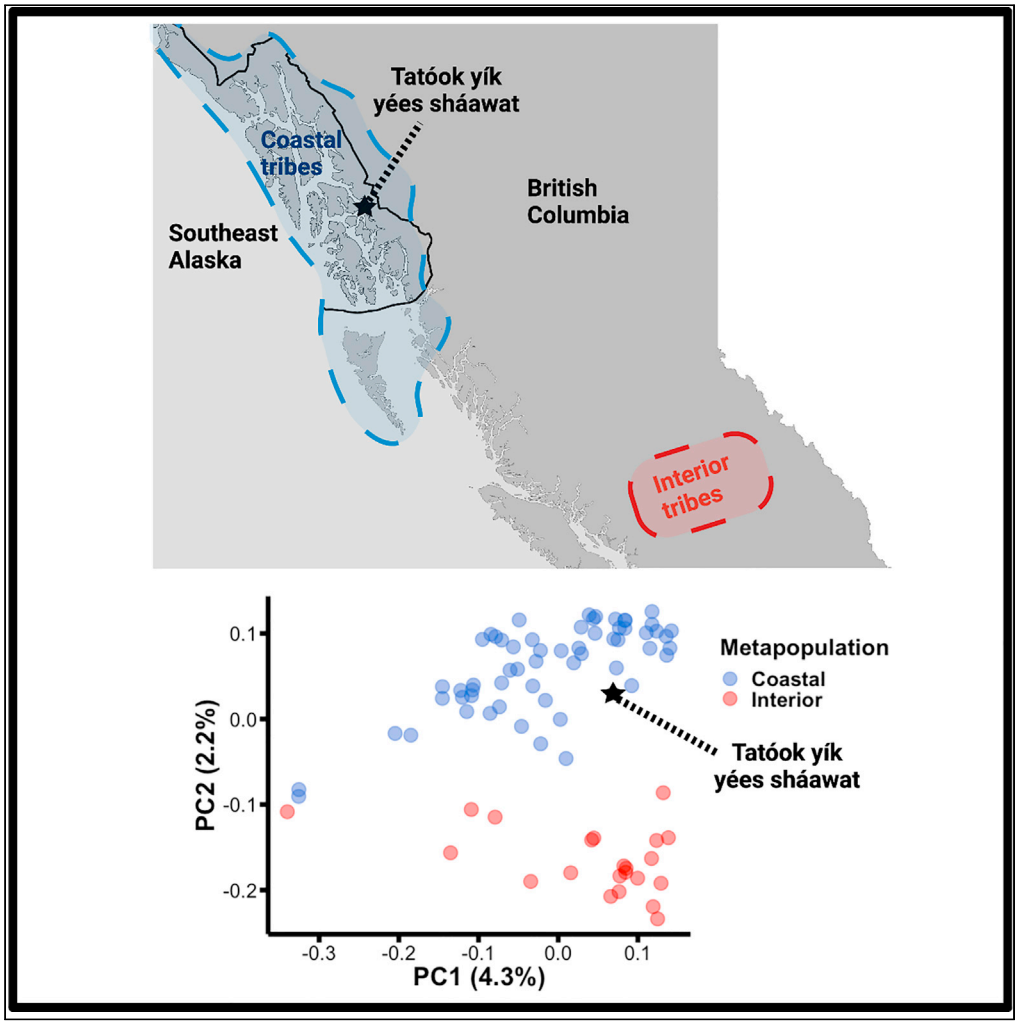
iScience

CellPress
OPEN ACCESS

Article

A paleogenome from a Holocene individual supports genetic continuity in Southeast Alaska



Alber Aqil, Stephanie Gill, Omer Gokcumen, ..., Jane L. Smith, Timothy T. Heaton, Charlotte Lindqvist

cl243@buffalo.edu

Highlights

We obtained genomic data from a ~3,000-year-old female from Southeast Alaska

With community engagement, the individual was named *Tatóok yík yées sháawat* (*TYYS*)

TYYS is most closely related to the Indigenous peoples of the Pacific Northwest Coast

The *Saqqaq* (Paleo-Inuit) genome may harbor Northern Native American ancestry

Aqil et al., iScience *26*, 106581
May 19, 2023 © 2023 The Author(s).
https://doi.org/10.1016/j.isci.2023.106581

iScience

CellPress
OPEN ACCESS

Article

A paleogenome from a Holocene individual supports genetic continuity in Southeast Alaska

Alber Aqil,[1] Stephanie Gill,[1] Omer Gokcumen,[1] Ripan S. Malhi,[2] Esther Aaltséen Reese,[3] Jane L. Smith,[4] Timothy T. Heaton,[5] and Charlotte Lindqvist[1,6,*]

SUMMARY

Many specifics of the population histories of the Indigenous peoples of North America remain contentious owing to a dearth of physical evidence. Only few ancient human genomes have been recovered from the Pacific Northwest Coast, a region increasingly supported as a coastal migration route for the initial peopling of the Americas. Here, we report paleogenomic data from the remains of a ~3,000-year-old female individual from Southeast Alaska, named *Tatóok yík yées sháawat* (*TYYS*). Our results demonstrate at least 3,000 years of matrilineal genetic continuity in Southeast Alaska, and that *TYYS* is most closely related to ancient and present-day northern Pacific Northwest Coast Indigenous Americans. We find no evidence of Paleo-Inuit (represented by *Saqqaq*) ancestry in present-day or ancient Pacific Northwest peoples. Instead, our analyses suggest the *Saqqaq* genome harbors Northern Native American ancestry. This study sheds further light on the human population history of the northern Pacific Northwest Coast.

INTRODUCTION

Many details concerning the early population histories of the Indigenous peoples of North America remain unknown. Studies of ancient and modern genomes, however, have in recent years provided considerable insights into the timing and routes of the entrance from Siberia into the Americas. In particular, it has been suggested that there have been at least three distinct waves of migration: the first wave, which contributed ancestry to all non-Inuit Indigenous peoples of the Americas[1,2]; a second wave, which included the Paleo-Inuit (previously referred to as Paleo-Eskimos), a people of the Dorset culture who reportedly came to the Americas from Siberia ~6 thousand years ago (ka)[3,4]; and the third wave, which included the Neo-Inuit (previously known as Neo-Eskimos), a people of the Thule culture who settled in the Arctic ~1 ka, possibly replacing the Paleo-Inuit, and giving rise to the present-day Inuit.[3]

The first wave of migration into the Americas has been studied in great detail. For example, genomic studies have shown that a small group splintered off the larger East Asian population ~30 ka and subsequently split into two populations by ~24 ka.[5] Thereafter, each of these two populations interbred with the first people of Siberia, known as the Ancient North Siberians.[5,6] One of these admixed populations, called Paleo-Siberians, stayed in Siberia and became ancestors to the Koryak and Chukchi people (the speakers of the Chukotka-Kamchatkan languages); the second admixed population ultimately entered the Americas in the first wave of migration.[5–7] This latter lineage went through millennia of isolation in Beringia,[2,8,9] the landmass that extended from the Lena River Valley in Siberia to the Yukon Territory in Canada, linking Asia and North America during the Last Glacial Maximum.[10,11] From this isolated ancestral population, at least three lineages emerged[12,13]: the Ancestral Native Americans, who moved south of the ice sheets and became ancestors of non-Inuit Indigenous Americans; the Ancient Beringians, who did not move farther south than Alaska and have not been identified in studies of human samples sometime after ~9 ka[7,14]; and a hypothetical "Population A", known only from traces of ancestry left behind in Mesoamerican populations.[13,14]

There is increasing evidence that the Ancestral Native Americans migrated southward along the Pacific Northwest Coast, which provided a deglaciated, ecologically viable pathway after ~17–15 ka.[15,16] This entryway was open earlier than the interior corridor between the Cordilleran and Laurentide ice sheets that is thought to have become viable ~13 ka.[17,18] While moving south along the Pacific Northwest Coast,

[1]Department of Biological Sciences, University at Buffalo, Buffalo, NY 14260, USA

[2]Department of Anthropology and Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

[3]Wrangell Cooperative Association, Wrangell, AK 99929, USA

[4]USDA-Forest Service, Tongass National Forest, Petersburg, AK 99833, USA

[5]Department of Earth Sciences, University of South Dakota, Vermillion, SD 57069, USA

[6]Lead contact

*Correspondence: cl243@buffalo.edu

https://doi.org/10.1016/j.isci.2023.106581

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

the Ancestral Native Americans likely split into the Southern Native American (SNA) and Northern Native American (NNA) lineages ~15.5 ka.[1,7] Following the split, the NNA lineage likely expanded to broadly inhabit northern North America, including the northern Pacific Northwest Coast.[7,14]

The northern Pacific Northwest Coast surrounds the Gulf of Alaska, lapping around the islands of the Alaskan and British Columbia coasts.[19] Today, the region is home to the Tlingit, Haida, Tsimshian, Nisga'a, and Salishan-speaking peoples. It has been proposed that these northern coastal populations have a recent shared origin with inland populations of the Pacific Northwest, such as the Splatsin and Stswecem'c from interior British Columbia.[20] However, the timing of the divergence between the coastal and inland populations remains unknown. Additionally, it has been suggested that the Paleo-Inuit contributed ancestry to some of these populations of the Pacific Northwest[21]; there is, however, no consensus as to whether or when this admixture took place.[12]

Despite being a key interchange region for both the initial peopling of the Americas and later migrations, both southward and northward along the Pacific Northwest Coast, relatively few studies have been conducted about the genetic ancestry of the Indigenous peoples of the northern Pacific Northwest. This is due, in part, to the limited physical evidence from human remains and few paleogenomes thus far analyzed. Despite these issues, it has been shown that mid- to late-Holocene genomes obtained from individuals on the British Columbia coast[22–24] show affinity with genomes of present-day Indigenous Americans in the region, providing evidence of genetic continuity therein. Specifically, maternal genetic continuity for at least 5,500 years in coastal British Columbia, and the Tsimshian territory in particular, has been demonstrated.[24] In addition, exome sequences from 25 ancient individuals from the Prince Rupert Harbor region in the Tsimshian Territory of coastal British Columbia exhibit genetic affinity with modern Tsimshian people.[22]

The only ancient genome reported from Southeast Alaska so far, the ~10,300 calendar years before present (cal BP) *Shukáa Kaa* individual, shows a greater genetic affinity with Northern Native Americans than with other populations, suggesting 10,000 years of regional genetic continuity.[23] However, the hypothesis that it belongs to a lineage (perhaps Ancient Beringians) that was basal to both NNA and SNA lineages has not been rejected unequivocally.[23,25] Moreover, *Shukáa Kaa*'s link to present-day tribes in the region has not been established. Furthermore, the ultra-low genomic coverage of the sample, limiting the number of sites available for comparative analyses, and its unique mitochondrial haplotype that is not shared with present-day populations in the region do not preclude a hypothesis of population turnover after the early Holocene.

