Current Biology 19, 1–8, January 13, 2009 ©2009 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2008.11.058

Article

Distinctive Paleo-Indian Migration Routes from Beringia Marked by Two Rare mtDNA Haplogroups

Ugo A. Perego,[1,2,9] Alessandro Achilli,[1,3,9] Norman Angerhofer,[2] Matteo Accetturo,[1] Maria Pala,[1] Anna Olivieri,[1] Baharak Hooshiar Kashani,[1] Kathleen H. Ritchie,[2] Rosaria Scozzari,[4] Qing-Peng Kong,[5,6] Natalie M. Myres,[2] Antonio Salas,[7] Ornella Semino,[1] Hans-Jürgen Bandelt,[8] Scott R. Woodward,[2] and Antonio Torroni[1,*]

[1]Dipartimento di Genetica e Microbiologia
Università di Pavia
27100 Pavia
Italy

[2]Sorenson Molecular Genealogy Foundation
Salt Lake City, UT 84115
USA

[3]Dipartimento di Biologia Cellulare e Ambientale
Università di Perugia
27100 Perugia
Italy

[4]Dipartimento di Genetica e Biologia Molecolare
Università La Sapienza
00185 Rome
Italy

[5]State Key Laboratory of Genetic Resources and Evolution
Kunming Institute of Zoology
Chinese Academy of Sciences
650223 Kunming, Yunnan
China

[6]Laboratory for Conservation and Utilization of Bio-resource
Yunnan University
650223 Kunming, Yunnan
China

[7]Unidade de Xenética
Instituto de Medicina Legal
Facultad de Medicina
Universidad de Santiago de Compostela
15782 Santiago de Compostela, Galicia
Spain

[8]Department of Mathematics
University of Hamburg
20146 Hamburg
Germany

Summary

Background: It is widely accepted that the ancestors of Native Americans arrived in the New World via Beringia approximately 10 to 30 thousand years ago (kya). However, the arrival time(s), number of expansion events, and migration routes into the Western Hemisphere remain controversial because linguistic, archaeological, and genetic evidence have not yet provided coherent answers. Notably, most of the genetic evidence has been acquired from the analysis of the common pan-American mitochondrial DNA (mtDNA) haplogroups. In this study, we have instead identified and analyzed mtDNAs belonging to two rare Native American haplogroups named D4h3 and X2a.

Results: Phylogeographic analyses at the highest level of molecular resolution (69 entire mitochondrial genomes) reveal that two almost concomitant paths of migration from Beringia led to the Paleo-Indian dispersal approximately 15–17 kya. Haplogroup D4h3 spread into the Americas along the Pacific coast, whereas X2a entered through the ice-free corridor between the Laurentide and Cordilleran ice sheets. The examination of an additional 276 entire mtDNA sequences provides similar entry times for all common Native American haplogroups, thus indicating at least a dual origin for Paleo-Indians.

Conclusions: A dual origin for the first Americans is a striking novelty from the genetic point of view, and it makes plausible a scenario positing that within a rather short period of time, there may have been several entries into the Americas from a dynamically changing Beringian source. Moreover, this implies that most probably more than one language family was carried along with the Paleo-Indians.

Introduction

When and from where did the first Americans arrive, and what migratory routes did they follow? Scientists from several disciplines continue to search for answers to these questions, but, despite new important evidence, the debate concerning the peopling of the Americas is far from resolved [1–3].

Recent data, based on genetic evidence and archaeological and environmental records, have proposed that humans entered the Americas from Beringia as early as 15,000 years ago and that the dispersal occurred along the deglaciated Pacific coastline [4–6]. Mitochondrial DNA (mtDNA) data presented in the scientific literature over the past two decades indicate that the current native populations of North, Central, and South America harbor significant variation from their Asian counterparts and that they belong to four common "pan-American" haplogroups (A2, B2, C1, and D1), which are found all over the double continent, and five minor lineages (C4c, D2a, D3, D4h3, and X2a) [7–15].

As for the pan-American haplogroups A2, B2, C1, and D1, a considerable number of complete Native American mtDNA sequences have become available recently [7, 8, 10, 16]. This has allowed the identification of three different subgroups of C1 (C1b, C1c, and C1d) as ancestral founders [7, 10] and, most importantly, has provided divergence values that are suggestive of a concomitant post-Last Glacial Maximum expansion from Beringia for all pan-American haplogroups. However, because of the use of different calibration and estimation approaches, resulting in time values that range from ~13.5 thousand years ago (kya) to ~19.0 kya [7, 8, 10], these studies have not resolved the controversy concerning the timing for the human entry into the Americas. Moreover, the lower part of this range has been recently narrowed by new archaeological findings, attesting to a human presence prior to 14 kya in both North and South America [4–6, 17].

*Correspondence: torroni@ipvgen.unipv.it
[9]These authors contributed equally to this work

Among the minor haplogroups, D3 and D2a are restricted to Native American populations of northern North America. Haplogroup D3 is confined to the Eskimos [18, 19], whereas D2a is detected in Na-Dene, Eskimos, and Aleuts [10, 20, 21], with a distribution that parallels that of A2a, one of the known subgroups within A2 [10, 21]. The absence of these haplogroups in more southern Amerindian groups and the concomitant detection in Siberian populations [21, 22] support the scenario that Na-Dene, Eskimos, and Aleuts arose, at least in part, from Beringian or Alaskan genetic sources that were geographically and/or temporally distinct from the source(s) that gave rise to the first Americans (Paleo-Indians) [9, 10, 13, 21]. Separate genetic stocks of Beringian origin contributing to the formation of the Na-Dene and the Eskimo-Aleuts would fit most linguistic classification models, whereas a single migratory event accounting for the origin of all other Native American populations from North, Central, and South America is linguistically extremely controversial [23, 24], but overall widely accepted by human geneticists [7, 10, 15].

Adding to the uncertainty regarding entry times and migratory paths is the poor representation of the other minor haplogroups in the published data. For instance, only 12 complete sequences for haplogroup X2a and two for haplogroup D4h3 have been reported in the literature to date [8, 10, 15, 25], leaving important unanswered questions concerning their origins and dispersal into the American continent.

Haplogroup X2a was first recognized as a fifth Native American haplogroup in 1996 [9] and then reported at a very high frequency (25%) in the Ojibwas of Manitoulin Island (northern Ontario) [26]. Later studies found X2a at lower frequencies also among the Nuu-Chah-Nulth, the Sioux, the Yakima, and even the Navajo [14, 27], although the presence in the Navajo (Southern Na-Dene) is most probably due to recent admixture with other northern Native American groups. Unlike in the case of all other Native American haplogroups, a close molecular counterpart for X2a has not been found in Asians [25], suggesting that its X2 ancestor became lost in Asians after entry in Beringia, most probably because of genetic drift.

