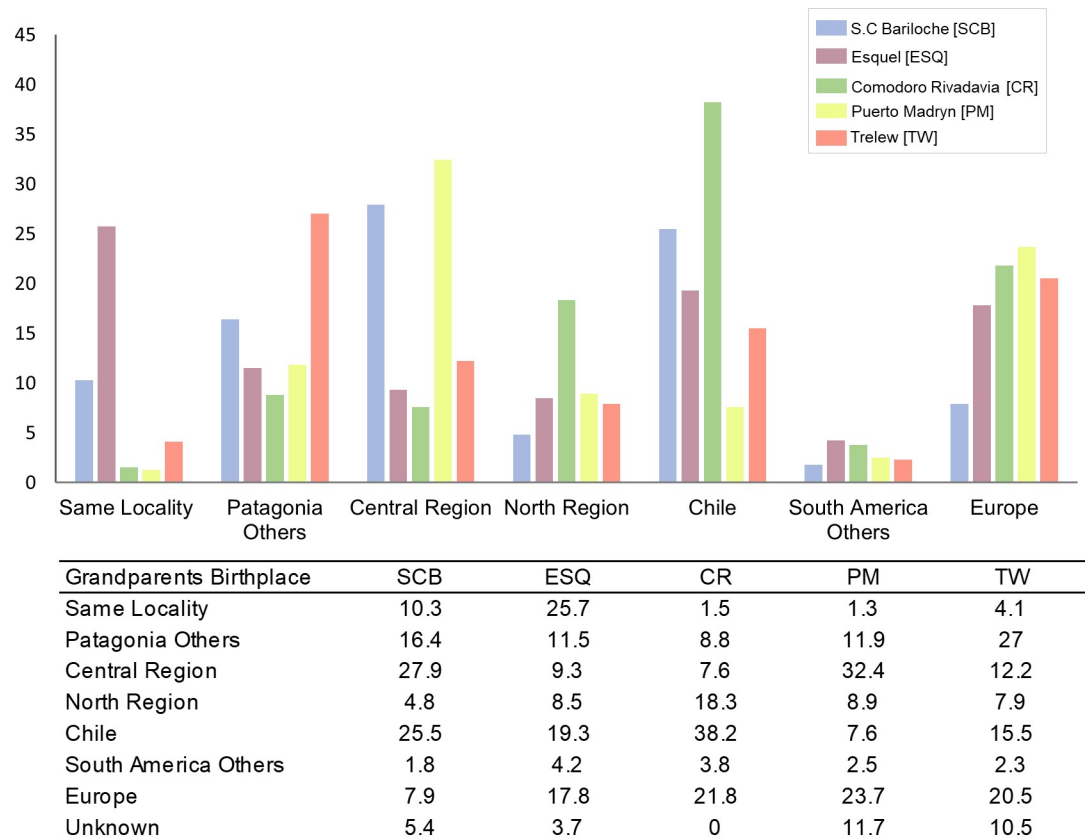


Fig 4. Proportion (%) of grandparents' birthplace by locality.

<https://doi.org/10.1371/journal.pone.0214830.g004>

region is most likely related to the demand of qualified laborers for the aluminum industry. Puerto Madryn and Trelew, 65 km distant from each other, share the same foundational origin of Welsh, Spanish and Italian roots. Despite the common historical origin and geographic proximity, Trelew showed the highest proportion of grandparents born in other Patagonian localities (27%), likely attracted to Trelew by labor sources in the textile industry and public administration [9].

The five cities sampled here have less than 160 years of existence and relative low population densities (i.e., Esquel has 27.5 inhabitants/km²). Therefore, migrations have had an important impact in the growth of their populations.

Origins and ancestry

The information about grandparents' birthplace (region/country) was compared with the proportions of Native American, European and African ancestries of the participants, to investigate a possible association between them. For this analysis, 85 individuals with the 4 grandparents born in the same geographic place were selected. These samples were grouped according to the grandparents' origin in five major regions: Europe, Chile, North Argentina, Central Argentina, and South Argentina.

