What uniparental genes tell us about the prehistoric human colonization of the Americas

Vicente M Cabrera[1,2*]

[1]*Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, San Cristóbal de La Laguna, Spain*

[*]Corresponding author: vicente.vca811@gmail.com

ABSTRACT

Introduction From the perspective of uniparental markers, the view of the human prehistoric settlement of America is that it resulted from a single main migration after the Last Glacial Maximum, following a period of prolonged stasis in Beringia. These pioneers rapidly expanded across a continent previously uninhabited by humans, forming isolated population groups that became genetically distinct. Ancient DNA studies and whole genome analyses have confirmed this view, revealing a remarkable complexity in the ancestries of the involved populations.

Methods In this work, I have reanalyzed all publicly available mitochondrial DNA and Y-chromosome haplogroups in the American population using simple phylogenetic and phylogeographic methodologies.

Results The results provide a genetic perspective opposed to the currently established model. The arrival of the American settlers occurred more than 30,000 years ago, preceding the LGM. As genomic studies have uncovered, at least two Asian populations contributed to the ancestry of the immigrant population. Low population density and climatic deterioration led to a long period of demographic eclipse, during which small bands of hunter-gatherers made long journeys in search of favorable niches. After the LGM, the climate improved, and demographic expansions occurred in multiple independent centers. The founding and expansion ages of the uniparental haplogroups indicate that these centers were in South America, particularly the Colombian isthmus, the Andean region, the Southern Cone, and the Amazon. Subsequent dispersals occurred in North America, one involving mitochondrial haplogroups A2, C1, and D1, but not B2, and another, of lesser magnitude, represented by the expansion of haplogroups C4c and X2a.

Conclusions This work offers a previously unexplored model for the colonization of the Americas.

Keywords Mitochondrial DNA, Y-chromosome, American human settlement

INTRODUCTION

The interpretation of the prehistoric peopling of the Americas has been a multidisciplinary enterprise [1,2]. From a molecular genetics perspective, the first foundations of these interpretations are based on the information provided by the uniparentally inherited mitochondrial DNA (mtDNA) and the non-recombining portion of the Y-chromosome (NRY) on extant populations [3,4]. As ancient DNA (aDNA) samples are a mixture of endogenous aDNA and contaminating DNA, results on human samples were first questioned. Fortunately, new sequencing technologies have overcome this problem [5], and current aDNA-based analyses have added an invaluable timing perspective to these interpretations [6]. Later, technological improvements have made it possible to sequence entire genomes in present-day and fossil individuals. Analyses of these genomes have revealed extinct ancestries and genetically well-differentiated ancestries from the past that, admixed in different proportions, persist in current Native American populations along the American continent. This new information has improved previous hypotheses and formulated new ones [7].

Thus, some conclusions have achieved general acceptance: 1) The Asiatic origin of the Native American ancestors; 2) The first migration, in one or several waves, occurred as a rapid movement towards the south of the double continent; 3) The initial low population density led to strong genetic differentiation geographically structured; 4) Subsequent migrations reshaped the genetic landscape of North and South America. Necessary for the first was the entrance from Beringia of several migratory waves that originated the settlement of Paleo- and Neo-Eskimos; 5) Time comparisons of population genetic profiles since post-Neolithic to present-day have uncovered an essential genetic continuity (although see Llamas et al. [8].

However, other propositions are still under debate, such as: 1) the precise Asian origin of the Native American ancestors; 2) the routes followed by these pioneers as they expanded throughout the Americas; and, perhaps, the most contentious question, 3) when these first migrations occurred. Although the most recent studies based on complete genomes favor a post-Glacial dispersal [9], with or without a prolonged standstill in Beringia [10], the first works with mitochondrial DNA favored an early entry, around 30 kya before the Last Glacial Maximum (LGM), avoiding the problem of having to cross the ice sheets of North America [11,12].

These age discrepancies are due to the molecular evolutionary rates used in each case. Unfortunately, the rate of molecular evolution is not constant; it depends on several parameters, one of which is the adequate population size (Ne), which has changed dramatically from the Paleolithic to present-day [13]. Therefore, like radiocarbon dating, molecular rates must be calibrated to obtain more accurate ages. One way to do this is by using ancient genomes and/or mitogenomes, but in the case of the American colonization, all the extracted aDNA postdates the LGM [6]. The option is to use evolutionary rates calibrated with Paleolithic fossils to date Paleolithic events [14]. Despite their utility, whole-genome studies barely mention

uniparental markers [15]. However, see [9,16]. The underlying reason for this abandonment is that these markers only offer monogenic insights, and only the set of all genes provides congruent results. In my opinion, this assumption is flawed. The leading cause of genetic differentiation in estimating phylogenies and genetic distances between populations or species is isolation. Isolation affects whole genomes and does not discriminate by sex. Populations differentiate genetically, due to the loss of previous polymorphisms, by fixation or extinction, and the accumulation of new polymorphisms by mutation in a time-dependent way. Thus, any DNA fragment, including those coding for uniparental markers, should be valid to realize genetic comparisons whenever they have accumulated enough genetic differentiation. The generalized use of DNA barcoding in molecular taxonomy [17] validates my previous assertion. Isolation is broken by migration/introgression, affecting uniparental markers and autosomes differently. Whether uniparental markers persist only in their original populations or both endure in the original and recipient populations, their coalescence node—the distance to their most recent common ancestor—will remain unchanged. However, if the introgressed fragments are not thoroughly masked for mixed autosomes, they will influence the genetic distances obtained between the populations/species involved. A paradigmatic example is the controversial phylogenetic relationships obtained between modern humans and their closest relatives, Neanderthals and Denisovans. While mtDNA and NRY sequences showed modern humans and Neanderthals as sister lineages [18,19], autosomes joined Denisovan and Neanderthal as the closest pair, leaving modern humans as an out-group [18]. Contrary to the widespread acceptance that the proper relationship between hominids is that obtained with the autosomes, I have suggested that the valid phylogeny is deduced from the uniparental markers. In contrast, the one obtained from the autosomes resulted from later introgressions between the three groups of hominins [20]. Taking into account the above-mentioned particularities of the uniparental markers, in this work, I perform a thorough reanalysis of all published mitogenomes and the most recently published Y-chromosome results, using phylogenetic and phylogeographic tools, to reconsider and shed new light on the hypotheses formulated to explain the prehistoric colonization of the New World by Asian peoples. The main conclusions obtained are: 1) that the first migration involved at least two Asian populations with most probable central and eastern Asian origins; 2) that this wave(s) occurred before 30 Kya; 3) that those pioneer colonizers did not stop in Beringia but kept trekking across the Americas leaving few genetic traces in past and present-day native American populations; 4) That after the LGM, small and relatively isolated resilient groups radiated from different areas to occupy new exploitable territories both in North and South America; 5) that these secondary expansive waves occurred first in southern-an meso-America than in northern North America; and 6) that gene flow existed among geographically adjacent groups.

MATERIAL AND METHODS

Material

For the mtDNA phylogenetic and phylogeographic analyses, I searched for partial and complete mitogenomes in the NCBI GenBank, www.ncbi.nlm.nih.gov/genbank/; Mitomap, www.mitomap.org/MITOMAP; EMPOP, https://empop.online; Ian Logan 2023, www.ianlogan.co.uk/sequences_by_group/haplogroup_select.htm; Genome Warehouse (GWH) public repository, https://bigd.big.ac.cn/gwh; and AmtDB, http://www.amtdb.org. Additionally, for phylogeographic analyses only, I also searched YFull and MTree. The NRY phylogenetic and phylogeographic analyses were based on the ISOGG database, http://www.isogg.org/tree/; YFull YTree v13.03.00; YTree; and data compiled by other authors[21-26]. For brevity, the Y-chromosome lineages in the text are referred to by the letter of their basic haplogroup and their terminal mutation.

Phylogeny

Maximum parsimony intraspecific phylogenetic trees [27] were constructed manually, and haplogroup nomenclature followed the PhyloTree database Building 17, http://www.phylotree.org. Haplogroup classification was confirmed using HaploGrep 3 3.2.1, https://haplogrep.i-med.ac.at [28]. Newly identified subgroups not classified in PhyloTree were provisionally named by the first letter and digit of their primary haplogroup, followed by an asterisk and consecutive numbers. Both labels are shown on the respective tree when a more formal nomenclature is available in YFullMTree.

Evolutionary rates

It is well known that the mtDNA evolutionary rate experiences a time-dependent effect, slowing down as we go back in time [29]. The main reason for this effect has been linked to differences in population size, with the smallest populations during Paleolithic times and exponential growth in recent times [30]. To account for this anomaly, I used an estimated Paleolithic mitogenome substitution rate of 1.5 x 10^-8 (95% CI: 0.42 - 9.07 x 10^-8) per site per year (assuming a mtDNA genome length of 16,500 base pairs), meaning one substitution occurs roughly every 4,040 years [14], to estimate haplogroup coalescent ages for the early prehistoric settlement of the Americas. Using a similar approach, a proposed mean evolutionary rate of 0.47 x 10^-9 (95% CI: 0.21 - 0.95 x 10^-9) per site per year was used to calculate coalescent ages of the Y-chromosome [14].

Ages

I have used three approaches to estimate haplogroup coalescent ages: 1) applying statistics rho [31] and Sigma [32] on individuals within haplogroups, along with the evolutionary rates mentioned above; 2) utilizing clades within main haplogroups to calculate a pooled mean, then applying normal averaging and standard error statistics for mutation rates across haplogroups; 3) since the sample sizes of these clades vary significantly, I calculated weighted averages and weighted standard errors to apply with the same normal statistics; 4) a common discussion in genetics is that haplotypes at low frequency might become much more common within a population, although this is generally an exception; typically, rare haplotypes—those at the tails of a distribution—are likely to go extinct due to genetic drift, while haplotypes near the mean tend to increase in frequency. Given that the Native American founding haplogroups have experienced genetic drift over time, I used the most divergent lineages within them as proxies for their past means to estimate coalescence ages [14,30].

Phylogeography

For phylogeographic purposes, the American continent has been divided into seven geographic regions: 1) North Northern America (NNAm), including Alaska and Canada; 2) North America, covering the entire United States (Nam); 3) Mesoamerica, comprising Mexico, Belize, Guatemala, El Salvador, Honduras, and Nicaragua (Mam); 4) Northeastern South America or Isthmus-Colombian area, including Costa Rica, Panama, and Colombia; 5) Northwestern South America (NWSAm), also known as the Andean region, including Ecuador, Peru, and Bolivia; 6) Central South America (CSAm), including Brazil, Paraguay, and Uruguay; 7) Southern South America (SSAm), including Argentina and Brazil. Additionally, within these regions, distinctions are sometimes made between inland and coastal Pacific and Atlantic areas and between Highland Andes and Amazonian zones. I have implemented some straightforward algorithms described below to support the phylogenetic and phylogeographic analyses.

Genetic distances

In addition to the molecular distance based on the number of mutations between two homologous DNA sequences (FST), I have implemented a phylogeographic distance (DXY) based on genetic affinity (IXY) between regions X and Y. This distance is measured by the frequency with which they form sister branches within a haplogroup tree (nXY) relative to how often each region participates as a sister branch within the measured haplogroup (nX and nY).

$$I_{xy} = \frac{2n_{xy}}{nx + ny}$$

Then D_{XY} is calculated as $1 - I_{XY}$.