The control-region motif for haplogroup D4h3, maintaining the nomenclature proposed by Tamm [10], was first identified in the Cayapa of Ecuador nearly a decade ago [28]. In 2007, the 10,300-year-old skeletal remains from the On Your Knees Cave on Prince of Wales Island (Alaska) were found to harbor the same control-region motif, thus indicating that an additional haplogroup, whose geographical distribution was mainly along the Pacific coast, was indeed present in early Native Americans [29].

In this study, to shed light on the origin of Paleo-Indians, we used an approach that is different from those employed in past mtDNA studies. Instead of analyzing mtDNAs belonging to the common pan-American haplogroups, we completely sequenced and analyzed a large number of mtDNAs belonging to the rare and poorly known haplogroups D4h3 and X2a, revealing that each marked a distinct entry path from Beringia, which contributed to the formation of Paleo-Indians.

Results

To identify mtDNAs belonging to the rare Native American haplogroups D4h3 and X2a, we searched the Sorenson Molecular Genealogy Foundation (SMGF) mtDNA database [30] for control-region mutational motifs matching, at least partially, the diagnostic motifs for D4h3 (152-489-16223-16241-16301-16342-16362) and X2a (153-195-200-225-16183C-16189-16213-16223-16278-16519). For haplogroup X2a, some additional mtDNAs were obtained from an Ojibwa sample collection [26]. A total of 55 unrelated mtDNAs, from either Native American populations or American populations of mixed origin, potentially belonging to D4h3 (n = 44) and X2a (n = 11) were identified and completely sequenced.

A phylogenetic analysis of these mitochondrial genomes and 14 additional complete sequences available in the literature was performed, with the tree rooted with a published L3a mtDNA sequence (Table S1 available online). The branches of the tree formed by the D4h3 and X2a complete sequences are illustrated in Figure 1. All mutations observed in each genome are reported together with the roots of the pan-American haplogroups A2, B2, C1b, C1c, C1d, and D1 in Figure S1.

Haplogroup D4h3

One of the 44 sequenced D4h3 mtDNAs was actually from Eastern China (Qingdao, Shandong province) and to date is the only reported mtDNA from Asia [31] with a control-region mutational motif (16301–16342) similar to that seen in Native American D4h3 mtDNAs. The complete sequence of the Chinese mtDNA (#01 in Figure 1 and Figure S1) confirmed its D4h3 membership. In fact, the D4h3 branch shows an initial deep split into two sister subclades, one encompassing all D4h3 mtDNAs from the Americas (D4h3a) and the other consisting only of the sequence from China (D4h3b), thus posing the upper limit for the most recent common female Asian ancestor of D4h3. The Native American branch D4h3a is mostly found in South America, with the exception of eight samples from Mexico and two found in California. Several basal branches characterize D4h3a. Among these, D4h3a1 and D4h3a2, found only in Chile, have accumulated a large amount of internal variation, whereas D4h3a3, defined only by a control-region deletion at nucleotide position (np) 71 (Figure S1), is slightly less differentiated but shared between a single Chilean representative and a northern subcluster (Mexico and California).

Haplogroup X2a

As for haplogroup X2a, all but one of the sequenced mitochondrial genomes harbored the distinguishing X2a coding-region motif 8913-12397-14502. The exception was one of the Ojibwa sequences (#47 in Figure 1 and Figure S1), which did not cluster either with X2a or any of the known Old World X2 branches (X2b–X2f) [25, 32]. This novel X2 branch has been named X2g, and its presence in Native Americans most probably indicates an additional and very rare Native American founder. The two internal X2a branches—X2a1 and X2a2—can be easily distinguished on the basis of their private control-region mutations (143-16093 and 225-16254C, respectively). Thus, whereas most X2a mtDNAs in the Great Lakes and the Great Plains regions belong to X2a1 and others are members of the less common X2a2, some mtDNAs observed at the western fringe of the X2a distribution area (Nuu-Chah-Nulth and Yakima) might constitute a branch of X2a distinct from the former two inasmuch as they lack at least one characteristic X2a1 and X2a2 mutation [14, 33, 34]. Such a geographic pattern suggests that the X2a root was the founder sequence and that the X2a branches arose in situ.

Expansion Times for D4h3a, X2a, and the Pan-American Haplogroups into the Americas

As evident from Figure 1, maximum-likelihood (ML) coding-region divergences for haplogroups D4h3a and X2a show a significant overlap, with values of $\sim$0.00022 $\pm$ 0.00003 and

