SUPPLEMENTARY INFORMATION

Local agricultural transition, crisis and migration in the Southern Andes

Ramiro Barberena; Pierre Luisi; Paula Novellino; Augusto Tessone; Daniela Guevara; Angelina García; Elizabeth A. Nelson; Petrus le Roux; Claudia Herrera; Graciela Coz; Matías Candito; Maria Lopopolo; Maël Le Corre; Lorena Becerra-Valdivia; Miren Iraeta Orbegozo; Gaétan Tressières; Gustavo Lucero; Marcelo Cardillo; Julia Merler Carbajo; Gabriela Da Peña; Jorge Suby; Maguelonne Roux; María Eugenia de Porras; Candela Acosta Morano; Claudia Mallea; Lumila Menéndez; María Fernanda Quintana; María Laura López; Andrés Troncoso; Julie Luyt; Kerryn Gray; Francisca Santana-Sagredo; Ludovic Orlando; Víctor Durán; Judith Sealy; Etienne Patin; Lluis Quintana-Murci; Hannes Schroeder; Nicolás Rascovan

This PDF includes

List of Figures

List of Tables

Captions for Supplementary Data

Supplementary Data S1. Isotopic results for human samples from UV and site BB6 (results produced by other research teams are indicated).

Supplementary Data S2. $^{87}Sr/^{86}Sr$ results for rodent samples from the southern Andes of Argentina and Chile.

Supplementary Data S3. Information on sequencing effort.
Column A (Warning): LowCov flag when the final Depth of Coverage for the individual is below 0.01; Cont_XXX flag when contamination estimates was above 5% for method XXX.
Columns B-D (Individual, Sample and Library IDs): IDs used for archaeogenetics work and throughout the manuscript.
Column E (Skeleton Element Type): Tooth or Petrous bone.
Column F (archaeologistID): ID used for archaeological work and in previous publications.
Column G (Site): Archaeological Site.
Column H (Laboratory): Laboratory where ancient DNA work was performed. When two laboratories are listed, the first corresponds to decontamination and sampling, while the second corresponds to the subsequent steps.
Column I (Strandness): Library Strandness.
Column J (UDG): UDG treatment (none, half or full).
Columns K-L (#bp trimmed in 5' and 3’, respectively): Number of nucleotides removed on each read end to avoid integrating deaminated sites in the analyses.
Columns M and N (Lib.GenomicSex and Sample.GenomicSex, respectively): Sex inferred using only the DNA data for one library, or all libraries for each individual, respectively.
Column O (% Endogenous): Percentage of collapsed reads mapping to the human reference genome.
Column P (% Duplicates): Percentage of duplicated collapsed mapped reads.
Column Q (MeanReadLength): Mean deduplicated collapsed mapped read length
Columns R-AC: Breadth of Coverage (BoC) and Depth Of Coverage (DoC): proportion of sites covered by at least one read and mean number of reads per site, respectively. BoC and DoC are provided for each library individually, and when merging all libraries for each individual, as well as when considering different lists of variants (whole autosomal genome, only autosomal sites on the 1240K SNP panel, and whole mitochondrial genome).
Columns AD-AT: Output from different contamination estimation methods.

Supplementary Data S4. Individuals included in different genomic analyses. Individuals included in each type of analysis. Group refers to the groupings used for non-individual level analyses (for UV individuals, different groupings were used). Those groups were subsequently assigned to major groups, according to geographical ranges and period: Early Holocene (EH; more than 8,500 BP), Middle Holocene (MH; between 8,500 and 3,500 BP), Late Holocene (LH; less than 3,500 BP), and Modern Major group. SG and 1240K in column names refer to analyses performed considering only sites from these SNP panels (and adding TVs to the column names when only considering transversions).

Supplementary Data S5. Kinship results. a. KIN results when using the 1240K SNP panel**. b.** KIN results when using the SG SNP panel**. c.** READv2 results when using the 1240K SNP panel**. d.** READv2 results when using the SG SNP panel**. e.** Final assignation of relatedness degree according to results from **a-d**

Supplementary Data S6. *f*-statistics analyses. For each comparison type we provide the results of different *f*-statistic analyses performed on two SNP panels, using all positions or only transversions.
a. Results for f_3(Ind1,Ind2; Mbuti), where Ind1 and Ind2 are two individuals provided as Source1 and Source2.
b. Results from Multidimensional Scaling analyses from genetic distances of the form 1-f_3, as shown in **a.**
c. Results for f_4(Mbuti,Modern;UV,SA), where Modern and SA are a modern and an ancient South American group, respectively, and UV stands for all analyzed Uspallata Valley individuals grouped together.
d. Results for f_4(Mbuti,SNA; UV1, UV2), where SNA is an ancient American group assigned to South Native American Lineages (SNA), and UV1 and UV2 stands for two groups in Uspallata Valley: Hunter-Gatherers (LH-HG), Migrant Farmers (LH-MF), Migrant Outlier Farmer (LH-MFout), Local Farmer from Potrero Las Colonias), and Local Farmers.
e. Results for f_4(Mbuti,UV;SNA1, SNA2), where SNA1 and SNA2 are two ancient American groups assigned to South Native American Lineages (SNA), and UV stands for all analyzed Uspallata Valley individuals grouped together.
f. Results for f_4(Mbuti,North/Central;UV,South), where North/Central is an ancient group assigned from Central or North America, South is an ancient group from South America, and UV stands for all analyzed Uspallata Valley individuals grouped together.
g. Results for f_4(Mbuti, SNA;LH-HG,LH-F), f_4(Mbuti, LH-HG;SNA,LH-F), and f_4(Mbuti, LH-F;SNA,LH-HG), where SNA is an ancient American group assigned to South Native American Lineages (SNA), and LH-HG and LH-F are the Hunter-Gatherers and Farmers from Uspallata Valley,respectively.

Supplementary Data S7. Conditional heterozygosity and effective population size.
a. Results for Conditional Heterozygosity analyses on the 1240K SNP panel. **b.** Results for Conditional Heterozygosity analyses on the SG SNP panel. **c.** Effective population size (Ne) estimates from hapROH performed using the 1240K SNP panel.

A. Ethics statement and community engagement actions

The sampling of the archaeological human remains was made under the required permits for archaeological research from Argentina (Provincial Law 6.034). The human remains studied in this work are placed in the Museo de Ciencias Naturales y Antropológicas Juan Cornelio Moyano, which is the official repository of the Province of Mendoza. The sampling and exportation of the samples used for ancient DNA (aDNA), isotope analyses, and ^{14}C dating were carried out with the required export permits to perform these procedures, granted in by the authorities of the Cultural Heritage of the Province of Mendoza. With this primary permit, a second export permit was requested from the National Registry of Archaeological Sites, Collections, and Objects (RENYCOA) depending from the Instituto Nacional de Antropología y Pensamiento Latinoamericano (INAPL), the national jurisdiction body that works with provinces and security forces to implement measures aimed at protecting archaeological heritage and evaluating and authorizing requests for the study of archaeological remains in laboratories abroad. The export permits from RENYCOA for the samples analyzed in this study are the following: EX-2021-18124171, EX-2021-18137399, EX-2020-85504578, EX-2020-84737313, EX-2020-86333220, EX-2022-35037747, EX-2022-35044700, EX-2022- 66036837, and 2023-47477466).

In addition to adhering to formal regulations, further steps were taken that involved advising from Indigenous Huarpe community members to ensure its ethical development. Within the framework of this project, the teams led by NR and RB work in conjunction with the Huarpe communities Guaytamari and Llahué Xumec from the territory of Uspallata Valley, engaging in activities to discuss our findings and the narratives surrounding them. Some of these engagements were publicized, such as the one documented here: https://www.mendoza.gov.ar/prensa/investigadores-del-instituto-pasteur-de-francia-realizan-investigaciones-sobre-poblaciones-prehispanicas-de-mendoza/

This was formalized in the temporary multivocal museum exhibit "*Hane Pecne* (Madre Algarrobo): tejiendo historias entre arqueología y pueblos originarios en los Andes", held at the Museo de Ciencias Naturales y Antropológicas Juan Cornelio Moyano between May 17th and November 1st of 2024.

https://www.mendoza.gov.ar/prensa/llega-la-muestra-hane-pecne-tejiendo-historias-entre-arqueologia-y-pueblos-originarios-en-los-andes-al-museo-moyano/. This exposition was guided by combined teams of students of Archaeology at the Universidad Nacional de Cuyo (Mendoza) and members of the indigenous communities (Fig. S1).

Fig. S1. Inauguration of the *Hane Pecne* exposition.
Omtas (leaders) Claudia Herrera and Graciela Coz at the inauguration of the *Hane Pecne* (Madre Algarrobo) exposition. Museo de Ciencias Naturales y Antropológicas Juan Cornelio Moyano, May 17th, 2024).

Successive meetings between the academic scholars and the Guaytamari and Llahué Xumec communities, ensuring that their narratives, vocabulary and terminology were respectful of Indigenous concerns. Three members of these communities (CH, GC, MCan) contributed to the interpretation of results and participated in the writing of the manuscript, and are thus co-authors of the present study. During this process, RB, PL and NR provided Spanish translations of the manuscript, including explanations facilitating meaningful exchange and ensuring accessibility for all spanish-speaking collaborators (Supplementary Text S1).

B. Description of the geological landscape, archaeological contexts, and sampling design

B.1 Geological landscape and bioavailable strontium

The study area (Main Fig. 1) encompasses Uspallata, an intermontane valley located in northwestern Mendoza Province between two morpho-structural geological units—Cordillera Frontal and Precordillera. UV is connected to other longitudinal valleys to the north (Calingasta and Iglesia), comprising a ~400 km-long biogeographic corridor[1]. Situated at an altitude of ~1900–2200 masl, UV can be occupied year-round, unlike the higher areas surrounding it, where the winter snow-cover limits human occupation to the summer season. Hence, this valley is a key region for tracking human migrations across and along the southern Andes.

The geology of the southern Andes is especially suited for tracking local residence and immigration through strontium isotopes ($^{87}Sr/^{86}Sr$) due to the marked diversity of bedrock age and composition in a restricted spatial scale (Fig. S2). The study area of the strontium project encompasses Central Argentina and Chile and is located in the transition between the Pampean flat slab, characterized by a shallow subduction angle resulting in a volcanic gap, and the normal subduction zone of central Chile and Argentina[2]. This segment corresponds to the northernmost active volcanism of the Southern Volcanic Zone of the Andes. In this latitudinal block, the Andes are composed of seven major morpho-structural provinces that provide our framework for the sampling and analysis of bioavailable strontium. The geological framework has been presented and analyzed in detail in previous work[3–5]. Broadly, rock age increases from the Central Valley in Chile (Western Principal Cordillera), assigned to the Oligocene-Miocene periods (<40 my), to the Paleozoic Precordillera (>300 my) in western Argentina, which flanks the UV and contributes to its highly radiogenic signature. This geological background is reflected in our $^{87}Sr/^{86}Sr$ results, showing values that increase eastwards from Principal Cordillera (~0.70393) and reaching its maximum ratios in the Precordillera east of UV (~0.71011) (Supplementary Data S2)[4].

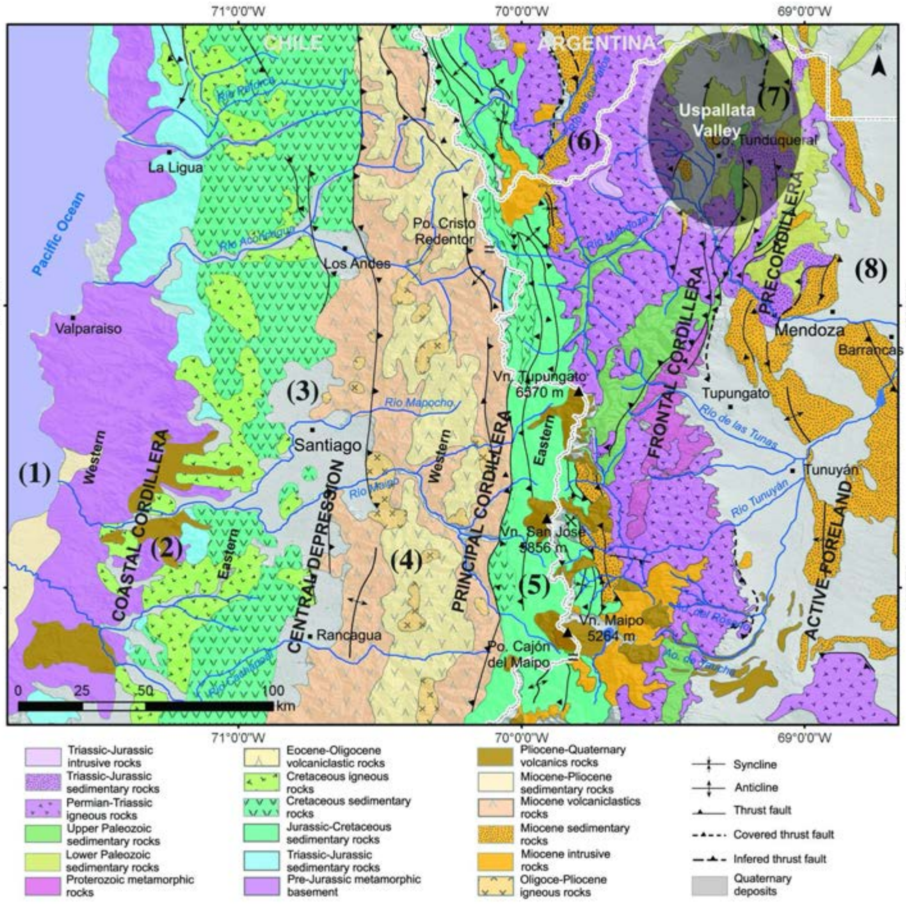

Fig. S2. Geological structure of the Uspallata Valley and neighboring areas.
References: (1) Pacific Ocean; (2) Coastal Cordillera; (3) Central Depression-Central Chilean Valley; (4) Western Principal Cordillera; (5) Eastern Principal Cordillera; (6) Frontal Cordillera; (7) Precordillera; and (8) Active Foreland or Quaternary lowlands.

B.2 Archaeological contexts

Table S1 presents a synthesis of the properties and specific studies of the archaeological sites analyzed. Nearly 200 individuals were retrieved from burials between 1200 and 600 cal years BP at seven clustered archaeological sites in UV. By including the hunter-gatherer site Barrancas B6 (BB6), emplaced in the eastern lowlands 70 km from UV, our total sample sums 233 individuals which span the last ca. 2300 years. We present additional contextual details for each site next –chronologically organized, from the oldest to the most recent–.

The sites Túmulo II (T-II), Túmulo I (T-I), Túmulo III (T-III), Potrero Las Colonias (PLC), and Agua de las Avispas (A.Av.) were all excavated by the pioneer researcher Carlos Rusconi since the 1930's[6–9]. All these remains are preserved in the Museo de Ciencias Antropológicas y Naturales Juan Cornelio Moyano (Mendoza, Argentina). While the now outdated techniques of excavation applied by Rusconi limit the quality of available contextual information, particularly in the case of the ossuaries PLC and T-III, this researcher was highly skilled and the produced information can be considered robust in terms of the general arrangements and presence or absence of grave goods.

B.2.1 Barrancas B6 (BB6)

BB6 is a multiple cemetery located 70 km from UV in the eastern lowlands adjacent to the modern city of Mendoza[10–12]. This site was excavated by coauthors PN and VD between 2009 and 2011. Based on six direct dates for human remains (Table S1), we establish that the site formed between 2275 and 2165 cal years BP with a median interval of 94 years, spanning between three and four human generations. The site contains the remains of 34 individuals of different ages. Interestingly, 16 out of 17 adults are male. A large part of these individuals were interred in spatial association following a fan pattern (Fig. S3). The isotopic results for $\delta^{13}C_{coll.}$, $\delta^{13}C_{ap.}$ and $\delta^{15}N$ indicate a C_3 trophic chain with a null maize consumption. The regional context, chronological position and isotopic results suggest a hunter-gatherer subsistence including wild plants and animals. We present genomic results for six individuals.

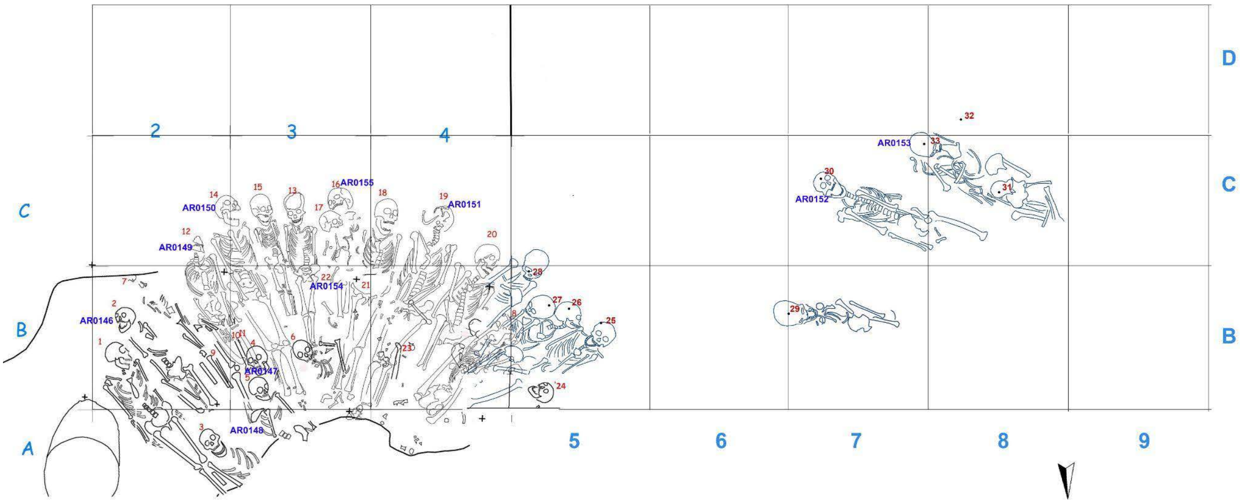

Fig. S3. Plan-view of the burials at the hunter-gatherer BB6 site.
IDs in purple for individuals with genomic, isotopic, and chronological information.

B.2.2 Túmulo II (T-II)

T-II is a primary burial containing the remains of 10 individuals of different age and sex (Table S1, Fig. S4). Three direct dates place the formation of this burial between 1000-1300 cal years BP. Stable isotopes show a very high C_4 consumption in the individuals from T-II, clearly identifying them as the earliest farmers in UV (see Main Text). The $^{87}Sr/^{86}Sr$ signatures recovered from 14 samples belonging to nine individuals (five were duplicated to analyze childhood/adulthood) match the local baseline for UV, identifying them as local. We present genomic results for three individuals.

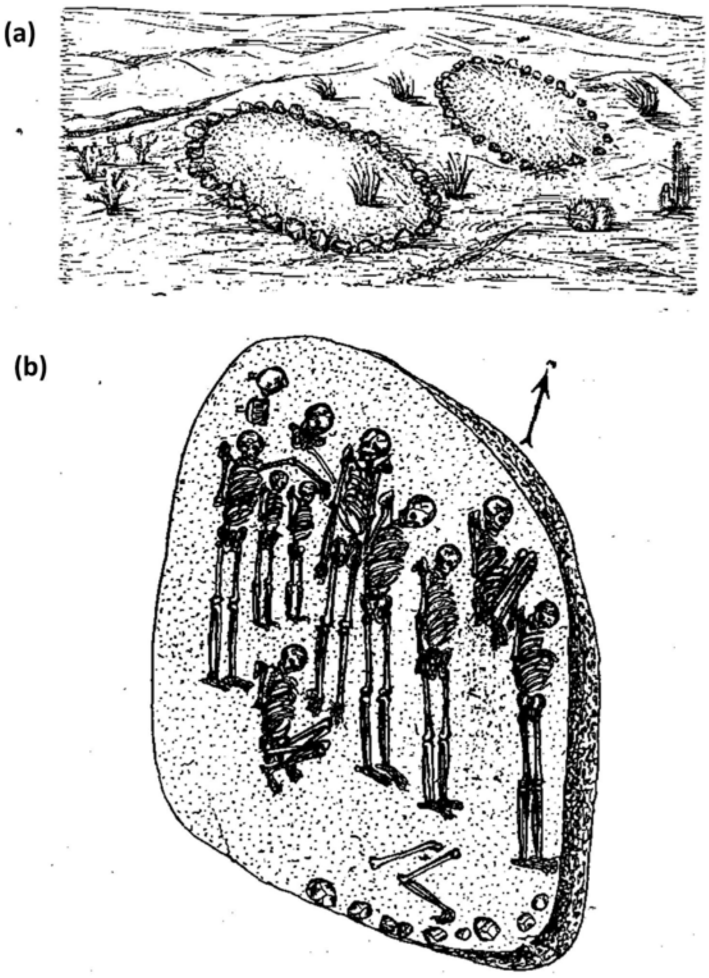

Fig. S4. Archaeological context for Túmulo I and Túmulo II sites.
a. Image of the sites Túmulo I and Túmulo II; **b.** Detail of the burial Túmulo I (modified from[7], pp. 363, 387).

T-II included an assemblage of grave goods that include two ceramic vessels of local cultural styles in relation with the male adult individual 245 (AR0185), one lithic lip plug–tembetá–recovered in association with the female adult individual 239 (AR0180), and a triangular lithic projectile point deposited with the adult individual 241 of undetermined sex (Fig. S5). There is no evidence that this point produced the death of the individual and may have been part of the mortuary assemblage.

Fig. S5. Grave goods of the Túmulo II site.
Detailed study from[13]: **a.** subglobular black and red vessel with polished surface and presence of oligist iron; **b.** small black vessel with burnished and engraved surface with geometric motifs; **c.** triangular lithic projectile point associated with individual 241.

B.2.3 Túmulo I (T-I)

The site T-I is located adjacent to T-II within UV (Fig. S4a). It is a large cemetery with the remains of 28 individuals deposited as primary burials. Three dates for human remains indicate a span of site formation between 960 and 850 cal years BP with a median interval of 120 years (~5 generations). The results for stable isotopes for seven individuals show a largely C_3 signature with little C_4 input, compatible with the consumption of wild herbivores such as guanaco–*Lama guanicoe*– and C_3 plants[1]. The $^{87}Sr/^{86}Sr$ signatures recovered from 14 samples from 14 individuals match the local baseline for UV, identifying them as local (Supplementary Data S1).

B.2.4 Potrero Las Colonias (PLC)

The site PLC has an Minimum Number of Individuals (MNI) of 124 individuals and is the largest from UV and northern Mendoza Province[7,14,15]. PLC has been described as an ossuary in a small burial pit where the remains of multiple individuals were deposited in an unorganized manner (see image and artistic recreation by the excavator in Fig. S6a, and stratigraphic profile in Fig. S6b). Importantly, despite the very large number of individuals recovered, no grave goods were associated. The site has a relatively balanced representation by age and sex (Table S1). 14 dates for different individuals allow bracketing the formation of the site between the modelled medians of ~680 and 600 cal BP, respectively, with a duration of 55-100 years–median of 75 years–(see main text). The $^{87}Sr/^{86}Sr$ signatures recovered from 48 samples from 32 individuals match the local baseline for UV, identifying them as non-locals (Supplementary Data S1), while one individual (AR0520) has a local signal throughout life. Stable isotopes show intensive C_4 (maize) farming diets for all the samples analyzed (n= 52). We produced genomic results for 36 individuals.

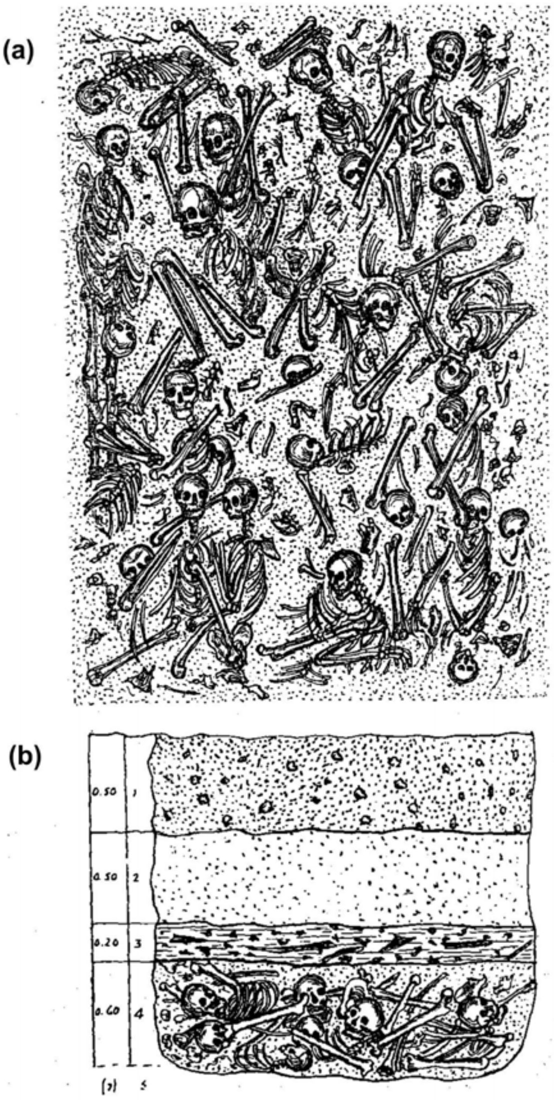

Fig. S6. Potrero Las Colonias (PLC) cemetery site.
a. artistic sketch of the burial pattern by the excavator Carlos Rusconi[9]; **b.** stratigraphic profile of PLC showing the deposit with human remains (modified from[7]).

B.2.5 Túmulo III (T-III)

T-III site is an ossuary where human remains were deposited without associated grave goods. It has an MNI of 26 individuals, and thus is the second largest site in UV after PLC. In terms of the archaeological context, this site is comparable to PLC. Adults and subadults are almost equally represented (Table S1), and intriguingly, only female adults were determined (n=9). 14 samples from 14 individuals were analyzed for ^{87}Sr/^{86}Sr, producing a coverage of 54%. Of these, 11 are local (79%) and three non-locals (21%). We produced paleodietary determinations for three of these individuals and dates for five (Table S1).

B.2.6 Usina Sur 2 (US2)

The site US2 was excavated by VD and RB's team in 2017. It is located adjacent to the classic site Usina Sur that was previously excavated[16–18], but they do not form part of the same burial structure[14]. US2 contains the partial remains of four individuals including two subadults and two adults, one of which is a woman and the other is unknown (Table S1). These remains were found eroding in a small ravine (Fig. S7). No grave goods were found associated, but since the site was eroding at the time of the rescue excavation, this cannot be ruled out. We processed two samples for ^{87}Sr/^{86}Sr and paleodiet, as well as one radiocarbon date. The strontium results showed non-local origins.

Fig. S7. Small ravine with site US2 (2017).

B.2.7 Agua de las Avispas (A.Av.)

A.Av. corresponds to an individual valley recovered by Rusconi near the UV, adjacent to the Precordillera (Main Fig. 1). We produced a direct date and a ^{87}Sr/^{86}Sr determination for this individual that showed a non-local signal. No grave goods were found associated with it.

B.2.8 Barrio Ramos I (BRI)

BRI is a multiple burial with the remains of six individuals of different ages and sex (Fig. S8a). The site has been dated to the initial Inka occupation of UV[19,20] and contains diverse grave goods including one lithic projectile point, bone points, one bone spatula, shell beads, one ceramic vessel (Fig. S8b). Three individuals from this site were isotopically analyzed by duplicates to reconstruct life-histories. We present genomic results for two individuals.

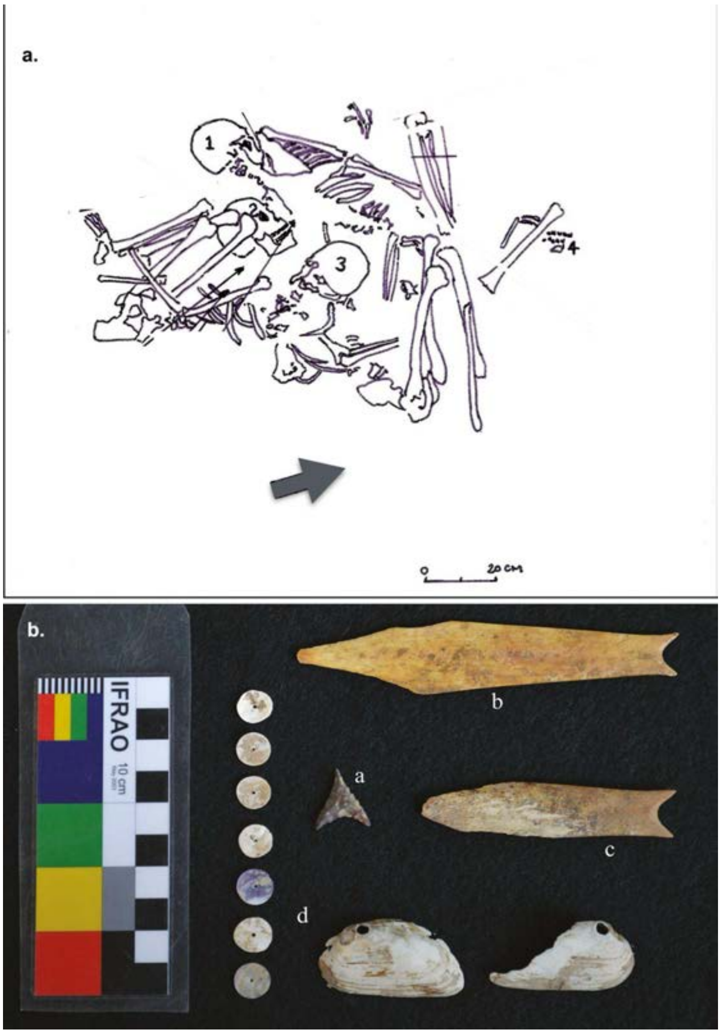

Fig. S8. Barrio Ramos I (BRI) site.
a. Arrangement of the six individuals buried in BRI; **b.** Grave goods in BRI: a. lithic projectile point, b. and c. bone projectile points, d. shell beads of *Diplodon* sp. (modified from[19]).

B.3 Migration or post-mortem transport?

An important issue regarding the mortuary practices involved in the formation of PLC lies in the possibility of the post-mortem transport of the remains to UV. This would likely be associated with total or partial modifications of the dead bodies in order to facilitate transport of such a large number of individuals[21–24].

There is contextual and taphonomic information that refutes this alternative. Firstly, at a contextual level, Rusconi's data show that several skeletons from PLC were discovered articulated in flexed position within a single multiple burial (Fig. S6a). This could reflect one of several burial events at this redundantly used burial site. At an anatomical level, most parts are well represented and preserved, including small bones of the hands and feet, which are among the first to disarticulate[25]. In addition, our detailed taphonomic analysis of the remains (see below) did not record cut-marks in the bones, which would necessarily result from a manipulation of the bodies for transportation. Finally, the representation of all the age groups, ranging from perinates and neonates, infants, children and adults of all ages[26,27]. Overall, building on the archaeological context, anatomical composition, and taphonomic properties, it can be solidly inferred that these burials composed by tens of individuals were not subject to significant post-mortem transport[21,23,28].