Here, we present a low-coverage genome from an ancient female individual dated to ~3,000 cal BP, excavated from Lawyer's Cave in Tlingit Territory on the Alaskan mainland east of Wrangell Island in the Alexander Archipelago of Southeast Alaska. The remains were found along with shell beads and a ~3,000-year-old bone awl. Under advisement from the Wrangell Cooperative Association (WCA), the federally recognized tribe associated with the area, we refer to this individual as *Tatóok yík yées sháawat* ("Young lady in cave"). To date, this is only the second ancient individual from Southeast Alaska (after *Shukáa Kaa*) that has been genetically confirmed as humans. By comparing the mitochondrial and nuclear genomic data from this Holocene individual with available data from other ancient and modern Indigenous American individuals, our aim was to investigate the complex population genetic history and structure of the Pacific Northwest.

RESULTS AND DISCUSSION

Dating, coverage, and sex determination for *Tatóok yík yées sháawat*

The sample from *Tatóok yík yées sháawat* (hereafter referred to as *TYYS*) is a small skeletal fragment of a humerus with a diameter of about 3 cm that was uncovered in Lawyer's Cave, located along Blake Channel on the southeastern Alaskan mainland east of Wrangell Island in the Alexander Archipelago (Figure 1). Radiocarbon dating of *TYYS* returned a date of 3,425 $\pm$ 50 ^{14}C years BP and the calibrated median date was estimated to be 2,950 cal BP ranging from 2,500 to 3,379 cal BP. The δ^{13}C value was −12.5‰, which falls within the range of stable carbon isotope data from Thule Culture individuals in southwestern Greenland and marine mammals,[26] indicating a diet almost entirely consisting of marine protein.

Shotgun Illumina sequencing and mitochondrial genome hybridization capture yielded a complete mitochondrial genome with an average coverage depth of 112X. The width and depth coverage of the nuclear

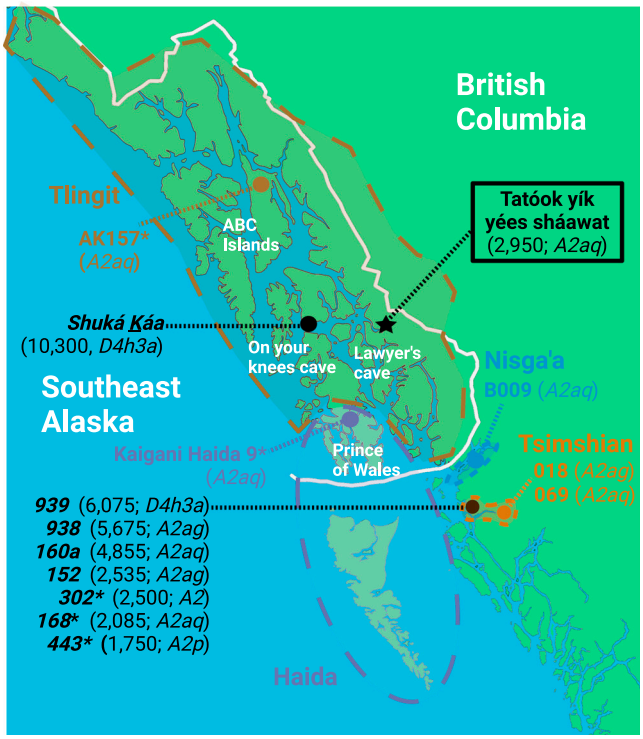


Figure 1. Map of the northern Pacific Northwest Coast and location of ancient and modern individuals

The territorial boundaries are rough demarcations (obtained from[24,31]). Ancient individuals and modern individuals that were used in mitochondrial DNA analysis are shown with black and colored text and symbols, respectively. Sample IDs of Individuals for whom the complete mitogenome was not available are appended with an asterisk. Each individual's mitochondrial haplotype and age (for ancient individuals) are shown in parentheses.

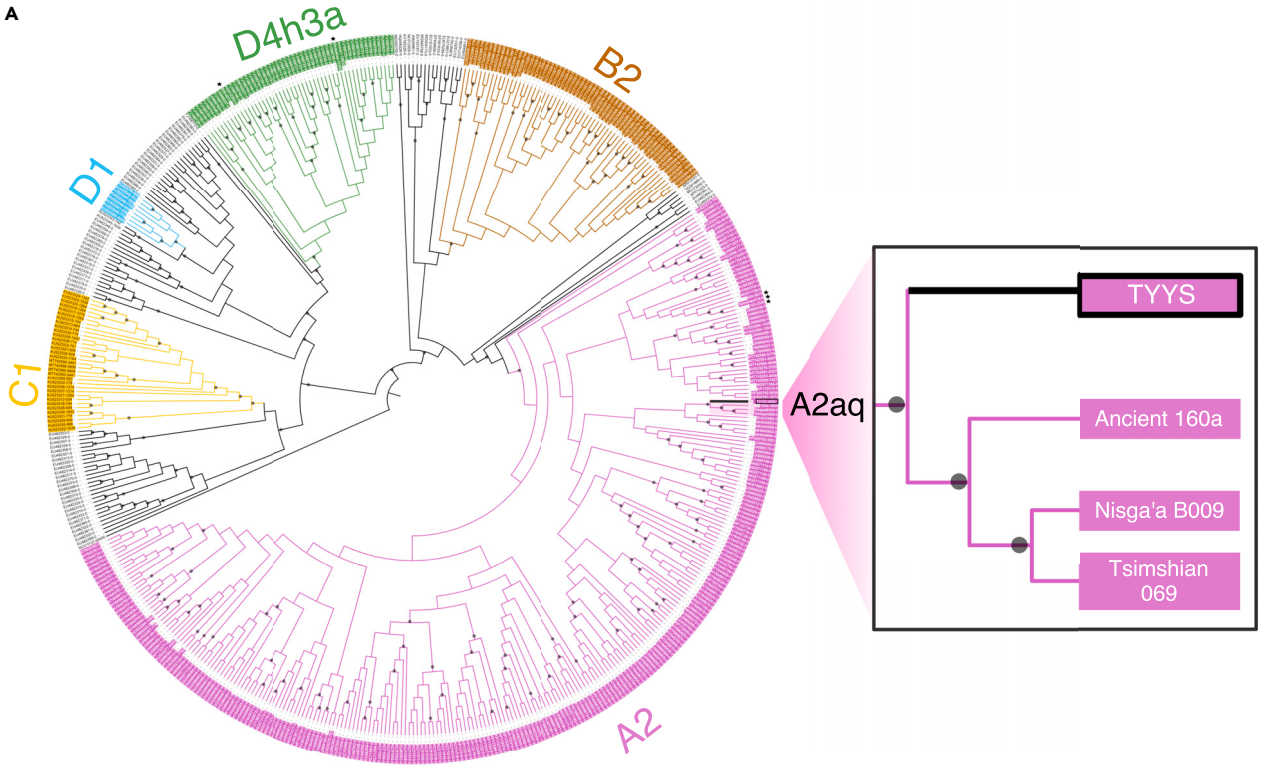
genome were 0.14 and 1.1X, respectively (Table S1). As anticipated for an ancient sample, the analysis of *TYYS*'s nuclear genome showed an increased rate of cytosine deamination at the 5′ end of the reads (Figure S1). We identified this individual as a female.

Tatóok yík yées sháawat belongs to mitochondrial haplotype A2aq, also found in present-day individuals from Southeast Alaska

The American founder mitochondrial haplogroups (A2, B2, C1, and D1) are today found throughout North, Central, and South America.[27–29] Haplogroup A2 remains the most frequently reported mitochondrial DNA (mtDNA) lineage in northern North America,[24] and it is particularly common along the Pacific Northwest Coast, reaching frequencies >90% among the Haida and Tlingit people.[30,31] Of 30 previously reported ancient individuals obtained from the Pacific Northwest Coast, 28 belonged to mtDNA haplogroups (or subhaplogroups) that have also been found among present-day Indigenous peoples in the region.[22,24,32,33] The only exceptions are the ~10,300 cal BP year-old *Shuká Káa*, the only ancient human genome from Southeast Alaska published so far, and the ~6,075 cal BP year-old *939* individual from British Columbia, who both belonged to haplotype D4h3a. This haplogroup, which is also shared with the 12,600 cal BP *Anzick-1* individual, remains largely undiscovered among present-day North Americans,[23,24] although it may be found among present-day Chumash.[33]

To date, temporal matrilineal genetic continuity in Southeast Alaska has not been reported. It is of note that matrilineal descent is particularly important in the context of the northern Pacific Northwest Coast because groups in this region have an exogamous matrilineal clan system, wherein a person's clan and moiety status is tied to the mother.[34–36] A strong association between maternal clan structure and mtDNA haplogroups has been found among these tribes today.[31]

To investigate whether ~2,950-year-old *TYYS* from Southeast Alaska exhibits a mitochondrial lineage that is also observed among present-day Indigenous peoples from the Pacific Northwest Coast, we constructed


