Ancient individuals from the North American Northwest Coast reveal 10,000 y of regional genetic continuity

John Lindo[a], Alessandro Achilli[b], Ugo A. Perego[b], David Archer[c], Cristina Valdiosera[d], Barbara Petzelt[e], Joycelynn Mitchell[e], Rosita Worl[f], E. James Dixon[g,h], Terence E. Fifield[i,j,1], Morten Rasmussen[k,l], Eske Willerslev[l], Jerome S. Cybulski[m,n,o], Brian M. Kemp[p,2], Michael DeGiorgio[q,r,2], and Ripan S. Malhi[s,2]

[a]Department of Human Genetics, University of Chicago, Chicago, IL 60637; [b]Department of Biology and Biotechnology "L. Spallanzani," University of Pavia, 27100 Pavia, Italy; [c]Prince Rupert, BC, Canada V8J 3Y3; [d]Department of Archaeology and History, La Trobe University, Melbourne, VIC 3086, Australia; [e]Metlakatla Treaty Office, Prince Rupert, BC, Canada V8J 3P6; [f]Rosita Worl Sealaska Heritage Institute, Juneau, AK 99801; [g]Maxwell Museum of Anthropology, University of New Mexico, Albuquerque, NM 87131; [h]Department of Anthropology, University of New Mexico, Albuquerque, NM 87131; [i]University of Alaska Southeast, Ketchikan, AK 99901; [j]Tongass National Forest, US Department of Agriculture Forest Service, Ketchikan, AK 99901; [k]Department of Genetics, School of Medicine, Stanford University, Stanford, CA 94305; [l]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, DK-1350 Copenhagen K, Denmark; [m]Canadian Museum of History, Gatineau, QC, Canada K1A 0M8; [n]University of Western Ontario, London, ON, Canada N6A 3K7; [o]Simon Fraser University, Burnaby, BC, Canada V5A 1S6; [p]Department of Anthropology, University of Oklahoma, Norman, OK 73019; [q]Department of Biology, Pennsylvania State University, University Park, PA 16802; [r]Institute for CyberScience, Pennsylvania State University, University Park, PA 16802; and [s]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL 61820

Edited by Anne C. Stone, Arizona State University, Tempe, AZ, and approved February 28, 2017 (received for review December 12, 2016)

Recent genomic studies of both ancient and modern indigenous people of the Americas have shed light on the demographic processes involved during the first peopling. The Pacific Northwest Coast proves an intriguing focus for these studies because of its association with coastal migration models and genetic ancestral patterns that are difficult to reconcile with modern DNA alone. Here, we report the low-coverage genome sequence of an ancient individual known as "Shuká Kää" ("Man Ahead of Us") recovered from the On Your Knees Cave (OYKC) in southeastern Alaska (archaeological site 49-PET-408). The human remains date to ∼10,300 calendar (cal) y B.P. We also analyze low-coverage genomes of three more recent individuals from the nearby coast of British Columbia dating from ∼6,075 to 1,750 cal y B.P. From the resulting time series of genetic data, we show that the Pacific Northwest Coast exhibits genetic continuity for at least the past 10,300 cal y B.P. We also infer that population structure existed in the late Pleistocene of North America with Shuká Kää on a different ancestral line compared with other North American individuals from the late Pleistocene or early Holocene (i.e., Anzick-1 and Kennewick Man). Despite regional shifts in mtDNA haplogroups, we conclude from individuals sampled through time that people of the northern Northwest Coast belong to an early genetic lineage that may stem from a late Pleistocene coastal migration into the Americas.

ancient DNA | paleogenomics | Native American | indigenous | peopling

The initial peopling of the Northwest Coast has received much attention because of its proximity to Beringia and associated implications for an initial coastal migration into the Americas (1–3). Genetic clues for the peopling of the Northwest Coast, however, may be obscured by later demographic events in the region. Studies based on mtDNA and Y-chromosomal markers suggest that populations in the region likely experienced admixture from other groups that entered the region after the initial peopling (4–6). Studies using genome-wide data (7–9) inferred ancient gene flow into North America likely stemming from subsequent movements after the initial settlement. However, because of the limited genomic data from populations in this geographic region, those studies leave questions regarding the degree of temporal genetic continuity of Northwest Coast populations.

In the Americas, the oldest thus far whole genome stems from Anzick-1, dating back to ∼12,600 calendar (cal) y B.P. and reportedly associated with Clovis technology (10, 11). Anzick-1 has proven to be surprising in a broader genetic sense, showing greater affinity with Central and South American groups than with Northern groups, despite the ancient burial existing in North America (but comparative indigenous populations from the United States are currently lacking). Shuká Kää, unearthed from On Your Knees Cave (Prince of Wales Island, AK), is not associated with Clovis culture but instead, is associated with a maritime tradition consistent with a coastal migration model and has been dated at ∼10,300 cal y B.P. (3). Shuká Kää exhibited the same mitochondrial haplogroup as Anzick-1 (12), suggesting a link in maternal lineage. Approximately 300 km southeast of the On Your Knees Cave archaeological site is Lucy Island off the coast of British Columbia, Canada. This ■■■ is the location of an individual, cataloged as 939, who died ∼6,075 cal y B.P. (13).

Significance

The peopling of the Americas has been examined on the continental level with the aid of SNP arrays, next generation sequencing, and advancements in ancient DNA, all of which have helped elucidate evolutionary histories. Regional paleogenomic studies, however, have received less attention and may reveal a more nuanced demographic history. We present genome-wide sequences of individuals from the northern Northwest Coast covering a timespan of ∼10,000 y and show that continental patterns of demography do not necessarily apply on the regional level. Compared with existing paleogenomic data, we show that geographically linked population samples from the Northwest Coast exhibit an early ancestral lineage and find that population structure existed among Native North American groups as early as the late Pleistocene.

Author contributions: J.L., B.M.K., M.D., and R.S.M. designed research; J.L., C.V., M.R., B.M.K., and R.S.M. performed research; J.L., E.W., B.M.K., M.D., and R.S.M. contributed new reagents/analytic tools; J.L., A.A., U.A.P., M.D., and R.S.M. analyzed data; D.A. and J.S.C. provided archaeological context; B.P., J.M., R.W., and T.E.F. organized community engagement; R.W. and E.J.D. interpreted results; and J.L., A.A., D.A., B.P., J.M., R.W., E.J.D., T.E.F., J.S.C., B.M.K., M.D., and R.S.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the NCBI Sequence Read Archive (accession no. PRJNA288803).

[1]Retired.

[2]To whom correspondence may be addressed. Email: bmkemp@ou.edu, mxd60@psu.edu, or malhi@illinois.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1620410114/-/DCSupplemental.

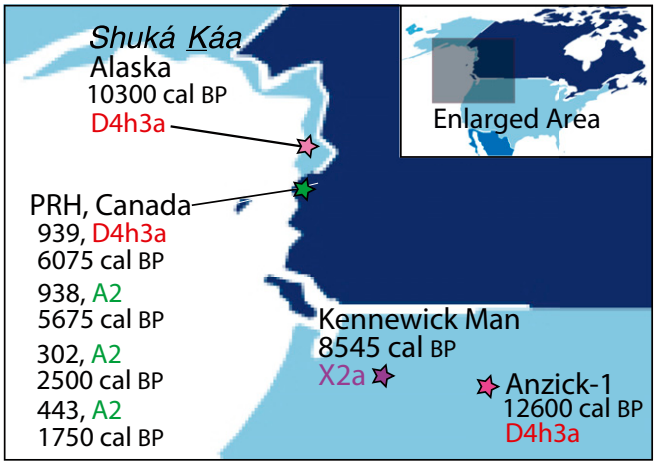


Fig. 1. Sampling locations of ancient samples and their associated mtDNA haplogroups.

Individual 939 displays genetic affinity to Pacific Northwest coast groups, such as the Coast Tsimshian (henceforth Tsimshian), that currently live in the same region, but it is difficult to reject 939 as ancestral to both North and South American groups (9).

The only other ancient genome from North America is the Ancient One (also known as Kennewick Man), unearthed in the US state of Washington and dating back to ~8,545 cal y B.P. (14). Kennewick Man also displays surprising results as an early Holocene individual who resided in the Pacific Northwest. His mtDNA belongs to the northern North America limited haplogroup X2a, but his nuclear genome shows affinities with Central and South American populations, similar to patterns observed for Anzick-1. However, a direct ancestry test shows the greatest link to living individuals from the Confederated Tribes of the Colville Reservation, a Native population living in the same geographic region as Kennewick Man (14). On a broader scale, numerous areas of the Americas exhibit patterns consistent with genetic continuity of peoples in the same geographic region over time (9).

To test hypotheses related to different demographic scenarios for the peopling of the Northwest Coast, we generated a low-coverage genome (including the complete mitochondrial genome) for Shuká Káa from Alaska (Fig. 1 and Table 1). In addition, we generated two ancient low-coverage genomes, 302 and 443, from Prince Rupert Harbor (PRH), British Columbia (Fig. 1 and Table 1) dating to 2,500 and 1,750 cal y B.P., respectively. Along with previously described genomes from the Americas, we test two hypotheses about the peopling of the Northwest Coast. First, we test whether the people of this geographic region show temporal genetic continuity dating back to at least 10,300 cal y B.P. Second, we test whether the ancestors of the Northwest Coast experienced additional gene flow in the mid-Holocene to further explore the previously observed shift in mtDNA haplogroups on the Northwest Coast (13).

Community Engagement

It is important to note that the interactions between scientists and indigenous community members associated with this study were and continue to be respectful. Shuká Káa is the indigenous name given to the ancient individual found in On Your Knees Cave on Prince of Wales Island in southeast Alaska. His remains were identified in 1996, the same year in which the Ancient One was unearthed from the banks of the Columbia River near Kennewick, WA. However, unlike the antagonistic relationships that were to develop over the handling of the Ancient One's remains, T.E.F., a US Forest Service archaeologist, and other researchers engaged with Tlingit- and Haida-speaking communities in Alaska and developed strong working relationships with community members (details are in Datasets S1–S4). With appropriate tribal engagement and discussions, our analyses were conducted on the last remaining tissue subsampled from Shuká Káa's molars for DNA analysis before his repatriation to the Tlingit and reburial in his ancestral land.

Farther south, J.S.C. and R.S.M. established a partnership with the Metlakatla and Lax Kw'alaams First Nations in 2007 to aid in the study of the population histories of those communities. They are located in the PRH region of British Columbia. As part of the active partnership, J.S.C. and R.S.M. visit the communities on a regular basis to develop research studies and discuss interpretations of results as well as manuscripts written for peer review publication. The First Nations agreed to allow destructive DNA methods of samples of ancestral individuals analyzed in this study.

Results

A Mitochondrial Genome Reassessment. The complete mitochondrial genome of the Shuká Káa individual belongs to haplogroup D4h3a and was compared with 52 modern (available in GenBank) and 2 ancient D4h3 mitochondrial genomes (939 and Anzick-1) (11, 13). The resulting tree (Fig. S1) clearly shows that Anzick-1 is ancestral to the entire D4h3a clade, whereas the ancient Northwest Coast mitochondrial genomes belong to two different subbranches known as D4h3a9 and D4h3a12, with the latter here defined and encompassing a modern sample of an individual currently living in Bolivia (*SI Text*).

Today, the haplogroup D4h3a is virtually absent in northern North America. To the contrary, the mitochondrial genomes of the more recent ancients from the Northwest Coast (443 and 302) are classified as A2 (Table 1), the most commonly reported mitochondrial haplogroup of native North America. Thus, based on the mtDNA data alone, it might be plausible that the native people of the northern Northwest Coast experienced a drastic change in their mtDNA gene pool in a rather short period, possibly because of additional gene flow in the mid-Holocene (mitochondrial genome change hypothesis). However, considering that the mtDNA haplogroup frequencies are likely to change

Table 1. Sequencing results and whole-genome coverage

Sample	Age ^{14}C B.P.	Age (cal y B.P.)	mtDNA haplogroup	Sex*	Sequencing strategy	Libraries sequenced	Mean read depth	Whole-genome coverage
Shuká Káa	9,200 ± 50	10,344 ± 83	D4h3a	M	Genome enrichment	8	2.85	0.059
443	1,820 ± 55	1,750 ± 70	A2d	M	Genome enrichment	2	4.01	0.560
302	2,440 ± 75	2,498 ± 142	A2p	F	Genome enrichment	2	4.74	0.701
939†	5,710 ± 40	6,075 ± 185	D4h3a7	F	N/A	N/A	5.15	0.343

F, female; M, male.

*Sex was determined from sequence reads using the method described by Skoglund et al. (27).

†Previously described by Raghavan et al. (9). Cal year conversion from radiocarbon dates was achieved with the method described by Fairbanks and coworkers.

radically because of drift in small populations over time (15), such a mitochondrial genome discontinuity might simply be the result of the limited number of complete mitochondrial genomes analyzed from the area, particularly from ancient individuals (16). Thus, because the mitochondrial genome can describe only part of the ancestral genetic history of the Northwest Coast, we extended the analyses to the entire genome to test alternative hypotheses.

Autosomal Genome Assessment. We used outgroup f_3 statistics to assess the shared ancestry among the ancient individuals and 169 worldwide populations (9). Outgroup f_3 statistics of a worldwide dataset show that all four ancient individuals (Shuká Káa, 939, 443, and 302) display greater affinity with Native American groups than with other worldwide populations (Fig. 2*A* and Fig. S3). Ranked outgroup f_3 statistics suggest that 939, 443, and 302 tend to share greatest affinity with Northwest Coast groups, whereas Shuká Káa ostensibly shows closer affinity to groups farther south (Fig. S4). However, because of the low coverage of the Shuká Káa sample, the relationship is not statistically significant.

To further elucidate the relationship among the ancient individuals of the Northwest Coast and their relationship to modern populations, we examined maximum likelihood trees created with TreeMix (17). C/T and G/A polymorphic sites were removed from the dataset to guard against the most common forms of postmortem DNA damage (18). We observe that 302 and 443 form a sister clade to the modern Tsimshian (masked for European ancestry) (Fig. S5 *A* and *B*, respectively). Individual 939 is an outgroup to both North and South Americans (Fig. S5*C*) as is Shuká Káa (Fig. 3*B*). However, adding a migration event introduces an edge connecting Europeans and Shuká Káa, which leads to Shuká Káa forming a clade with the Tsimshian and Athabascan (Fig. S6). The signal may represent Native American dual ancestry (19) or be a result of possible contamination (Table S1).

