Ancient and modern mitogenomes from Central Argentina: new insights into population continuity, temporal depth and migration in South America

Angelina García,[1,2,†] Rodrigo Nores,[1,2, †,*] Josefina M. B. Motti,[3] Maia Pauro,[1,2] Pierre Luisi,[1] Claudio M. Bravi,[4] Mariana Fabra,[1,2] Anna L. Gosling,[5] Olga Kardailsky,[5] James Boocock,[6,7] Neus Solé-Morata,[8] Elizabeth A. Matisoo-Smith,[5,‡] David Comas,[8,‡] and Darío A. Demarchi[1,2,‡]

[1]Universidad Nacional de Córdoba, Facultad de Filosofía y Humanidades, Departamento de Antropología, Córdoba, Argentina.

[2]CONICET, Instituto de Antropología de Córdoba, IDACOR, Universidad Nacional de Córdoba, Córdoba, Argentina.

[3]Universidad Nacional del Centro de la Provincia de Buenos Aires, FACSO (NEIPHPA), CONICET, Quequén, Argentina.

[4]Instituto Multidisciplinario de Biología Celular, IMBICE, (CCT La Plata CONICET/CICPBA/Universidad Nacional de La Plata), La Plata, Argentina.

[5]Department of Anatomy, University of Otago, Dunedin, New Zealand.

[6]Department of Biochemistry, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand.

[7]Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA.

[8]Institut de Biologia Evolutiva (CSIC-UPF), Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona, Spain.

[†]These authors contributed equally to this work

[‡]These authors contributed equally to this work

*Correspondence to Rodrigo Nores: Museo de Antropología, Av. Hipólito Yrigoyen 174, 5000, Córdoba, Argentina. Tel/FAX: +543514331058. Email: rodrigonores@ffyh.unc.edu.ar

ABSTRACT

The inverted triangle shape of South America places Argentina territory as a geographical crossroads between the two principal peopling streams that followed either the Pacific or the Atlantic coasts, which could have then merged in Central Argentina. Although the genetic diversity from this region is therefore crucial to decipher past population movements in South America, its characterization has been overlooked so far. We report 92 modern and 22 ancient mitogenomes spanning a temporal range of 5,000 years, which were compared to a large set of previously reported data. Leveraging this dataset representative of the mitochondrial diversity of the subcontinent, we investigate the maternal history of Central Argentina populations within a wider geographical context. We describe a large number of novel clades within the mitochondrial DNA tree, thus providing new phylogenetic interpretations for South America. We also identify several local clades of great temporal depth with continuity until present that stem directly from the founder haplotypes, suggesting that they originated in the region and expanded from there. Moreover, the presence of lineages characteristic of other South American regions reveals the existence of gene flow to Central Argentina. Finally, we report some lineages with discontinuous distribution across the Americas, which suggest the persistence of relic lineages likely linked to the first population arrivals. The present study represents to date the most exhaustive attempt to elaborate a Native American genetic map from modern and ancient complete mitochondrial genomes in Argentina and provides relevant information about the general process of settlement in South America.

INTRODUCTION

The peopling of the Americas has been extensively studied and discussed from a range of different disciplines including linguistics, history, archeology, and population genetics, among others. For more than three decades, mitochondrial DNA (mtDNA) has been widely used as a powerful tool to address this issue (1–4). As a result, there is a growing body of evidence that describes the existing variability within South America of the Native American specific haplogroups A2, B2, C1, D1, and D4h3a (5–10). While studies of mtDNA control region are still the most prevalent in the literature, the development of high-throughput sequencing technologies allows the sequencing of complete mtDNA genomes in ancient and modern samples. The identification of locally differentiated sub-haplogroups from whole mitochondrial genome sequences increases the possibility to interpret population differentiation within the continent and to develop demographic models at a regional scale.

It has been suggested that South America was initially populated following two main routes along the Pacific and Atlantic coastlines [(11) and bibliography therein)]. On the one hand, the earliest archaeological sites are located near the Pacific coast and date to the Late Pleistocene-Early Holocene transition, as documented by sites in Chile, Peru and Ecuador (11–16). There is also archaeological evidence of an early occupation of the Andean highlands (17) that could be interpreted as a signal of an additional inland Andean route. However, the archaeological record and aDNA studies show an interconnection between highlands and coastal groups (17,18), suggesting that the western Andean occupation occurred through the Pacific route. On the other side, the earliest South American sites near the Atlantic are located in the middle courses of the rivers, indicating the entry into the

4

inland following the river shores (19) through least cost routes (20,21). The rise of sea level after the last glaciation displaced the Atlantic coastline tens or hundreds of kilometers inland, raising the possibility that earlier sites on the east coast are today under water. Evidence of this early human occupation in the Southern Cone of South America is registered in the Argentinean Pampas region, in the site Arroyo Seco 2, which is dated to 12,170 ^{14}C ybp (14,060 cal ybp) (22). The inverted triangle shape of the South American continent places Argentina at the meeting point between these two principal peopling streams that followed the Atlantic and the Pacific routes.

This research addresses, for the first time, the analysis of modern and ancient mitogenomes from Central Argentina (hereafter CA), defined in this paper as the territory encompassing Córdoba province, the Sierras in San Luis province, and the Southwest of Santiago del Estero province (Figure 1). Given its geographical position as a crossroads, CA has a crucial importance for understanding the demographic processes that took place during the peopling of the Southern Cone of South America. It may further have served as a major route of dispersal to surrounding geographic areas and, consequently, maintained gene flow with those populations, as suggested by archeological (23), morphological (21,24,25), and molecular evidence (6,26,27).

Archaeological research has revealed that the CA region has been inhabited since the Pleistocene-Holocene transition, at least 11,000 ybp (28,29). These records are characterized by the presence of fishtail projectile points, a lithic design that is widespread in early occupations of the Eastern and Western slopes of South America (30–32). After a period of scarce archaeological evidence, the signal of human occupation reappears by 8,000 ybp (33), characterized by sites with a technology of lanceolate or foliate projectile

points, known as "Ayampitín" that has been identified as a signal of Andean influence

(30,34–36). During the Middle Holocene (8,000 - 5,500 ybp) a series of changes appear in

the archaeological record: reduction in the size of lithic projectile points and acquisition of

a triangular shape (37,38), decrease in mobility, intensification of production, and

diversification of residential site types in different regions, suggesting an increase in

population (35,39). By 3,000 ybp these changes were adopted throughout the region, albeit

with noticeable regional variations (40). For this period, foreign rocks have been identified

as prime material for manufacturing tools (35), and this, combined with the appearance of

clam shells in burials (41), indicates the development of long-distance interaction networks.

Between 1,500 and 1,000 ybp, a process of diversification in the way-of-life arose among

CA communities, in terms of exploitation of resources, occupation of environments, and

population growth. Moreover, biological changes were observed from the analysis of

craniofacial morphology (42), secular variation in height (43), and mitochondrial

haplogroups of ancient populations (44). Between 1,200 and 1,000 ybp the practice of

small-scale cultivation was established (45) as a complement to foraging. During the last

centuries before the Spanish conquest, there were signals of stress due to population

growth, resource scarcity, or environmental pressure; all of which continued later, with the

arrival of the Spaniards (46–48). In the colonial period, many populations of CA settled

along the so-called "Camino Real", which was the main route to "Alto Perú" (present-day

Andean Bolivia). This major passage for people and goods prompted important socio-

cultural and demographic exchanges. The CA populations along the "Camino Real"

occupied a strategic position for trade, both legal and illegal, between Buenos Aires, Chile

and Alto Peru (49).

The evolutionary history of present-day human populations of CA has been investigated using non-recombinant uniparental markers (mtDNA and Y-chromosome) (26,27,50) and autosomal markers (51). These studies indicate that, after five centuries of colonization and cultural exchange, the modern populations of CA maintain most of their maternal Native American gene pool (~80%). Genetic distances based on mtDNA control region sequences assigned to the Native American haplogroups show that there is no significant genetic differentiation among CA populations, probably due to their common origin and sustained gene flow (27).

Focusing on a larger geographical scale, haplogroup-level analyses suggested that present-day populations of CA have genetic affinities with populations from Patagonia (26), while studies of mtDNA control region sequences demonstrated similarities with the populations from West-Central Argentina (27,52,53). By contrast, morphological studies suggest biological relationships between CA and other populations from surrounding regions, which differ according to the variables and/or configurations of the samples used (54,55). Furthermore, according to the traditional model for the peopling of CA, supported by archeological evidence, the region should be closely linked to the central Andean region (56–59). In addition to this model, there are different and opposing hypotheses regarding the origin of the CA populations: West-Central Argentina (60,61), Patagonia (24,62), and North-East Argentina (20,21) have been suggested as source areas. Nonetheless, the contrasting hypotheses concerning the origin and genetic affinities of CA populations have never been explored with high-resolution genetic markers such as whole mitochondrial genome sequences.

Several questions remain open about demographic processes in CA within a more general context: from where did human populations originate; to what extent did this region maintain gene flow with populations from surrounding regions; and whether it has served as a major dispersal route to other territories.

The aim of this study is to reconstruct the maternal history of the populations from CA through the description of the spatial and temporal patterns of molecular diversity within this region. Additionally, we provide relevant information about the general process of settlement of the Southern Cone of South America, and describe new clades within the mitochondrial phylogenetic tree.

RESULTS

Haplogroup Characterization

To survey the mitogenome variation in CA, a total of 114 modern and ancient complete mitochondrial genomes (Figure 1) were successfully sequenced with ~100% mitogenome coverage. The 92 modern and 22 archaeological mitogenomes were sequenced to an average coverage depth of 1027x (120x-2615x) and 80.1x (14.3x-189x), respectively (Table S1). Damage profiles and estimates of mitochondrial contamination for the 22 archaeological samples indicate authentic endogenous ancient DNA (Table S1b). We found no significant differences in the frequency distribution of haplogroups between modern and archaeological samples ($p = 0.229 \pm 0.005$) according to the Exact Test (63). From the total dataset, 19 individuals were assigned to haplogroup A2 (17%; 3 ancient and 16 modern), 25

to B2 (22%; 8 and 17), 25 to C1b (35%; 7 and 18), 5 to C1c (35%; 0 and 5), 10 to C1d (35%; 1 and 9), and 30 to D1 (26%, 3 and 27). The major haplogroups of Native American origin found in the modern mitogenomes show similar frequencies to a larger present-day sample set from CA (27), indicating that the sample analyzed here is largely representative (Table S2).