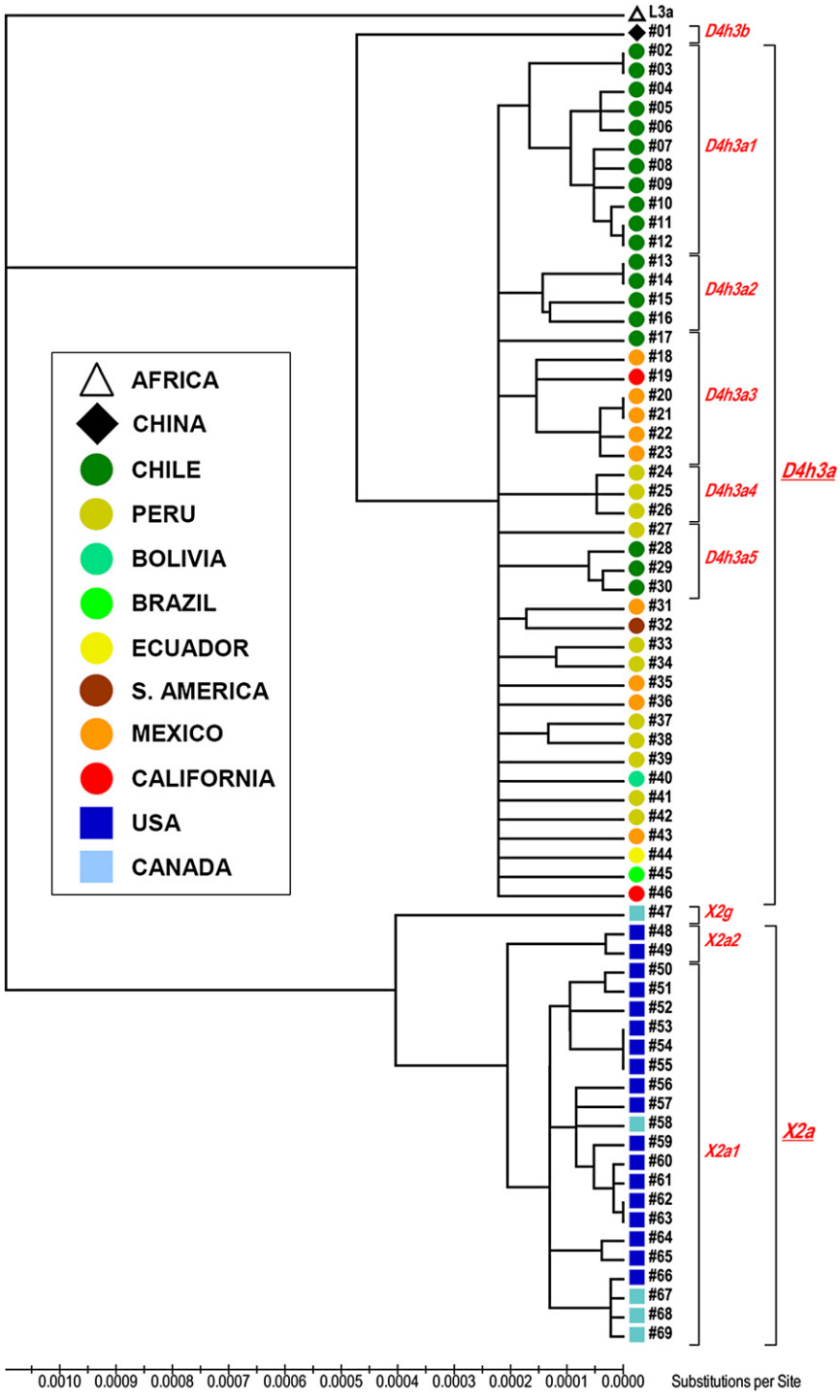

Figure 1. Schematic Phylogeny of Complete MtDNA Sequences Belonging to Haplogroups D4h3a and X2a

The tree was rooted with a published L3a sequence [54]. A maximum-likelihood (ML) time scale is shown on its bottom. The exact values for important (sub)clades are also available in Table 1 together with averaged distance (ρ) of the haplotypes of a clade to the respective root haplotype, accompanied by a heuristic estimate of standard error (SE [σ]). All sequences within D4h3a and X2a branches are new except for #32, #44, #48, #52–56, #60, #61, #63, #65, #66, and #69 (Table S1). Two additional novel clades, here defined as D4h3b (#01) and X2g (#47), are also included. Details concerning phylogeny construction and mutational motifs of each sequence are provided in Figure S1.

0.00021 ± 0.00006 substitutions per site, respectively. This is also confirmed when the average distance of the haplotypes from the root of each haplogroup (ρ statistics) is computed (Table 1).

Even when standard errors are considered, the largely overlapping sequence divergences observed for haplogroups D4h3 and X2a indicate a concomitant or very close temporal expansion of the two haplogroups, from either the same Beringian source population or different yet related Beringian sources. In either case, D4h3a and X2a did not expand alone, but did so together with at least a subset of the other Native American haplogroups. To evaluate this issue in detail, we included 276 complete coding-region sequences belonging to A2, B2, C1b, C1c, C1d, and D1 (Table S2) in the phylogeny. The resulting tree, again rooted on L3a, is reported in Figure 2, in which each Native American clade has been condensed in a triangle with the base and the height representing the haplogroup internal variation and the ML branch length, respectively. The ancestral tip of all the triangles correspond to similar values ranging between 0.00021 and 0.00025 substitutions per site (Table 2). A close similarity was also obtained for the average sequence divergences, which we calculated by considering all coding-region base substitutions as well as only synonymous transitions (Table 3). Because the mutation rate of Mishmar [35] is probably an overestimate [7, 8], mainly because of partial saturation of some synonymous mutations, and that of Kivisild [36] represents an underestimate [7, 10], we would now recommend taking an intermediate global coalescence time of modern human mtDNA as a reference point for the internal calibration. Accordingly, we converted the substitution distances from the root of each haplogroup into time estimates by using averaged time calibrations (see Experimental Procedures) corresponding to 4610 years per coding-region substitution and 7650 years per synonymous transition (Table 3 and Figure S1). The average age of the Native American clades that we obtained by applying the two different approaches is approximately 16 thousand years (ky) (16.7 and 15.5 ky, respectively), with the ages of the single clades falling into the range from 14.2 to 18.7 ky (except for the outlier age of C1d).

Spatial Distribution of Haplogroups D4h3a and X2a

Intriguingly, the similarity in coalescence times is not reflected in terms of spatial distribution. The phylogeography of the uncommon Native American haplogroups D4h3a and X2a in

Table 1. Divergence Estimates for Haplogroups D4h3 and X2 and Their Subclades

Haplogroups/ Subhaplogroups	Number of mtDNAs	Maximum Likelihood Substitutions per site	SE	ρ Statistics[a] ρ	σ
D4h3	46	0.00047	0.00009	8.870	2.261
D4h3a	45	0.00022	0.00003	3.911	0.580
D4h3a1	11	0.00017	0.00004	4.000	1.408
D4h3a2	4	0.00014	0.00006	2.500	1.061
D4h3a3	7	0.00022	0.00003	2.286	0.904
D4h3a4	3	0.00005	0.00003	0.333	0.333
D4h3a5	4	0.00022	0.00003	5.500	1.732
X2	23	0.00040	0.00010	6.087	1.964
X2a	22	0.00021	0.00006	3.091	1.097
X2a2	2	0.00013	0.00004	0.500	0.500
X2a1	20	0.00009	0.00003	2.050	0.650
X2a1a	6	0.00008	0.00003	1.500	0.687
X2a1b	8	0.00003	0.00003	1.750	0.848

[a] The ρ estimate was computed as average number of base substitutions in the mtDNA coding region (between nps 577 and 16,023) from the ancestral sequence type, with a σ error derived from an estimate of the genealogy [18].

the double continent is strikingly different, with the former being found along the Pacific coast, including in the skeletal remains from Alaska dated 10,300 years ago [29], and the latter being restricted to northern North America, with no instances detected south of the United States. A search for the D4h3 control-region motif in the SMGF mtDNA database [30] and in the literature allowed an evaluation of its geographic distribution both in general mixed populations of national states and autochthonous Native American groups. The mtDNA #01 in Figure 1 remains the only case found in Asia, indicating a frequency lower than 10^{-4}. In contrast, D4h3a—the Native American counterpart—is less rare (Figure 3 and Table S3). This haplogroup is detected at low frequencies in modern and ancient population samples of both North and South America, with an overall higher frequency and peaks along the Pacific coast and the western side of the Andes. An analogous query for the X2a control-region motif confirmed that this haplogroup is confined to northern North America, with a frequency peak in the Great Lakes area (Figure 3 and Table S3).