A
D4h3a
B2
D1
C1
A2
A2aq
TYYS
Ancient 160a
Nisga'a B009
Tsimshian
069

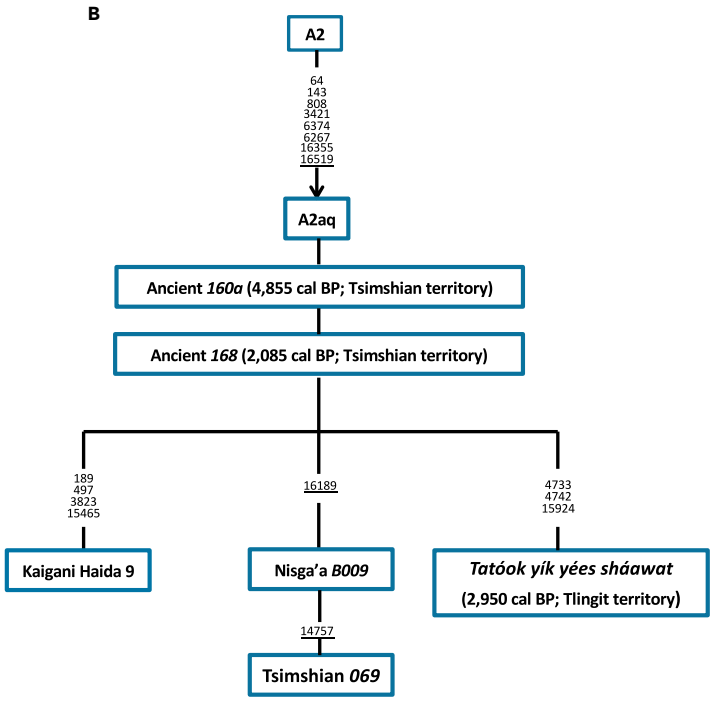

B
A2
64
143
808
3421
6374
6267
16355
16519
A2aq
Ancient 160a (4,855 cal BP; Tsimshian territory)
Ancient 168 (2,085 cal BP; Tsimshian territory)
189
497
3823
15465
16189
4733
4742
15924
Kaigani Haida 9
Nisga'a B009
Tatóok yík yées sháawat
(2,950 cal BP; Tlingit territory)
14757
Tsimshian 069

Figure 2. Maternal ancestry of *TYYS*

(A) Bayesian phylogenetic tree based on 523 complete mitochondrial genomes. The branches leading to genomes representing the founder haplogroups of the Americas are shown in different colors. Black dots at nodes indicate a posterior probability >0.9. The branches leading to the haplotype A2aq are highlighted with purple shading and expanded in the box to the right of the tree. The branch leading to *Tatóok yík yées sháawat* (*TYYS*) is shown in bold black. *TYYS* belongs to the A2aq clade that comprises three other individuals: a present-day Tsimshian individual, an ancient individual from Tsimshian territory, and a present-day Nisga'a individual.

(B) The coordinates are based on Cambridge Reference Sequence. Mutations shown here are transitions. Recurrent mutations are underlined. Note that the mitochondrial DNA sequence for Ancient 168 is unavailable. Also note that *AK157* (which likely belongs to A2aq as well) is not shown here because we do not have the information for all the relevant SNPs in *AK157* to place it in the phylogeny. However, we know that it belongs to haplogroup A2 and that it contains the 16355 mutation, which is not diagnostic of any haplotype other than A2aq within the haplogroup A.

a dataset containing 523 complete ancient and modern mitochondrial genomes (see Figure 2A), including sequences from *TYYS*, *Anzick-1*,[37] *Shuká Káa*,[23] and ancient individuals sampled from the Pacific Northwest Coast of British Columbia.[24] Figure 1 shows a map of the northern Pacific Northwest Coast along with the locations of individuals from whom mitogenomes used in this study were obtained.

Our phylogenetic analyses, using BEAST (Bayesian Evolutionary Analysis Sampling Trees),[38] showed that the clades representing the mitochondrial haplogroups A2, B2, C1, D1, and D4h3a were supported by posterior probabilities >0.9. *TYYS* formed a clade, with a posterior probability of 1.0, that also included three other whole mitogenomes sampled from the Pacific Northwest Coast: a living Nisga'a individual (Nisga'a B009), a living Tsimshian individual (Tsimshian 069), and a 4,855 cal BP ancient individual from Dodge Island in the Tsimshian territory (*160a*).[24] All three of these samples belong to the mitochondrial haplotype A2aq, a sublineage of the haplogroup A2. We note that Cui et al. 2013 named this haplotype A2ah; however, this haplotype is now recognized as A2aq,[39] while A2ah refers to a different haplotype that is common in South America.[40] We used HaploGrep2[41] to confirm that *TYYS*'s mitochondrial haplotype is indeed A2aq. Using the radiocarbon-dated ages from the ancient samples for divergence time calibration, the age of the most recent common ancestor of these four A2aq whole mitogenomes (*TYYS*, Ancient *160a*, Tsimshian 069, and Nisga'a B009) was estimated to be ~9,056 years BP (95% Highest Posterior Density: 5,535, 13,286 years BP). Among the individuals for whom the complete mtDNA was unavailable (and therefore not part of the phylogenetic analysis), an ~2,085 cal BP individual, *168*, from Tsimshian territory in British Columbia[22] and three modern-day Kaigani Haida individuals from Southeast Alaska[24,31] also harbor the haplotype A2aq (Figure 2B and Table S2). Based on mutations in the hypervariable control region, a present-day Tlingit individual belonging to the A2 haplogroup, *AK157*, was also shown to carry a diagnostic mutation (16355T) that places it in the A2aq haplotype.[30] We note that the Kaigani Haida people moved north to Prince of Wales Island in Southeast Alaska from Haida Gwaii in British Columbia in the early 18th century,[19,34] suggesting that either this haplotype is shared between the Haida and Tlingit peoples, or later obtained by the Haida through gene flow with the Tlingit. Moreover, the fact that both ~2,950 cal BP *TYYS* and the present-day *AK157* individual share a common mitochondrial haplotype, leads us to hypothesize matrilineal genetic continuity in Southeast Alaska for at least ~3,000 years. We caution that the small number of ancient genomes from Southeast Alaska prevents formal testing of this hypothesis. Nevertheless, it is worth noting that by 3,000 years ago, the cultures in this region appear to have been largely the same as those observed at the time of European contact.[19,42] It is, therefore, possible that *TYYS* was part of a population that was both culturally and biologically ancestral to the present-day Tlingit people.

The presence of the A2aq lineage in both the Tsimshian and the Tlingit territories over the last few thousand years raises the question as to where in the Pacific Northwest Coast this haplotype first appeared ~9 ka (the age estimate of its most recent common ancestor). One possible scenario is the initial appearance of this haplotype in the Tsimshian territory in British Columbia, followed by a northward migration to the Tlingit territory in Southeast Alaska sometime before 3,000 years ago (the age of *TYYS*). The fact that the oldest ancient sample belonging to this lineage comes from Tsimshian territory supports this scenario. Moreover, this scenario is also consistent with the Tlingit oral tradition that most Tlingit clans trace their origins to the Tsimshian coast, around the mouth of the Skeena River, from which they migrated northward.[30,34,43–45] Indeed, the original migration stories of both the Raven and the Eagle clans included this northward movement.[46] An alternative scenario to explain the presence of A2aq in both Tsimshian and Tlingit territories could be a movement of people southward from Tlingit to Tsimshian territory. Under this scenario, a southward movement must have occurred at least ~4,855 years ago, a limit imposed by the age of the oldest known ancient individual from Tsimshian territory with the A2aq haplotype (*160a*). Another possible explanation is that people (along with the A2aq haplotype) moved into the Tlingit and Tsimshian territories independently from a third location that already harbored the A2aq haplotype. It is

important to point out, however, that inference on the origin of this haplotype is limited by the very small sample size studied here, and any interpretations may change with increased data.

Tatóok yík yées sháawat shows the greatest genetic affinity to Northern Native Americans

To date, all pre-colonial ancient individuals from the Pacific Northwest Coast, including *Shuká Káa*, *939*, *443*, and *302*, spanning the last ~10 thousand years, have been reported to show the greatest nuclear genetic affinity to Northern Native Americans (NNA).[2,23,24] Therefore, we hypothesized that, like other Pacific Northwest Coast samples, the *TYYS* individual from Southeast Alaska would also show high genetic affinity to NNA. To test this hypothesis, we performed Principal Component Analysis (PCA) using phylogenetically instructive nuclear genomic SNPs from a panel of modern Eurasian and Indigenous American populations, including present-day Tlingit people, as well as ancient individuals from the Americas. Comparing PC1 against PC2 (Figure 3A), and more clearly, PC2 against PC3 (Figure 3B), we observed that *TYYS* indeed grouped among present-day Northern Native American populations, clustering closely with the three ancient individuals from British Columbia, as well as *The Ancient One* (also known as *Kennewick Man*), a ~8,500 cal year BP individual from Washington State.[47] Hence, our nuclear genomic data also show evidence for a close genetic affinity of this ancient Southeast Alaska individual with ancient and modern Indigenous individuals from North America. Although clustering among Northern Native American populations, *Shuká Káa* is a more distant relative to these other ancient Pacific Northwest individuals.

Tatóok yík yées sháawat is more closely related to Pacific Northwest coastal tribes than it is to inland tribes

A genetic divergence between present-day coastal and inland Pacific Northwest populations has been reported.[20] To test whether *TYYS* shows a greater affinity to one group relative to the other, we performed PCA using a panel of modern and ancient individuals from the Pacific Northwest (Figure 4). Our results show that *TYYS*, as well as other ancient coastal Pacific Northwest individuals, have a greater affinity to present-day coastal northern Pacific Northwest populations (Tlingit, Nisga'a, Haida, and Tsimshian) than to inland populations (Splatsin and Stswecem'c).

It is possible that the apparently higher affinity of *TYYS* to coastal versus inland Pacific Northwest populations is due to differential post-colonial gene flow into the two groups. Indeed, it has been shown that the inland groups have a greater amount of East Asian gene flow than coastal groups and vice-versa for European gene flow.[20] Compared to Central and South America, significant contact with Eurasians took place relatively recently in the Pacific Northwest, beginning with Russia's fur trade operation in the 1700s; later, Scandinavian and East Asian immigrants followed.[19] To explore the admixture of the Pacific Northwest populations, and to gauge whether *TYYS* shares an ancestry component with tribes of the Pacific Northwest, we performed model-based clustering of the individuals in our dataset using *ADMIXTURE*[48] (Figures 5 and S2). At K = 7, which was found to have the best predictive accuracy given its lowest cross-validation error, *TYYS*, *939*, *443*, and *302* are composed entirely of an ancestry component almost exclusively found among Northern Native Americans. This observation demonstrates that the genome of *TYYS*, as well as those of other ancient Holocene individuals, are indeed free of European admixture. Expectedly, present-day Pacific Northwest and Athabascan-speaking Indigenous Nations showed high levels of European gene flow. As such, any conclusion about the differential relatedness of *TYYS* to different subsets of these tribes may be confounded by this post-colonial gene flow. We, therefore, masked the data from modern Pacific Northwest individuals for European, East-Asian, and African admixture.