Population Diversity Statistics

Similar genetic diversity levels (Hd and π) were detected in contemporary and ancient samples, although the former was slightly higher (Table 1). This result was consistent with the high diversity levels found with mtDNA complete control region (CR) sequences in a separate study (27).

Neutrality tests (Tajima's D and Fu's F(s)) did not reach significance for the ancient samples, whereas they did for modern samples. Significant negative values were observed for both tests with modern sequences, indicating population growth, although the Tajima's D was only marginally significant (Table 1). This is likely because Fu's F(s) test is more sensitive to recent population expansions than Tajima's D (64). To further explore this issue, we tested the hypothesis of neutrality in mtDNA variation from the modern set of samples by estimating both parameters, but separating the samples by haplogroup and partitioning the mitogenomes into CR, coding region and whole mitogenome components. From this analysis, we observed statistically significant values for most cases, mainly for the Fu's F(s), further supporting population expansion. The exception was haplogroup D1, for which Tajima's D test showed non-significant statistical values for the three partitions (Table S3).

We observed no significant genetic differentiation between ancient and contemporary population samples: AMOVA's $F_{ST} = 0.018$ (p = 0.112) and Weir and Cockerham's $F_{ST} = 0.020$ (p = 0.114). Since the ancient sample includes individuals from different dates and locations, hence cannot be considered to be a real population, this result suggests that population continuity in the region should be considered as only prospective.

Phylogenetic and phylogeographic diversity in Central Argentina

We compiled ~3,700 sequences, including all available mitogenomes of Native American origin, both modern and ancient, to describe phylogenetic relationships between mitogenomes from CA and those from the rest of the Americas. Through the reconstruction of maximum-parsimony trees for each Pan-American haplogroup (A2, B2, C1b, C1c, C1d, and D1; Figures S1-S5, and S8), we identified a large number (N=85) of novel clades and subclades, at different levels of branching. We also modified the definition of some clades (B2z, B2am, D1g5, and C1d1e), correcting the diagnostic motif previously established according to PhyloTree Build 17 (65) and other publications (see Supplementary Section S1 and Figures S2-S5 and S8).

Each clade was clasified as belonging to one of three categories—local, recent, and relic — according to the phylogenetic and phylogeographic relationships identified within CA and between other regions of the Americas. These are described in the three following subsections.

To facilitate the interpretation of the results, we define the geographic regions of South America to which we refer throughout this study. CA constitutes a transitional region between the eastern valleys of the Andean chains, here called West-Central Argentina

10

(WCA), the Pampas-Patagonian lowlands in the east and south, and the plains of the Gran Chaco toward the north and northeast. The Salado river in Santiago del Estero province serves as the border between CA and the Gran Chaco, while there are no natural boundaries between CA and the rest of the surrounding areas. The Paraná and Paraguay rivers constitute the eastern geographical limit of the Gran Chaco with the subtropical forests of the Southern Cone, including the North-East Argentina (NEA) region, while the 500 m.a.s.l. contour line to the west separates the Gran Chaco from the Andean region of North-West Argentina (NWA).

Local Clades

Local clades refer to those clusters that exhibit local differentiation, and therefore encompass mainly CA samples. Phylogeographic and/or phylogenetic analyses revealed that some clades (A2bb, A2bc, B2ak, C1b16, C1d1b, D1g5a, and D1j) had links with other regions of the Southern Cone, while for other local clades (A2ax, A2ay, A2ba, $C1b_{+207}$, C1b17, C1c10, C1d1e, and D1v) no such link could be identified. In what follows, we summarize the main results concerning these clades (additional details are provided in Supplementary Section S2 and Figures S6-S7). Phylogeographic analyses for local clades show that $A2_{+150}$ (A2bb and A2bc), $B2_{+16142}$ (B2ak1), and $C1b_{+146}$ (C1b16) are specific to CA and WCA, although they are occasionally found in the Central Andes, Gran Chaco, Patagonia, Bolivia, and Chile (Figure 2). The remarkable correspondence between higher frequency and greater diversity is consistent with the hypothesis that these clades trace their origins to CA and WCA and subsequently spread from there.

The estimated times to the most recent common ancestor for the major haplogroups ranged from ~17,000 (A2) to ~12,000 (C1c) ybp, which are similar to those obtained in previous studies (66,67). Within them, the coalescence ages estimated for the local CA lineages varied enormously, with values ranging from ~11,000 ybp (D1j'w) to ~1,200 ybp (D1v) (Table S4; Figure 3). The oldest dates are consistent with the archaeological evidence that dates the peopling of CA back to at least 11,000 ybp [(32,36), and literature cited within)].

Recent Clades

Recent clades refer to those clades that have low frequency in CA and are more frequent and diverse in other regions of South America. The presence of these clades in the region is most likely explained by recent migration processes because (i) none of the ancient CA mitogenomes fall into them, and (ii) modern CA haplotypes belonging to them present few or no mutational steps from those described in other regions. In this study, we defined recent migrations as those demographic movements that have occurred since the 15th century, although there is no clear temporal break (68).

It is widely documented from the archaeological and historical records that the Inca empire expansion first, and the European conquest later, caused demographic changes that affected the CA populations. The displacement of entire communities was a state policy of the Inca empire, employed to reaffirm their domination over newly annexed territories (69). Once the territory was occupied by the Spaniards, they established "Indian villages", where Native people from different regions were grouped together (70). In addition to these coordinated forced relocations, many Indigenous groups voluntarily moved away from their

original territories, avoiding Inca or European domination, and occupied unconquered neighboring regions such as the Gran Chaco lowlands (71). Recent migration processes also include all those movements of individuals or families that occurred in the 20th century, mainly to Córdoba province, which is the second most populated province in Argentina and has many industries that attract workers from other regions (72).

In what follows, we summarize the main results concerning this category. The geographic distribution and genetic diversity observed for recent clades allow us to infer the sources of migrations into CA. First, the presence of A2bd, B2aj1a, B2z, D1f4, and D1u in CA is most likely explained by the recent migration from the Central Andes.

In addition, we identified a new clade called B2b2b, which has a wide distribution as it occurs at low frequency in modern populations from Amazonian Bolivia, NWA, Gran Chaco, Chile, and CA, while it accounted for 7 out of 19 individuals from the archaeological site of Pampa Grande, NWA (73). Since this clade occurred at very low frequency in relation to its wide geographical dispersion, it is very difficult to make a firm judgement about its specific origin. Within it, we define the sub-branch B2b2b1 that was found only in CA and Gran Chaco mitogenomes, probably representing a rare and highly derived local lineage of B2b2b that likely spread along the Pacific route. Furthermore, clades A2az, B2b15, C1b18, and C1b19, which are marginally represented in CA populations, also indicate a recent link between CA, Gran Chaco and NEA. In a wider geographical context, the clades B2e1 and the newly defined C1b15 and B2am show a genetic link between CA and Amazonia, Paraguay, and NEA (Supplementary Section S3).

Finally, it is not straightforward to interpret the presence of clades A2p and B2j in CA (Figures S1-S2), as their known continuous distribution is restricted to Mexico, Central America, the Caribbean and Venezuela. Given their absence in the intermediate populations of Ecuador, Peru and Brazil, we propose that the CA sequences represent southern isolated occurrences that may be related to late migrations in colonial times associated with movements of soldiers, merchants, priests, Crown officials, etc. with their families, servants and/or slaves. In Supplementary Section S3 we detail other lineages included in the Recent Clades category.

Relic Clades

CA relic clades are those that nest in internal branches of previously described clades that span broad and diverse geographic regions, but are separated by several mutational steps from them. A possible explanation for their presence in CA is that they are local survivors of clades that used to have wider geographic distributions. We identified two clades, C1b3a and C1c9a, that appeared at low frequency in CA but differed from the recent clades with discontinuous distribution by having well-differentiated haplotypes.

C1b3 has a disjointed geographical distribution in the United States (US) and South America (Figure S3). The three US sequences belong to contributors of Mexican or Hispanic origin (74–76), while the South American sequences come from admixed individuals from Argentina and Peru (77,78). A phylogeographical analysis of $C1b_{+201-16093-16192}$ (C1b3a) (Figure S9) shows a distribution in Argentina centered in CA, WCA and

NWA, with sporadic cases in Gran Chaco and NEA, Northern Chile (79), and Lima, Peru (80).

The new clade C1c9 is differentiated into three sub-branches and has a broad geographical distribution in the American continent (Figure S4). The first of the three sub-branches, C1c9a, is divided into two branches, one exclusive to CA and the other to the Gran Chaco, that represent the southernmost distribution of C1c9 lineages. The second clade, C1c9b, is restricted to southern North America (75,81), while the third one is represented by a unique well-differentiated haplotype from Peru (77). The haplotype network for the southern branch (C1c$_{+16169}$) does not exhibit any nodal haplotype but, instead, many unique haplotypes (Figure S10A). The isofrequency map based on the C1c9a diagnostic polymorphism 16169 shows a very low frequency for this clade, which is present in populations of the Gran Chaco, NWA and, to a lesser extent, in CA (Figure S10B).

Single Branches: Allochthonous Lineages or Extinction?