The overlapping sequence divergences indicate that the Beringian haplogroup D4h3a started spreading in America at approximately the same time as X2a. The rapid dispersion of D4h3a southward along the Pacific coast is supported by two deep subclades (D4h3a1 and D4h3a2), found exclusively in Chile, and by its spatial distribution. This pattern differs somewhat from those of the pan-American Native American haplogroups in that it is restricted to near-coastal areas in the west of the double continent, but it fully supports the hypothesis that the Pacific coast was the major entry and diffusion route for the early Paleo-Indians [4, 6]. In contrast, the other rare Native American haplogroup, X2a, despite a similar expansion time, is restricted to northern North America, with a focus in the Great Lakes and the Great Plains regions.

Discussion

How can the different geographical distributions of haplogroups D4h3a and X2a be explained? A priori, we can envision two possible scenarios. First, haplogroup X2a could have entered into the Americas by following the Pacific coastal path just like D4h3a and the pan-American haplogroups.

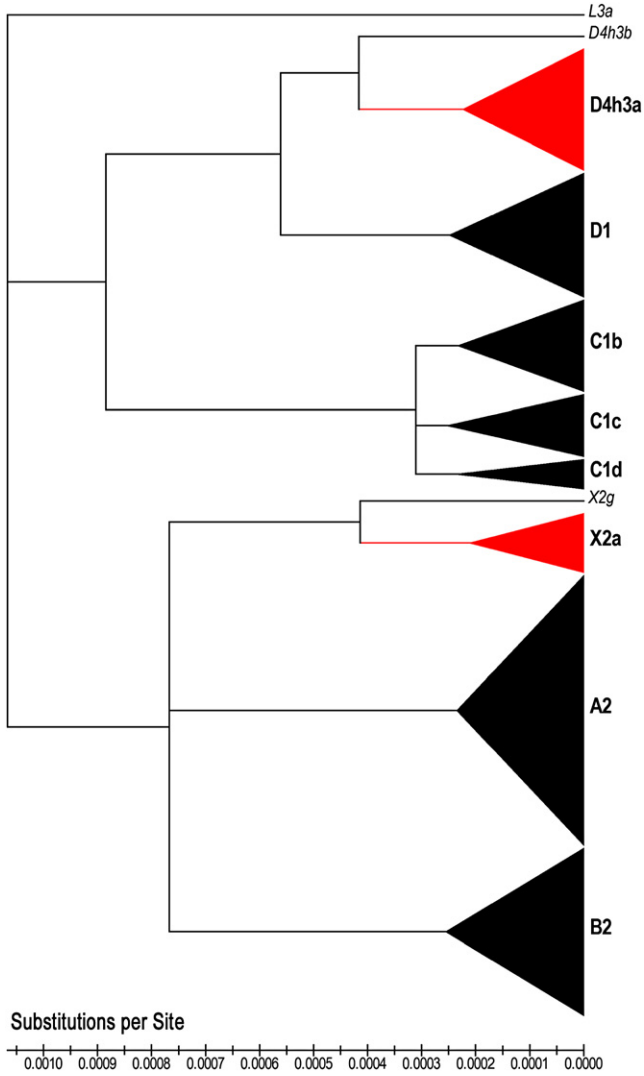

Figure 2. Evolutionary Relationships of 344 Complete Native American Coding-Region Sequences

The sequences belong to the pan-American haplogroups (A2, B2, C1b, C1c, C1d, and D1; Table S2) and to the uncommon Native American clades D4h3a and X2a (Table S1). The evolutionary history was inferred by hand with the maximum-parsimony method, with all the available pan-American coding-region sequences included. The tree is drawn to scale. The evolutionary distances were computed with the maximum-likelihood method (see Experimental Procedures) and are in the units of number of base substitutions per site. All positions containing gaps and ambiguous data were eliminated from the data set. The triangle width is proportional to the number of sequences, thus representing the internal haplogroup variation.

However, in contrast to the other mtDNA haplogroups, it would have disappeared from the wave front of the expanding population at a very early stage, thus remaining confined to some coastal enclave from where it expanded exclusively toward the east after the complete retreat of the Cordilleran ice sheet. This scenario appears rather unlikely, given that it would require the disappearance of a lineage from the front of an expansion wave involved in the colonization of a previously unpopulated territory, a demographic situation in which the loss of lineages due to genetic drift is minimized [37].

Alternatively, haplogroup X2a might have arrived from Beringia through a path that was different from that followed

Table 2. Maximum-Likelihood Divergence of Relevant Nodes in the Native American Phylogeny of Figure 2

Node/Clade	Number of mtDNAs	Substitutions per site[a]	SE
L3	346	0.00107	0.00011
M	160	0.00088	0.00011
D	92	0.00056	0.00009
D4h3	46	0.00042	0.00007
D4h3a	45	0.00022	0.00003
D1	46	0.00025	0.00002
C1	68	0.00031	0.00005
C1b	36	0.00023	0.00003
C1c	23	0.00025	0.00003
C1d	9	0.00023	0.00004
N	185	0.00077	0.00010
X2	23	0.00041	0.00009
X2a	22	0.00021	0.00006
A2	100	0.00023	0.00002
B2	62	0.00025	0.00002

[a] All positions containing gaps and missing data were eliminated from the data set. There were a total of 15,427 positions in the final data set.

by the pan-American haplogroups. According to environmental and paleoecological data, such a path existed and was represented by the ice-free corridor between the Laurentide and Cordilleran ice sheets, which opened approximately 15 kya [5] or possibly was never completely closed [38]. Through such a corridor, where some glacial-refuge areas have been recently identified [39], X2a could have moved from Beringia directly into the North American regions located east of the Rocky Mountains. This latter scenario would imply that the X2a expansion in America occurred in the Great Plains region, where the terminal part of the glacial corridor ended, and is in complete agreement with both the extent of diversity and distribution of X2a observed in modern Native American populations.

In contrast, haplogroup D4h3 clearly signifies the coastal route. However, the fact that it is very rare in North America and mainly found in South America with deep variation needs some explanation. One could envision that the front wave of the initial coastal-settlement process had most or all pan-American mtDNA haplogroups in its package, including haplogroup D4h3a. Then, the back wave(s) with a slightly shifted haplogroup composition but no D4h3a mtDNAs might have had a kind of push effect in North America (and predominantly founded various riverine offshoots, as simulated by Fix [40]) and led to the settlement along the western coast of South America. This explanation would thus lower the impact of subsequent drift in a strict one-wave model as the sole factor explaining different regional mtDNA-haplogroup compositions and different physical traits in space and time.