Using the masked data, we performed the outgroup f_3 statistic[49] to measure the shared genetic drift between *TYYS* and various Siberian and American populations, relative to the Han Chinese. In particular, we calculated outgroup f_3 of the form (X, *TYYS*; Han), where X is an American or Siberian population. With the masked data, we observed higher f_3 values when X is an American population than when it is a Siberian population, meaning *TYYS* is more closely related to Indigenous peoples of the Americas than Siberians. Among Indigenous peoples of the Americas, higher f_3 values were observed for the NNA populations compared to SNA populations, implying *TYYS*'s closer relationship to Native North Americans than South Americans. Further, among the NNA populations, higher values are observed for Pacific Northwest coastal populations than for Pacific Northwest inland populations (Figure 6A), confirming that *TYYS* is more closely related to Pacific Northwest coastal individuals than to inland individuals. That *TYYS* showed the greatest genomic affinity to coastal Pacific Northwest individuals was also observed with f_2 (Figure S3). Additionally, we calculated f_4 statistics[50] of the form (Han, *TYYS*; X, Y), where X and Y are North American

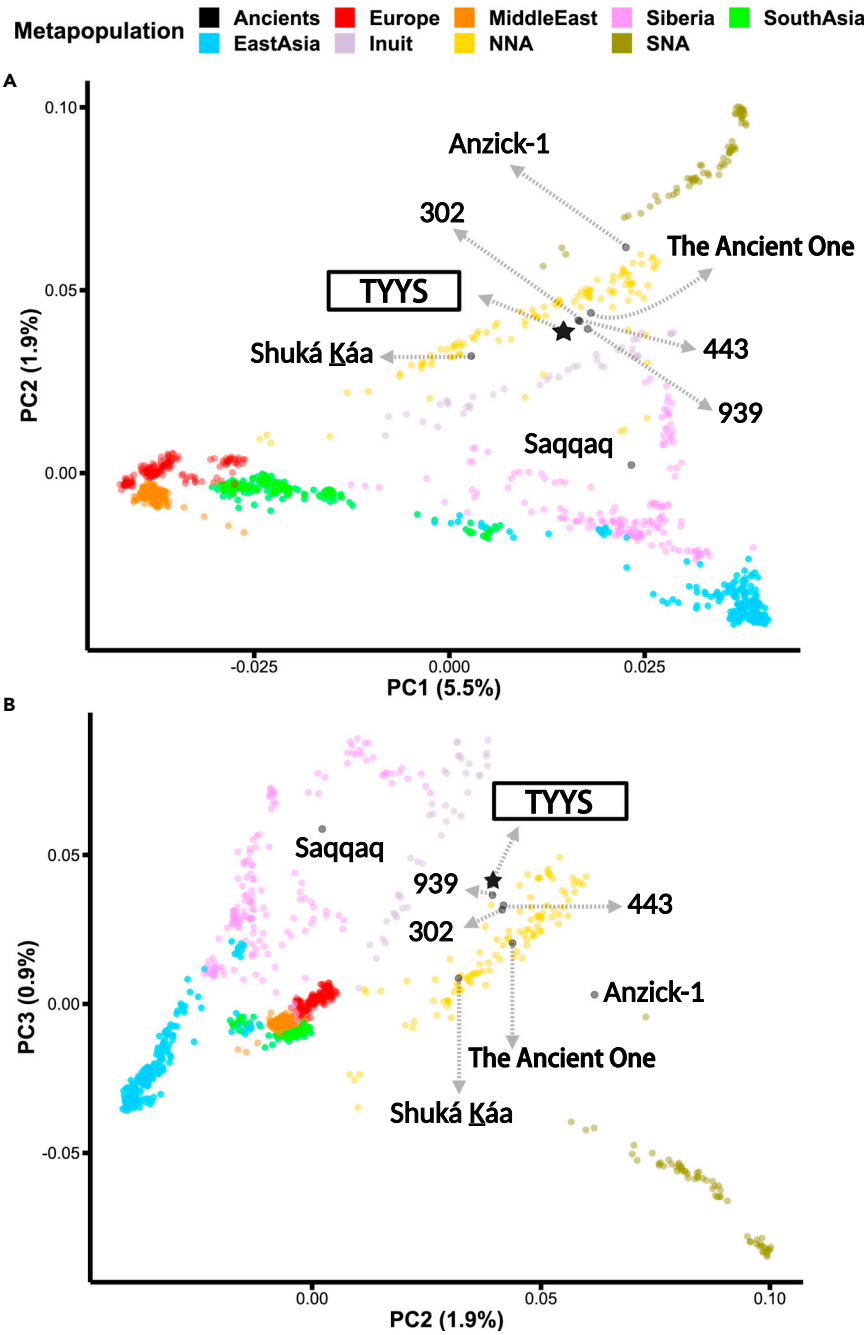


Figure 3. Principal Component Analysis (PCA) plots based on nuclear genome-scale data

PCA from a global panel of 1,107 Eurasian and Indigenous American individuals. NNA stands for Northern Native Americans and SNA refers to Southern Native Americans. Here, NNA are represented by both coastal and inland people from the northern Pacific Northwest. Ancient individuals from North America, including *Tatóok yík yées sháawat* (*TYYS*), group with present-day North American populations.

(A) Principal components PC1 vs. PC2.

(B) Principal components PC2 vs. PC3.

populations. We expect this to be positive if Y is more closely related to *TYYS* than X, and negative if X is more closely related to *TYYS* than Y. Based on f_4 results, we again observed that Pacific Northwest coastal tribes were more closely related to *TYYS* than were the inland tribes (Figure 6B). We observed similar results for ancient individuals *939*, *443*, and *302*, but not for *Shuká Káa*, *The Ancient One*, *Anzick-1*, and *Saqqaq*, a

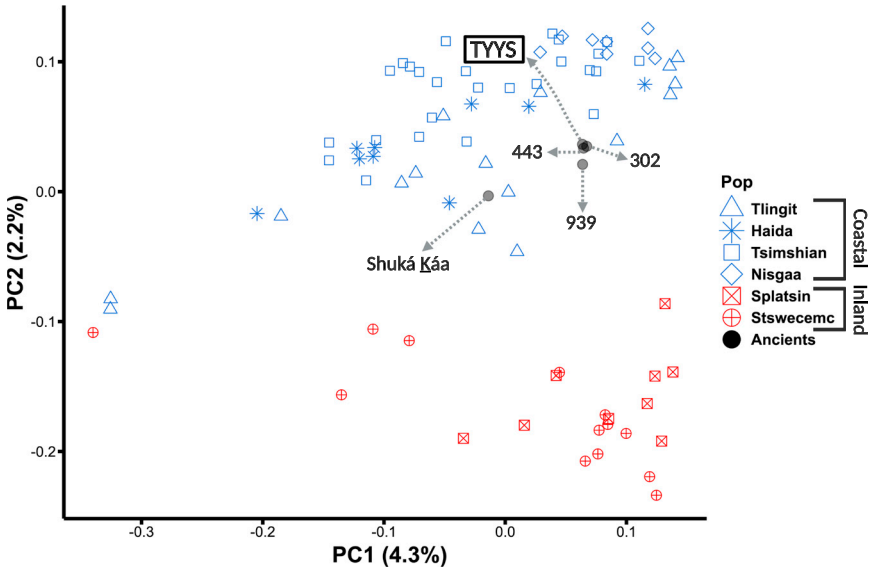


Figure 4. Principal Component Analysis (PCA) of individuals from the northern Pacific Northwest

PCA plot based on the nuclear genome-scale data from a panel of 86 ancient and modern individuals from the northern Pacific Northwest. Individuals belonging to coastal tribes are shown in blue; those belonging to inland tribes are shown in red. There appears to be a separation between inland and coastal peoples. Ancient individuals from the Pacific Northwest Coast, including *Tatóok yík yées sháawat* (*TYYS*), cluster with present-day coastal individuals, to the exclusion of the inland peoples.

4,000-year-old individual from Greenland[4] (Figures S4 and S5). Since the age of the *939* individual is ~6 ka, these results suggest that the divergence between coastal and inland PNW populations may have occurred prior to 6,000 years ago. Because the coastal art and aesthetic styles associated with the Northwest Coast began to appear around 5,000 years ago,[42,51] it seems reasonable that the creators of this art were the ancestors of the Pacific Northwest coastal tribes, but not of the inland tribes.

The greater relatedness of the Pacific Northwest ancient individuals to coastal versus inland tribes is likely not a result of higher residual (post-masking) admixture in inland tribes than in coastal tribes. Firstly, this is because our results are robust to the choice of outgroup. Secondly, if the apparent closer relatedness of Pacific Northwest ancient individuals to coastal groups was a consequence of differential residual admixture in coastal versus inland groups, we would have seen similar results for *The Ancient One*, which we do not. It should be noted that although *Shuká Káa* shows the greatest genetic affinity to Northern Native Americans, and possibly even coastal groups (Figure 4), given the very low coverage of his genome, we are unable to definitively ascertain that he is more closely related to the coastal Pacific Northwest peoples than he is to the inland peoples of the Pacific Northwest (Figures S4G–S4H).