We define single branches as those represented by one or two individuals that do not belong to any internal clade already defined in the phylogenetic trees. We observed a relatively low number of such single haplotypes that were not related to any others. This was the case for 7% of the total sample, including 5 modern and 3 ancient individuals (Figures S2-S5). Focusing on the SNPs in the CR, we searched for matches between CA and other South American populations and found none or only distantly related sequences (detailed information can be found in Supplementary Section S4), so the presence of these single haplotypes in modern CA individuals is not consistent with a recent migration

scenario. Therefore, we suggest that they are representatives of clades that underwent

extinction (in the case of ancient samples), or that are in the way of being lost by genetic

drift (in the case of modern samples), possibly affected by the reduction of the Indigenous

population due to epidemics, warfare and slavery related to European colonization (82).

DISCUSSION

To date, most studies of complete mitochondrial genome sequences in the Americas have

focused on specific haplogroups (7,10,67,83,84), and on understanding the human

evolutionary history at broad geographical and temporal scales (10,66,67,85). Only a few

publications have documented regional patterns of variation in order to address questions

linked to a shorter time scale or interactions among closer populations (86–88).

In the present work, we focused on the sampling of a restricted geographical region, Central

Argentina, but with a relatively large time transect. With this sampling strategy, we could

define new clades through phylogenetic analyses of whole mtDNA genomes from ancient

and modern populations. We were also able to revise phylogenetic interpretations of

previously characterized clades (see Supplementary Section S1 and Figures S2-S5 and S8).

Based on our results, we observed within each of the Pan-American haplogroups a wide

range of differentiation into derived branches, including 85 newly defined sub-haplogroups,

of which 15 are specific to CA. From the definition of these new clades and the analysis of

mitochondrial diversity and geographic distribution, we formulated alternative hypotheses

to explain the main evolutionary processes that led to the formation of the populations in CA.

Local clades as evidence of great temporal depth and population continuity in Central Argentina

Most of the local clades are defined here for the first time. They represent 64% of all the CA samples analyzed, and for most, the distribution is centered in CA and/or WCA, supporting the hypothesis of a close relationship between both regions (27,89). For several local clades, the coalescence ages were estimated to date between approximately 10,800 and 8,500 ybp, revealing a great temporal depth (Figure 3 and Table S4). These dates are consistent with the archaeological evidence that indicates human presence in the region since the Pleistocene-Holocene transition (32,36).

Although some ancient samples analyzed here show private mutations that are absent in present-day populations, most of them (19/22) belong to clades that also include modern CA samples (Figures S1-S5). Moreover, most of those ancient haplotypes (17/22) nest inside branches of local clades having great temporal depth, suggesting long-term population continuity in the region. Based on radiocarbon dating, it is possible to establish a genetic link between modern people and individuals that lived at least 5,000 ybp, (the mean calibrated age of the oldest sample included in the clade B2ak1; sample #199, Table S1b). Further support for the continuity hypothesis includes both the frequency distribution of haplogroups and AMOVA's F_{ST} results that show no significant differentiation between modern and archaeological samples. To discard a potential bias due to sample size

difference between both groups, we also applied Weir and Cockerham's F_{ST} index, and we report non-significant *P*-values estimated through random permutations.

To further investigate population continuity in CA, we constructed two mitogenomic trees, one of them using only the newly generated lineages (Figure 5) and the other including these data together with an extensive comparative dataset (Figure S11). The trees show that shared branches between ancient and modern haplotypes bifurcate along the stem of Native American clades, at times spanning over a wide range. Our results contrast markedly with those obtained by Llamas et al. (66), who sequenced 92 whole mitochondrial genomes from pre-Columbian South American inhabitants, mostly from Peruvian archaeological sites. The phylogenetic analysis using their novel lineages and a comparative dataset of 206 modern Native American mitogenomes from both Indigenous and admixed populations showed that none of the ancient haplotypes shared a common ancestor with a modern haplotype more recently than ~9,000 ybp, suggesting that ancient sequences represent unique branches from the stem of Native American haplogroups. Conversely, we find that the bifurcations between ancient and modern samples in the mitogenomic trees span a large time range between 6,500 to 500 ybp, much more recently than 9,000 ybp (Figures 5 and S11). The apparent contradiction between our results and those of Llamas et al. (2016) is likely attributable to different regional dynamics and/or the more comprehensive database used here, which provides a more representative spatial overlap between archaeological sites and present-day population locations. Analyzing the genetic diversity at a fine regional scale, leveraging modern and ancient samples that cover the same restricted geographical area proves to be essential to discriminate between extinction (irrecoverable loss of genetic

diversity at some time) and population continuity (i.e. stable genetic diversity across time) scenarios (90).

Haplogroup B2: extinction or recent gene flow?

Although haplogroup B2 represents only 12% of the Native American lineages in modern CA populations (27), it is much more frequent in archaeological samples, reaching 42% in the mountain region of Córdoba province (44). The high diversity but low frequency observed in modern populations for B2 has been previously attributed to the extinction of ancient lineages (91). Under this evolutionary scenario, intermediate haplotypes would represent genetic "gaps" that accentuate the difference between lineages in the phylogenetic tree (92). However, the phylogenetic analysis based on mitogenomes shows that 52% (13/25) of modern B2 samples nest in clades that are characteristic of other regions. Therefore, the high diversity found for the B2 haplogroup in modern CA populations is better explained by recent gene flow caused either by the expansion of the Inca empire, the impact of European conquest and/or, by workforce arrivals from surrounding areas during the 20th century.

Nonetheless, we also describe a new clade within B2, called B2ak, with important temporal depth (8,945 ybp). This local clade exhibits great internal genetic diversity, and although it is highly represented among ancient samples (ranging between ~550 to ~5,000 ybp), it is still found in present-day populations. Together, these patterns reveal that different demographic processes are involved in the evolutionary history of the B2 haplogroup. This

example demonstrates the importance of considering internal branches in mtDNA studies, beyond studies at the haplogroup level.

Although the evidence does not support the hypothesis of massive extinction of lineages due to the impact of European colonization, it is indisputable that the European conquest led to a series of transformations that affected the entire "Indigenous world", by which most of these societies were disrupted and eventually "disappeared" (93). However, other groups survived and persisted, some of them reduced in "Indian villages" within the system of Hispanic domination (46), and many other individuals, especially women, were incorporated into the colonial society.

Genetic relationships between neighboring regions

We performed phylogenetic and phylogeographic analyses with modern and ancient mitogenomes, including an extensive database of mitogenomes and mtDNA control region sequences from the Americas, to analyze genetic links that reflect the interactions of CA populations with other regions in South America.

Relationships with West-Central Argentina: Despite the lack of published mitogenomes from WCA, the phylogeographic analysis based on control region sequences shows that most of the CA local clades are also highly represented in modern samples from WCA, indicating a possible common origin and/or sustained gene flow between those regions, as supported by morphological (42,55), linguistic (94), and ethno-historical (60,61) evidence.

Further analysis of WCA mitogenomes will be needed to understand the nature of the links between both regions.

Relationships with Central Andes and North-West Argentina: In total, 7% of the CA samples belong to Central Andes specific clades. These are generally haplotypes separated by just a few mutational steps from those published for samples from Peru, Bolivia, Chile, and NWA, indicating a recent link with these regions. The presence of Andean haplotypes probably reflects gene flow associated with the Inca expansion or post-conquest processes. It is worth mentioning that B2aj, a characteristic haplotype of the Andean Plateau populations (95) that spread successfully into NWA from Bolivia (10), was only found in a single modern individual from Córdoba province. This finding, together with the fact that most of ancient B2 samples nest into local clades, allows us to review the proposal of Nores et al. (44) that the high frequency of B2 found in past populations after 1,200 ybp from Córdoba province was due to significant contacts with populations from the Andean region.

Relationships with Gran Chaco and North-East Argentina: As mentioned above, some clades marginally represented among the CA samples are also present in the Gran Chaco and NEA. In all cases, they share similar haplotypes, indicating a recent contact between these regions, with gene flow probably in both directions. For example, the presence of the lineage $C1d1b_{+16259-16271-16311}$ in archaeological samples from Santiago del Estero (96) and Córdoba (sample #122, dated to ~2,300 ybp) provinces as well as in modern inhabitants of Gran Chaco suggests gene flow between both regions (96). The archaeological evidence suggests movements of Gran Chaco groups into the south of Santiago del Estero province by at least 1,000 ybp (97,98). For the Mar Chiquita Lagoon (Northeast of the Córdoba province), contact between these regions since at least 2,000 ybp is supported by

ethnohistorical (99), bioanthropological (100), and archaeological evidence

(41,45,101,102). Ethnohistorical evidence further supports the relationship with Gran

Chaco as they describe post-contact south to north movements of Indigenous people who

found in the Gran Chaco a refuge from Spanish conquerors (71), and as a result of 19th

century movements to occupy new territories for breeding (103).

Relationships with the Pampas: As only two mitogenomes have been published for the

Pampean region of Argentina, it is therefore difficult to decipher genetic links of CA with

this region. On the one hand, these two sequences, reported from the Arroyo Seco 2 site

and dated to more than 7,000 ybp (66), are not related to CA sequences. On the other hand,

HVRI sequences from an archaeological site in the Pampas-Patagonia transition revealed

the presence of one individual (out of 13) carrying a D1j haplotype, thus suggesting a

possible relation with CA (104). A recent study of Native American ancestry with a large

set of autosomal markers assigned admixed individuals from both Central and Pampas

regions to the same cluster, which also includes individuals belonging to Native American

communities from the Gran Chaco, the Amazonia, and Northern Andes (105). Thus, more

studies of ancient samples from the Pampas will be needed to elucidate the nature and

extent of the genetic links and better understand the biological and cultural relatedness

between this region and CA, suggested by craniometric analysis (106) and shared burial

traditions (107).

Relationships with Patagonia: In this research, we have found no evidence linking Central

Argentina to Patagonia. While a single CA sample was grouped with others from northern

Patagonia, defining the sub-haplogroup C1d1f, this sample represents an example of the

migratory process to CA during the last decades, according to genealogical information.