When searching for sister branches between X and Y regions within a haplogroup, I observed cases where one sequence from one of the involved regions served as a direct outgroup to the pair, indicating it as an ancestral relative to the other region, which was considered derived. The significance of this relationship was tested using observed-expected chi-square statistics.

Additionally, for phylogeographic purposes, I have examined each major Native American founder haplogroup to identify the regions where their oldest sub-haplogroups are found and the areas from which they initially expanded.

Finally, to identify the Asian centers where the ancestors of the major American founder clades may have originated, I searched for the most ancient Asian sequences (less mutated) from the Asian-American shared phylogenetic nodes and the regions from which these sequences come.

Other statistics

Analysis of molecular variance (AMOVA), correlation/linear regression, and graphical Principal Coordinates Analysis (PCoA) were performed using the GenAlEx 6.5 software [33]. Descriptive normal statistics, 95% confidence intervals for means, and two-tailed t-tests were conducted with GraphPad https://www.graphpad.com/quickcalcs/. Weighted means and standard errors were calculated using the Weighted calculator (https://calculatodo.com).

False positive correction

Since I performed multiple statistical tests for each founder haplogroup, I applied a Bonferroni p-value adjustment. However, because the Bonferroni correction is very conservative, I also used the Benjamin and Hochberg [34] false discovery rate (FDR) method as a less restrictive alternative

---

RESULTS

All the phylogenetic and phylogeographic mtDNA results described below have been obtained from the trees presented in the supplementary figures (Supplemental_Fig_S1.xls (HgA2), Fig_S2.xls (HGB2), Fig_S3.xls (HgC1), Fig_S4.xls (HgD), and Fig_S5.xls (HgX2), and all the NRY results described below have been obtained from the trees already published by other authors [21-26,35].

The most probable origin of the American Asian ancestors

The main Native American mtDNA haplogroup founders can be divided into two groups: 1) those belonging to macro-Haplogroup M (C1, C4c, D1, D2a1, D4e1b, and D4h3a), and 2) those that belong to macro-Haplogroup N (A2, B2, and X2a).

The same grouping occurs when we look at the number of mutations accumulated in their basal stems that link them to the common nodes with Asian lineages. While C1, C4c, D1, D2a1 and D4h3a have two mutations, and D4e1b only one; A2 and B2 have five mutations and X2a four mutations (Table 1).

American haplogroup	Mutations	Coalescent Age (Kya)	Asian haplogroup	Asian region
A2	5	20,2	A-a1	NW China (Xinjiang)
A10a	1	4,04	A10	Western Siberia
B2	5	20,2	B4b1a	Altai
C1	2	8,08	C4	Eastern Asia
C4c	2	8,08	C4a'b'c	Eastern Asia
D1	2	8,08	D4	Eastern Asia
D4e1b	1	4,04	D4e1	Eastern Asia
D2a1	2	8,08	D2a'b	Northeastern Siberia
D4h3a	2	8,08	D4h3	NW China (Xinjiang)
X2a	4	16,16	X2a'j	Tajikistan

Table 1 Mutations on American basal branches to their common nodes with Asia

An unpaired t-test shows that the difference between groups is highly significant ($t = 7.2164$; $df = 5$; two-tailed P value = 0.0008). Furthermore, it is well known that macro-Haplogroup M is dominant in southern and eastern Asia [36], while

macro-Haplogroup N is more frequent in western Eurasia [37]. Consequently, the same occurs with the Asian ancestral sequences/haplogroups that show the shortest distances with the Native American haplogroups. It has already been demonstrated that the closest relatives of the macro-Haplogroup M American lineages are in eastern Asia/southeastern Siberia [38,39]. Some further commentaries are pertinent for these lineages: Hg C1 comprises four common and well spread sub lineages, three of them (C1b, C1c, C1d) practicably exclusive of the Americas and one (C1a) with a wide Asian range, and several rare clades/isolates, the most ancient of which is represented by two identical sequences from Brazil [40] and Paraguay [41] that carries twelve mutations in its basal stem (Supplementary_Fig_S3.xls). Curiously, the other isolates have been detected in northwestern Eurasia instead of East Asia: C1e in Iceland [42], western Russia [43], and western China [44]; C1f in Tajikistan [45] and C1g in Mesolithic remains from western Russia [46]. As the oldest clades are found in the New World, the most parsimonious conclusion is that all the C1 branches in Eurasia resulted from retro migrations from America. The oldest Native American samples harboring C1b lineages are USR1 from Alaska [47] and I11974 from Chile [48], both around 12,000 years old; for C1c it is MW291667 from Argentina [49], approximately 8,500 years old; and for C1d it is Lapa01, 9,600 years old, from Brazil [48]. Haplogroup D1 has the oldest samples in Yucatan, Mexico, represented by a Late Pleistocene skeleton (HN5/48) around 12 ky old [50], in Spirit Cave, Nevada, around 11 kya [9], an approximately 10 ky old LLP.S2.E1 sample from Argentina [49], and the Sumidouro4, 9,645 years old, sample from Brazil [9]. Until recently, the American haplogroup D4h3a was separated from its Asian root by six mutations, but recent aDNA studies [51] have rescued a Chinese sequence that shares four of the six basal substitutions with the American clade (Supplementary_Fig_S4.xls). The remains, from which this sequence was extracted, were unearthed at the Amur region, in northeastern China, and are around 14,000 years old. The oldest samples harboring D4h3a sequences in the New World are the 12,600-year-old Anzick1 from Montana [52], the 10,300-year-old Shuká Káa from Alaska [53], and the 10,400-year-old specimens from Lagoa Santa in Brazil [9].

About the macro-Haplogroup N, for the New World haplogroup A2, the closest Asian sequences are from haplogroup A-a1 (BZL_M16A, BZL_M16B) and belong to historic samples from Xinjiang, western China [44]. For haplogroup B2, it is AY519494, included in the Asian haplogroup B4b (Supplementary_Fig_S2.xls), with the Altai region as its geographic origin [54]. Finally, for the American haplogroup X2a, the closest Asian relatives are into haplogroup X2j, that groups samples from Egypt and Europe (Supplementary_Fig_S5.xls) but, for geographic proximity, I consider as the closest Asian sequence a partial one (sample 40) from Tajikistan [55], as it carries 16179 and 16357 substitutions that are terminal mutations in the Egyptian X2j [56] mitogenomes (Supplementary_Fig_S5.xls). For haplogroups A2 and B2 the most ancient sequences, around 14,000 years old, were extracted from human coprolites found in Paisley Caves, Oregon [57], next oldest for A2 is Lapa15

(around 8,700 years old) from Lapa do Santo, Brazil [48], and for B2 is USR2 (approximately 12,000 years old) from Upward Sun River site in Alaska [47]. The most ancient Native American sample with an X2a haplotype is the Kennewick Man from Washington, which is around 9,000 years old [58].
A way to reconcile the close arrivals of the Asian ancestors carrying the Native American M and N lineages with their significant differences in the number of mutations accumulated on their basal stems is to assume that the route followed across Asia of the N carriers was longer and slower than that of the M carriers.

Finally, if we compare the number of lineages carried by the M and N groups to their respective population sizes upon arrival in America, we could estimate that the contribution of the eastern Asian contingent was 62.5% and that of western Eurasia 37.5%.

The NRY does not show American colonization patterns as clear as those of mtDNA, but using its phylogeny and phylogeography in the Americas and Asia I have considered the ancestors of the American clades C-MPB373, that has its closest haplotypes in the eastern Asian outgroup clade C-L1373 [23,59], QF746 that includes the Q-YP1475 subclade represented by the Saqqaq ancient sample [60] and has as sister subclade Q-B143 represented by northeastern Siberian Koryaks [61], and Q-FT9713, that has its closest sequences in Mongolia and northeastern China (https://www.yfull.com/tree/Q) as of eastern Asian origin, while C-P39/Z30536, having as outgroup the Central Asian subclade C-F1756 [21], Q-Z780, that includes the Native American ancient sample Anzick1 [52] and has Q-L330 of Central Asia ascription [62,63] as sister clade, Q-M3, that has as out group Q-L804 with a northern European geographic range [64], and the rare Q-Y1150 and Q-M378 [22] as of western Eurasian origin. Unlike mtDNA, most American male lineages appeared to have their ancestors in western Eurasia. Interestingly, the C-MPB373 South American lineage is phylogenetically an outgroup of C-F1699 that comprises numerous western Eurasian lineages, including the Canadian C-P39/Z30536 branch [23]. Also of interest is that the most abundant American clade, Q-M3 [25], has as a sister branch Q-L804, which is specific to northern Europe [22]. The oldest sample belonging to the eastern group of lineages is Lapa01, from Brazil, which is approximately 8,840 years old [48]. For the western group, Anzick1 from Montana, with an approximate age of 12,600 years, represents the oldest North American sample [22]. In contrast, the oldest sample (I11974) in South America was found in Los Rieles, Chile, and is approximately 10,500 years old [48].

The global ages of the first arrival and expansions of the Native American parental ancestors

A global average of the different mean mutational data, calculated for each American mitochondrial haplogroup using various approaches (Table 2), give mean mutational values of 7.11 (95% CI: 6.34 - 7.88) substitutions at the arrival in

324 America of the founding lineages, and a mean of 4.42 (95%: 3.98 – 4.85)
325 substitutions for their initial expansions into the American continent, which, in
326 turn, produce mean ages of 28.72 (95%: 25.61 – 31.84) kya and of 17.86 (95% CI:
327 16.08 – 19.59) kya for the founding and expansion events respectively.
328

HG	Phylogeny	Mean (Age Kya)	Type	N	95% CI: Mutation/Age(Kya)
A2	Founding	7.55 (30,50)	Normal	1119 (Indiv)	7.37-7.73 / 29,77-31,23
	Expansion	4.63 (18,70)	"	1083 (Indiv)	4.46-4.79 / 18,02-19,35
	Founding	7.25 (29,29)	Weighted	118 (Clades)	6.62-7.88 / 26,74-31,83
	Expansion	4.35 (17,57)	"	118 (Clades)	4.04-4.66 / 16,32-18,83
	Coalescent	9.13 (36,88)	Max (19 nt)	16 (Branches)	8.46-9.79 / 34,18-39,55
B2	Founding	6.68 (26,99)	Normal	1081 (Indiv)	6.51-6.85 / 25,30-27,68
	Expansion	4.68 (18.90)	"	1070 (Indiv)	4.51-4.85 / 18,22-19,00
	Founding	6.62 (26,75)	Weighted	80 (Clades)	6.21-7.03 / 25,09-28,40
	Expansion	4.69 (18,95)	"	80 (Clades)	4.28-5.11 / 17,29-20,64