Principal components analysis reveals a tight clustering of 939, 443, and 302, which also overlap with modern North American indigenous populations (Fig. 2*C* and Fig. S7). Shuká Káa falls in close proximity but overlaps with both North and South American groups. The admixture clustering analysis shows a more complicated pattern,

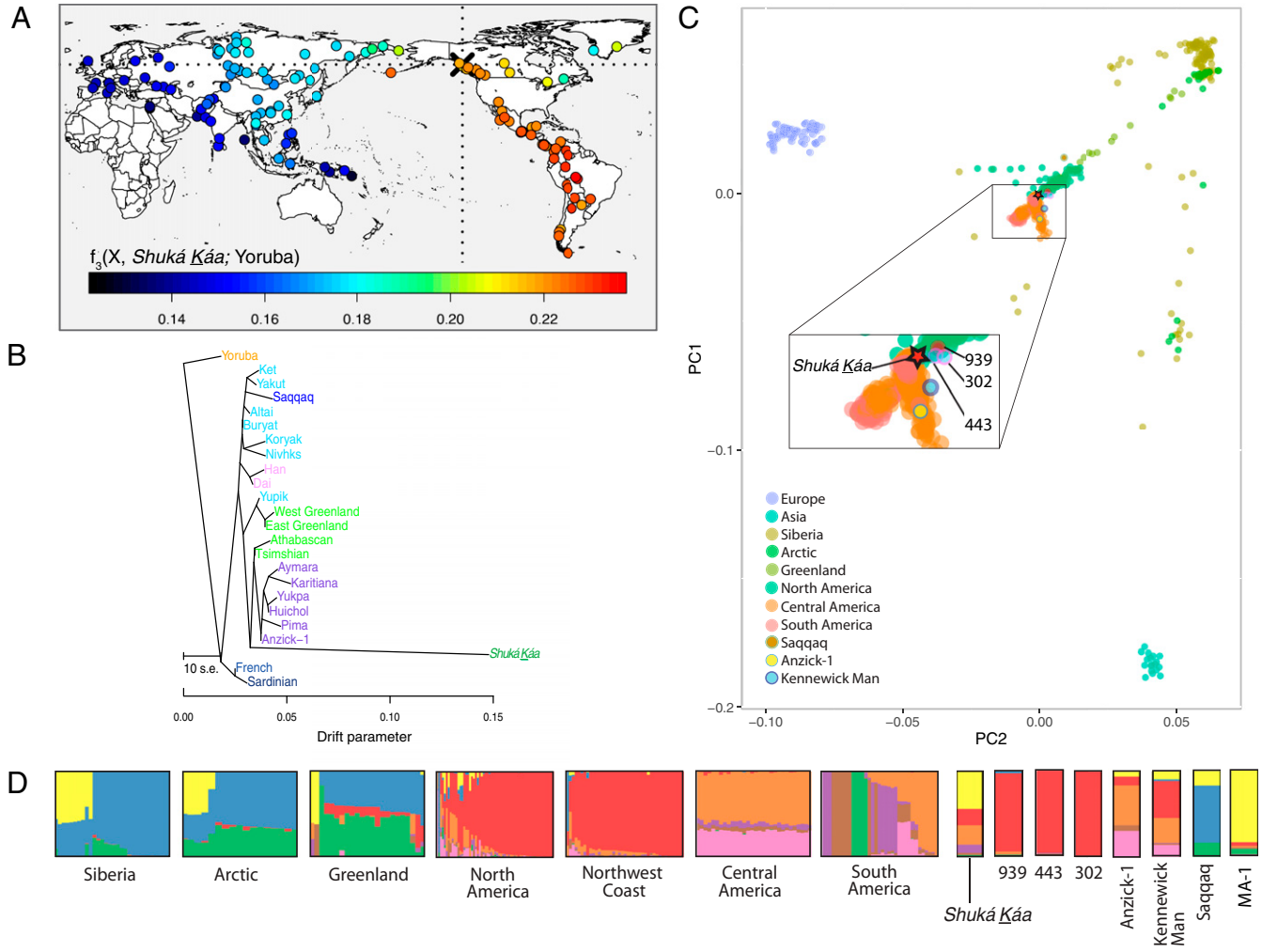


Fig. 2. Genetic affinity of Shuká Káa and the other Northwest Coast prehistoric humans to global and regional indigenous populations. (*A*) Heat map represents the outgroup f_3 statistics estimating the amount of shared genetic drift between Shuká Káa and each of 156 contemporary populations since their divergence with the African Yoruban population. (*B*) Maximum likelihood tree generated by TreeMix using whole-genome sequencing data from Raghavan et al. (9) and with the Tsimshian genome masked for European ancestry. (*C*) Principal components analysis projecting Shuká Káa, 939, 302, 443, Anzick-1 (11), Saqqaq (30), and Kennewick Man (14) onto a set of non-African populations from Raghavan et al. (9), with Native American populations masked for non-native ancestry. (*D*) Cluster analysis generated by ADMIXTURE for a set of indigenous populations from the Americas, Siberia, the Arctic, and Greenland and the Anzick-1, Kennewick, MA-1 (19), Saqqaq, Shuká Káa, 939, 302, and 443 samples. The number of displayed clusters is $K = 8$, which was found to have the best predictive accuracy given the lowest cross-validation index value.

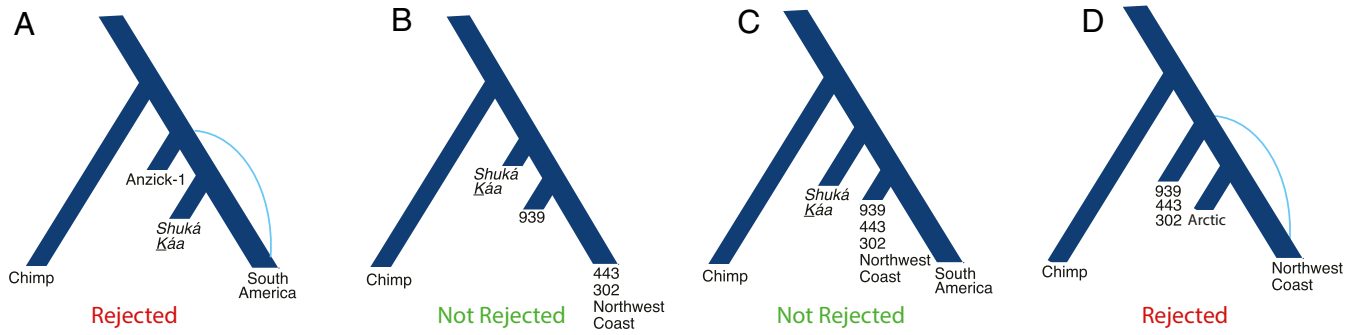


Fig. 3. Hypothetical scenarios for the regional peopling of the Northwest Coast. (*A*) Scenario tested by the *D* statistic where Anzick-1 is basal to both Shuká Káa and South America, which is rejected, indicating a closer affinity to South America. (*B*) Scenario tested by the *D* statistic where Shuká Káa is basal to 939 and both contemporary and ancient Northwest Coast individuals, which is not rejected. (*C*) Scenario tested by the *D* statistic where Shuká Káa is basal to ancient and modern Northwest Coast and South America, which is not rejected. (*D*) Scenario tested by the *D* statistic where 939, 443, and 302 are basal to the Arctic (Yup'ik and Inuit) and contemporary Northwest Coast populations, which is rejected, indicating a closer affinity to contemporary Northwest Coast populations. Supporting *D* statistics are listed in Table S2.

where Shuká Káa exhibits mainly components identified in North American and Siberian/Arctic individuals as well as smaller fractions found in South American populations (Fig. 2*D*). However, individuals 939, 302, and 443 all exhibit a major component found in North American populations.

To further test the hypothesis that people of the Northwest Coast exhibit a close genetic relationship with ancient individuals from the same region, we used the *D* statistic (20). Given the TreeMix admixture results between Shuká Káa and European populations, we performed a contamination correction to the *D* statistic as described in the work by Raghavan et al. (19) using observed *D* statistics with European populations (*SI Text*). Hypothetical scenarios based on the *D* statistics are depicted in Fig. 3. The *D* statistic does not support a scenario of genetic continuity between Anzick-1 and Shuká Káa with respect to South Americans (Fig. 3*A* and Table S2). The relationship of Shuká Káa, however, is more complex when examined with specific North American ancient and modern groups. Comparing Shuká Káa with 939 and the more recent ancient individuals from the Northwest Coast (443 and 302), we cannot reject an equally diverged relationship with respect to Shuká Káa (Fig. 3*B* and Table S2). However, we see the same relationship when the comparison is performed between both ancient and modern Northwest Coast individuals and individuals from South America, where Shuká Káa is basal to both groups (Fig. 3*C* and Table S2). Individual 939 displays a similar pattern with TreeMix and *D* statistics (Fig. S5*C* and Table S2), where the individual seems basal to both the Northwest Coast and South America. However, the admixture results show a predominately "North American" component, and contamination-corrected *D* statistics for 939 indicate a significant relationship toward the Northwest Coast (Table S2, tests 17 and 18).

The *D* statistic did not reveal a signal of gene flow between Arctic populations (Inuit and Yup'ik) and either the modern or ancient Northwest Coast populations compared with Shuká Káa (Table S2, tests 19–28). However, comparing the more recent ancient individuals, the tree was rejected with 939, 302, and 443, indicating greater affinity toward the Northwest Coast populations than the Arctic (Fig. 3*D* and Table S2).

Because certain *D* (Table S2, tests 8 and 9) and f_3 (Fig. S4*D*) statistics yielded nonsignificant results, we next wanted to examine whether the basal relationship that Shuká Káa exhibited to Northwest Coast and South American populations could be caused by its age relative to the split time of those groups. To address this hypothesis, we simulated genetic data with FastSimCoal2 (21), which allowed us to sample Shuká Káa 10,300 y in the past (*SI Text*). We considered one scenario (scenario 1), in which Shuká Káa is on the branch leading to the Northwest Coast, and another scenario (scenario 2), in which the sample was on a branch that diverged earlier than the split of the Northwest Coast and South American populations (Fig. S8*A*). Results for 1,000 simulated replicates under each scenario are plotted in Fig. S8*B*, indicating that only a small fraction of simulated replicates from scenario 1 could reject the null hypothesis that Shuká Káa is equally diverged to the Northwest Coast and South American lineages, although the simulations placed Shuká Káa on the Northwest Coast branch. However, the reason for this lack of power may be because of the amount of data (which we controlled to yield a similar number of *D* statistic informative sites as the empirical data). We, therefore, also considered a set of simulations where we increased the expected number of *D*-statistic informative sites by an order of magnitude. Results from these simulations (Fig. S8*C*) show that the clear majority of simulated replicates from scenario 1 could reject the null hypothesis that Shuká Káa is equally diverged to the Northwest Coast and South American lineages, with Shuká Káa having higher affinity to the Northwest Coast. Furthermore, results

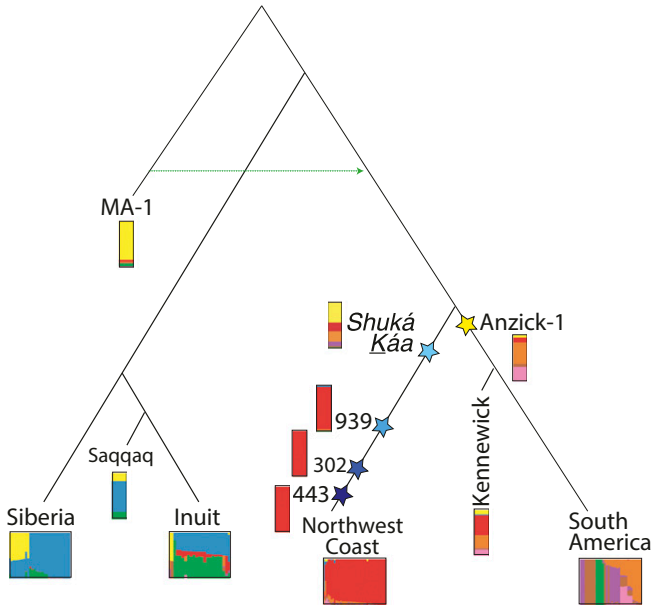


Fig. 4. Shuká Káa in relation to other Native American groups. Schematic showing Shuká Káa placed on the branch leading to North Americans, which is supported by simulation-based *D* statistics.

for scenario 2 indicate that the null hypothesis is generally not rejected as expected from the simulated scenario (Fig. S8 *B* and *C*).

We next considered models that examine a less direct relationship between Shuká Káa and the more recent ancient individuals in addition to the spread of the D4h3a mtDNA haplogroup across the Americas. Because Anzick-1 from Montana predates Shuká Káa (~12,600 vs. ~10,300 y B.P.) and shares the same mtDNA haplogroup, we performed *D* statistics to assess their relationship. The tree with Shuká Káa and 939 as sisters relative to Anzick-1 is not rejected (Fig. S9*A*). The trees with modern Northwest Coast (mtDNA haplogroup A2) as sister to either 443 (mtDNA haplogroup A2) or 302 (mtDNA haplogroup A2) relative to 939 (mtDNA haplogroup D4h3a) are also not rejected (Fig. S9*B*). These results are inconsistent with a change in the overall gene pool of the Northwest Coast after the early Holocene, although it led to the observation of very different mitochondrial haplogroups detected over time in the region.

We also explored the relationship of Shuká Káa to Kennewick Man (14), who was unearthed in Washington and dates to ~8,545 cal y B.P. Although Kennewick Man does not share the same mtDNA haplogroup with Shuká Káa, they coexisted within 1,700 y of each other (about 68 generations). In all tests with the *D* statistic, Shuká Káa displayed a basal relationship to both ancient and modern Northwest Coast populations with respect to Kennewick Man (Fig. S9*C* and Table S2).