Another clade that deserves discussion is D1g5, a lineage found in two ancient and three modern samples of Central Argentina. Despite D1g is a clearly Patagonian clade (8), D1g5 is the most widespread subclade within D1g, ranging from Tierra del Fuego to the extreme north of Argentina (108). Moreover, only D1g5a, which accounts for four out of the five D1g5 CA samples, is present at lower latitudes, while the rest of the subclades within D1g and D1g5 are found almost exclusively in the central-southern part of Chile and Argentina (8,108). Thus, the presence of D1g5a in CA is unlikely to reflect a link with Patagonia. In addition, clades C1b13, B2i2, and D4h3a5, which are highly represented in Patagonia (7,8), have not been found among the CA samples.

Overall, our results do not support previous claims pointing to a possible link between the CA and the Patagonian region based on archaeological (35) and morphological evidence (21,24,109), as well as mtDNA haplogroup distribution (26,44). The apparent genetic similarity is due to the high frequency of C and D haplogroups in both regions; however, after increasing the resolution with complete mitochondrial sequences, we show that these regions do not share the same sub-haplogroups.

Demographic model for Central Argentina: a genetic variation mosaic

The results obtained in the present work reveal a mosaic of genetic diversity in CA with several local clades, most of them having great temporal depth, and demonstrate that a dense sampling scheme of a particular region allows to reach sufficient resolution of the genetic diversity to better understand population dynamics. In this sense, it is possible to identify genetic signals of population movements, either leading to gene flow among

populations that occurred in the last centuries at different intensities and that have changed the mitochondrial variability found in modern CA populations, or to partial population replacements that occurred in more remote times.

Although it is not possible to know exactly the processes that led to the presence of foreign clades in CA, the archaeological and historical evidence allow us to interpret the presence of Andean lineages as a result of major forced population movements into CA, caused by Inca and/or European conquest. In the case of clades shared with the Gran Chaco and NEA regions, they are probably the result of individual or kin-structured migration with these regions. Also, some isolated haplotypes could have arrived in the region as a consequence of individual or family migrations in historic or modern times. Further investigation would be needed to evaluate alternative hypotheses.

Based on our findings, we propose a demographic model for the population of CA within the general process of peopling of the Southern Cone of South America (Figure 6). According to this demographic model, CA is considered as the center of origin and dispersal of many lineages, some of which were probably lost through genetic drift or extinction, but most having survived, revealing maternal genetic continuity between ancient and modern populations.

In this work, we found that local clades from CA are not nested inside any other South American clade with a wide spatial distribution. As a consequence, it is not possible to discern whether the ancient populations from CA arrived from the Pacific or Atlantic routes, or both. Whatever their entry route, the first inhabitants reached CA during the Pleistocene-Early Holocene transition, as evidenced by the archaeological record and our

TMRCA estimates for local clades. In addition, the presence of some lineages with discontinuous distributions (C1b3 and C1c9) suggests the genetic persistence of relic groups that arrived in remote times, probably with the first founders. Following early settlements, significant and distinctive mitochondrial diversity emerged over millennia. By 8,000 ybp, the archaeological record reports a greater occurrence of residential sites in several regions of CA (39,110–112), which is consistent with population growth during that time. Furthermore, lithic typology changed around 7,500 years ago, associated with the appearance of triangular points. This new technology was established around 5,000 ybp and reached a wider territorial expansion that continued until around 2,000 ybp (38). The genetic data we present here provides further support to the archaeological evidence of population size increase. Neutrality tests (Tajima's D and Fu's F(s)) hold significant negative values for modern mitogenomes, indicating population expansion, while the Bayesian Skyline Plot (BSP; Figure 4) shows a rather sharp increase in population size between 7,500–5,000 ybp. In agreement, the TMRCA for the outermost nodes of the local clades (A2bb, A2bc, B2ak1 and B2ak1a) also suggests a population growth between 7,200–4,200 ybp (Figure 3 and Table S4). Finally, BSP analysis shows that a smooth demographic contraction occurred immediately after the population growth, and lasted up to ~2,500 ybp. This was followed by a long period of demographic stability until present (Figure 4). Finally, for recent centuries, our results point to gene flow from surrounding regions, mainly from Central Andes, and to a lesser extent from Gran Chaco, NEA and Subtropical Forests of South America.

In conclusion, given the deep temporality of the local CA clades and the high levels of molecular diversity, which do not differ between modern and ancient mitogenomes, the

central region of Argentina should no longer be seen as a mere "crossroads" where the Andean, Patagonian, Gran Chaco, and NEA components met. On the contrary, the populations of CA reveal themselves as having their own identity in genetic terms.

MATERIALS AND METHODS

Modern samples

Complete mtDNA sequences were obtained from 92 unrelated individuals randomly chosen amongst those carrying Native American lineages from a set of 812 biological samples collected in Central Argentina (27). Specifically, this modern data set includes individuals from the provinces of Córdoba (N=52), Santiago del Estero (N=16), and San Luis (N=24) (Figure 1). All participants gave written consent to be part of this nonprofit scientific investigation. Information gathered in the field regarding birthplaces of the subjects and their parents indicates that the inhabitants of the studied populations present low to moderate mobility and are restricted to their places of origin by at least two or three generations in most of the cases. Only seven of the 92 individuals declared to have a maternal ancestor who came from a different region than Central Argentina (see the Table S1a and Supplementary Section S5). All associated information concerning modern samples is provided in Table S1a. Laboratory methods for analyses of modern samples (DNA extraction, preparation of Illumina sequencing libraries, massively parallel sequencing, and raw data analysis), are detailed in Supplementary Section S6.

Archaeological Samples

DNA extraction was performed from bone or dental remains of 22 individuals of archaeological origin, preserved in the Museo de Antropología (Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba), and from different public and private museums in Córdoba province. The samples were originally obtained from different archaeological sites located in Córdoba province (Figure 1). Macroscopic analysis of archaeological remains showed a general appearance of good preservation, with very little bone loss and postmortem fragmentation (25). Radiocarbon dating of 20 of these individuals span ~4,000 years of local history, from 459 ± 40 to 4450 ± 80 uncalibrated ^{14}C ybp. All associated information concerning archaeological samples (including site location, current collection location, and specimen description) is provided in Table S1b. Laboratory methods of ancient samples (DNA extraction, preparation of Illumina sequencing libraries, mtDNA hybridization capture enrichment, massively parallel sequencing, and raw data analysis), are detailed in Supplementary Section S6.

Ethical statement

This study was performed following ethical guidelines for working with human remains, treating these deceased people with due respect. Our research program with modern populations and ancient human remains was approved by the Ethics Committee of the CEMIC (Comité de Ética en Investigación, Centro de Educación Médica e Investigaciones Clínicas "Norberto Quirno"). Skeletal samples were exported with the proper permits from Córdoba province and Argentina Governments.

Data Edition and Haplogroup Assignment

Polymorphic sites were defined by differences from the rCRS (113) using mitoSAVE. This software converts data within the VCF into mtDNA haplotypes using phylogenetically-established nomenclature for the human mitochondrial genome. Each position was only considered if there was a 10X minimum coverage. The data were verified manually by checking BAM files in IGV (Integrative Genomics Viewer) 2.3.34 (114). Once the debug and edited files were generated, .hsd files were uploaded to HaploGrep 2 (115) to assign mitochondrial haplogroups. Finally, all haplogroup assignments were manually confirmed by checking diagnostic positions as described in PhyloTree Build 17 (65).

Modern sample haplogroups obtained through NGS were verified by comparing with the previously Sanger-generated control region sequences. The data edition also implicated the exclusion for all calculations of the transitions at the hypervariable position 16519, and any length variations in the C-stretch between nucleotides 16181-16193, 303-315 and 513-524. The mitogenomes in fasta format were aligned using the R package DECIPHER (116,117), sequentially calling functions AlignSeqs and AdjustAlignment. The alignments were checked manually in HaploGrep 2 (115).

Genetic Diversity

Different measures of genetic diversity (sequence number, polymorphic sites, sequence diversity, and nucleotide diversity), analysis of molecular variance (AMOVA), pairwise genetic distances (Φst and pairwise difference), and Exact Test of population differentiation were performed using Arlequin 3.5.1.3 (118). We also computed Weir and Cockerham's

F_{ST} (119) with *pp.fst* function from the R package *hierfstat,* and its *P*-value was estimated randomly permuting sample groups 10,000 times.

Tests of neutral selection such as Tajima's D (64) and Fu's F(s) (120) were run using a bioinformatic pipeline (121). Statistical significance was estimated either from the expected distribution (for Tajima's D) or from 10,000 simulations (for Tajima's D and Fu's F(s)). Simulations were run using the three following steps: 1- Generate 10,000 gene tree replicates with ms (122), containing the same number of sequences as the tested sample and with mutation parameter ($4N_0\mu$) equal to its Watterson estimator (θ_W) (123); 2- For each gene tree replicate, generate the sequences with Seq-Gen (124), following the HKY model with the number of nucleotides for the genomic region analysed as the total sequence length l, and the value θ_W/l as the branch length scale parameter; 3- For each sequence sample thus obtained, calculate Tajima's D and Fu's F(s) scores as for the real data. *P*-values were then calculated as the number of simulations exhibiting a score lower than that observed for real data.

Phylogenetic Analysis

We built maximum-parsimony trees, one for each of the haplogroups A2, B2, C1b, C1c, C1d, and D1 using the mtPhyl software (https://sites.google.com/site/mtphyl/home), which were hand-modified with reference to PhyloTree Build 17 (65). In total, 278 mitogenomes were included in the analysis (164 published and 114 from this study). The database is detailed on Table S5. Novel clades and subclades were defined when they included a minimum of two different sequences and shared at least one stable mutation. Mitogenomes used to build the phylogenetic trees were selected from an in-house database containing

~3,700 sequences, which includes all the publicly available mitogenomes of Native

American origin, both modern and ancient. The use of a large dataset allowed exploring

phylogenetic affinities between our samples and ancient and modern mitogenomes of the

Americas. Because the capital cities of South American countries, such as Buenos Aires

and Lima, disproportionately concentrate a large part of the nation's populations due to

migrations of the last century (125), influencing inevitably the current lineages distribution

(126), they were not considered when discussing distribution/origin of lineages.