329
330

	Coalescent	6.42 (25,93)	Max (16 nt)	14 (Branches)	5.12-7.72 / 20,68-31,17			
C1	Founding	6.50 (26,26)	Normal	887 (Indiv)	6.33-6.67 / 25.57-26,95			
	Expansion	4.36 (17,61)	"	873 (indiv)	4.17-4.55 / 16,84-18,38			
	Founding	6.33 (25,57)	Weighted	138 Clades)	6.05-6.61 / 24,44-26,71			
	Expansion	4.01 (16,20)	"	138 (Clades)	3.71-4.31 / 14,99-17,41			
	Coalescent	6.22 (25,13)	Max (16 nt)	42(Branches)	5.54-6.90 /22,38-27,88			
C4c	Founding	6.32 (25,53)	Normal	34 (Indiv)	5.74-6.91 /23,19-27,92			
	Expansion	4.32 (17,45)	"	34 (Indiv)	3.74-4.91 / 15,11-19,84			
	Founding	3.48 (14,06)	Weighted	8 (Clades)	2.85-4.11 /11,51-16,60			
	Expansion	1.30 (5,25)	"	8 (Clades)	0.67-1.93 /2,71-7,80			
	Coalescent	4.71 (19,03)	Max (9 nt)	3 (Branches)	2.47-6.96 /9,98-28,12			
D1	Founding	7.00 (28,28)	Normal	496 (Indiv)	6.73-7.27 / 27,19-29,37			
	Expansion	4.97 (20,08)	"	463 (Indiv)	4.68-5.27 / 18,91-21,29			
	Founding	7.02 (28,36)	Weighted	58 (Clades)	6.53-7.51/26,38-30,34			

	Expansion	4.97 (20,08)	"	58 (Clades)	4.48-5.46 / 18,10-22,06			
	Coalescent	7.14 (28,86)	Max (16 nt)	28(Branches)	6.67-7.62 / 26,93-30,79			
D4h3a	Founding	6.20 (25,05)	Normal	80 (Indiv)	5.57-6.83 / 22,50-27,59			
	Expansion	4.24 (17,13)	"	64 (Indiv)	3.63-4.85 /14,67-19,59			
	Founding	6.27 (25,33)	weighted	13 (Clades)	5.73-6.81 / 23,15-27,51			
	Expansion	4.24 (17,13)	"	13 (Clades)	3.80-4.68 / 15,35-18,91			
	Coalescent	7.77 (31.39)	Max (12 nt)	12(Branches)	6.15-9.39 / 24,85-37,93			
X2a	Founding	7.83 (31,63)	Normal	41 (Indiv)	7.11-8.55 / 28,72-34,54			
	Expansion	3.83 (15,47)	"	40 (indiv)	3.11-4.55 / 12,56-18,38			
	Founding	6.83 (27,59)	Weighted	2 (Clades)	5.37-8.29 /21,69-33,49			
	Expansion	2.83 (11,43)	"	2 (Clades)	1.37-4.29 / 5,53-17,33			
	Coalescent	11.00 (44,44)	Max (15 nt)	2(Branches)	-39.82-61.82/-160,87-249,75			

Table 2. Mutational means and coalescent mean ages for the main American mtDNA haplogroups

The founding ages for the mtDNA Native American haplogroups, exposed above, coincide with the approximately 30 kya proposed in the first mtDNA studies

[11,12,38], and with the calibrated ages calculated by Rieux et al. [65], based on ancient mitogenomes. However, recent extant and ancient mitogenome studies [6,8] and whole genome-based studies [7,9] strongly support post-LGM ages for the settlement of the Americas. However, this is partly inspired by the most conservative archaeological ages proposed for the American settlement.
The age for the first population expansions calculated here, 17.86 kya, after the rigors of the LGM had lessened on the continent, does not differ much from those proposed by most authors. The difference is that this expansion coincides with most authors' arrival in America proper. In contrast, here, the arrival in America occurred much earlier, around 30 kya, before the glacial period had its worst effects.

Concerning the NRY, it deserves mention that, assuming a first colonization of the Americas about 15 kya Poznik et al. [66], estimated a substitution rate for the Y-chromosome of 8.2 x 10-10 (95% CI: 7.2 x 10-10 - 9.2 x 10-10) per site, year. However, doubling the time to 30 kya, as proposed here based on mtDNA, the substitution rate would be 4.1 x 10-10 per site, year which is very similar to that obtained by Francalacci et al. [67] of 5.3 x 10-10 (95% CI: 4.2 X10-10-7.0 x 10-10) per site, year and both estimates falling within the limits of the mean 4.7 x 10-10 (95% CI: 2.1 $\times$ 10-10 - 9.5 $\times$ 10-10) per site, year proposed by myself based on past population sizes obtained under phylogenetic considerations [14]. These results contrast with most NRY younger coalescence ages estimated by other authors [23,25,26,68].
As YFull Tree v13.03.00 (YTree) has the most actualized NRY phylogeny and the most extensive phylogeography, it is increasingly used as a data source in scientific articles, as in the present work. The YFull substitution rate [69] is identical to the one calculated by Poznik et al. [66] under the premise that the human American settlement occurred around 15 kya,; therefore, I have halved their substitution rate values when I used them. Under these requirements, NRY haplogroup C clades would arrive in the Americas in separate waves, C-MPB373 first at about 32.2 kya (95% CI: 28.8 - 35.6 kya), and C-P39/Z30536 second at about 26.2 kya (95% CI: 22.8 - 29.8 kya). The three main NRY haplogroup Q clades have very similar ages around 30 kya: Q-F746, 31.8 kya (95% CI: 28.4 - 35.2); Q-Z780, 31.0 kya (95% CI: 29.0 -33.0); Q-M3, 29.8 (27.0 - 32.8). The age of arrival to the Americas of the rare clades Q-M1150 (30.0 kya (95% CI:26.6 - 33.4 kya) and Q-M378, 21.4 kya (95% CI: 17.6 - 25.4 kya), that include European and Asian lineages, is only tentative (https://www.yfull.com/tree/Q).

Regional heterogeneity for founding and expansion events

Initially, I compared the founding and expansion ages of native haplogroups between North and South American subcontinents (Table 3). Except for haplogroup C1, the results show significantly older ages in South America for both founding and expansion events.

Haplogroup	t-value	df	p-value	oldest region
A2 - F	3.0397	1119	0.0024	South America
A2 -E	4.8326	1105	<0.0001	South America
B2 - F	1.0604	1099	0.2892	ns
B2 - E	3.1395	1083	0.0017	South America
C1 - F	1.4777	499	0.1401	ns
C1- E	0.4722	482	0.6370	ns
D1 - F	5.7700	498	<0.0001	South America
D1 - E	2.2473	466	0.0251	South America
D4h3a - F	3.0245	69	0.0035	South America
D4h3a - E	1.0115	59	0.3159	ns

Table 3. Differences in Founding(F) and expansion(E) ages between the North America and South America subcontinents

More detailed comparison between geographic regions, within and among subcontinents, using mean molecular age comparisons, pairwise genetic distances and Pearson correlation tests are described in supplementary tables (Supplemental Table S1.1 to S1.7 for A2; Supplemental Table S2.1 to S2.7 for B2; Supplemental Table S3.14, .7, .10, .13, .16 for C1b; Supplemental Table S3.2, .5, .8, .11, .14, .17 for C1c; Supplemental Table S.3.3, .6, .9, .12, .15, .18 for C1d; Supplemental Table S3.19 for C4c; Supplemental Table S4.1 to S4.6 for D1; Supplemental Table S4.7 to S4.9 for D4h3, and Supplemental Table S5.1 to S5.3 for X2a).

Haplogroup A2

Globally, A2 is the most abundant haplogroup, representing approximately 44% of the Native American population, but it has an uneven geographic distribution. The oldest founding ages for A2 (Supplemental_Table_S1.5) are in Brazil and the Colombia-Isthmus area (≈ 34 kya), expanding first in Colombia around 23 kya. From the Colombian Isthmus region, significant latitudinal clines exist to Siberia (Slope −13.37; R^2 0.8396; P-value 0.0102), and to Tierra del Fuego (Slope 23.49; R^2 0.8091; P-value 0.0377). In addition, notice that in the most ancient samples of southern South America, mtDNA haplogroup A and B lineages were totally absent or very rare [49,70-73] while they were detected in the most ancient samples recovered in northwestern North America more than 14 kya [57], and also were prominent in ancient samples from the Isthmus-Colombian area [74]. Another geographic peculiarity of haplogroup A2 is its significant abundance in the eastern regions compared to the western ones [40,75-77]. Notably, the oldest A2 haplotypes are found in North America and Brazil (Table S1.4). Furthermore, a secondary southward migration of some North American sub-lineages is also detectable. One of these lineages is A2f, which is significantly older in North America (≈35,000 ya) than in Mexico (≈23,000 ya) (t=3.6172; df 22; p-value 0.0015).
For haplogroup A2, the shortest molecular genetic distances between regions (Table S1.3) were found for the pairs CSAm-SSAm (0.657), NWSAm-NESAm (0.682), and NCAm-NNAm (0.681). When these distances are shown in a PCoA (Fig. 1), the first principal coordinate places NESAm in a central position, with the remaining regions of North- and South-America positioned to its left and right, respectively. The second principal coordinate separates the areas of each subcontinent as latitudinal clines, reflecting the Pearson correlation results.

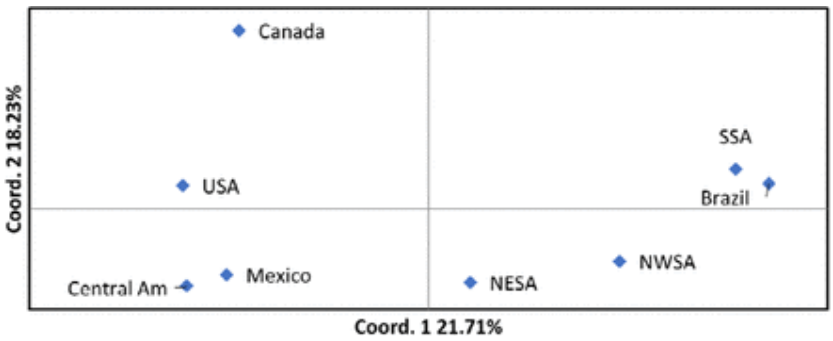

Fig. 1 Hg A2 PCoA

Haplogroup B2

Haplogroup B2 is the second-largest Native American haplogroup, comprising approximately 24.0% of the global distribution. The oldest founding (approximately 30,000 years ago) and expansion (approximately 21,000 years ago) regional values for this haplogroup are found in South and Central America. As one moves northward from these southern regions, the values decrease in a clinal trend, following a northwest-southeast orientation to reach Central and

North America. A significant negative correlation exists between the molecular ages of haplogroup B2 and latitude (Slope: −13.57; R^2: 0.5455; P-value: 0.0061).

Unlike haplogroup A2, haplogroup B2 is either absent or rare in northern North America, which mirrors its low frequency in northeastern Siberia [38]. Moreover, in contrast to haplogroups A2 and D1, haplogroup B2 is more prevalent in western continental regions than in eastern ones [78,79]. The oldest B2 haplotypes have been identified in the Andean region.
The shortest genetic molecular distances were observed between the South South American (SSAm) and Central South American (CSAm) groups (0.492), as well as between the SSAm and Northwest South American (NWSAm) pairs (0.500A PCoA bidimensional graphic representation of these distances shows NWSAm having a central position, with SSAm and CSAm at its left and NESAm and all northern American regions at its right, following a latitudinal cline as observed with the Pearson correlation (Fig. 2). (Supplementary_Table_ S2.3).