Site	**Archaeological context**	**MNI**	**Age classes**		**Sex**			**N ^{14}C dates**	**Samples ^{87}Sr/^{86}Sr**	**Individuals ^{87}Sr/^{86}Sr**	**Samples paleodiet**	**Genomic study**
			SA	**A**	**M**	**F**	**ND**					
Túmulo I (T-I)	Cemetery with primary burials	28	9 (32%)	19 (68%)	4 (14.3%)	9 (32.1%)	15 (53.6%)	3	14	14	7	–
Túmulo II (T-II)	Cemetery with primary burials and diverse associated grave goods	10	3 (30%)	7 (70%)	3 (30%)	3 (30%)	4 (40%)	3	14	9	15	3
Túmulo III (T-III)	Ossuary with no associated grave goods	26	14 (54%)	12 (46%)	0	9 (34.6%)	17 (65.4%)	5	14	14	3	—
Potrero Las Colonias (PLC)	Cemetery with no associated grave goods	124	54 (43,6%)	70 (56,4%)	28 (22.6%)	35 (28.2%)	61 (49.2%)	14	50	33	52	35
Barrio Ramos I (BRI)	Multiple primary burial, abundant Inka-period grave goods	6	3 (50%)	3 (50%)	2 (30%)	1 (20%)	3 (50%)	2	6	3	6	2
Usina Sur 2 (US2)	Multiple primary burial, no associated grave goods	4	2 (50%)	2 (50%)	0	1 (25%)	3 (75%)	1	2	2	2	—
Agua de las Avispas (A.Av.)	Isolated remains	1	0	1 (100%)	1 (100%)	0	0	1	1	1	1	—
Barrancas B6 (BB6)	Cemetery in subterranean pit	34	17 (50%)	17 (50%)	16 (47%)	1 (3%)	17 (50%)	6	20	14	20	6
Total		233	83	113	35	42	119	**35**	121	88	106	47

Table S1. Archaeological contexts and interdisciplinary analyses presented in this study.
References for context: T-I, [7,8,14]; T-II, [7,8,13]; T-III, [7,8,14]; PLC, [7,8,15]; BRI, [19]; US2, [29]; A.Av., [8]; BB6, [10].

B.4 Sampling design and results

B.4.1 Human remains

As part of an exploratory design, we firstly produced an initial isotopic sampling of the sites in UV[29]. This first sampling is expanded here considering variation in age, sex, and archaeological context. This dataset spans a temporal transect of 1,700 years, ranging from hunter-gatherer to incipient, mixed, and intensive farmers. For UV at large and the adjacent BB6 site, we have currently produced strontium isotopic information for 88 individuals from an estimated MNI of 233, (38% coverage). As mentioned in the main text, we have also produced duplicated sampling targeting childhood (teeth or petrous bones) and adulthood (bones) for a total of 32 adult individuals that allow assessing life-history changes, producing an overall number of 121 samples of human remains analyzed for $^{87}Sr/^{86}Sr$. In addition, there is paleodietary information for 106 human samples, of which 89 were produced by our team and 17 come from the literature[30,31] (Supplementary Data S1). Basically, all of the sites included in this study have been intensively sampled for isotopes and dating.

We attempted to generate genome-wide ancient DNA (aDNA) data from 46 individuals selected from the four largest of these sites using “Shotgun” sequencing (see Supplementary Material Section E.). This genomic dataset encompasses hunter-gatherers (n=6) to incipient, mixed, and intensive farmers (n=40), and includes 18 females and 17 males as determined by their genomic sex (Supplementary Material Section E.5.). Individuals with low-quality genomic data (n=13) were excluded (Methods). The final dataset for population genetic analyses included 33 unrelated (Supplementary Material Section E.7.) individuals with depth of coverage between 0.0216x and 5.5877x (median 0.873x; Supplementary Data S3).

B.4.2 Bioavailable strontium

The transect for bioavailable strontium extends longitudinally for over 350 km between the Pacific Ocean in Chile and the eastern lowlands in Argentina (Main Fig. 1). We sampled small rodents–with restricted home ranges–and plants, but found that rodents offered a more robust proxy for the study of human ranges of mobility[4]. The rodent samples on which our strontium isoscape is based include archaeological and modern samples—from contexts not affected by recent anthropic activity[32]. The geological regions described above provided the units for bioavailable strontium sampling. 91 rodent samples were analyzed, distributed as follows: (1) Pacific coast= 13; (2) Coastal Cordillera= 5; (3) Western Principal Cordillera= 12; (4) Eastern Principal Cordillera= 8; (5) Frontal Cordillera= 13; (6) Precordillera and mountain valleys (includes UV)= 27; (7) Active Foreland= 13 (Supplementary Data S2).

C. Bioarchaeological work

The remains from all the sites (Table S1) were analyzed from a bioarchaeological and paleopathological perspective. However, most of the results presented here will be focused on the site PLC, since the majority of the skeletons analyzed were found in this site.

C.1 Demographic profiles

The commingled characteristics of the sites PLC and T-III, comprising the largest part of our sample, combined with outdated recovery techniques applied in the 1930’s by Rusconi[7,9], lead to mixed assemblages where the integrity of the specimens for each individual has been largely lost. To accurately estimate the Minimal Number of Individuals (MNI)[33,34], we began by separating specimens from adults and subadults. The analysis was adjusted by considering additional variables such as morphology, size, robusticity and presence of distinctive morphologic features in the remains[26].

For the estimation of age of death we considered diagnostic elements from adults and subadults, while sex estimations were conducted only for adults based on the structure of the skull and os coxae. To reconstruct the age at death profile, we divided the samples into Subadults (prenatal stage to the fusion of the epiphyses of the postcranial skeleton, ca. 20 years) and Adults (>20 years)[26,35]. For subadult individuals, the sequence of tooth formation and eruption was considered, following established criteria[27], the maximum length of the long bones (humerus and femur) and the patterns of obliteration of the apophysis and epiphysis[35], according to these categories: fetal (before birth–40 weeks); infant (0–4.9 years); child (5–14.9 years); adolescent (15–19.9 years). In the case of adult individuals, morphological changes in the pubic symphysis and the auricular and retroauricular surface of the ileum were assessed[36,37]. Cranial suture fusions were also assessed following[26], given its good preservation and associated sex information. We applied the following age intervals: peri-neonate, 38 weeks of gestation-1 month; infant, 1 month-4 years; child 4-11 years; adolescent, 11-17 years; young adult I, 18-23 years; young adult II, 24-35 years; adult, 36-49 years; older adult, >49 years.

From the PLC site, we quantified 875 diagnostic bone specimens from cranial and postcranial anatomical elements. The quantitative analysis determines a MNI of 124 individuals composed of 70 adults and 54 subadults. The distribution by age categories was as follows: one perinate, ten infants, eight children, three adolescents, two young adults I, four young adults II, thirty-eight adults and seventeen older adults (Fig. S9). It is important to note that all categories are represented. Sex estimations on the preserved crania from adults show the presence of 8 females (38%) and 13 males (62%). Based on the sample of os coxae, there are 30 females (43%), 24 males (34%) and 16 undetermined individuals (23%) (Table S1). The site T-III contains 14 subadults (54%) and 12 adults (46%), nine of which are women (75%) and three indeterminate (25%).

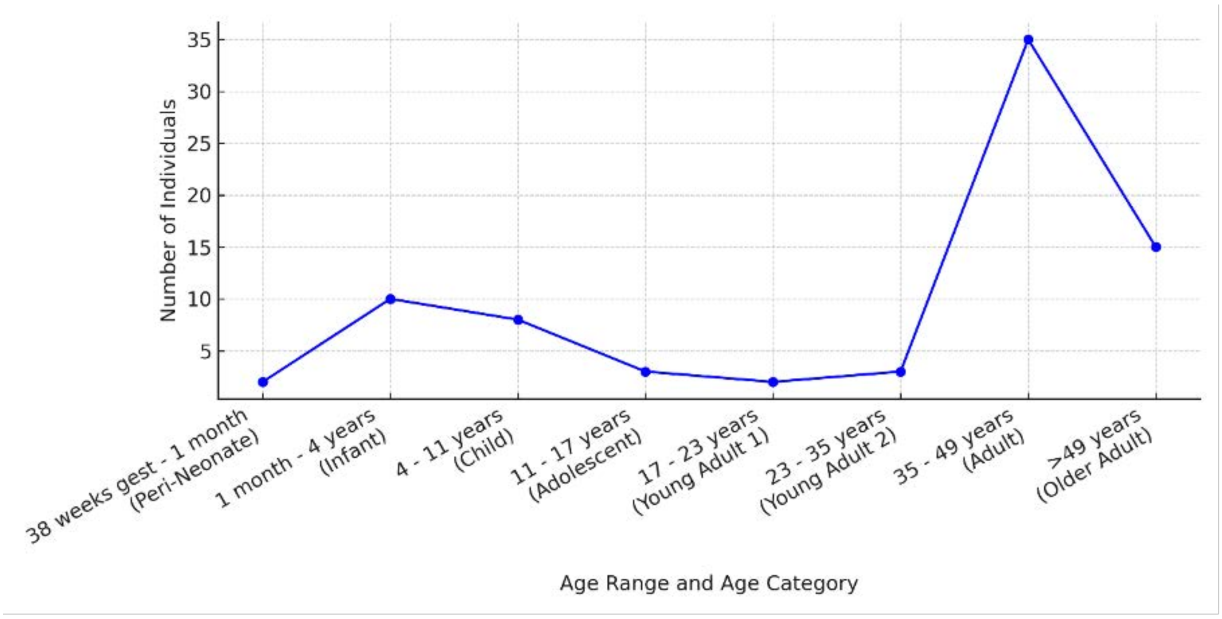

Fig. S9. Demographic profile of the site PLC.

C.2 Stress, trauma, and infections

We recorded the presence of porotic lesions in crania from adults and subadults with over 70% of completeness, and in humeri and femurs with over 70% of the proximal segment preserved. To localize lesions we followed a standard methodology[38]: cranial vault (porotic hyperostosis, PH), cranial orbits (cribra orbitalia), in femora (cribra femoralis), and humeri (cribra humeralis). The severity and healing condition of the lesions at time of death was assessed based on qualitative and quantitative criteria[38].

In PLC, 25 of 32 (78%) adult skulls with adequate analytical conditions exhibited PH, as did 9 of 10 (90%) subadult skulls (Table S2). The orbital cribra was recorded in only 3 of 27 (11%) adult skulls, and in 6 of 9 subadult skulls (67%). In the postcranial skeleton of adults, from a total of 65 femurs, CF was recorded in three (5%) with mild severity and in a condition of active scarring at the time of death. In subadults, cribra humeralis was observed in 4 of 18 (22%) elements, and cribra femoralis in 14 of 16 (87.5%) elements, along with nonspecific bone infections (Fig. S10). Linear enamel hypoplasias (LEH) were recorded in the anterior dentition (incisors and canines) following standard methodologies[39,40]. Twelve of 22 (54%) crania, mandibles and maxillae analyzed showed at least one LEH. Of the 22 skulls and mandibles analysed, 12 (54%) exhibited bilateral LEH, mainly in canines. In most of the individuals with LEH, at least one event was recorded. However, up to three events of LEH were identified in two individuals[14].

Site	Archaeological context	Adults			Subadults			
		PH	CO	LEH	PH	CO	CH	CF
BB6	Hunter-gatherers, locals	62	12	44	-	-	-	-
T-II	Early intensive farmers, locals	85	0	0	50	100	-	86
T-I	Mixed diets, locals	50	0	0	20	33	40	50
T-III	Late intensive farmers, locals and few migrants	-	-	-	-	38	0	17
PLC	Late intensive farmers, migrants	78	11	54	90	67	22	87
BRI	Mixed diets, locals	0	0	33	-	-	0	0

Table S2. Lesion abundance (in %) of systemic stress in sites from UV and BB6 site. Data from[14,41]. PH, porotic hyperostosis; CO, cribra orbitalia; LEH, linear enamel hypoplasia; CH, cribra humeralis; CF, cribra femoralis.

Regarding infectious diseases, macroscopic observation of bone elements was performed with the naked eye, 3x and 10x magnifying glass. In cases of pseudopathology or equifinality[42], a 10x binocular magnifying glass was used. Systematic recording followed criteria proposed by these refs[43–45]. Subsequently, analysis and differential diagnosis was performed, based on the available paleopathological and clinical information and the application of the adapted method for paleopathological analysis of the "Istanbul Protocol for the Effective Investigation and Documentation of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment"[46]. This protocol establishes degrees of diagnostic certainty as listed: not consistent, consistent, highly consistent, typical of and diagnostic[47]. We present a number of examples in Fig. S10.

Fig. S10. Examples of bone infections at the site PLC.
a. Cranium from individual PLC-296 with woven bone formation with porosity; **b.** Left humerus with an osteolytic lesion with remodelled margins on the posterior aspect of the proximal metaphysis and woven bone formation on the posterolateral aspect of the proximal third and the crest for the *triceps brachii*; **c.** Right fibula exhibiting woven bone formation with porosity on the surface of the interosseous membrane; **d.** Cranium with woven bone formation with porosity from individual AR0524 (PLC-301).

Traumatic lesions, including cranial compressions and cranial and post-cranial fractures, were described according to the element affected, its laterality, and the kind of tissue involved (lamellar, woven or mixed bone tissue). The cases considered as resulting from interpersonal violence showed either characteristics compatible with the use of weapons or the presence of projectiles associated with the lesions. Fractures were recorded as complete or incomplete, and aligned or not[44]. The ante or post-mortem character was determined based on the macroscopic formation of woven or lamellar tissue[44,48]. Perimortem lesions were identified when possible due to alterations related to violence (i.e. lesions suspected to be caused by weapons or projectiles). The lesions were quantified in relation with the total of elements and MNI for each site. In the case of PLC, only five lesions compatible with trauma were recorded (4% of the MNI) (Fig. S11).

The traumatic lesions were recorded in three crania (two females and one male, 5% of the MNE) and two long bones (0.3% of MNE). The first cranium showed two compressions of reduced size (circumferences of 13.7 and 8.7mm) in both parietals, both were fully remodelled and lacking fracture lines. The cranium of the second individual had a depression of 16mm of circumference without fractures. The last cranium showed a fracture in the right nasal bone.

Regarding the lesions on long bones, one affected the anterior distal portion of a right radium with a modification of the morphology and linearity in a posterior sense. Considering its location, close to the carpals, it is compatible with a Colles fracture[49]. The other long bone is a left cubit with an augment of the diameter of the distal diaphysis and a remodelled fracture line with lamellar tissue and a mild angulation in a posterior sense. This is compatible with a Parry fracture associated with a defensive action during a violent attack[50] (Fig. S11).

Intensive maize cultivation, as measured by $\delta^{13}C_{ap.}$, is associated with higher systemic stress in the farmers, whether locals (T-II) or migrants (PLC). Both sites display a high prevalence of PH, CO, and CF (Table S2). However, the migrants from PLC show significantly higher abundance in almost all indicators of systemic nutritional stress across age classes (adults: PH and LEH; sub-adults: PH, CO, and CF). In contrast, the sites BB6 (hunter-gatherers), TI and BRI (diachronic local populations with mixed C_3-C_4 diets) show a lower prevalence of these indicators of systemic stress.

There is robust evidence of malnutrition during childhood and adulthood in the local (T-II) and migrant (PLC) farmer populations. This is reflected in very high abundances of PH, CO, and CF, and also in the representation of LEH only for the migrants, which is a nonspecific indicator of nutritional deficiencies during crown formation[40,51,52]. This is more evident when compared against the hunter-gatherers from BB6 and with groups with mixed subsistence strategies from different periods (T-I, BRI). In PLC, particularly, there is also a relatively high incidence of bone infections that, while nonspecific (Fig. S10), would be compatible with the confirmed *Mycobacterium tuberculosis*. While a direct causal relationship cannot be assumed, previous studies have associated elevated frequencies of CO and PH with the occurrence of chronic respiratory infections such as tuberculosis[38,53]. On the other hand, levels of conflict were very low during all the agropastoral period, as shown by the scarce evidence of traumatic lesions, not all of which can be confidently related with interpersonal violence.

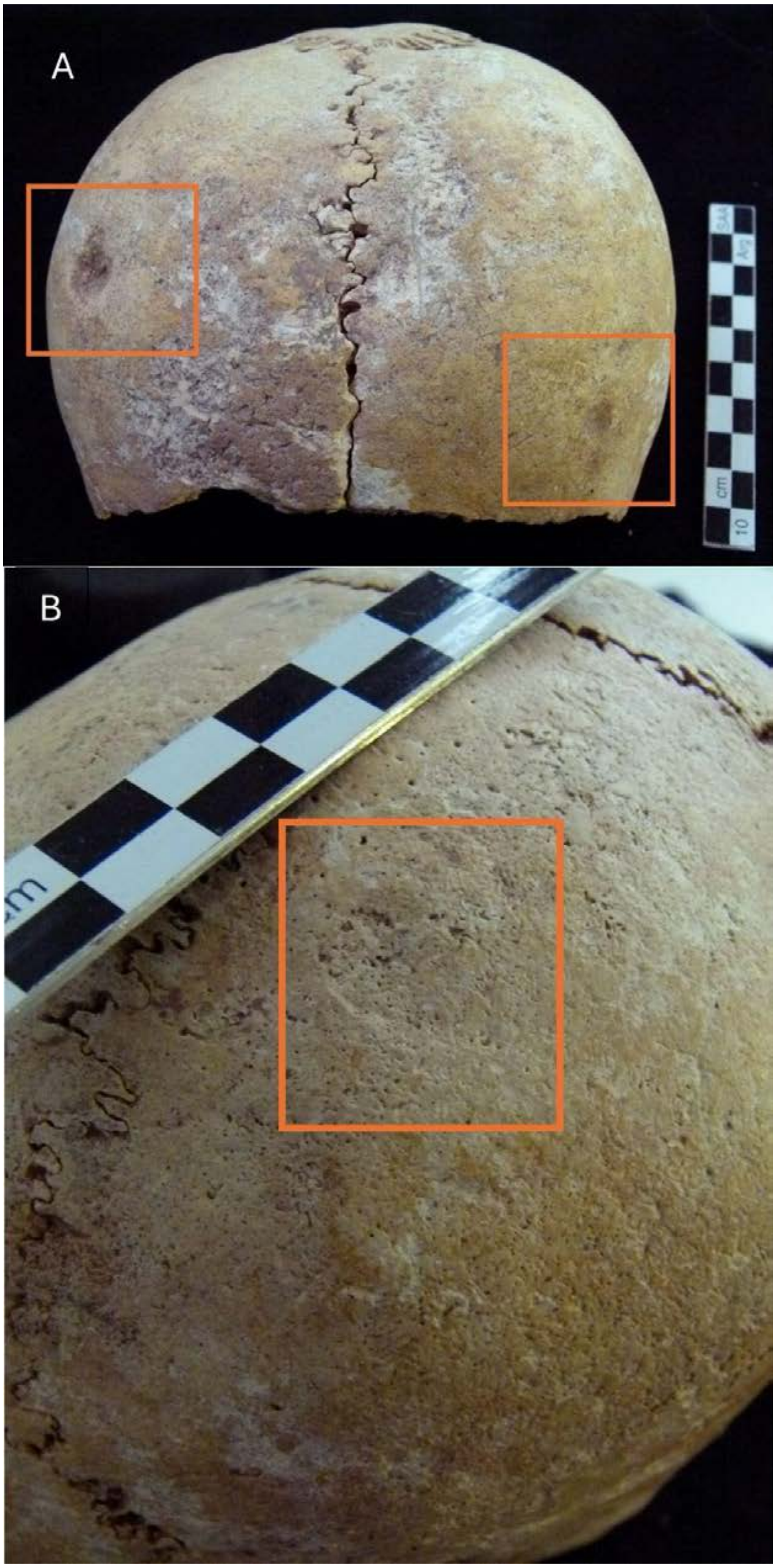

Fig. S11. Evidence of bone trauma at the PLC site.
a. cranium from the adult individual PLC-286 with two compressions of reduced size in both parietals; **b.** cranium from the adult individual PLC-807 (AR0186) with one compression.

D. Isotopic Work

D.1 Strontium isotopes

D.1.a Laboratory procedures

All strontium isotope analytical work was done in the MC-ICP-MS Facility, Dept. of Geological Sciences, University of Cape Town. Powdered samples were weighed into 7 ml Savillex PFA beakers, 2-3 ml of 65% HNO3 added and the closed beakers kept at 140ºC for an hour. Following complete sample dissolution, the beakers were opened, and the samples dried. The samples were taken up in 1.5 ml of 2M HNO_3 in preparation for elemental separation chemistry. The Sr fraction was isolated after methods detailed in[54] using Triskem International Sr.Spec. resin, dried down and re-dissolved in 0.2% HNO_3. Solutions were analyzed as ±200 ppb Sr in 0.2% HNO_3 using a Un Instruments NuPlasma MC-ICP-MS and all $^{87}Sr/^{86}Sr$ data presented are referenced to bracketing analyses of the international strontium isotope standard SRM987, using an $^{87}Sr/^{86}Sr$ value for this standard of 0.710255[55]. Isobaric interference of ^{87}Rb on ^{87}Sr was corrected using the measured signal for ^{85}Rb and the natural $^{85}Rb/^{87}Rb$ ratio, while the effect of instrumental mass fractionation was corrected using the exponential law and an $^{86}Sr/^{88}Sr$ value of 0.1194. Repeat analyses of an in-house carbonate reference material NM95 as unknown during processing of samples for this study yielded an $^{87}Sr/^{86}Sr$ result (0.708912 ± 0.000037; n=17), in agreement with long-term data for this material (0.708911 ± 0.000040 2; n=414; over >8 years). Total procedural strontium blanks were typically better than ~250 pg elemental Sr, and therefore negligible.

D.1.b Diagenesis

The possibility of diagenetic strontium uptake from the soil cannot be completely ruled out, particularly in the bone samples, but we consider that the patterns in the data are not the product of diagenesis. We have conducted a pilot study with Ca/P and U/Ca elemental concentrations[3]. Ca/P values range between 1.9 and 2.1, within the range of values characteristic of modern hydroxyapatite[56,57] and the results obtained for U/Ca are consistent with little diagenetic contamination. This is consistent with general expectations for minimal diagenesis in recent samples from very dry climates, like those presented here[58]. The main isotopic finding presented here, namely that the migrant bone and teeth/petrous samples from the window of migration at PLC produce an almost identical non-local signal, could not be produced by contamination with the widely different strontium from the soil of the depositional environment. The independent genomic and chronometric results presented here are coherent with these geographical assignments, lending further support to the interpretations presented.

The human collagen samples showed good preservation with C/N values within 3.0 and 3.4 (3.2±0.01), %N within 11 and 15.9 (14.7±1) and %C within 31.9 and 45.41 (40.6±2.6) (Supplementary Data S1)[59,60]. Only three samples presented anomalous C/N, %C or %N, and were excluded from the analysis (Supplementary Data S1). Regarding $\delta^{13}C_{ap.}$ values, preservation criteria have not been used extensively and the assumption of reliability is largely based on the extraction protocols. In this research, we assumed that if collagen is well-preserved, then bioapatite is also well-preserved[61,62].

D.1.c Bioavailable strontium isoscape: Random forest approach

We built the isoscape using a global dataset of bioavailable $^{87}Sr/^{86}Sr$ from soil, plant and animal samples across the world[63]. This dataset was updated with additional published $^{87}Sr/^{86}Sr$, particularly those specific to South America, and the 54 local rodent samples analyzed in this study to improve both regional and local prediction. Additional references are listed in Supplementary Data S2. To ensure that the recorded $^{87}Sr/^{86}Sr$ values accurately reflected local environmental conditions, we filtered the dataset to retain only samples from plants, soils, and low-mobility animals. The Random Forest was trained on 12,026 samples from 5,388 sites. Random Forest does not assume any specific data distribution or homoscedasticity and can handle complex relationships within the data[64].

To predict the spatial distribution of bioavailable $^{87}Sr/^{86}Sr$, the Random Forest integrates various lithological, environmental, and climatic variables that likely influence $^{87}Sr/^{86}Sr$ at each sampling site. These variables include bedrock and terrane age, predicted $^{87}Sr/^{86}Sr$ from bedrock[63], and predictor variables at the levels of topography, soil properties, climate, and atmospheric deposition. A complete list of the predictors is provided in Table S3. Using R v4.3[65], we extracted these variables at each $^{87}Sr/^{86}Sr$ sampling site. When a variable was unavailable at a given site due to low raster resolution or coverage, we assigned the nearest available value instead. The Random Forest model was then applied to the $^{87}Sr/^{86}Sr$ dataset and its associated covariates. First, we removed highly correlated variables ($R > 0.9$) and applied a variable selection algorithm specifically designed for Random Forest (VSURF R package)[66] to eliminate irrelevant and redundant predictors. The remaining covariates were used to train the final Random Forest model. We set the number of trees to 3,000 and determined the number of predictors per tree as the total number of covariates divided by three[64]. Model performance was evaluated using root mean squared error (RMSE) and ten-fold cross-validation (repeated five times), while node impurity (a measure of how effectively the trees split the dataset at each node) was used to assess predictor importance. Finally, we applied the model to the covariate rasters to predict $^{87}Sr/^{86}Sr$ values across the landscape and we used a quantile random forest regression to generate a 68.27% prediction interval and estimate the spatial uncertainty (standard deviation) associated with the predictions[67].

The bioavailable $^{87}Sr/^{86}Sr$ isoscape along with the associated spatial uncertainty is displayed in Fig. S12. The Random Forest model demonstrates an RMSE of 0.0034 and an explanatory power of 68% of the variance. The final Random Forest model was constructed using eight predictors (Fig. S13). It includes lithological variables (maximum age of the bedrock, predicted $^{87}Sr/^{86}Sr$ value of the bedrock, terrane age), mean annual temperature, and atmospheric deposition variables (dust, volcanic, black carbon, and fossil fuel deposition). The central and southern regions along the coast demonstrate low $^{87}Sr/^{86}Sr$ values (<0.708), while the northern coastal region rises up to 0.718. A global eastward trend shows an increase in $^{87}Sr/^{86}Sr$ values up to 0.714, reflecting increasing bedrock age[4]. Strontium values from rodents sampled in the study area present a good agreement with the prediction of the isoscape at their sampling site ($n=90$, $R^2=0.93$), highlighting the good predictive power of the isoscape.

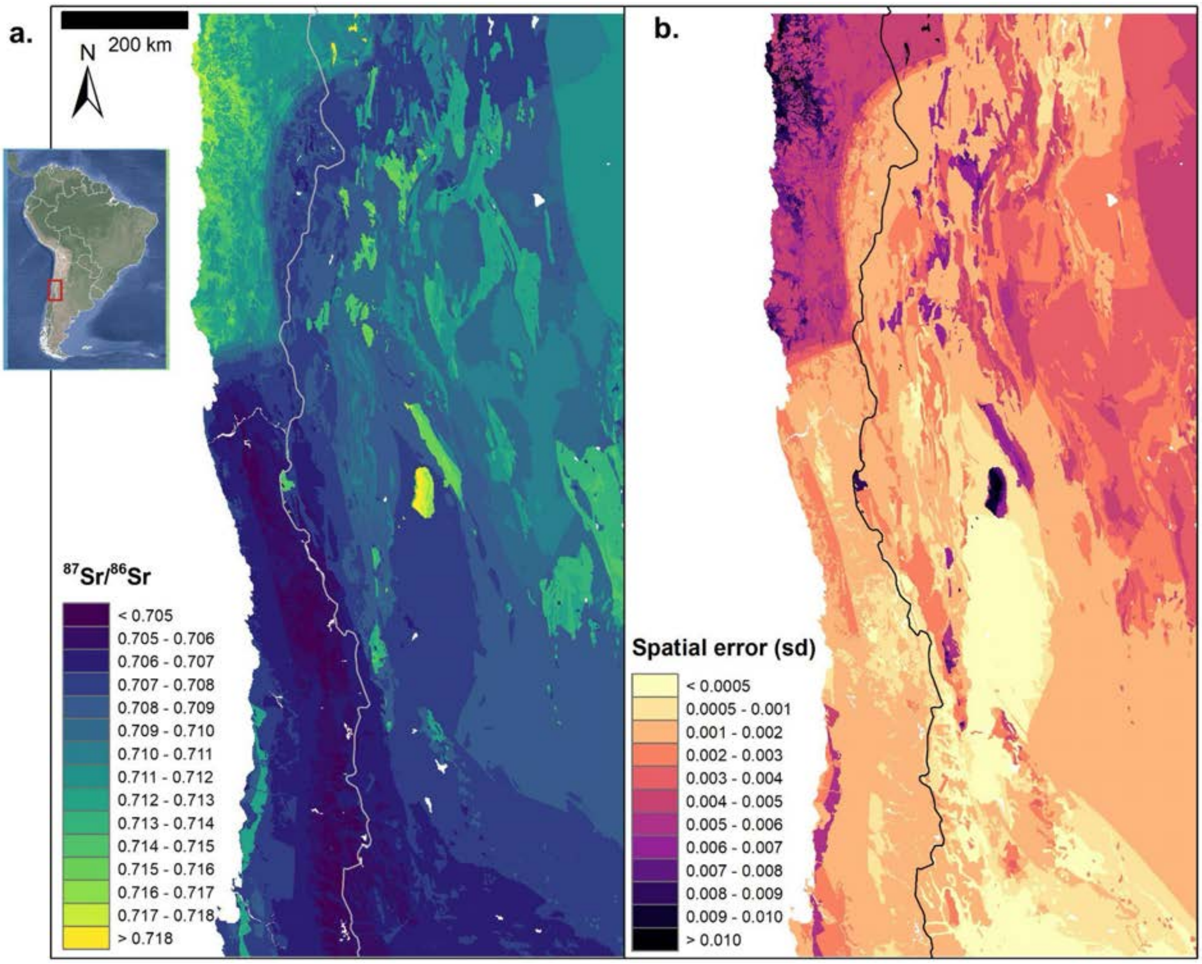

Fig. S12. $^{87}Sr/^{86}Sr$ isoscape for the southern Andes of Argentina and Chile.
a. Bioavailable $^{87}Sr/^{86}Sr$ isoscape for the southern Andes of Argentina and Chile; **b.** Associated spatial error.

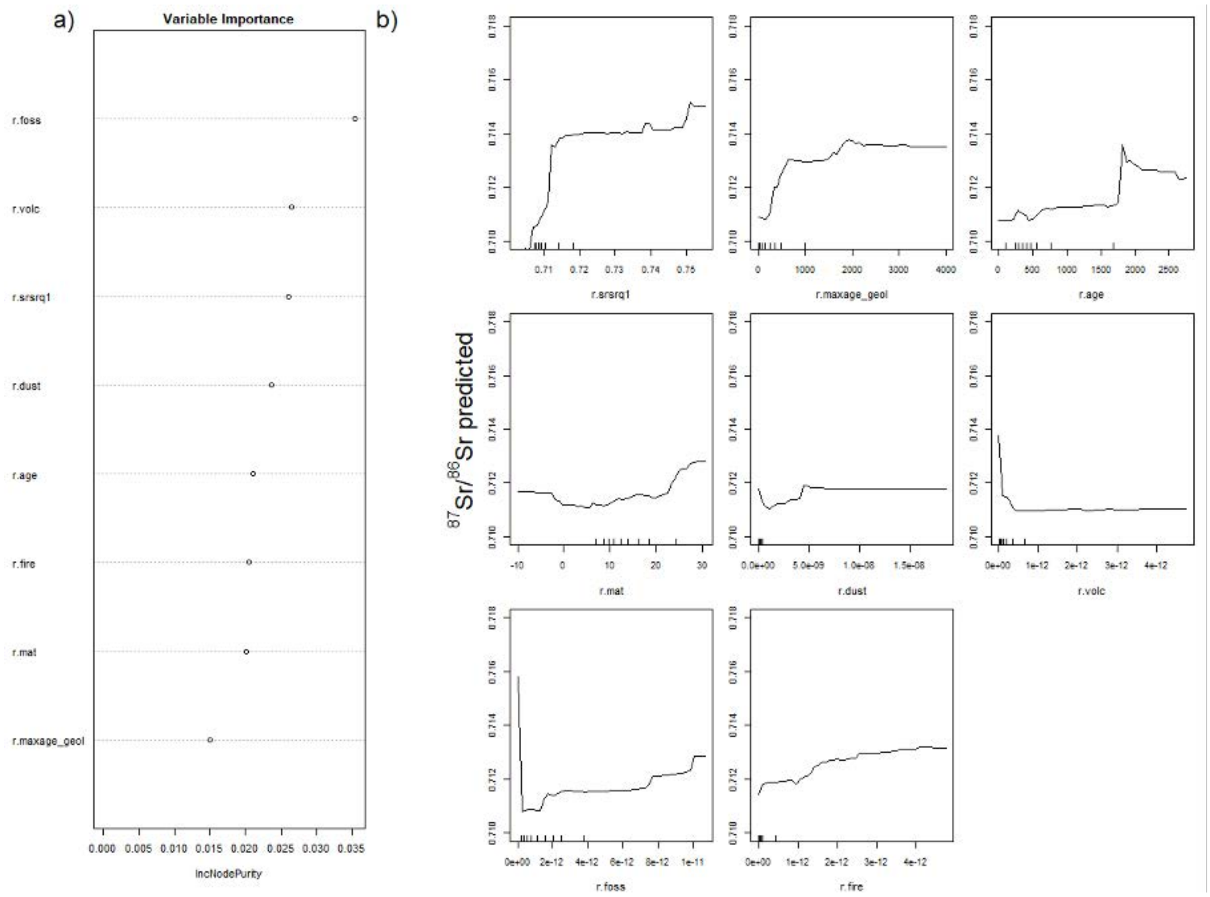

Fig. S13. Random forest regression results.
Variable importance and partial dependence plots of the random forest regression. **a.** Variable importance plot depicting the influence of the different predictors on the predicted bioavailable ^{87}Sr/^{86}Sr. **b.** Partial dependence plots depicting the relationship between the predictors retained for the final model and the predicted bioavailable ^{87}Sr/^{86}Sr. Hash marks on the x-axes correspond to the deciles of the predictor distribution. Description of the predictors is provided in Table S3.