Not surprisingly, the results showed that the number of individuals with grandparents from Europe was highly associated to the European ancestry in the population. The average of European ancestry among individuals with all grandparents from Europe was 96.8% (Fig 5). This is in agreement with previous observations by Avena et al. [18], suggesting that genealogical

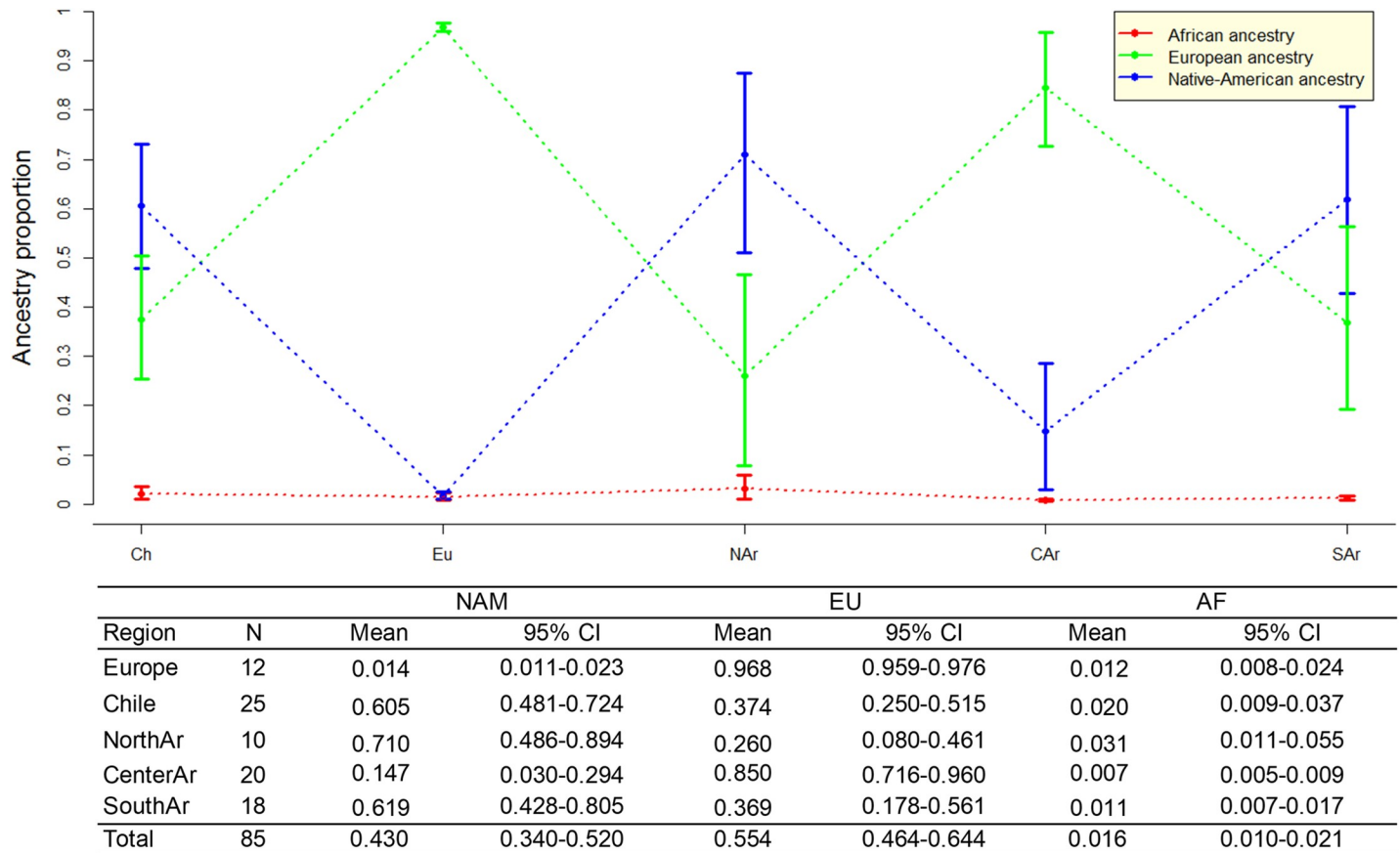


Fig 5. Average individual ancestry and the associated 95% CI obtained in samples with four grandparents born in the same geographic region. The blue line represents the Native American component (NAM), the green line the European component (EU) and the red line the African component (AF).

<https://doi.org/10.1371/journal.pone.0214830.g005>

information, such as the number of grandparents from Europe, could be a strong predictor of genetic ancestry in samples from Argentina.

Also, a clear association was observed between the Native American ancestry and the number of grandparents born in Chile. The average Native American ancestry among individuals with four grandparents from Chile was 60.5%. These findings are in accordance with the relatively high prevalence of the Native American ancestry (range 46.9%-49.5%) reported in Chilean Patagonia [39].

Comparing the three main regions of Argentina, an association was also observed between the grandparents' birthplace and the average European and Native American ancestries of the individuals. As expected, a high correlation between European ancestry and birthplaces in the Central region of Argentina was observed. The average European ancestry among individuals with all grandparents born in the Central region was 85% (Fig 5). As previously mentioned, the large number of European immigrants who arrived in Argentina between 1860 and 1950 mainly settled in Buenos Aires city and in the rest of the Pampean region [2,3].