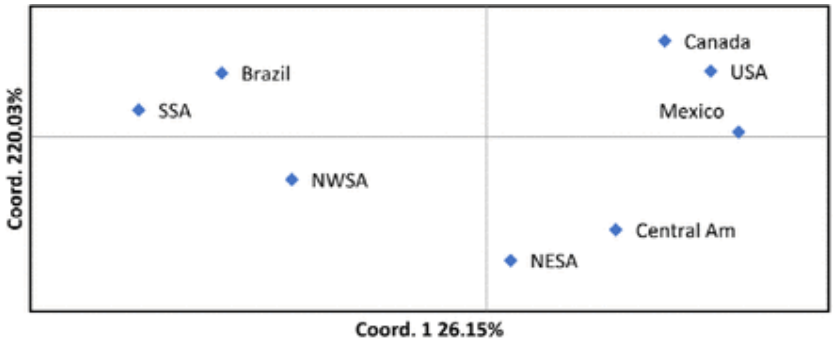

Fig. 2 HgB2 PCoA

Haplogroup C1

Haplogroup C1 was initially considered a homogenous clade; however, its three main native branches reflect distinct phylogeographic movements that warrant particular attention.

For haplogroup C1b, the oldest founding and expansion regional ages (see Supplementary Table S3.13) are found in Northeast South America (approximately 29,000 to 17,000 years ago), with younger ages extending down to Tierra del Fuego and up to Canada. The data shows a significant trend (Slope −12.20; R^2 0.8876; P-value 0.0166). These phylogeographic characteristics are somewhat similar to those of haplogroup A2. However, the longitudinal differentiation seen in the A2 clade is not present in C1b. The oldest haplotypes for C1b have been identified in northern South America, particularly in Ecuador and Colombia, and extend into Mexico (see Supplementary Table S3.10). Additionally, the shortest genetic distance between regions (refer to Supplementary Table S3.7) is observed between Northwest South America (NWSAm) and Northeast South America (NESAm)

with a distance of 0.522, followed by NWSAm and Southern South America (SSAm) at 0.605, and SSAm to Central South America (CSAm) at 0.654.

In this context, the PCoA (Principal Coordinates Analysis) graphic representation emphasizes the close relationship between NWSAm-Mexico and the USA-Canada regions, positioning NESAm as the central point from which movements occurred both southward across Brazil and northward into Central America (see Fig. 3).

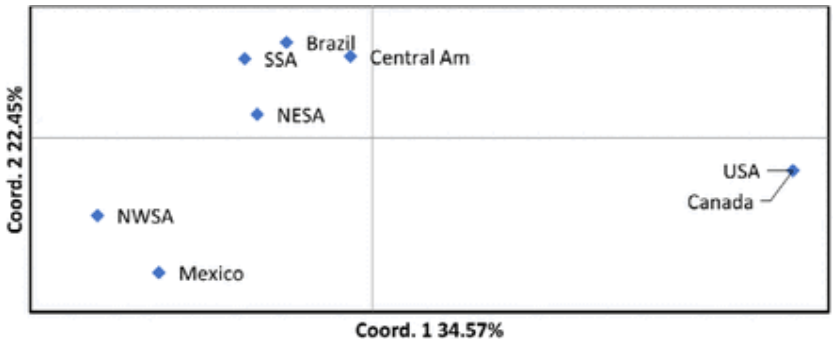

Fig. 3 HgC1b PCoA

For Hg C1c, the earliest regional ages associated with founding and expansion (Supplementary_Table_ S3.14) are centered in NWSAm (≈ 34 kya / 26 kya). From this center, populations with younger ages follow southern and northern geographic transects ending in SSAm and Canada, respectively. The latter shows a significant correlation with latitude for both founding (Slope −14.79; R^2 0.8055; P-value 0.0152) and expansion ages (Slope −11.71; R^2 0.9629; P-value 0.0005). These results partially resemble those obtained for Hg B2 in its North American subcontinent range. Congruent with its Andean center, the oldest haplotypes for C1c are found in NWSAm and Mexico (Supplementary_Table_ S3.11). The shortest genetic distances within Hg C1c (Supplementary_Table_S3.8) indicate the existence of three geographic centers: the first represented by the CSAm-CAm pair (0.571), the second by the SSAm-NWSam pair (0.630), and the third by the NWSAm-Mexico pair (0.644). This pattern could be explained by supposing NWSAm as the leading center that spread downward to Argentina and Chile and upward to Mexico, and a later center that, from Brazil, reached Central America, erasing the signal of the previous route from NWSAm to Mexico in this region. In this respect, it is worth mentioning that the Y-chromosome Q-M925, mainly Mesoamerican, has close phylogenetic ties with the Brazilian Q-Y26547 subclade [25]. This complex pattern is partially reflected in the PCoA plot (Fig. 4), which shows a latitudinal cline from NWSA to Canada passing through NESAm, with the unusual inclusion of SSAm in it, and an independent close relationship between Central America and Brazil.

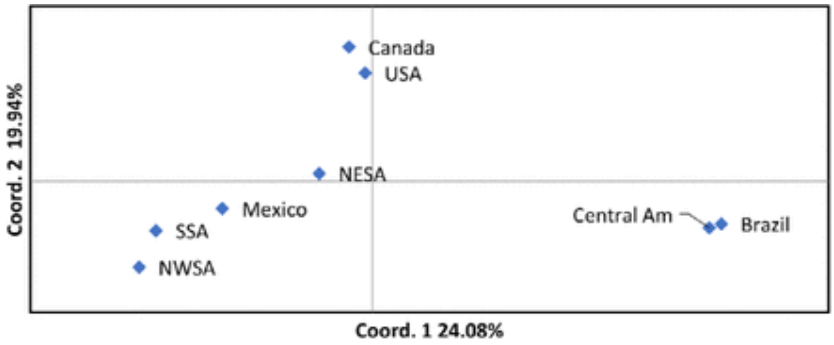

Fig. 4 Hg C1c PCoA

Hg C1d, with the oldest founding and expansion ages (Supplemental_Table_S3.15) in CSAm and SSAm (approximately 38 kya / 33 kya), is the oldest C1 branch. In fact, it is the oldest Native American clade. Although not reaching statistical significance, C1d shows a geographic trend of decreasing population ages from North America to Central America (Supplemental_Table_S3.15). This pattern resembles the one observed for Hg A2. Additionally, it exhibits a significant latitudinal correlation from SSAm to Central America, with the oldest founding (Slope −16.65; R^2 0.7649; P-value 0.0415) and expansion ages (Slope −13.29; R^2 0.6985; P-value 0.0382) in southern South America. This geographic trend within the South American subcontinent mirrors the pattern seen in Hg B2. The oldest C1d haplotypes are found in CSAm and SSAm (Supplemental_Table_S3.12), supporting the idea of an ancestral southern South American center of expansion for this clade. For C1d, the closest genetic distances (Supplemental_Table_S3.9) are between the SSAm-CSAm pair (0.591) and the NWSAm-NESAm pair (0.592). The latter reflects the close genetic relationship between these regions, also seen in the C1b clade. The PCoA bi-dimensional representation of these distances (Fig. 5) shows a central group linking SSAm and CSAm with Central America, and two distant groups: one on the left joining NWSAm and NESAm, and another on the right grouping the northern North American populations and Mexico. The independent and close genetic relationship between the South American regions and Central America resembles what was found for C1c, which was explained by a later expansion from Brazil to Central America. In contrast, the close genetic affinity between NWSAm and NESAm mirrors what was observed in C1b.

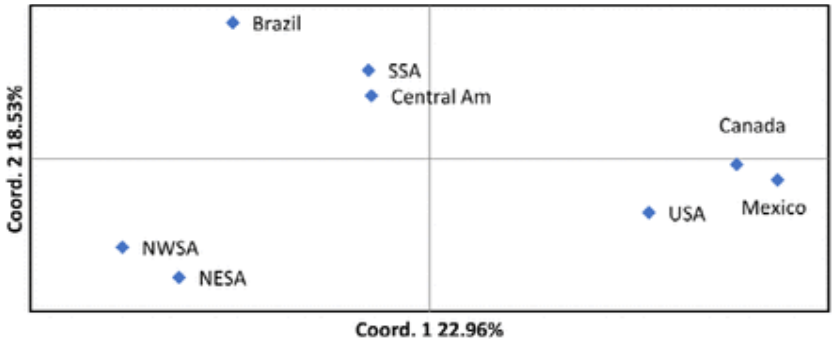

Fig. 5 HgC1d PCoA

Haplogroup C4c

The founding and expansion ages for C4c are approximately 25 and 17,000 years ago, respectively (Table 2), but ages for its main branches vary significantly (Supplemental_Table_S3.19). The most developed branch, C4c1, has founding and expansion ages of about 16 and 12,000 years ago, following the LGM. In contrast, the less common C4c2 sister branch has a founding age of 31 years ago, aligning more closely with other main Native American haplogroups. Still, its expansion age is very recent, at roughly 3,000 years ago. C4c is a minor haplogroup with a limited geographic distribution that, aside from a sporadic find in Colombia and an ancient sample from Brazil, Loca Do Suin_9100 years old [80], is confined to NNAm (Supplemental_Fig_S3). Therefore, the unusual results observed for this lineage may result from limited sampling and/or strong drift effects.

Haplogroup D1

For D1, the oldest founding ages (around 31,000 years ago) are found in SSAm and CSAm (Supplementary_Table_S4.5). The earliest expansion ages are also identified in these regions but at different times (CSAm approximately 23,000 years ago; SSAm approximately 18,000 years ago). Regarding regional D1 founding ages, there is a significant latitudinal correlation from SSAm to Central America (slope −16.65; R^2 0.5431; P-value 0.0370), while a non-significant decreasing trend is observed from Canada to Central America (Supplementary_Table_S4.5). The oldest D1 haplotypes are primarily located in NWSAm and later in SSAm (Supplementary Table S4.4). The shortest genetic distances in D1 are between SSAm-CSAm (0.618), USA-Canada (0.667), and those involving NWSAm with SSAm (0.768) and with NESAm (0.759). A PCoA plot (Fig. 6) visually illustrates these genetic relationships, placing all Mesoamerican and northern South American populations in a central position, distinctly separated from two groups—one consisting of northern North American populations and the other of southern South American populations, which are very distant from each other.

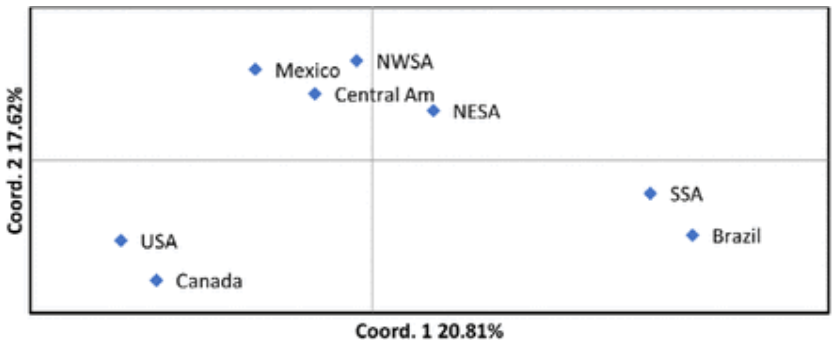

Fig. 6 Hg D1 PCOA

The D1 phylogeographic data described above closely resemble the patterns observed for haplogroups B2 and C1d.