Variables	Description	Source
r.m1	median bedrock model	[63]
r.srsrq1	1st quartile bedrock model	[63]
r.srsrq3	3rd quartile bedrock model	[63]
r.meanage_geol	mean GLiM age (Myrs)	[68]
r.minage_geol	minimal GLiM age (Myrs)	[68]
r.maxage_geol	maximal GLiM age (Myrs)	[68]
r.age	terrane age (Myrs)	[69]
r.mat	mean annual temperature (°C)	[70]
r.map	mean annual precipitation (mm.yrs^{-1})	[70]
r.pet	global potential evapo-transpiration	[71]
r.ai	global aridity index	[71]
r.salt	Simulation of sea salt deposition (g.m^{-2}.yr^{-1})	[72]

r.dust	dust deposition (g.m^{-2}.yr^{-1})	[72]
r.fire	Black carbon deposition (kg.m^{-2}.s^{-1})	[72]
r.foss	Fossil Fuel (kg.m^{-2}.s^{-1})	[72]
r.volc	Volcanic deposition (kg.m^{-2}.s^{-1})	[73]
r.dist	Distance to the coast (km)	[74]
r.elevation	shuttle radar topography mission (m)	[75]
r.bouguer	bouguer anomaly	[76]
r.GUM	global unconsolidated sediment map	[77]
r.cec	cation exchange capacity (mmol(c)/kg)	[78]
r.ph	soil pH (H2O, x10)	[78]
r.phkcl	soil pH (KCl, x10)	[78]
r.clay	clay (g/kg)	[78]
r.ocs	organic carbon stocks (t/ha)	[78]
r.bulk	bulk density (cg.cm-3)	[78]

Table S3. Auxiliary variables used in the random forest analysis.
The bedrock model[63] estimates the distribution of predicted ^{87}Sr/^{86}Sr values within lithological units based on their age and type[79]. GLiM: Global Lithological Map.

D.1.d Geographical assignment of the human remains

Spatial assignment of the human samples was done on the isoscape to assess the origin and document the mobility of the individuals. Bulk ^{87}Sr/^{86}Sr values of each sample were assigned to the isoscape using a Bayesian spatial assignment approach[80]. The Bayesian approach allows for assignment on continuous-probability surfaces inherently accounting for the probabilistic nature of the model and the associated uncertainties. Analyses were done in R, using the assignR package[81].

The set of analyzed samples for the geographic assignment based on the random forest isoscape corresponds to the ‘migration window’ samples of the sites with non-local staples. This includes the samples displaying the signal relevant for the last years of life of each individual (i.e. bones for adults and teeth or petrous bones for subadults). In the case of PLC, providing for 83% of the migrant individuals, this set of samples is strikingly homogeneous in terms of ^{87}Sr/^{86}Sr values, emphasizing that these non-local individuals likely shared the geographic range. The 30 samples from PLC assigned to the ‘migration window’ have a mean ^{87}Sr/^{86}Sr value of 0.707284±0.00011, showing that there is minimum spread around the mean.

We present the assignment of the following two samples selected to account for the observed variation within the migrant population: AR0510 (PLC-471k, petrous bone of a subadult individual) with a ^{87}Sr/^{86}Sr value of 0.7072341, which is indistinguishable from the mean value for the migration set (Main Fig. 4c). Here we also present the assignment of sample AR0514 (PLC-LPD#2) corresponding to another migrant adult with a slightly more radiogenic signal (^{87}Sr/^{86}Sr=0.7073646) (Fig. S14). The assignments of these two individuals are largely coincident.

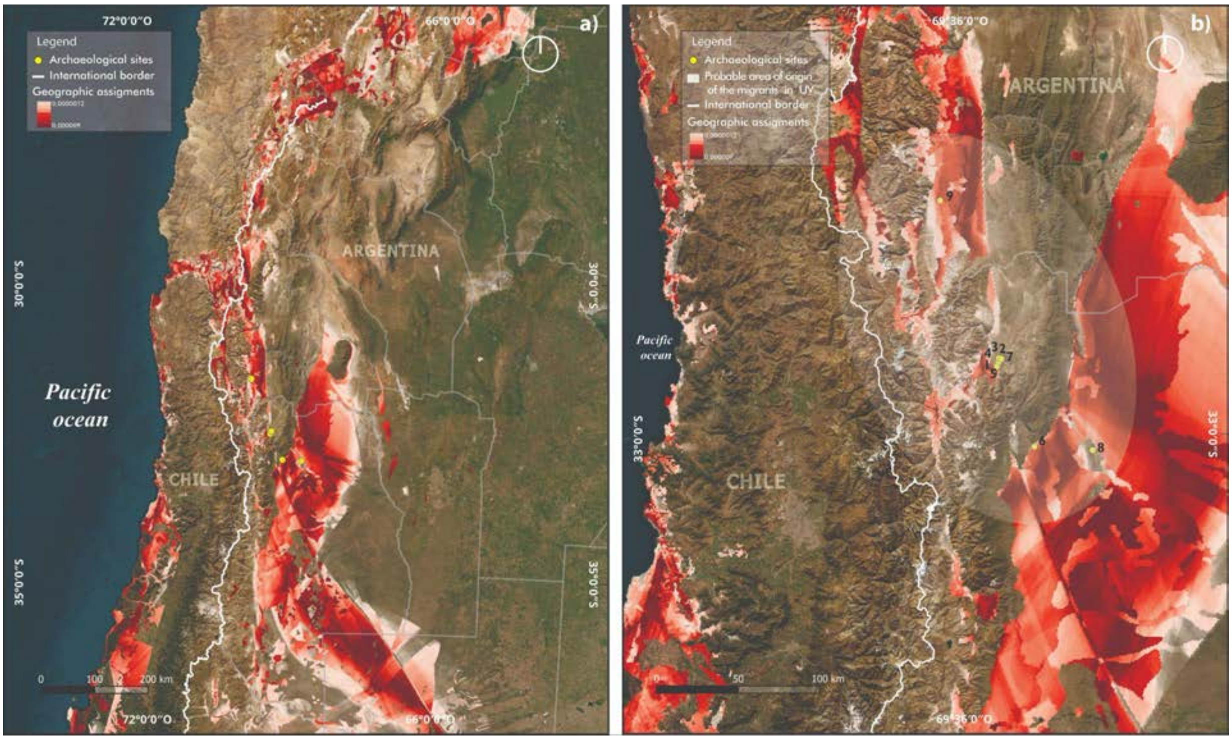

Fig. S14. Geographical assignment of individual AR0514 from random forest isoscape Geographical assignment of the individual AR0514 (PLC) based on the random forest isoscape for the southern Andes. **a.** Sample assigned at the largest scale provided by the isoscape; **b.** Zoom into a regional scale.

D.2 Light isotopes (carbon and nitrogen)

D.2.a Laboratory procedures

Bone collagen was prepared using the pseudomorph method described in[82]. For bone collagen analyses, each sample was mechanically cleaned and weighed to enable determination of the collagen yield. The samples were then soaked overnight at room temperature in a defatting solution of chloroform, methanol and water (2.0:1.0:0.8 v:v). Most lipids were removed by this solution and floated to the top of the liquid. The defatting process was repeated if large quantities of lipids were observed. All samples were then rinsed with distilled water. Samples were then soaked in dilute HCl (0.2M) at room temperature to remove the apatite from the bone. The acid was changed every day for between five to ten days, depending on the size of the sample and the density of the bone. This resulted in a translucent, flexible pseudomorph of (mostly) collagen of the same shape and size as the original bone fragment. Each pseudomorph was soaked in 0.1M NaOH overnight to remove base-soluble contaminants such as humic acids and any remaining lipids[83].

Next, the samples were soaked in distilled water, changed daily until its pH remained neutral. The sample was then freeze-dried and then allowed to equilibrate with the atmosphere (overnight) before being weighed to obtain the mass of collagen extracted. This is expressed as a percentage of the initial weight of the bone sample to give collagen yield. For each sample, approximately 0.5 mg of extracted collagen was weighed into a tin cup. The samples were combusted in a Flash EA 2000 series elemental analyzer (Thermo Finnigan, Milan, Italy). The gases were passed through to a Delta Plus V IRMS (Thermo Electron, Bremen, Germany), via a Conflo IV gas control unit (Thermo Finnigan, Bremen, Germany). Standards used were in-house standards Chocolate ($\delta^{13}C$= -17.75‰; $\delta^{15}N$= 4.31‰); New MG ($\delta^{13}C$= -20.94‰; $\delta^{15}N$=

4.7‰; Sucrose ($\delta^{13}C$= -10.6‰); Valine ($\delta^{13}C$= -26.8‰; $\delta^{15}N$= 12.14‰). Each standard had been calibrated against international standard materials NBS 21, IAEA N1 and N2 and standards exchanged with other laboratories. The reproducibility of repeated measurements of standard materials was ≤0.2‰.

In order to extract enamel and bone apatite, a hand-held rotary drill fitted with a 1.2 mm diameter diamond-tipped dental drill bit was used to collect approximately 10 mg of enamel powder. The tooth was drilled from the occlusal surface to the cervix to collect enamel representative of the entire period of crown formation in order to average out seasonal variations. Bone fragments were ground in a liquid nitrogen-cooled SPEX mill and sieved through 180 and 106 µm mesh. Only the 179-107 µm fraction was analyzed[83]. The enamel was treated following the established method for tooth enamel pre-treatment[84]. Samples were treated with 1.75% sodium hydroxide for 45 minutes, rinsed three times with distilled water, and then reacted with 0.1M acetic acid for 15 minutes. Finally, they were rinsed three times with distilled water, freeze-dried[85]. The bone powder was treated with 1.75% sodium hypochlorite for 3 hours, rinsed three times with distilled water, then reacted with 0.1M acetic acid for 15 minutes. Samples were rinsed again and finally freeze dried. Approximately 2 mg of each apatite powder (both enamel and bone apatite) were reacted with 100% phosphoric acid in a ThermoFinnigan (Germany) model II gas bench and the resultant CO_2 gas passed into a ThermoFinnigan Delta Plus XP isotope ratio mass spectrometer. Precision was monitored by repeated analysis of international standard materials (IAEA-CO1 and NBS18) in each run. $^{13}C/^{12}C$ ratios were reported in the standard notation relative to the PeeDee Belemnite (PDB) standard in parts per mil (‰). Precision of repeated analyses of standard materials was <0.2‰ for $\delta^{13}C$.

Finally, using the formula below, the percentage C_4 was calculated in accordance with[86] as follows: $\%C_4$= $[(25\ (\delta^{13}C_{ap.}\ 9.7))/15] \times 100$, with 25 representing the ‰ C_3 end-member, $\delta^{13}C_{apatite}$ representing the observed isotopic ratio, 9.7 representing $\Delta^{13}C_{apatite\text{-}diet}$, and 15 being the expected difference between the C_3 and C_4 dietary end-members. We present the descriptive statistics of the % of C_4 input per site in Table S4.

Site	PLC	T-III	T-I	BB6	T-II	BRI	US2	A.Av.
N	50	7	7	17	8	6	1	1
Mean	67.56	63.81	29.64	27.11	56.26	41.65	49.1	55.16
Standard deviation	5.91	8.93	9.47	7.01	12.59	15.76	-	-
Standard error	0.83	3.37	3.57	1.7	4.45	6.43	-	-
Min	52.3	56.2	20.1	17.3	32.7	28.5	-	-
Max	79.1	79.3	46.7	42.5	72.5	72.8	-	-

Table S4. Descriptive statistics of the percentual C_4 dietary input per site.

D.2.b. Comparative analysis of Andean paleodiets

A bibliographic compilation of Central Andean and Central-South Andean human samples was done in order to produce Main Fig. 3b. Samples from Peru, Bolivia, Chile, and Argentina with adequate isotope data and sample sizes were included, particularly the availability of $\delta^{13}C$ in bioapatite as a proxy of the total diet[87]. This compilation contextualizes the values of UV and, in particular, PLC site, within the Andean stable isotopic record. The data is compiled in Table S5.

N	Site	Country	N	$\delta^{13}C_{ap.}$	SD	N	$\delta^{13}C_{col.}$	SD	Source
1	Conchopata	Peru	19	-3.9	1.0	10	-10.7	1.2	[88]
2	Maraynitoq	Peru	2	-2.3	0.8	70	-9.7	1.0	[89]
3	Trigopampa	Peru	4	-3.2	0.5	11	-11.1	1.0	[89]
4	Ancón LMH	Peru	9	-5.1	0.8	9	-10.2	0.5	[90]
5	Beringa MH	Peru	8	-11.9	1.7	10	-14.9	1.1	[91]
6	SARNC Inland LIP	Chile	3	-7.7	0.8	8	-12.5	1.8	[92]
7	Aconcagua valley	Chile	19	-6.6	0.8	19	-11.6	0.8	[93]
8	Los Amarillos	Argentina	11	-4.9	0.5	11	-10.5	0.5	[94]
9	Rio Atuel 1	Argentina	4	-8.3	0.8	4	-14.2	0.9	[30]
10	PLC	Argentina	34	-4.9	0.9	34	-12.4	1.6	[29], Present study
11	T-I	Argentina	7	-10.9	1.4	7	-16.1	1.3	[29], Present study
12	T-II	Argentina	8	-6.8	1.8	12	-12.8	1.7	[29]
13	T-III	Argentina	8	-5.8	1.3	8	-11.8	0.7	[29]
14	Paredones	Peru	9	-4.3	1.4	6	-10.8	1.0	[95]
15	Tablada Lurin 1	Peru	10	-5.7	1.5	5	-11.2	1.6	[96]
16	Viru-Gallinazo	Peru	4	-6.3	0.8	6	-11.8	1.6	[97]
17	Pica 8 LIP	Chile	33	-5.6	2.4	34	-10.4	2.1	[98]
18	Machu Picchu	Peru	27	-6.6	3.3	13	-12.6	2.6	[99,100]
19	Río Muerto M43a	Peru	12	-7.6	1.1	13	-12.6	1.1	[86]
20	Río Muerto M70a	Peru	20	-7.1	1.8	20	-11.9	1.7	[86]
21	BB6	Argentina	18	-11.3	1.1	18	-18.0	0.6	[3,30], Present study
22	Titicaca-Tiwanaku	Bolivia	65	-8.7	2.9	65	-14.9	3.0	[101]

23	BRI	Argentina	6	-9.1	2.4	6	-14.7	1.6	29

Table S5. Stable isotope results for sites along the South Central and Southern Andes.

E. Archaeogenetic work

E.1. Laboratory procedures for ancient DNA work

E.1.a. Sampling, DNA extraction and library preparation

Sampling, DNA extraction and library preparation were conducted in three different ancient DNA facilities (Supplementary Data S3) located at the Globe Institute, University of Copenhagen, Danmark; the Center of Anthropology and Genomics of Toulouse (CAGT), France; and the Institute of Genomics, University of Tartu, Estonia, following established workflows designed to reduce the risk of contamination.

Protocols followed at the Globe Institute, University of Copenhagen

Sampling at the Globe Institute

Before sampling for DNA extraction, we decontaminated the remains by wiping them with a cloth dipped in 5% sodium hypochlorite solution. Next, we conducted a visual assessment of the preservation status, using the guidelines set forth in reference[102] to classify the samples as either 'poor' or 'good' quality. The teeth were sliced transversally at the cemento-enamel junction using a Dremel diamond-coated cutting disc as in reference[103] and 50-100 mg of cementum was crushed to a coarse powder using a pestle and a mortar.

The temporal bones were decontaminated using the same procedures as the teeth to remove any superficial environmental contamination. To maximize endogenous DNA recovery, the innermost region, the pars petrosa, was isolated following sampling protocols similar to those described in reference[104]. Once extracted, the pars petrosa was ground into 50–100 mg of bone powder using a sterilized mortar and pestle. The resulting powder was then collected in a 5 mL Eppendorf tube for subsequent aDNA extraction.

DNA extraction at the Globe Institute

All samples were extracted using a modified silica-based method adapted from established protocols[105–107]. To enhance endogenous aDNA recovery, the process began with a pre-digestion step. Bone powder (50-150 mg) was combined with 4.5 mL of 0.45 M EDTA (ThermoFisher, 15575020), 62.5 µl of 0.25 mg/ml proteinase K, and 437.5 µl of UV-sterilized molecular-grade water. The mixture was incubated at 37°C for 20 minutes, after which it was centrifuged at 2000 rpm for 1 minute to discard the supernatant, leaving behind the bone pellet. The pellet was then resuspended in 4 mL of fresh digestion buffer and incubated overnight at 37°C on a rotor for thorough breakdown. On the second day, the digested material underwent DNA binding. The supernatant was transferred to 40 mL of binding buffer (BB1), which contained 5 M guanidine hydrochloride (Gu-HCl), 100 mM sodium acetate (NaOAc), 20 mM sodium chloride (NaCl), and 30% isopropanol. To facilitate DNA adhesion, 50 µl of a pre-mixed silicon dioxide (SiO2) suspension was added to the solution. Since the soil chemistry and burial environment influence the chemical properties of each sample, pH adjustments were performed to maintain levels between 4 and 5 using pH indicator strips and 30% hydrochloric acid (HCl). The samples were then incubated at room temperature in darkness for 1 hour. Once incubation was complete, centrifugation at 2000 rpm for 5 minutes was conducted, after which the supernatant was carefully removed, ensuring that Gu-HCl waste was disposed of in accordance with safety guidelines. The silica pellets were subjected to a series of washing steps. First, they were resuspended in 800 µl of chilled 80% ethanol (VWR, 437433T) and then transferred into 1.5 mL

DNA Lo-bind tubes. To maximize recovery, any remaining pellet material was collected by rinsing the original falcon tubes with an additional 200 µl of 80% ethanol. The samples were then centrifuged at 2000 rpm for 5 minutes, and the ethanol supernatant was discarded. This washing step was repeated, followed by two additional centrifugation rounds at 10,000 rpm for 1 minute each, with the tubes rotated 180 degrees between spins to enhance drying. For elution, the dried pellets were resuspended in 60 µl of elution buffer (EB) and incubated at 37°C on a shaker for 10 minutes. Afterward, the eluate underwent two rounds of centrifugation at 12,000 rpm for 1 minute, with a tube rotation in between. The final DNA extract was transferred to a fresh 1.5 mL Lo-bind tube, and DNA concentration was quantified using the Qubit dsDNA HS Assay Kit (ThermoFisher, Q32854).

To reduce uracil-related DNA damage, a partial uracil glycosylase (UDG) treatment[108] was applied. Each 20 µl DNA sample was treated with 3 µl of USER Enzyme (2 U/µl), along with 6 µl of Tango buffer (10×), 0.24 µl of dNTPs (25 mM), and 0.6 µl of ATP (100 mM). The reaction was maintained at 37°C for 30 minutes. To halt the uracil excision, 3 µl of UGI (2 U/µl) was subsequently added to the reaction mixture.

Single-stranded Library preparation at the Globe Institute

For the preparation of single-stranded libraries, we followed the approach outlined by[109], which optimizes the efficient recovery of aDNA. In summary, after purifying and eluting the USER-treated aDNA, the ligation step was carried out using diluted P5 and P7 oligos along with splints, adjusted according to the quantified DNA from the extracts. This method effectively reduces the formation of adapter dimers and other artifacts. Following adapter ligation, qPCR analysis was conducted to determine the appropriate number of cycles required for each indexing PCR reaction. During indexing PCR, 21 µl of each non-amplified aDNA library was combined with 4 µl of 10 µM IDT dual unique index pairs. The reaction was performed using KAPA HiFi HotStart ReadyMix (Roche Diagnostics, cat KK2801) under the following thermal cycling conditions: an initial denaturation at 98°C for 45 seconds (one cycle), followed by a variable number of cycles of denaturation, annealing, and elongation, optimized based on Ct values from qPCR, at 98°C for 30 seconds, 65°C for 30 seconds, and 72°C for 30 seconds, respectively. A final elongation step at 72°C for one minute was included for a single cycle. Notably, while the indexing reaction was set up in a cleanroom, the subsequent indexing and post-amplification steps were conducted in a separate facility exclusively designated for handling amplified DNA. Finally, the amplified libraries were purified and pooled for shotgun sequencing.

Protocols followed at the CAGT

Sampling at the CAGT

All laboratory work was processed in the ancient CAGT laboratory (CNRS UMR 5288/Université Paul Sabatier) of Toulouse in order to avoid any modern and cross contaminations. All material used during those experiments were UV-ved beforehand. We cut root apex after brushing gently the surface with a clean disc using a dremel tool. Chunks were powdered with the Mixer Mill MM200 (retsch) Micro-dismembrator and then stored in a 2 mL sterile tube for DNA extraction.

DNA extraction at the CAGT

The extraction method was performed using Seguin-Orlando et al., 2021 protocol[110], inspired by Gamba et al., 2016[111], with some modifications for pathogen DNA recovery. We used 60 to 85 mg of teeth powdered and added 963 µL of lysis buffer (0,5M EDTA, 0,25mg/mL proteinase K and 0,5% sodium lauroyl sarcosinate) in 2mL tubes and started first digestion step for 1h at 37°C on Thermo Scientific Tube Revolver (Dutscher). After centrifugation at 8000 rpm for 1 min, the supernatant was stored and a fresh buffer was added again to digest powder left during overnight digestion at 42°C. Next day, we took 200 µL of each solution for DNA purification with Qlagen@ kit. We mixed it with 1,25 mL of PB and centrifuged at 8000 rpm for 1 min before releasing the liquid. This was done three times in order to pass all the mix on the column. Then, we added 600 µL of PE for washing, we centrifuged at the same speed and repeated a second time at 13000 rpm to dry the column. DNA present on the column was eluted with 26 µL of Elution Buffer supplemented with 0,05% Tween 20 (EBT). We used 22,8 µL of DNA extracted to perform a USER® (NEB®) treatment during 3 hours at 37°C as mentioned in previous studies[111]. For some samples that required more DNA free for new library building in deep sequencing approach, we used the lysate left which was stored at -20°C. The purification step was performed with High Pure Viral Nucleic Acid Kit, provided by ROCHE@, allowing to purify high volumes. We used columns and reagents (9.5 mL of PB and 600 µL of PE per sample) from the Qiagen kit in order to be in the same conditions as before and eluted with 46 µL of EBT.

Double-stranded library preparation at the CAGT

Library construction was performed in the same conditions as extraction in the ancient laboratory of CAGT. We used the triple-indexed double-stranded library method as described in Rohland et al., 2015[108] and modified later by Fages et al., 2019[112]. A unique couple of P5 and P7 barcodes is involved during the ligation step, leading to their junction in 3' regions of DNA fragments. The addition of a 6bp Illumina index[113] during PCR amplification leads to the construction of an indexed and barcoded library ready for sequencing. PCRs runs were performed using 2 µL of barcoded library in a reaction volume of 25 µL, containing 17,9µL of H_2O, 0.4 µL of AccuPrime™ Pfx DNA polymerase (one unit), 2.5 µL 10X Accuprime™ Pfx reaction mix, 1µL of BSA (20mg/µL), 0,2µL of inPE1 primer (25 µM) and 1µL of indexing primer (5 µM). It is important to note that if the PCR mix was prepared in the ancient lab, the reaction occurred in another building where only "modern DNA" is manipulated. Purification steps using Agencourt Ampure XP beads were made by pooling two amplifications together with a pre-established level of beads (1:1.2 or 1:1.4 as DNA:beads ratio), then PCR products were eluted with 15 µL of EBT. DNA fragment length and concentration were determined with the Invitrogen and Agilent technologies, respectively Qubit HS dsDNA assay and Tapestation 4200 instrument using a High sensitivity D1000 ScreenTapeAssay.

Protocols followed at the Institute of Genomics, University of Tartu

Sampling at the Institute of Genomics

All pre-PCR work was performed in the dedicated ancient DNA facilities of the Institute of Genomics, University of Tartu, following stringent decontamination protocols to prevent contamination. Sampling of skeletal remains, including petrous bones and the malleus (e.g., sample AR0507_1.i709.i514), was performed using a blade disc under sterile conditions. Samples were processed following a standardized decontamination procedure involving immersion in 6% sodium hypochlorite (NaOCl), multiple MilliQ water rinses, and 70% ethanol treatment. The samples were air-dried under UV exposure for at least two hours in a dedicated laminar flow hood before DNA extraction.

DNA extraction at the Institute of Genomics

DNA was extracted using a high-volume silica-based method optimized for ultra-short aDNA fragments, based on protocols modified from ref[105] and previously published by Keller and colleagues. Approximately 50–100 mg of powdered bone was incubated in a lysis buffer containing 0.5 M EDTA, proteinase K (0.25 mg/mL), and 0.5% sodium lauroyl sarcosinate. A two-step digestion was performed: an initial 1-hour incubation at 37°C, followed by an overnight digestion at 42°C. Extracts were purified using silica spin columns (Qiagen or Roche), with final elution in 50 µL of EB buffer supplemented with 0.05% Tween-20. A USER (uracil-Specific Excision Reagent) treatment was applied for partial UDG treatment to reduce deamination damage while preserving terminal damage patterns for authentication.

Double-stranded library preparation at the Institute of Genomics

Double-stranded libraries were prepared following a double-indexing protocol optimized for ancient DNA, as described in ref[113], with subsequent modifications tailored for ultra-short DNA fragments. Library construction was carried out in a dedicated cleanroom using a half-UDG treatment to reduce deamination-induced errors while preserving terminal damage for authentication. Unique dual indices were incorporated during PCR amplification using

AccuPrime™ Pfx DNA polymerase. Indexing reactions were conducted in a separate post-PCR facility to prevent contamination. Final libraries were purified using Agencourt AMPure XP beads and quantified with the Qubit dsDNA HS Assay Kit (ThermoFisher). Fragment length distributions were assessed using an Agilent TapeStation 4200 system with the High Sensitivity D1000 ScreenTape Assay.

E.1.b Capture enrichment for MTBC DNA

We enriched *Mycobacterium tuberculosis* complex (MTBC) DNA from two samples, AR0185 and AR0187, using a custom RNA-based probe set from myBaits Custom Kit (Arbor Biosciences) which includes Lineage 4 strains and members of the animal adapted group (*M. pinnipedii*, *M. microti*, *M. africanum*, *M. canettii, M. caprae*). Probes were designed as 60 bp long with 4x tiling density and masking of low complexity regions.

To start, we split each indexed, adapter-ligated library into two reactions and amplified them using the KAPA HiFi HotStart ReadyMix to a minimum of 200 ng per ul. We then used 7 ul of the amplified products for capture to have two capture reactions for AR0185 and one capture reaction for AR0187. Given the ancient nature of the samples we followed the High Sensitivity Protocol (v5.02, March 2022)[114] optimized for degraded and ancient DNA with some adaptations for our research. We carried out hybridization for 20 hours at 60° Celsius, adhering to the protocol recommended for bait-target sequence divergence between 15-20%. Following hybridization, we amplified the purified enriched libraries using KAPA HiFi HotStart[115], running 16 cycles. The amplified product was purified using QIAgen PCR Purification Kit (QIAGEN) resulting in a final elution volume of 25ul. The final elution was split with 15ul to be sequenced. The remaining 10ul of purified capture product was concentrated to 7ul and used for a second round of capture. All steps were repeated as before with exception of the final amplification which was amended to 10 cycles.

Finally, we purified the enriched libraries using MinElute columns and quantified the final product using a Qubit 4 Fluorometer and the Qubit dsDNA Broad Range Assay Kit (Thermo Scientific) and assessed fragment size distribution using a Fragment Analyzer (Agilent Technologies) to confirm successful enrichment and suitable fragment profiles for sequencing.

E.1.c Next-generation sequencing

For the first round of screening for human DNA and pathogens, we performed a low depth (<20M reads) shotgun sequencing of pooled libraries prepared at the three aDNA facilities. Sequencing was done through the Biomics platform of Institut Pasteur on either an Illumina NovaSeq 6000 150PE from the Novogene, Cambridge, UK or an Illumina NovaSeq X 150PE from the ICM Institute in Paris, France. Libraries with high human DNA content (>5%) and targeted TB-enriched libraries were then shotgun sequenced at high depth on either the same machine or an Illumina NovaSeq X 150PE machine.

E.2. Sequence data processing

E.2.a Adapter removal, mapping and deduplication (steps 1 - 4)

To process FASTQ files for human DNA analyses, the following steps were performed using the nf-core/eager workflow[116] v2.4.2.

1. For the case of human data:
Adapter removal[117] v2.3.2 was run with parameters: --qualitymax 41 --collapse --preserve5p --trimns --trimqualities --adapter-list list_adapt_for_removal_paired.txt --minlength 30 --minquality

2	--minadapteroverlap	1.

For the case of pathogen screening and TB analyses:
Adapter removal v2.3.2 was run with these parameters: --collapse --minlength 30 --mm 3 --trimns --trimqualities --minquality 2 --qualitybase 33

2. The collapsed reads thus obtained were aligned with Burrows-Wheeler Aligner[118] v0.7.17-r1188 against GRCH37 Human Reference Genome (downloaded from UCSC Genome Browser; https://hgdownload.soe.ucsc.edu/goldenPath/hg19/bigZips/hg19.fa.gz), with the aln function (BWA-ALN) and parameters -n 0.01 -l 1024 -k 2 -o 2. Subsequently, alignments were generated given single-end reads using *samse* function. To recover mitochondrial DNA (mtDNA) data, the reads were aligned leveraging circularmapper[119] against the revised Cambridge Reference Sequence[120].

3. The aligned reads were filtered using SAMtools[121] v1.12 with parameters -F4 -q 25.

4. The reads were then deduplicated using Dedup[119] v0.12.8 with default parameters.

E.2.b. Indels realignment and trimming reads ends (steps 5 - 7)

5. We realigned the Indels with *RealignerTargetCreator* and *IndelRealigner* from GenomeAnalysisTK[122] v3.6, and MD tags were subsequently generated using *calmd* from SAMtools.
6. For each library, we manually assigned a number of base pairs to be trimmed on 5' and 3' ends of the reads (Supplementary Data S3) by visually inspecting the damage plots generated with damageprofiler[123] v0.4.9.
7. Using these values, trimming was performed with *trimBam* from bamUtil[124] v1.0.15.

E.2.c. Contamination estimates (step 8)

We obtained contamination estimations with 5 methods:

8.1. Schmutzi[125] v1.5.6 was applied on the BAM files for mtDNA obtained at step 5, with either setting the number of bp to be considered to 3 for half-UDG and 7 for fully UDG treated libraries (Schmutzi_auto; not calculated for fully UDG treated libraires), or with a customized number for each library being the maximum of bp to be trimmed in 5' and 3' reads ends determined at step 6 (Supplementary Data S3).