In contrast, the Native American ancestry increased and the European decreased when the four grandparents were born in the North (NAM = 71%) or in the South (NAM = 61.9%) regions of the country (Fig 5). These results reflect the historical and demographic information described in previous section, and is consistent with the genetic studies carried out on biparental [13,16,18,19,22] and uniparental markers [4,6-9,42].

When Kruskal-Wallis test was assessed, significant differences ($H \geq 43.495$; $p < 0.001$) were observed for the Native American and European ancestry distributions among participants with four grandparents born in the three main regions of Argentina, Europe or Chile. However, no significant differences ($H = 10.219$; $p = 0.045$) were detected for the African component. In accordance with the ancestry frequencies obtained for each region (Fig 5), the *post-hoc* analysis showed no significant differences ($p > 0.05$), for the Native American and European ancestries, between donors with four grandparents born in Europe and in Central Argentina (NAM $p = 0.126$, EU $p = 0.288$), neither between donors with four grandparents born in Chile, North and South Argentina (NAM $p = 0.288$ – 0.490 , EU $p = 0.231$ – 0.476). These results are consistent with the high average European ancestry ($\geq 85\%$) in the samples with four grandparents born in Central Argentina and in Europe. Also, similar average Native American ancestry ($\geq 60\%$) was observed among participants with four grandparents born in Chile, North and South Argentina (Fig 5).

Geospatial representation of the ancestry

The results obtained showed a clear variation in ancestry proportions mainly associated to the geographical origin of the samples. In order to obtain a spatial representation of this variation, we performed a geostatistical modelling based on the geographic coordinates for the birthplaces of the participants and their individual ancestry. For this representation, 392 individuals born in North, Central and South Argentina, and in Chile were included (41 samples from other countries or with incomplete birthplace information were excluded). The distribution of the individual genetic ancestry among the samples included in this analysis varied from 0.4 to 99% Native American, 0.6 to 99.2% European, and 0.3 to 23.4% African contribution.

In agreement with previous studies [13,16,18,19,21,22,39], the maps show the highest levels of Native American ancestry in individuals born in the North-Western region (NAM = 58.3%, 95% CI: 37.1–75.9%) and in Argentinian and Chilean Patagonia (NAM = 46.4%, 95% CI: 41.6–51.2%) (Fig 6A). As observed by Muzzio et al. [22] and Avena et al. [18], the European ancestry was higher in individuals born in the North East and Central Argentina (EU = 80.6%, 95% CI: 75.6–85.6%) (Fig 6B). Finally, the African ancestry is poorly represented across the country, with the highest frequency observed in northern Argentina (AF = 4.3%, 95% CI: 1.9–7.1%) (Fig 6C). This is also in agreement with previous studies carried out in this region [13,16,18,22,42].

Genetic distance analysis

Using data from the 46 AIM-Indels, pairwise F_{ST} genetic distances were calculated between the five Patagonian populations from this study, and samples from six Argentinian provinces [16], six South American populations from Brazil [25], Bolivia [26] and Colombia [27,28], as well as the European, African and Native American HGDP-CEPH reference populations [23] (S5 Table).

Pairwise F_{ST} genetic distances showed significant differences (p -values < 0.000005) between the Argentinian populations and the parental and other South American populations except between Buenos Aires and South Brazil. Puerto Madryn showed low distances and no significant differences with Buenos Aires ($F_{ST} = 0.0017$, p -value = 0.1715) and La Pampa ($F_{ST} = 0.0012$, p -value = 0.2423). This is in agreement with the higher European ancestry registered in these populations, and with the birthplace of the grandparents of Puerto Madryn individuals, most born in the Central region of the country (Fig 4 and S1 Fig). On the other hand, no significant differences were observed between Esquel, Comodoro Rivadavia, Santa Cruz, and Tucumán (non-differentiation p -values > 0.0385 ; for a significance level of 0.00026, after