Our nuclear genome-scale results are consistent with our inference of at least 3,000 years of matrilineal genetic continuity in Southeast Alaska from mitochondrial DNA analysis. That ancient individuals from the Pacific Northwest Coast dated to 6,000 years BP are most closely related to modern individuals from the coast to the exclusion of inland groups is also consistent with the Tlingit contention that their ancestors have been the custodians of Southeast Alaska since "time immemorial".[46] It is indeed plausible, although it remains unresolved with the available data, that the divergence between coastal and interior groups significantly predates 6,000 years BP. Even though it cannot be determined with the available data, the placement of *Shuká Káa* among the coastal individuals in PCA (Figure 4) could indicate that the divergence dates back even further. It is, therefore, likely that the emergence of the mitochondrial haplotype A2aq, which we dated to ~9,000 years BP, occurred *in situ* in the northern Pacific Northwest coast.

Gene flow between the Paleo-Inuit and Northern Native Americans

Although it has been proposed that the Paleo-Inuit contributed ancestry to present-day PNW populations, it remains a contentious issue.[12,21] The general ancestry group that is represented by the Paleo-Inuit and

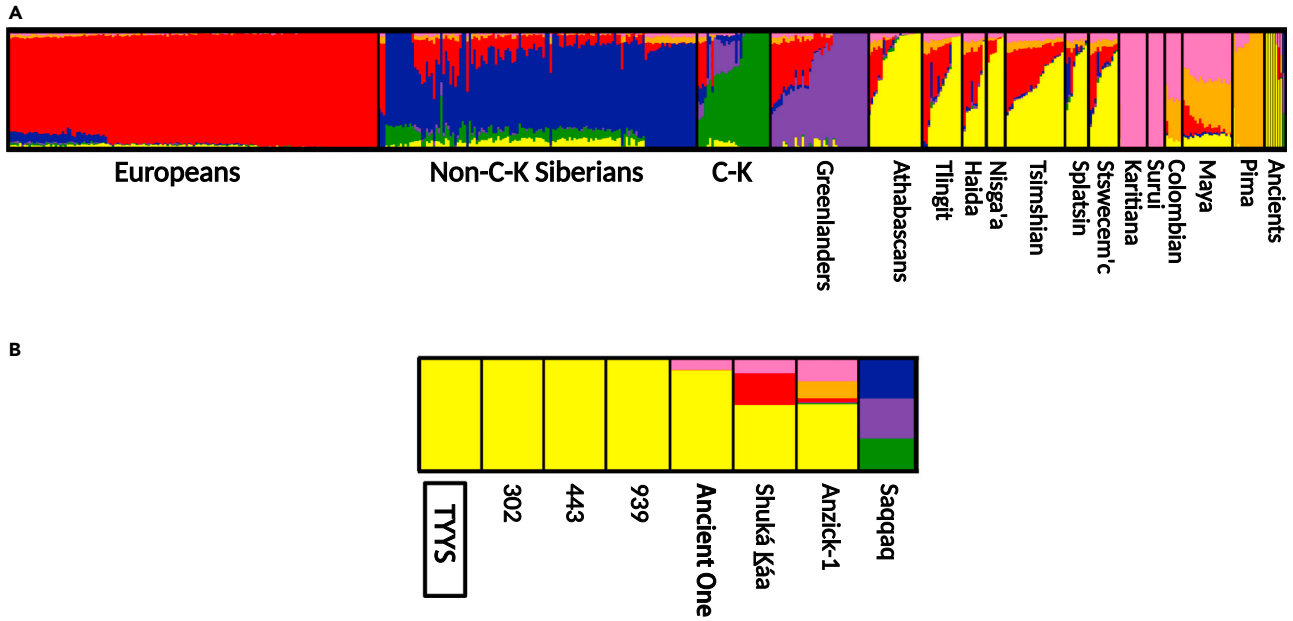


Figure 5. ADMIXTURE clustering analysis

Cluster analysis using ADMIXTURE for a set of European, Siberian, American, and Greenlandic populations, and a set of relevant ancient individuals. K = 7 clusters are displayed here because it was found to have the best predictive accuracy given its lowest cross-validation error.

(A) Results for all individuals in the panel. C-K refers to speakers of Chukotka-Kamchatkan languages (represented here by the Koryak and Chukchi peoples).

(B) A zoomed-in view of the results for ancient individuals in the panel (labeled as "Ancients" in A.). We note that *Tatóok yík yées sháawat* (*TYYS*), like 302, 443, and 939, is composed entirely of the ancestry component that is observed at high levels only in the Indigenous North American populations (in yellow color), while the present-day Pacific Northwest populations are admixed with Europeans (red).

the speakers of the Chukotka-Kamchatkan languages (the Koryak and Chukchi peoples) in Siberia has been termed "Proto-Paleo-Eskimo" (PPE).[21] We sought to investigate 1) whether the present-day PNW populations harbor PPE ancestry, and 2) whether the ancient PNW individuals harbor PPE ancestry.

In order to answer the first question, we used *qpAdm* modeling,[52,53] seeking to model present-day PNW populations as a combination of representatives of SNA/NNA and PPE (represented by *Saqqaq* and the speakers of Chukotka-Kamchatkan languages) ancestries. Except for certain cases where *Anzick-1* is one source, all other models were rejected, and consequently, we found no evidence of PPE admixture into present-day PNW populations, at least based on our dataset (Tables S3 and S4). We note that it was recently suggested that Saqqaq may not serve as the best proxy for gene flow into the NNA because the PPE source that contributed ancestry to some North American groups may have been more closely related to the Koryak than to the Paleo-Inuit[88]; however, our results are consistent when the Koryak and Chukchi peoples are used a source.

To test whether *TYYS*, along with other ancient individuals from the Americas (*443*, *302*, *939*, *The Ancient One*, and *Anzick-1*), show signatures of the Paleo-Inuit ancestry, we first calculated D-statistics[50] of the form D(Han, *Saqqaq*; X, Y), where X is an unadmixed SNA population and Y is an ancient individual (Figure 7A). The results indicate that ancient individuals (particularly those from the NNA, including *The Ancient One* from 8,500 ka) show closer affinity to the Paleo-Inuit, represented by *Saqqaq*,[4] than do present-day SNA populations. To confirm that this genetic affinity stems from PPE gene flow into ancient individuals, we again used *qpAdm* to model each ancient genome as a combination of representatives of SNA/NNA, and PPE. Such models were rejected indicating that these ancient individuals do not harbor PPE ancestry (Table S5). Instead, when *Saqqaq* was modeled as a combination of PPE (represented by speakers of Chukotka-Kamchatkan languages) and SNA/NNA ancestries, we find that *Saqqaq* can be modeled as a combination of PPE and NNA ancestries, but not as a combination of PPE and SNA ancestries (Table S6). Therefore, our results suggest that PPE (represented by both *Saqqaq* and speakers of Chukotka-Kamchatkan languages) may not have contributed ancestry to present-day or ancient PNW peoples, but instead, the *Saqqaq* genome harbors both PPE and NNA ancestries (Figure 7B). It is important to note, however, that results from different studies are inconsistent with one another, and more work is warranted on this issue.

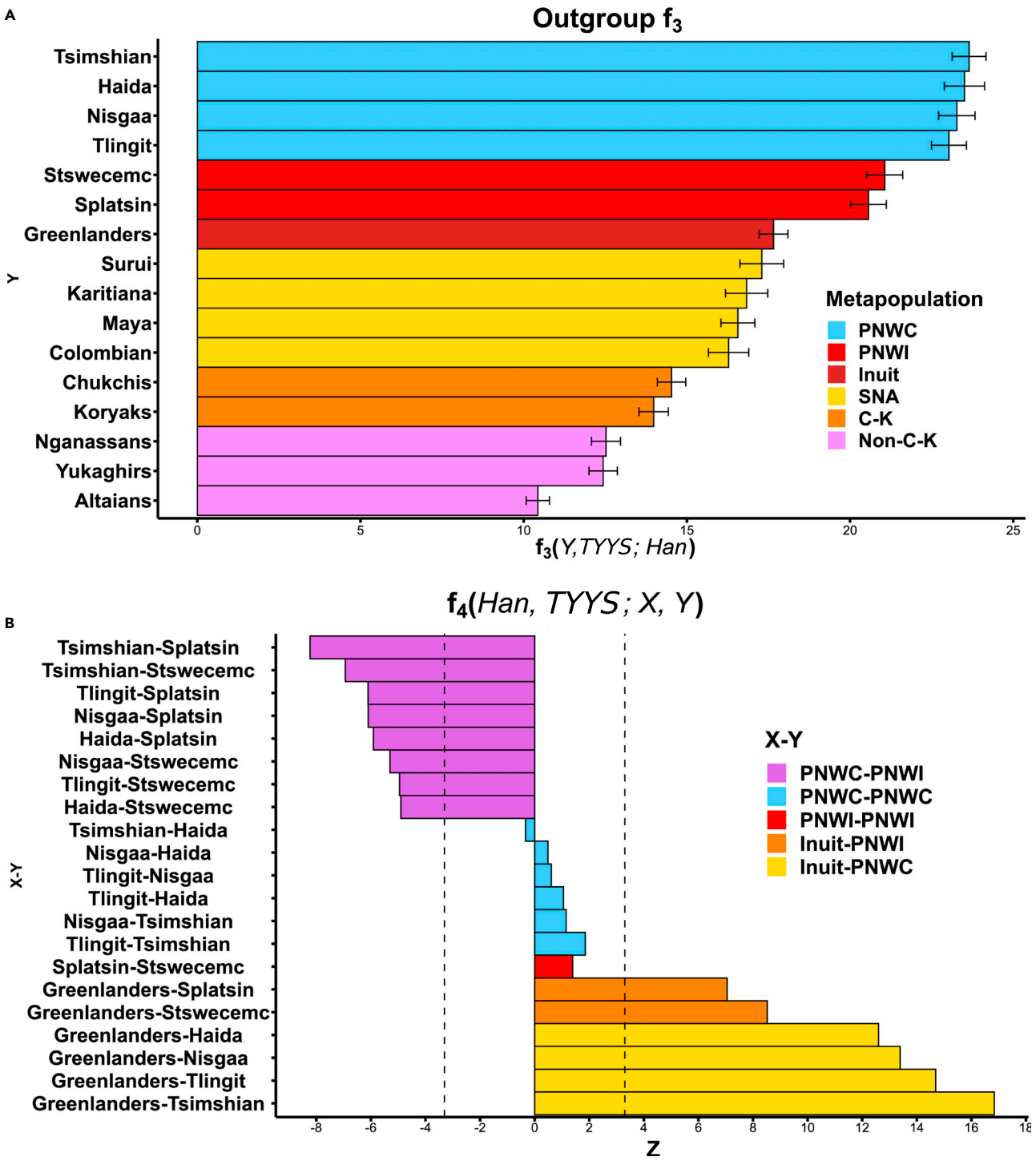


Figure 6. Results from f_3 and f_4 analyses

PNWC and PNWI refer to Pacific Northwest coastal and inland populations, respectively. C-K refers to speakers of the Chukotko-Kamchatkan languages. Non-C-K are Siberian populations that are not C-K speakers. *TYYS* refers to *Tatóok yík yées sháawat*. The error bars represent standard error.