8.2. To run ContamMix[126] v1.0-10, we first performed the following procedure from the BAM files for mtDNA obtained at step 5. We rescaled the BAM files with mapdamage2[127] with the length to be considered in 5' and 3' reads ends as determined at step 6. We then converted the BAM files to consensus sequence fasta files applying successively *bamtofastq* function in bedTools[128] v2.29.2 mia[129] v1.0 and ma. We then ran MAFFT[130] on this sequence together with the sequences of 311 contaminants provided[126] to generate a FASTA file used as input for –malnFn parameter in ContamMix. We finally used the consensus fasta to remap the reads with BWA, thus obtaining a bam file used as input for --samFn parameter in ContamMix.

8.3. Haplocheck[131] v1.3.3 was run on the VCF files obtained with mtDNA Variant Detection[132] v2.0.0-rc13 from BAM files for mtDNA generated at step 7.

8.4. Nuclear Contamination with ANGSD[133] v0.935 was applied only for Males through nf-core/eager workflow.

8.5. ContamLD[134] was leveraged to obtain nuclear contamination from BAM files obtained at step 6, using the 1240K SNP panel and comparing the individual of interest to the 1000 Genomes

Project[135] population with which it shares most genetic affinity, as estimated through f_3-outgroup(Individual, 1000 Genomes Population; Mbuti) with Admixtools[136] v6.

Contamination estimates of the libraries produced from the individuals with Depth of Coverage greater than 0.01X –i.e., those included in our analyses (Supplementary Data S3)– did not show any significant contamination. Nuclear Contamination, ContamLD and Schmutzi estimates were below 3% for all libraries. Five libraries (AR0504.1, AR0504_1i710507, AR0505.1, AR0512_1i708511) had ContamMix estimates over 3%, but below 6%. Contamination values were obtained by Haplocheck for only two libraries (AR0505.1 and AR0512_1i708511), which showed ~7%. Altogether, contamination estimates across five methods were below 3% for most libraries, and we did not detect any library with consistent contamination estimates greater than 5% with more than one method.

E.2.d. Generating a unique alignment file for each individual (step 9)

9. When several libraries were generated for the same individual, the alignment files were obtained merging the alignment files for those using the *merge* function from SAMtools.

E.2.e. Genotyping nuclear genome (step 10)

10. From BAM files obtained for nuclear DNA generated at step 7, genotypes were called with pileupCaller (https://github. com/stschiff/sequenceTools) v.1.4.0.5 after applying *mpileup* from SAMtools with q and Q parameters set to 30, and disabling disable per-Base Alignment Quality (-B). We called genotypes for two autosomal SNP panels using all positions or only transversions (TVs): 1240K[137] (1.15M SNPs, 0.22M TVs) and Shotgun[134] (SG; 5.6 M SNPs; 1.82M TVs).

E.2.f. Generating consensus mtDNA sequence (step 11)

11. To generate mitogenomes, we applied ANGSD with options -GL 1 -doGIF 2 -doMajorMinor 1 -doFasta 2 -doCounts 1. The FASTA thus obtained were then manually corrected after visual inspection of the corresponding BAM files with Integrative Genomics Viewer[138] 2.3.34.

E.3. Patterns of ancient DNA degradation

All libraries exhibited ancient DNA damage patterns. Reads length distribution presents a peak below 100bp for all libraries, and below 50bp for most of them (Fig. S15). For double stranded libraries that were non fully UDG treated, we also observed an excess of C-to-T and G-to-A substitutions in 3’ and 5’ ends of reads, respectively, while for single-stranded, we observed an excess of C-to-T substitutions at both ends of the reads (Fig. S16).

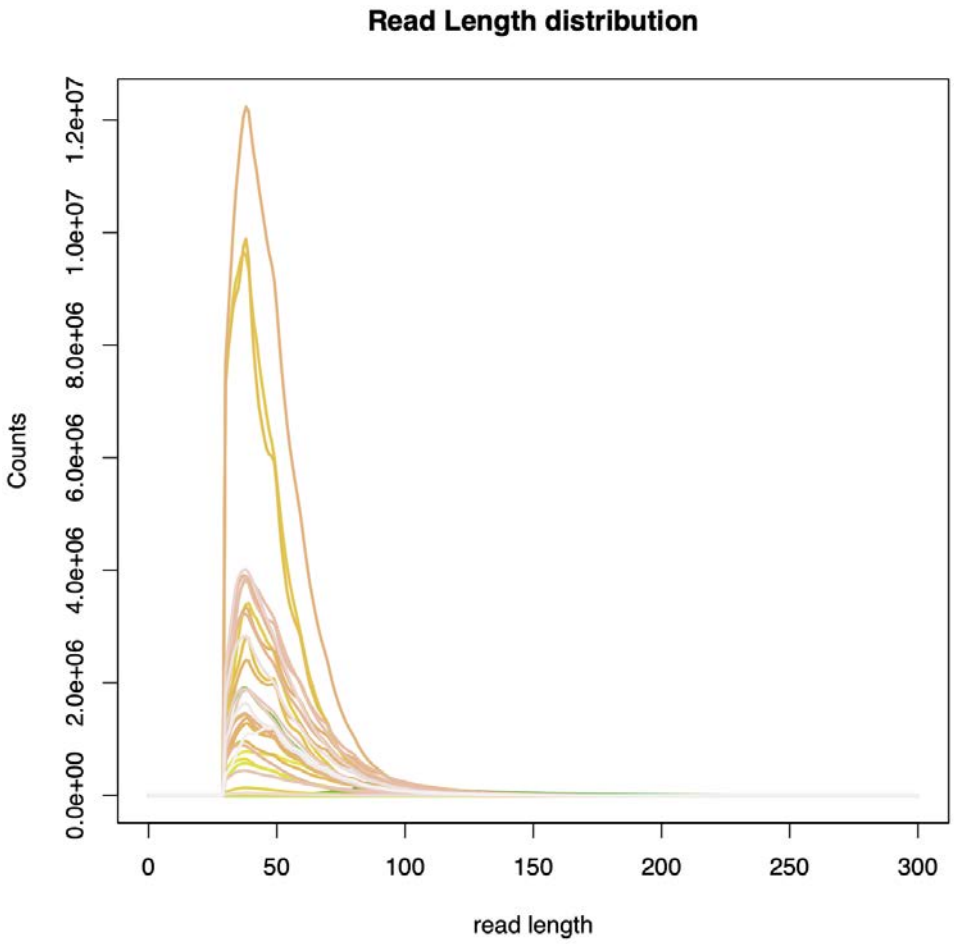

Fig. S15. Distribution of sequencing read lengths.
Frequency (y--axis) with respect to length of the deduplicated collapsed mapped reads in number of nucleotides position (x-axis) (one line per library).

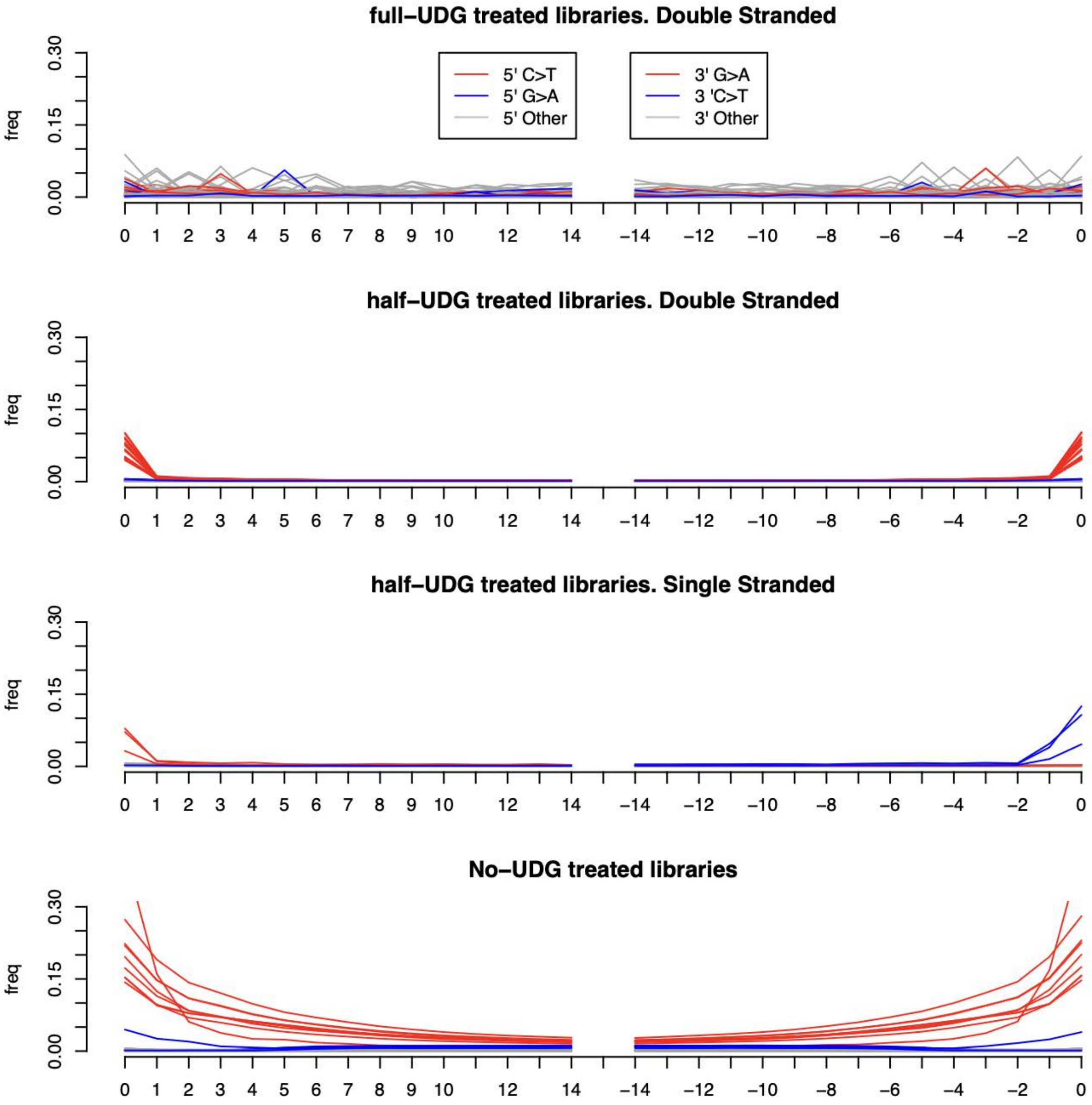

Fig. S16. Misincorporation patterns in sequencing reads.
Misincorporation frequency (y--axis) with respect to the position (x-axis) from the 5' and 3' ends of the collapsed reads (three lines –blue, red, and grey, for G-to-A, C-to-T, and other substitutions, respectively– per library). Libraries are shown according to the UDG treatment and strandness, from top to bottom: full UDG treatment/double stranded, half-UDG treatment/double stranded, half-UDG treatment/single stranded, no UDG treatment/double stranded.

E.4. Reference Panels from published nuclear genomic data

E.4.a. Publicly available ancient NGS data

Following the same procedure described in Supplementary Section D.2., we reprocessed 332 ancient individuals[139–158]. We used the raw FASTQ files when available, however most available sequencing data consists of the aligned reads, and in this case, we converted first BAM files into FASTQ.

E.4.b. Publicly available modern NGS data

We included Next-Generation sequencing whole genome data for 206 individuals[153,159–161] in order to analyze UV ancient genomes in a broader geographical context, and to approximate genetic variation in the populations thought to have contributed to present-day South Americans. This consisted in (i) the Chane from Argentina (n=1), the Piapoco (n=6) from Columbia, the Aymara (n=1), Matzes (n=12), Moches (n=28), Quechua (n=3) from Peru, the Huichol (n=1); Mayan (n=21), Mixe (n=3), Mixtec (n=2), Pima (n=13) and Zapotec (n=2) from Mexico, and the Karitiana (n=11), Surui (n=8) and Yukpa (n=1) from Brazil; (ii) individuals from Chopccas (n=26), Cusco (n=16), Iquitos (n=16), Trujillo (n=16), and Uros (n=9) in Peru ; (iii) and 10 Mbuti Hunter-Gatherers from Central Africa (n=10). The 206 accessed genomes were sequenced at a depth of coverage of ~10-60×.

Genotype calling was performed jointly on all published genomes, which were all accessed as FASTQ files. FASTQ files were checked for integrity using the FASTQC software[162] and multiQC[163] v.1.9. FASTQ files were converted to unmapped BAM files (uBAM), read groups were added and Illumina adapters were marked with Picard Tools version 2.20.1 (https://broadinstitute.github.io/picard). Read pairs were mapped to the hs37d5 human reference genome[164] using the Burrows-Wheeler Aligner[118] (BWA-MEM) v. 0.7.17 and duplicates were marked with Picard. We computed the per-sample depth of sequencing coverage (DP) and the number of reads mapping to the human reference genome using SAMtools[121] v.1.10. We estimated the percentage of DNA contamination using VerifyBamID2[165]. Samples with a DNA contamination proportion (FREEMIX estimate) >3% were excluded. For the remaining samples, the mapping quality of sequencing reads was capped with capmq (https://github.com/mcshane/capmq), considering the proportion of DNA contamination estimated by VerifyBamID2.

Variant calling was performed following the GATK Best Practices recommendations (https://software.broadinstitute.org/gatk/best-practices/) using GATK v.4.1.2.0. All samples were genotyped individually using HaplotypeCaller and gVCF files were combined with CombineGVCFs. Multi-sample joint genotyping was performed using the GenotypeGVCFs tool with the option -allSites, to create a multisample VCF that includes invariant sites. We split the multisample VCF into two VCFs: one with only invariant sites and one with only autosomal variant sites. Variant sites were first filtered using GATK 'VQSR' with a truth sensitivity cut-off of 99.5 ('--ts_filter_level 99.5') for SNPs and 99 for insertions-deletions (indel). Using BCFtools v.1.11 (http://www.htslib.org/), we set as missing all genotypes of variant or invariant positions with (i) a DP < sample mean(DP) – sample SD(DP) or (ii) a DP > sample mean(DP) + sample SD(DP), and (iii) a genotype quality (GQ/RGQ) < 30 [166]. Variants with genotype call rate < 95% were then discarded. Hardy-Weinberg equilibrium was tested with RUTH, which accounts for population structure[167], and we discarded variants showing a P-value < 10^{-4}, considering the four first PCs. To identify and exclude cryptically related individuals, kinship coefficients were estimated using KING[168] v.2.2.7. To maximize sample size, we iteratively excluded the individual who showed second-degree relatedness or higher with the largest number of individuals, until no pairs of related individuals remain. In the case of ties, we excluded the individual with the lowest depth of sequencing coverage.

E.4.c. Publicly available Modern Genotyping Data

We included genotyping data for 713 individuals[157,169–171]. The masked data to focus only on positions for which both alleles are assigned to genetic ancestry Indigenous to the Americas for individuals genotyped in[172] were obtained from the same reference, while those for [157,169,171] were retrieved from[169].

E.4.d. Reference Panels leveraged in each analysis

Using the ancient and modern genotype data thus retrieved from the literature, combined with the newly generated data, we constructed different reference panels as described in Supplementary Data S4. The individuals from the reference panels were grouped mostly according to the metadata associated with the Allen Ancient DNA Resource[173] v54, but also consulting the groupings used in subsequent articles[139,142,143,146,149,151,155,156,158] (Supplementary Data S4; Figs. S17-18). We differentiated between groups whose individuals were sequenced through Shotgun (SG) and in-solution capture (Capt). Groups were subsequently assigned to major geographical ranges, and divided according to period: Early Holocene (EH; before 8,500 BP), Middle Holocene (MH; between 8,500 and 3,500 cal years BP), Late Holocene (LH; after 3,500 cal years BP, and before European invasion ~500 cal years BP), and Modern.

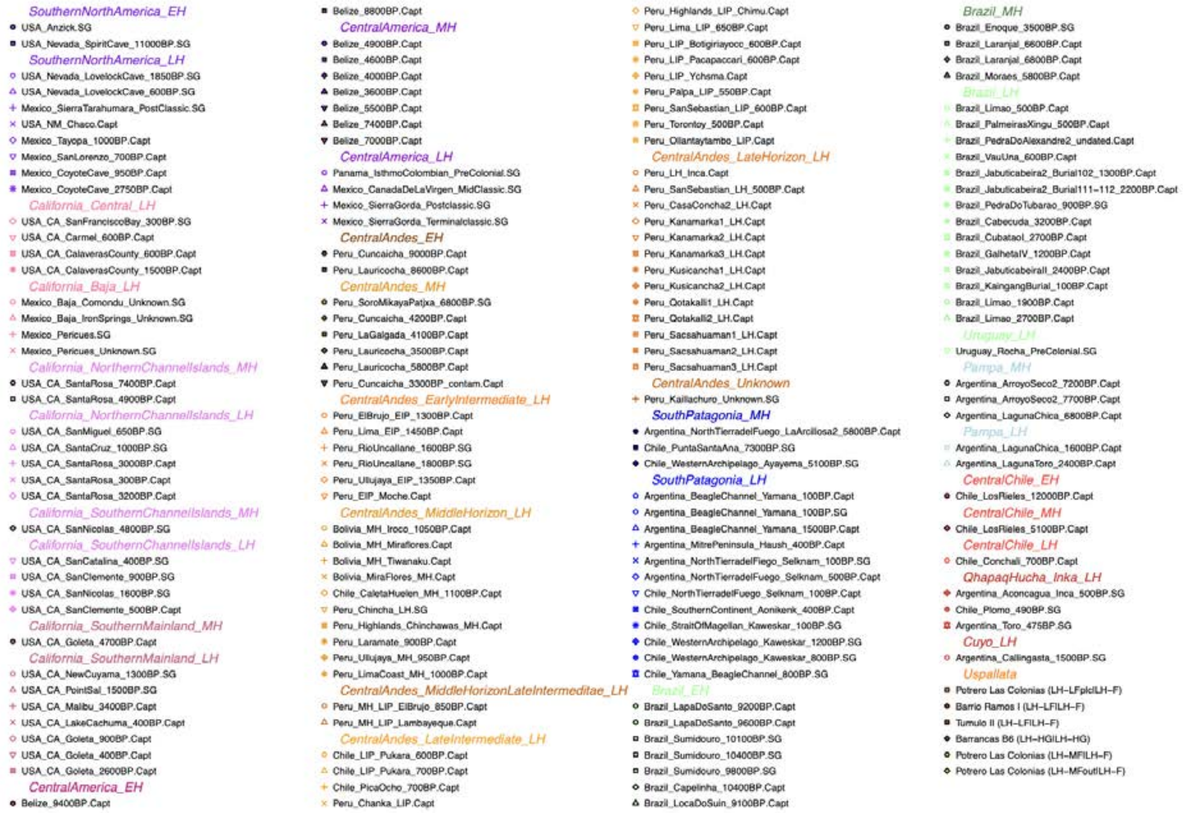

Fig. S17. Ancient groups included in the reference panel for paleogenomics analyses. Representation of published and newly sequenced genomes included in genetic structure analyses**.** Groups were assigned to a major group according to the geographic region and period. SG and Capt: for groups whose individuals were sequenced through Shotgun and, in-solution capture, respectively. EH, MH, LH: Early, Middle and Late Holocene, respectively.

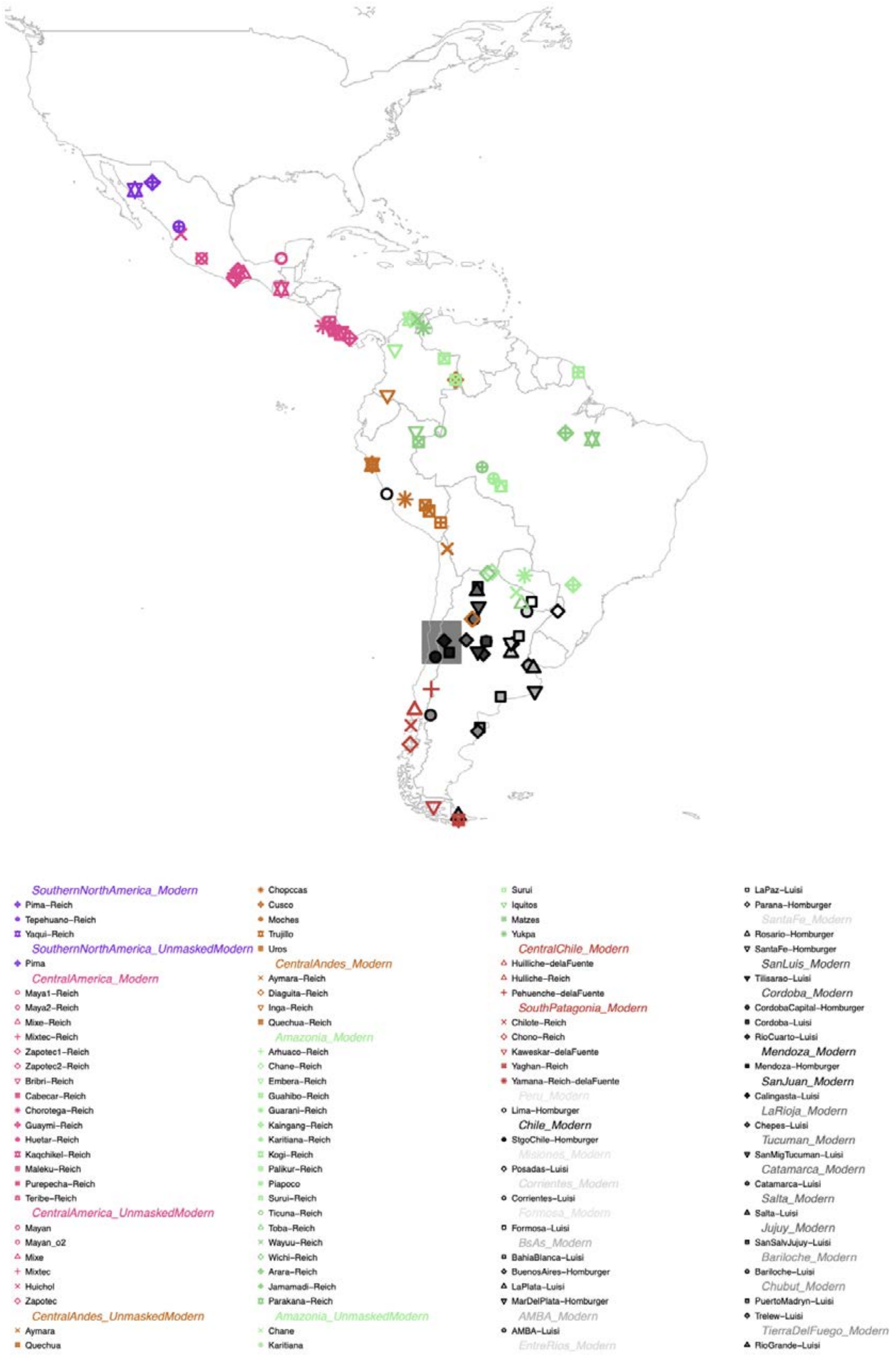

Fig. S18. Modern groups included in the reference panel for paleogenomics analyses. Geographical location of the published modern individuals included in genetic structure analyses. Names including Reich, Luisi, de la Fuente or Homburger indicate individuals whose genotyping data was reported in [157,170,171,174], and for which we used masked genotype data for genetic ancestry Indigenous to the Americas. All the other modern groups include individuals whose

NGS sequencing data from the whole-genome was used (data from [153,159–161] reprocessed by us). Groups whose individuals do not self-identify as Indigenous are represented in grey shades, while groups represented in other colors were reported as Indigenous in the original publications.

E.5. Genomic sex determination

We compared the ratio between the Depth of Coverage (DoC) observed for sexual chromosomes to DoC on autosomal chromosomes. Hereafter, we name those ratios ratX and ratY, for chromosome X and Y, respectively. We assigned the genomic sex applying the following rules:

- Male, when ratX was comprised between 0.35 and 0.65, and ratY was greater than 0.35
- Female, when ratX was greater than 0.65 and ratY was lower than 0.35
- Undefined, otherwise

Genomic sex was thus inferred for each library separately (using BAM files for nuclear DNA obtained at step 7 in Supplementary Material Section E.2.), as well for each individual (BAM files obtained at step 9 in Supplementary Material Section E.2. when the DoC observed for the corresponding BAM file was greater than 0.0001X (Fig. S19; Supplementary Data S3)

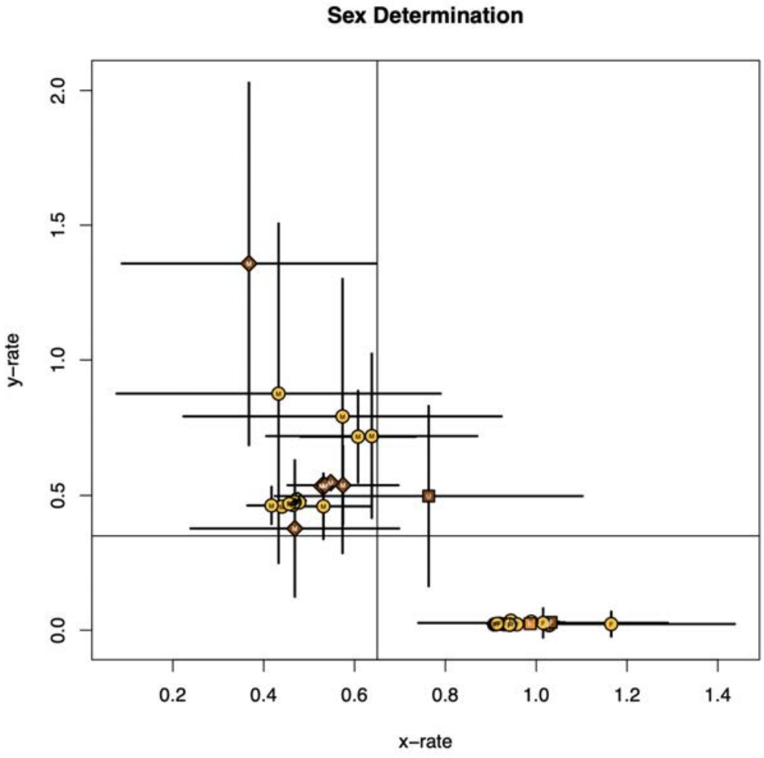

Fig. S19. Genomic Sex Determination for UV and BB6 individuals.
The ratio of coverage on chromosome Y (chromosome X) over the coverage on autosomal chromosomes is plotted on the y-axis (x-axis), with segments representing ± 2 standard errors. Points are colored according to the group assigned, with the inner letter according to genomic sex assigned (F for Female, M for Male and U for Undefined). 44 (out of 46) individuals with Depth of Coverage greater than 0.00001X are shown.

E.6 Uniparental markers analyses

E.6.a Methods for uniparental markers analyses

Maternal haplogroup assignment and mtDNA phylogeny

We used mtPhyl v 4.015 (2013; https://sites.google.com/site/mtphyl/home) with, as input, FASTA files of the consensus mtDNA sequences (obtained at step 11 of Supplementary Material

Section E.2.) to build maximum parsimony trees combining 38 new mitogenomes generated in this work and all closely related published mitogenomes from South America, as listed in Table S6, including all the sub-haplogroups or tree branches within A2, B2, C1b, and C1c. For this analysis hypervariable position 16519 and length variations at positions 303–315, 514–525, and 16181–16193 were removed. The sub-haplogroups in the trees were finally annotated with reference to HaploGrep[132] v2 and PhyloTree[175], and previously published trees[176–178]. We assigned new sub-haplogroups (B2ak1b3, B2b16a and C1b37) when they encompassed a minimum of three different haplotypes sharing at least one stable mutation not recurrent in the tree. The trees are shown in Fig. S20, and haplogroup assignment in Table S6.

We performed a Bayesian analysis using BEAST[179] v1.10.4 to reconstruct past population dynamics. The analysis included 24 newly generated mitogenomes that reached the quality requirements described above and derived from unrelated individuals up to the 2nd-degree (see Supplementary Material Section E.7.). Multiple sequence alignment of the data set was constructed with the online service of MAFFT[130] using the progressive FFT-NS-2 setting. We trimmed the alignments by excluding the position and length variations mentioned above. BEAST was run with the following parameters: HKY evolution model, substitution rate reported in Soares et al. (2009)[180], and strict clock model. The latter taking into account the nature of human mtDNA evolution, as similar generation times and DNA repair mechanisms are expected[181]. As calibration points, we used the radiocarbon dates of ancient individuals when available, and the median observed at the archaeological site otherwise (dates as shown in Main Fig. 4b). We considered a generation time of 25 years. BEAST was then run with 100 million states, each with a sampling frequency of 10,000. Finally, Tracer[182] v1.7.2 was used to ensure convergence of the MCMC chains and an Estimated Sample Size (ESS). We thus obtained a population size estimate over time for each sampled iteration (Main Fig. 5c). Finally, we reconstructed the consensus phylogenetic tree with TreeAnnotator[179], using default parameters, and plotted it using *ape* package[183] in R (Fig. S21).

Paternal haplogroup assignment

To assess paternal haplogroups to the 17 individuals assigned to male genomic sex, yLeaf[184], was run with parameters *-r 1 -q 30 -b 90 -t 1 -pos WGS_hg19.txt*. 16 paternal haplogroups were thus inferred, and we report its equivalence in ISOGG nomenclature (https://docs.google.com/spreadsheets/d/1UY26FvLE3UmEmYFiXgOy0uezJi_wOut-V5TD0a_6-bE/edit#gid=1934392066) following ref.[185] (Table S7).

E.6.a. Supplementary results from uniparental markers analyses

Links with other American groups inferred from mitogenomes

To establish the maternal lineage links between UV individuals and the rest of the Americas, we used an in-house database of modern and ancient mitogenomes from the Americas. At present, this database consists of more than 4.000 samples and encompasses most of the regions of the American continent.

After an exhaustive revision of all these publicly available mitogenomes, it was possible to identify a clear link to Central Argentina and LH Pampa regions, revealed by the topology within B2ak sub-lineage, a clade previously described by refs[176,177] for individuals from those regions. The recent description of individuals carrying B2ak haplogroup in Northern Patagonia[178] allows us to expand geographically this link (Tamburrini, 2024). However, these mitogenomes could not be considered in the tree because they have not been published yet. UV individuals nest in the internal branches of B2ak, and exhibit great internal genetic diversity, thus suggesting a link with a great temporal depth in a broad geographical range, with local differentiation of UV populations (Fig. S20).

A link between UV, CA and LH La Pampa is also observed within the sub-lineages B2b16, pre-A2ay, and C1b37. B2b16 and pre-A2ay encompass each a unique UV individual and several individuals from CA and La Pampa regions. On the contrary, C1b37, a numerically important clade (n=28), is exclusively represented by UV individuals, except for the presence of a single CA individual. This shared pattern for B2b16, pre-A2ay, and C1b37 can be interpreted as gene flows among UV, CA and LH La Pampa, rather than a link with great temporal depth –as described for B2ak.

The rare Q-Z780 American paternal haplogroup is observed at high frequency in UV

We observed nine individuals (eight LH-F migrants and 1 LH-HG) carrying the Q-Z780 paternal haplogroup. One LH-HG and one LH-F migrant are basals, while six LH-F migrants are within the sub-haplogroup Q-Z781 (Table S7). This is a striking result in light of the low frequency of Q-Z780 in the Americas[185]. This founder lineage, with a temporal depth estimated to 19,300 BP, is, despite its low prevalence, widely distributed in the continent as it is reported in Central America, Mexico, Colombia, Peru, Bolivia, Brazil, Paraguay, and Argentina[185]. The basal Q-Z781 is the most represented sub-lineage of Q-Z780 with a divergence time also estimated to 19,300 BP[185]. While its basal state is only reported in Mexican individuals in ref.[185], its sub-branch Q-YP937, characteristic of South America, was dated to 18,700 cal years BP[185]. The persistence of this rare haplogroup in UV could thus reflect a deep temporal divergence of UV populations combined with genetic drift in the founder population.