Haplogroup D4h3a

For D4h3a, the oldest founding age (approximately 30,000 years ago) and expansion age (roughly 22,000 years ago) occurred in CSAm (Supplementary_Table_S4.8). Age comparisons between regions are only significant when the oldest expansion ages of SSAm and CSAm are compared to those of NNAm and NWSAm (Supplementary Table S4.9). When living separately, one can observe that NESAm is isolated; the oldest D4h3a haplotypes are found in NWSAm, Mexico, and consistently in Chile (Supplementary_Table_S4.7). It has been suggested that D4h3a indicates a Pacific southward expansion [81]. However, these results do not support that idea. It appears that D4h3a was more common in the past and was most frequent in eastern regions compared to today.

Haplogroup X2

Haplogroup X2a is a minor Native American haplogroup with a limited geographic range confined to northern North America [81]. The estimated ages for the origin and spread of haplogroup X2a are approximately 29 years ago and 16 years ago, respectively (Supplementary_Table_S5.1), with the oldest haplotypes found in the X2a1 branch in samples from the North American Great Plains. These ages are notably older in North America than Canada (Supplementary Table S5.3). However, the discovery of a divergent X2 haplotype in Canada [81], which may define a new Native American haplogroup called X2g (Supplemental_Fig_S5), is noteworthy. Although the current distribution of X2a, both ancient and modern, is concentrated in the northeastern part of North America [81], with sporadic cases in Mexico [82,83], its oldest known sample, the Kennewick Man–approximately 9000 years old and carrying a basal X2a haplotype–was excavated in northwestern Washington state [58]. This finding raises the possibility that the X2a clade may have originated in the northern Pacific region [81].

NRY haplogroups

On a global scale, two major parental lineages with Pan-American distribution have been identified (Q-CTS1780/Z780 and Q-M3), both within haplogroup Q, with the first occurring at lower frequencies everywhere compared to the second. Additionally, two minor lineages belonging to haplogroup C, with a subcontinental range (C-MPB373 and C-P39/Z30536), are considered Native American founder lineages. The southern C-MPB373 has been consistently reported in present-day Colombians and Equatorians [4,35,84,85]. The northern C-P39/Z30536 has been found in Canada and northern North America, reaching Mexico [86], showing a northern to southern gradient of decreasing diversity [62]. Although phylogenetically Q-

CTS1780/Z780 is the oldest clade, the demographic expansion following the LGM mainly occurred among males within the Q-M3 lineage that branched off simultaneously, mainly in Mesoamerica (Q-M925) and South America (Q-M848), forming well-differentiated regional subclades [22-26,35,68,75,84,87-89]. Some of these minor Q subclades are also noteworthy due to their geographic distributions and STR diversity. One such lineage is Q-MEH2, identified in northeast Siberian Koryaks. Its close relationship to the extinct Paleo-Eskimo Saqqaq [60] suggests it could trace a recent migration from Siberia to America across the Bering Strait [90]. However, Q-MEH2 has recently been found in Nahua populations from southern Mexico, and based on its STR variability, it appears to have arrived much earlier in Mexico than in northern America [24]. Another notable example is the Mexican Q-PV3 lineage, initially considered a minor branch of Q-L54, but its divergent haplotypic variability suggests it might represent another founding Native American lineage [24]. This subcontinental differentiation of male ancestries supports diverse demographic histories in North and South America [87]. However, before drawing conclusions, it is necessary to examine what NRY ancient DNA analyses indicate, especially since post-Columbian admixture heavily impacted the Native American male genetic pool in both North [91] and South America [92]. The picture was similar. Although Q-CTS1780 might have been more common in the past across both subcontinents [22,48,80], Q-M3 was also the dominant NRY lineage in all ancient DNA data examined [6]. Nonetheless, the detection of today's extinct lineages in the past [25,48] supports the idea that firm founding and drift effects occurred in NRY Native American lineages. These findings do not align well with the hypothesis of a rapid and recent migratory wave from the north to explain the prehistoric colonization of South America.

DISCUSSION

The early arrival of Native uniparental lineages in the Americas, over 30 kya, as proposed here, renders the hypotheses of a prolonged [93,94] or brief [23] standstill in Beringia, a previous expansion of their ancestors into Asia [21,59,95], or the existence of a Pacific Coastal Route around 20 kya [96-98], unnecessary.
The archaeological evidence for humans in the Siberian Arctic around 30 kya [99] and possibly since 45 kya [100] supports the idea of an early entrance into America by its native ancestors. More challenging is gaining acceptance from archaeologists for human occupation in the American continent before and around the LGM, but it looks pretty likely that this occurred around 16-18 kya at Monte Verde, Chile [101], at the Santa Elina site in Central Brazil during the LGM, at Chiquihuite Cave in Central Mexico about 30 kya [102], in the Colorado Plateau more than 36 kya [103], at White Sands, New Mexico, during a stratigraphic record spanning from 24 kya to 17 kya [104], and at Bluefish Caves, Canada, approximately 24 kya [105].

638 Ancient genome studies have characterized the ancestral Native American
639 population as a mixture of northern Eurasians and northern East Asians [106]. The
640 same dual contribution was previously suggested using NRY data [4,84,107], and a
641 similar conclusion has been reached here based on mtDNA data. Genomic research
642 indicates that the ancestral Native Americans arrived in the New World as an
643 already admixed group [9]. However, although coalescence ages do not show
644 significant differences among haplogroups, the NRY phylogeny is more consistent
645 with two successive waves: the first carrying C-MPB373 and Q-Z780 lineages, and
646 the second comprising carriers of C-P39/Z30536 and Q-M3 lineages. Evidence for
647 mtDNA is less conclusive, but the fact that the basal branches connecting
648 haplogroups C and D to their respective Asian roots are significantly shorter than
649 those of A and B could support the idea that C and D arrived earlier in America.
650 Another clue might be found in the Southern Cone of South America, the region
651 farthest from the continent's entry point. Ancient and modern samples from this
652 area show that haplogroups C and D were the first and predominant lineages
653 reaching the region—still the most common today [70-73,108,109]. The dominance of
654 haplogroups C and D has also been observed in Amazonian populations from
655 Brazil [110], suggesting an early southward migration. Furthermore, mtDNA diversity
656 for haplogroups C1 and D1 increases from north to south, while North American
657 populations exhibit higher diversity levels for haplogroups A2 and B2 [111]. In North
658 America, haplogroups A, B, and X have expanded across populations in a clinal
659 pattern, whereas there is no evidence of such spread for haplogroups C and D [112].
660 These patterns are consistent with two successive migration waves. However,
661 genetic drift may also explain them, as samples from other regions show the
662 absence of some founder mtDNA haplogroups. For example, the Kuna from
663 Panama lack C and D [113], while the Ayoreo from Bolivia show no A or B
664 haplogroups [114]. Based on the distribution of two rare mtDNA haplogroups, D4h3
665 and X2a, with the first spreading along the Pacific coast and the second entering
666 via the ice-free corridor, two migratory waves from Beringia have been proposed
667 [81]. Nevertheless, for D4h3, the authors noted that its rarity in North America and
668 higher frequency and variation in South America require further explanation [81].
669 Another hypothesis, based on genome studies, suggests that after arriving in
670 North America, the native ancestral population split into two distinct branches
671 known as South Native Americans (SNA) and North Native Americans (NNA) [9], or
672 their approximate equivalents, Ancestry A (ANC-A) and Ancestry B (ANC-B),
673 respectively [96]. While the northern ancestry is mainly limited to Canada and eastern
674 North American regions, the southern ancestry is detectable in the oldest sampled
675 genomic sequence, Anzick-1, which is approximately 12700 years old from
676 Montana [52], and in the Spirit Cave sample from Nevada, around 10700 years old [9],
677 extending southward to Mesoamerica and South America [9]. There are single-
678 parental markers that share the geographic range of NNA genomic ancestry, such
679 as mtDNA haplogroups X2a [115] and C4c [116], as well as NRY lineages C-P39 [23], Q-
680 Y4276 [22]. The situation becomes much more complex when explaining the

transcontinental distribution of the southern component of the Native American genome. The initial dual explanation involving southward Pacific coastal expansion—where successive populations could be seen as mixtures of SNA and NNA components in varying proportions [96]—was proven insufficient [48]. First, the uniqueness of the northernmost SNA component, represented by the Clovis-associated Anzick-1 sample (SNA1), could not account for the genomic makeup of many South American populations, leading to the need to add a new, dominant non-Clovis SNA2 component, partly represented by Spirit Cave ancestry, to better explain the genomes of many southern populations. Approaching Mesoamerica, the Mexican Mixe population required gene flow from an unsampled population for accurate representation [9]. Furthermore, several undiscovered genomic ancestries needed to be included to properly model the structure of various South American regional populations. An interesting case is the presence of a divergent ancestry related to Australasian people found in Amazonian tribes such as Surui, Karitiana, and Xavante [117,118]. This component, called "population Y," was confirmed in ancient Brazilian samples like Sumidouro-5, approximately 10000 years old [9]. At least one mtDNA haplotype, likely from haplogroup B—whose southernmost lineage aligns well with the Australasian ancestry—was necessary to accompany this component. However, identifying the ancestral Asian population from which it entered the New World remains challenging. Basic lineages of haplogroup B appeared in eastern Asia around 40,000 years ago, as shown by the Chinese Tianyuan sample [119]. The Australasian ancestry shared by ancient South American samples and Tianyuan is also notable [120], suggesting it could have arrived with populations deeply rooted in eastern Asia with ancient Australasian ties, or with carriers of mtDNA haplogroup B from central Eurasia. The Isthmus-Colombian region also showed genomic complexity, with a Pleistocene ancestry reaching Panamanians and remaining confined to the Isthmus [74]. Hunter-gatherers from the Colombian Altiplano also carried an unknown basal lineage from the initial South American genomic radiation [16]. In northwest South America, a previously unrecognized ancestry was distinctive of northern Ecuador and Colombian populations [121], while Native Peruvians from the Lake Titicaca area showed a Siberian component also found in ancient North Americans like Anzick-1, Kennewick, and Saqqaq [53]. The complex ancestries uncovered in South America have led some scholars to caution against oversimplifying the rapid post-glacial colonization model, suggesting that a northward expansion from South America to explain the Pacific migration wave would require a more intricate history [48]. Others propose that the dispersal pattern of SNA south of the continent's ice sheets involved complex admixture events among earlier populations [9]. The uniparental markers analyzed in this study offer an alternative hypothesis, potentially integrating all genomic data from a different perspective.