(A) Outgroup f_3 statistics. Note that outgroup f_3 has very large values here because ADMIXTOOLS arbitrarily sets the f_3 denominator to 0.001.

(B) f_4 statistics. The magnitude along the x axis represents the Z-value associated with the f_4 statistic. A negative value along the x axis denotes that the associated f_4 value is negative; a positive value along the x axis denotes that the f_4 value is positive. The vertical dotted lines correspond to the Bonferroni-corrected p value threshold of 0.00048 (0.01/21). Results from both A and B indicate that *TYYS* shows the closest relationship to the Pacific Northwest coastal peoples (to the exclusion of the inland peoples).

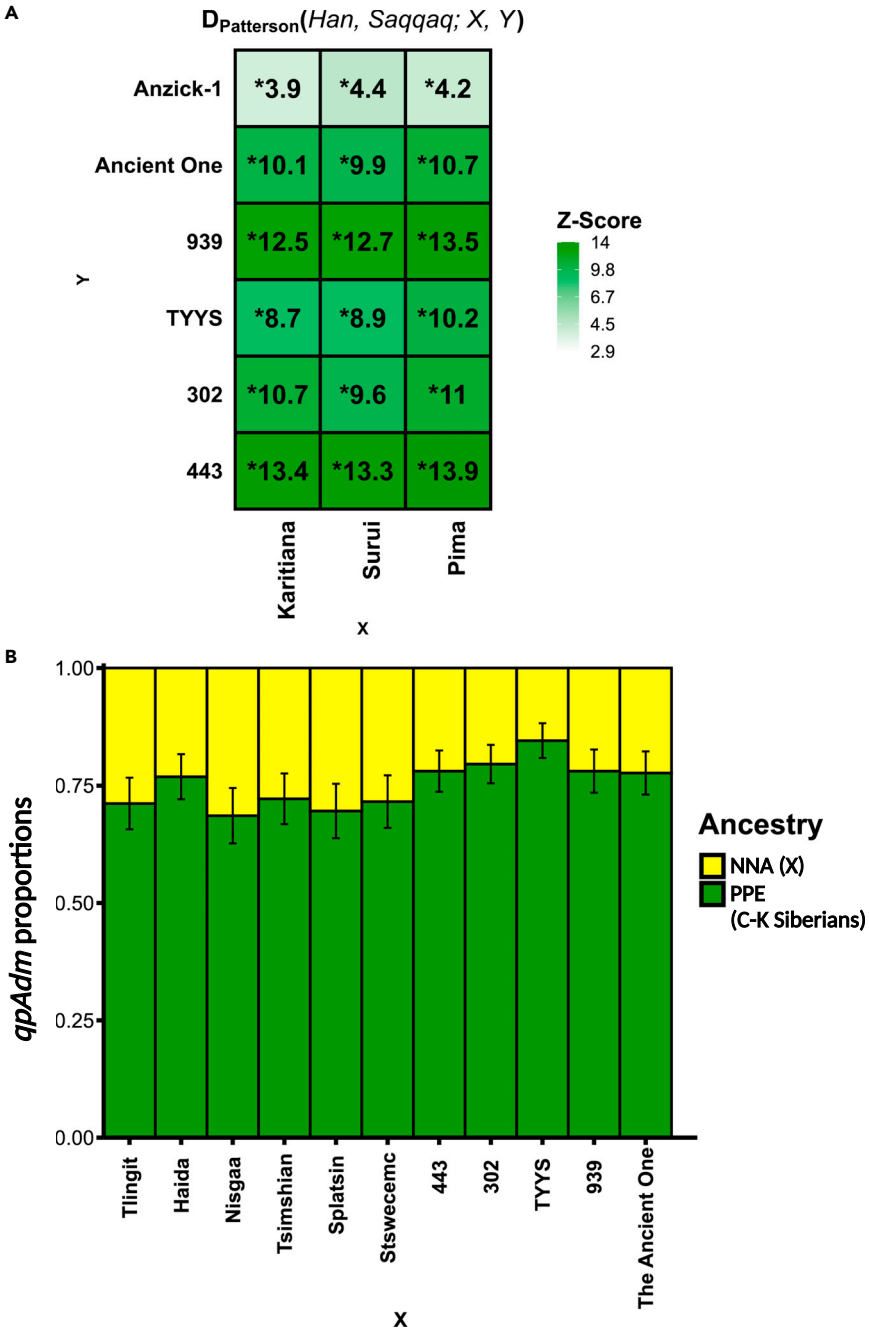


Figure 7. D-statistics and *qpAdm* analyses of admixture

(A) Patterson's D of the form (Han, Saqqaq; X, Y); X is one of the unadmixed SNA populations listed along the x axis; Y is one of the ancient individuals listed along the Y axis, Bonferroni-corrected p value threshold is 0.00056 (0.01/18). Each box in the figure states the Z-score corresponding to the relevant D-statistic. Darker colors correspond to higher Z-scores. Boxes containing statistically significant results are marked with an asterisk. *TYYS* refers to *Tatóok yík yées sháawat*.

(B) *qpAdm* analysis. The *Saqqaq* genome is modeled as a combination of the "Proto-Paleo-Eskimo" (PPE) ancestry represented by C-K Siberians (speakers of Chukotka-Kamchatkan languages), and various representatives of Northern Native Americans (NNA) listed along the x axis. The error bars represent standard error.

Conclusions

In this study, we sequenced the mitochondrial and nuclear genomes from a ~2,950-year-old ancient female, *TYYS*, uncovered from a cave in present-day Tlingit territory in Southeast Alaska. Based on *TYYS*'s mitochondrial DNA, we showed that this individual is most closely related to present-day individuals from the Tlingit, Haida, Nisga'a, and Tsimshian territories along the northern Pacific Northwest Coast, as well as to a ~4,850-year-old individual from the Tsimshian territory. We note that this is consistent with the Tlingit oral tradition that the Tlingit trace their roots from the present-day Tsimshian territory.[30,34,43–46] We showed that the A2aq maternal lineage has persisted in the Tlingit territory for at least ~3,000 years and that this haplotype emerged ~9,000 years ago. These results are also consistent with another Tlingit oral tradition that places them in the region during a volcanic activity at Mt. Edgecumbe,[54] which last erupted ~4,500 years ago.[55] Based on *TYYS*'s nuclear genome, this individual is more closely related to Pacific Northwest coastal individuals than to inland peoples. We find that the split between the coastal and inland peoples of the northern Pacific Northwest took place before ~6,000 years ago. Since a Pacific Northwest Coast culture very similar to that at the time of European contact had become established by 3,000 years BP, it seems likely that *TYYS* was part of a population that was both culturally and genetically ancestral to the present-day Tlingit people. The stable isotope $\delta^{13}C$ value of −12.5‰ from *TYYS* and a similar value from the ~6,075-year-old *939* individual fall within the range of stable carbon isotope data from Thule Culture individuals in southwestern Greenland and marine mammals,[26] indicating a diet almost entirely consisting of marine protein, further lending support to a presence of a northwest coast culture at this time. Lastly, we found that neither *TYYS* nor other ancient genomes used in this study harbor Proto-Paleo-Eskimo (PPE) ancestry. More surprisingly, we found that the *Saqqaq* genome harbors NNA ancestry. Our results shed light on the history of the Indigenous peoples from the northern Pacific Northwest Coast, and the possible pre-European contact of ancient individuals in this region with the Paleo-Inuit. We expect future paleogenomic analyses may offer more complete insights into the population history of the highly diverse Pacific Northwest, including the hypothesis of the Pacific Northwest Coast as a gateway for the initial peopling of the Americas.

Limitations of the study

Our study is limited by the fact that we have access to DNA from only one ancient individual from Southeast Alaska. Most analyses performed in this study rely on only a handful of previously published ancient genomes from the Pacific Northwest Coast, along with TYYS. Therefore, conclusions about the lack of Paleo-Inuit ancestry in ancient individuals from the Pacific Northwest Coast are not definitive and warrant further study. Additionally, because ancient individuals from the Pacific Northwest Coast at various time points are unavailable, claims of matrilineal genetic continuity in the PNW with regard to the haplotype A2aq relies on the assumption that the haplotype was not lost and reintroduced sometime after 3,000 years ago.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Archaeological context and sex determination
  - Community engagement
  - DNA extraction and sequencing
  - Trimming and mapping
  - Mitochondrial genome analysis
  - Variant calling
  - Dataset for nuclear DNA analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - PCA
  - ADMIXTURE

- Masking for alleles found in Europeans
- *f*- and D-statistics
- Conceptual concerns in testing for admixture between PPE and first people

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106581.

ACKNOWLEDGMENTS

The authors are particularly grateful to the Wrangell Cooperative Association for their support of this project. The authors also thank Tongass National Forest archaeologists Gina Esposito and Tiffany Brunson, soil scientist Jackie de-Montigny, University of South Dakota student field assistants Frank Andy Klock, Nathan Carter, Brandon Silver, Louis Rezac, Clarissa Ford, Alex Santos and Christy Heaton. The authors are thankful to Flavio da Silva Coelho for help with radiocarbon dating calibrations, Victor A. Albert for valuable comments on the manuscript, and Dilek Koptekin for insightful discussions during the review process. This research was supported with funding from the National Science Foundation (grant nos. 1556565, 1854550, 9870343, and 0208247) and the National Geographic Society (grant no. 6212-98).