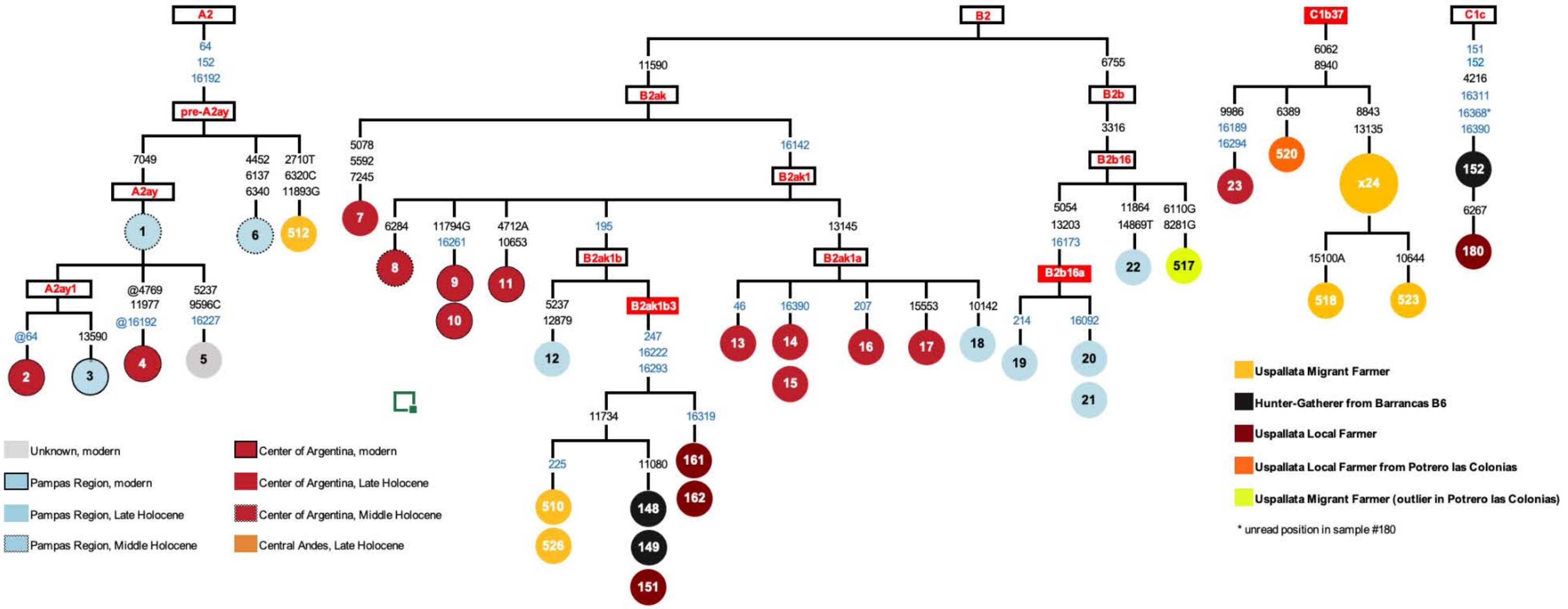

Fig. S20. Maximum parsimony trees from mitogenomes in a South American context.

a. The Maximum parsimony trees for 38 mitogenomes newly sequenced, along with already published mitogenomes from South American individuals. Filled color corresponds to a geographical region, and border type to a time period. Mutations are shown on the branches; they are transitions unless the base change is explicitly indicated. The prefix @ indicates the reversion of a mutation occurring earlier in the phylogeny. Recurrent mutations in the phylogeny are underlined. Sub-haplogroups defined for the first time in this study (C1b37 and B2b16a) are highlighted with a red fill. Meta-data concerning the mitogenomes analyzed are provided in Table S6.

Node #	GenBank/sample ID	Haplogroup	Archaeological Site / Region / Period	Ref.
1	OP605624	A2ay	Laguna Toro / Pampas Region / Middle Holocene	176
2	MW886172	A2ay1	Santa Fe / Central Argentina / Modern	177
3	KF548612	A2ay1	Buenos Aires / Pampas Region / Modern	DS
4	MW886192	A2ay	Santiago del Estero / Central Argentina / Modern	177
5	LS43	A2ay	Unknown / America / Modern	186
6	B2S19	A2ay (pre)	Arroyo Seco 2 / Pampas Region / Middle Holocene	187
7	MW886104	B2ak	Loma Bola / Central Argentina / Late Holocene	177
8	MW886116	B2ak1	Alpa Corral / Central Argentina / Middle Holocene	177
9	MW886180	B2ak1	Santa Rosa de Conlara / Central Argentina / Modern	177
10	MW886210	B2ak1	Villa Dolores / Central Argentina / Modern	177
11	MW886135	B2ak1	Jovita / Central Argentina / Modern	177
12	OP605622	B2ak1	El Guanaco 1 / Pampas Region / Late Holocene	176
13	MW886103	B2ak1a	El Vado / Central Argentina / Late Holocene	177
14	MW886112	B2ak1a	Santa Rosa / Central Argentina / Late Holocene	177
15	MW886114	B2ak1a	Santa Rosa / Central Argentina / Late Holocene	177
16	MW886117	B2ak1a	Potrero de Garay / Central Argentina / Late Holocene	177
17	MW886123	B2ak1a	Rosca Yaco / Central Argentina / Late Holocene	177
18	OP605619	B2ak1a	El Guanaco 1 / Pampas Region / Late Holocene	176
19	OP605617	B2b16a	Mar Chiquita 1 / Pampas Region / Late Holocene	176
20	OP605620	B2b16a	El Guanaco 1 / Pampas Region / Late Holocene	176
21	OP605621	B2b16a	El Guanaco 1 / Pampas Region / Late Holocene	176
22	SC50_L761	B2b16	Laguna Chica / Pampas region / Late Holocene	188
23	MW886115	C1b37	Ayampitín / Central Argentina / Late Holocene	177
512	AR0512	pre-A2ay	Uspallata Migrant Farmer	PS
510	AR0510	B2ak1b	Uspallata Migrant Farmer	PS
526	AR0526	B2ak1b	Uspallata Migrant Farmer	PS
148	AR0148	B2ak1b	Hunter-Gatherer from Barrancas B6	PS
149	AR0149	B2ak1b	Hunter-Gatherer from Barrancas B6	PS
151	AR0151	B2ak1b	Uspallata Local Farmer	PS

Node #	GenBank/sample ID	Haplogroup	Archaeological Site / Region / Period	Ref.
161	AR0161	B2ak1b	Uspallata Local Farmer	PS
162	AR0162	B2ak1b	Uspallata Local Farmer	PS
517	AR0517	B2b16	Uspallata Migrant Farmer (outlier in PLC)	PS
520	AR0520	C1b37	Uspallata Local Farmer from PLC	PS
x24	AR0185; AR0191; AR0192; AR0194; AR0195; AR0500and513; AR0501; AR0503; AR0504; AR0505; AR0506; AR0507; AR0508; AR0509; AR0511; AR0514; AR0515; AR0516; AR0519; AR0521; AR0522; AR0524; AR0527; AR0528	C1b37	Uspallata Migrant Farmer	PS
518	AR0518	C1b37	Uspallata Migrant Farmer	PS
523	AR0523	C1b37	Uspallata Migrant Farmer	PS
152	AR0152	C1c12	Hunter-Gatherer from Barrancas B6	PS
180	AR0180	C1c12	Uspallata Local Farmer	PS
NC	AR0502	C1b	Uspallata Migrant Farmer	PS

Table S6. Mitogenomes used Maximum Parsimony phylogenetic analyses.

Node # refers to the number associated with each mitogenome in Fig. S20. DS: Direct Submission to GenBank by Greenspan B. et al. on the 28th of July 2013 (Family Tree DNA - Genealogy by Genetics). PS: Present Study. NC: Not Considered in Fig. S20 for its insufficient quality, but for which haplogroup could be estimated with Haplogrep[132] v2.

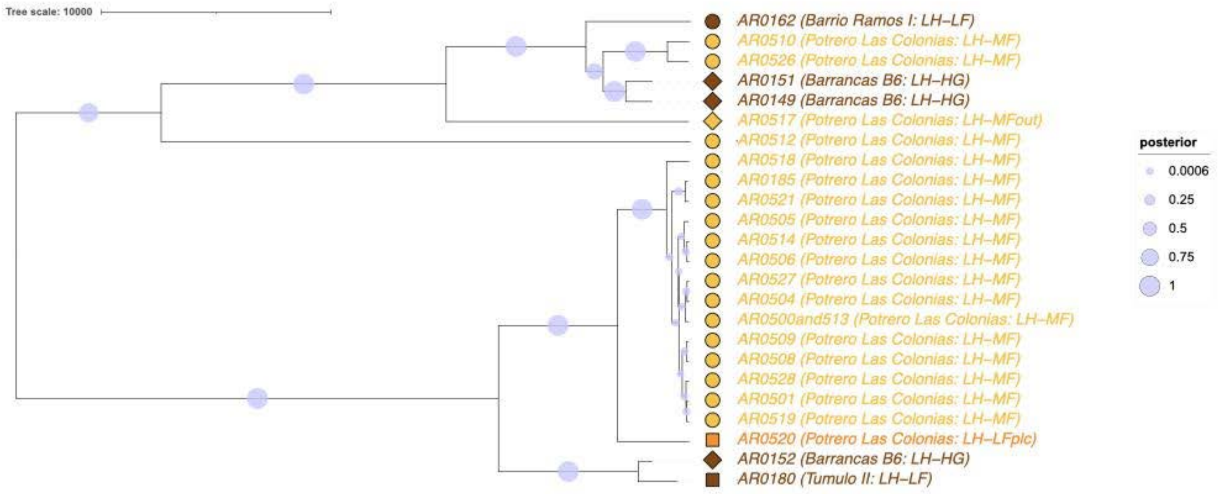

Fig. S21. Consensus bayesian phylogenetic tree from mitogenomes in UV and BB6.
Only unrelated up to the 2nd-degree individuals from the present study were considered. Light blue circles size represent branches' posterior probabilities. The branch lengths are provided in years. Leaves represent mitogenomes from individuals encoded according to the group assigned. LH-HG: Hunter-Gatherer; LH-MF: Migrant Farmer; LH-Mout: Migrant Farmer outlier, LH-LF: Local Farmer; LH-LFplc: Local Farmer from Potrero Las Colonias.

Group	Individual	DoC	Haplogroup Name	Alternative Name
LH-HG	AR0149	2.74027	Q1b1a1a1	Q-M3 -> Q-M848
LH-HG	AR0151	4.72093	Q1b1a1a	Q-M3
LH-HG	AR0152	1.73892	Q1b1a2~	Q-Z780
LH-MF	AR0501	3.14702	Q1b1a2	Q-Z780
LH-MF	AR0505	17.5587	Q1b1a2a3~	Q-Z780 -> Q-Z781 -> Q-YP837
LH-MF	AR0509	11.2554	Q1b1a2~	Q-Z780
LH-MF	AR0510	13.9287	Q1b1a1a1	Q-M3 -> Q-M848
LH-MF	AR0511	7.91074	Q1b1a2a3~	Q-Z780 -> Q-Z781 -> Q-YP837
LH-MF	AR0512	18.8519	Q1b1a2a3~	Q-Z780 -> Q-Z781 -> Q-YP837
LH-MF	AR0514	6.6656	Q1b1a2a~	Q-Z780 -> Q-Z781
LH-MFout	AR0517	3.57415	Q1b1a1a1	Q-M3 -> Q-M848
LH-MF	AR0518	14.7137	Q1b1a2a3~	Q-Z780 -> Q-Z781 -> Q-YP837
LH-MF	AR0519	14.0409	Q1b1a1a	Q-M3
LH-MF	AR0522	7.78054	Q1b1a2a~	Q-Z780 -> Q-Z781
LH-MF	AR0524	0.166013	NA	NA
LH-MF	AR0526	11.1729	Q1b1a1a1	Q-M3 -> Q-M848
LH-MF	AR0527	7.54687	Q1b1a1a1	Q-M3 -> Q-M848

Table S7. Paternal haplogroup assignment.

DoC: Depth of Coverage on chromosome Y; LH-MF: Late Holocene Migrant Farmer; LH-LF: Late Holocene Local Farmer; LH-MFout: Late Holocene Migrant Farmer Outlier.

Genetic continuity between LH-HG and LH-F observed from uniparental lineage diversity.

In both the consensus phylogenetic tree including only UV unrelated individuals (Fig. S21) and the Maximum Parsimony tree which includes all UV individuals as well as Southern Cone individuals (Fig. S20), both estimated from full-length mitogenomes, we observe that two LH-F are located within the same clades than the three LH-HG, within C1c and B2ak1 haplogroups. Moreover, we observed both Q-M3 and Q-Z70 paternal haplogroups in LH-HG (2 and 1 individuals, respectively) and LH-F (6 and 9 individuals, respectively). Altogether, the presence of the rare[185] Q-Z780 paternal lineage in LH-HG and LH-F, and shared maternal haplogroups in both groups would be consistent with the scenario of genetic continuity between hunter-gatherers and farmers in UV.

Uniparental haplogroup diversity suggest matrilocality

The LH-F migrant group is the only one with enough male individuals to compare the paternal and maternal haplogroup diversity. We observe that this group holds an uneven uniparental lineage composition (Tables S6-7). Indeed, while we observed 27 (out of 30) individuals carrying the same maternal lineage (newly defined C1b32), eight and four individuals carry Q-M3 and Q-Z780 paternal lineages, respectively. Moreover, when we compared the maternal lineage composition among groups, C1b32 was not observed among local LH-F and LH-HG. The two only migrant LH-F with the B2ak1 haplogroup –observed also in local LH-F and LH-HG– are located in a different sub-branch as they carry a mutation (225A) not present in the other groups. The contrast between the heterogeneity of male uniparental haplogroups and the high uniformity of female haplogroups in LH-F migrant groups aligns more parsimoniously with a matrilineal social organization.

E.7. Kinship analysis

For each pair of newly sequenced individuals with DoC greater than 0.01X, we generated Kinship estimations with READv2[189]and KIN[190] using positions from the SG or 1240K SNP panels. For each pair of individuals, we manually assigned the degree of relatedness after visualizing the indexes generated by each method (i.e. Average pairwise P0 for READv2 and k0 and coefficient of relatedness for KIN) on each SNP panel, crossed with the number of SNPs used. To account for uncertainty in the final assignation, we used the following categories to describe the degree of relatedness: Parent-Child, Siblings, Second Degree, Second Degree / maybe Third Degree, Third Degree, Third Degree / maybe Unrelated, Unrelated / maybe Third Degree, and Unrelated (Fig. S22; Supplementary Data S5).

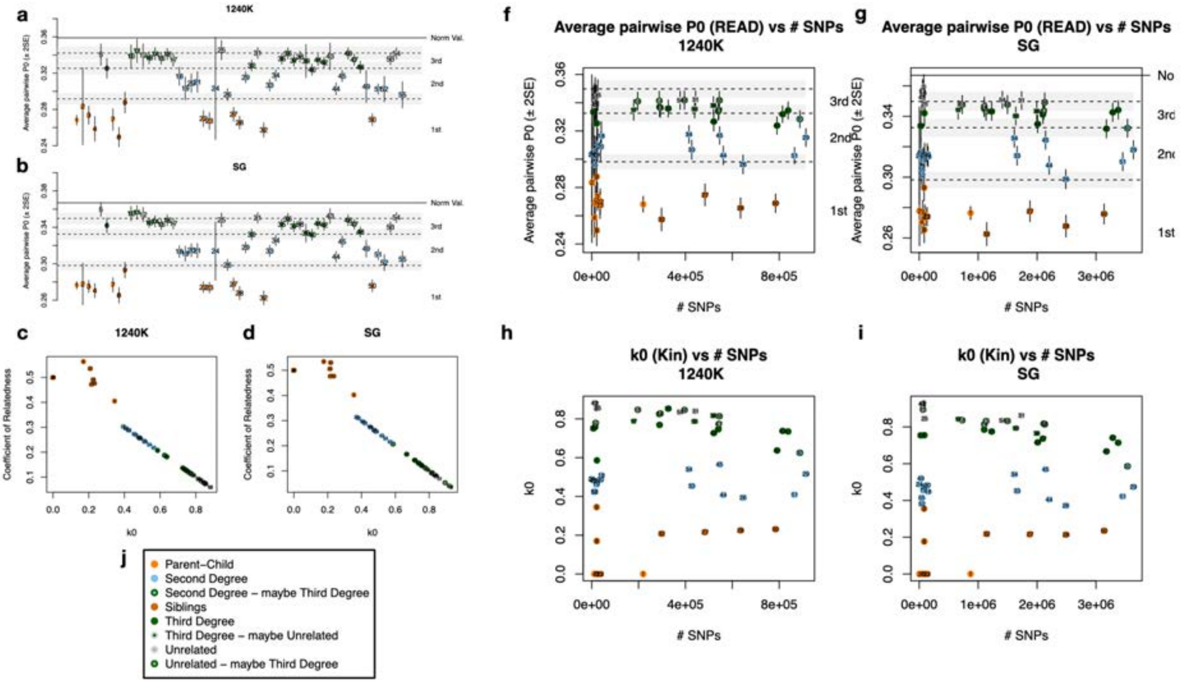

Fig. S22. Final kinship assignment from genomic data.
For the 54 pairs of individuals for which we observed at least one estimation of relatedness out of the 4 obtained (2 methods and 2 SNP panels), we visualized metrics provided by the methods (Average P0 for READv2 in **a, b,** on 1240K and SG SNP panels, respectively; Coefficient of Relatedness for KIN in **c** and **d,** on 1240K and SG SNP panels, respectively). We also visualized how Average P0 (**f** and **g**) and k0 (**h** and **i)** may be affected by the number of SNPs leveraged for the estimation. Based on these plots, we assigned a degree of relatedness, and points are shown according to the legend in **j.** Numbers associated with each individual pair are given in Supplementary Data S5.

Population genetic diversity and structure analyses require individuals to be unrelated. We therefore removed some individuals to avoid any relationship up to the 2nd-degree, by prioritizing removing first individuals involved in the higher number of pairwise relationships, and in case of ties, individuals with lower DoC.

E.8. Population diversity and structure analyses from autosomal markers

E.8.a. Methods for population diversity and structure analyses from autosomal markers

f_3-outgroup

We computed pairwise genetic distances between individuals with more than 50,000 SNPs with a genotype on the considered SNP panel, of the form *1 - f_3(Individual1, Individual2; Mbuti)* using Admixtools[136] v6 (Supplementary Data S6a). From the distance matrix distance, we performed Multidimensional Scaling analyses in R with *cmdscale* package (Supplementary Data S6b; Fig. S24a-h; Main Figs. 2b-c for 1240K SNP panel).
We also computed pairwise genetic distances between individuals of the form *1 / f_3(Individual1, Individual2; Mbuti)* to compute Neighbour-Joining trees with *ape* package[183] in R (Fig. S24i-l).

ADMIXTURE

We obtained genetic ancestry proportion estimates for individuals with more than 30,000 SNPs using ADMIXTURE software[191], on a set of LD-pruned SNPs (pruning with indep-pairwise flag in plink[192] v1.9 with default parameters) with less than 50% of missing genotypes. For each number of putative ancestral populations *K* varying between 2 and 15, we performed 10 independent runs, and we retained the one exhibiting the highest likelihood. Results are presented in Fig. S25 for ancient individuals only, and in Fig. S26 for modern and ancient individuals.

We explored the current geographic distribution of a specific genetic ancestry component (named Central Southern Cone, see Main text and Supplementary Material E.8.b.) from the results of ADMIXTURE analyses including both ancient and modern individuals for *K*=15. We represent the within group distribution of this component proportions (boxplots; Fig. S27a), and the geographic distribution of the median per group (scatter-pie charts on a map; Main Fig. 2d for Southern Cone only and Fig. S27b for the Americas)

Grouping for *f_4-statistics*

As mentioned before, the individuals from the reference panels were grouped mostly according to the metadata associated with the Allen Ancient DNA Resource[173] v54, but also consulting the groupings used in subsequent articles[139,142,143,146,149,151,155,156,158] (Supplementary Data S4; Figs. S17-18). Groups were subsequently assigned to major geographical ranges, and divided according to period: Early Holocene (EH; before 8,500 BP), Middle Holocene (MH; between 8,500 and 3,500 BP), Late Holocene (LH; after 3,500 BP, and before European invasion ~500 BP), and Modern.

The individuals from Uspallata Valley (UV) were grouped in three different ways:
1. All UV individuals together
2. Late Holocene-Hunter Gatherers (LH-HG; encompassing individuals from BB6) and Late Holocene-Farmes (LH-F; encompassing individuals from BR-I, T-II and PLC)
3. LH-HG as above, and LH-F divided into 4 subgroups: Migrant Farmers (LH-MF; containing all but one migrants from PLC), Migrant Outlier Farmer (LH-MFout; containing one migrant from PLC with distinctive genetic ancestry as observed in Admixture and MDS analyses), Local Farmer from PLC (LH-LFplc; the unique sequenced non-migrant individual from), and Local Farmers (LH-LF; containing Locals from BR-I and T-II).

f_4-statistics

We computed different D-statistics of the form *f_4(Mbuti, Group1; Group2, Group3)* using Admixtools[136] v6, filtering out individuals with less than 50,0000 SNPs genotyped on the SNP panel considered (except when a threshold of 30,000 SNPs is explicitly mentioned). In what follows, we describe how Group1, Group2 and Group3 were chosen to test different hypotheses.

To test whether one Modern Group is genetically closer to UV individuals or another Ancient South American group, we performed *f_4(Mbuti, Modern Group; UV, Ancient South American Group)*, where UV are all UV individuals together. In this case, we only leveraged the 1240K SNP panel, and we kept individuals with more than 30,000 SNPs with a genotype (instead of 50,000 for other f_4 analyses). The results (Supplementary Data S6c) are shown in beeswarm-style plots generated with *beeswarm* package in R(Fig. S28).

To test whether any Ancient American groups were attracted to one of the UV groups (LH-HG, LH-MF, LH-MFout, LH-LF, LH-LFplc), we performed *f_4(Mbuti, Ancient American Group; UV1 ,UV2)*, where UV1 and 2 is a pair of UV groups. The results ((Supplementary Data S6d) are shown in beeswarm-style plots (Fig. S29).

To test whether UV was closer to one or other Ancient American Group, we performed *f_4(Mbuti, UV; Ancient American Group 1, Ancient American Group 2),* where UV are all UV individuals together. The results (Supplementary Data S6e) were plotted in the form of heatmaps (Fig. S30) generated in R-base.

To test whether any Central or North American groups were closer to Uspallata groups than to other ancient South American groups, we performed *f_4(Mbuti, Central or North American Group; UV, Ancient South American Group),* where UV are all UV individuals together. The results (Supplementary Data S6f) are shown in beeswarm-style plots (Fig. S31) generated with ggplot2[193] and ggbeeswarm (https://github.com/eclarke/ggbeeswarm) packages in R.

To test whether LH-HG and LH-F were closer to each other as compared to other ancient American groups, we performed *f_4(Mbuti, LH-HG; Ancient American Group, LH-F)* and *f_4(Mbuti, LH-F; Ancient American Group, LH-HG)*. To test whether LH-HG and LH-F form a clade as compared to other Ancient American groups we performed *f_4(Mbuti, Ancient American Group; LH-HG, LH-F)*. The results (Supplementary Data S6g) are shown in beeswarm-style plots (Main Fig. 3c and Fig. S32).

Consideration concerning *f*-statistics on different SNP panels

Many SNPs in 1240K are monomorphic in the Americas, and thus are uninformative to study within continent population diversity[194]. In order to increase the number of informative sites, one could only use SNP panels that contain higher numbers of SNPs, such as the SG SNP panel (5.6 M positions). However, individuals sequenced through in-solution capture targeting the 1240K SNP panel, are assigned missing genotypes for most positions of the SG SNP panel that do not overlap with the 1240K SNP panel. This explains why the differences of *f*-statistics computed for the same comparison but with the SG or the 1240K SNP panel are greater when individuals sequenced through in-solution capture included in the analysis (Fig. S23). Therefore, to avoid systematic bias induced by differences in the genotyping panel, we removed genomes produced by 1240K capture from the analyses based on the SG SNP panel. However, since the spatio-temporal range covered by Shotgun-sequenced individuals is still limited (99 with shotgun data vs 328 with 1240K-capture data), the main results throughout the article are based on the 1240K SNP panel to privilege the number of ancient individuals (and groups) in our analyses.

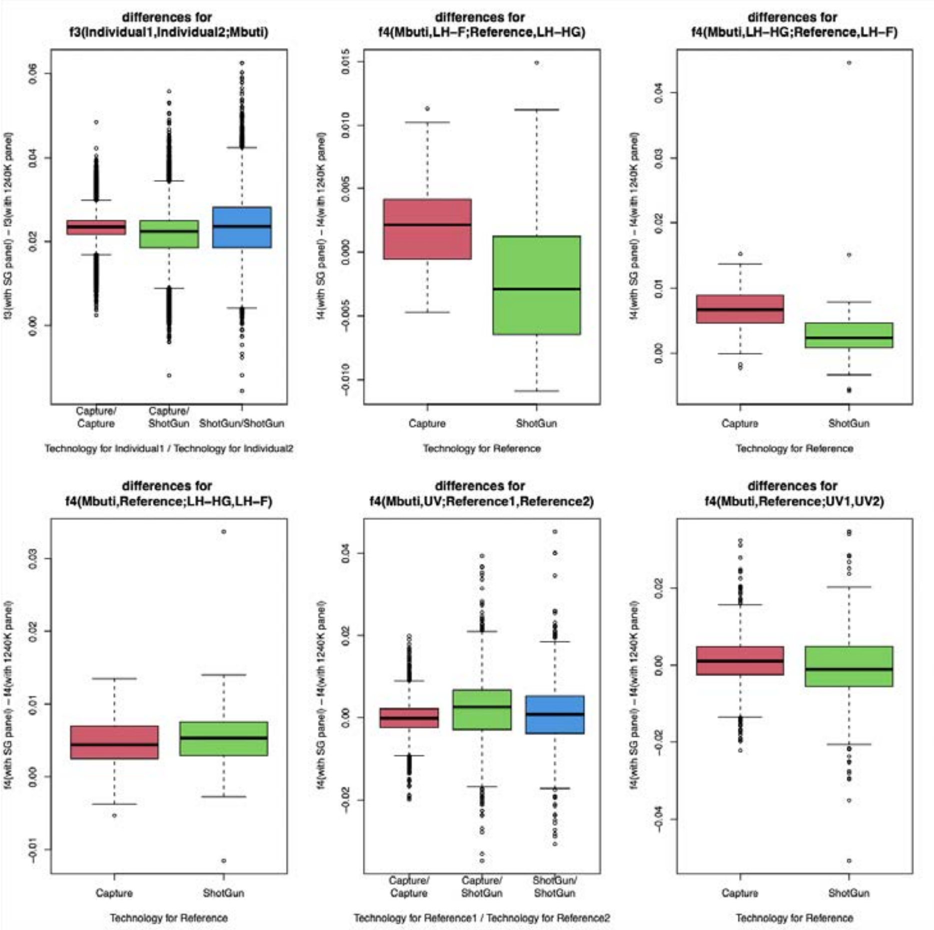

Fig. S23. Differences of f-statistics computed on different SNP panels.
Each plot shows the distribution of the differences of f-statistics for the same comparisons but performed on the 1240K or SG SNP panels. Comparisons are grouped whether they involve none, one, or two groups for which individuals have been sequenced through in-solution capture.

E.8.b. Supplementary results for population diversity and structure analyses from autosomal markers

A specific genetic ancestry component observed for the Central Southern Cone

Our Multidimensional Scaling (MDS) Analyses based on genetic distances among pairs of individuals of the form f_3*(Individual1, Individual2; Mbuti)* show that, regardless of the SNP panel used, Uspallata individuals form a cluster in one of the four first MDS dimensions (Fig. S24). All, except two PLC individuals are located at the most extreme position in that dimension, and with the two remaining PLC individuals being at a more intermediate position together with BB6, T-II and BR-I individuals, as well as two individuals from Calingasta Valley previously sequenced[151], and a bit more distant, the individual from El Toro archaeological site, situated at the southern San Juan province and associated with Qhapaq Hucha ceremony[158].

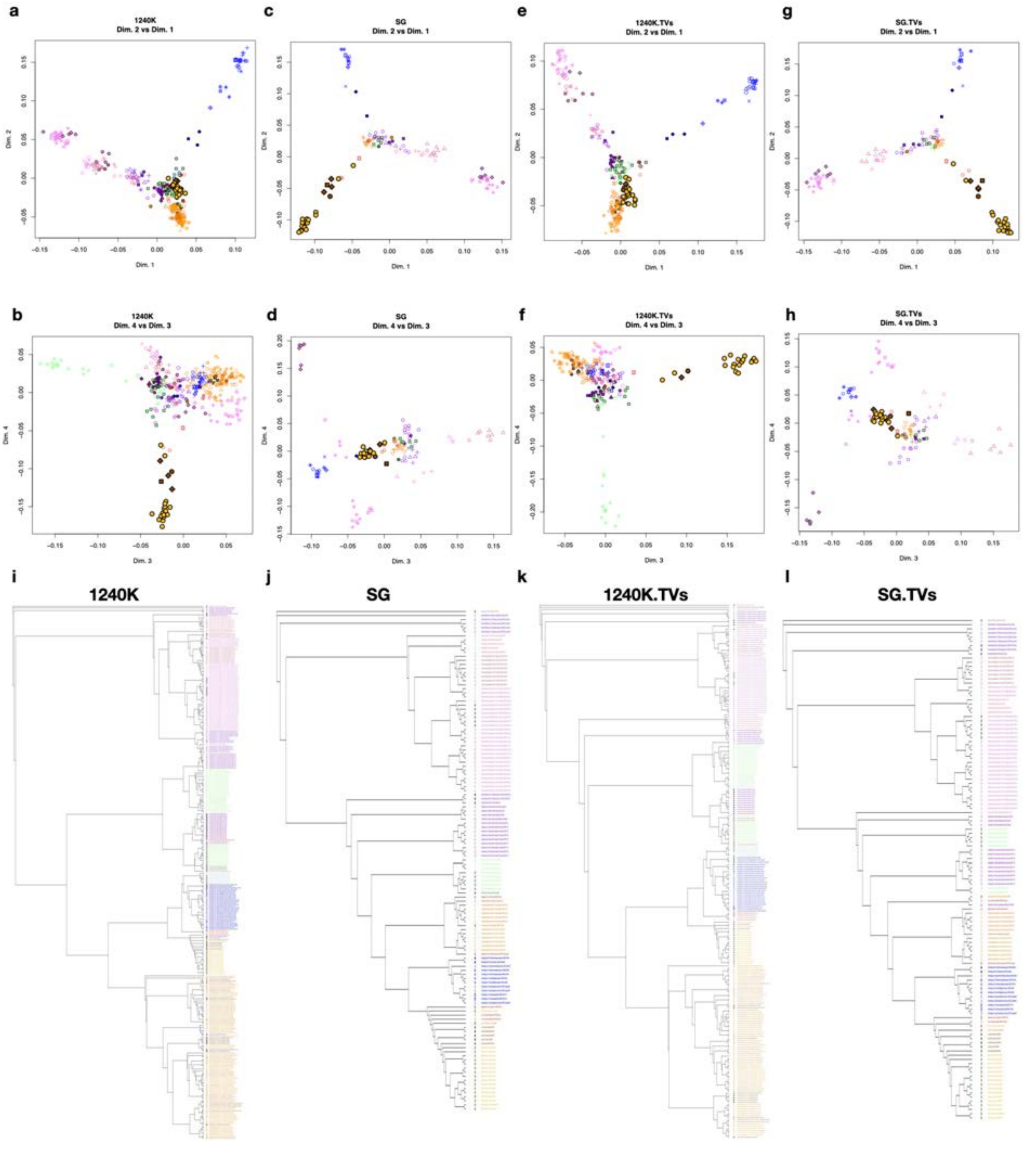

Fig. S24. CSC genetic component detected regardless of the SNP panel considered. Multidimensional Scaling Analyses **(a-h)** and Neighbour-Joining trees **(i-l)** from pairwise genetic distances among individuals of the form on *1-f_3(Individual1, Individual2; Mbuti), and 1/f_3(Individual1, Individual2; Mbuti)*, respectively, considering all autosomal positions on the 1240K (**a, b, i**) and the SG SNP panels (**c, d, j**), or only the transversions from both panel (**e, f, k** and **g, h**, **l,** respectively) The first and second MDS dimensions are shown as x and y-axes, respectively (**a, c, e** and **g**). The third and fourth MDS dimensions are shown as x and y-axes, respectively (**b, d, f** and **h**). Points are represented as shown in Main Fig. 2a, and the reference of icons and colors used to identify ancient groups can be consulted in Fig. S17. This figure is associated with Supplementary Data S6a and Supplementary Data S6b. Panels **a** and **b** are also shown in Main Figs. 2b-c, and they are included here again for a simpler interpretation in relation to other panels of the present figure.