If we accept that Asian populations gave rise to Native Americans before 30 kya, then the story of human settlement in America can be viewed differently. It is known that since their arrival, the climate gradually worsened, covering the Arctic

724 and Antarctic regions with ice (Clark et al. 2009)[122], clearing the Amazon and
725 Central American rainforests (Häggi et al. 2017)[123], and transforming most of the
726 land into forests and dry steppes (University of Geneva, Switzerland et al. 2001)[124].
727 During this tough period, humans had to survive in small, isolated groups and
728 adapt to different glacial refugia, leaving few archaeological traces. When the
729 climate improved, these groups grew in number and expanded into new, now
730 habitable areas. Thus, unlike the rapid southern wave of colonization proposal, the
731 American continent was not empty of humans. The post-LGM colonization was not
732 a single or multiple waves spreading from northern North America to settle
733 Mesoamerica and then South America. Instead, it involved more or less
734 simultaneous radiations from various regions, most of which, based on uniparental
735 markers, are centered in South America. One major expansion center for mtDNA is
736 the northwest Colombian-Isthmus area, where younger haplogroup A2 and C1b
737 clines are seen stretching north and south. A similar distribution of A2 frequencies
738 has also been observed in this region [125]. Another potential core is in the Andean
739 region, where mtDNA haplogroup C1c shows its oldest origins and expansion
740 ages, with younger clines extending toward Canada and down to SSAm. This
741 pattern mirrors haplogroup B, which expanded from Chile and the Andean region
742 into Mesoamerica. However, it appears that at least part of the B2 cline results
743 from displacement by other mtDNA lineages in North America, as B was present
744 early on in northeast Asia as an ancestral B lineage [119] and in Alaska as a basal B2
745 lineage [47], before nearly disappearing from both areas. The highest frequencies
746 and genetic diversity of haplogroup B2 are found in the Andean Altiplano [126,127],
747 the Bolivian Piedmont [128], and the Argentinian Gran Chaco [109]. Although the classic
748 explanation for the B2 cline suggests descent from California [127], South American
749 origins have also been proposed, from the Peruvian Andes [129] or the Chilean
750 Southern Cone [130]. Similarly, the C1d haplogroup, like B2, shows its most ancient
751 origins and expansion in the southern Cone, spreading northward with younger
752 ages toward Central America, indicating a significant geographic pattern. In North
753 America, some lineages of C1d show an expansion history that ends in Central
754 America, paralleling trends observed in haplogroup A2 sub-lineages. C1a [94], a
755 northeast Asiatic branch of C1, along with minor European sub-clades C1e [42] and
756 C1f [46], suggests a later reverse migration from America [94]. Using the evolutionary
757 rate proposed here, the founding age of C1a in Asia is about 26.02 kya (95% CI:
758 18.58-33.49 kya), signaling a return to Asia during worsening weather conditions.
759 Its spread across Asia around 13.9 kya (95% CI: 6.46-21.37 kya) points to a
760 Mesolithic migration, similar to the pattern seen in the northeast European C1f
761 clade [46]. This early American retro-migration might also be supported by the
762 phylogeny and phylogeography of the Y-chromosome haplogroups C-L1373 [23,25]
763 and Q-L54 [22,23].
764 Another mtDNA haplogroup indicating the southern Cone region as a key center of
765 expansion is D1, which shows a significant negative correlation from this area to
766 Central America. Its mean coalescent ages are younger as you move upward,

although there is a longitudinal bias leaning toward central and eastern South American regions. D1 is most common in the eastern parts of the Andean region, with ancient regional clades in Central Argentina [131] and the southern Cone [132]. Similarly, for A2 and C1d, a southward expansion from Canada to Central America is observed for D1. These expansion patterns might explain the B2 retraction seen in North America. The evidence here does not support the idea that mtDNA haplogroup D4h3a signals a Pacific coastal route for peopling of the Americas (81), later backed by genomic studies (9). The earliest ages for the founding and expansion of D4h3a are found in SSAm and CSAm, and are significantly older than those in NWSAm and NNAm (Supplementary_Table_S4.9_ S4.9). Additionally, although D4h3a is absent today, it has been found in ancient samples from Alaska, Shuká-Káa, 10.3 kya [53], from British Columbia, Canada, 939, 6.5 kya [133], and Montana, Anzick 1, 12.6 kya [52]. It was also detected in ancient remains from the Brazilian Amazon, Lagoa Santa, 10 kya [9]. This pattern closely resembles that of B2. Therefore, rather than modeling a southward Pacific Coastal expansion, both markers seem to be remnants of a broader past occupation displaced by more recent human migrations.

A specific area in the Amazon Basin (CSAm) deserves special mention. It participated in all the mtDNA radiation centers discussed earlier and is known as a significant reservoir of deep mtDNA lineages [134,135]. It stands out for the ancient presence of very deep-rooted Y-chromosome lineages such as C-L1373* [25] and for the detection of lineages that were previously thought to be unique to the northern subcontinent, like Q-Y4303* [22]. This genetic diversity is also supported by whole-genome studies on ancient specimens [9] and current populations [136,137]. This genetic landscape aligns with the region's linguistic richness. Notably, out of 71 identified linguistic isolates across the continent, over 45% are in Brazil, compared to 15% in North and Mesoamerica combined [138]. Future archaeological findings will likely uncover human settlements much older than currently recognized. Furthermore, the differences in genetic and linguistic diversity between the northern (NNA) and southern (SNA) American branches make the idea of simultaneous colonization of both subcontinents from a northward expansion unlikely.

Supplementary information

Figures S1 Haplogroup A2 and A10 phylogenetic trees

Figures S2 Haplogroup B2 and B4b1 phylogenetic trees

Figures S3 Haplogroup C1a, C1b, C1c, C1d, C1e, C1f, and C4c phylogenetic trees

Figures S4 Haplogroup D1 and D4h phylogenetic trees

Figures S5 Haplogroup X2 phylogenetic tree

Tables S1 Haplogroup A2 statistical comparisons

Tables S2 Haplogroup B2 statistical comparisons

Tables S3 Haplogroup C1b, C1c, C1d, and C4c statistical comparisons

Tables S4 Haplogroup D1 and D4h3 statistical comparisons

Tables S5 Haplogroup X2a statistical comparisons

.

Availability of data and materials

All the sequences and haplotypes used in this study are available in the public domain. Databases and references are detailed in the text.

Competing interests

The author declares no conflicts of interest regarding this work.

Founding

Not applicable

Acknowledgements

Not applicable

REFERENCES

1. Becerra-Valdivia, L. & Higham, T. The timing and effect of the earliest human arrivals in North America. *Nature* **584**, 93–97 (2020).

2. Fiedel, S. J. The Peopling of the New World: Present Evidence, New Theories, and Future Directions. *Journal of Archaeological Research* **8**, 39–103 (2000).

3. Schurr, T. G. The Peopling of the New World: Perspectives from Molecular Anthropology. *Annu. Rev. Anthropol.* **33**, 551–583 (2004).

4. Karafet, T. M. *et al.* Ancestral Asian Source(s) of New World Y-Chromosome Founder Haplotypes. *The American Journal of Human Genetics* **64**, 817–831 (1999).

5. Palencia-Madrid, L. & De Pancorbo, M. M. Next generation sequencing as a suitable analysis for authentication of degraded and ancient DNA. *Forensic Science International: Genetics Supplement Series* **5**, e338–e340 (2015).

6. Bisso-Machado, R. & Fagundes, N. J. R. Uniparental genetic markers in Native Americans: A summary of all available data from ancient and contemporary populations. *American J Phys Anthropol* **176**, 445–458 (2021).

7. Willerslev, E. & Meltzer, D. J. Peopling of the Americas as inferred from ancient genomics. *Nature* **594**, 356–364 (2021).

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

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

10. Faught, M. K. Where was the PaleoAmerind standstill? *Quaternary International* **444**, 10–18 (2017).

11. Bonatto, S. L. & Salzano, F. M. Diversity and Age of the Four Major mtDNA Haplogroups, and Their Implications for the Peopling of the New World. *The American Journal of Human Genetics* **61**, 1413–1423 (1997).

847 12. Torroni, A., Neel, J. V., Barrantes, R., Schurr, T. G. & Wallace, D. C. Mitochondrial
848 DNA 'clock' for the Amerinds and its implications for timing their entry into North
849 America. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1158–1162 (1994).
850 13. Cabrera, V. M. Counterbalancing the time-dependent effect on the human
851 mitochondrial DNA molecular clock. *BMC Evol Biol* **20**, 78 (2020).
852 14. Cabrera, V. M. Early male and female footprints of modern humans across Eurasia
853 and Australasia. *Medical Research Archives* **12**, (2024).
854 15. Gusareva, E. S. *et al.* From North Asia to South America: Tracing the longest
855 human migration through genomic sequencing. *Science* **388**, eadk5081 (2025).
856 16. Krettek, K.-L. *et al.* A 6000-year-long genomic transect from the Bogotá Altiplano
857 reveals multiple genetic shifts in the demographic history of Colombia. *Science*
858 *AdvAnceS* (2025).
859 17. Hubert, N. & Hanner, R. DNA Barcoding, species delineation and taxonomy: a
860 historical perspective. *DNA Barcodes* **3**, (2015).
861 18. Posth, C. *et al.* Deeply divergent archaic mitochondrial genome provides lower
862 time boundary for African gene flow into Neanderthals. *Nat Commun* **8**, 16046
863 (2017).
864 19. Petr, M. *et al.* The evolutionary history of Neanderthal and Denisovan Y
865 chromosomes. *Science* **369**, 1653–1656 (2020).
866 20. Cabrera, V. M. Revisiting the Hominin Phylogeny: An Alternative Introgression
867 Scenario to Reconcile the Uniparental and autosomal DNA Topologies. *Jpeb* **9**,
868 (2021).
869 21. Wei, L.-H. *et al.* Paternal origin of Paleo-Indians in Siberia: insights from Y-
870 chromosome sequences. *Eur J Hum Genet* **26**, 1687–1696 (2018).

871 22. Grugni, V. *et al.* Analysis of the human Y-chromosome haplogroup Q characterizes
872 ancient population movements in Eurasia and the Americas. *BMC Biol* **17**, 3
873 (2019).
874 23. Pinotti, T. *et al.* Y Chromosome Sequences Reveal a Short Beringian Standstill,
875 Rapid Expansion, and early Population structure of Native American Founders.
876 *Current Biology* **29**, 149-157.e3 (2019).
877 24. Gómez, R. *et al.* Y chromosome diversity in Aztlan descendants and its
878 implications for the history of Central Mexico. *iScience* **24**, 102487 (2021).
879 25. Colombo, G. *et al.* Overview of the Americas' First Peopling from a Patrilineal
880 Perspective: New Evidence from the Southern Continent. *Genes* **13**, 220 (2022).
881 26. Paz Sepúlveda, P. B. *et al.* Human Y chromosome sequences from Q Haplogroup
882 reveal a South American settlement pre-18,000 years ago and a profound
883 genomic impact during the Younger Dryas. *PLoS ONE* **17**, e0271971 (2022).
884 27. Bandelt, H. J., Forster, P. & Rohl, A. Median-joining networks for inferring
885 intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37–48 (1999).
886 28. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in
887 the era of high-throughput sequencing. *Nucleic Acids Res* **44**, W58–W63 (2016).
888 29. Henn, B. M., Gignoux, C. R., Feldman, M. W. & Mountain, J. L. Characterizing the
889 Time Dependency of Human Mitochondrial DNA Mutation Rate Estimates.
890 *Molecular Biology and Evolution* **26**, 217–230 (2008).
891 30. Cabrera, V. M. Human molecular evolutionary rate, time dependency and
892 transient polymorphism effects viewed through ancient and modern
893 mitochondrial DNA genomes. *Sci Rep* **11**, 5036 (2021).