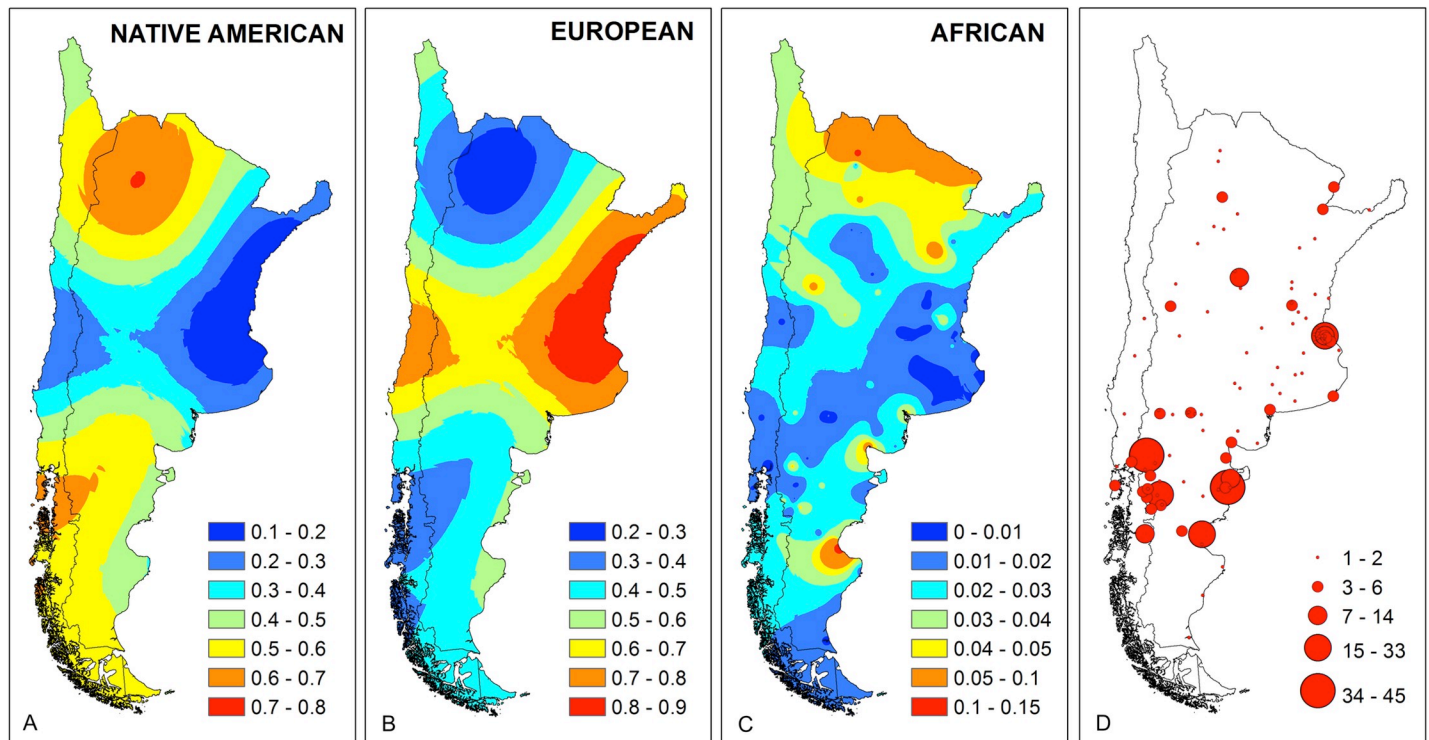


Fig 6. Maps of Argentina and Chile showing the distribution of Native American (A), European (B) and African (C) individual ancestry estimated in individuals who inhabit in Argentinian Patagonia, together with the distribution of the sampled individuals (D) by birthplace and density (red dots).

<https://doi.org/10.1371/journal.pone.0214830.g006>

applying the Bonferroni correction). This is consistent with the relatively high Native American ancestry present in these populations from the North and South regions (S1 Fig).

The pairwise F_{ST} matrix (S5 Table) was used to represent the genetic distances in a multidimensional-scaling plot (Fig 7).

In accordance with the ancestry estimates, Argentinian populations exhibited an intermediate position between European and Native American reference groups. Puerto Madryn and Trelew are closer to Central region populations, while Esquel, San Carlos de Bariloche and Comodoro Rivadavia group together with other Patagonian populations and with Tucumán.

In summary, the location of Patagonian populations in the MDS plot closely resembles the results based on the ancestry estimates and are also in accordance with our previous genetic distances analysis based on uniparental and autosomal markers carried out in central Patagonia populations [7,8,19].

Conclusion

This study constitutes the first analysis of the genetic ancestry in different populations of the Argentinian Patagonia region, complementing previous information obtained on uniparental and biparental markers. The appropriate representation of the studied samples, both in size and collection strategy, supports the obtained results. Likewise, the 46 AIMs analyzed prove to be robust and reliable for the estimation of population admixture, which has been supported by its strong association with genealogical information, mainly for European ancestry.