Similar observations are derived from ADMIXTURE analyses including only ancient individuals. Indeed, these analyses identify a specific genetic component linked to UV and Calingasta from K=4 onwards (shown in yellow component in Fig. S25). At *K=14,* this component is ~100% in all LH-F migrants, while it only reaches ~50-60% in other UV and Calingasta individuals. It is also observed at ~20% in El Toro individual (further north of Calingasta who died in Inka times), ~15% LH Central Chile (Conchalí, ~700 BP) and La Pampa (Laguna Toro, ~2400 BP; Laguna Chica, ~1600 BP) individuals, and ~5-8% in MH La Pampa (Arroyo Seco 2, ~7700 BP; Laguna Chica, ~6800BP) and Central Chile (Los Rieles, ~5100 BP) individuals. These results suggest that this genetic ancestry component originally named "Central Western Argentina" by Luisi et al.[169] was more broadly expanded across the Southern Cone, and we suggest to name it Central Southern Cone (CSC).

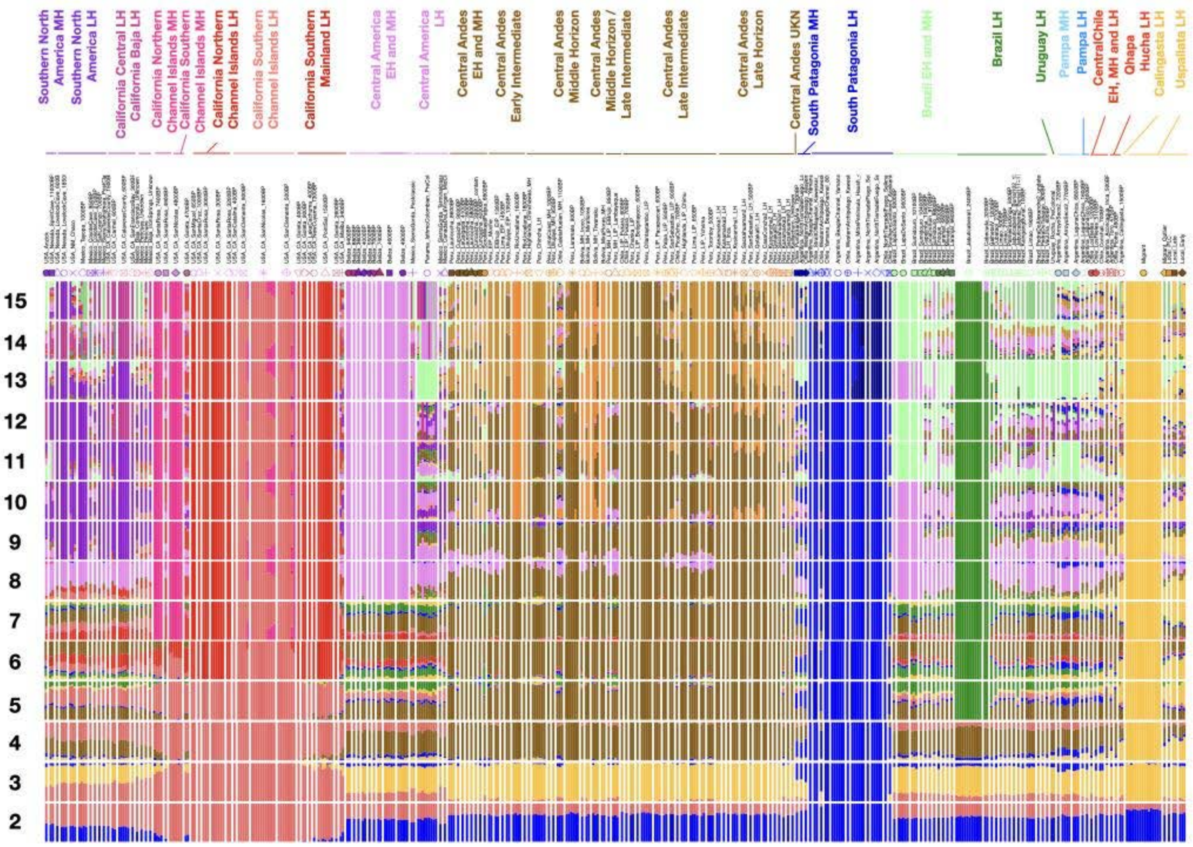

Fig. S25. Admixture analysis performed on ancient American genomes.
Analyses were performed using the 1240K SNP panel, filtering out individuals with less than 30,000 positions with known genotype. For each number of ancestral populations considered (*K* from 2 to 15), we performed 10 independent runs and only the one exhibiting maximum likelihood is shown.

Genetic continuity with present-day populations

First, Admixture analyses including modern individuals (Fig. S26) show that the specific genetic component maximized in ancient UV and Calingasta individuals, is present in some significant proportions in modern individuals from Argentina and Chile (*K*=10 and 11). The fact that it was observed in individuals from Santiago de Chile, was originally attributed by scholars to the large-scale deportation of Huarpe people from Argentina as forced labor in Chile during the colonial period[169]. However, local Huarpe communities, based on their oral traditions, interpreted these results as evidence of their long-standing presence across the Andes.

Our ADMIXTURE analyses demonstrate a remote link between the genetic diversity observed in Central Chile (Huilliche and Pehuenche) and in UV and Calingasta. Indeed, for *K*=12 onwards, this component divides into two sub-components: one maximized in Central Chile, for both modern Indigenous and ancient groups, and which segregates at high proportions in Santiago de Chile's cosmopolitan population; and a second one that is maximized in UV ancient individuals. We checked the contribution of the latter (using ADMIXTURE results for *K*=15), in individuals according to their geography (Main Fig. 1d, Fig. S27), and it is mostly observed in Argentinean individuals, particularly in the provinces of San Juan, Mendoza, La Rioja, Tucumán, Catamarca, San Luis, and Córdoba (i.e. in Central Western, North Western and Central regions of Argentina). These results reinforce the suitability of the name suggested for this genetic ancestry component (Central Southern Cone; CSC).

Through f_4-statistics of the form f_4*(Mbuti, Modern group; UV, Ancient South America group)*, we confirmed that CSC genetic component mostly segregates today in Central and Central Western Argentina regions, as modern groups from the Argentinean provinces mentioned above tend to be genetically closer to ancient UV individuals than to other ancient South American groups (Fig. S28).

Notably, these results definitely dismiss the misinterpretation of Luisi et al. (2020)[169]' results by Pedersen et al. (2022)[151], who stated that current populations in San Juan are unrelated to the original Huarpes indigenous groups in the region. Notably, such a misleading claim was mostly sustained by the results obtained from a very low-quality genomic data from a single individual (SJ-Nit-2; not passing our quality filters, and thus not included in the present study). Our results reflect a story that is much more consistent with the family histories and the oral traditions transmitted by local communities concerning their remote link with UV territory.

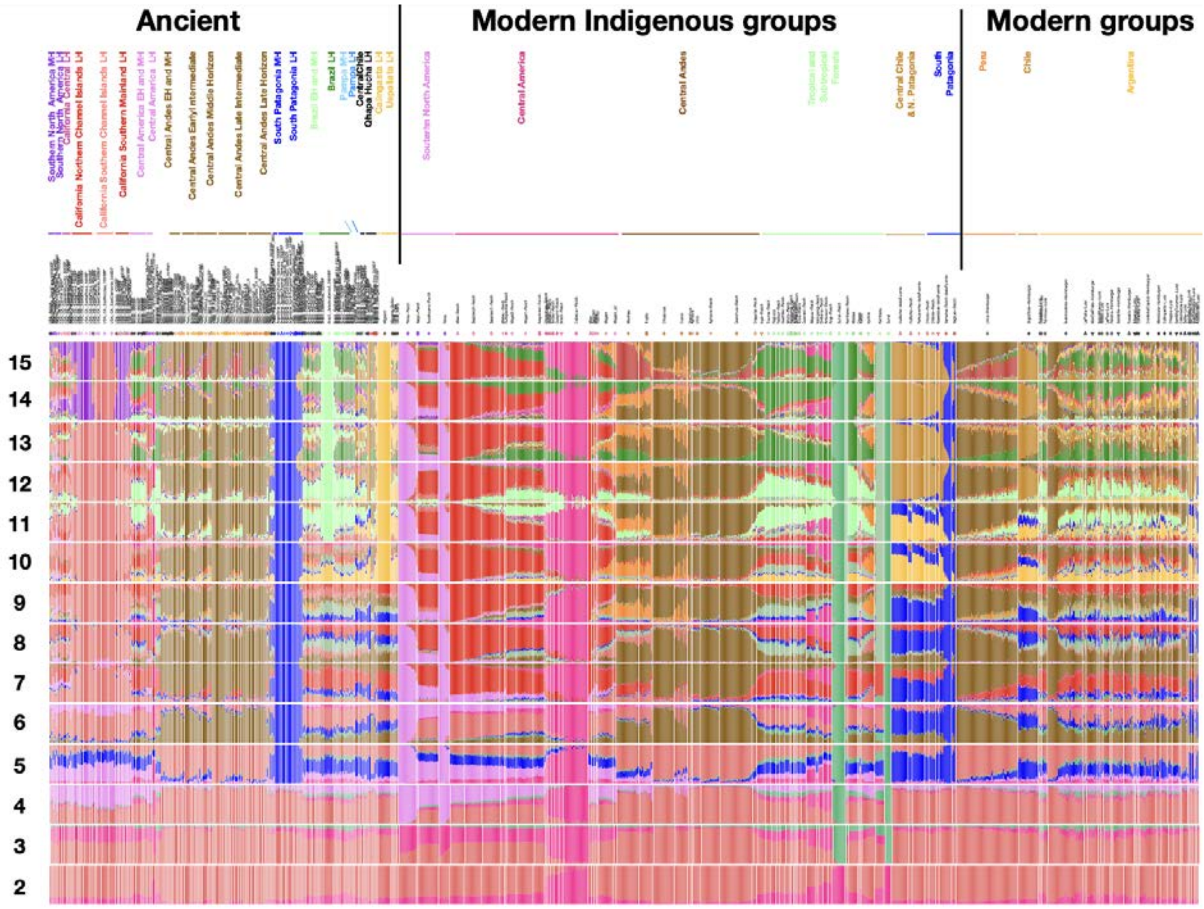

Fig. S26. Admixture analyses with Modern individuals.
Analyses were performed filtering out individuals with less than 30,000 positions with known genotype. For each number of ancestral populations considered (*K* from 2 to 15), we performed 10 independent runs and only the one exhibiting maximum likelihood is shown.

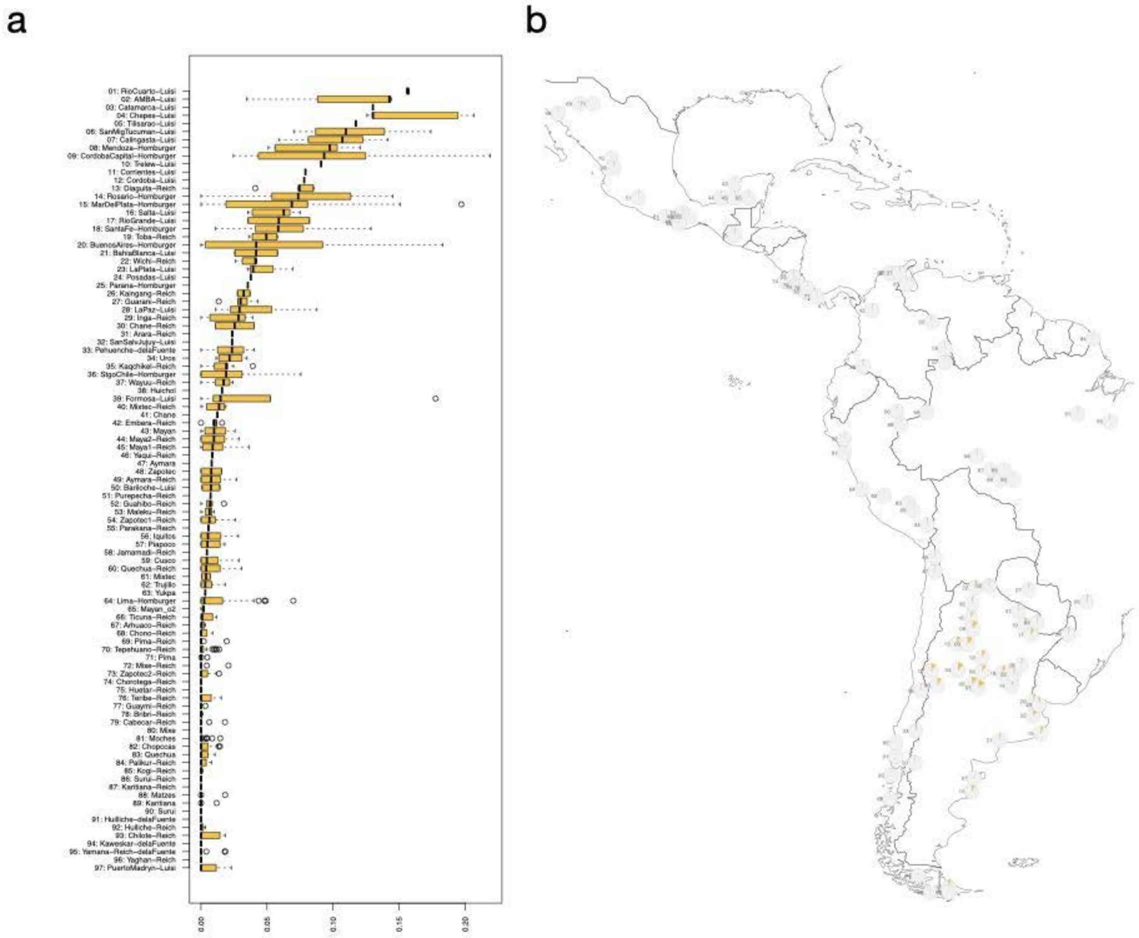

Fig. S27. CSC genetic component in modern populations.
Geographic distribution of the ancestry genetic component identified as Central Southern Cone (CSC; shown in yellow in Fig. S26) in modern groups across the Americas. The proportions were obtained from the ADMIXTURE analysis including ancient and modern groups with *K*=15. **a.** Distribution of the genetic ancestry proportion within each modern group. **b.** Pie-charts represent the median proportions for CSC (yellow) and the sum of all other components (grey), per modern group. Pie-charts are located at the geographical coordinates reported in the original articles[157,171,172,174] (some were manually shifted for better visualization). The number associated to each group is equivalent in **a** and **b**. Panel **b** is the geographical extension northward of the results shown in Main Fig 1d.

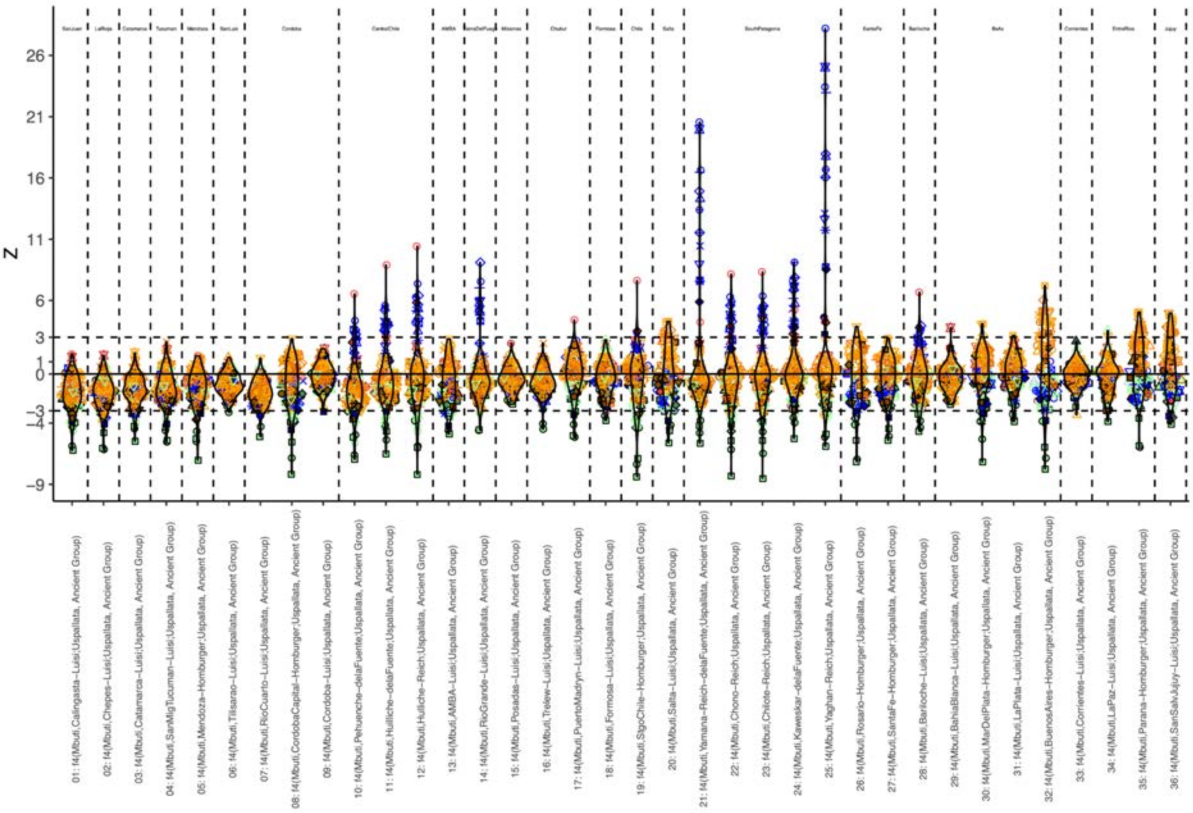

Fig. S28. ***f_4(Mbuti, Modern Group; UV, Ancient South American Group).***
Z-scores derived from *f_4(Mbuti, Modern Group; UV, Ancient South American Group)* are shown by grouping along the x-axis by modern groups, and points are represented according to the ancient South American Group. Modern groups are ordered with increasing median of the Z-scores across all *f_4* comparisons involving those groups. Dashed lines for |Z|=3 are used as thresholds for significance. Points are represented as shown in Main Fig. 2*a*, and the reference of icons and colors used to identify ancient groups can be consulted in Fig. S17. This figure, associated with Supplementary Data S6c, shows that modern Individuals from Western Argentina exhibit the highest affinity with Ancient UV individuals as compared to other Ancient South American groups.

Structure among UV groups

The individuals from UV do not form a compact cluster in the MDS analyses (Main Fig. 2c; Fig. S24), and their genetic ancestry proportions show some variability in the ADMIXTURE analyses (Fig. S25). Based on these observations, along with archaeological context, we categorized them into five groups: pre-farming hunter-gatherers (LH-HG), migrant farmers (LH-MF), a genetically distinct migrant farmer (LH-MFout), local farmers (LH-LF), and a local farmer from PLC (LH-LFplc), as previously described. To assess the most appropriate strategy to group them when comparing to other ancient American populations, we performed *f_4(Mbuti, Ancient American Group; UV1, UV2)* analyses. The results show that all ancient American groups form a distinct clade in relation to all pairs of UV groups (Fig. S29). This confirms that, when investigating the evolutionary history in UV by studying its connections to other ancient American populations, it is valid to treat all UV individuals as a single group. Additionally, these analyses indicate that migrant LH-F and local LH-F are more closely related to each other than to LH-HG, supporting our decision to classify them together as LH-F when examining genetic continuity with LH-HG. Despite their genetic dissimilarity, we decided to also include in this group the two anomalous individuals from PLC (LH-LFplc and LH-MFout) to avoid forcing grouping choices based solely on genetic criteria (i.e., not accounting for archaeological contexts).

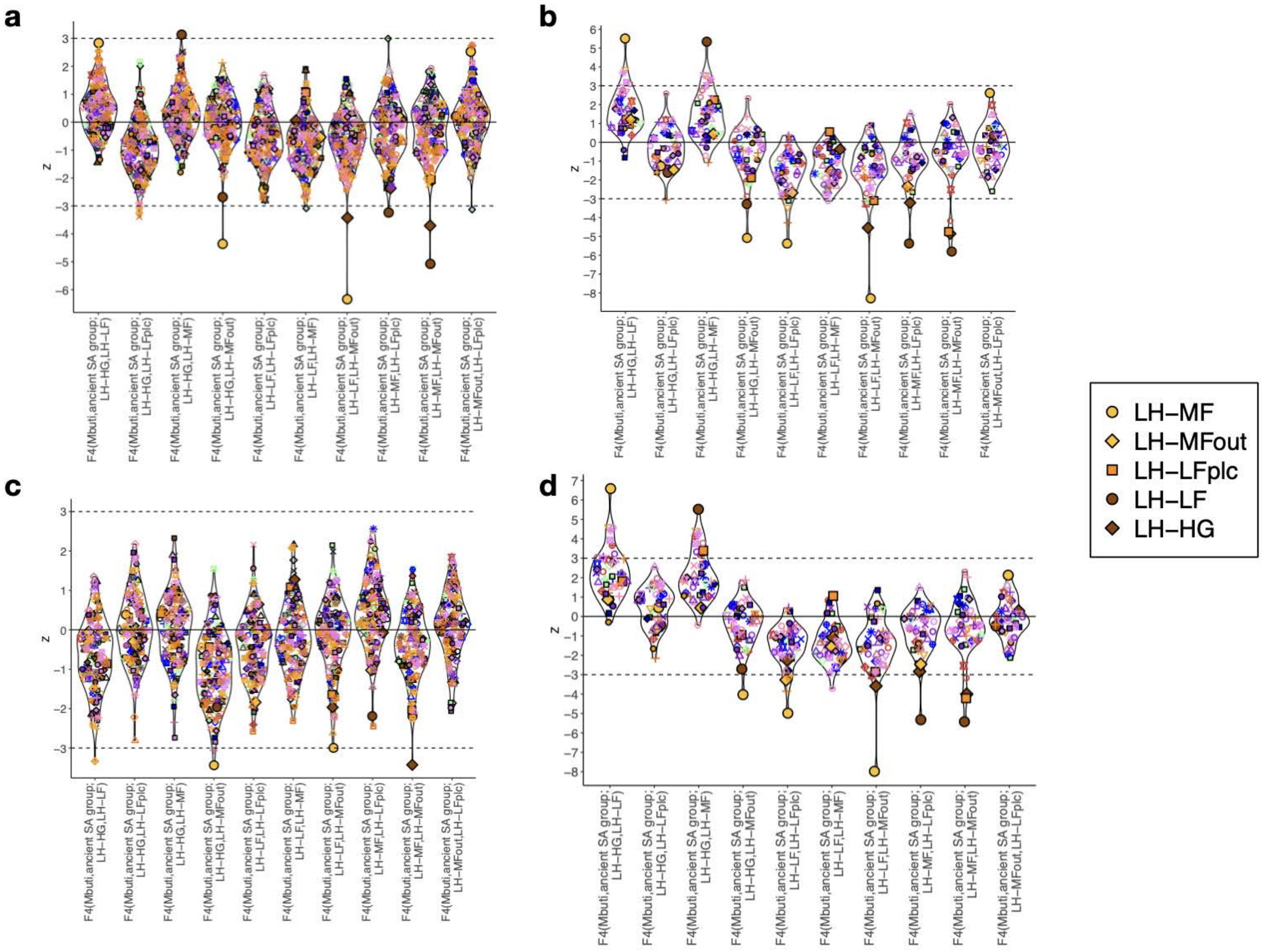

Fig. S29. ***f_4*(Mbuti, Ancient American Group; UV1, UV2).**
Z-scores derived from f_4(Mbuti, Ancient American Group; UV1, UV2) are shown, with pairs of UV1 and UV2 along the x-axis. Points are represented according to the Ancient American Group. Larger points are when the Ancient American Group is one group from UV, following the code shown in legend. Dashed lines for |Z|=3 are used as thresholds for significance. f_4 were computed considering all autosomal positions on the 1240K (**a**) and the SG SNP panels (**b**), or only the transversions from both panels (**c, d** respectively). Points are represented as shown in Main Fig. 2a, and the reference of icons and colors used to identify ancient groups can be consulted in Fig. S17. This figure, associated with Supplementary Data S6d, shows that UV individuals can be assigned to one unique group to study the potential origin of the genetic ancestry component specific to UV, and its links with ancient American groups.

Genetic links with other American groups

The results derived from ADMIXTURE and MDS from f_3-based distances show that the genetic diversity observed for UV and Calingasta individuals fall within the South American diversity. Moreover, the CSC genetic component is also observed in our ADMIXTURE analyses in non-negligible proportions in Central Chile and La Pampa individuals from the Late Holocene (LH) and Middle Holocene (MH, ~5-15%, see above). At *K=15,* several Calingasta and UV individuals (1 Calingasta, 1 LH-HG, 1 LH-MFout, and 1 LH-Fplc) exhibit a ~100% proportion estimates of a new component, which seems specific to groups that inhabited territories located in today's Uruguay and South East Brazil. This component also reaches 100% in a few LH individuals from Central America and California Baja. However, this result is poorly supported by other analyses described below, and ADMIXTURE results should not be overinterpreted[195], particularly because we are using aDNA data.

f_4-statistics analyses of the form *f_4(Mbuti, UV, Ancient American Group1, Ancient American Group2)* show that UV individuals are genetically closer to ancient groups from Calingasta and El Toro (both in San Juan province, close in the north of UV; Main Fig. 2c) than to any other ancient American group (Fig. S30; Supplementary Data S6e). They also exhibit high genetic affinity with LH Pampa and Central Chile groups. UV individuals are also genetically closer to ancient Central Andes groups, than to groups from Brazil, Uruguay and Patagonia, although most *Z*-scores do not reach significance when including LH groups from the latter regions.

Given the lack of a clear genetic affinity between UV individuals and any early South American groups with available genomic data (i.e., from the Middle Holocene or earlier), we explored whether they exhibit specific genetic affinity with any ancient Central and North American population. To do so, we performed f_4-statistics analyses of the form *f_4(Mbuti, ancient North or Central American group; UV; ancient South American group)* (Fig. S31, Supplementary Data S6f), systematically testing all possible pairwise comparisons between 51 North or Central American groups and 105 South American groups, yielding a total of 5,355 comparisons (Supplementary Data S6f). To facilitate visualization, we grouped the 51 North and Central American populations into 13 broader spatio-temporally defined groups (Fig. S16), which were established after running the analyses. This allowed us to summarize the results in violin plots (Fig. S31), where each violin represents the distribution of f_4 values for all South American groups relative to a given North or Central American larger group. If a larger group contains multiple original populations (e.g., five of the 51), each South American group appears in the violin plot as many times as the number of comparisons performed (i.e., five). In this representation, positive f_4 values indicate that the North or Central American group is genetically closer to the South American group than to UV, while negative values suggest a stronger affinity between UV and the North American group.
As expected, almost all comparisons did not reach significance ($|Z| < 3$), reinforcing the idea that UV genetic diversity falls within the one observed in South America. However, some exceptions are worth mentioning. First, some cases involving Sumidouro groups (EH Brazil as South American group) and Spirit Cave (as North American group) are significantly negative ($Z < -3$), indicating higher affinity of UV to Spirit Cave than to EH groups in Brazil. Furthermore, this is not limited to Spirit Cave, since several comparisons involving Sumidouro groups (and particularly the group assigned to 10,100BP), show North and Central American groups closer to UV ($Z < -3$) than these early Brazilian groups. Second, the other few significant negative tests ($Z < -3$) involve in almost all cases MH La Pampa and South Patagonia groups (as South American group) and LH Californian groups (as North America group), suggesting that UV has tighter evolutionary links with these Californian groups than to early groups in La Pampa and South Patagonia. This contrasts with the comparisons involving early Central Chile groups (EH and MH individuals from Los Rieles archaeological site), which show that all North and Central American groups form a single clade–*i.e.,* are equidistant– in relation to UV and early Central Chile ($|Z| < 3$). Together, these results suggest that UV populations, along with the CSC component, may originate –at least a proportion of its ancestry– from a yet uncharacterized migration wave of a distinct South American group, whose common ancestry with other South American populations traces back to Central and North America.

Additionally, some comparisons involving the LH Central Andes group are significantly positive ($Z > 3$), when some Californian or Central American groups are included. The same holds true when replacing the LH Central Andes group with the LH Central Chile group from the Conchalí archaeological site (although most comparisons reaching significance are limited to those involving MH Central America groups). These results suggest that the reported gene flow[152] that occurred at the end of the Middle Holocene between populations from Central Andes and other groups carrying Californian- and/or Central American-related ancestry, may have also reached Central Chile, but has not impacted UV.

Neighbour-joining trees reconstructed from f_3-based distances show that one Calingasta individual clusters with Central American groups, and a MH individual from Belize clusters with the remaining UV and Calingasta individuals (Fig. S24i). Additionally, UV and Calingasta individuals form a clade with South Patagonia and MH La Pampa individuals (Fig. S24i and Fig. S24k). Although these neighbour-joining trees are based on recombinant genetic diversity, which summarizes the combined effects of multiple gene flow events among groups—making their interpretation complex—they nonetheless highlight genetic affinities between UV and Calingasta with Central American groups, reinforcing the f_4 results described above in this section.

Altogether, our analyses did not allow deciphering the potential origin of the specific genetic ancestry component observed in UV and Calingasta, present also in LH Pampa and Central Chile groups. They suggest, nonetheless, a remote origin, which was independent of the lineages observed in EH and MH Brazil, MH Pampa, South Patagonia and Central Andes. Despite its relevance, understanding the origins of this ancestry is beyond the scope of this study, and would require new ancient genomes from individuals that lived in regions not yet counting with aDNA data, particularly those from most inland regions throughout South America. Such data would help to get a better understanding of the timing and regions where the different divergence events occurred during the rapid human spread in South America.