Discussion

These data support a shared ancestry for the indigenous peoples of the Northwest Coast dating back to at least ~10,300 cal y B.P. The individual supporting this scenario is Shuká Káa, who belongs to mtDNA haplogroup D4h3a. Although both Anzick-1 and 939 also belong to this mtDNA haplogroup, later ancient and modern individuals of the Northwest Coast do not (13, 22). Despite belonging to different mtDNA haplogroups, Shuká Káa exhibits a close nuclear DNA relationship with 302 (~2,500 cal y B.P.) and 443 (~1,750 cal y B.P.), both of whom belong to mtDNA haplogroup A2, which is observed today at high frequencies among modern Northwest Coast populations (22). Thus, based on the limited mtDNA data alone, the hypothesis of a change in the genetic composition of the Northwest Coast appearing by the time of 302 and 443 might be plausible. However, our more extended genome-wide analysis does not support a change in the genetic composition of the Northwest Coast appearing by the time of 302 and 443, both of whom form a sister clade with the Tsimshian (Fig. S5). Instead, we observe a trend of genetic continuity through time, which is exemplified by individual 939 who displays affinities with both the more recent Northwest Coast ancient individuals and Shuká Káa. Individual 939 shares the same mtDNA haplogroup D4h3a as Shuká Káa while belonging to the predominant North American ancestry component observed in 443 and 302 (Fig. 2*D*). Furthermore, previously published complete mitochondrial genome data from individual 938 (13), who is similar in age and found on the same island site as 939, exhibit an A2 mtDNA haplogroup. These results indicate that the two haplogroups were already present in the ancestral population and are not the result of later gene flow into the area.

Also of interest is the evidence for the placement of Shuká Káa on different lineages than Anzick-1 and Kennewick Man (Fig. S9*D*). Despite their shared North American geographies and time periods as well as evidence that a single peopling event occurred in the late Pleistocene (23), our analyses show that population structure existed in the late Pleistocene of North America. These data are concordant with the late Pleistocene archaeological record documenting distinct contemporaneous archaeological cultures in different regions of North America (24).

We find that the placement of Shuká Káa (based on *D*-statistic simulations) is consistent with residing on a branch with modern day indigenous people from the Northwest Coast, whereas Anzick-1 fits separately on a branch that leads to the southern lineage, including populations from Central and South America (Fig. 4). Our result suggests that Shuká Káa was part of a population closely related to the ancestors that gave rise to the current populations of the northern Northwest Coast.

Table 2. Ancient overlapping sites used for autosomal-based analyses Q:33, 34

■■■	f_3	PCA and ADMIXTURE	TreeMix
Shuká Káa	245 (minimum); 1,906 (median); 1,914 (maximum)	2,327	1,044
302	7,656 (minimum); 56,400 (median); 56,500 (maximum)	59,205	29,592
443	9,881 (minimum); 75,000 (median); 75,170 (maximum)	79,233	35,620

On a broader scope, ADMIXTURE analysis revealed a component that dominates contemporary North American populations, which increases over time regarding the Northwest Coast (Fig. 2*D*). We also observe a small fraction of this component in the 24,000-y-old MA-1 individual from Siberia (19) (Fig. 2*D*). This result suggests that the North American component reflects an early ancestral lineage that has drifted through time to high frequency on the Northwest Coast.

It should be noted that, because of the highly degraded nature of the DNA extracted from Shuká Káa coupled with the lack of additional available sample, the autosomal analyses had less than the optimal number of overlapping sites (Table 2). This ■■■ is likely the cause for the lack of significance in both the ranked f_3 and *D* statistics when regarding Shuká Káa. In light of this ■■■, projections that consider missing data and contamination corrections were used to help mitigate the low number of sites available (*SI Text*). Despite these limitations, the analyses presented here, taken as a whole and compared with higher-depth samples, paint a sensible and convincing picture of the peopling of the Northwest Coast.

We conclude that the Northwest Coast exhibits an ancestral lineage that stems from the initial peopling of the region. The observed temporal change of mtDNA haplogroups in the area was probably caused by sampling, and no clear signs of gene flow into the area after the first settlement have been identified. Shuká Káa, who lived some 10,300 y ago, was part of an ancestral population that may have first populated the region but was distinct from ancestral populations related to Anzick-1. Although we cannot use our data to identify the specific ancestral location of this coastal lineage, the facts that Shuká Káa lived over 10,000 y ago and that Anzick-1 is related to the South American lineage suggest that the ancestral population likely existed north of the continental ice before the deglaciation of the interior corridor sometime after 12,600 y B.P. (23). This inference is supported by both the ADMIXTURE analysis showing a large component of ancestry of Shuká Káa shared with Siberian populations (the yellow component) and previous analysis by Verdu et al. (25) that suggests a closer relationship to Northeast Asian groups by Pacific Northwest populations than Central and South American populations.

The collaborative approach of this study shows an example of how indigenous community members and scientific researchers can work together in a positive and mutually beneficial way. In addition, the results presented here reveal the power of regional studies to elucidate demographic complexities, shedding light on the peopling of the Northwest Coast and the early ancestral lineages of North America.

Methods

DNA Extraction and Library Preparation. We used standard ancient DNA extraction methods following stringent guidelines to work with ancient human remains and conducted these in dedicated ancient DNA laboratories. DNA was extracted from three teeth belonging to individuals Shuká Kåa, 302, and 443. Furthermore, each DNA extract was converted into Illumina libraries (*SI Text*).

Genome Enrichment. We captured from eight libraries of Shuká Kåa DNA with the MyBaits whole-genome enrichment kit enhanced with protocol modifications recommended for the study of ancient DNA (*SI Text*). The eight captured libraries from each individual were pooled and sequenced (single end) on four lanes of an Illumina HiSeq 2000 run. Two captured libraries from 302 and 443 were pooled and run on two additional lanes.

Contamination Estimates. Contamination estimates using the mitochondrial genome were run on all three samples using the Scmutzi program described by Renaud et al. (26). The method jointly estimates present day human contamination and reconstructs the endogenous mitochondrial genome by considering both deamination patterns and fragment length distributions. Because Shuká Kåa and 443 were typed as male using the method described by Skoglund et al. (27), contamination based on the X chromosomes was also performed for these samples using the method described by Korneliussen et al. (28) and applied through the ANGSD software suite (www.popgen.dk/angsd/index.php/ANGSD). The Shuká Kåa sample did not have sufficient coverage along the X chromosome to perform the estimate.

f_3 And *D* Statistics. To test the genetic affinity of the ancient individuals with global populations, we performed f_3 outgroup statistics using the method outlined by Patterson et al. (20). We also examined the genetic affinities of each individual and various populations using ranked f_3 statistics, which are shown in Fig. S4. To examine the relationship between the ancient individuals (Shuká Kåa, 939, 443, 302, and Anzick-1), we performed an ABBA-BABA test or *D* statistic (29) using the definition used by ANGSD (28). The chimpanzee genome was used as an outgroup sequence. The tests also included the whole genome of a contemporary Tsimshian (9), which was masked for European ancestry, and all comparisons with Shuká Kåa used a correction for European contamination as did several comparisons with 939 (*SI Text*). To guard against potential bias from DNA damage in the ancient individuals, transitions were not considered during the tests.

ACKNOWLEDGMENTS. We thank Cara Monroe for assistance in the laboratory at Washington State University and Timothy H. Heaton for his paleontological work that resulted in the analysis of Shuká Kåa. This project was made possible through the active partnerships of the Lax Kw'alaams and Metlakatla First Nations and the Sealaska Heritage Institute. The field research was conducted in partnership with the Klawock Cooperative Association (Tribe) and the Craig Community Association (Tribe). The field work and preliminary analyses were supported by National Science Foundation Grants OPP-99-04258, 97-22858, BCS-1413551, and BCS-1518026; the USFS Tongass National Forest; and the National Geographic Society. Research for this analysis was funded by the Office of the Vice Chancellor of Research, University of Illinois at Urbana–Champaign; the Canadian Museum of History in Gatineau, Quebec, Canada; and Pennsylvania State University startup funds. Portions of this research were conducted with the Advanced CyberInfrastructure computational resources provided by The Institute for CyberScience at Pennsylvania State University.

1. Fladmark KR (1979) Routes: Alternate migration corridors for early man in North America. *Am Antiq* 44:55.
2. Erlandson JM, et al. (2007) The kelp highway hypothesis: Marine ecology, the coastal migration theory, and the peopling of the Americas. *J Island Coast Archaeol* 2:161–174.
3. Dixon EJ, Heaton TH, Lee CM (2014) Evidence of maritime adaptation and coastal migration from southeast Alaska. *Kennewick Man the Scientific Investigation of an Ancient American Skeleton*, ed Owsley DW (Texas A&M Univ Press).
4. Schurr TG, Sherry ST (2004) Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: Evolutionary and demographic evidence. *Am J Hum Biol* 16:420–439.
5. Dulik MC, et al.; Genographic Consortium (2012) Y-chromosome analysis reveals genetic divergence and new founding native lineages in Athapaskan- and Eskimoan-speaking populations. *Proc Natl Acad Sci USA* 109:8471–8476.
6. Achilli A, et al. (2013) Reconciling migration models to the Americas with the variation of North American native mitogenomes. *Proc Natl Acad Sci USA* 110:14308–14313.
7. Reich D, et al. (2012) Reconstructing Native American population history. *Nature* 488: 370–374.
8. Raghavan M, et al. (2014) The genetic prehistory of the New World Arctic. *Science* 345:1255832.
9. Raghavan M, et al. (2015) POPULATION GENETICS. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349:aab3884.
10. Lahren L, Bonnichsen R (1974) Bone foreshafts from a clovis burial in southwestern montana. *Science* 186:147–150.
11. Rasmussen M, et al. (2014) The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* 506:225–229.
12. Kemp BM, 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.
13. Cui Y, et al. (2013) Ancient DNA analysis of mid-holocene individuals from the Northwest Coast of North America reveals different evolutionary paths for mitogenomes. *PLoS One* 8:e66948.
14. Rasmussen M, et al. (2015) The ancestry and affiliations of Kennewick Man. *Nature* 523:455–458.
15. Helgason A, et al. (2009) Sequences from first settlers reveal rapid evolution in Icelandic mtDNA pool. *PLoS Genet* 5:e1000343.
16. Tackney JC, et al. (2015) Two contemporaneous mitogenomes from terminal Pleistocene burials in eastern Beringia. *Proc Natl Acad Sci USA* 112:13833–13838.
17. Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* 8:e1002967.
18. Briggs AW, et al. (2007) Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci USA* 104:14616–14621.
19. Raghavan M, et al. (2014) Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505:87–91.
20. Patterson N, et al. (2012) Ancient admixture in human history. *Genetics* 192:1065–1093.
21. Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLoS Genet* 9:e1003905.
22. Schurr TG, et al.; Genographic Consortium (2012) Clan, language, and migration history has shaped genetic diversity in Haida and Tlingit populations from Southeast Alaska. *Am J Phys Anthropol* 148:422–435.
23. Pedersen MW, et al. (2016) Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537:45–49.
24. Dixon EJ (2015) Late Pleistocene colonization of North America from northeast Asia: New insights from large-scale paleogeographic reconstructions. *Mobility and Ancient Society in Asia and the Americas* (Springer), pp 169–184.
25. Verdu P, et al. (2014) Patterns of admixture and population structure in native populations of Northwest North America. *PLoS Genet* 10:e1004530.
26. Renaud G, Slon V, Duggan AT, Kelso J (2015) Schmutzi: Estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol* 16:224.
27. Skoglund P, Stora J, Gotherstrom A, Jakobsson M (2013) Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J Archaeol Sci* 40:4477–4482.
28. Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics* 15:356.
29. Green RE, et al. (2010) A draft sequence of the Neandertal genome. *Science* 328:710–722.
30. Rasmussen M, et al. (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463:757–762.
31. Dixon EJ (1999) *Bones, Boats, and Bison* (Univ of New Mexico Press).
32. Heaton TH, Grady F (2003) The late Wisconsin vertebrate history of Prince of Wales Island, Southeast Alaska. *Ice Age Cave Faunas of North America*, ed Schubert BW.
33. MacDonald GF, Cybulski JS (2001) Introduction: The Prince Rupert Harbour Project. *Perspectives on Northern Northwest Coast Prehistory*, ed Cybulski JS (Canadian Museum of Civilization, Quebec, Gatineau, QC, Canada), pp 1–23.
34. Archer D (2001) Village patterns and the emergence of ranked society in the Prince Rupert area. *Perspectives on Northern Northwest Coast Prehistory*, ed Cybulski JS (Canadian Museum of Civilization, Quebec, Gatineau, QC, Canada), pp 203–222.
35. Barta JL, Monroe C, Kemp BM (2013) Further evaluation of the efficacy of contamination removal from bone surfaces. *Forensic Sci Int* 231:340–348.
36. Moss ML, Judd KG, Kemp BM (2014) Can salmonids (Oncorhynchus spp.) be identified to species using vertebral morphometrics? A test using ancient DNA from Coffman Cove, Alaska. *J Archaeol Sci* 41:879–889.
37. Lindo J, et al. (2016) A time transect of exomes from a Native American population before and after European contact. *Nat Commun* 7:13175.
38. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010:pdb.prot5448.
39. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
40. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.
41. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.
42. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29:1682–1684.
43. Rasmussen M, et al. (2011) An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334:94–98.
44. Prüfer K, et al. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505:43–49.
45. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2:e190.
46. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19:1655–1664.
47. Sikora M, et al. (2014) Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe. *PLoS Genet* 10:e1004353.
48. Andrews RM, et al. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 1

Q: 1_Please contact PNAS_Specialist.djs@sheridan.com if you have questions about the editorial changes, this list of queries, or the figures in your article. Please include your manuscript number in the subject line of all email correspondence; your manuscript number is 201620410.

Q: 2_Please (i) review the author affiliation and footnote symbols carefully, (ii) check the order of the author names, and (iii) check the spelling of all author names, initials, and affiliations. Please check with your coauthors about how they want their names and affiliations to appear. To confirm that the author and affiliation lines are correct, add the comment "OK" next to the author line. This is your final opportunity to correct any errors prior to publication. Misspelled names or missing initials will affect an author's searchability. Once a manuscript publishes online, any corrections (if approved) will require publishing an erratum; there is a processing fee for approved erratum.

Q: 3_Please review and confirm your approval of the short title: Genetic continuity on the Northwest Coast. If you wish to make further changes, please adhere to the 50-character limit. (NOTE: The short title is used only for the mobile app and the RSS feed.)

Q: 4_Please review the information in the author contribution footnote carefully. Please make sure that the information is correct and that the correct author initials are listed. Note that the order of author initials matches the order of the author line per journal style. You may add contributions to the list in the footnote; however, funding should not be an author's only contribution to the work.

Q: 5_Your article will appear in the following sections of the journal: Social Sciences (Anthropology) and Biological Sciences (Anthropology). Please confirm that this is correct.

Q: 6_You have chosen not to pay an additional $1450 (or $1100 if your institution has a site license) for the PNAS open access option. Please confirm this is correct and note your approval in the margin.