The different analyses performed allowed appreciating the intraregional differences in urban Argentinian Patagonia, even in a same province, where Puerto Madryn shows the highest European contribution, while Comodoro Rivadavia and Esquel preserve strong native

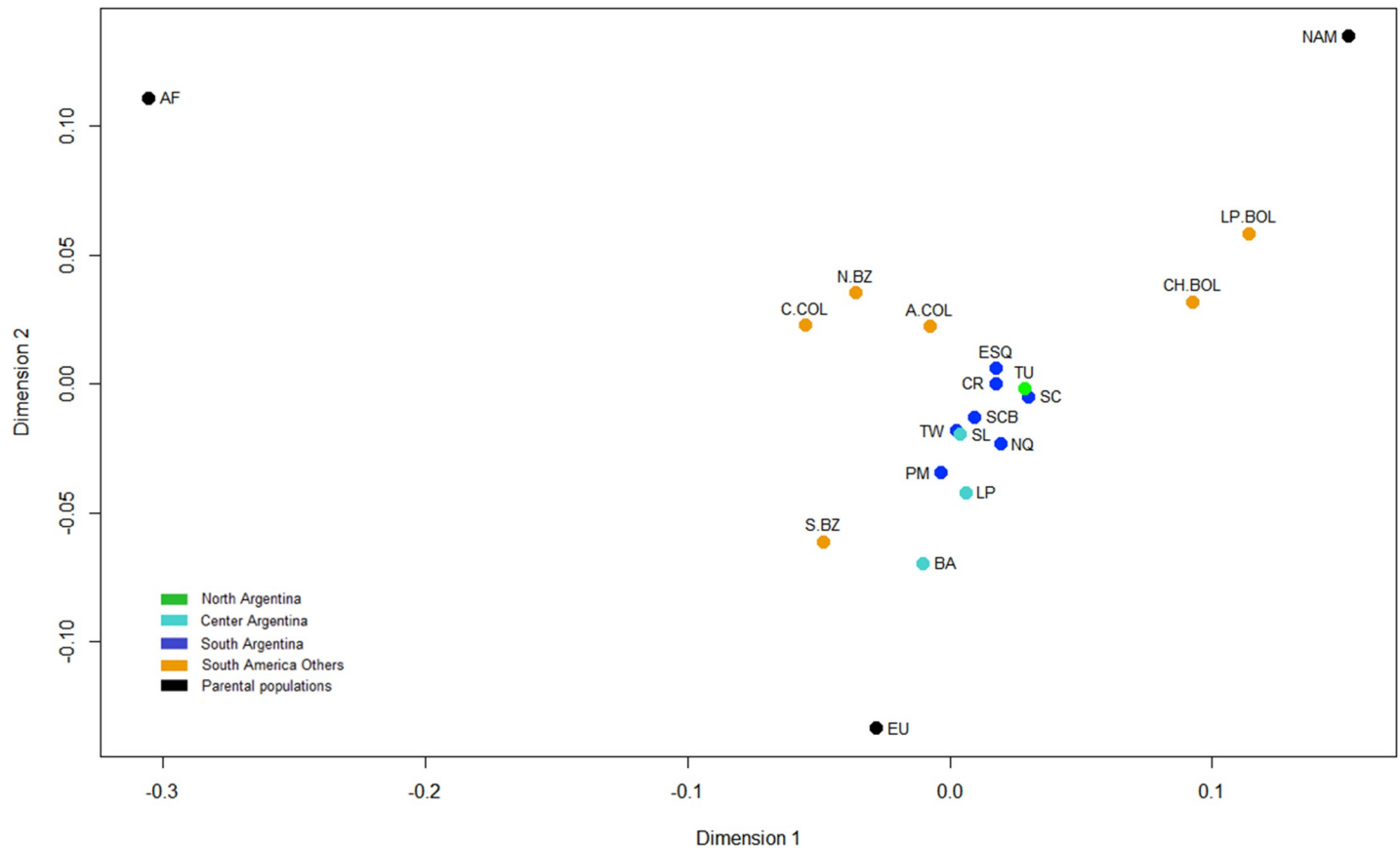


Fig 7. MDS plot of the F_{ST} pairwise genetic distances between eleven Argentine populations, other six South American populations and the three HGDP-CEPH parental populations. Stress = 0.10644. Population references are in [S5 Table](#).

<https://doi.org/10.1371/journal.pone.0214830.g007>

roots. Nevertheless, as it was demonstrated in this study, the Native American contribution have different origins in each region and even in each locality, highlighting the importance of knowing the origin of the participants and their ancestors for the correct interpretation and contextualization of the genetic information.

Moreover, the differences observed among the five localities of the same region and of the same province, as in the case of Chubut, highlight the importance of taking into account the local historical demographic process in the analysis of the genetic composition of admixed populations. Accordingly, the admixture results, combined with the genealogical information, reveal that intra-regional variations are in agreement with the different geographic origin of the participants and their grandparents', and with different local and social-demographic histories.