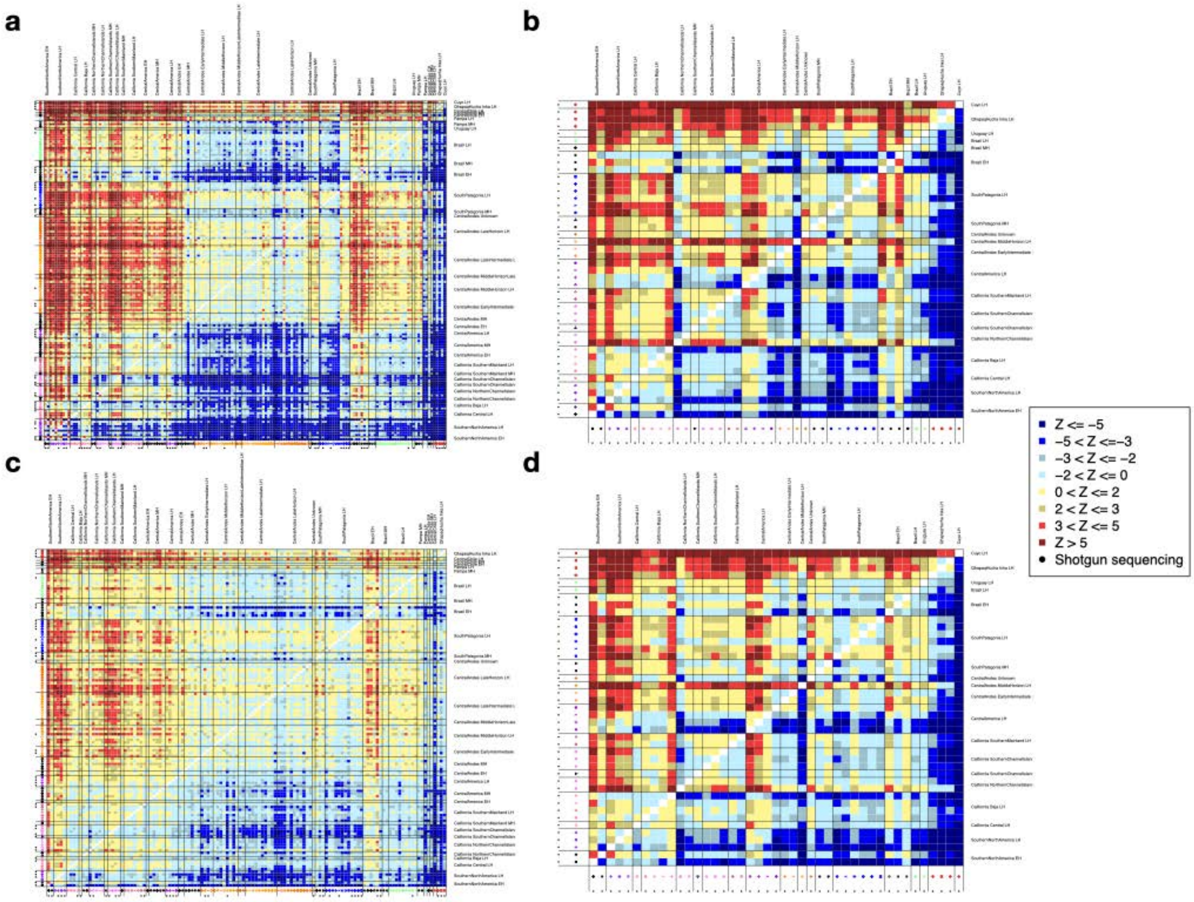

Fig. S30. ***f_4Mbuti, Uspallata; Ancient American Group 1, Ancient American Group 2).*** Heatmaps summarizing Z-scores derived from *f_4(Mbuti, Uspallata; Ancient American Group 1, Ancient American Group 2)*, where Ancient American Groups 1 and 2 are represented by point type and colors on the x and y axes, respectively. Ancient groups whose individuals were sequenced through shotgun are marked with black dots. Ancient American groups are ordered by major geographical ranges and periods. Z-scores greater and lower than 0 are colored in shades of yellow to red and light blue to dark blue, respectively. Ancient San Juan groups are from Calingasta Valley (Cuyo Late Holocene) and from El Toro archaeological site (associated with Qhapaq Hucha ceremony), represented with a red dot and a red star of David, respectively. f_4 were computed considering all autosomal positions on the 1240K (**a**) and the SG SNP panels (**b**), or only the transversions from both panels (**c, d** respectively). This figure, associated with Supplementary Data S6e, shows that UV individuals are closer to ancient groups from Calingasta and El Toro (San Juan province) than to any other ancient American group.

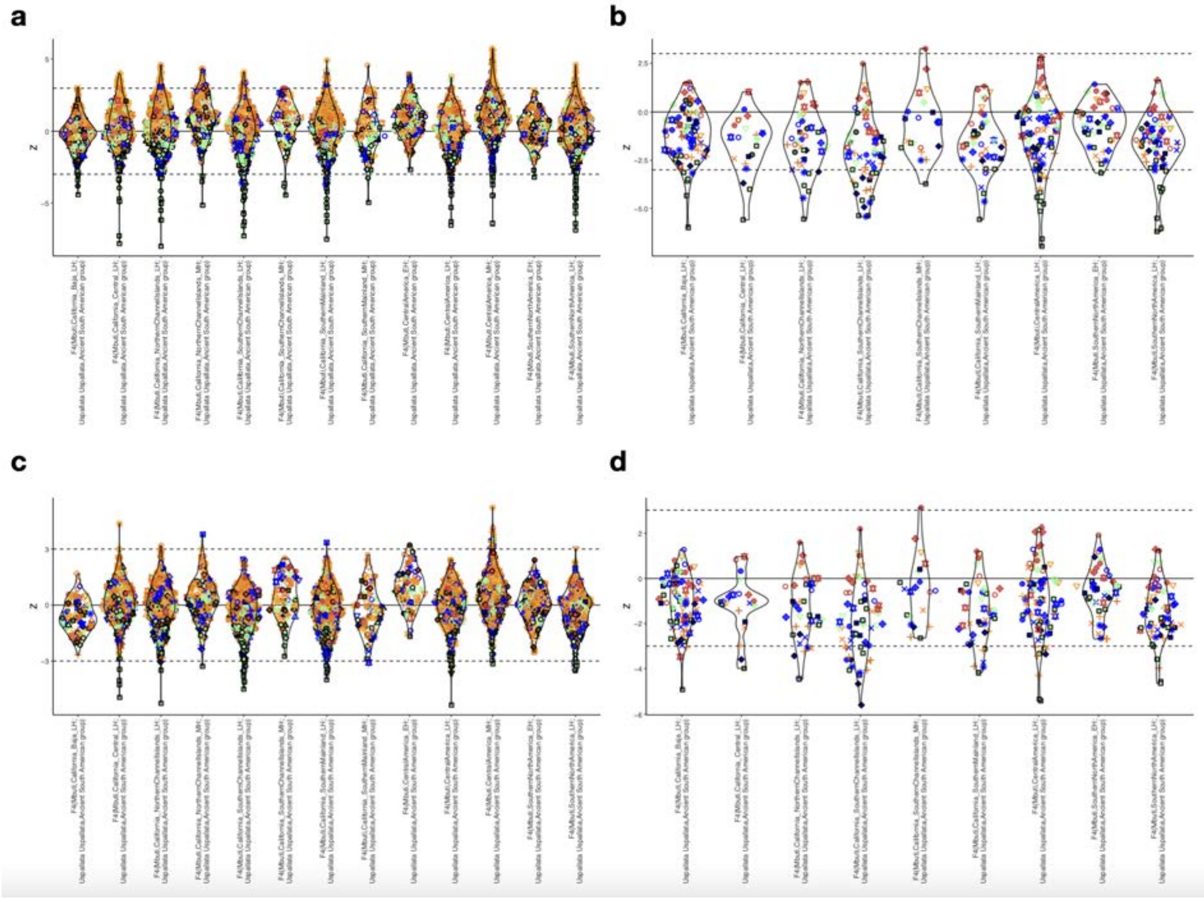

Fig. S31. ***f_4(Mbuti, Ancient Central/North American Group; UV, Ancient South American Group).***

To explore whether UV exhibits specific genetic affinity with any Central or North American group, Z-scores derived from *f_4(Mbuti, Ancient Central or North American Group; UV, Ancient South American Group)* are shown by grouping the Ancient Central and North American groups along the x-axis by geographical range and period. Points are represented according to the Ancient South American Group. In one column, South American Groups appear N times when the grouping on the x-axis includes N Central or North American groups. Dashed lines for |Z|=3 are used as thresholds for significance. f_4 were computed considering all autosomal positions on the 1240K (**a**) and the SG SNP panels (**b**), or only the transversions from both panels (**c, d** respectively). Points are represented as shown in Main Fig. 2a, and the reference of icons and colors used to identify ancient groups can be consulted in Fig. S17. This figure, associated with Supplementary Data S6f, does not allow to identify clear genetic links with ancient groups from any specific Central or North American region.

<u>Genetic continuity between LH-HG and LH-F</u>

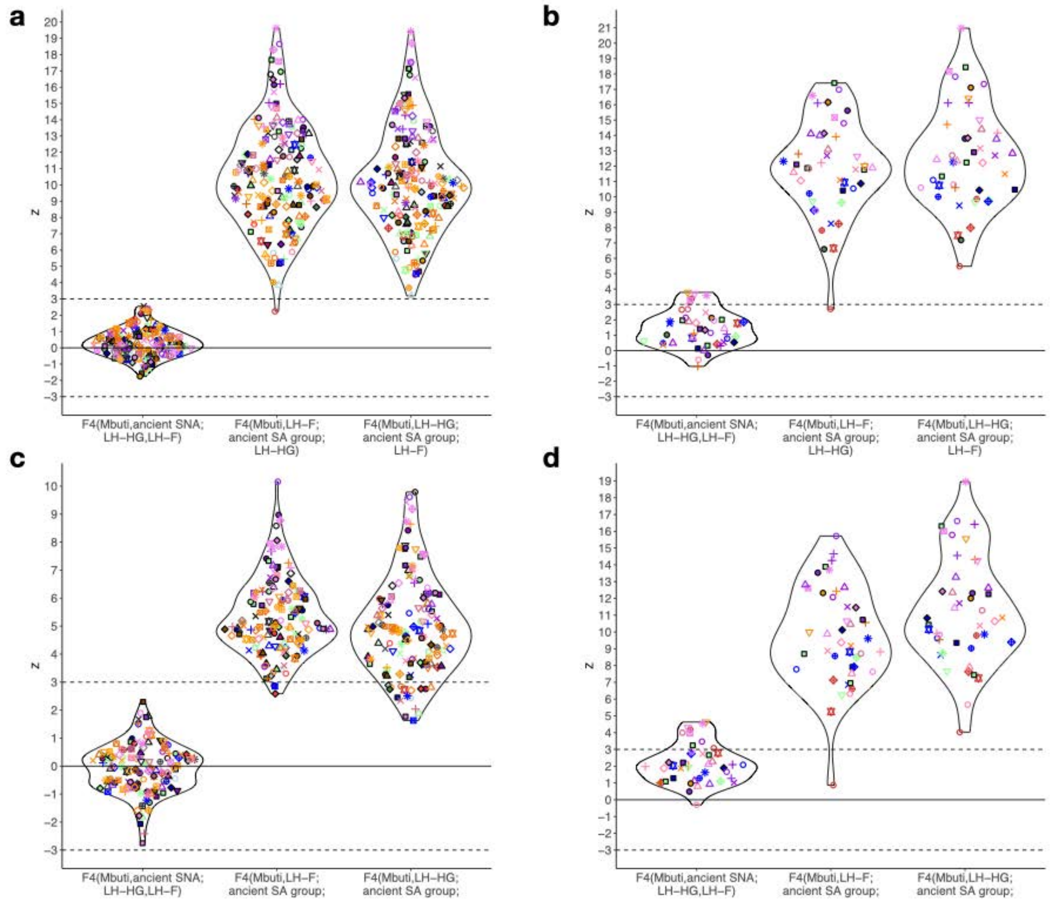

Fig. S32. Genetic continuity between LH-HG and LH-F in Uspallata Valley.
We performed different f_4-statistics analyses involving one Ancient SNA group, LH-F and LH-HG, leveraging 1240K (**a,c**) or SG (**b,d**) SNP panels considering all positions (**a,b**) or only transversions (**c,d**). Each point represents a comparison, with the point representing the Ancient SNA group involved in the f_4-statistics, while the comparisons are grouped according to the position of the Ancient SNA Group, LH-HG and LH-F in the formula. The y-axis is for the Z-score derived from the f_4. Dashed lines for |Z|=3 are used as thresholds for significance. f_4 were computed considering all autosomal positions on the 1240K (**a**) and the SG SNP panels (**b**), or only the transversions from both panels (**c, d** respectively). Points are represented as shown in Main Fig. 2a, and the reference of icons and colors used to identify ancient groups can be consulted in Fig. S17. This figure is associated with Supplementary Data S6g. Panel **a** is the same as in Main Fig. 3c but it is shown here again for a simpler interpretation in relation to other panels of the present figure.

E.9. Inferences related to effective population size from autosomal markers

Conditional heterozygosity

To estimate genetic diversity in a group, we computed conditional heterozygosity using POPSTATS (https://github.com/pontussk/popstats) with the September 26, 2018 version with default settings. Conditional heterozygosity is obtained by sampling a random allele from each of two randomly chosen individuals at a known panel of polymorphisms[196]. We performed these analyses on groups with at least two individuals with more than 50,000 genotyped SNPs on the considered SNP panel (1240K or SG; Supplementary Data S7a and Supplementary Data S7b, respectively; Main Fig. 5a).

Effective population size (*Ne*) from Run of Homozygosity (hapROH)

To identify Runs of Homozygosity (ROH), we used hapROH[197] with the default parameters to estimate the length of segments in homozygosity for individuals with more than 400,000 SNPs on the 1240K SNP panel, for which the method is calibrated. From the hapROH estimates, we estimated effective population size (*Ne*) for each group, removing inbred individuals, defined as those with more than 50 Runs of Homozygosity longer than 20cM (Supplementary Data S7c; Main Fig. 5b).

Ne variation across time (hapNe)

We estimated Ne variation across time for LH-MF with hapNe[198] with default parameters, using the 1240K SNP panel, for which the method is calibrated, and considering only individuals with more than 50,000 genotyped SNPs (Main Fig. 5c).

I).E.10 Pathogen screening and MTBC phylogenetic analyses

Screening shotgun aDNA data datasets for the presence of pathogens

All adapter-trimmed, merged sequencing reads were screened for evidence of ancient pathogen DNA using the HOPS v0.33 pipeline[197]. Reads were aligned against a custom RefSeq Genome database (constructed on November 3, 2017) consisting of representative viral, bacterial, and eukaryotic genomes. Only alignments with 90% sequencing identity or greater were considered. MALT, integrated into the HOPS pipeline, was run with the following parameters: -id 90 -lcaID 90 -m BlastN -at SemiGlobal -topMalt 1 -sup 1 -mq 100 -verboseMalt 1 -memoryMode load.
HOPS evaluates metagenomic taxonomic assignments using multiple authenticity criteria, including read abundance, edit distance distributions, and damage patterns characteristic of ancient DNA[199]. Based on these criteria reads from three individuals (AR0185, AR0501, AR0515) aligned to the *Mycobacterium tuberculosis* complex (MTBC) node. These candidate MTBC reads were subjected to additional validation through reference-based mapping to the MTBC ancestor reference genome[200] using BWA aln v0.7.18 with parameters suited for partial UDG treated libraries (-l 32; -n 0.1; -q 37). Alignments were visually inspected in MEGAN (citation) and IGV (citation) to assess read distribution and mismatch patterns. Weak positive hits were excluded based on poor genomic coverage, a high mismatched rate, and a high number of duplicate versus unique reads. Environmental mycobacterial background was considered in authenticating MTBC reads given the conservation of regions across mycobacterial genomes.

Additional pathogen candidates identified by HOPS included taxa such as *Clostridium spp., Klebsiella pneumoniae, Streptococcus sp.* which are common environmental microbes known to exist in soils and water. These were not confirmed as authentic infections due to their ubiquity and lack of aDNA authentication. Similarly, hits to *Schistosoma mansoni* were excluded due reference genome issues including the complex structure and repetitive elements which can complicate taxonomic assignments in metagenomic datasets[201]. Taxa consistent with oral microbiome communities (*Parvimonas micra, Porphyromonas gingivalis*, *Tanerella forsythia*, *Streptococcus mutans*, *Treponema denticola, Viellonella parvula*) were identified but will be presented in future publications.

Building of a reference phylogenetic tree of MTBC

To construct a robust phylogenetic framework for *Mycobacterium tuberculosis* complex (MTBC), we assembled a curated dataset of 79 modern and 6 ancient publicly available high-coverage MTBC genomes[202], representative of the known global diversity across major lineages. These genomes were selected from previous publications and included representatives of all human- and animal-infecting lineages, including *M. tuberculosis*, *M. africanum*, *M. bovis*, and *M. pinnipedii*. Raw sequencing reads were processed using a standardized pipeline developed by our lab: reads were adapter-trimmed with AdapterRemoval v2.3.1[203] and mapped to the reconstructed MTBC ancestor genome using BWA[204] aln v0.7.18 under stringent parameters (-l 1024, minimum mapping quality 30). Duplicates were removed with Picard v2.23.8 and damaged sites were rescaled using mapDamage v2.2.1[205] (for ancient strains only) to minimize the impact of postmortem DNA damage in ancient genomes. Variant calling and genotyping were performed with GATK v4.1.9.0, requiring a minimum of 10 supporting reads per site in modern samples and 4 reads for ancient strains. In ancient metagenomic samples (6 public and ours) we applied some additional filters before generating the consensus fasta file, which are described below. Consensus sequences were generated with BCFtools by integrating called variants into the MTBC ancestor backbone, and genes present in most strains were extracted based on a curated GFF annotation. Genes with >10% missing data (Ns) in a given sample were masked, and genes missing in >50% of samples were excluded. The retained genes were concatenated with GoAlign v0.3.4 into a multiple sequence alignment, which was used to build a maximum likelihood phylogeny with RAxML-ng v1.0.2 (100 bootstrap replicates, under the GTR+FO+G4+I model. This tree was also made including AR0185, which counted with sufficient genome coverage for a de novo reconstruction.

Filtering and phylogenetic analyses of AR0185 and public ancient strains 281U, 386U, 54U, 58U, 64U, 82U

For the individuals with MTBC-positive metagenomic signals, raw fastq reads were trimmed and quality-filtered using AdapterRemoval v2.3.1 and mapped to the computationally reconstructed genome of the MTBC ancestor[200] using BWA v0.7.18 and SAMtools v1.21. Mapping was performed using the following parameters: bwa aln -l 1024 -t 20, followed by filtering for reads with mapping quality >30 and removal of unmapped reads.

Previous work[201] and our own analyses (Nelson et al., manuscript in preparation) show that this mapping strategy often results in substantial off-target environmental noise, particularly problematic at low coverage. To mitigate this issue, we implemented a taxonomic filtering strategy in which only reads classified as *Mycobacterium tuberculosis* by the metagenomic taxonomic classifier KrakenUniq v1.0.2[206] for taxonomic classification. Retained *M. tuberculosis*-assigned reads were then extracted from the BAM files for variant calling. Variant calling was performed using GATK v4.1.9.0 (HaplotypeCaller and GenotypeGVCFs) with the MTBC ancestor genome as reference, a minimum base quality score of 20, and sample ploidy set to 1. We applied stringent post-calling filters to remove indels, heterozygous sites, and conserved positions common to all *Mycobacterium* species. A consensus sequence was reconstructed with BCFtools (commit f27f849) by placing REF or ALT alleles at genotyped positions, and using 'N' at uncalled positions, as described in the previous section.

To minimize the impact of sequencing and PCR errors in the low-coverage data of AR0185—where most positions were supported by only one or two reads—we excluded private variants unique to this strain and genotyped only positions known to be variable in other high-coverage MTBC genomes. This strategy, commonly applied in ancient DNA research and similarly used in our human genome analyses, helps prevent false variants that would artificially extend its terminal branch. As a result, the branch length of AR0185 in the phylogenetic tree is zero, thus reflecting only its placement, based solely on phylogenetically informative variants that are shared with other strains. The resulting FASTA sequence for sample AR0185 was included in a multiple sequence alignment together with 86 high-coverage ancient and modern MTBC genomes representative of the known global diversity and a maximum likelihood phylogenetic tree was computed as described above. Sample AR0185 was placed robustly within the *M. pinnipedii* clade, together with the other previously reported ancient strains from South America.

F. Radiocarbon chronology and Bayesian modelling

F1. Methods for radiocarbon chronology and Bayesian modelling

Calibration and Bayesian modelling were undertaken using the SHCal20 calibration curve[207] in OxCal 4.4[208]. Chronometric data from migrant individuals from each site were included in single-phase, uniform models [individually (Figs. S33-S35) and for the greater UV model (Main Fig. 4b). Kinship (Supplementary Material Section E.7) data was introduced to the PLC model as prior information using calculated values for the number of years (with uncertainty) that separate the dates of death (DOD) between two genetically related individuals[209]. This is 35±26 years for 2nd-3rd degree relationships, with the value incorporated as Interval(N(35, 26)) within a Phase. For individuals where the difference between the dated event and the age at death exceeds 10 years—considering that calibration curve resolution for the period is decadal[207]—corrections were applied to yield the date of death. This applies to five adults (age at death: 35-49 years) dated using the petrous bone, which reflects the first 2-3 years of life[210,211], as well as an additional adult with the same age at death who was dated through a premolar (reflecting the first 9 years of life[27]).

Considering the dated event versus the age at death, a correction of 32-49 years was applied to the five pretrous-dated adults (subtracting 3 years to the lower bound of 35), and 28-49 years to the premolar-dated adult (subtracting 9 years to the lower bound of 35). These corrections were incorporated into the modelling through an adapted Charcoal outlier model[212]. By default, this model [Outlier_Model("Charcoal", Exp(1, -10, 0), U(0, 3), "t")] assumes that measurements are older (enforcing a prior outlier probability of 1 or 100%; see Main Fig. 4b), applying an exponential distribution [Exp(1, -10, 0)] with a time constant (scale) of 1-1,000 years [represented as 100 to 103 or U(0, 3)]. The latter was narrowed to U(1.5, 1.7) to reflect 32-49 years, and U(1.4, 1.7) for 28-49 years. All age estimates are here noted at 95.4% credible/confidence intervals (CI) and rounded to 5 years.

F.2 Site-specific Bayesian models and sensitivity testing

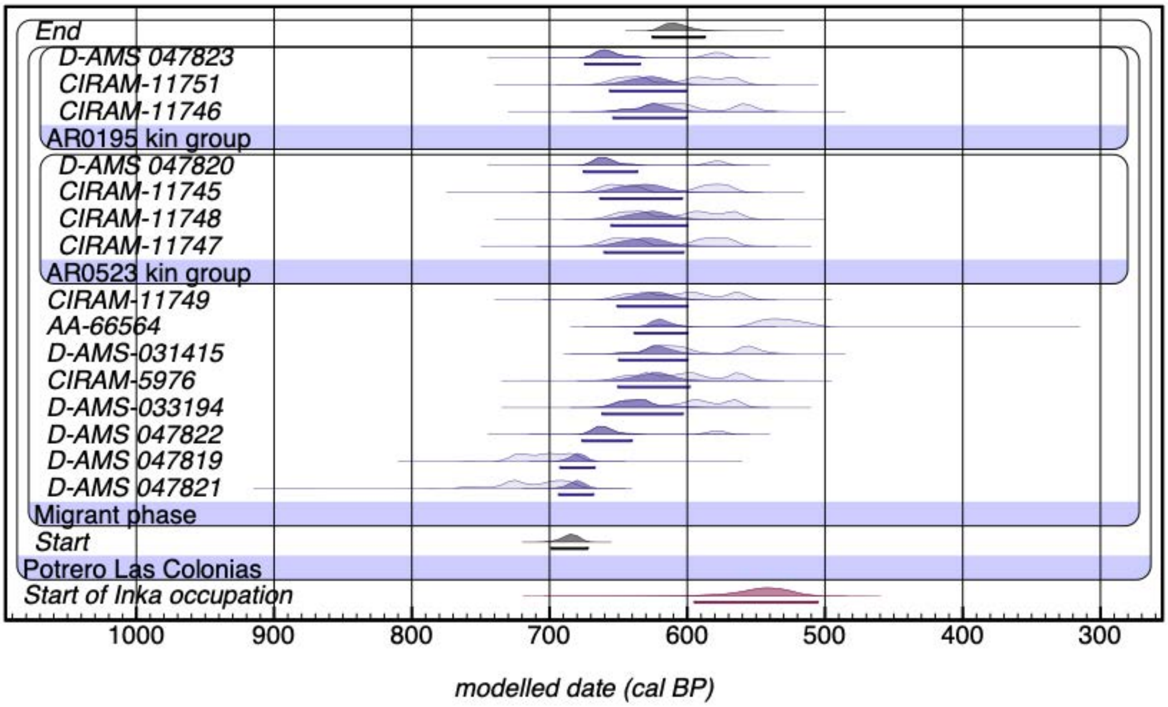

Fig. S33. Bayesian model for the migrant phase at PLC.
The analyses were performed including kinship data (DOD values) as prior information. Bars underneath each distribution denote 95.4% CI. The start of Inka occupation in the region[20], is included at the bottom in red. The start and end estimates are 700-670 cal BP (median 685 cal BP) and 625-585 cal BP (median 610 cal BP), respectively, with a likely duration of between 55 to 100 (median 75 years).

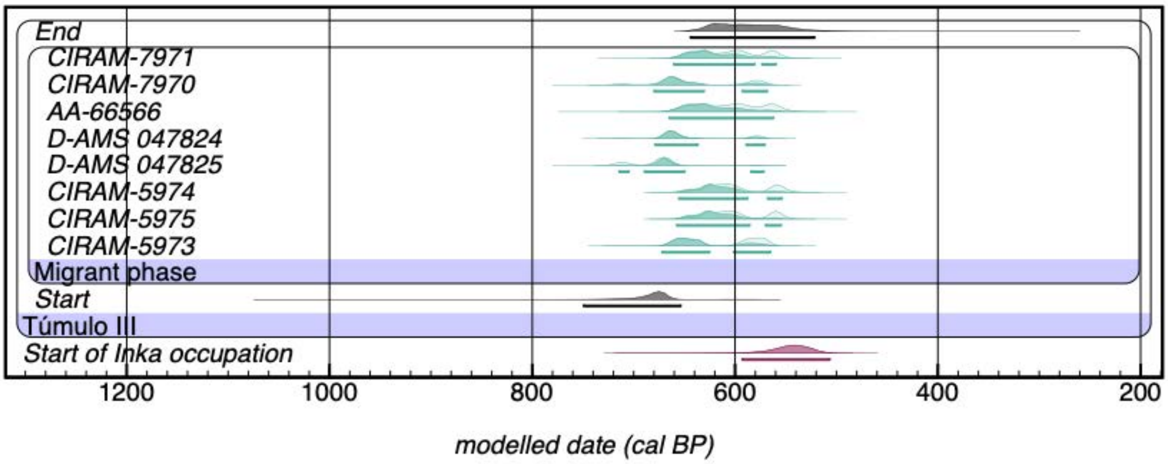

Fig. S34. Bayesian model for the migrant phase at Túmulo III.
The radiocarbon dates are in green following colouring in Main Fig. 4b. Bars underneath each distribution denote 95.4% CI. The start of Inka occupation in the region, as calculated in[20], is included at the bottom in red. The start and end estimates are 750-650 cal BP (median 680 cal BP) and 645-520 cal BP (median 590 cal BP), respectively, with a likely duration of between 20 to 195 (median 90 years).

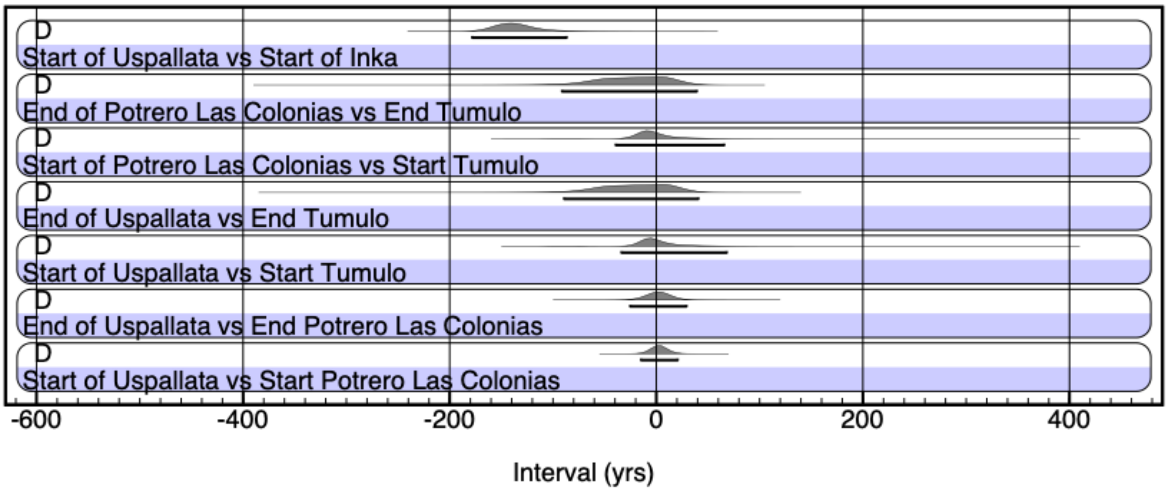

Fig. S35. Consistency of the three Bayesian models.
Probability density functions for the difference ('D') between the start and end boundaries for the three models created (Main Fig. 4b; Figs. S33-34) and the start of the Inka occupation in the region, as calculated by[20]. Black bars underneath each distribution denote 95.4% CI. Results show that apart from the start of the migrant phase at UV and the start of the Inka occupation, there is no significant difference between the modelled outputs as the distributions include zero at 95.4% CI.

G. Paleoclimate work and archives utilized

A detailed review of published papers focused on paleoecological/paleoclimatological research in the subtropical Andes (29°-33°S) was carried out in order to provide the environmental/climatic context for archaeology. While there is an increasing body of recent papers in this area, we considered high Andean records (> 3000 masl) with a sub-centennial to decadal resolution for the focus time period. We also included the lithic influx data from a sediment core located in the Peruvian offshore which reflects El Niño Southern Oscillation (ENSO) activity, since precipitation increases in the subtropical Andes while positive phases of ENSO when coupled with the rainy season (winter)[213] (Main Fig. 5d).

Paleoecological records show strengthened sub-centennial climate variability on a macro-regional scale between 1400-500 years BP with a decreasing trend in precipitation reflected in the main regional records, coinciding with a reduction in the frequency of ENSO events. An increase of cold tolerant and dry taxa pollen percentages in LCA and a predominance of longer than present ice temporal cover in LCH indicate colder and drier than present conditions in this period. A slight amelioration occurred between 850-500 cal yrs BP accompanied by stormier conditions in LCB and a shift to higher ENSO frequency, coinciding with the migrations. Since the migrants were nearly sedentary maize farmers, these conditions would have represented a challenge in terms of climate and at the level of the geomorphic dynamics of the soils suitable for artificially irrigated farming. Drier and more unstable climatic conditions than before under a high variability associated with enhanced ENSO activity prevailed north of 32°S, encompassing UV and the putative areas of migrant origin. The results are given in (Main Fig. 5d).