Q: 7_Please verify that all callouts for supporting information (SI) in text are correct. Note, however, that the hyperlinks for SI callouts will not work until the article is published online. In addition, SI that is not composed in the main SI PDF (appendices, datasets, movies, and "Other Supporting Information Files") have not been changed from your originally submitted file and so are not included in this set of proofs. The proofs for any composed portion of your SI are included in this proof as subsequent pages following the last page of the main text. If you did not receive the proofs for your SI, please contact **PNAS_Specialist.djs@sheridan.com**.

Q: 8_Please check the order of your keywords and approve or reorder them as necessary. Note that PNAS allows up to five keywords; please do not add new keywords unless you wish to replace others.

Q: 9_Per PNAS style, certain compound terms are hyphenated when used as adjectives and unhyphenated when used as nouns. This style has been applied consistently throughout where (and if) applicable.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 2

Q: 10_Please confirm whether all units/divisions/departments/laboratories/sections have been included in the affiliations line for each footnote symbol or add if missing. PNAS requires smallest institutional unit(s) to be listed for each author in each affiliation. Please confirm or correct affiliation addresses.

Q: 11_Please confirm that locations and ZIP/postal codes inserted in the affiliations are correct.

Q: 12_Please provide department or section and institution in affiliation c.

Q: 13_Please provide department or section in affiliations e, f, i, and m-o.

Q: 14_Please note that the author contribution "C.V., M.R., and E.W. edited manuscript" has been deleted per PNAS style.

Q: 15_Please indicate whether the sequences have been deposited in NCBI Sequence Read Archive or another publicly accessible database before your page proofs are returned. It is PNAS policy that the data be deposited BEFORE the paper can be published.

Q: 16_PNAS mandates clarification of ambiguous pronoun antecedents. That is, a pronoun, such as "this" or "which," must refer to a specific word or phrase, not the general idea expressed in the previous clause or sentence. Please provide an appropriate noun after "■■■."

Q: 17_SOM changed to Datasets S1–S4 in section Community Engagement. Please verify.

Q: 18_PNAS does not allow claims of priority or primacy; therefore, the phrase "for the first time" has been deleted.

Q: 19_The URL popgen.dk/angsd has been updated to www.popgen.dk/angsd/index.php/ANGSD. Please verify.

Q: 20_PNAS prefers that any and all numbered composed pieces of Supporting Information be cited in the article's main text so that they may be integrated into the full text (HTML) presentation of the article online. The following components are not cited in your article: Figs. S2 and S10. Please cite these components in numerical order in the main text if possible.

Q: 21_In the Acknowledgments, please check the names of persons thanked for assistance/contributions, names of granting agencies, grant numbers, and initials of grant recipients for accuracy.

Q: 22_PNAS articles should be accessible to a broad scientific audience. As such, please spell out "USFS."

Q: 23_Because the references are not in PubMed, please verify information in refs. 1, 27, and 36.

Q: 24_Is this a one-page article, a letter, or a published abstract? If it is an unpublished abstract, then it must be removed from the reference list and placed in a footnote to the text. Please indicate which type of article it is, and it will be fixed accordingly in the proofs. Please verify in ref. 1.

Q: 25_Because the reference is not in PubMed, please verify journal name and information in ref. 2.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 3

Q: 26_Please provide page range and publisher's city, state, and country in ref. 3.

Q: 27_Please provide editors and publisher's city, state, and country in ref. 24.

Q: 28_Please provide publisher's city, state, and country in ref. 31.

Q: 29_Please provide page range and publisher's name, city, state, and country in ref. 32.

Q: 30_Please define "N/A" in Table 1 legend.

Q: 31_Reference number 30 changed to 27 in Table 1 legend to match Skoglund et al. Please verify.

Q: 32_Please provide correct reference number for Fairbanks and coworkers in Table 1.

Q: 33_Please provide column 1 heading in Table 2.

Q: 34_PNAS articles should be accessible to a broad scientific audience. As such, please spell out "PCA."

Supporting Information

Lindo et al. 10.1073/pnas.1620410114

SI Text

Community Engagement for Analysis of Shuká Kåa. Because this study provides a positive example of how scientific researchers and indigenous community members can engage with each other, we are providing details of the communications between the two that can be used as an example.

When human skeletal remains were identified at On Your Knees Cave on July 4, 1996 during permitted paleontological investigations on the Thorne Bay Ranger District of the Tongass National Forest, the researchers halted their investigations and notified the local Forest Service authorities. T.E.F., then a US Forest Service archaeologist, after inspecting the remains and the site, notified tribes that might consider this area traditional territory. During the week after the finding of ancestral remains and discussion with potential researchers and tribal culture specialists, T.E.F., representing the Tongass National Forest, engaged with representatives of three tribes: the Klawock Cooperative Association (KCA), the Craig Community Association (CCA), and the Organized Village of Kake (OVK).

Members of the KCA, the CCA, and the Sealaska Heritage Institute (SHI) discussed whether to allow scientific research of the ancestral remains. The OVK opted to defer to the KCA and the CCA on this issue. Other Prince of Wales Island tribes were notified of the ancestral remains but chose not to participate in the decision process. In July of 1996, the KCA and the CCA drafted and adopted resolutions giving approval to radiocarbon date and analyze the artifacts and human remains identified at On Your Knees Cave (Datasets S1–S4). In 2004, B.M.K. engaged with the KCA, the CCA, and the SHI to discuss the potential for ancient DNA analysis. After extensive discussions, although not unanimous, the Tlingit supported the DNA study because of their belief that indigenous ancestors lived in southeast Alaska "since time immemorial" and also, because of the Tlingit cultural value of Haa Shuká that the current generation is tied to both ancestral and future generations and that the ancient individual offered himself to learn about the past. The SHI adopted a resolution to allow for destructive analysis of a molar from the ancient individual for DNA analysis (Datasets S1–S4). Through acumen of scientific and legal arenas, Tlingit community members developed strong working relationships with scientists and government managers that resulted in positive and mutually beneficial relationships. The scientists and tribes created a partnership where tribal representatives were involved in nearly all aspects of the scientific research project.

In 2007, Kemp et al. (12) published mtDNA and Y-chromosome analysis of the ancient individual, and the ancestral remains were returned to the indigenous communities. The name Shuká Kåa is the Tlingit name given to the ancient person by a Council of Elders in 2008 shortly before the remains were reburied. A marble headstone donated by Sealaska Corporation with Shuká Kåa's presumed date of birth and death marked the grave where he was buried on September 25, 2008. The KCA and Craig Tribal Association (CTA; formerly known as the CCA), with the support of the SHI, Sealaska Corporation, and the Tongass National Forest, celebrated Shuká Kåa's life on September 26 and 27, 2008. The Shuká Kåa Honor Ceremony was attended by over 600 guests, including the research scientists and tribal, local, state, and federal dignitaries.

In 2014, B.M.K. engaged with R.W. on using the last remains of the extracted molar for genomic analysis in this study. The results of this study were shared and discussed with members of the SHI, the KCA, and the CTA. Since the time of the burial, the spelling of the name of the ancient individual was corrected from Shuká Kåa to Shuká Kåa.

Archaeological Context.

Shuká Kåa. The skeletal remains of Shuká Kåa are dated to ~9,200 ± 50 ^{14}C y B.P. (12, 31) and were unearthed from On Your Knees Cave (Site 49-PET-408) located on northern Prince of Wales Island, AK. The spatial distribution of the remains within the cave suggests that the individual was not intentionally buried but instead, was deposited or redeposited in the cave, possibly as a result of accidental death and postdepositional taphonomic agents (3). The paleontological record of the southeast Alaskan coast suggests that large areas were refugia during the last glacial maximum, with continual use starting at about 17,200 y B.P. (32). Humans may have made use of the cave as early as 12,000 y B.P.

Isotope analysis of the bone collagen revealed a long-term diet of marine foods, with little sustenance derived from land sources (3). The stone tools occurring in the same stratigraphic level as the human remains but not directly associated with the individual were manufactured with materials originating from nearby islands and at least one mainland source. This evidence suggests that the population associated with Shuká Kåa comprised maritime-adapted coastal navigators who participated in established trade networks between adjacent islands and the mainland (3).

Lucy Island and PRH ancient individuals. The archaeological context of the ancient individuals recovered from the PRH region, ranging in age from ~6,500 to 1,750 cal y B.P., shows cultural continuity between the ancient history of the region and the Coast Tsimshian communities described in historical times. Two of the ancient individuals (302 and 443) derive from the PRH excavation sites (33), whereas individual 939 (previously described in refs. 9 and 13) stems from the Lucy Islands site 19 km west of PRH. Seasonal occupations in the region date back to almost 10,000 y B.P. The Lucy Islands also have some of the earliest evidence for housing in the form of a rectangular depression dating to about 5,400 y B.P. The human remains recovered from the site, including individual 939, were found adjacent to the house depression and are the same age or slightly older. The house depression is similar to more recent rectangular depressions found in the PRH region dating from 2,500 to 1,500 y B.P. (34). They likely were precursors of the ethnohistorical house structures of the Coast Tsimshian.

DNA Extraction and Library Preparation. DNA extractions were completed in ancient DNA laboratories at the University of Illinois Urbana–Champaign, Washington State University, and the University of Copenhagen. Ancient DNA extractions and PCR amplification setups were completed in the ancient DNA laboratory facility at the University of Illinois. The ancient DNA laboratory is a positively pressured clean room with hepafiltered air. The clean room contains an anteroom, and air flows from the ancient DNA laboratory to the anteroom to the hallway. Personnel working in the ancient DNA laboratory wear disposable hairnets, facemasks, laboratory coveralls, and booties. All equipment, reagents, and consumables are dedicated for use in the ancient DNA laboratory. The ancient DNA laboratory is routinely cleaned with bleach, and all containers are wiped with DNA Away before placed in the ancient DNA laboratory. Personnel are restricted from entering the ancient DNA after being in a contemporary DNA laboratory. A database containing mitochondrial control region sequences is maintained of all personnel working in

the ancient laboratory. Negative controls were used with every DNA extraction and PCR setup to detect contamination originating from reagents or inadvertently caused during the procedures. Also, a series of negative controls is routinely performed in the ancient DNA laboratory to assess for general contamination of the laboratory.

Teeth were used from each ancient individual for the extraction. Each tooth was soaked in 6% sodium hypochlorite for 3 min, rinsed three times with UV-irradiated molecular-grade water, and dried in a UV Crosslinker for 10 min to remove surface contamination. Approximately 0.20 g tooth powder was incubated in 4 mL demineralization/lysis buffer (0.5 M EDTA, 33.3 mg/mL Proteinase K, 10% *N*-lauryl sarcosine) for 24 h at 37 °C. The digested sample was then concentrated to ∼250 μL using Amicon centrifugal filter units. After concentration, the digest was run through silica columns using the MinElute Qiagen PCR Purification Kit (Qiagen, Hilden, Germany) and eluted in 60 μL DNA extract.

Washington State University laboratory. The Washington State University laboratory was the former location of the ancient DNA laboratory of B.M.K. Subsamples of tissue weighing ∼54–121 mg were carefully removed from subsections of Shuká Kåa's first and second molars reserved for genetic analyses before his repatriation to the Tlingit and reburial. These samples of molars were decontaminated by submersion in 6% (wt/vol) sodium hypochlorite for 4 min (35). The sodium hypochlorite was poured off, and the tissue was rinsed twice by submersion in DNA-free water. DNA was extracted from the tissue following the modified protocol of Kemp et al. (12) described by Moss et al. (36). Because this individual was previously determined to belong to mitochondrial haplogroup D, each extract was screened for the definitive marker of the haplogroup (i.e., AluI site loss at nucleotide position 5,176) following the work by Kemp et al. (12). After an extract was confirmed to contain mtDNA belonging to haplogroup D and be free of contamination (i.e., exhibiting no visual trace of nonhaplogroup D mtDNA on the gels after enzymatic digestion of the amplicons with AluI), it was sent to the University of Illinois for additional processing.

Centre for GeoGenetics, Copenhagen, Denmark. A separate Shuká Kåa sample was extracted and underwent shotgun sequencing at the Centre for GeoGenetics in Copenhagen, Denmark using the protocols described in the work by Rasmussen et al. (11). The sample was not subjected to enrichment or the uracil-specific excision reagent specified below to permit detection of postmortem DNA damage patterns (Fig. S1).

Library build. Libraries were created using the New England Biolabs Ultra Kit (E7370S) following the manufacturer's protocol with the modifications performed in the work by Lindo et al. (37). DNA purification was conducted with the MinElute Reaction Cleanup Kit (Qiagen). The libraries were constructed with the NEB USER Enzyme, which excises uracil bases that may be a consequence of deamination. Library amplification was done in two steps. The first round of amplification used the kit's reagents and protocol with 12 cycles of 10 s at 98 °C, 30 s at 65 °C, and 30 s at 72 °C. For the second round, we achieved a sufficient DNA concentration for the whole-genome enrichment (∼500 ng), without excessive amplification, by creating four PCR reactions from the initial amplified product and then pooling them before applying the Qiagen MinElute PCR Cleanup Kit. For the second PCR, we created a 50-μL reaction using 0.2 μM primers P5 (5′-AATGATACGGCGACCACCGA) and P7 (5′-CAAGCAGAAGACGGCATACGA) (38), 5 μL initial PCR, 25 μL Phusion High-Fidelity PCR Master Mix with HF Buffer (New England Biolabs), 3% DMSO (New England Biolabs), and 0.2 mg/mL BSA (New England Biolabs). PCR conditions were as follows: 4 min at 98 °C and 10 cycles of 10 s at 98 °C, 30 s at 62 °C, and 30 s at 72 °C with a final extension at 72 °C for 10 min. Library fragment sizes were confirmed via a BioAnalyzer High Sensitivity assay to be above 130 bp.

Genome Enrichment and Illumina Sequencing. For each of the PRH ancient samples (302 and 443), two whole-genome enrichments were performed using the MyBaits Whole Genome Capture Kit (Mycroarray). These individual captured libraries were pooled and run on two lanes (single end; 100 bp) of the Illumina HiSeq 2000 at the High-Throughput Sequencing Division of the W. M. Keck Biotechnology Center at the University of Illinois Urbana–Champaign. Shuká Kåa underwent eight whole-genome enrichments and was sequenced on four lanes of the Illumina HiSeq 2000. For the capture, the manufacturer's protocol was used with the following modifications: the hybridization temperature was decreased to 50 °C for 24 h, the Qiagen MinElute PCR Cleanup Kit was used instead of beads, a final heat elution from the Streptavidin beads was performed with molecular biology-grade water at 90 °C for 2 min, and the postcapture amplification involved 12 cycles. Because of the degraded nature and low endogenous DNA content (∼0.22%) of the Shuká Kåa sample, multiple rounds of PCR were needed after both library construction and enrichment.