It is tempting to relate the different early entries to the archaeological record of Beringia and North America. Perhaps the two inferred dispersal events were responsible for the emergence of the two major Paleo-Indian traditions of North America—fluted-point industries in eastern U.S. and stemmed-point industries in western U.S. [41]. Greater Beringia also carried well distinguishable archaeological traditions manifest in the Allerød, viz. the Nenana complex (after the earlier Swan Point microblade industry) and the Sluiceway-Tuluaq complex in the east and the early Ushki complex of Kamchatka [5]. This would suggest that Beringia at the time (and also earlier) was settled by a nonhomogeneous population group. It would, however, be premature—without direct evidence of ancient DNA findings from those areas—to see this division as exactly corresponding to the two entry routes of the first migrants.

Conclusion

The phylogeography of the rare Native American haplogroups D4h3a and X2a, when evaluated at the level of complete mitochondrial genomes, indicates that at least two migration routes from Beringia were used at approximately the same time by the Paleo-Indians. The Pacific coastal path probably played the major role in the peopling of the double continent, but the entry through the ice-free corridor between the Laurentide and Cordilleran ice sheets also had a significant impact, at least for the colonization of northern North America. However, a precise quantification of the relative inputs in North America will be possible only after identification and evaluation of the

Table 3. Averaged Substitution Distance from Relevant Nodes in the Native American Phylogeny of Figure 2

		All Coding-Region Base Substitutions				Only Synonymous Transitions			
Node/Clade	Number of mtDNAs	ρ[a]	σ[b]	T[c] (ky)	ΔT (ky)	ρ	σ	T[c] (ky)	ΔT (ky)
L3	346	16.812	1.990	77.4	9.2	9.266	1.524	70.9	11.7
M	160	13.550	2.054	62.4	9.5	6.644	1.372	50.8	10.5
D4	92	8.772	1.604	40.4	7.4	4.457	1.130	34.1	8.6
D4h3	46	8.870	2.261	40.9	10.4	4.783	1.734	36.6	13.3
D4h3a	45	3.911	0.580	18.0	2.7	1.867	0.374	14.3	2.9
D1	46	3.674	0.431	16.9	2.0	2.130	0.315	16.3	2.4
C1	68	4.250	0.702	19.6	3.2	2.222	0.699	17.0	5.3
C1b	36	3.583	0.824	16.5	3.8	2.222	0.699	17.0	5.3
C1c	23	3.739	0.671	17.2	3.1	2.304	0.565	17.6	4.3
C1d	9	2.111	0.533	9.7	2.5	1.000	0.333	7.6	2.5
N	185	12.059	1.859	55.5	8.6	5.941	1.213	45.4	9.3
X2	23	6.087	1.964	28.0	9.0	3.217	1.387	24.6	10.6
X2a	22	3.091	1.097	14.2	5.1	2.227	1.047	17.0	8.0
A2	100	3.420	0.289	15.8	1.3	1.960	0.232	15.0	1.8
B2	62	4.065	0.371	18.7	1.7	2.161	0.270	16.5	2.1
All Native Clades[d]	343	3.618	1.505	16.7	6.9	2.026	1.201	15.5	9.2

[a] The average number of base substitutions in the mtDNA coding region (between positions 577 and 16,023) from the root sequence type.
[b] Standard error calculated from an estimate of the genealogy [18].
[c] Estimate of the time to the most recent common ancestor of each cluster, using a corrected age estimate of approximately 4610 years per substitution in the whole coding region and approximately 7650 years per synonymous transition (see Experimental Procedures).
[d] Including A2, B2, C1b, C1c, C1d, D1, D4h3a, and X2a.

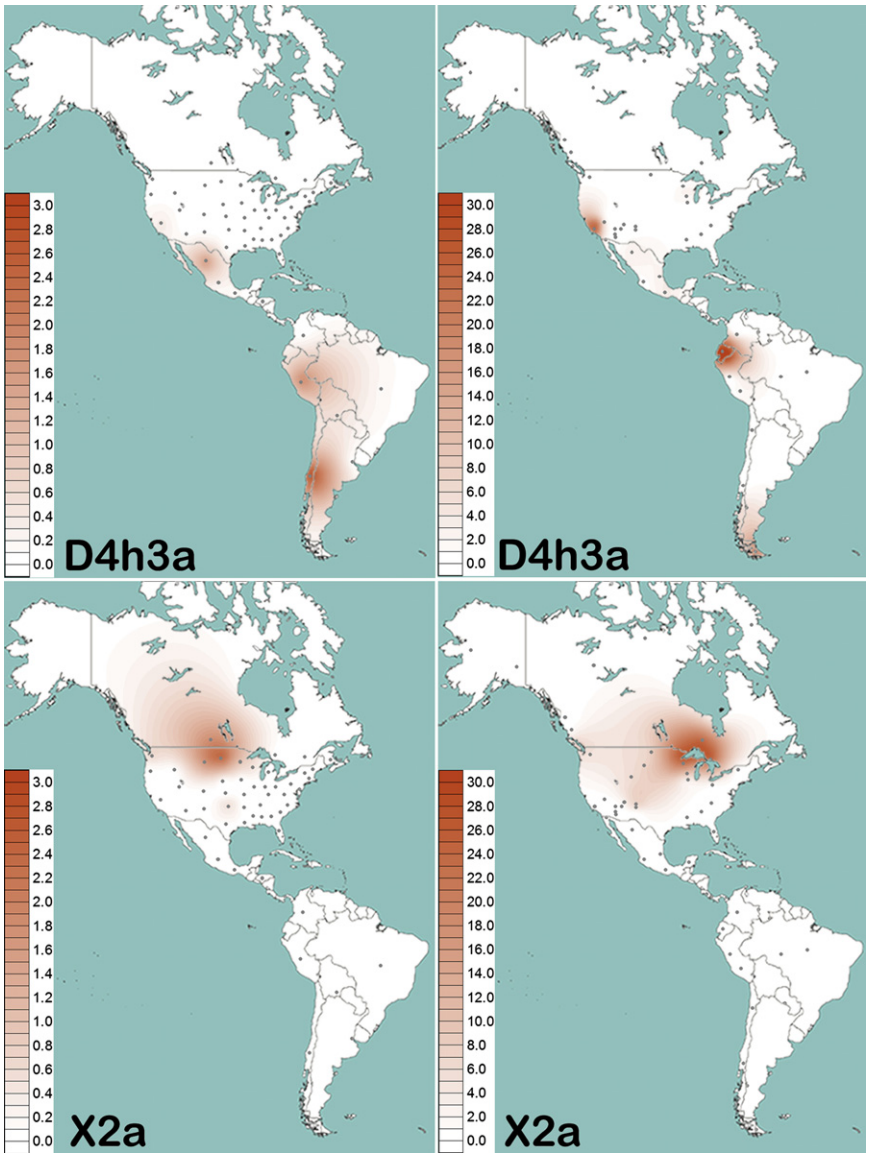

Figure 3. Spatial-Frequency Distributions of Haplogroups D4h3a and X2a

The upper maps show the frequency distributions of haplogroup D4h3a, and the lower maps depict X2a frequencies. On the left are shown data from general mixed populations of national states, and on the right are those from Native American groups. Note that the frequency scales (%) employed for general mixed populations and Native American groups are different. The dots indicate the location of the population samples included in each survey (Tables S3 and S4). Frequency maps of haplogroups were obtained as in Olivieri et al. [55].