I. Supplementary References

1. Barberena, R. *et al. Diversity in Socioecological Niches in the Andes* (DISENIA): an isotope-based project. *Antiquity* **97**, e16 (2023).
2. Ramos, V. A. & Folguera, A. Andean flat-slab subduction through time. *Geol. Soc. Lond. Spec. Publ.* **327**, 31–54 (2009).
3. Barberena, R. *et al.* Scale of human mobility in the southern Andes (Argentina and Chile): A new framework based on strontium isotopes. *Am. J. Phys. Anthropol.* **164**, 305–320 (2017).
4. Barberena, R. *et al.* Bioavailable Strontium, Human Paleogeography, and Migrations in the Southern Andes: A Machine Learning and GIS Approach. *Front. Ecol. Evol.* **9**, 584325 (2021).
5. Barberena, R. *et al.* Bioavailable Strontium in the Southern Andes (Argentina and Chile): A Tool for Tracking Human and Animal Movement. *Environ. Archaeol.* **26**, 323–335 (2021).
6. Rusconi, C. Ritos funerarios de los indígenas prehistóricos de Mendoza. *An. Soc. Científica Argent.* **CXLIII**, 97–114 (1947).
7. Rusconi, C. *Poblaciones Pre y Posthispanicas de Mendoza. Volumen I. Etnografía*. (Mendoza, 1961).
8. Rusconi, C. *Poblaciones Pre y Posthispanicas de Mendoza. Volumen II. Antropología*. (Provincia de Mendoza, Mendoza, 1962).
9. Rusconi, C. *Poblaciones Pre y Posthispanicas de Mendoza. Volumen III. Arqueología*. (Provincia de Mendoza, Mendoza, 1962).
10. Novellino, P. S. *et al.* Sitios de entierro de la localidad arqueológica Barrancas, Maipú (Mendoza, Argentina). *Intersecc. En Antropol.* **14**, 271–277 (2013).
11. Barberena, R. *et al.* Esferas de movilidad, sistemas de parentesco e isótopos: Una exploración comparativa para el norte de Mendoza. *Chungará* **54**, 419–438 (2022).
12. Suby, J. A., Novellino, P., Da Peña, G. & Pandiani, C. D. Elongated odontoid process in late Holocene skeletal remains from B6 archaeological site, Mendoza, Argentina. *Int. J. Paleopathol.* **22**, 86–91 (2018).
13. Da Peña Aldao, G., Novellino, P. & Frigolé, C. Caracterización del Túmulo II (Uspallata, Mendoza): Actualización del análisis del contexto funerario. *Comechingonia* **20**, 21–49 (2016).
14. Guevara Batllori, M. D. Paleodemografía y paleopatología humana en los Andes Centrales de Argentina: Estudio bioarqueológico de contextos funerarios del valle de Uspallata (Mendoza) desde el AD 700 al 1500. (Universidad Nacional de Cuyo, Mendoza, 2025).
15. Guevara, D. *et al.* Estructura demográfica, dieta y migración en los Andes del sur: nuevos análisis del sitio Osario Potrero Las Colonias, Uspallata (Mendoza, Argentina). *Intersecc. En Antropol.* **23**, 67–82 (2022).
16. Schobinger, J. El enterratorio de Uspallata-Usina: estudio de su ajuar funrario. *An. Arqueol. Etnología* **XXIX–XXXI**, 67–89 (1974).
17. Schobinger, J. Arqueología del Valle de Uspallata. Provincia de Mendoza (Sinópsis preliminar). *Relac. Soc. Argent. Antropol.* **5**, 71–84 (1971).
18. Bárcena, J. R. Antropología física del Valle de Uspallata, Provincia de Mendoza. *An. Arqueol. Etnología* **29–31**, 109–183 (1976).
19. Durán, V. *et al.* Barrio Ramos I. Funebria y modos de vida en el inicio del período de dominación inca del valle de Uspallata (Mendoza, Argentina). *Relac. Soc. Argent. Antropol.* **XLIII**, 55–86 (2018).
20. Marsh, E. J., Kidd, R., Ogburn, D. & Durán, V. Dating the Expansion of the Inca Empire: Bayesian Models from Ecuador and Argentina. *Radiocarbon* **59**, 117–140 (2017).
21. Knüsel, C. J. & Schotsmans, E. M. J. *The Routledge Handbook of Archaeothanatology: Bioarchaeology of Mortuary Behaviour*. (Routledge, London, 2022). doi:10.4324/9781351030625.
22. Schroeder, S. Secondary disposal of the dead: cross-cultural codes. *World Cult.* **12**, 77–93 (2001).

23. Martínez, G., Flensborg, G. & Bayala, P. D. Human corpse manipulation and the body as symbol: A case study from the Eastern Pampa–Patagonia transition (Argentina) during the Final Late Holocene. *J. Anthropol. Archaeol.* **31**, 215–226 (2012).

24. Berón, M., Di Donato, R. M. & Markán, A. Leather funerary packages: Mortuary practices and differential preservation in a Late Holocene prehispanic cemetery (Pampean region, Argentina). *Quat. Int.* **278**, 51–62 (2012).

25. Duday, H., Le Mort, F. & Tillier, A. M. Archaeothanatology and funeral archaeology. Applicaton to the study of primary single burials. *Anthropologie* **LII**, 235–246 (2014).

26. Buikstra, J. & Ubelaker, D. *Standards for Data Collection from Human Skeletal Remains*. vol. 44 (Archaeological Survey Research Series, Arkansas, 1994).

27. AlQahtani, S. J., Hector, M. P. & Liversidge, H. M. The London atlas of human tooth development and eruption. *Am. J. Phys. Anthropol.* **142**, 481–490 (2010).

28. Schroeder, S. Secondary Disposal of the Dead: Cross-Cultural Codes. *World Cult.* **12**, 77–93 (2001).

29. Barberena, R. *et al.* Multi-isotopic and morphometric evidence for the migration of farmers leading up to the Inka conquest of the southern Andes. *Sci. Rep.* **10**, 21171 (2020).

30. Gil, A. F. *et al.* Isotopic evidence on human bone for declining maize consumption during the little ice age in central western Argentina. *J. Archaeol. Sci.* **49**, 213–227 (2014).

31. Gil, A. F. *et al.* Stable isotopes and maize consumption in central western Argentina. *Int. J. Osteoarchaeol.* **19**, 215–236 (2009).

32. Andreasen, R. & Thomsen, E. Strontium Is Released Rapidly From Agricultural Lime–Implications for Provenance and Migration Studies. *Front. Ecol. Evol.* **8**, (2021).

33. Grayson, D. K. *Quantitative Zooarchaeology. Topics in the Analysis of Archaeological Faunas*. (Academic Press, New York, 1984).

34. Gifford-Gonzalez, D. *An Introduction to Zooarchaeology*. (Springer International Publishing, Cham, 2018). doi:10.1007/978-3-319-65682-3.

35. Scheuer, L. & Black, S. *Developmental Juvenile Osteology*. (Academic Press, London, 2000).

36. Meindl, R. & Lovejoy, C. Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral anterior sutures. *Am. J. Phys. Anthropol.* **68**, 57–66 (1985).

37. Brooks, S. & Suchey, J. Skeletal Age Determination Based on the Os Pubis: A Comparison of the Acsádi-Nemeskéri and Suchey-Brooks Methods. *Hum. Evol.* **5**, 227–238 (1990).

38. O'Donnell, L., Hill, E. C., Anderson, A. S. A. & Edgar, H. J. H. Cribra orbitalia and porotic hyperostosis are associated with respiratory infections in a contemporary mortality sample from New Mexico. *Am. J. Phys. Anthropol.* **173**, 721–733 (2020).

39. Larsen, C. S. Biological Changes in Human Populations with Agriculture. *Annu. Rev. Anthropol.* **24**, 185–213 (1995).

40. Hillson, S. *Teeth*. (Cambridge University Press, New York, 2005).

41. Acosta, C. B. Un acercamiento al pasado en Barrancas, Mendoza: estudio de dieta y estrés nutricional en el Holoceno tardío. (Universidad Nacional de Cuyo, Mendoza, 2025).

42. Klaus, H. D. & Lynnerup, N. Chapter 5 - Abnormal Bone: Considerations for Documentation, Disease Process Identification, and Differential Diagnosis. in *Ortner's Identification of Pathological Conditions in Human Skeletal Remains (Third Edition)* (ed. Buikstra, J. E.) 59–89 (Academic Press, San Diego, 2019). doi:10.1016/B978-0-12-809738-0.00005-3.

43. DeWitte, S. N. Differential survival among individuals with active and healed periosteal new bone formation. *Int. J. Paleopathol.* **7**, 38–44 (2014).

44. Redfern, R. & Roberts, C. A. Trauma. in *Ortner's Identification of Pathological Conditions in Human Skeletal Remains* (ed. Buikstra, J. E.) 211–284 (Academic Press, New York, 2019).

45. Weston, D. A. Nonspecific Infection in Paleopathology: Interpreting Periosteal Reactions. in *A*

Companion to Paleopathology 492–512 (John Wiley & Sons, Ltd, 2011). doi:10.1002/9781444345940.ch27.

46. Office of the High Commissioner for Human Rights. *Istanbul Protocol: Manual on the Effective Investigation and Documentation of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*. https://www.ohchr.org/en/publications/policy-and-methodological-publications/istanbul-protocol-manual-effective-0 (2022).
47. Appleby, J., Thomas, R. & Buikstra, J. Increasing confidence in paleopathological diagnosis – Application of the Istanbul terminological framework. *Int. J. Paleopathol.* **8**, 19–21 (2015).
48. Lovell, N. Trauma analysis in paleopathology. *Yearb. Phys. Anthropol.* **40**, 139–170 (1997).
49. Mays, S. A. A palaeopathological study of Colles' fracture. *Int. J. Osteoarchaeol.* **16**, 415–428 (2006).
50. Judd, M. A. The parry problem. *J. Archaeol. Sci.* **35**, 1658–1666 (2008).
51. Armelagos, G. J., Goodman, A. H., Harper, K. N. & Blakey, M. L. Enamel hypoplasia and early mortality: Bioarcheological support for the Barker hypothesis. *Evol. Anthropol. Issues News Rev.* **18**, 261–271 (2009).
52. Nikiforuk, G. & Fraser, D. The etiology of enamel hypoplasia: A unifying concept. *J. Pediatr.* **98**, 888–893 (1981).
53. Masson, M. *et al.* Osteological and Biomolecular Evidence of a 7000-Year-Old Case of Hypertrophic Pulmonary Osteopathy Secondary to Tuberculosis from Neolithic Hungary. *PLoS ONE* **8**, e78252 (2013).
54. Pin, C., Briot, D., Bassin, C. & Poitrasson, F. Concominant Separation of strontium and samarium–neodymium for isotopic analysis in silicate samples, based on specific extraction chromatography. *Anal. Chim. Acta* **298**, 209–217 (1994).
55. Waight, T., Baker, J. & Peate, D. Sr isotope ratio measurements by double-focusing MC-ICPMS: techniques, observations and pitfalls. *Int. J. Mass Spectrom.* **221**, 229–244 (2002).
56. Price, T. D., Blitz, J., Burton, J. & Ezzo, J. A. Diagenesis in prehistoric bone: Problems and solutions. *J. Archaeol. Sci.* **19**, 513–529 (1992).
57. Bentley, R. A. Strontium Isotopes from the Earth to the Archaeological Skeleton: A Review. *J. Archaeol. Method Theory* **13**, 135–187 (2006).
58. Kendall, C., Eriksen, A. M. H., Kontopoulos, I., Collins, M. J. & Turner-Walker, G. Diagenesis of archaeological bone and tooth. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **491**, 21–37 (2018).
59. DeNiro, M. J. Postmortem preservation and alteration of in vivo bone collagen isotope ratios in relation to palaeodietary reconstruction. *Nature* **317**, 806–809 (1985).
60. van Klinken, G. Bone Collagen Quality Indicators for Palaeodietary and Radiocarbon Measurements. *J. Archaeol. Sci.* **26**, 687–695 (1999).
61. France, C. A. M., Sugiyama, N. & Aguayo, E. Establishing a preservation index for bone, dentin, and enamel bioapatite mineral using ATR-FTIR. *J. Archaeol. Sci. Rep.* **33**, 102551 (2020).
62. Pezo-Lanfranco, L. & Colonese, A. C. The role of farming and fishing in the rise of social complexity in the Central Andes: a stable isotope perspective. *Sci. Rep.* **14**, 4582 (2024).
63. Bataille, C. P., Crowley, B. E., Wooller, M. J. & Bowen, G. J. Advances in global bioavailable strontium isoscapes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **555**, 109849 (2020).
64. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
65. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing (2023).
66. Genuer, R., Poggi, J.-M. & Tuleau-Malot, C. VSURF function | R Documentation. https://www.rdocumentation.org/packages/VSURF/versions/1.1.0/topics/VSURF (2019).
67. Funck, J., Bataille, C., Rasic, J. & Wooller, M. A bio-available strontium isoscape for eastern Beringia: a tool for tracking landscape use of Pleistocene megafauna. *J. Quat. Sci.* **36**, 76–90 (2021).

68. Hartmann, J. & Moosdorf, N. The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochem. Geophys. Geosystems* **13**, (2012).

69. Mooney, W. D., Laske, G. & Masters, T. G. CRUST 5.1: A global crustal model at 5° × 5°. *J. Geophys. Res. Solid Earth* **103**, 727–747 (1998).

70. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).

71. Zomer, R. J., Trabucco, A., Bossio, D. A. & Verchot, L. V. Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agric. Ecosyst. Environ.* **126**, 67–80 (2008).

72. Chien, C. *et al.* Effects of African dust deposition on phytoplankton in the western tropical Atlantic Ocean off Barbados. *Glob. Biogeochem. Cycles* **30**, 716–734 (2016).

73. Brahney, J. *et al.* Ecological changes in two contrasting lakes associated with human activity and dust transport in western Wyoming. *Limnol. Oceanogr.* **60**, 678–695 (2015).

74. Bataille, C. P. *et al.* Triple sulfur-oxygen-strontium isotopes probabilistic geographic assignment of archaeological remains using a novel sulfur isoscape of western Europe. *PLOS ONE* **16**, e0250383 (2021).

75. Jarvis, A., Reuter, A., Nelson, A. & Guevara, E. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database. CGIAR CSI Consort. *Spat. Inf.* 1–9 (2008).

76. Balmino, G., Vales, N., Bonvalot, S. & Briais, A. Spherical harmonic modelling to ultra-high degree of Bouguer and isostatic anomalies. *J. Geod.* **86**, 499–520 (2012).

77. Börker, J., Hartmann, J., Amann, T. & Romero-Mujalli, G. Terrestrial Sediments of the Earth: Development of a Global Unconsolidated Sediments Map Database (GUM). *Geochem. Geophys. Geosystems* **19**, 997–1024 (2018).

78. Poggio, L. *et al.* SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**, 217–240 (2021).

79. Bataille, C. P. *et al.* A geostatistical framework for predicting variations in strontium concentrations and isotope ratios in Alaskan rivers. *Chem. Geol.* **389**, 1–15 (2014).

80. Wunder, M. B. Determining geographic patterns of migration and dispersal using stable isotopes in keratins. *J. Mammal.* **93**, 360–367 (2012).

81. Ma, C., Vander Zanden, H. B., Wunder, M. B. & Bowen, G. J. assignR: An r package for isotope-based geographic assignment. *Methods Ecol. Evol.* **11**, 996–1001 (2020).

82. Sealy, J. C., Johnson, M., Richards, M., Nehlich, O., *et al.* Comparison of two methods of extracting bone collagen for stable carbon and nitrogen isotope analysis: Comparing whole bone demineralization with gelatinization and ultrafiltration. *J. Archaeol. Sci.* **47**, 64–69 (2014).

83. Ambrose, S. H. Preparation and characterization of bone and tooth collagen for isotopic analysis. *J. Archaeol. Sci.* **17**, 431–451 (1990).

84. Lee-Thorp, J. *et al.* Exploring problems and opportunities offered by down-scaling sample sizes for carbon isotope analyses of fossils. *Bull Soc Geol Fr* **168**, 767–773 (1997).

85. Luyt, J., Hare, V. J. & Sealy, J. The relationship of ungulate δ13C and environment in the temperate biome of southern Africa, and its palaeoclimatic application. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **514**, 282–291 (2019).

86. Somerville, A. D. *et al.* Diet and gender in the Tiwanaku colonies: Stable isotope analysis of human bone collagen and apatite from Moquegua, Peru. *Am. J. Phys. Anthropol.* **158**, 408–422 (2015).

87. Ambrose, S. & Norr, L. Experimental Evidence for the Relationship of Carbon Isotope Ratios of Whole Diet and Dietary Protein to Those of Bone Collagen and Carbonate. in *Prehistoric Human Bone: Archaeology at the Molecular Level* (eds. Lambert, J. B. & Grupe, G.) 1–38 (Springer-Verlag, Berlin, 1993).

88. Finucane, B. C., Agurto, P. M. & Isbell, W. H. Human and animal diet at Conchopata, Peru: stable isotope evidence for maize agriculture and animal management practices during the Middle Horizon. *J. Archaeol. Sci.* **33**, 1766–1776 (2006).

89. Finucane, B. C. Maize and Sociopolitical Complexity in the Ayacucho Valley, Peru. *Curr. Anthropol.* **50**, 535–545 (2009).

90. Slovak, N. M. & Paytan, A. Fisherfolk and farmers: Carbon and nitrogen isotope evidence from Middle Horizon Ancón, Peru. *Int. J. Osteoarchaeol.* **21**, 253–267 (2011).

91. Tung, T. A. & Knudson, K. J. Stable isotope analysis of a pre-Hispanic Andean community: Reconstructing pre-Wari and Wari era diets in the hinterland of the Wari empire, Peru. *Am. J. Phys. Anthropol.* **165**, 149–172 (2018).

92. Alfonso-Durruty, M., Troncoso, A., Larach, P., Becker, C. & Misarti, N. Maize (Zea mays) consumption in the southern andes (30°-31° S. Lat): Stable isotope evidence (2000 BCE-1540 CE). *Am. J. Phys. Anthropol.* **164**, 148–162 (2017).

93. Falabella, F., Planella, M. T. & Tykot, R. H. El Maíz (Zea mays) en el Mundo Prehispánico de Chile Central. *Lat. Am. Antiq.* **19**, 25–46 (2008).

94. Killian Galván, V. A. Models for paleodietary research: Three case-studies from arid and semi-arid environments in Northwest Argentina. *J. Archaeol. Sci. Rep.* **18**, 608–616 (2018).

95. Tung, T. A., Dillehay, T. D., Feranec, R. S. & DeSantis, L. R. G. Early specialized maritime and maize economies on the north coast of Peru. *Proc. Natl. Acad. Sci.* **117**, 32308–32319 (2020).

96. Tykot, R. H., Burger, R. L. & Van Der Merwe, N. J. The Importance of Maize in Initial Period and Early Horizon Peru. in *Histories of Maize* 187–197 (Elsevier, 2006). doi:10.1016/B978-012369364-8/50266-7.

97. Ericson, J. E., West, M., Sullivan, C. H. & Krueger, H. W. The development of maize agriculture in the Viru Valley, Peru. in *The Chemistry of Prehistoric Human Bone* (ed. Price, T. Douglas, M.) 68–104 (Cambridge University Press, Cambridge, 1989).

98. Santana-Sagredo, F. *et al.* ‘White gold’ guano fertilizer drove agricultural intensification in the Atacama Desert from ad 1000. *Nat. Plants* 1–7 (2021) doi:10.1038/s41477-020-00835-4.

99. Turner, B. L. & Armelagos, G. J. Diet, residential origin, and pathology at Machu Picchu, Peru. *Am. J. Phys. Anthropol.* **149**, 71–83 (2012).

100. Burger, R., Lee-Thorp, J. & Van Der Merwe, N. J. Rite and crop in the Inca state revisited: an isotopic perspective from Machu Picchu and beyond. in *The 1912 Yale Peruvian Scientific Expedition Collections from Machu Picchu* (eds. Burger, R. L. & Salazar, L.) 119–137 (Yale University Press, New Haven, 2003).

101. Berryman, C. A. & Blom, D. E. Transforming Food in the Tiwanaku Heartland: Maize and the Rise and Fall of Sociopolitical Complexity in the Southern Lake Titicaca Basin. in *Foodways of the Ancient Andes: Transforming Diet, Cuisine, and Society* (eds. Blom, D. E. & Alfonso-Durruty, M. P.) 111–132 (University of Arizona Press, 2023).

102. Hansen, H. B. *et al.* Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. *PLOS ONE* **12**, e0170940 (2017).

103. Damgaard, P. B. *et al.* Improving access to endogenous DNA in ancient bones and teeth. *Sci. Rep.* **5**, 11184 (2015).

104. Pinhasi, R. *et al.* Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. *PLOS ONE* **10**, e0129102 (2015).

105. Dabney, J. & Meyer, M. Extraction of Highly Degraded DNA from Ancient Bones and Teeth. in *Shapiro, B., Barlow, A., Heintzman, P., Hofreiter, M., Paijmans, J., Soares, A. (eds) Ancient DNA. Methods in Molecular Biology* vol. 1963 (Humana Press, New York, 2019).

106. Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth. *Nat. Protoc.* **2**, 1756–1762 (2007).

107. Allentoft, M. E. *et al.* Population Genomics of Stone Age Eurasia. 2022.05.04.490594 Preprint at https://doi.org/10.1101/2022.05.04.490594 (2022).

108. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20130624 (2015).

109. Kapp, J. D., Green, R. E. & Shapiro, B. A Fast and Efficient Single-stranded Genomic Library

Preparation Method Optimized for Ancient DNA. *J. Hered.* **112**, 241–249 (2021).

110. Seguin-Orlando, A. *et al.* Heterogeneous Hunter-Gatherer and Steppe-Related Ancestries in Late Neolithic and Bell Beaker Genomes from Present-Day France. *Curr. Biol.* (2021) doi:10.1016/j.cub.2020.12.015.

111. Gamba, C. *et al.* Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Mol. Ecol. Resour.* **16**, 459–469 (2016).

112. Fages, A. *et al.* Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. *Cell* **177**, 1419-1435.e31 (2019).

113. Meyer, M. & Kircher, M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).

114. Arbor Biosciences. myBaits® Hybridization Capture for Targeted NGS – Manual v5.02: High Sensitivity Protocol. Arbor Biosciences, March 2022.

115. Roche. KAPA HiFi HotStart ReadyMix PCR Kit – Technical Data Sheet. Roche Sequencing and Life Science, Basel, Switzerland. Updated 2020.

116. Fellows Yates, J. A. *et al.* Reproducible, portable, and efficient ancient genome reconstruction with nf-core/eager. *PeerJ* **9**, e10947 (2021).

117. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).

118. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

119. Peltzer, A. *et al.* EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).

120. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147–147 (1999).

121. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

122. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

123. Neukamm, J., Peltzer, A. & Nieselt, K. DamageProfiler: fast damage pattern calculation for ancient DNA. *Bioinformatics* **37**, 3652–3653 (2021).

124. Jun, G., Wing, M. K., Abecasis, G. R. & Kang, H. M. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* **25**, 918–925 (2015).

125. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).

126. Fu, Q. *et al.* A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes. *Curr. Biol.* **23**, 553–559 (2013).

127. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).

128. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

129. Lichtenstein, F., Antoneli, F. & Briones, M. R. S. MIA: Mutual Information Analyzer, a graphic user interface program that calculates entropy, vertical and horizontal mutual information of molecular sequence sets. *BMC Bioinformatics* **16**, 409 (2015).

130. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).

131. Weissensteiner, H. *et al.* Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome Res.* **31**, 309–316 (2021).

132. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).

133. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).

134. Nakatsuka, N. *et al.* ContamLD: estimation of ancient nuclear DNA contamination using breakdown of linkage disequilibrium. *Genome Biol.* **21**, 199 (2020).

135. The 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

136. Patterson, N. *et al.* Ancient Admixture in Human History. *Genetics* **192**, 1065–1093 (2012).

137. Rohland, N. *et al.* Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. *Genome Res.* **32**, 2068–2078 (2022).

138. Robinson, J. T. Integrative genomics viewer. (2011).

139. Campelo dos Santos, A. L. *et al.* Genomic evidence for ancient human migration routes along South America's Atlantic coast. *Proc. R. Soc. B Biol. Sci.* **289**, 20221078 (2022).

140. Bongers, J. L. *et al.* Integration of ancient DNA with transdisciplinary dataset finds strong support for Inca resettlement in the south Peruvian coast. *Proc. Natl. Acad. Sci.* **117**, 18359–18368 (2020).

141. Capodiferro, M. R. *et al.* Archaeogenomic distinctiveness of the Isthmo-Colombian area. *Cell* (2021) doi:10.1016/j.cell.2021.02.040.

142. Ferraz, T. *et al.* Genomic history of coastal societies from eastern South America. *Nat. Ecol. Evol.* **7**, 1315–1330 (2023).

143. Kennett, D. J. *et al.* South-to-north migration preceded the advent of intensive farming in the Maya region. *Nat. Commun.* **13**, 1530 (2022).

144. Kennett, D. J. *et al.* Archaeogenomic evidence reveals prehistoric matrilineal dynasty. *Nat. Commun.* **8**, 14115 (2017).

145. Lindo, J. *et al.* The genetic prehistory of the Andean highlands 7000 years BP though European contact. *Sci. Adv.* **4**, eaau4921 (2018).

146. Lindo, J. *et al.* The genomic prehistory of the Indigenous peoples of Uruguay. *PNAS Nexus* **1**, pgac047 (2022).

147. Moreno-Mayar, J. V. *et al.* Early human dispersals within the Americas. *Science* **362**, eaav2621 (2018).

148. Nakatsuka, N. *et al.* Ancient genomes in South Patagonia reveal population movements associated with technological shifts and geography. *Nat. Commun.* **11**, 3868 (2020).

149. Nakatsuka, N. *et al.* Genetic continuity and change among the Indigenous peoples of California. *Nature* 1–8 (2023) doi:10.1038/s41586-023-06771-5.

150. Nakatsuka, N. *et al.* A Paleogenomic Reconstruction of the Deep Population History of the Andes. *Cell* (2020) doi:10.1016/j.cell.2020.04.015.

151. Pedersen, M. W. *et al.* Ancient Human Genomes and Environmental DNA from the Cement Attaching 2,000-Year-Old Head Lice Nits. *Mol. Biol. Evol.* **39**, msab351 (2022).

152. Posth, C. *et al.* Reconstructing the Deep Population History of Central and South America. *Cell* **175**, 1185-1197.e22 (2018).

153. Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **349**, aab3884 (2015).

154. Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**, 225–229 (2014).

155. Salazar, L. *et al.* Insights into the genetic histories and lifeways of Machu Picchu's occupants. *Sci. Adv.* **9**, eadg3377 (2023).

156. Villa-Islas, V. *et al.* Demographic history and genetic structure in pre-Hispanic Central Mexico.

Science **380**, eadd6142 (2023).

157. De La Fuente, C. *et al.* Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. *Proc. Natl. Acad. Sci.* **115**, (2018).

158. de la Fuente Castro, C. *et al.* The Genomic and Cultural Diversity of the Inka Qhapaq Hucha Ceremony in Chile and Argentina. *Genome Biol. Evol.* **16**, evae196 (2024).

159. Harris, D. N. *et al.* Evolutionary genomic dynamics of Peruvians before, during, and after the Inca Empire. *Proc. Natl. Acad. Sci.* **115**, E6526–E6535 (2018).

160. Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).

161. Bergström, A. *et al.* Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, (2020).

162. Andrews, S. FastQC: a quality control tool for high throughput sequence data.

163. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

164. Genovese, G., Handsaker, R. E., Li, H., Kenny, E. E. & McCarroll, S. A. Mapping the Human Reference Genome's Missing Sequence by Three-Way Admixture in Latino Genomes. *Am. J. Hum. Genet.* **93**, 411–421 (2013).

165. Zhang, F. *et al.* Ancestry-agnostic estimation of DNA sample contamination from sequence reads. *Genome Res.* **30**, 185–194 (2020).

166. Choin, J. *et al.* Genomic insights into population history and biological adaptation in Oceania. *Nature* **592**, 583–589 (2021).

167. Kwong, A. M. *et al.* Robust, flexible, and scalable tests for Hardy–Weinberg equilibrium across diverse ancestries. *Genetics* **218**, iyab044 (2021).

168. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).

169. Luisi, P. *et al.* Fine-scale genomic analyses of admixed individuals reveal unrecognized genetic ancestry components in Argentina. *PLOS ONE* **15**, e0233808 (2020).

170. Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).

171. Homburger, J. R. *et al.* Genomic Insights into the Ancestry and Demographic History of South America. *PLOS Genet.* **11**, e1005602 (2015).

172. Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian population history. *Nature* **461**, 489–494 (2009).

173. Swapan, M. & Reich, D. The Allen Ancient DNA Resource (AADR): A curated compendium of ancient human genomes. https://doi.org/10.7910/DVN/FFIDCW.

174. Luisi, P. *et al.* Fine-scale genomic analyses of admixed individuals reveal unrecognized genetic ancestry components in Argentina. *PLOS ONE* **15**, e0233808 (2020).

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

176. Motti, J. M. B. *et al.* Ancient mitogenomes from the Southern Pampas of Argentina reflect local differentiation and limited extra-regional linkages after rapid initial colonization. *Am. J. Biol. Anthropol.* **181**, 216–230 (2023).

177. García, A. *et al.* Ancient and modern mitogenomes from Central Argentina: new insights into population continuity, temporal depth and migration in South America. *Hum. Mol. Genet.* **30**, 1200–1217 (2021).

178. Tamburrini, C. Diversidad genética de muestras arqueológicas humanas del nordeste de la provincia del Chubut (Patagonia argentina) durante el Holoceno (6000-200AP). *Arqueología* **30**, 14390 (2024).

179. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

180. Soares, P. *et al.* Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *Am. J. Hum. Genet.* **84**, 740–759 (2009).

181. Ho, S. Y. W. & Duchêne, S. Molecular-clock methods for estimating evolutionary rates and timescales. *Mol. Ecol.* **23**, 5947–5965 (2014).

182. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).

183. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

184. Ralf, A., Montiel González, D., Zhong, K. & Kayser, M. Yleaf: Software for Human Y-Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. *Mol. Biol. Evol.* **35**, 1291–1294 (2018).

185. Paz Sepúlveda, P. B. *et al.* Human Y chromosome sequences from Q Haplogroup reveal a South American settlement pre-18,000 years ago and a profound genomic impact during the Younger Dryas. *PLOS ONE* **17**, e0271971 (2022).

186. Strobl, C. *et al.* Evaluation of mitogenome sequence concordance, heteroplasmy detection, and haplogrouping in a worldwide lineage study using the Precision ID mtDNA Whole Genome Panel. *Forensic Sci. Int. Genet.* **42**, 244–251 (2019).

187. Llamas, B. *et al.* Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* **2**, e1501385 (2016).

188. Roca-Rada, X. *et al.* Ancient mitochondrial genomes from the Argentinian Pampas inform the early peopling of the Southern Cone of South America. *iScience* **24**, 102553 (2021).

189. Alaçamlı, E. *et al.* READv2: advanced and user-friendly detection of biological relatedness in archaeogenomics. *Genome Biol.* **25**, 216 (2024).

190. Popli, D., Peyrégne, S. & Peter, B. M. KIN: a method to infer relatedness from low-coverage ancient DNA. *Genome Biol.* **24**, 10 (2023).

191. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

192. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015-0047–8 (2015).

193. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).

194. De La Fuente Castro, C., Motti, J., Arencibia, V. & Luisi, P. Tales from the end of the world: three decades of paleogenetic research in Patagonia. *Hum. Popul. Genet. Genomics* 1–34 (2024) doi:10.47248/hpgg2404010003.

195. Lawson, D. J., Van Dorp, L. & Falush, D. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat. Commun.* **9**, 3258 (2018).

196. Skoglund, P. *et al.* Genetic evidence for two founding populations of the Americas. *Nature* **525**, 104–108 (2015).

197. Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* **12**, 5425 (2021).

198. Fournier, R., Tsangalidou, Z., Reich, D. & Palamara, P. F. Haplotype-based inference of recent effective population size in modern and ancient DNA samples. *Nat. Commun.* **14**, 7945 (2023).

199. Hübler, R. *et al.* HOPS: automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biol.* **20**, 280 (2019).

200. Comas, I. *et al.* Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nat. Genet.* **42**, 498–503 (2010).

201. Berger, D. J. *et al.* Whole-genome sequencing of Schistosoma mansoni reveals extensive diversity with limited selection despite mass drug administration. *Nat. Commun.* **12**, 4776 (2021).

202. Bos, K. I. *et al.* Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* **514**, 494–497 (2014).

203. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).

204. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

205. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).

206. Breitwieser, F. P., Baker, D. N. & Salzberg, S. L. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. *Genome Biol.* **19**, 198 (2018).

207. Hogg, A. G. *et al.* SHCal20 Southern Hemisphere Calibration, 0–55,000 Years cal BP. *Radiocarbon* **62**, 759–778 (2020).

208. Bronk Ramsey, C. Bayesian Analysis of Radiocarbon Dates. *Radiocarbon* **51**, 337–360 (2009).

209. Sedig, J. W., Olalde, I., Patterson, N., Harney, É. & Reich, D. Combining ancient DNA and radiocarbon dating data to increase chronological accuracy. *J. Archaeol. Sci.* **133**, 105452 (2021).

210. Kontopoulos, I. *et al.* Histological and stable isotope analysis of archeological bones from St. Rombout's cemetery (Mechelen, Belgium): Intrasite, intraindividual, and intrabone variability. *Int. J. Osteoarchaeol.* **32**, 1142–1156 (2022).

211. Jørkov, M. L. S., Heinemeier, J. & Lynnerup, N. The petrous bone—A new sampling site for identifying early dietary patterns in stable isotopic studies. *Am. J. Phys. Anthropol.* **138**, 199–209 (2009).

212. Bronk Ramsey, C. Dealing with Outliers and Offsets in Radiocarbon Dating. *Radiocarbon* **51**, 1023–1045 (2009).

213. Montecinos, A. & Aceituno, P. Seasonality of the ENSO- related rainfall variability in central Chile and associated circulation anomalies. *J. Clim.* **16**, (2003).