Finally, the heterogeneity among the populations analyzed emphasizes the need of additional studies involving rural populations to obtain an integral picture of Argentinian Patagonia, which is important for the correct interpretation of the genetic information in different studies, such as in medical, forensic and anthropological fields.

Supporting information

S1 Fig. Continental ancestry estimated for 11 Argentinian populations (5 from present work and 6 from Toscanini *et al.* [16]) grouped in the three main regions of Argentina. (TIF)

S1 Table. List of 46 AIM-Indels genotypes obtained in five urban Argentinian Patagonia populations.

(XLSX)

S2 Table. Allele frequencies for 46 AIM-Indels in five urban Argentinian Patagonia populations.

(XLSX)

S3 Table. Observed and expected values of heterozygotes and *P*-values for the exact test of Hardy-Weinberg equilibrium.

(XLSX)

S4 Table. Average ancestry observed in the three main regions of Argentina for 24, 46 and 99 AIM sets.

(XLSX)

S5 Table. Genetic distances (*F_{ST}*) between 17 South American populations, and 3 parental populations (lower diagonal), and corresponding non-differentiation *P*-values (upper diagonal).

(XLSX)

Acknowledgments

The authors greatly thank to the participants, the Directors of the Hemotherapy Services (Laura Jones, María Del Carmen Ríos Part, Silvina Fleischer, Guillermo Manera, Claudia Tedeschi and Marcelo Furque), the hemotherapys' technical teams for their support in collecting the samples, and Freya Chappell for carefully reviewing the writing.

Author Contributions

Conceptualization: María Laura Parolin, Ulises F. Toscanini, Leonor Gusmão.

Data curation: María Laura Parolin, Ulises F. Toscanini, Cintia Llull, Gabriela L. Berardi, Leonor Gusmão.

Formal analysis: María Laura Parolin, Ulises F. Toscanini, Irina F. Velázquez, Alfredo Holley, Camila Tamburrini, Noela Sánchez Carnero, Leonor Gusmão.

Funding acquisition: María Laura Parolin, Francisco R. Carnese, José L. Lanata, Néstor G. Basso, Leonor Gusmão.

Investigation: María Laura Parolin, Ulises F. Toscanini, Leonor Gusmão.

Methodology: María Laura Parolin, Irina F. Velázquez, Cintia Llull, Gabriela L. Berardi, Alfredo Holley, Camila Tamburrini, Noela Sánchez Carnero, Lucas F. Arce, Rui Pereira, Leonor Gusmão.

Project administration: María Laura Parolin, Leonor Gusmão.

Resources: María Laura Parolin, Ulises F. Toscanini, Sergio Avena, Francisco R. Carnese, José L. Lanata, Néstor G. Basso, Rui Pereira, Leonor Gusmão.

Supervision: Ulises F. Toscanini, Leonor Gusmão.

Visualization: María Laura Parolin, Ulises F. Toscanini, Leonor Gusmão.

Writing – original draft: María Laura Parolin, Ulises F. Toscanini, Leonor Gusmão.

Writing – review & editing: María Laura Parolin, Ulises F. Toscanini, Irina F. Velázquez, Alfredo Holley, Camila Tamburrini, Sergio Avena, Francisco R. Carnese, José L. Lanata, Néstor G. Basso, Leonor Gusmão.