AUTHOR CONTRIBUTIONS

C.L. designed the study; S.G. extracted the DNA and performed initial analyses; A.A. and C.L. analyzed the data; R.S.M. provided data; E.A.R. organized collaborative opportunities between the Wrangell Cooperative Association, the USFS and C.L. for culturally-appropriate review of data, tribal examination of the site and naming of *Tatóok yík yées shàawat*; T.H.H. and J.S. performed the cave explorations and provided the sample and paleontological context; A.A., O.G., and C.L. wrote the manuscript with contributions from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: October 10, 2022
Revised: February 27, 2023
Accepted: April 3, 2023
Published: April 8, 2023

SUPPORTING CITATIONS

The following reference appears in the supplemental information: Meyer et al., 2012.[89]

REFERENCES

1. Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.V., Rojas, W., Duque, C., Mesa, N., et al. (2012). Reconstructing Native American population history. Nature *488*, 370–374. https://doi.org/10.1038/nature11258.

2. Raghavan, M., Steinrücken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., Albrechtsen, A., Valdiosera, C., Ávila-Arcos, M.C., Malaspinas, A.-S., et al. (2015). POPULATION GENETICS. Genomic evidence for the Pleistocene and recent population history of Native Americans. Science *349*, aab3884. https://doi.org/10.1126/science.aab3884.

3. Raghavan, M., DeGiorgio, M., Albrechtsen, A., Moltke, I., Skoglund, P., Korneliussen, T.S., Grønnow, B., Appelt, M., Gulløv, H.C., Friesen, T.M., et al. (2014). The genetic prehistory of the new world arctic. Science *345*, 1255832. https://doi.org/10.1126/science.1255832.

4. Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., et al. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature *463*, 757–762. https://doi.org/10.1038/nature08835.

5. Sikora, M., Pitulko, V.V., Sousa, V.C., Allentoft, M.E., Vinner, L., Rasmussen, S., Margaryan, A., de Barros Damgaard, P., de la Fuente, C., Renaud, G., et al. (2019). The population history of northeastern Siberia since the Pleistocene. Nature *570*, 182–188. https://doi.org/10.1038/s41586-019-1279-z.

6. Yu, H., Spyrou, M.A., Karapetian, M., Shnaider, S., Radzevičiūtė, R., Nägele, K., Neumann, G.U., Penske, S., Zech, J., Lucas, M., et al. (2020). Paleolithic to bronze age Siberians reveal connections with first Americans and across eurasia. Cell *181*, 1232–1245.e20. https://doi.org/10.1016/j.cell.2020.04.037.

7. Moreno-Mayar, J.V., Potter, B.A., Vinner, L., Steinrücken, M., Rasmussen, S., Terhorst, J., Kamm, J.A., Albrechtsen, A., Malaspinas, A.-S., Sikora, M., et al. (2018). Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans.

Nature *553*, 203–207. https://doi.org/10.1038/nature25173.

8. 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. https://doi.org/10.1371/journal.pone.0000829.

9. Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S.M., Rohrlach, A., et al. (2016). Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci. Adv. *2*, e1501385. https://doi.org/10.1126/sciadv.1501385.

10. Jakobsson, M., Pearce, C., Cronin, T.M., Backman, J., Anderson, L.G., Barrientos, N., Björk, G., Coxall, H., de Boer, A., Mayer, L.A., et al. (2017). Post-glacial flooding of the Bering Land Bridge dated to 11 cal ka BP based on new geophysical and sediment records. Clim. Past *13*, 991–1005. https://doi.org/10.5194/cp-13-991-2017.

11. Elias, S.A., Short, S.K., Nelson, C.H., and Birks, H.H. (1996). Life and times of the Bering land bridge. Nature *382*, 60–63. https://doi.org/10.1038/382060a0.

12. Willerslev, E., and Meltzer, D.J. (2021). Peopling of the Americas as inferred from ancient genomics. Nature *594*, 356–364. https://doi.org/10.1038/s41586-021-03499-y.

13. Raff, J. (2022). Origin: A Genetic History of the Americas (Grand Central Publishing).

14. Moreno-Mayar, J.V., Vinner, L., de Barros Damgaard, P., de la Fuente, C., Chan, J., Spence, J.P., Allentoft, M.E., Vimala, T., Racimo, F., Pinotti, T., et al. (2018). Early human dispersals within the Americas. Science *362*, eaav2621. https://doi.org/10.1126/science.aav2621.

15. Lesnek, A.J., Briner, J.P., Lindqvist, C., Baichtal, J.F., and Heaton, T.H. (2018). Deglaciation of the Pacific coastal corridor directly preceded the human colonization of the Americas. Sci. Adv. *4*, eaar5040. https://doi.org/10.1126/sciadv.aar5040.

16. Walcott, C.K., Lesnek, A.J., Baichtal, J.F., Briner, J.P., Licciardi, J.M., Lesnek, A.J., and Licciardi, J.M. (2022). Cosmogenic ages indicate No MIS 2 refugia in the alexander Archipelago, Alaska. Geochronology *4*, 191–211. https://doi.org/10.5194/gchron-4-191-2022.

17. Heintzman, P.D., Froese, D., Ives, J.W., Soares, A.E.R., Zazula, G.D., Letts, B., Andrews, T.D., Driver, J.C., Hall, E., Hare, P.G., et al. (2016). Bison phylogeography constrains dispersal and viability of the Ice Free Corridor in western Canada. Proc. Natl. Acad. Sci. USA *113*, 8057–8063. https://doi.org/10.1073/pnas.1601077113.

18. Pedersen, M.W., Ruter, A., Schweger, C., Friebe, H., Staff, R.A., Kjeldsen, K.K., Mendoza, M.L.Z., Beaudoin, A.B., Zutter, C., Larsen, N.K., et al. (2016). Postglacial viability and colonization in North America's ice-free corridor. Nature *537*, 45–49. https://doi.org/10.1038/nature19085.

19. Ames, K.M., and Maschner, H.D.G. (2000). Peoples of the Northwest Coast: Their Archaeology and Prehistory (Thames and Hudson)).

20. Verdu, P., Pemberton, T.J., Laurent, R., Kemp, B.M., Gonzalez-Oliver, A., Gorodezky, C., Hughes, C.E., Shattuck, M.R., Petzelt, B., Mitchell, J., et al. (2014). Patterns of admixture and population structure in native populations of Northwest North America. PLoS Genet. *10*, e1004530. https://doi.org/10.1371/journal.pgen.1004530.

21. Flegontov, P., Altınışık, N.E., Changmai, P., Rohland, N., Mallick, S., Adamski, N., Bolnick, D.A., Broomandkhoshbacht, N., Candilio, F., Culleton, B.J., et al. (2019). Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America. Nature *570*, 236–240. https://doi.org/10.1038/s41586-019-1251-y.

22. Lindo, J., Huerta-Sánchez, E., Nakagome, S., Rasmussen, M., Petzelt, B., Mitchell, J., Cybulski, J.S., Willerslev, E., DeGiorgio, M., and Malhi, R.S. (2016). A time transect of exomes from a Native American population before and after European contact. Nat. Commun. *7*, 1–11. https://doi.org/10.1038/ncomms13175.

23. Lindo, J., Achilli, A., Perego, U.A., Archer, D., Valdiosera, C., Petzelt, B., Mitchell, J., Worl, R., Dixon, E.J., Fifield, T.E., et al. (2017). Ancient individuals from the North American Northwest Coast reveal 10,000 years of regional genetic continuity. Proc. Natl. Acad. Sci. USA *114*, 4093–4098. https://doi.org/10.1073/pnas.1620410114.

24. Cui, Y., Lindo, J., Hughes, C.E., Johnson, J.W., Hernandez, A.G., Kemp, B.M., Ma, J., Cunningham, R., Petzelt, B., Mitchell, J., 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. https://doi.org/10.1371/journal.pone.0066948.

25. Reich, D. (2018). Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past (Oxford University Press).

26. Erle Nelson, D., Lynnerup, N., and Arneborg, J. (2012). A First Isotopic Dietary Study of the Greenlandic Thule Culture. Noat, pp. 51–64. https://doi.org/10.3721/037.004.s305.

27. Perego, U.A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., Hooshiar Kashani, B., Ritchie, K.H., Scozzari, R., Kong, Q.-P., et al. (2009). Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Curr. Biol. *19*, 1–8. https://doi.org/10.1016/j.cub.2008.11.058.

28. 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. https://doi.org/10.1046/j.1469-1809.2003.00049.x.

29. Roca-Rada, X., Politis, G., Messineo, P.G., Scheifler, N., Scabuzzo, C., González, M., Harkins, K.M., Reich, D., Souilmi, Y., Teixeira, J.C., et al. (2021). Ancient mitochondrial genomes from the argentinian pampas inform the early peopling of the southern cone of South America. iScience *24*, 102553. https://doi.org/10.1016/j.isci.2021.102553.

30. De Los Santos, R., Monroe, C., Worl, R., Worl, R., Schroeder, K.B., and Kemp, B.M. (2019). Genetic diversity and relationships of Tlingit moieties. Hum. Biol. *91*, 95–116. https://doi.org/10.13110/humanbiology.91.2.03.

31. Schurr, T.G., Dulik, M.C., Owings, A.C., Zhadanov, S.I., Gaieski, J.B., Vilar, M.G., Ramos, J., Moss, M.B., and Natkong, F.; 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. https://doi.org/10.1002/ajpa.22068.

32. Monsalve, M.V., Stone, A.C., Lewis, C.M., Rempel, A., Richards, M., Straathof, D., and Devine, D.V. (2002). Brief communication: molecular analysis of the Kwäday Dän Ts'finchi ancient remains found in a glacier in Canada. Am. J. Phys. Anthropol. *119*, 288–291. https://doi.org/10.1002/ajpa.10116.

33. Kemp, B.M., Malhi, R.S., McDonough, J., Bolnick, D.A., Eshleman, J.A., Rickards, O., Martinez-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. https://doi.org/10.1002/ajpa.20543.

34. Emmons, G.T., and De Laguna, F. (1991). The Tlingit Indians. Anthropological Papers of the American Museum of Natural History No. 70 (University of Washington Press; New York: American Museum of Natural History).