Demographic Analysis

We performed a Bayesian skyline analysis using BEAST v1.8.3 to reconstruct past

population dynamics. BEAST runs were performed only with the 113 full mtDNA genomes

from Central Argentina. For these analyses, we used GTR nucleotide substitution model,

the best substitution model estimated with MrMTgui v1.0, and applied the corresponding

substitution rates reported previously in Soares et al. (127). We assumed that every branch

in a phylogenetic tree evolves according to the same evolutionary rate and therefore we

used a strict clock. BEAST was then run with 100,000,000 iterations, sampling each

10,000[th] iteration. We thus obtained a population size estimate over time for each sampled

iteration. Finally, we reconstructed the consensus phylogenetic tree with TreeAnnotator,

using default parameters, and plotted it using TreeStat v.18.3.

Haplogroup and Sub-haplogroup Coalescence Age Estimation

Coalescence times were estimated using a Bayesian approach. In addition to the 113 novel

modern and ancient Native American mitogenomes from Central Argentina populations

(we excluded an undated ancient sample), we included a representative sample (N=824) of

modern and ancient Native American mitogenomes from South America, from the
following publications: Fagundes et al. (85) (N=47); Bodner et al. (5) (N=45); de Saint
Pierre, Gandini, et al. (7) (N=46); Llamas et al. (66) (N=92); Barbieri et al. (9) (N=116);
Brandini et al. (67) (N=223); Sans, Mones, et al. (128) (N=5); Sevini et al. (78) (N=124);
Simão et al. (88) (N=90). Some of these samples (N=8) were excluded from the analysis
due to an uncertain assignment in a phylogenetic tree performed to quickly test the
usefulness of each downloaded fasta file. We additionally included all the published
mitogenomes contained in the branches of interest for which the coalescence age was
estimated, and these also were used for phylogenetic analysis (N=44).

This set of mitogenomes were chosen so the Bayesian analyses were computationally
feasible while the dataset remained representative for both genetic diversity and
geographical range. BEAST was then run on this 937 sequences alignment with
300,000,000 iterations, sampling each 10,000th iteration, HKY evolution model, applying
the corresponding substitution rates reported previously in Soares et al. (127) and using a
strict clock. Using a burn-in of 30,000,000 iterations, we then obtained the TMRCA with
Tracer and reconstructed the consensus phylogenetic tree with TreeAnnotator with default
parameters. The tree was finally plotted with iTOL online tool (129).

Phylogeographic Analysis

For comparative purposes, we included in the analysis 7,042 mtDNA CR published
sequences from other populations of South America (Table S6a). Narrowing the mtDNA
range to the CR allows us to build a more extensive database and to analyze
phylogeographic patterns with accuracy. CR sequences between rCRS np 16032-16544 and

51-555 were used. Polymorphic sites 16182C, 16183C, and 16519 as well as indels variations at 16184-16193, 303-315, and 515-524 were not considered due to their high mutation rates (127).

Haplotype networks. For selected sub-haplogroups ($A2_{+150}$, $B2_{+16142}$, $C1b_{+146}$, $C1b_{+201-16093-16192}$, $C1c_{+16169}$, $C1d1_{+195}$, and $D1_{+152-16311}$), median joining networks were estimated with Network 4.6.1.3 (Fluxus Technology). Maximum parsimony was used as the post-processing option, using the three criteria proposed by Crandall and Templeton (130), based on predictions of the coalescence theory: frequency, topology, and geography. The database used for median joining networks is detailed in Table S6b.

Frequency maps. Based on the frequency of these selected sub-haplogroups we built isoline maps, following the Kriging procedure using Surfer 7 (Golden Software Inc., Golden, Colorado). Isolines start at a frequency of 1%, lower frequencies are not drawn on the map. The populations considered are detailed on Table S6c, and frequencies of each sub-haplogroup for each population are on Table S6d.

CONFLICT OF INTEREST STATEMENT

All authors have no relationships or activities that could appear to have influenced the submitted work.

DATA AVAILABILITY STATEMENT

The mitogenomes generated during this study are available in GenBank database (accession numbers in progress).

REFERENCES

1. Kivisild, T. (2015) Maternal ancestry and population history from whole mitochondrial genomes. *Investig. Genet.*, **6**, 3.

2. Schurr, T. G., Ballinger, S. W., Gan, Y. Y., Hodge, J. A., Merriwether, D. A., Lawrence, D. N., Knowler, W. C., Weiss, K. M. and Wallace, D. C. (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Am. J. Hum. Genet.*, **46**, 613–23.

3. Torroni, A., Schurr, T. G., Cabell, M. F., Brown, M. D., Neel, J. V., Larsen, M., Smith, D. G., Vullo, C. M. and Wallace, D. C. (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.*, **53**, 563–90.

4. Merriwether, D. A., Clark, A. G., Ballinger, S. W., Schurr, T. G., Soodyall, H., Jenkins, T., Sherry, S. T. and Wallace, D. C. (1991) The structure of human mitochondrial DNA variation. *J. Mol. Evol.*, **33**, 543–555.

5. Bodner, M., Perego, U. A., Huber, G., Fendt, L., Rock, A. W., Zimmermann, B., Olivieri, A., Gomez-Carballa, A., Lancioni, H., Angerhofer, N., et al. (2012) Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes. *Genome Res.*, **22**, 811–820.

6. García, A., Pauro, M., Nores, R., Bravi, C. M. and Demarchi, D. A. (2012) Phylogeography of mitochondrial haplogroup D1: An early spread of subhaplogroup

D1j from Central Argentina. *Am. J. Phys. Anthropol.*, **149**, 583–590.

7.   de Saint Pierre, M., Gandini, F., Perego, U. A., Bodner, M., Gómez-Carballa, A., Corach, D., Angerhofer, N., Woodward, S. R., Semino, O., Salas, A., et al. (2012) Arrival of Paleo-Indians to the Southern Cone of South America: New Clues from Mitogenomes. *PLoS One*, **7**, e51311.

8.   de Saint Pierre, M., Bravi, C. M., Motti, J. M. B., Fuku, N., Tanaka, M., Llop, E., Bonatto, S. L. and 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.

9.   Barbieri, C., Sandoval, J. R., Valqui, J., Shimelman, A., Ziemendorff, S., Schröder, R., Geppert, M., Roewer, L., Gray, R., Stoneking, M., et al. (2017) Enclaves of genetic diversity resisted Inca impacts on population history. *Sci. Rep.*, **7**, 1–12.

10.   Gómez-Carballa, A., Pardo-Seco, J., Brandini, S., Achilli, A., Perego, U. A., Coble, M. D., Diegoli, T. M., Álvarez-Iglesias, V., Martinón-Torres, F., Olivieri, A., et al. (2018) The peopling of South America and the trans-Andean gene flow of the first settlers. *Genome Res.*, **28**, 767–779.

11.   Prates, L., Politis, G. G. and Perez, S. I. (2020) Rapid radiation of humans in South America after the last glacial maximum: A radiocarbon-based study. *PLoS One*, **15**, e0236023.

12.   Dillehay, T. D., Ramirez, C., Pino, M., Collins, M. B., Rossen, J. and Pino-Navarro, J. D. (2008) Monte Verde: Seaweed, Food, Medicine, and the Peopling of South

America. *Science*, **320**, 784–786.

13.   Sandweiss, D. H. (1998) Quebrada Jaguay: Early South American Maritime

       Adaptations. *Science*, **281**, 1830–1832.

14.   Keefer, D. K. (1998) Early Maritime Economy and El Ni&ntilde;o Events at

       Quebrada Tacahuay, Peru. *Science*, **281**, 1833–1835.

15.   Stothert, K. E., Piperno, D. R. and Andres, T. C. (2003) Terminal Pleistocene/Early

       Holocene human adaptation in coastal Ecuador: the Las Vegas evidence. *Quat. Int.*,

       109–**110**, 23–43.

16.   Dillehay, T. D. (1997) *Monte Verde: A Late Pleistocene Settlement in Chile, Volume

       II: The Archaeological Context*; Smithsonian Institution, Washington DC.

17.   Rademaker, K., Hodgins, G., Moore, K., Zarrillo, S., Miller, C., Bromley, G. R. M.,

       Leach, P., Reid, D. A., Álvarez, W. Y. and Sandweiss, D. H. (2014) Paleoindian

       settlement of the high-altitude Peruvian Andes. *Science (80-.).*, **346**, 466–469.

18.   Nakatsuka, N., Lazaridis, I., Barbieri, C., Skoglund, P., Rohland, N., Mallick, S.,

       Posth, C., Harkins-Kinkaid, K., Ferry, M., Harney, É., et al. (2020) A Paleogenomic

       Reconstruction of the Deep Population History of the Andes. *Cell*, **181**, 1131-

       1145.e21.

19.   Miotti, L. L. (2006) La fachada atlántica, como puerta de ingreso alternativa de la

       colonización humana de América del Sur durante la transición

       Pleistoceno/Holoceno. *II Simposio Internacional El Hombre Temprano en América*,

       Instituto Nacional de Antropología e Historia, Mexico, pp. 155–188.

20.   Anderson, D. G. and Gillam, J. C. (2000) Paleoindian Colonization of the Americas: Implications from an Examination of Physiography, Demography, and Artifact Distribution. *Am. Antiq.*, **65**, 43–66.

21.   Fabra, M., Laguens, A. G. and Demarchi, D. A. (2007) Human colonization of the central territory of Argentina: Design matrix models and craniometric evidence. *Am. J. Phys. Anthropol.*, **133**, 1060–1066.

22.   Politis, G. G., Gutiérrez, M. A., Rafuse, D. J. and Blasi, A. (2016) The Arrival of Homo sapiens into the Southern Cone at 14,000 Years Ago. *PLoS One*, **11**, e0162870.