894 31. Forster, P., Harding, R., Torroni, A. & Bandelt, H.-J. Origin and Evolution of Native
895 American mtDNA Variation: A Reappraisal. *Am. J. Hum. Genet.* (1996).
896 32. Saillard, J., Forster, P., Lynnerup, N., Bandelt, H.-J. & Nørby, S. mtDNA Variation
897 among Greenland Eskimos: The Edge of the Beringian Expansion. *The American*
898 *Journal of Human Genetics* **67**, 718–726 (2000).
899 33. Peakall, R. & Smouse, P. E. GenAlEx 6.5: genetic analysis in Excel. Population
900 genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537–
901 2539 (2012).
902 34. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and
903 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*
904 *Series B: Statistical Methodology* **57**, 289–300 (1995).
905 35. Villaescusa, P. *et al.* A Y-chromosomal survey of Ecuador's multi-ethnic population
906 reveals new insights into the tri-partite population structure and supports an early
907 Holocene age of the rare Native American founder lineage C3-MPB373. *Forensic*
908 *Science International: Genetics* **51**, 102427 (2021).
909 36. Marrero, P., Abu-Amero, K. K., Larruga, J. M. & Cabrera, V. M. Carriers of human
910 mitochondrial DNA macrohaplogroup M colonized India from southeastern Asia.
911 *BMC Evol Biol* **16**, 246 (2016).
912 37. Fregel, R., Cabrera, V., Larruga, J. M., Abu-Amero, K. K. & González, A. M. Carriers
913 of Mitochondrial DNA Macrohaplogroup N Lineages Reached Australia around
914 50,000 Years Ago following a Northern Asian Route. *PLoS ONE* **10**, e0129839
915 (2015).
916 38. Starikovskaya, Y. B., Sukernik, R. I., Schurr, T. G., Kogelnik, A. M. & Wallace, D. C.
917 mtDNA Diversity in Chukchi and Siberian Eskimos: Implications for the Genetic

History of Ancient Beringia and the Peopling of the New World. *The American Journal of Human Genetics* **63**, 1473–1491 (1998).

39. Derenko, M. *et al.* Origin and Post-Glacial Dispersal of Mitochondrial DNA Haplogroups C and D in Northern Asia. *PLoS ONE* **5**, e15214 (2010).

40. Simão, F. *et al.* The Ancestry of Eastern Paraguay: A Typical South American Profile with a Unique Pattern of Admixture. *Genes* **12**, 1788 (2021).

41. Avila, E. *et al.* Full mtDNA genome sequencing of Brazilian admixed populations: A forensic-focused evaluation of a MPS application as an alternative to Sanger sequencing methods. *Forensic Science International: Genetics* **42**, 154–164 (2019).

42. Ebenesersdóttir, S. S. *et al.* A new subclade of mtDNA haplogroup C1 found in icelanders: Evidence of pre-columbian contact? *American J Phys Anthropol* **144**, 92–99 (2011).

43. Järve, M. *et al.* Shifts in the Genetic Landscape of the Western Eurasian Steppe Associated with the Beginning and End of the Scythian Dominance. *Current Biology* **29**, 2430-2441.e10 (2019).

44. Wang, W. *et al.* Ancient Xinjiang mitogenomes reveal intense admixture with high genetic diversity. *Sci. Adv.* **7**, eabd6690 (2021).

45. Peng, M.-S. *et al.* Mitochondrial genomes uncover the maternal history of the Pamir populations. *Eur J Hum Genet* **26**, 124–136 (2018).

46. Der Sarkissian, C. *et al.* Mitochondrial Genome Sequencing in Mesolithic North East Europe Unearths a New Sub-Clade within the Broadly Distributed Human Haplogroup C1. *PLoS ONE* **9**, e87612 (2014).

940 47. Tackney, J. C. *et al.* Two contemporaneous mitogenomes from terminal
941 Pleistocene burials in eastern Beringia. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13833–
942 13838 (2015).
943 48. Posth, C. *et al.* Reconstructing the Deep Population History of Central and South
944 America. *Cell* **175**, 1185-1197.e22 (2018).
945 49. Roca-Rada, X. *et al.* Ancient mitochondrial genomes from the Argentinian Pampas
946 inform the early peopling of the Southern Cone of South America. *iScience* **24**,
947 102553 (2021).
948 50. Chatters, J. C. *et al.* Late Pleistocene Human Skeleton and mtDNA Link
949 Paleoamericans and Modern Native Americans. *Science* **344**, 750–754 (2014).
950 51. Mao, X. *et al.* The deep population history of northern East Asia from the Late
951 Pleistocene to the Holocene. *Cell* **184**, 3256-3266.e13 (2021).
952 52. Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis
953 burial site in western Montana. *Nature* **506**, 225–229 (2014).
954 53. Lindo, J. *et al.* The genetic prehistory of the Andean highlands 7000 years BP
955 though European contact. *SCIENCE ADVANCES* (2018).
956 54. Starikovskaya, E. B. *et al.* Mitochondrial DNA Diversity in Indigenous Populations
957 of the Southern Extent of Siberia, and the Origins of Native American
958 Haplogroups. *Annals of Human Genetics* **69**, 67–89 (2005).
959 55. Ovchinnikov, I. V., Malek, M. J., Drees, K. & Kholina, O. I. Mitochondrial DNA
960 variation in Tajiks living in Tajikistan. *Legal Medicine* **16**, 390–395 (2014).
961 56. Kujanová, M., Pereira, L., Fernandes, V., Pereira, J. B. & Černý, V. Near Eastern
962 Neolithic genetic input in a small oasis of the Egyptian Western Desert. *American J*
963 *Phys Anthropol* **140**, 336–346 (2009).

964 57. Gilbert, M. T. P. *et al.* DNA from Pre-Clovis Human Coprolites in Oregon, North
965 America. *Science* **320**, 786–789 (2008).
966 58. Rasmussen, M. *et al.* The ancestry and affiliations of Kennewick Man. *Nature* **523**,
967 455–458 (2015).
968 59. Sun, J. *et al.* Post-last glacial maximum expansion of Y-chromosome haplogroup
969 C2A-L1373 in northern Asia and its implications for the origin of Native Americans.
970 *American J Phys Anthropol* **174**, 363–374 (2021).
971 60. Rasmussen, M. *et al.* Ancient human genome sequence of an extinct Palaeo-
972 Eskimo. *Nature* **463**, 757–762 (2010).
973 61. Karmin, M. *et al.* A recent bottleneck of Y chromosome diversity coincides with a
974 global change in culture. *Genome Res.* **25**, 459–466 (2015).
975 62. Dulik, M. C. *et al.* Mitochondrial DNA and Y Chromosome Variation Provides
976 Evidence for a Recent Common Ancestry between Native Americans and
977 Indigenous Altaians. *The American Journal of Human Genetics* **90**, 229–246
978 (2012).
979 63. Flegontov, P. *et al.* Palaeo-Eskimo genetic ancestry and the peopling of Chukotka
980 and North America. *Nature* **570**, 236–240 (2019).
981 64. Hallast, P. *et al.* The Y-Chromosome Tree Bursts into Leaf: 13,000 High-Confidence
982 SNPs Covering the Majority of Known Clades. *Molecular Biology and Evolution* **32**,
983 661–673 (2015).
984 65. Rieux, A. *et al.* Improved Calibration of the Human Mitochondrial Clock Using
985 Ancient Genomes. *Molecular Biology and Evolution* **31**, 2780–2792 (2014).
986 66. Poznik, G. D. *et al.* Sequencing Y Chromosomes Resolves Discrepancy in Time to
987 Common Ancestor of Males Versus Females. *Science* **341**, 562–565 (2013).

988 67. Francalacci, P. *et al.* Low-Pass DNA Sequencing of 1200 Sardinians Reconstructs
989 European Y-Chromosome Phylogeny. *Science* **341**, 565–569 (2013).
990 68. Jota, M. S. *et al.* New native South American Y chromosome lineages. *J Hum*
991 *Genet* **61**, 593–603 (2016).
992 69. Adamov, D., Guryanov, V., Karzhavin, S., Tagankin, V. & Urasin, V. Defining a New
993 Rate Constant for Y-Chromosome SNPs based on Full Sequencing Data. (2015).
994 70. Lalueza, C. Lack of founding Amerindian mitochondrial DNA lineages in extinct
995 aborigines from Tierra del Fuego-Patagonia. *Human Molecular Genetics* **6**, 41–46
996 (1997).
997 71. Moraga, M. L. *et al.* Mitochondrial DNA polymorphisms in Chilean aboriginal
998 populations: Implications for the peopling of the southern cone of the continent.
999 *Am. J. Phys. Anthropol.* **113**, 19–29 (2000).
1000 72. Nakatsuka, N. *et al.* A Paleogenomic Reconstruction of the Deep Population
1001 History of the Andes. *Cell* **181**, 1131-1145.e21 (2020).
1002 73. Postillone, M. *et al.* Mitochondrial DNA Diversity and Evolutionary History of
1003 Native Human Populations of Argentinean Northwest Patagonia. *hub* (2020)
1004 doi:10.1353/hub.2017.0069.
1005 74. Capodiferro, M. R. *et al.* Archaeogenomic distinctiveness of the Isthmo-Colombian
1006 area. *Cell* **184**, 1706-1723.e24 (2021).
1007 75. Kemp, B. M. *et al.* Genetic analysis of early holocene skeletal remains from Alaska
1008 and its implications for the settlement of the Americas. *American J Phys Anthropol*
1009 **132**, 605–621 (2007).

1010 76. Raff, J. A., Bolnick, D. A., Tackney, J. & O'Rourke, D. H. Ancient DNA perspectives
1011 on American colonization and population history. *Am. J. Phys. Anthropol.* **146**,
1012 503–514 (2011).
1013 77. Rothhammer, F., Fehren-Schmitz, L., Puddu, G. & Capriles, J. Mitochondrial DNA
1014 haplogroup variation of contemporary mixed South Americans reveals prehistoric
1015 displacements linked to archaeologically-derived culture history. *American J Hum*
1016 *Biol* **29**, e23029 (2017).
1017 78. Bisso-Machado, R., Bortolini, M. C. & Salzano, F. M. Uniparental genetic markers in
1018 South Amerindians. *Genet. Mol. Biol.* **35**, 365–387 (2012).
1019 79. Brandini, S. *et al.* The Paleo-Indian Entry into South America According to
1020 Mitogenomes. *Molecular Biology and Evolution* **35**, 299–311 (2018).
1021 80. Ferraz, T. *et al.* Genomic history of coastal societies from eastern South America.
1022 *Nat Ecol Evol* **7**, 1315–1330 (2023).
1023 81. Perego, U. A. *et al.* Distinctive Paleo-Indian Migration Routes from Beringia
1024 Marked by Two Rare mtDNA Haplogroups. *Current Biology* **19**, 1–8 (2009).
1025 82. Peñaloza-Espinosa, R. I. *et al.* Characterization of mtDNA Haplogroups in 14
1026 Mexican Indigenous Populations. *hub* **79**, 313–320 (2007).
1027 83. Bodner, M. *et al.* The Mitochondrial DNA Landscape of Modern Mexico. *Genes* **12**,
1028 1453 (2021).
1029 84. Roewer, L. *et al.* Continent-Wide Decoupling of Y-Chromosomal Genetic Variation
1030 from Language and Geography in Native South Americans. *PLoS Genet* **9**,
1031 e1003460 (2013).
1032 85. Mezzavilla, M., Geppert, M., Tyler-Smith, C., Roewer, L. & Xue, Y. Insights into the
1033 origin of rare haplogroup C3* Y chromosomes in South America from high-density