Genome Mapping and Shuká Kåa Special Processing. Raw data from the Illumina HiSeq 2000 platform was base called with CASAVA 1.8.2. Sequences were demultiplexed with a requirement for a full match of six nucleotide indexes that were used for library preparation. Illumina adapter sequences were trimmed using Trimmomatic-0.32 (39) with a minimum length of 25 and by removing leading and trailing quality or N bases below a quality score of three. Five bases were additionally trimmed at each end of the reads to minimize the effects of postmortem DNA damage. Trimmed reads were aligned to the human reference genome hg19 using Bowtie2-2.2.5 (40) with a local realignment option and a seed length set to 1,000. SAMtools-1.2 (41) was used to sort and remove duplicate reads based on mapping positions. Because of the low average read depth of the PRH and Shuká Kåa individuals, genotypes were not called. The resulting alignment files were further filtered for downstream statistical analysis using a minimum mapping quality of 30 and a minimum base quality of 30.

The Shuká Kåa sequences exhibited extensive damage in both tranversions and transitions at both the 3′ and 5′ ends, which was likely caused by the heavy amplification needed to proceed with enrichment and sequencing. The sample, therefore, was trimmed for reference (hg19) mismatches that exceeded a ratio of 0.1 using BamUtil's filter option (genome.sph.umich.edu/wiki/BamUtil:_filter) before proceeding with all downstream analyses.

DNA Damage Patterns. DNA damage was assessed by examining the C → T substitution rates toward sequencing starts and complementary increase in G → A rates toward read ends using MapDamage 2.0 (42). A specific pattern of DNA damage has been identified in other ancient DNA studies (30, 43). These studies show a pattern of increased DNA damage at the beginning and end of degraded DNA fragments. MapDamage was run on the shotgun sequencing runs without trimming. The results show signatures of DNA, which suggest that the PRH ancient and Shuká Kåa sequences originate from ancient DNA templates and do not originate from modern contaminants (Fig. S1).

TreeMix Using Whole-Genome Sequencing Data. TreeMix (17) was applied to the dataset to generate maximum likelihood (ML) trees and admixture graphs from allele frequency data. For each ancient sample, reads with mapping quality below 30 and nucleotides with quality below 20 were not considered. A single allele was sampled from each ancient sample and intersected with the combined whole-genome sequence dataset used in the work by Raghavan et al. (9). Sites in the Tsimshian individual were

masked for European ancestry using the exact masking of Raghavan et al. (9). Sites used for each individual were Shuká Kúa, 1,044; 939, 20,270; 302, 29,592; and 443, 35,620.

The Yoruban population was used to root the tree (with the –root option). We accounted for linkage disequilibrium by grouping M adjacent sites (with the –k option), and we chose M, such that a dataset with L sites will have approximately $L/M \approx 20{,}000$ independent sites. At the end of the analysis (i.e., number of migrations), we performed a global rearrangement (with the –global option). We considered admixture scenarios with $m = 0$ to $m = 2$ migration events. Each migration scenario was run with 100 replicates, and the replicate with the highest likelihood was chosen to represent the ML tree ($m = 0$) or graph ($m > 0$) for the given migration scenario.

Fig. 2*B* and Fig. S5 display the results for the ML tree with no admixture ($m = 0$) events. Here, both 302 and 443 fall as sisters to the contemporary Tsimshian (Fig. S5 *A* and *B*). Both 939 (Fig. S5*C*) and Shuká Kúa (Fig. 2*B*) form an outgroup to North and South American groups. However, with one migration event, gene flow appears between Shuká Kúa and European populations (Fig. S6). This result may be a product of sample contamination. Adding additional migration events with 302 and 443 does not change the positions of 302 and 443 relative to the Tsimshian.

Outgroup f_3 Statistics. To test the genetic affinity of the ancient individuals with global populations, we preformed outgroup f_3 statistics (19) using the method outlined by Patterson et al. (20). For each ancient sample, reads with mapping quality below 30 and nucleotides with quality below 20 were not considered. A single allele was sampled from each ancient sample and intersected with the SNP chip dataset in the work by Raghavan et al. (9) that included 2,081 contemporary individuals sampled from across the world. Sites in Native Americans were masked for European ancestry using the exact masking used by Raghavan et al. (9). We computed outgroup f_3 statistics of the form f_3(A, M; Yoruba), where A is one of four ancient individuals, M is a contemporary non-African population from the SNP chip dataset of Raghavan et al. (9), and Yoruba is an African population representing the outgroup. Each individual had the following overlap with the modern SNP data: Shuká Kúa, 245 (minimum), 1,906 (median), and 1,914 (maximum); 939, 5,720 (minimum), 43,310 (median), and 43,360 (maximum); 302, 7,656 (minimum), 56,400 (median), and 56,500 (maximum); and 443, 9,881 (minimum), 75,000 (median), and 75,170 (maximum).

The f_3(A, M; Yoruba) statistic measures the amount of shared genetic drift between A and M since their divergence with Yoruba. Lower values indicate less shared genetic drift, and higher values indicate more shared genetic drift. Contemporary non-African populations M for which an ancient sample A has the highest f_3(A, M; Yoruba) represent populations for which the ancient sample has highest genetic affinity.

The genetic affinities of four ancient individuals with different combinations of non-African populations are depicted in Fig. 2*A* and Fig. S3. We also examined the genetic affinities of each individual and various populations by ranking outgroup f_3 statistics, which are shown in Fig. S4. The two least ancient samples from PRH (302 and 443) display highest affinity to populations from the Northwest Coast (Fig. S4 *A* and *B*), reflecting the findings from TreeMix, in which 302 and 443 fall with the Athabascan and Tsimshian clade (Fig. S5 *A* and *B*). The other ancient sample from PRH (939) shows highest affinities to populations from southwestern American and the Northwest Coast (Fig. S4*C*), which is consistent with the TreeMix results, in which 939 falls ancestral to all of the Native American samples (Fig. S5*C*). The Shuká Kúa sample from Alaska shows highest affinity to individuals from southwestern, Central, and South America (Fig. S4*D*), although the error bars are wide because of the samples' low coverage.

Sequence Data-Based *D* Statistic. To examine the relationship between the ancient individuals (Shuká Kúa, 939, 443, and 302) and other populations, we performed ABBA-BABA tests or D statistic (29) using the definition used by ANGSD (28). The statistic tests the tree [((*P*1, *P*2), *P*3), *O*], where O is an outgroup to populations *P*1, *P*2, and *P*3. If the null hypothesis, where $D = 0$, cannot be rejected, then there is no evidence for gene flow either between *P*1 and *P*3 or between *P*2 and *P*3. If there is a significant deviation from zero, then we can reject the hypothesis of the tree [((*P*1, *P*2), *P*3), *O*], and the data are consistent with gene flow between *P*1 and *P*3 or between *P*2 and *P*3 depending on the sign of the D. A Z score was used to determine the significance of the test, whereby an absolute value greater than three ($|Z| > 3$) was used as a critical value (29). The chimpanzee genome was used for the outgroup. The tests also included the masked whole genome of a contemporary Tsimshian (T60) (9), which was masked for European ancestry, and the whole genome of a Karitiana individual (HGDP00998) (44). We also used the genomes of Athabascan_1, Esk_17, and Greenlander_1 from ref. 9 to guard against DNA damage from the ancient individuals; transitions were not considered during the tests.

***D*-Statistic Contamination Correction.** Given the TreeMix result for Shuká Kúa with one migration event (Fig. S6), where gene flow is approximated between the ancient individual and a European population, we performed a contamination correction before proceeding with the D statistic. We also performed the correction with 939 with scenarios that approached significance. We used the sequence of a Great Britain (GBR) individual from the 1,000 Genomes Project (HG00118) as a representative of the population presumed to be the source of the contamination. We then followed the method performed by Raghavan et al. (19). Let $D_{Shuká\ Kúa}$ denote a D statistic involving the potentially contaminated Shuká Kúa sample, and let D_{GBR} denote the identical D statistic substituting GBR for the Shuká Kúa sample. Furthermore, let $D^*_{Shuká\ Kúa}$ denote the D involving the Shuká Kúa sample if it was not contaminated. Assuming a contamination rate of c, we can write the D statistic for the potentially contaminated Shuká Kúa sample as $D_{Shuká\ Kaa} = (1-c)D^*_{Shuká\ Kúa} + cD_{GBR}$. Rearranging, we can estimate the D statistic corrected for contamination as

$$D^*_{Shuká\ Kúa} = \frac{D_{Shuká\ Kúa} - cD_{GBR}}{1-c}.$$

For the contamination rates, we used the mtDNA-based estimate of 2.16% ($c = 0.0216$). We observed that the tests that involved Northwest Coast populations exhibited significant changes after the correction was performed, suggesting that the contamination rate was affecting our results that were near significance (Table S2).

Simulation-Based *D* Statistics. Because D statistics did not reject scenarios in which Shuká Kúa was basal to the split of the Northwest Coast and South America, we wanted to test whether we had power to reject the null hypothesis that Shuká Kúa had equal affinity to the Northwest Coast and South America. We considered one scenario (scenario 1), in which Shuká Kúa was on the branch leading to the Northwest Coast populations, and another scenario (scenario 2), in which Shuká Kúa was on a separate branch that diverged from the ancestor of the Northwest Coast and South American lineages (Fig. S8*A*).

Simulations were performed using FastSimCoal2 (21). We simulated four sequences: one from chimpanzee, one from Shuká Kúa, one from the Northwest Coast lineage, and one from the South American lineage. The sequence from Shuká Kúa was

sampled at 10.3 kya. We assumed a divergence time of 5 My of chimp with humans and a split time of 15 kya of the Northwest Coast and South American lineages. In addition, for scenario 2, the separate lineage containing Shuká Kåa diverged from the Northwest Coast and South American lineages 18 kya. We assumed a constant effective size of 10,000 diploid individuals across the tree, per-site per-generation recombination and mutation rates of 2.5×10^{-8}, and a generation time of 25 y. We performed 1,000 simulated replicates under each scenario, where each replicate was based on 200 independent 100-kb genomic regions. These parameters yielded a similar number of *D*-statistic informative sites as we observed in our empirical dataset with Shuká Kåa. For each set of 200 genomic regions (a simulated replicate), we computed the *D* statistics as well as a *Z* score based on a block bootstrap.

Principal Component Analysis. Principal component analysis was performed on a subset of individuals from the dataset of Raghavan et al. (9), which excluded African populations, using EIGENSOFT (45) and included 199,285 sites. Each ancient individual had the following overlap with the modern dataset: Shuká Kåa (2,327); 302 (59,205); 443 (79,233); and 939 (43,441). Native American populations in the dataset were masked for nonnative ancestry. Shuká Kåa, Anzick-1, Kennewick Man, 939, 443, 302, and Saqqaq were projected onto the components inferred from these sets of contemporary individuals by using the "lsqproject" option of smartpca. Heterozygote genotypes were converted to homozygous before the analysis by randomly sampling reads if positions were covered by multiple reads in each ancient individual. To prevent DNA damage from skewing the results, C → T and G → A transitions were removed from both the contemporary and ancient individuals.

ADMIXTURE Analysis. We performed model-based clustering analysis using the ML approach implemented in ADMIXTURE (46). We ran ADMIXTURE on a reference panel of 917 individuals from the work by Raghavan et al. (9), only including populations from North America, Central America, South America, Siberia, the Arctic, and Greenland. The analysis included 199,285 sites. Each ancient individual had the following overlap with the modern dataset: Shuká Kåa (2,327); 302 (59,205); 443 (79,233); and 939 (43,441). Native American populations in the dataset were masked for nonnative ancestry. We assumed $K = 3$ to $K = 15$ ancestral clusters. For each value of *K*, 10 replicate runs were performed, and the run with the greatest likelihood was selected for additional analysis. Clusters $K = 2$ through $K = 7$ are presented in Fig. S10.

Because the ancient individuals had varying overlap with the contemporary reference panel, we projected the ancient individuals onto the ancestral cluster allele frequencies inferred from the contemporary individuals with the same method described in ref. 47. This method also prevents the ancient genotypes from affecting the clustering solutions of the contemporary individuals. The low-coverage ancient genotypes were called in the same manner as in the principal coordinate analysis. C → T and G → A transitions were removed from both the contemporary and ancient individuals.

Mitochondrial Genome Analyses. The complete mitochondrial genomes of 3 ancient and 52 modern individuals belonging to the haplogroup D4h3a were phylogenetically compared by building a tree under the maximum parsimony criterion (Dataset S1) as previously described (6), and it is the only method that allows for the identification of exact motifs of new subhaplogroups. Coalescence ages were estimated by using both ML and Bayesian Evolutionary Analysis Sampling Trees (BEAST) computations. In particular, the ML estimates were calculated using PAML 4.5, in which the entire mitochondrial genome was partitioned into control and coding regions and analyzed under the HKY85-type substitution model with Γ-distributed rates. The clock hypothesis was confirmed by a likelihood ratio test. We used the same model (HKY85 with Γ-distributed rates plus invariant sites) and a strict clock for the Bayesian computation performed with BEAST 1.8.3. Radiocarbon dates of ancient specimens were used as priors, and the entire D4h3 clade was set as monophyletic using rCRS (48) as an outgroup and approximating the maximum root eight (corresponding to the L3 node) at ~65 kya (95% confidence interval = 60–70 kya) as in previous studies (6). The chain length was set at 10^7, disregarding the first 10% of data and logging the parameter values every 10^4 states. The final trace was explored using Tracer 1.6. The same BEAST settings were used to estimate the ages of haplogroup D4h3 (and its subclades) and infer population size changes in the population(s) carrying the D4h3 haplogroup through a Bayesian skyline plot.