References

1. Fejerman L, Carnese FR, Goicoechea AS, Avena SA, Dejean CB, Ward RH. African ancestry of the population of Buenos Aires. *Am J Phys Anthropol.* 2005; 128: 164–170. <https://doi.org/10.1002/ajpa.20083> PMID: 15714513
2. Devoto F. Historia de la inmigración en la Argentina. 1st ed. Buenos Aires, Argentina: Sudamericana Press; 2003.
3. Carnese FR, Avena SA, Parolin ML, Postillone MB, Dejean CR. Gene admixture estimation through genetic markers and genealogical data in a sample from Buenos Aires Metropolitan Area. In: Gibbon S, Ventura Santos R, Sans M, editors. *Racial Identities, Genetic Ancestry and Health in South America: Brazil, Argentina, Uruguay and Colombia.* New York: Palgrave Macmillan Academic Publishers; 2011. pp. 177–194.
4. Avena SA, Parolin ML, Boquet M, Dejean CB, Postillone MB, Alvarez Trentini Y, et al. Mezcla génica y linajes uniparentales en Esquel (Pcia. de Chubut). Su comparación con otras muestras poblacionales Argentinas. *J Basic Appl Genet.* 2010; 21: 01–14.
5. Bandieri S. Historia de la Patagonia. 2nd ed. Buenos Aires, Argentina: Sudamericana Press; 2011.
6. Avena SA, Parolin ML, Dejean CB, Ríos Part MC, Fabrikant G, Goicoechea AS, et al. Mezcla génica y linajes uniparentales en Comodoro Rivadavia (Provincia de Chubut, Argentina). *Rev Arg Antrop Biol.* 2009; 11: 25–41.
7. Parolin ML, Avena SA, Fleischer S, Pretell M, Di Fabio Rocca F, Rodríguez DA, et al. Análisis de la diversidad biológica y mestizaje en la ciudad de Puerto Madryn (Prov. de Chubut, Argentina). *Rev Arg Antrop Biol.* 2013; 15: 61–75.
8. Parolin ML, Baillet G, Bravi CM, Sala C, Schwab M, Paz P, et al. Variabilidad genética y autopercepción del mestizaje en la población Cordillerana de San Carlos De Bariloche (Río Negro, Arg.). XIV Congreso de la Asociación Latinoamericana de Antropología Biológica, Tacuarembó, Uruguay. ALAB Abstract Book; 2016. pp. 130–131.
9. Tamburrini C. Mestizaje, migraciones y cultura: Diversidad genética y linajes uniparentales en la localidad de Trelew (Prov. de Chubut). Bachelor thesis. Universidad Nacional de la Patagonia Press; 2016.
10. Jurado Medina LS, Parolin ML, Ramallo V, Sepúlveda PB, Real LE, Avena SA, et al. En busca de la herencia americana en poblaciones mestizas. Análisis comparativo entre muestras masculinas de Patagonia, Noroeste y Cuyo. XII Jornadas Nacionales de Antropología Biológica. Corrientes, Argentina. JNAB Abstract Book; 2015. pp. 82.
11. Avena SA, Goicoechea A, Bartomioli M, Fernandez V, Cabrera A, Dugoujon JM, et al. Mestizaje en el sur de la región pampeana (Argentina). Su estimación mediante el análisis de marcadores proteicos y moleculares uniparentales. *Rev Arg Antrop Biol.* 2007; 9: 59–76.
12. García A and Demarchi DA. Linajes parentales amerindios en poblaciones del norte de Córdoba. *Rev Arg Antrop Biol.* 2006; 8: 57–71.
13. Corach D, Lao O, Bobillo C, Van Der Gaag K, Zuniga S, Vermeulen M, et al. Inferring continental ancestry of argentineans from Autosomal, Y-chromosomal and mitochondrial DNA. *Ann Hum Genet.* 2010; 74: 65–76. <https://doi.org/10.1111/j.1469-1809.2009.00556.x> PMID: 20059473.
14. Salas A, Jaime JC, Alvarez-Iglesias V, Carracedo A. Gender bias in the multiethnic genetic composition of Central Argentina. *J Hum Genet.* 2008; 53: 662–674. <https://doi.org/10.1007/s10038-008-0297-8> PMID: 18560750.
15. Carvajal-Carmona LG, Ophoff R, Service S, Hartiala J, Molina J, Leon P, et al. Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. *Hum Genet.* 2003; 112: 534–41. <https://doi.org/10.1007/s00439-002-0899-8> PMID: 12601469.
16. Toscanini U, Gusmão L, Berardi G, Gómez A, Pereira R, Raimondi E. Ancestry proportions in urban populations of Argentina. *Forensic Sci Int Genet, Supplement Series.* 2011; 3: e387–e338. <https://doi.org/10.1016/j.fsigs.2011.09.055>.
17. Parolin ML, Real LE, Martinazzo LB, Basso NG. Population genetic analyses of the Powerplex® Fusion kit in a cosmopolitan sample of Chubut Province (Patagonia Argentina). *Forensic Sci Int Genet.* 2015; 19: 221–222. <https://doi.org/10.1016/j.fsigen.2015.07.020> PMID: 26275612.
18. Avena S, Via M, Ziv E, Perez-Stable EJ, Gignoux CR, Dejean C, et al. Heterogeneity in genetic admixture across different regions of Argentina. *PLoS ONE.* 2012; 7(4): e34695. <https://doi.org/10.1371/journal.pone.0034695> PMID: 22506044.