23.   Laguens, A. G., Demarchi, D. A. and Fabra, M. (2007) Un estudio arqueológico y bioantropológico de la colonización humana en el sector sur de las Sierras Pampeanas. In Olmedo, E., Ribero, F. (eds.), *Debates actuales en Arqueología y Etnohistoria*, Río Cuarto, pp. 119–134.

24.   Fabra, M., Laguens, A. G. and Demarchi, D. A. (2005) Análisis intra e inter poblacional de rasgos craneanos no métricos en aborígenes prehispánicos del actual territorio de Córdoba. *Rev. Argentina Antropol. Biológica*, **7**, 47–65.

25.   Fabra, M. (2014) *Historia de las poblaciones prehispánicas del sector austral de las Sierras Pampeanas: variabilidad morfológica y modelos arqueológicos.*; Secretaria de Ciencia y Tecnología. Facultad de Filosofía y Humanidades. Universidad Nacional de Córdoba, Córdoba.

26.   García, A. and Demarchi, D. A. (2009) Incidence and Distribution of Native

American mtDNA Haplogroups in Central Argentina. *Hum. Biol.*, **81**, 59–69.

27. García, A., Pauro, M., Bailliet, G., Bravi, C. M. and Demarchi, D. A. (2018) Genetic variation in populations from central Argentina based on mitochondrial and Y chromosome DNA evidence. *J. Hum. Genet.*, **63**, 493–507.

28. Rivero, D. E. (2012) La Ocupacion Humana Durante La Transicion Pleistoceno-Holoceno (11,000–9000 a.P.) En Las Sierras Centrales De Argentina. *Lat. Am. Antiq.*, **23**, 551–564.

29. Cornero, S., Neves, W. and Rivero, D. E. (2014) Nuevos aportes a la cronología de las ocupaciones tempranas en las Sierras de Córdoba. La Gruta de Candonga (Córdoba, Argentina). *Relac. - Soc. Argent. Antropol.*, **39**, 285–292.

30. Laguens, A. G., Pautassi, E., Sario, G. and Cattaneo, R. (2007) ELS1, a Fishtail Projectile-Point Site from Central Argentina. *Curr. Res. Pleistocene*, **24**, 55–57.

31. Cattáneo, R., Izeta, A. D. and Caminoa, J. M. (2016) A Fishtail Projectile Point from the Southern Pampean Hills, Characato, Córdoba, Argentina. *PaleoAmerica*, **2**, 274–276.

32. Rivero, D. E., Cornero, S., Truyol, G. and Neves, W. (2019) Human cranium of Candonga Cave site and its implications for the initial peopling of South America. *HOMO*, **70**, 283–295.

33. Rivero, D. E. (2010) La transición Pleistoceno-Holoceno (11.000 – 9.000 AP) en las Sierras de Córdoba (Rep. Argentina). *Arqueología*, **16**, 175–189.

34. González, A. R. (1960) La estratigrafía de la gruta de Intihuasi (Pcia. de San Luis, Rep. Argentina) y sus relaciones con otros sitios precerámicos de Sudamérica. *Publ. Inst. Antrop. Lingüíst. Folk.*, **I**, 5–296.

35. Laguens, A. G., Giesso, M., Bonnin, M. and Glascock, M. (2007) Interacciones a larga distancia entre cazadores-recolectores el Holoceno medio en Intihuasi, San Luis, Argentina. In Olmedo, E., Ribero, F. (eds.), *Debates actuales en Arqueología y Etnohistoria*, Río Cuarto, pp. 135–142.

36. Rivero, D. E., Pastor, S. and Heider, G. (2018) The Tigre Projectile Point in Central Argentina: Implications for the Initial Peopling of the Region. *PaleoAmerica*, **4**, 68–70.

37. Menghín, O. and González, A. R. (1954) Excavaciones arqueológicas en el yacimiento de Ongamira, Córdoba (Rep. Arg.). Nota preliminar. *Notas Mus. La Plata*, **XVII**.

38. Berberián, E., Pastor, S., Rivero, D., Medina, M., Recalde, A., Lopez, L. and Roldán, F. (2008) Ultimos avances de la investigación arqueológica en las Sierras de Córdoba. *Comechingonia. Rev. Arqueol.*, **11**, 143–174.

39. Rivero, D. E. (2007) Ecología de cazadores-recolectores del sector central de las Sierras de Córdoba (Rep. Argentina). *Ecología de cazadores-recolectores del sector central de las Sierras de Córdoba (Rep. Argentina)*; BAR Publishing, Oxford, (2007) .

40. Pastor, S. (2007) Arqueología del Valle de Salsacate y pampas de altura adyacentes

(Sierras Centrales de Argentina). Una aproximación a los procesos sociales del período prehispánico tardío (900-1573 d.C.). Universidad Nacional de La Plata.

41.   Fabra, M., Gordillo, S. and Piovano, E. L. (2012) Arqueomalacología en las costas de Ansenuza: análisis de una almeja nacarífera (Anodontites trapesialis) hallada en contexto funerario del sitio El Diquecito (Laguna Mar Chiquita, Córdoba). *Arqueologia*, **18**, 257–266.

42.   Fabra, M. and Demarchi, D. A. (2012) Morfología craneofacial y estructura genética en poblaciones del centro de Argentina. *Rev. Argentina Antropol. Biológica*, **14**, 45–56.

43.   Loupias, L. and Fabra, M. (2019) Estudio de la variación secular en la talla de poblaciones humanas del centro de Argentina durante el Holoceno tardío. *Jangwa Pana*, **18**, 396–419.

44.   Nores, R., Fabra, M. and Demarchi, D. A. (2011) Variación temporal y espacial en poblaciones prehispánicas de Córdoba. Análisis de ADN antiguo. *Rev. Mus. Antrop.*, **4**, 187–194.

45.   Tavarone, A., Colobig, M. de los M. and Fabra, M. (2019) Late Holocene plant use in lowland central Argentina: Microfossil evidence from dental calculus. *J. Archaeol. Sci. Reports*, **26**, 101895.

46.   Castro Olañeta, I. (2003) Transformaciones y continuidades de sociedades indígenas en el sistema colonial. El caso del pueblo de indios de Quilino a principios del siglo XVII. Universidad Nacional de Córdoba.

47. Laguens, A. G. and Bonnin, M. (2009) *Sociedades indígenas de las Sierras Centrales. Arqueología de Córdoba y San Luis*; Universidad Nacional de Córdoba, Córdoba.

48. González, C. and Fabra, M. (2018) Desgaste dental y hábitos dietarios en poblaciones arqueológicas del centro de Argentina. *Arqueología*, **24**, 133–159.

49. Assadourian, C., Beato, G. and Chiaramonte, J. (2005) *Historia Argentina. Tomo II. De la Conquista a la Independencia.*; 2a edición.; Paidós, Buenos Aires.

50. Pauro, M., García, A., Nores, R. and Demarchi, D. A. (2013) Analysis of Uniparental Lineages in Two Villages of Santiago del Estero, Argentina, Seat of Pueblos de Indios in Colonial Times. *Hum. Biol.*, **85**, 699–719.

51. García, A., Dermarchi, D. A., Tovo-Rodrigues, L., Pauro, M., Callegari-Jacques, S. M., Salzano, F. M. and Hutz, M. H. (2015) High interpopulation homogeneity in Central Argentina as assessed by Ancestry Informative Markers (AIMs). *Genet. Mol. Biol.*, **38**, 324–331.

52. Motti, J. M. B. (2012) Caracterización de linajes maternos en la población actual del noroeste y centro-oeste argentinos. Universidad Nacional de La Plata.

53. Pauro, M. (2015) Análisis molecular de linajes uniparentales en poblaciones humanas del centro de Argentina. Universidad Nacional de Córdoba.

54. Fabra, M. (2006) Relaciones biológicas craneales entre poblaciones prehispánicas de la Región Central del territorio argentino y otras del Cono Sur sudamericano. *An. Arqueol. Etnol.*, **61**–**62**, 123–145.

55. Fabra, M. and Demarchi, D. A. (2009) Variabilidad craneofacial en poblaciones del sector austral de las sierras pampeanas: aportes desde la morfometría geométrica. *Relac. - Soc. Argent. Antropol.*, **34**, 297–320.

56. Outes, F. (1911) Los tiempos prehistóricos y protohistóricos en la Provincia de Córdoba. *Rev. Mus. La Plata*, **XVII**, 261–374.

57. González, A. R. (1944) Algunas observaciones sobre los caracteres antropológicos de los primitivos habitantes de Córdoba. *Publ. Inst. Antrop. Lingüíst. Folk..*, **31**, 1185–90.

58. Serrano, A. (1945) *Los comechingones*; Serie Aborígenes Argentinos Vol. I, Instituto de Arqueología, Lingüística y Folklore, Universidad Nacional de Córdoba.

59. Rivero, D. E. and Berberián, E. E. (2006) El poblamiento inicial de las Sierras Centrales de Argentina. Las evidencias arqueológicas tempranas. *Cazad.- recolectores cono sur*, **1**, 127–138.

60. Canals Frau, S. (1944) El grupo huarpe-comechingón. *An. Inst. Etnogr. Am.*, **V**.

61. Canals Frau, S. (1953) *Poblaciones indígenas de la Argentina*; Editorial Sudamérica, Buenos Aires, Argentina.

62. Marcellino, A. J. and Colantonio, S. E. (1993) Relaciones morfológicas de los aborígenes prehispánicos del territorio argentino. VII: La región serrana de Córdoba. *Sep. la Rev. la Junta Prov. Hist. Córdoba*, **15**, 113–135.

63. Raymond, M. and Rousset, F. (1995) GENEPOP (Version 1.2): Population Genetics

Software for Exact Tests and Ecumenicism. *J. Hered.*, **86**, 248–249.

64. Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

65. van Oven, M. and Kayser, M. (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.*, **30**, E386–E394.

66. Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S. M., Rohrlach, A., et al. (2016) Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.*, **2**, e1501385.

67. Brandini, S., Bergamaschi, P., Fernando Cerna, M., Gandini, F., Bastaroli, F., Bertolini, E., Cereda, C., Ferretti, L., Gómez-Carballa, A., Battaglia, V., et al. (2018) The Paleo-Indian entry into South America according to mitogenomes. *Mol. Biol. Evol.*, **35**, 299–311.

68. Cabana, G. S. and Clark, J. J. (2011) *Rethinking Anthropological Perspectives on Migration*; Cabana, G. S., Clark, J. J. (eds.); University Press of Florida.

69. Lorandi, A. M. (1991) Evidencias en torno a los Mitmaqkuna incaicos en el N.O. argentino. (Seguido de un apéndice sobre tecnología cerámica a cargo de Beatriz Cremonte). *Anthropologica*, **9**, 211–243.

70. Tell, S. and Castro Olañeta, I. (2011) El registro y la historia de los pueblos de indios de Córdoba entre los siglos XVI y XIX. *Rev. Mus. Antrop.*, **4**, 235–248.

71. Farberman, J. (2016) La construcción de un espacio de frontera. Santiago del Estero, el Tucumán y el Chaco desde el Prehispánico Tardío hasta mediados del siglo XVIII. *Rev. Mus. Antrop.*, **9**, 187–198.

72. Instituto Nacional de Estadística y Censos (2012) *Censo Nacional de Población, Hogares y Viviendas 2010: censo del Bicentenario. Resultados definitivos, Serie B nᵒ 2*; (INDEC), I. N. de E. y C. (ed.); Buenos Aires.

73. Carnese, F. R., Mendisco, F., Keyser, C., Dejean, C. B., Dugoujon, J. M., Bravi, C. M., Ludes, B. and Crubézy, E. (2010) Paleogenetical study of pre-Columbian samples from Pampa Grande (Salta, Argentina). *Am. J. Phys. Anthropol.*, **141**, 452–462.

74. Just, R. S., Diegoli, T. M., Saunier, J. L., Irwin, J. A. and Parsons, T. J. (2008) Complete mitochondrial genome sequences for 265 African American and U.S. "Hispanic" individuals. *Forensic Sci. Int. Genet.*, **2**, 45–48.

75. Zheng, H.-X., Yan, S., Qin, Z.-D. and Jin, L. (2012) MtDNA analysis of global populations support that major population expansions began before Neolithic Time. *Sci. Rep.*, **2**, 745.

76. King, J. L., Larue, B. L., Novroski, N. M., Stoljarova, M., Seo, S. B., Zeng, X., Warshauer, D. H., Davis, C. P., Parson, W., Sajantila, A., et al. (2014) High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. *Forensic Sci. Int. Genet.*, **12**, 128–135.

77. Tito, R., Polo, S. and Lewis, C. (2012) Unpublished. *GenBank direct submission*.

78. Sevini, F., Vianello, D., Barbieri, C., Iaquilano, N., De Fanti, S., Luiselli, D., Franceschi, C. and Franceschi, Z. (2014) Human mitochondrial genomes reveal population structure and different phylogenies in Gran Chaco (Argentina). Unpublished. *GenBank direct submission*.

79. Gómez-Carballa, A., Moreno, F., Álvarez-Iglesias, V., Martinón-Torres, F., García-Magariños, M., Pantoja-Astudillo, J. A., Aguirre-Morales, E., Bustos, P. and Salas, A. (2016) Revealing latitudinal patterns of mitochondrial DNA diversity in Chileans. *Forensic Sci. Int. Genet.*, **20**, 81–88.

80. Messina, F., Di Corcia, T., Ragazzo, M., Sanchez Mellado, C., Contini, I., Malaspina, P., Ciminelli, B. M., Rickards, O. and Jodice, C. (2018) Signs of continental ancestry in urban populations of Peru through autosomal STR loci and mitochondrial DNA typing. *PLoS One*, **13**, e0200796.

81. Kivisild, T., Shen, P., Wall, D. P., Do, B., Sung, R., Davis, K., Passarino, G., Underhill, P. A., Scharfe, C., Torroni, A., et al. (2006) The role of selection in the evolution of human mitochondrial genomes. *Genetics*, **172**, 373–387.

82. Livi-Bacci, M. (2006) The Depopulation of Hispanic America after the Conquest. *Popul. Dev. Rev.*, **32**, 199–232.

83. Perego, U. A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., Kashani, B. H., Ritchie, K. H., Scozzari, R., Kong, Q. P., et al. (2009) Distinctive Paleo-Indian Migration Routes from Beringia Marked by Two Rare mtDNA Haplogroups. *Curr. Biol.*, **19**, 1–8.

84. Achilli, A., Perego, U. A., Lancioni, H., Olivieri, A., Gandini, F., Kashani, B. H., Battaglia, V., Grugni, V., Angerhofer, N., Rogers, M. P., et al. (2013) Reconciling migration models to the Americas with the variation of North American native mitogenomes. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 14308–14313.

85. Fagundes, N. J. R., Kanitz, R., Eckert, R., Valls, A. C. S., Bogo, M. R., Salzano, F. M., Smith, D. G., Silva, W. A., Zago, M. A., Ribeiro-dos-Santos, A. K., et al. (2008) Mitochondrial Population Genomics Supports a Single Pre-Clovis Origin with a Coastal Route for the Peopling of the Americas. *Am. J. Hum. Genet.*, **82**, 583–592.

86. Barbieri, C., Heggarty, P., Castrì, L., Luiselli, D. and Pettener, D. (2011) Mitochondrial DNA variability in the Titicaca basin: Matches and mismatches with linguistics and ethnohistory. *Am. J. Hum. Biol.*, **23**, 89–99.

87. Arias, L., Barbieri, C., Barreto, G., Stoneking, M. and Pakendorf, B. (2018) High-resolution mitochondrial DNA analysis sheds light on human diversity, cultural interactions, and population mobility in Northwestern Amazonia. *Am. J. Phys. Anthropol.*, **165**, 238–255.

88. Simão, F., Strobl, C., Vullo, C., Catelli, L., Machado, P., Huber, N., Schnaller, L., Huber, G., Xavier, C., Carvalho, E. F., et al. (2019) The maternal inheritance of Alto Paraná revealed by full mitogenome sequences. *Forensic Sci. Int. Genet.*, **39**, 66–72.

89. Motti, J. M. B., Schwab, M. E., Beltramo, J., Jurado-Medina, L. S., Muzzio, M., Ramallo, V., Bailliet, G. and Bravi, C. M. (2017) Diferenciación regional de poblaciones nativas de América a partir del análisis de los linajes maternos.

Intersecc. Antropol., **18**, 271–282.

90.   Sans, M., Figueiro, G., Hughes, C. E., Lindo, J., Hidalgo, P. C. and Malhi, R. S.
        (2015) A South American Prehistoric Mitogenome: Context, Continuity, and the
        Origin of Haplogroup C1d. *PLoS One*, **10**, e0141808.

91.   García, A., Pauro, M., Nores, R. and Demarchi, D. A. (2013) Diversidad del
        haplogrupo B2 en el centro de Argentina: ¿extinción de haplotipos intermedios? *XI
        Jornadas Nacionales de Antropología Biológica*, Universidad Maimónides y
        Facultad de Filosofía y Letras, Universidad de Buenos Aires, p. 183.

92.   Avise, J. C. (2000) *Phylogeography: The History and Formation of Species*; Harvard
        University Press (ed.); Cambridge MA.

93.   Piana de Cuestas, J. (1992) *Los indígenas de Córdoba bajo el régimen colonial
        (1570-1620)*; Córdoba, Argentina.

94.   Tovar, A. and Larrucea de Tovar, C. (1984) *Catálogo de las lenguas de América del
        Sur con clasificaciones, indicaciones tipológicas, bibliogrqfía y mapas*; Gredos
        (ed.); Madrid.

95.   Gayà-Vidal, M., Moral, P., Saenz-Ruales, N., Gerbault, P., Tonasso, L., Villena, M.,
        Vasquez, R., Bravi, C. M. and Dugoujon, J. M. (2011) MtDNA and Y-chromosome
        diversity in Aymaras and Quechuas from Bolivia: Different stories and special
        genetic traits of the Andean Altiplano populations. *Am. J. Phys. Anthropol.*, **145**,
        215–230.

96.   Tavella, M. P., Demarchi, D. A. and Nores, R. (2020) Diversidad genética en restos

humanos de contextos arqueológicos tardíos de la provincia de Santiago del Estero a partir del estudio de ADN mitocondrial. *Rev. Argent. Antropol. Biol.*, **22**, 013.

97. Lorandi, A. M. (1978) El desarrollo cultural prehispanico en Santiago del Estero, Argentina. *J. Soc. des Am.*, **65**, 63–85.

98. Taboada, C. (2017) Espacio, cultura material y procesos sociales tardíos en la llanura santiagueña. Modelo para pensar las poblaciones de la región. In Ventura, B., Ortiz, G., Cremonte, M. (eds.), *Arqueología de la vertiente oriental Surandina. Interacción macro-regional, materialidades, economía y ritualidad*, Sociedad Argentina de Antropología, Ciudad Autónoma de Buenos Aires, pp. 237–266.

99. Castro Olañeta, I. (2013) "Donde estan situados los mas yndios de la jurisdiccion desta ciudad". Un acercamiento etnohistórico a las encomiendas y pueblos de indios del Río Salado. *Surandino Monográfico*, **3**, 1–23.

100. Fabra, M. and Demarchi, D. A. (2013) Análisis morfogeométrico aplicado al estudio de los patrones espaciales y temporales de variación morfológica craneofacial en poblaciones del centro de Argentina. *Cuad. Inst. Nac. Antropol. Pensam. Latinoam.*, **1**, 87–101.