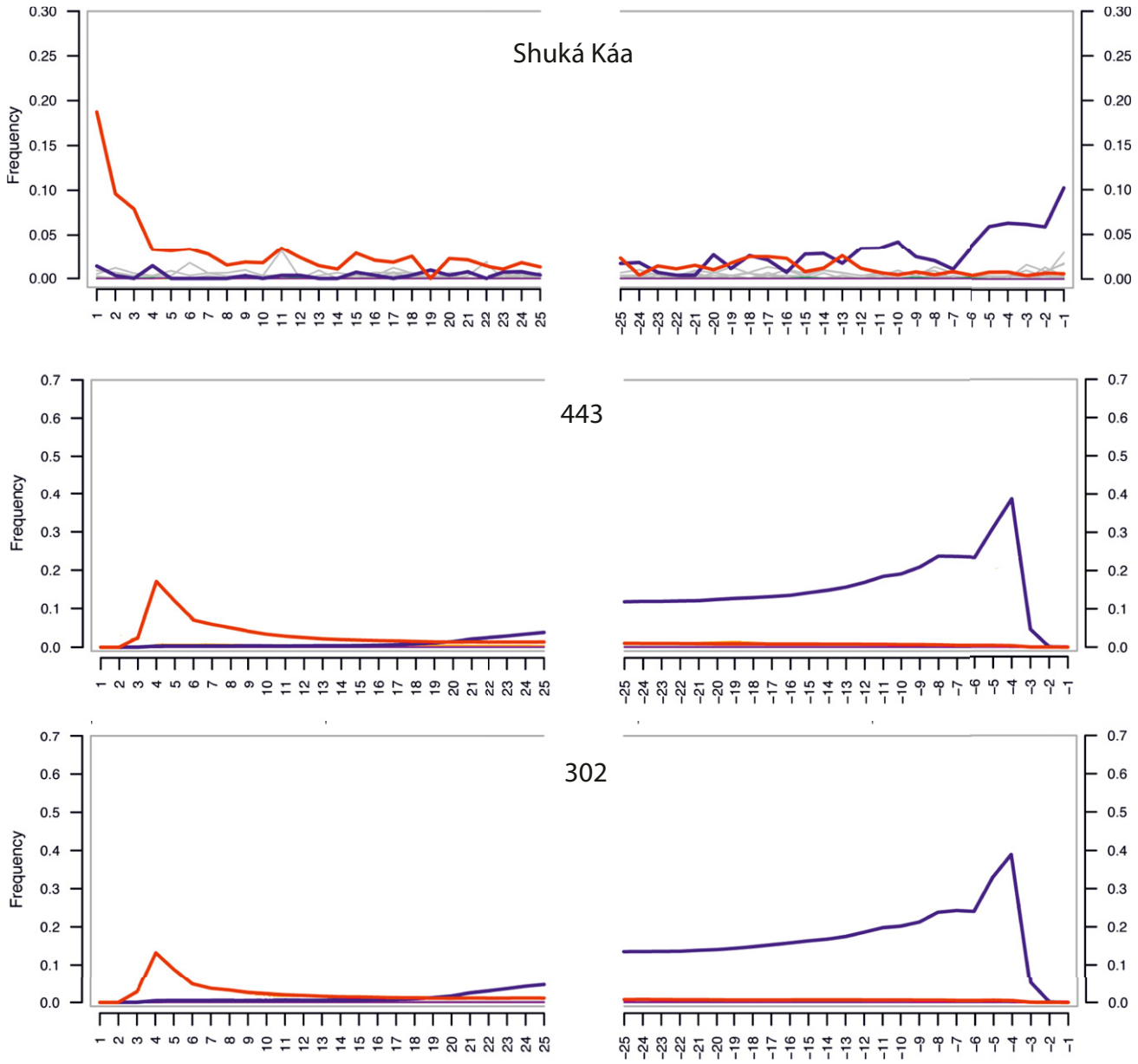
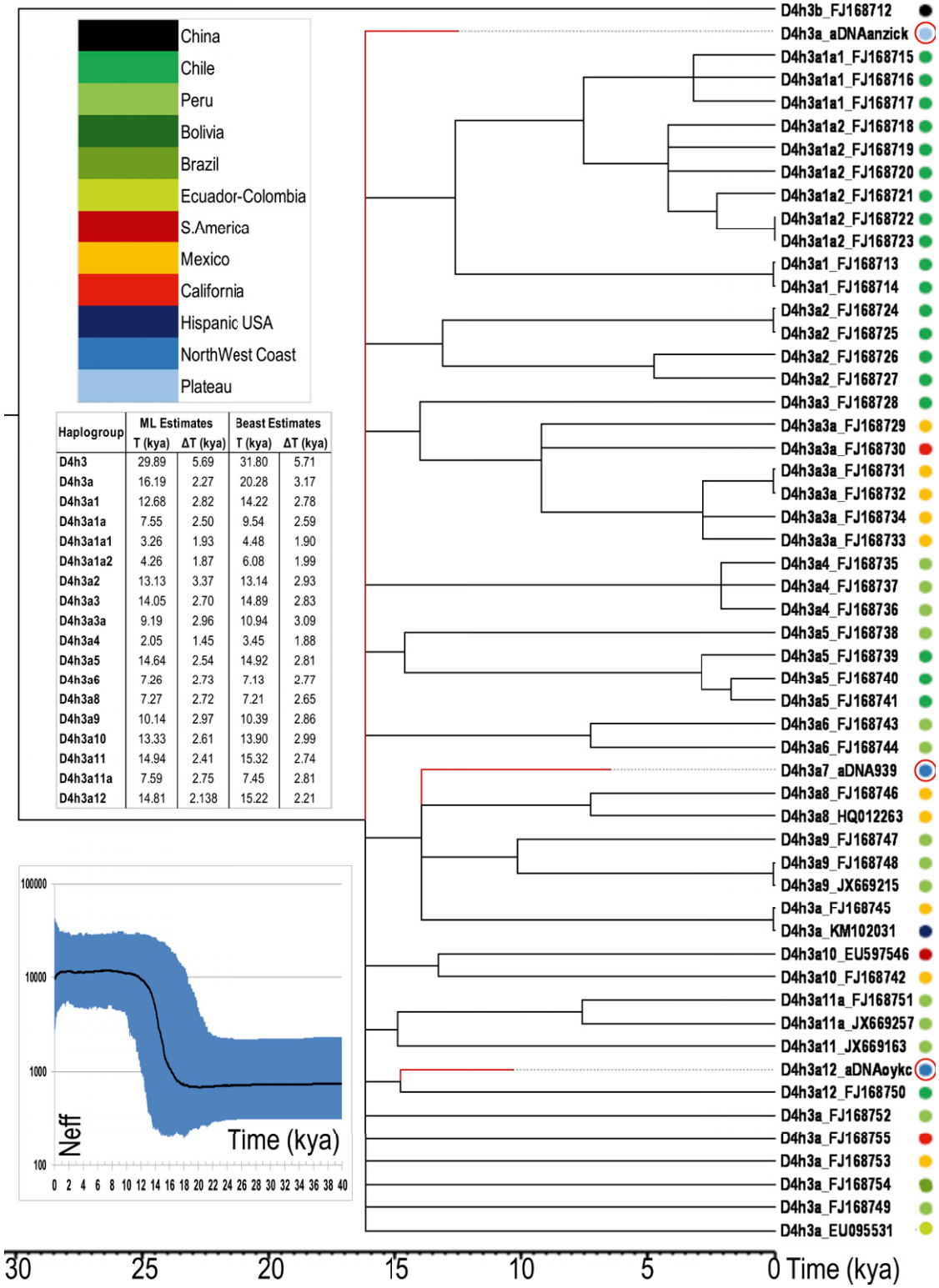


Fig. S1. Phylogeny of mtDNA haplogroup D4h3. The topology was inferred by maximum parsimony. An ML timescale is indicated below the tree. *Insets* show the geographic origin of samples, age estimates for the main nodes, and the Bayesian skyline plot. Neff indicates the effective number of females. Ancient mitochondrial genomes are circled in red.



Haplogroup	ML Estimates		Beast Estimates	
	T (kya)	ΔT (kya)	T (kya)	ΔT (kya)
D4h3	29.89	5.69	31.80	5.71
D4h3a	16.19	2.27	20.28	3.17
D4h3a1	12.68	2.82	14.22	2.78
D4h3a1a	7.55	2.50	9.54	2.59
D4h3a1a1	3.26	1.93	4.48	1.90
D4h3a1a2	4.26	1.87	6.08	1.99
D4h3a2	13.13	3.37	13.14	2.93
D4h3a3	14.05	2.70	14.89	2.83
D4h3a3a	9.19	2.96	10.94	3.09
D4h3a4	2.05	1.45	3.45	1.88
D4h3a5	14.64	2.54	14.92	2.81
D4h3a6	7.26	2.73	7.13	2.77
D4h3a8	7.27	2.72	7.21	2.65
D4h3a9	10.14	2.97	10.39	2.86
D4h3a10	13.33	2.61	13.90	2.99
D4h3a11	14.94	2.41	15.32	2.74
D4h3a11a	7.59	2.75	7.45	2.81
D4h3a12	14.81	2.138	15.22	2.21

Fig. S2. DNA damage patterns for the ancient individuals. Random subset of all mapped reads for each ancient individual. The mismatch frequency is relative to the reference as a function of read position: C to T in red and G to A in blue.

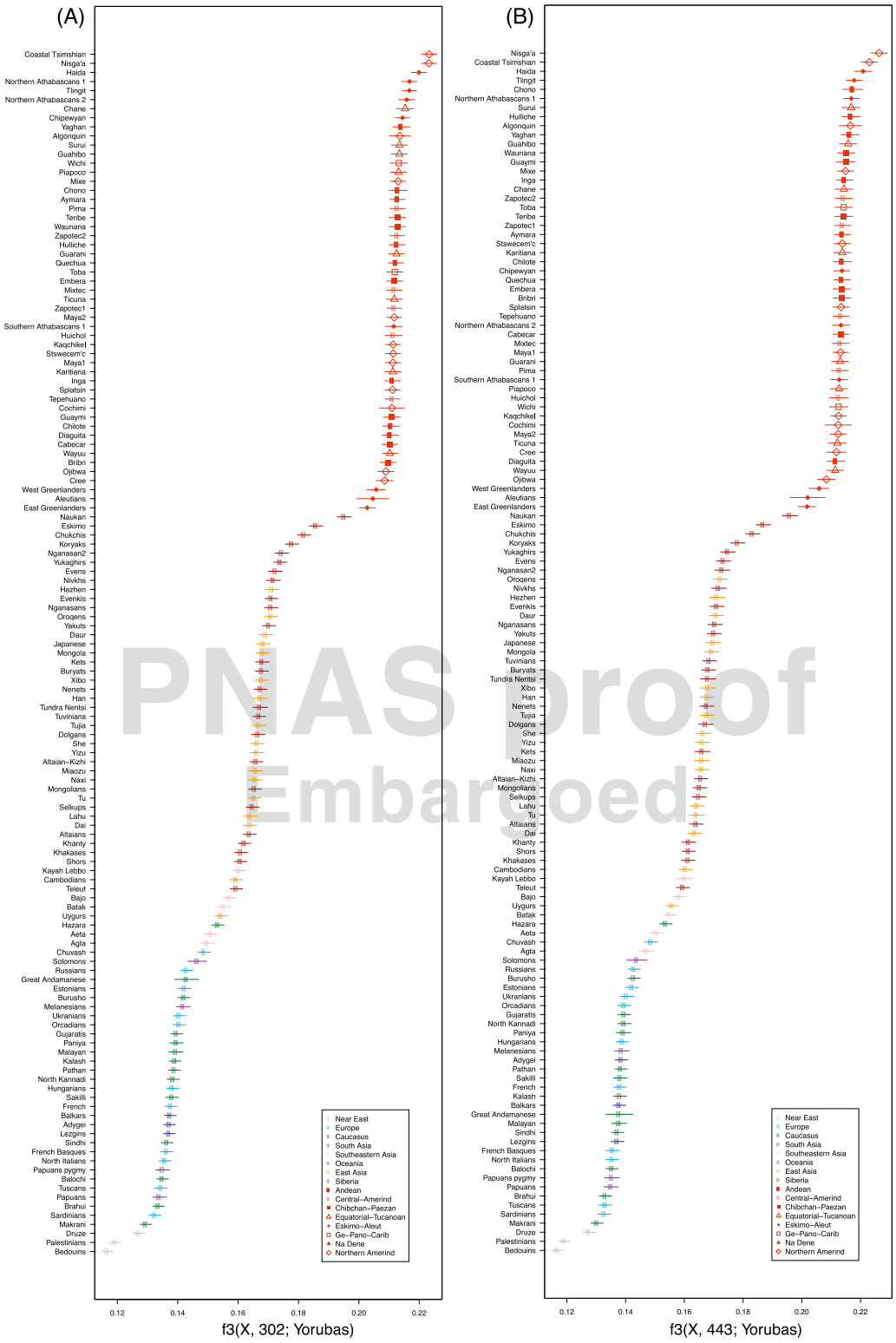


Fig. S3. (Continued)