19. Parolin M, Zanetti D, Calò C, Esteban E, Avena S, Carnese F, et al. Migrations, admixture and genetic diversity in Central Argentinian Patagonia: analysis of autosomal Alu polymorphisms. *J Basic Appl Genet.* 2017; 28: 29–42.
20. Seldin MF, Tian C, Shigeta R, Scherbarth HR, Silva G, Belmont JW, et al. Argentine population genetic structure: large variance in Amerindian contribution. *Am J Phys Anthropol.* 2007; 132: 455–462. <https://doi.org/10.1002/ajpa.20534> PMID: 17177183.
21. Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, et al. Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet.* 2008; 4: e1000037. <https://doi.org/10.1371/journal.pgen.1000037> PMID: 18369456.
22. Muzzio M, Motti JMB, Paz Sepulveda PB, Yee MC, Cooke T, Santos MR, et al. Population structure in Argentina. *PLoS ONE.* 2018; 13: e0196325. <https://doi.org/10.1371/journal.pone.0196325> PMID: 29715266.
23. Pereira R, Phillips C, Pinto N, Santos C, Dos Santos SE, Amorim A, et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. *PLoS ONE.* 2012; 7: e29684. <https://doi.org/10.1371/journal.pone.0029684> PMID: 22272242.
24. Santos C, Fondevila M, Ballard D, Banemann R, Bento AM, Børsting C et al. Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise. *Forensic Sci Int Genet.* 2013; 19: 56–67. <https://doi.org/10.1016/j.fsigen.2015.06.004> PMID: 26122263.
25. Saloum de Neves Manta F, Pereira R, Vianna R, Rodolfo Beuttenmüller de Araújo A, Leite Góes Gitaí D, et al. Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. *PLoS ONE.* 2013; 8: e75145. <https://doi.org/10.1371/journal.pone.0075145> PMID: 24073242.
26. Vullo C, Gomes V, Romanini C, Oliveira AM, Rocabado O, Aquino J, et al. Association between Y haplogroups and autosomal AIMs reveals intra-population substructure in Bolivian populations. *Int J Legal Med.* 2014; 129: 673–680. <https://doi.org/10.1007/s00414-014-1025-x> PMID: 24878616.
27. Ossa H, Aquino J, Sierra S, Ramóñez A, Carvalho EF, Gusmão L. Analysis of admixture in Native American populations from Colombia. *Forensic Sci Int Genet, Supplement Series.* 2015; 5: e332–e333. <https://doi.org/10.1016/j.fsigs.2015.09.132>.
28. Ossa H, Aquino J, Pereira R, Ibarra A, Ossa RH, Pérez LA, et al. Outlining the Ancestry Landscape of Colombian Admixed Populations. *PLoS ONE.* 2016; 11: e0164414. <https://doi.org/10.1371/journal.pone.0164414> PMID: 27736937.
29. Excoffier L and Lischer HE. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010; 10: 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x> PMID: 21565059.
30. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Softw.* 2008; 25: 1–18.
31. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics.* 2000; 155: 945–959. PMID: 10835412.
32. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. Available from: <https://www.r-project.org>.
33. Kruskal WH and Wallis WA. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association.* 1952; 47: 583–621.
34. Dunn O J. Multiple comparisons among means. *Journal of the American Statistical Association.* 1961; 56: 52–64.
35. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series,* 1900; 5: 157–175.
36. Matheron G. The theory of regionalized variables and its applications. Paris: Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau. Ecole des Mines de Paris; 1971.
37. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze;* 1936. 8: 3–62.
38. Bandieri S. Cruzando la cordillera: La frontera argentino-chilena como espacio social. In: Torres SB, editor. *La migración chilena en la Patagonia Austral en la primera mitad del siglo XX y su inserción en los centros urbanos de Comodoro Rivadavia, Rio Gallegos y Ushuaia.* Neuquén, Argentina: Centro de Estudios de Historia Regional (CEHIR). Facultad de Humanidades. Universidad del Comahue; 2001. pp. 421–458.
39. Eyheramendy S, Martinez FM, Vial C, Repetto GM. Genetic structure characterization of Chileans reflects historical immigration patterns. *Nat Commun.* 2015; 6: 6472. <https://doi.org/10.1038/ncomms7472> PMID: 25778948.

40. Vieira-Machado CD, Tostes M, Alves G, Nazer J, Martinez L, Wettig E, et al. Uniparental ancestry markers in Chilean populations. *Genet Mol Biol*. 2016; 39:573–579. <https://doi.org/10.1590/1678-4685-GMB-2015-0273> PMID: 27561109.
41. Finkelstein D, Novella MM. Actividades económicas y proceso de construcción social en las áreas andinas de Río Negro y Chubut. In: Bandieri S, Blanco G and Varela G, editors. *Hecho en Patagonia. La historia en perspectiva regional*. Neuquén: Ed. Educo—Universidad Nacional del Comahue; 2006. pp. 191–209.
42. Di Fabio Rocca F, Albeza MV, Postillone MB, Acreche N, Lafage L, Parolin ML, et al. Historia poblacional y análisis antropogenético de la ciudad de Salta. *ANDES (Salta)*. 2016; 27: 1–22.