101. González, C. and Fabra, M. (2019) La muerte, los ritos, lo sagrado: prácticas funerarias entre los habitantes de la costa sur de la Laguna Mar Chiquita y llanuras adyacentes en el holoceno tardío (Córdoba, Argentina). *Jangwa Pana*, **18**, 445–468.

102. Lamenza, G. N., Calandra, H. A. and Salceda, S. A. (2019) Arqueología de los ríos Pilcomayo, Bermejo y Paraguay. *Rev. del Mus. La Plata*, **4**, 481–510.

103. Sevini, F., Yao, D. Y., Lomartire, L., Barbieri, A., Vianello, D., Ferri, G., Moretti, E., Dasso, M. C., Garagnani, P., Pettener, D., et al. (2013) Analysis of Population Substructure in Two Sympatric Populations of Gran Chaco, Argentina. *PLoS One*, **8**, e64054.

104. Postillone, M. B., Martínez, G., Flensborg, G. and Dejean, C. B. (2020) First analysis of mitochondrial lineages from the eastern Pampa–Patagonia transition during the final late Holocene. *Am. J. Phys. Anthropol.*, **171**, 659–670.

105. Luisi, P., García, A., Berros, J. M., Motti, J. M. B., Demarchi, D. A., Alfaro, E., Aquilano, E., Argüelles, C., Avena, S., Bailliet, G., et al. (2020) Fine-scale genomic analyses of admixed individuals reveal unrecognized genetic ancestry components in Argentina. *PLoS One*, **15**, e0233808.

106. Fabra, M. (2013) Variación Epigenética Craneofacial Y Dinámica Evolutiva De Poblaciones Humanas Del Norte De La Región Pampeana Durante El Holoceno Tardío. *Cuad. Inst. Nac. Antropol. Pensam. Latinoam. – Ser. Espec.*, **1**, 87–101.

107. Bayón, C., Pupio, A., Frontini, R., Vecchi, R. and Scabuzzo, C. (2010) Localidad arqueológica Paso Mayor: Nuevos estudios 40 años después. *Intersecc. Antropol.*, **11**, 115–128.

108. Motti, J. M. B., Muñoz, A. S., Cruz, I., D´Angelo del Campo, M. D., Borrero, L. A., Bravi, C. M. and Guichón, R. A. (2019) Análisis de ADN mitocondrial en restos humanos del Holoceno tardío del sur de Santa Cruz. In Gómez Otero, A., Sandoval, A., Benegas, A. (eds.), *Arqueología de la Patagonia: El pasado en las arenas*,

Editorial IDEAUS, CONICET-CENPAT, Puerto Madryn, pp. 493–503.

109. Fabra, M. and Demarchi, D. A. (2011) Geographic Patterns of Craniofacial Variation in Pre-Hispanic Populations from the Southern Cone of South America. *Hum. Biol.*, **83**, 491–507.

110. Laguens, A. G., Bonnin, M. I., Giesso, M. and Glascock, M. D. (2007) Más allá del horizonte: cazadores-recolectores e intercambio a larga distancia en Intihuasi (provincia de San Luis, Argentina). *Intersecc. Antropol.*, **8**, 7–16.

111. Medina, M., Pastor, S. and Berberián, E. E. (2014) ''Es Gente Fazil de Moverse de una Parte a Otra''. Diversidad en las Estrategias de Subsistencia y Movilidad Prehispánicas Tardías (Sierras de Córdoba, Argentina). *Complutum*, **25**, 73–88.

112. Medina, M. and Rivero, D. E. (2016) Human Colonization, Diet Breadth and Niche Construction during Sierras of Córdoba Holocene (Argentina). In Muscio, H., Cardillo, M. (eds.), *Darwin's Legacy: The State of Evolutionary Archaeology in Argentina*, Archaeopress, pp. 67–81.

113. Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. and Howell, N. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.*, **23**, 147.

114. Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.

115. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G.,

Bandelt, H.-J., Kronenberg, F., Salas, A. and Schönherr, S. (2016) HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.*, **44**, W58–W63.

116. Wright, Erik, S. (2016) Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. *R J.*, **8**, 352.

117. R Core Team (2020) R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria*.

118. Excoffier, L. and Lischer, H. E. L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.*, **10**, 564–567.

119. Weir, B. S. and Cockerham, C. C. (1984) Estimating F -statistics for the analysis of population structure. *Evolution (N. Y).*, **38**, 1358–1370.

120. Fu, Y. X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

121. Pybus, M., Dall'Olio, G. M., Luisi, P., Uzkudun, M., Carreño-Torres, A., Pavlidis, P., Laayouni, H., Bertranpetit, J. and Engelken, J. (2014) 1000 Genomes Selection Browser 1.0: A genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res.*, **42**, 903–909.

122. Hudson, R. R. (2002) Generating samples under a Wright–Fisher neutral model of genetic variation. *Bioinformatics*, **18**, 337–338.

123. Watterson, G. A. (1975) On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.*, **7**, 256–276.

124. Rambaut, A. and Grass, N. C. (1997) Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Bioinformatics*, **13**, 235–238.

125. Dufour, D. L. and Piperata, B. A. (2004) Rural-to-urban migration in Latin America: An update and thoughts on the model. *Am. J. Hum. Biol.*, **16**, 395–404.

126. Catelli, M., Álvarez-Iglesias, V., Gómez-Carballa, A., Mosquera-Miguel, A., Romanini, C., Borosky, A., Amigo, J., Carracedo, Á., Vullo, C. and Salas, A. (2011) The impact of modern migrations on present-day multi-ethnic Argentina as recorded on the mitochondrial DNA genome. *BMC Genet.*, **12**, 77.

127. Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V. and Richards, M. B. (2009) Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *Am. J. Hum. Genet.*, **84**, 740–759.

128. Sans, M., Mones, P., Figueiro, G., Barreto, I., Motti, J. M. B., Coble, M. D., Bravi, C. M. and Hidalgo, P. C. (2015) The mitochondrial DNA history of a former native American village in northern Uruguay. *Am. J. Hum. Biol.*, **27**, 407–416.

129. Letunic, I. and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.*, **47**, W256–W259.

130. Crandall, K. A. and Templeton, A. R. (1993) Empirical tests of some predictions

from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 959–969.

LEGENDS TO FIGURES

Figure 1. Locations of the samples analyzed in the present study. The green circles represent modern populations with the diameter being proportional to the number of individuals analyzed (also given as text). The red circles correspond to the approximated location of archaeological samples, labeled with sample ID. References for populations abbreviated names: Santiago del Estero province: Sumampa (SUM), Villa Atamisqui (VA); Córdoba province: Córdoba capital (CBA), Amboy (AMB), Cruz Alta (CA), La Carlota (LCA), Jovita (JOV), Villa de Soto (VSO), La Para (LPA), San Carlos Minas (SCM), Villa Dolores (VDO), Chancaní (CHA), Río Cuarto (RIV); San Luis province: La Toma (LTM), Santa Rosa del Conlara (SRC), Tilisarao (TIL).

Figure 2. Phylogenetic, Phylogeographic and Isofrequency analyses for Central Argentina local lineages. From left to right within each panel, we present a summarized phylogenetic tree, the median joining network and the isoline map are represented. (A) Lineage $A2_{+150}$ includes 77 samples, (B) Lineage $B2_{+16142}$ (B2ak1) includes 54, and (C) Lineage $C1b_{+146}$ (C1b16) includes 119. The phylogeographic analyses were performed using mtDNA sequences from positions 16032-16544 and 051-555 positions. The black

arrows show the nodal haplotypes. The dots in the isoline maps indicate the geographic location of the populations included in the analysis. Corresponding references and population grouping at the regional level are detailed in Table S6.

Figure 3. Time to Most Recent Common Ancestor for founding haplogroups and Central Argentina local lineages. The coalescence ages are presented in Table S4. The TMRCA distributions were obtained across the sampled iterations from the BEAST procedure with the 937 mitogenomes, and are represented with both grey violin plots and white boxplots. Braces: (1) Internal branches for local clades, with estimated TMRCA consistent with a human presence since the Pleistocene-Holocene transition (~9,000 - 11,000 ybp; upper shaded area), and (2) outermost nodes for local clades that occurred during the major pulse of population growth (~7,200 - 4,200 ybp; lower shaded area).

Figure 4. Bayesian Skyline Plot of female effective population size, based on a generation time of 25 years. We obtained a population size estimate over time for each sampled iteration from the Bayesian skyline analysis using the 113 Central Argentina mitogenomes. Black line and shaded area represent the average and 95% credibility interval of these estimates over time, respectively. Vertical dashed lines represent the time range for changes reported in the archaeological record that are likely linked to the major pulse of population growth (see main text).

Figure 5. Phylogenetic relationships between modern and ancient mitogenomes. Dendrogram that includes all the mitogenomes from the present study, with ancient samples shown as bold lines. The stars highlight the most recent coalescence of an ancient sample with a modern one.

Figure 6. Demographic scenario for the peopling and settlement of Central Argentina. The arrows represent the relationships and migration routes of the clades discussed in the article.

Table 1. Diversity parameters estimated from mitogenomes for 92 modern and 22 ancient samples of Central Argentina. (N) Sample size; (n) Number of observed nucleotide sites; (K) number of different haplotypes; (S) Number of polymorphic (segregating) sites; (Hd) Gene diversity; (π) nucleotide diversity; (Tajima's D) test of selective neutrality; (Fu's F(s)) test of selective neutrality; (sd) standard deviation. Italic font indicates statistically significant (p<0.05).

Population	N	n	K	S	Hd ± sd	π ± sd	Tajima's D	Fu's F(s)
Central Argentina (modern DNA)	92	16569	84	334	0.9980±0.0020	0.0021±0.0010	*-1.4687 (p=0.0411)*	*-24.0992 (p=0.0007)*
Córdoba (ancient DNA)	22	16564	20	140	0.9910±0.0170	0.0020±0.0010	-0.2014 (p= 0.4663)	-1.7345 (p= 0.1974)