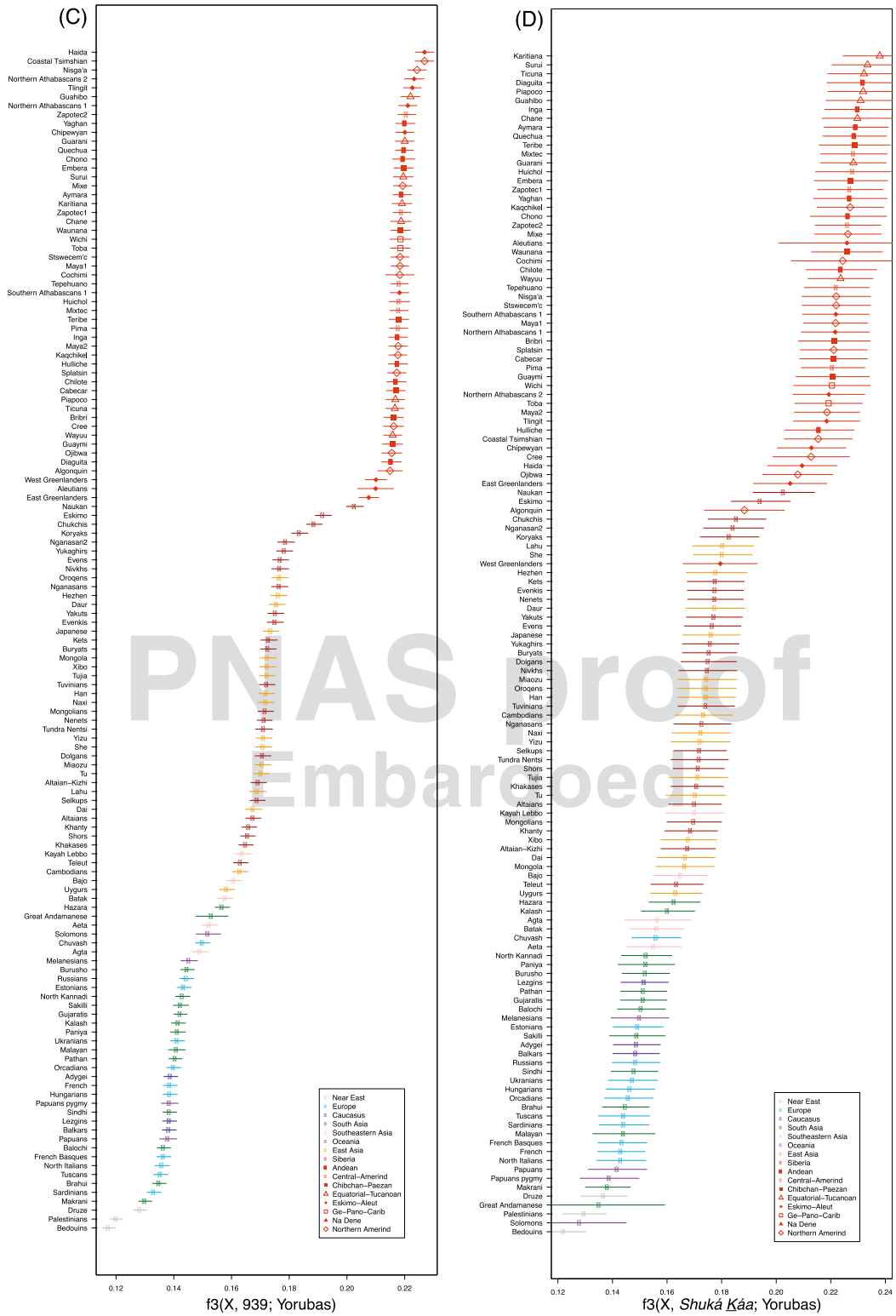


Fig. S3. Genetic affinity of the ancient individuals to contemporary human populations: (*A*) 302, (*B*) 443, and (*C*) 939 all show greater genetic affinity with Native American groups than other global populations. Heat maps summarize the outgroup f_3 statistics that estimate the amount of shared genetic drift between each ancient individual and each of 156 contemporary populations since their divergence with the African Yoruban population.

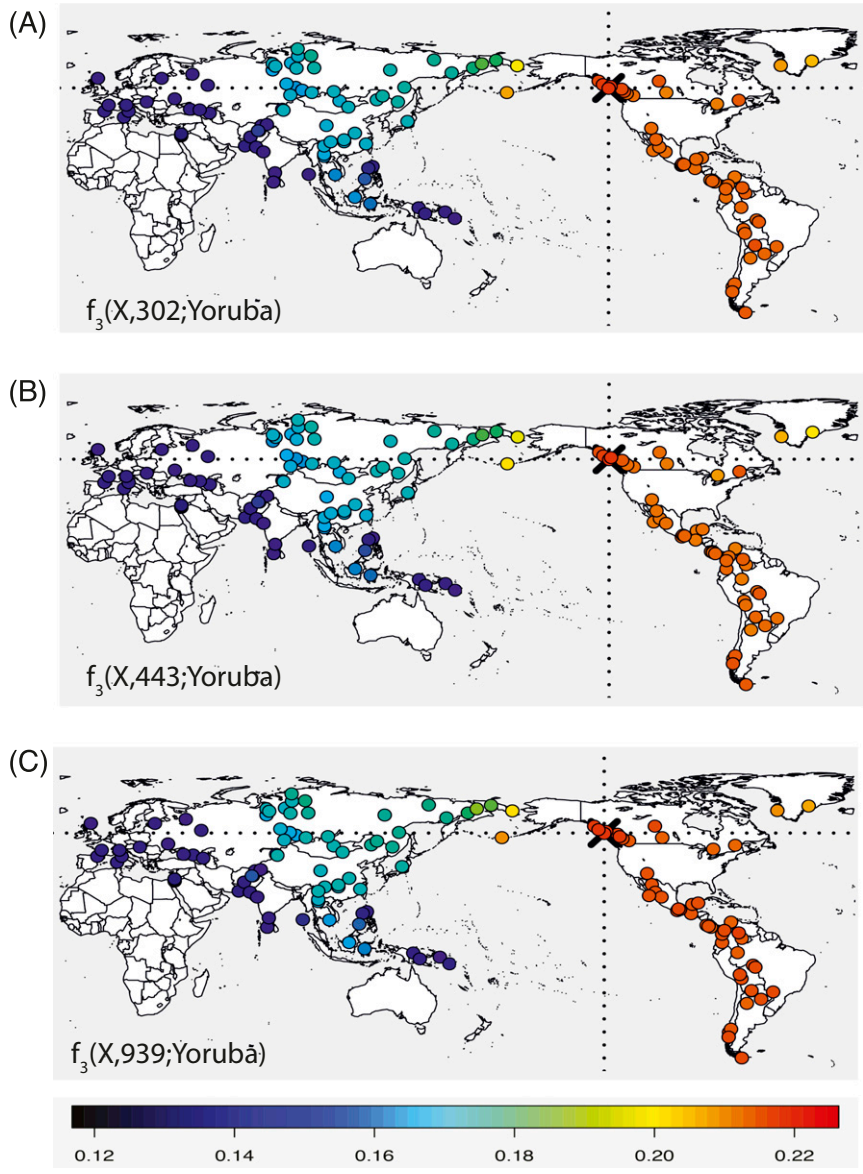


Fig. S4. Ranked outgroup f_3 statistics examining shared genetic history: (*A*) 302 and (*B*) 443 both show a close relationship to indigenous populations from the Northwest Coast, and (*C*) 939 also has close affinity to individuals from the Northwest Coast but also, the Yaqui from the American Southwest. (*D*) Shuká Káa shows a close affinity to populations from the American Southwest; however, the rank is not statistically significant.

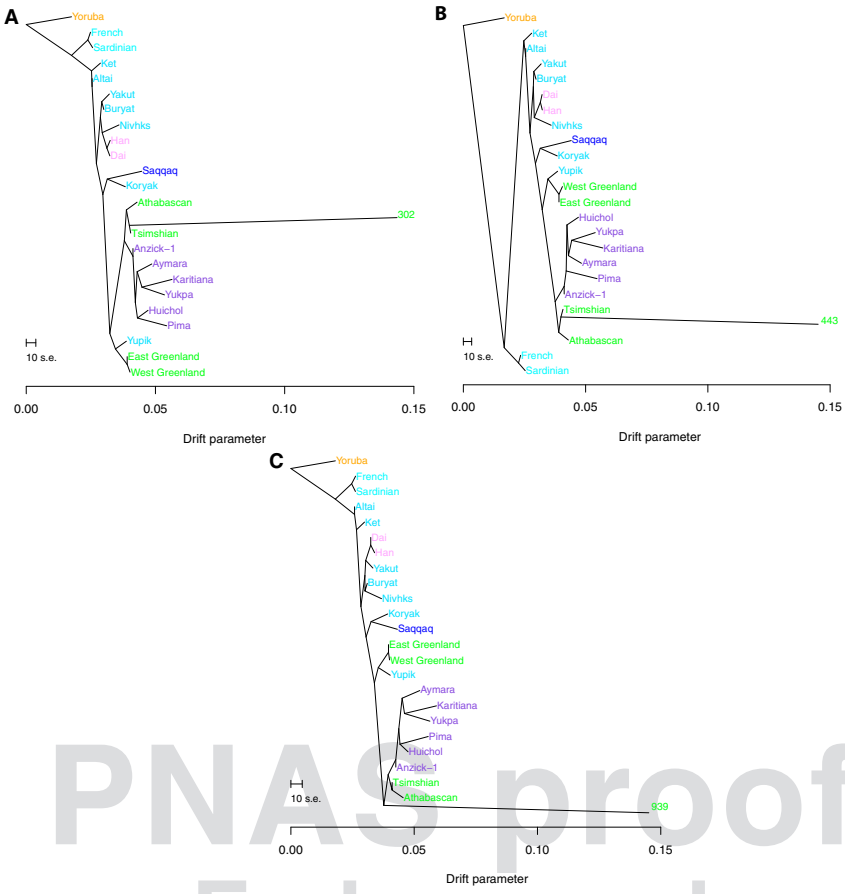


Fig. S5. ML trees. ML tree generated by TreeMix using whole-genome sequencing data, with the Tsimshian genome masked for European ancestry: (*A*) 302, (*B*) 443 and (*C*) 939.

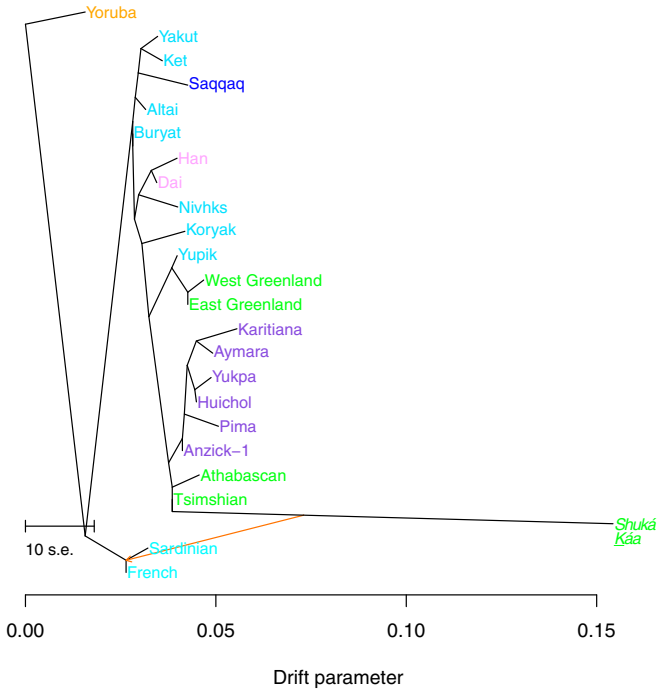


Fig. S6. ML tree with Shuká Káa and one migration. Samples and parameters are the same as in Fig. 2*B*. Adding the migration event causes Shuká Káa to fall into the Pacific Northwest clade (Athabascan and Tsimshian), with an edge connecting Shuká Káa and Europeans. This result may be a product of contamination (Table S2).

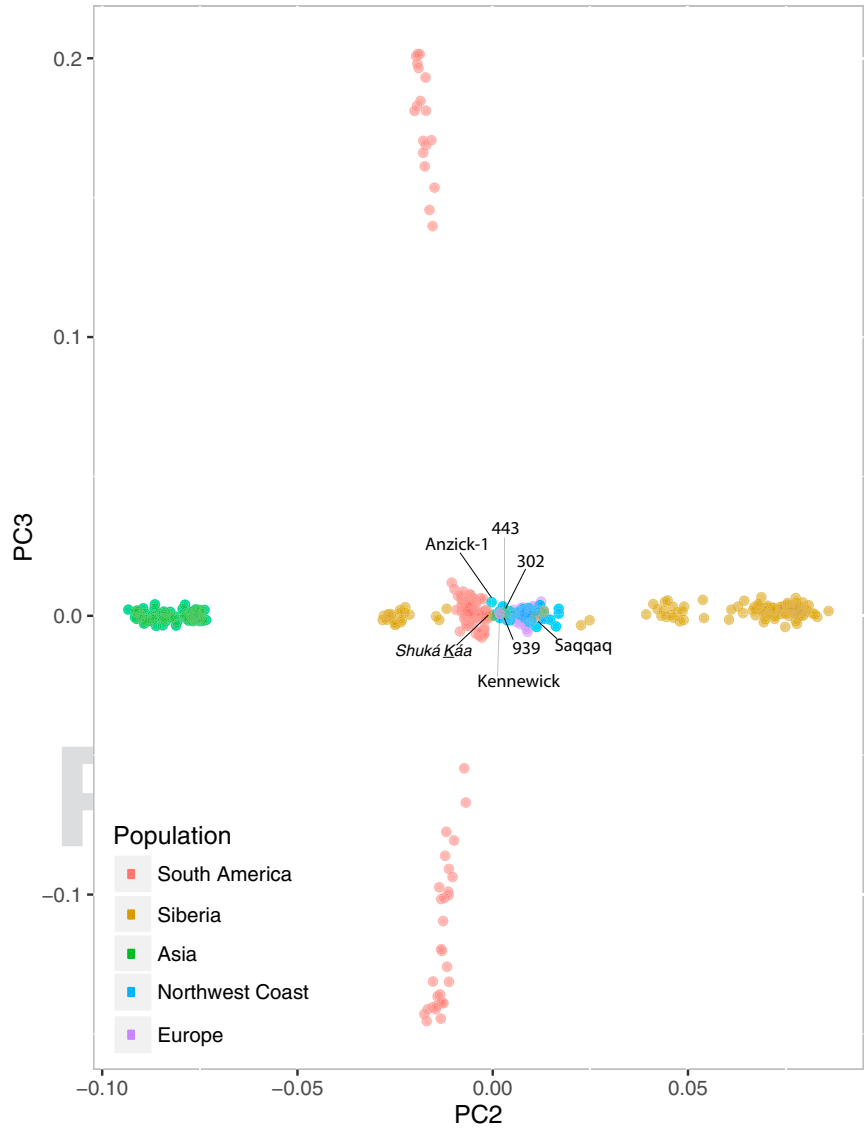


Fig. S7. Principal components (PC) analysis (PC2 and PC3). Shuká Káa, 939, 302, 443, Anzick-1 (11), Saqqaq (30), and Kennewick Man (14) are projected onto a set of non-African populations from the work by Raghavan et al. (9), with Native American populations masked for nonnative ancestry.