spatial patterns of the subclades within the common haplogroups A2, B2, C1b, C1c, C1d, and D1. Indeed, haplogroup D4h3a and X2a did not enter into the Americas alone. Thus, if some internal subclades of the pan-American haplogroups [7] arose before the expansion into the continent, their geographic distribution patterns should parallel those reported here for haplogroups X2a and D4h3a.

Overall, these mtDNA findings make plausible a scenario positing that within a rather short period of time during the pre-Bølling interstadial (15.8 to 14.9 kya [42]) and during the warmer Bølling interstadial (14.5 to 13.9 kya), the ice-free corridor may have opened for successful southward migration, whereas the Pacific coastal path may have been feasible somewhat earlier, but not before 17 kya, allowing for successive small-scale migrations of Beringian groups. Somewhat later, three short-lived cooling events (intra-Bølling cold period, Older Dryas, and intra-Allerød cold period [43]) may have triggered additional coastal north-to-south movements into (and within) North America from related and dynamically changing Beringian and North American sources. Parts of the coast and the interior of Beringia were probably settled with short interruptions, possibly of only a few hundred years, in response to relatively cooler and drier conditions. This would have caused differently composed population groups to retreat and expand periodically from core enclaves, sometimes intermingling with or integrating with distantly related or even unrelated population groups newly arrived from regions located west of Beringia.

The detection of two distinct entry paths for Paleo-Indians, both possibly exploited intermittently by a dynamic Beringian gene pool that, despite the narrow time window, was continuously reshaped not only by drift, but also gene flow, would also have an impact on some linguistic hypotheses. In fact, the scenario of a structured and temporally changing Beringian source population makes it most unlikely that only a single language family was carried along with the first Pleistocene groups of migrants [23, 44]. Some of these different languages may have been in close contact for several hundred years in Beringia and on the move into the Americas, rendering a distinction between contact features and inherited traits virtually impossible in retrospect after a few thousand years. Consequently, from a standard linguistic point of view, Greenberg's Amerind hypothesis might be regarded as a claim that can neither be validated nor dismissed [45].

The traditional three-wave model [46] and the now-popular three-stage model for the peopling of the Americas [8, 10, 15, 47, 48] that emphasizes a single origin are somewhat too simplistic to explain the initial and subsequent processes that eventually led to the settlement of the Americas. The use of simple models in simulation studies [40, 49, 50] cannot replace a more nuanced interpretation of archaeological findings and genetic variation. The mitochondrial DNA record suggests that the resident population of (greater) Beringia was never replaced completely but, on the other hand, did not develop in complete isolation until the mid-Holocene. The initial Pleistocene migrations into the Americas may have occurred within a narrow time window of no more than 2000 years, with a succession of temporally distinct movements along the coastal route and the ice-free corridor.

Experimental Procedures

Analysis of MtDNA Sequence Variation

Sequencing of entire mtDNAs and phylogeny construction were performed as previously described [51].

Mutation-Rate Estimate

We propose here for the first time a new mutation rate taking into account the previous estimates reported by Mishmar [35] for all coding-region base substitutions and by Kivisild [36] for only synonymous transitions. With three decimal digits used throughout, the rounded values were 5140 years per coding-region substitution and 6760 years per synonymous transition, respectively. The rho estimated (average distance of the haplotypes of a clade from the respective root) human coalescence times are then 202 kya according to Mishmar et al. [35] and 160 kya according to Kivisild et al. [36]. The postulated time obtained as the arithmetic mean of both estimates is approximately 181 ± 21 kya. Thus, ages estimated considering all the coding-region substitutions have to be decreased by a factor of $181/202 \approx 0.90$, whereas the estimates based only on synonymous transitions have to be increased by a factor of $181/160 \approx 1.13$. Given that $5140 \times 181/202 \approx 4610$ and $6760 \times 181/160 \approx 7650$, we obtained the averaged calibrations.

Testing the Molecular-Clock and Maximum-Likelihood Estimates

Assuming the HKY85 mutation model [52] (with indels ignored, as usual) with gamma-distributed rates (approximated by a discrete distribution with 32 categories), we used PAML 3.15 [53] to check the molecular-clock hypothesis by maximizing the likelihood over the branch lengths and the parameters κ (the parameter for the transition/transversion ratio in the HKY85 model) and α (the heterogeneity parameter in the gamma distribution), both with and without the clock enforced. Then we used a likelihood-ratio test to compare the maximized likelihoods. This analysis was performed on the tree of Figure 2 with a published L3a sequence [54] as an outgroup. We could not reject the null hypothesis that the molecular clock describes the evolution of these sequences. The generalized likelihood-ratio statistic was 343.61, which, assuming the asymptotic χ^2 (345) distribution as valid, yields a p value of 0.43 under the clock hypothesis. The maximum likelihood estimates of κ and α under the clock model were 35.14 ± 4.99 and 0.12 ± 0.01, respectively.

Accession Numbers

Fifty-five novel mtDNA sequences have been deposited in GenBank under accession numbers FJ168712–FJ168766.

Supplemental Data

Supplemental Data include one figure and four tables and can be found with this article online at http://www.current-biology.com/supplemental/S0960-9822(08)01618-7.

Acknowledgments

We are grateful to all the donors for providing biological specimens, to J. Edgar Gómez-Palmieri, Ann Turner, and everyone else at the Sorenson Molecular Genealogy Foundation for their assistance with the preliminary data, and to four anonymous reviewers for their useful comments and suggestions. This research received support from Progetti Ricerca Interesse Nazionale 2007 (Italian Ministry of the University) (to R.S., O.S., and A.T.), the Sorenson Molecular Genealogy Foundation (to S.R.W.), Compagnia di San Paolo (to O.S. and A.T.), and Fondazione Cariplo (to A.T.).