1034 autosomal SNP genotyping. *Forensic Science International: Genetics* **15**, 115–120
1035 (2015).
1036 86. Perez-Benedico, D. *et al.* Mayans: a Y chromosome perspective. *Eur J Hum Genet*
1037 **24**, 1352–1358 (2016).
1038 87. Bortolini, M.-C. *et al.* Y-Chromosome Evidence for Differing Ancient Demographic
1039 Histories in the Americas. *The American Journal of Human Genetics* **73**, 524–539
1040 (2003).
1041 88. Battaglia, V. *et al.* The First Peopling of South America: New Evidence from Y-
1042 Chromosome Haplogroup Q. *PLoS ONE* **8**, e71390 (2013).
1043 89. Geppert, M. *et al.* Identification of new SNPs in native South American
1044 populations by resequencing the Y chromosome. *Forensic Science International:*
1045 *Genetics* **15**, 111–114 (2015).
1046 90. Malyarchuk, B. *et al.* Ancient links between Siberians and Native Americans
1047 revealed by subtyping the Y chromosome haplogroup Q1a. *J Hum Genet* **56**, 583–
1048 588 (2011).
1049 91. Malhi, R. S. *et al.* Mitochondrial haplogroup M discovered in prehistoric North
1050 Americans. *Journal of Archaeological Science* **34**, 642–648 (2007).
1051 92. Resque, R. *et al.* Male Lineages in Brazil: Intercontinental Admixture and
1052 Stratification of the European Background. *PLoS ONE* **11**, e0152573 (2016).
1053 93. Hoffecker, J. F., Powers, W. R. & Goebel, T. The Colonization of Beringia and the
1054 Peopling of the New World. *Science* **259**, 46–53 (1993).
1055 94. Tamm, E. *et al.* Beringian Standstill and Spread of Native American Founders. *PLoS*
1056 *ONE* **2**, e829 (2007).

95. Ning, C. *et al.* The genomic formation of First American ancestors in East and Northeast Asia. Preprint at https://doi.org/10.1101/2020.10.12.336628 (2020).

96. Scheib, C. L. *et al.* Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **360**, 1024–1027 (2018).

97. Davis, L. G. & Madsen, D. B. The coastal migration theory: Formulation and testable hypotheses. *Quaternary Science Reviews* **249**, 106605 (2020).

98. Li, Y.-C. *et al.* Mitogenome evidence shows two radiation events and dispersals of matrilineal ancestry from northern coastal China to the Americas and Japan. *Cell Reports* **42**, 112413 (2023).

99. Pitulko, V. V. *et al.* The Yana RHS Site: Humans in the Arctic Before the Last Glacial Maximum. *Science* **303**, 52–56 (2004).

100. Pitulko, V., Pavlova, E. & Nikolskiy, P. Revising the archaeological record of the Upper Pleistocene Arctic Siberia: Human dispersal and adaptations in MIS 3 and 2. *Quaternary Science Reviews* **165**, 127–148 (2017).

101. Dillehay, T. D. *et al.* New Archaeological Evidence for an Early Human Presence at Monte Verde, Chile. *PLoS ONE* **10**, e0141923 (2015).

102. Ardelean, C. F. *et al.* Evidence of human occupation in Mexico around the Last Glacial Maximum. *Nature* **584**, 87–92 (2020).

103. Rowe, T. B. *et al.* Human Occupation of the North American Colorado Plateau ~37,000 Years Ago. *Front. Ecol. Evol.* **10**, 903795 (2022).

104. Holliday, V. T. *et al.* Paleolake geochronology supports Last Glacial Maximum (LGM) age for human tracks at White Sands, New Mexico. *Science AdvAnceS* (2025).

1080 105. Bourgeon, L., Burke, A. & Higham, T. Earliest Human Presence in North America
1081 Dated to the Last Glacial Maximum: New Radiocarbon Dates from Bluefish Caves,
1082 Canada. *PLoS ONE* **12**, e0169486 (2017).
1083 106. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of
1084 Native Americans. *Nature* **505**, 87–91 (2014).
1085 107. Lell, J. T. *et al.* The Dual Origin and Siberian Affinities of Native American Y
1086 Chromosomes. *The American Journal of Human Genetics* **70**, 192–206 (2002).
1087 108. García-Bour, J., Pérez-Pérez, A., Prats, E. & Turbón, D. MOLECULAR APPROACH TO
1088 THE PEOPLING OF THE AMERICAS BY SEQUENCING mtDNA FROM EXTINCT
1089 FUEGUIANS AND PATAGONS. *Chungará (Arica)* **32**, (2000).
1090 109. Demarchi, D. A., Panzetta-Dutari, G. M., Motran, C. C., De Los Angeles López De
1091 Basualdo, M. & Marcellino, A. J. Mitochondrial DNA haplogroups in Amerindian
1092 populations from the Gran Chaco. *American J Phys Anthropol* **115**, 199–203
1093 (2001).
1094 110. Merriwether, D. A. & Ferrell, R. E. The Four Founding Lineage Hypothesis for the
1095 New World: A Critical Reevaluation. *Molecular Phylogenetics and Evolution* **5**,
1096 241–246 (1996).
1097 111. Lee & Merriwether. Identification of Whole Mitochondrial Genomes from
1098 Venezuela and Implications on Regional Phylogenies in South America. *Human*
1099 *Biology* **87**, 29 (2015).
1100 112. Malhi, R. S., Schultz, B. A. & Smith, D. G. Distribution of Mitochondrial DNA
1101 Lineages among Native American Tribes of Northeastern North America. *Human*
1102 *Biology* **73**, 17–55 (2001).

1103 113. Batista, O., Kolman, C. J. & Bermingham, E. Mitochondrial DNA diversity in the
1104 Kuna Amerinds of Panama. *Human Molecular Genetics* **4**, 921–929 (1995).
1105 114. Dornelles, C. L. *et al.* Mitochondrial DNA and *Alu* insertions in a genetically
1106 peculiar population: The Ayoreo Indians of Bolivia and Paraguay. *American J Hum*
1107 *Biol* **16**, 479–488 (2004).
1108 115. Brown, M. D. *et al.* mtDNA Haplogroup X: An Ancient Link between
1109 Europe/Western Asia and North America? *The American Journal of Human*
1110 *Genetics* **63**, 1852–1861 (1998).
1111 116. Kashani, B. H. *et al.* Mitochondrial haplogroup C4c: A rare lineage entering
1112 America through the ice-free corridor? *American J Phys Anthropol* **147**, 35–39
1113 (2012).
1114 117. Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population
1115 history of Native Americans. *Science* **349**, aab3884 (2015).
1116 118. Skoglund, P. *et al.* Genetic evidence for two founding populations of the Americas.
1117 *Nature* **525**, 104–108 (2015).
1118 119. Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China.
1119 *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223–2227 (2013).
1120 120. Yang, M. A. *et al.* 40,000-Year-Old Individual from Asia Provides Insight into Early
1121 Population Structure in Eurasia. *Current Biology* **27**, 3202-3208.e9 (2017).
1122 121. Barbieri, C. *et al.* The Current Genomic Landscape of Western South America:
1123 Andes, Amazonia, and Pacific Coast. *Molecular Biology and Evolution* **36**, 2698–
1124 2713 (2019).
1125 122. Clark, P. U. *et al.* The Last Glacial Maximum. *Science* **325**, 710–714 (2009).

1126 123. Häggi, C. *et al.* Response of the Amazon rainforest to late Pleistocene climate
1127 variability. *Earth and Planetary Science Letters* **479**, 50–59 (2017).
1128 124. Ray, N. & Adams, J. M. A GIS-based Vegetation Map of the World at the Last
1129 Glacial Maximum (25,000-15,000 BP). *IA* (2001) doi:10.11141/ia.11.2.
1130 125. Keyeux, G., Rodas, C., Gelvez, N. & Carter, D. Possible Migration Routes into South
1131 America Deduced from Mitochondrial DNA Studies in Colombian Amerindian
1132 Populations. *Human Biology* **74**, 211–233 (2002).
1133 126. Gayà-Vidal, M. *et al.* mtDNA and Y-chromosome diversity in Aymaras and
1134 Quechuas from Bolivia: Different stories and special genetic traits of the Andean
1135 Altiplano populations. *American J Phys Anthropol* **145**, 215–230 (2011).
1136 127. Taboada-Echalar, P. *et al.* The Genetic Legacy of the Pre-Colonial Period in
1137 Contemporary Bolivians. *PLoS ONE* **8**, e58980 (2013).
1138 128. Corella, A., Bert, F., Pérez-Pérez, A., Gené, M. & Turbón, D. Mitochondrial DNA
1139 diversity of the Amerindian populations living in the Andean Piedmont of Bolivia:
1140 Chimane, Moseten, Aymara and Quechua. *Annals of Human Biology* **34**, 34–55
1141 (2007).
1142 129. Arencibia, V. *et al.* Novel B2 mitogenomes from Continental southern Patagonia's
1143 Late Holocene: New insights into the peopling of the Southern Cone. *American*
1144 *Journal of Biological Anthropology* **186**, e24822 (2025).
1145 130. Bravi, C. M., Motti, J. M. B. & García, A. Letter to the editor: A Southern Cone
1146 origin rather than Peruvian affinities for ancient Patagonian B2 mitogenomes.
1147 *American Journal of Biological Anthropology* **184**, e24934 (2024).

1148 131. García, A., Pauro, M., Nores, R., Bravi, C. M. & Demarchi, D. A. Phylogeography of
1149 mitochondrial haplogroup D1: An early spread of subhaplogroup D1j from Central
1150 Argentina. *American J Phys Anthropol* **149**, 583–590 (2012).
1151 132. Bodner, M. *et al.* Rapid coastal spread of First Americans: Novel insights from
1152 South America's Southern Cone mitochondrial genomes. *Genome Res.* **22**, 811–
1153 820 (2012).
1154 133. Lindo, J. *et al.* Ancient individuals from the North American Northwest Coast
1155 reveal 10,000 years of regional genetic continuity. *Proc. Natl. Acad. Sci. U.S.A.* **114**,
1156 4093–4098 (2017).
1157 134. De Souza, F. G. *et al.* Mitochondrial ancestry from complete mitogenomes
1158 highlights a lack of characterization of indigenous haplogroups in Brazilian
1159 Amazon population. *Commun Biol* **8**, 835 (2025).
1160 135. Di Corcia, T. *et al.* East of the Andes: The genetic profile of the Peruvian Amazon
1161 populations. *American J Phys Anthropol* **163**, 328–338 (2017).
1162 136. Ribeiro-dos-Santos, A. M. *et al.* Exome Sequencing of Native Populations From the
1163 Amazon Reveals Patterns on the Peopling of South America. *Front. Genet.* **11**,
1164 548507 (2020).
1165 137. Capodiferro, M. R. *et al.* The multifaceted genomic history of Ashaninka from
1166 Amazonian Peru. *Current Biology* **33**, 1573-1581.e5 (2023).
1167 138. Blench, R. The Role of Agriculture in Explaining the Diversity of Amerindian
1168 Languages. in *The Past Ahead. Language, Culture, and Identity in the Neotropics*
1169 13–38 (ActaUniversitatisUpsaliensis, Uppsala, 2012).

1170