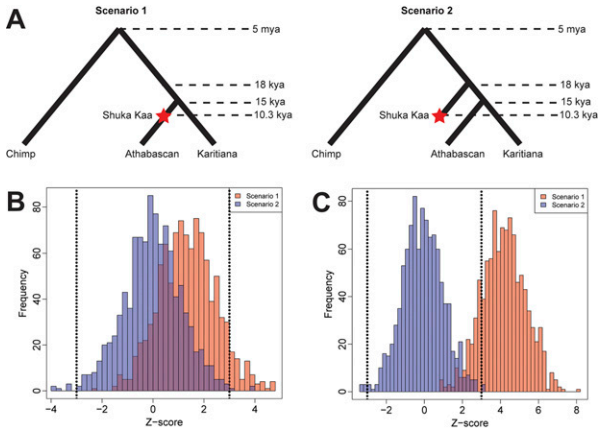


Fig. S8. Simulated *Z*-score distributions for different scenarios relating Shuká Káa to Northwest Coast and South American lineages. (*A*) Two scenarios relating Shuká Káa to Northwest Coast and South American linages. Scenario 1 places Shuká Káa on the branch leading to Northwest Coast groups. Scenario 2 places Shuká Káa ancestral to both Northwest Coast and South American groups. (*B*) Distributions of *Z* scores under each scenario, in which each distribution is computed from 1,000 independent simulated replicates and where each replicate is based on 200 independent 100-kb genomic regions. The number of *D*-statistic informative sites is similar to the number in our empirical dataset with Shuká Káa. Only a small fraction of simulated replicates for scenario 1 is able to reject the null hypothesis, although Shuká Káa is on the Northwest Coast branch. In contrast, the majority of replicates for scenario 2 cannot reject the null hypothesis as expected. (*C*) Distributions of *Z* scores under each scenario, in which each distribution is computed from 1,000 independent simulated replicates and where each replicate is based on 200 independent 100-kb genomic regions. The number of *D* statistic informative sites is an order of magnitude larger than in *B*. A large fraction of simulated replicates for scenario 1 is able to reject the null hypothesis, indicating that Shuká Káa has higher affinity to the Northwest Coast than to the South American lineage. Therefore, more data provided us with the power to reject the null hypothesis. In contrast, the majority of replicates for scenario 2 still cannot reject the null hypothesis as expected.

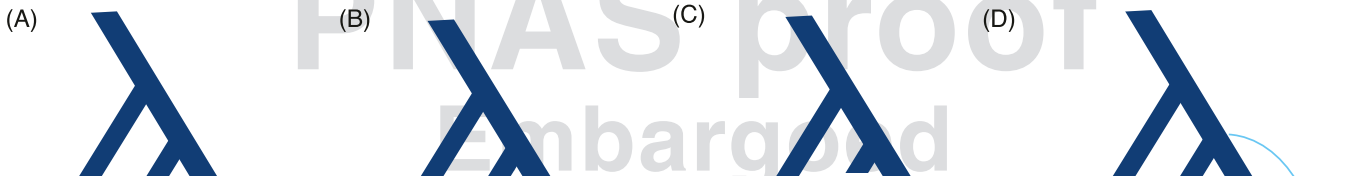


Fig. S9. Scenarios tested by the *D* statistic involving the spread of mitochondrial haplogroups in the Northwest Coast. *A* supports Anzick-1 being basal to both Shuká Káa and 939, which all share the D4h3a haplogroup. *B* supports Anzick-1 being basal to 939 and 443/302, which carry the mitochondrial haplogroup A2. *C* supports Shuká Káa as basal to Kennewick Man and the younger ancients from the Northwest Coast. *D* supports a closer relationship between Anzick-1 and Kennewick Man than to Shuká Káa. Table S2 shows sequence-based *D*-statistic results.

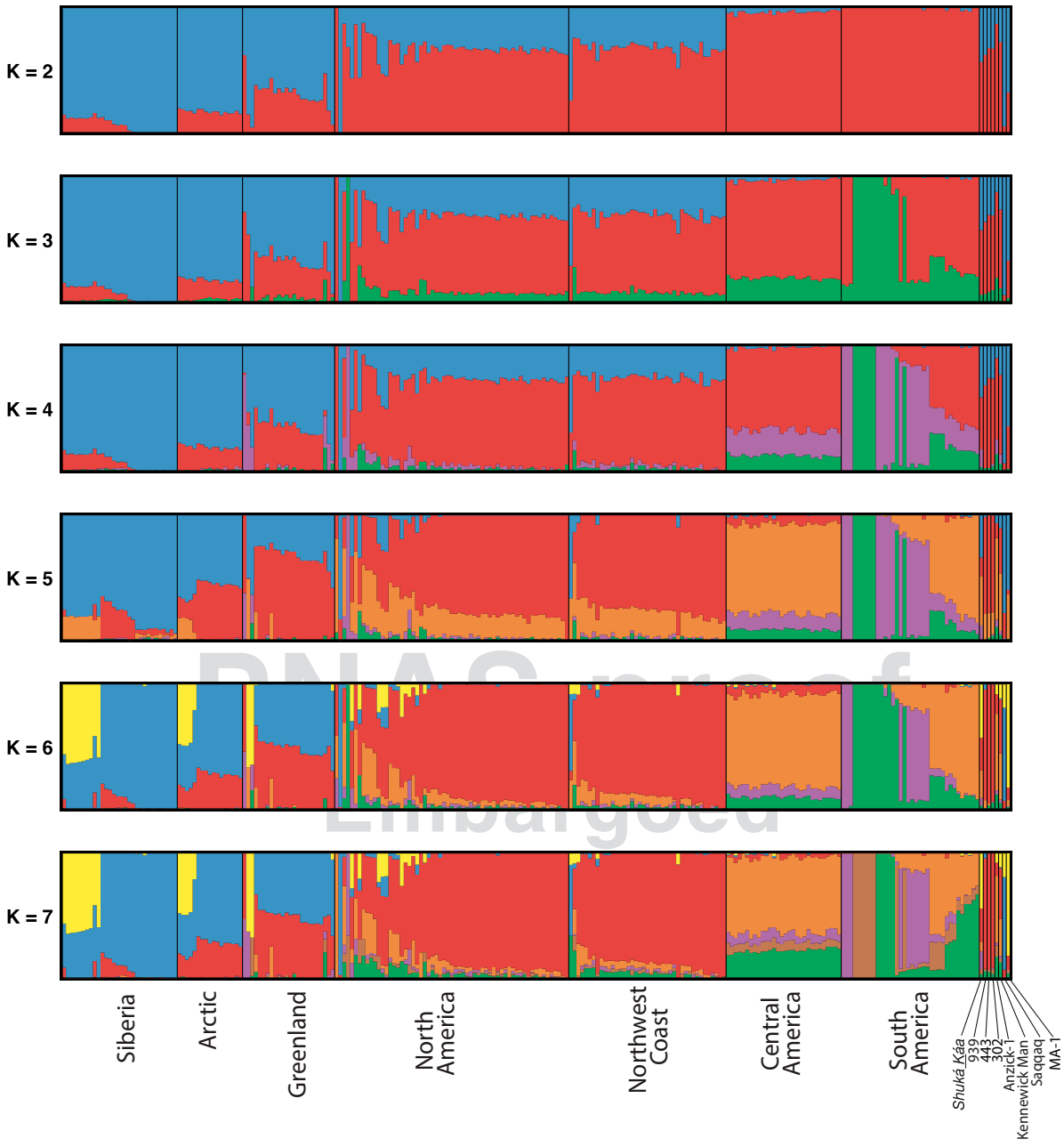


Fig. S10. Cluster analysis generated by ADMIXTURE. Set includes indigenous populations from the Americas, Siberia, and the Arctic, Greenland and the Anzick-1, Kennewick, Saqqaq, Shuká Káa, 939, 302, and 443 samples. The number of displayed clusters is $K = 2$ through $K = 7$.

Table S1. Contamination estimates based on mitochondrial and X-chromosomal data

Sample	Method	Mean coverage	Contamination estimate	*P* value	95% Confidence intervals
Shuká Káa	Mitochondria	101.46	0.02	N/A	0.01–0.03
Shuká Káa	X chromosome	Insufficient coverage	—	—	—
302	X chromosome	3.98	0.008		1.23×10^{-9}
302	Mitochondria	164.39	0.01	N/A	0.00–0.02
443	Mitochondria	183.90	0.01	N/A	0.00–0.02

Table S2. Sequence-based *D* statistics

Test	*Z*	*Y*	*X*	*D* statistic	SE	*Z* score	*D* statistic corrected*	SE	*Z* score	Figure
1	Anzick-1	Shuká Káa	443	0.07639752	0.01890633	4.040844	0.07425173	0.01126007	6.594251	Fig. S6
2	Anzick-1	Shuká Káa	302	0.1215492	0.02157855	5.632872	0.132017	0.02036344	6.48304	Fig. S6
3	Anzick-1	Shuká Káa	939	0.05583756	0.02655283	2.102885	0.05890426	0.03539308	1.664287	Fig. S6
4	Anzick-1	Shuká Káa	Kennewick	−0.1488178	0.03776425	−3.940706	−0.1399817	0.03768238	−3.714778	Fig. S6
5	Anzick-1	Shuká Káa	Karitiana	−0.1	0.01634114	−6.119523	−0.1091605	0.01603733	−6.806647	Fig. 3*A*
6	Shuká Káa	Tsimshian	Athabascan	−0.03193613	0.03147981	−1.014496	0.1039187	0.03577532	2.904699	—
7	Shuká Káa	Athabascan	Kennewick	−0.04731861	0.05731246	−0.825625	−0.0437018	0.0507274	−0.861502	—
8	Shuká Káa	Karitiana	Athabascan	−0.0223518	0.02228589	−1.002957	−0.02941176	0.04035663	−0.728796	Fig. 3*C*
9	Shuká Káa	Karitiana	Tsimshian	0.01579467	0.02425089	0.6513026	0.01526104	0.02447883	0.6234384	Fig. 3*C*
10	Shuká Káa	939	443	0.05988701	0.03436534	1.742657	0.03121808	0.04262305	0.7323999	Fig. 3*B*
11	Shuká Káa	939	302	0.1008264	0.04208303	2.395893	0.1383589	0.05131352	2.696321	Fig. 3*B*
12	Shuká Káa	939	Tsimshian	−0.09117221	0.0380607	−2.395443	−0.02724043	0.01451433	−1.876795	Fig. 3*B*
13	Shuká Káa	939	Kennewick	−0.04035874	0.06831446	−0.590779	0.005610852	0.08670631	0.0646166	Fig. S6
14	Shuká Káa	302	Kennewick	0.01503759	0.06298859	0.2387352	0.01526104	0.02447883	0.6234384	Fig. S6
15	Shuká Káa	443	Kennewick	−0.0437018	0.0507274	−0.861502	−0.04761905	0.132735	−0.358752	Fig. S6
16	Shuká Káa	939	Athabascan	−0.02339181	0.03708638	−0.630738	−0.04035874	0.06831446	−0.590779	Fig. 3*B*
17	939	Athabascan	Karitiana	0.03644538	0.00917138	2.973817	0.02750544	0.00736812	3.733041	—
18	939	Tsimshian	Karitiana	0.04521475	0.00956955	2.724854	0.03587975	0.00969878	3.699408	—
19	Shuká Káa	Tsimshian	Yup'ik	0.06519655	0.02374105	2.746152	0.07499347	0.01714268	4.374664	—
20	Shuká Káa	Tsimshian	Inuit	0.05819593	0.0220833	2.635291	0.08775409	0.02309288	3.800049	—
21	Shuká Káa	Athabascan	Inuit	0.03925417	0.02249508	1.745011	0.06118252	0.0178027	3.436698	—
22	Shuká Káa	Athabascan	Yup'ik	0.0343853	0.02322224	1.480705	0.06519655	0.02374105	2.746152	—
23	Shuká Káa	Inuit	302	−0.08787432	0.02235151	−3.931472	−0.01849764	0.00906124	−4.041402	—
24	Shuká Káa	Inuit	443	−0.07419712	0.02015663	−3.681027	−0.04598959	0.01015894	−4.527009	—
25	Shuká Káa	Inuit	939	0.02229055	0.02857812	0.7799864	0.0210728	0.03066293	0.6872402	—
26	Shuká Káa	Yup'ik	302	−0.08775409	0.02309288	−3.800049	−0.03443418	0.00787645	−4.371786	—
27	Shuká Káa	Yup'ik	443	−0.09149856	0.01999172	−4.576824	−0.04298268	0.00773691	−5.555533	—
28	Shuká Káa	Yup'ik	939	−0.01936483	0.0282451	−0.685599	−0.04068522	0.05322185	−0.764445	—
29	939	Inuit	Athabascan	−0.07921773	0.00851659	−9.301574	N/A	N/A	N/A	Fig. 3*D*
30	939	Yup'ik	Tsimshian	−0.07381749	0.00986385	−7.483635	N/A	N/A	N/A	Fig. 3*D*
31	302	Inuit	Athabascan	−0.09807331	0.00810703	−12.09731	N/A	N/A	N/A	Fig. 3*D*
32	302	Yup'ik	Tsimshian	−0.1061131	0.00963526	−11.013	N/A	N/A	N/A	Fig. 3*D*
33	443	Inuit	Athabascan	−0.08621938	0.00762381	−11.30922	N/A	N/A	N/A	Fig. 3*D*
34	443	Yup'ik	Tsimshian	−0.1005294	0.00975614	−10.30421	N/A	N/A	N/A	Fig. 3*D*

Q:6 The sequences from Yup'ik (Esk17), Athabascan (Athabascan 2), Inuit (Greenlander 2), and masked Tsimshian are from the work by Raghavan et al. (9); the sequences from the PaleoEskimos are from the work by Raghavan et al. (8). The Kennewick sequence is from the work by Rasmussen et al. (14), and the Karitiana sequence derives from the work by Rasmussen et al. (11). Significant *D* statistics are those associated with scores of $|Z| > 3$. For the *D* statistics of the form *D*(Chimp, (*Z*, (*Y*, *X*))), population *Z* is closer to population *X* than to population *Y* for $D < 0$, and it is closer to population *Y* than to population *X* for $D > 0$. Significant *D* statistics are color-coded blue, where a closer relationship is represented between populations in blue relative to the third population in white in the test [the Kennewick sequence from the work by Rasmussen et al. (14)].

Q:7 *The corrected *D* statistic was performed for tests including Shuká Káa and 939 as described in *SI Text*, *Sequence Data-Based* D *Statistic*.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (PDF)
Dataset S3 (PDF)
Dataset S4 (PDF)

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

Q: 1_In SI Text, supplemental documents citations were changed to Datasets S1–S4. Please verify.

Q: 2_Please clarify "corrected" information in sentence "Since the time of the burial, the spelling..."

Q: 3_PNAS uses roman type for restriction enzymes and sites. Please check throughout the proof and confirm/correct as necessary.

Q: 4_Figs. S1-S4 legends match Figs. S2, S1, S4, and S3. Please verify the figure order.

Q: 5_Please define "N/A" in Table S1 legend.

Q: 6_A-D deleted from Fig. S6 citations in Table S2, because they do not appear in the figure. Please verify.

Q: 7_Please verify information in Table S2 legend and footnote and define "N/A."