Received: October 24, 2008
Revised: November 25, 2008
Accepted: November 26, 2008
Published online: January 8, 2009

References

1. Hey, J. (2005). On the number of New World founders: A population genetic portrait of the peopling of the Americas. PLoS Biol. *3*, e193.
2. Schurr, T.G. (2004). The peopling of the New World: Perspectives from molecular anthropology. Annu. Rev. Anthropol. *33*, 551–583.
3. Waters, M.R., and Stafford, T.W., Jr. (2007). Redefining the age of Clovis: Implications for the peopling of the Americas. Science *315*, 1122–1126.
4. Dillehay, T.D., Ramirez, C., Pino, M., Collins, M.B., Rossen, J., and Pino-Navarro, J.D. (2008). Monte Verde: seaweed, food, medicine, and the peopling of South America. Science *320*, 784–786.
5. Goebel, T., Waters, M.R., and O'Rourke, D.H. (2008). The late Pleistocene dispersal of modern humans in the Americas. Science *319*, 1497–1502.
6. Gilbert, M.T.P., Jenkins, D.L., Götherström, A., Naveran, N., Sanchez, J.J., Hofreiter, M., Thomsen, P.F., Binladen, J., Higham, T.F.G., Yohe, R.M., II, et al. (2008). DNA from pre-Clovis human coprolites in Oregon, North America. Science *320*, 786–789.
7. Achilli, A., Perego, U.A., Bravi, C.M., Coble, M.D., Kong, Q.P., Woodward, S.R., Salas, A., Torroni, A., and Bandelt, H.J. (2008). The phylogeny of the four pan-American mtDNA haplogroups: Implications for evolutionary and disease studies. PLoS ONE *3*, e1764.
8. Fagundes, N.J., Kanitz, R., Eckert, R., Valls, A.C., Bogo, M.R., Salzano, F.M., Smith, D.G., Silva, W.A., Jr., Zago, M.A., Ribeiro-dos-Santos, A.K., et al. (2008). Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. Am. J. Hum. Genet. *82*, 583–592.
9. Forster, P., Harding, R., Torroni, A., and Bandelt, H.-J. (1996). Origin and evolution of Native American mtDNA variation: A reappraisal. Am. J. Hum. Genet. *59*, 935–945.
10. Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D.G., Mulligan, C.J., Bravi, C.M., Rickards, O., Martinez-Labarga, C., Khusnutdinova, E.K., et al. (2007). Beringian standstill and spread of Native American founders. PLoS ONE *2*, e829.
11. Torroni, A., Schurr, T.G., Cabell, M.F., Brown, M.D., Neel, J.V., Larsen, M., Smith, D.G., Vullo, C.M., and Wallace, D.C. (1993). Asian affinities and continental radiation of the four founding Native American mtDNAs. Am. J. Hum. Genet. *53*, 563–590.
12. Schurr, T.G., Ballinger, S.W., Gan, Y.Y., Hodge, J.A., Merriwether, D.A., Lawrence, D.N., Knowler, W.C., Weiss, K.M., and Wallace, D.C. (1990). Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am. J. Hum. Genet. *46*, 613–623.
13. Torroni, A., Schurr, T.G., Yang, C.C., Szathmary, E.J.E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowler, W.C., Lawrence, D.N., Weiss, K.M., et al. (1992). Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics *130*, 153–162.
14. Brown, M.D., Hosseini, S.H., Torroni, A., Bandelt, H.-J., Allen, J.C., Schurr, T.G., Scozzari, R., Cruciani, F., and Wallace, D.C. (1998). MtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? Am. J. Hum. Genet. *63*, 1852–1861.
15. Bandelt, H.-J., Herrnstadt, C., Yao, Y.-G., Kong, Q.-P., Kivisild, T., Rengo, C., Scozzari, R., Richards, M., Villems, R., Macaulay, V., et al. (2003). Identification of Native American founder mtDNAs through the analysis of complete mtDNA sequences: Some caveats. Ann. Hum. Genet. *67*, 512–524.
16. Hartmann, A., Thieme, M., Nanduri, L.K., Stempfl, T., Moehle, C., Kivisild, T., and Oefner, P.J. (2008). Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. Hum. Mutat., in press. Published online July 11, 2008. 10.1002/humu.20816.
17. Buchanan, B., Collard, M., and Edinborough, K. (2008). Paleoindian demography and the extraterrestrial impact hypothesis. Proc. Natl. Acad. Sci. USA *105*, 11651–11654.
18. Saillard, J., Forster, P., Lynnerup, N., Bandelt, H.-J., and Nørby, S. (2000). MtDNA variation among Greenland Eskimos: The edge of the Beringian expansion. Am. J. Hum. Genet. *67*, 718–726.
19. Helgason, A., Pálsson, G., Pedersen, H.S., Angulalik, E., Gunnarsdóttir, E.D., Yngvadóttir, B., and Stefánsson, K. (2006). MtDNA variation in Inuit populations of Greenland and Canada: Migration history and population structure. Am. J. Phys. Anthropol. *130*, 123–134.
20. Derbeneva, O.A., Sukernik, R.I., Volodko, N.V., Hosseini, S.H., Lott, M.T., and Wallace, D.C. (2002). Analysis of mitochondrial DNA diversity in the Aleuts of the Commander islands and its implications for the genetic history of Beringia. Am. J. Hum. Genet. *71*, 415–421.
21. Volodko, N.V., Starikovskaya, E.B., Mazunin, I.O., Eltsov, N.P., Naidenko, P.V., Wallace, D.C., and Sukernik, R.I. (2008). Mitochondrial genome diversity in arctic Siberians, with particular reference to the

evolutionary history of Beringia and Pleistocenic peopling of the Americas. Am. J. Hum. Genet. *82*, 1084–1100.

22. Derenko, M., Malyarchuk, B., Grzybowski, T., Denisova, G., Dambueva, I., Perkova, M., Dorzhu, C., Luzina, F., Lee, H.K., Vanecek, T., et al. (2007). Phylogeographic analysis of mitochondrial DNA in northern Asian populations. Am. J. Hum. Genet. *81*, 1025–1041.
23. Greenberg, J.H. (1987). Language in the Americas (Stanford, CA: Stanford University Press).
24. Campbell, L. (1997). American Indian Languages: The Historical Linguistics of Native America (New York: Oxford University Press).
25. Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., Tambets, K., Tolk, H.V., Parik, J., Loogväli, E.L., Derenko, M., Malyarchuk, B., et al. (2003). Origin and diffusion of mtDNA haplogroup X. Am. J. Hum. Genet. *73*, 1178–1190.
26. Scozzari, R., Cruciani, F., Santolamazza, P., Sellitto, D., Cole, D.E., Rubin, L.A., Labuda, D., Marini, E., Succa, V., Vona, G., et al. (1997). MtDNA and Y chromosome-specific polymorphisms in modern Ojibwa: Implications about the origin of their gene pool. Am. J. Hum. Genet. *60*, 241–244.
27. Smith, D.G., Malhi, R.S., Eshleman, J., Lorenz, J.G., and Kaestle, F.A. (1999). Distribution of mtDNA haplogroup X among Native North Americans. Am. J. Phys. Anthropol. *110*, 271–284.
28. Rickards, O., Martínez-Labarga, C., Lum, J.K., De Stefano, G.F., and Cann, R.L. (1999). MtDNA history of the Cayapa Amerinds of Ecuador: Detection of additional founding lineages for the Native American populations. Am. J. Hum. Genet. *65*, 519–530.
29. Kemp, B.M., Malhi, R.S., McDonough, J., Bolnick, D.A., Eshleman, J.A., Rickards, O., Martínez-Labarga, C., Johnson, J.R., Lorenz, J.G., Dixon, E.J., et al. (2007). Genetic analysis of early Holocene skeletal remains from Alaska and its implications for the settlement of the Americas. Am. J. Phys. Anthropol. *132*, 605–621.
30. SMGF: The Sorenson Molecular Genealogy Foundation Mitochondrial Database. (http://www.smgf.org).
31. Yao, Y.-G., Kong, Q.-P., Bandelt, H.-J., Kivisild, T., and Zhang, Y.P. (2002). Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am. J. Hum. Genet. *70*, 635–651.
32. Shlush, L.I., Behar, D.M., Yudkovsky, G., Templeton, A., Hadid, Y., Basis, F., Hammer, M., Itzkovitz, S., and Skorecki, K. (2008). The Druze: A population genetic refugium of the Near East. PLoS ONE *3*, e2105.
33. Shields, G.F., Schmiechen, A.M., Frazier, B.L., Redd, A., Voevoda, M.I., Reed, J.K., and Ward, R.H. (1993). MtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am. J. Hum. Genet. *53*, 549–562.
34. Ward, R.H., Frazier, B.L., Dew-Jager, K., and Pääbo, S. (1991). Extensive mitochondrial diversity within a single Amerindian tribe. Proc. Natl. Acad. Sci. USA *88*, 8720–8724.
35. Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A.G., Hosseini, S., Brandon, M., Easley, K., Chen, E., Brown, M.D., et al. (2003). Natural selection shaped regional mtDNA variation in humans. Proc. Natl. Acad. Sci. USA *100*, 171–176.
36. Kivisild, T., Shen, P., Wall, D.P., Do, B., Sung, R., Davis, K., Passarino, G., Underhill, P.A., Scharfe, C., Torroni, A., et al. (2006). The role of selection in the evolution of human mitochondrial genomes. Genetics *172*, 373–387.
37. Edmonds, C.A., Lillie, A.S., and Cavalli-Sforza, L.L. (2004). Mutations arising in the wave front of an expanding population. Proc. Natl. Acad. Sci. USA *101*, 975–979.
38. Fagan, B.M. (2004). The Great Journey: The Peopling of Ancient America, Updated Edition (Gainesville, FL: University Press of Florida).
39. Loehr, J., Worley, K., Grapputo, A., Carey, J., Veitch, A., and Coltman, D.W. (2006). Evidence for cryptic glacial refugia from North American mountain sheep mitochondrial DNA. J. Evol. Biol. *19*, 419–430.
40. Fix, A.G. (2005). Rapid deployment of the five founding Amerind mtDNA haplogroups via coastal and riverine colonization. Am. J. Phys. Anthropol. *128*, 430–436.
41. Tankersley, K.B. (2004). The concept of Clovis and the peopling of North America. In The Settlement of the American Continents: A Multidisciplinary Approach to Human Biogeography, C.M. Barton, G.A. Clark, D.R. Yesner, and G.A. Pearson, eds. (Tucson, AZ: The University of Arizona Press), pp. 49–63.
42. LaViolette, P.A. (2005). Evidence for a global warming at the Termination I boundary and its possible cosmic dust cause. arXiv:physics/0503158v1. (http://arxiv.org/abs/physics/0503158).
43. Yu, Z., and Eicher, U. (2001). Three amphi-Atlantic century-scale cold events during the Bølling-Allerød warm period. Géographie Physique et Quaternaire *55*, 171–179.
44. Gibbons, A. (1996). The peopling of the Americas. Science *274*, 31–33.
45. Kaufman, T., and Golla, V. (2000). Language groupings in the New World: Their reliability and usability in cross-disciplinary studies. In America Past, America Present: Genes and Language in the Americas and Beyond, C. Renfrew, ed. (Cambridge: McDonald Institute for Archaeological Research), pp. 47–57.
46. Williams, R.C., Steinberg, A.G., Gershowitz, H., Bennett, P.H., Knowler, W.C., Pettitt, D.J., Butler, W., Baird, R., Dowda-Rea, L., Burch, T.A., et al. (1985). GM allotypes in Native Americans: Evidence for three distinct migrations across the Bering land bridge. Am. J. Phys. Anthropol. *66*, 1–19.
47. Bonatto, S.L., and Salzano, F.M. (1997). A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc. Natl. Acad. Sci. USA *94*, 1866–1871.
48. Schroeder, K.B., Schurr, T.G., Long, J.C., Rosenberg, N.A., Crawford, M.H., Tarskaia, L.A., Osipova, L.P., Zhadanov, S.I., and Smith, D.G. (2007). A private allele ubiquitous in the Americas. Biol. Lett. *3*, 218–223.
49. Kitchen, A., Miyamoto, M.M., and Mulligan, C.J. (2008). A three-stage colonization model for the peopling of the Americas. PLoS ONE *3*, e1596.
50. Mulligan, C.J., Kitchen, A., and Miyamoto, M.M. (2008). Updated three-stage model for the peopling of the Americas. PLoS ONE *3*, e3199.
51. Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., et al. (2004). The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am. J. Hum. Genet. *75*, 910–918.
52. Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. *22*, 160–174.
53. Yang, Z., and Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. Mol. Biol. Evol. *14*, 717–724.
54. Torroni, A., Achilli, A., Macaulay, V., Richards, M., and Bandelt, H.-J. (2006). Harvesting the fruit of the human mtDNA tree. Trends Genet. *22*, 339–345.
55. Olivieri, A., Achilli, A., Pala, M., Battaglia, V., Fornarino, S., Al-Zahery, N., Scozzari, R., Cruciani, F., Behar, D.M., Dugoujon, J.M., et al. (2006). The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. Science *314*, 1767–1770.