35. Kaplan, S. (1986). Raven's journey: the world of Alaska's Native people. Archaeology *39*, 54–57.

36. de Laguna, F. (1990). Tlingit. In Handbook of North American Indians, *Vol. 7*, W. Suttles, ed. (Smithsonian Institution Press), pp. 203–239. *Northwest Coast*.

37. Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Jr., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., et al. (2014). The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature *506*, 225–229. https://doi.org/10.1038/nature13025.

38. Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior summarization in bayesian phylogenetics using tracer 1.7. Syst. Biol. *67*, 901–904. https://doi.org/10.1093/sysbio/syy032.

39. van Oven, M. (2015). PhyloTree Build 17: growing the human mitochondrial DNA tree. Forensic Sci. Int.: Genetics Supplement

Series *5*, e392–e394. https://doi.org/10.1016/j.fsigss.2015.09.155.

40. Taboada-Echalar, P., Alvarez-Iglesias, V., Heinz, T., Vidal-Bralo, L., Gómez-Carballa, A., Catelli, L., Pardo-Seco, J., Pastoriza, A., Carracedo, A., Torres-Balanza, A., et al. (2013). The genetic legacy of the pre-colonial period in contemporary Bolivians. PLoS One *8*, e58980. https://doi.org/10.1371/journal.pone.0058980.

41. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. *44*, W58–W63. https://doi.org/10.1093/nar/gkw233.

42. Davis, S.D. (1990). Prehistory of southeastern Alaska. In Handbook of North American Indians, *Vol. 7*, W. Suttles, ed. (Smithsonian Institution Press), pp. 197–202. *Northwest Coast*.

43. De Laguna, F. (1960). The Story of a Tlingit Community: A Problem in the Relationship between Archeological, Ethnological, and Historical Methods (Bureau of American Ethnology Bulletin).

44. Worl, R.F. (1998). Tlingit at.Oow: Tangible and Intangible Property.

45. Olson, R.L. (1967). Social Structure and Social Life of the Tlingit in Alaska (University of California Press).

46. Worl, R.K. (2019). The Native American Graves Protection and Repatriation Act: Integrating Science and Stories: Tlingit Ancient History. Society for Applied Anthropology (Sealaska Heritage Institute).

47. Rasmussen, M., Sikora, M., Albrechtsen, A., Korneliussen, T.S., Moreno-Mayar, J.V., Poznik, G.D., Zollikofer, C.P.E., de León, M.P., Allentoft, M.E., Moltke, I., et al. (2015). The ancestry and affiliations of Kennewick Man. Nature *523*, 455–458. https://doi.org/10.1038/nature14625.

48. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109.

49. Raghavan, M., Skoglund, P., Graf, K.E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford, T.W., Jr., Orlando, L., Metspalu, E., et al. (2014). Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature *505*, 87–91. https://doi.org/10.1038/nature12736.

50. Reich, D., Thangaraj, K., Patterson, N., Price, A.L., and Singh, L. (2009). Reconstructing Indian population history. Nature *461*, 489–494. https://doi.org/10.1038/nature08365.

51. Ackerman, R.E. (1992). Earliest stone industries on the north pacific coast of NorthNorth America. Arctic Anthropol. *29*, 18–27.

52. Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature *522*, 207–211. https://doi.org/10.1038/nature14317.

53. Harney, É., Patterson, N., Reich, D., and Wakeley, J. (2021). Assessing the performance of qpadm: a statistical tool for studying population admixture. Genetics *217*. https://doi.org/10.1093/genetics/iyaa045.

54. Thornton, T.F., and Hope, F. (1998). Traditional Tlingit Use of Sitka National Historical Park (US National Park Service).

55. Riehle, J.R., Brew, D.A., Reed, K.M., and Bartsch-Winkler, S. (1984). Explosive latest pleistocene (?) and holocene activity of the mount edgecumbe volcanic field, Alaska. US Geol. Surv. Circular *939*, 111–114.

56. Heaton, T.H., and Grady, F. (2003). The late Wisconsin vertebrate history of Prince of Wales Island, southeast Alaska. Ice age cave faunas of North America *2*, 17.

57. Heaton, T.H. (2002). Ice Age Paleontology of Southeast Alaska (University of South Dakota). http://orgs. usd. edu/esci/alaska/index. html.

58. da Silva Coelho, F.A., Gill, S., Tomlin, C.M., Heaton, T.H., and Lindqvist, C. (2021). An early dog from southeast Alaska supports a coastal route for the first dog migration into the Americas. Proc. Biol. Sci. *288*, 20203103. https://doi.org/10.1098/rspb.2020.3103.

59. Heaton, T.J., Köhler, P., Butzin, M., Bard, E., Reimer, R.W., Austin, W.E.N., Bronk Ramsey, C., Grootes, P.M., Hughen, K.A., Kromer, B., et al. (2020). Marine20—the marine radiocarbon age calibration curve (0–55,000 cal BP). Radiocarbon *62*, 779–820. https://doi.org/10.1017/RDC.2020.68.

60. Stuiver, M., Reimer, P.J., and Reimer, R.W. (2021). CALIB 8.2 [WWW program]. http://calib. org.

61. Schmuck, N., Reuther, J., Baichtal, J.F., and Carlson, R.J. (2021). Quantifying marine reservoir effect variability along the Northwest Coast of North America. Quat. Res. *103*, 160–181. https://doi.org/10.1017/qua.2020.131.

62. Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. *40*, 4477–4482. https://doi.org/10.1016/j.jas.2013.07.004.

63. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763. https://doi.org/10.1073/pnas.1314445110.

64. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88. https://doi.org/10.1186/s13104-016-1900-2.

65. Li, H., and Durbin, R. (2009). Making the leap: maq to BWA. Mass Genomics *25*, 1754–1760.

66. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). Genome Project data processing subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

67. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv. https://doi.org/10.48550/arXiv.1303.3997.

68. Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. *16*, 224. https://doi.org/10.1186/s13059-015-0776-0.

69. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

70. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193.

71. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

72. Katoh, K., Rozewicki, J., and Yamada, K.D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. *20*, 1160–1166. https://doi.org/10.1093/bib/bbx108.

73. Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. *4*, vey016. https://doi.org/10.1093/ve/vey016.

74. Chan, E.K.F., Timmermann, A., Baldi, B.F., Moore, A.E., Lyons, R.J., Lee, S.-S., Kalsbeek, A.M.F., Petersen, D.C., Rautenbach, H., Förtsch, H.E.A., et al. (2019). Human origins in a southern African palaeo-wetland and first migrations. Nature *575*, 185–189. https://doi.org/10.1038/s41586-019-1714-1.

75. Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V., and Richards, M.B. (2009). Correcting for purifying selection: an improved human mitochondrial molecular clock. Am. J. Hum. Genet. *84*, 740–759. https://doi.org/10.1016/j.ajhg.2009.05.001.

76. Letunic, I., and Bork, P. (2021). Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. *49*, W293–W296. https://doi.org/10.1093/nar/gkab301.

77. Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D., et al. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at bioRxiv. https://www.bioRxiv.org/content/10.1101/201178v2.

78. Bergström, A., McCarthy, S.A., Hui, R., Almarri, M.A., Ayub, Q., Danecek, P., Chen, Y., Felkel, S., Hallast, P., Kamm, J., et al. (2020). Insights into human genetic variation and population history from 929 diverse genomes. Science *367*, eaay5012. https://doi.org/10.1126/science.aay5012.

79. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093. https://doi.org/10.1534/genetics.112.145037.

80. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909. https://doi.org/10.1038/ng1847.

81. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190. https://doi.org/10.1371/journal.pgen.0020190.

82. Behr, A.A., Liu, K.Z., Liu-Fang, G., Nakka, P., and Ramachandran, S. (2016). pong: fast analysis and visualization of latent clusters in population genetic data. Bioinformatics *32*, 2817–2823. https://doi.org/10.1093/bioinformatics/btw327.

83. Delaneau, O., Marchini, J., and Zagury, J.-F. (2011). A linear complexity phasing method for thousands of genomes. Nat. Methods *9*, 179–181. https://doi.org/10.1038/nmeth.1785.

84. Maples, B.K., Gravel, S., Kenny, E.E., and Bustamante, C.D. (2013). RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am. J. Hum. Genet. *93*, 278–288. https://doi.org/10.1016/j.ajhg.2013.06.020.

85. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.

86. Maier, R., Flegontov, P., Flegontova, O., Changmai, P., and Reich, D. (2022). On the limits of fitting complex models of population history to genetic data. Preprint at bioRxiv. https://doi.org/10.1101/2022.05.08.491072.

87. Koptekin, D., Yüncü, E., Rodríguez-Varela, R., Altınışık, N.E., Psonis, N., Kashuba, N., Yorulmaz, S., George, R., Kazancı, D.D., Kaptan, D., et al. (2023). Spatial and temporal heterogeneity in human mobility patterns in holocene southwest Asia and the east mediterranean. Curr. Biol. *33*, 41–57.e15. https://doi.org/10.1016/j.cub.2022.11.034.

88. Nielsen, S.V., Vaughn, A.H., Leppälä, K., Landis, M.J., Mailund, T., and Nielsen, R. (2023). Bayesian inference of admixture graphs on Native American and Arctic populations. PLoS Genet. *19*, e1010410. https://doi.org/10.1371/journal.pgen.1010410.

89. Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., et al. (2012). A high-coverage genome sequence from an archaic denisovan individual. Science *338*, 222–226. https://doi.org/10.1126/science.1224344.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
PP-00294	This study	Tatóok yík yées sháawat (TYYS)
Chemicals, peptides, and recombinant proteins		
Sodium Hypochlorite Solution	Fisher Scientific	Cat. No. SS290-1
Proteinase K	Fisher Scientific	Cat. No. FEREO0492
EDTA, Disodium Salt, Dihydrate, Solution (0.5M)	Fisher Scientific	Cat. No. 40-551-00 ML
Guanidine hydrochloride, Molecular Biology Grade	Fisher Scientific	Cat. No. 50-305-KG
Isopropanol	Fisher Scientific	Cat. No. BP26181
Tween 20	Fisher Scientific	Cat. No. 65-520-650 ML
Sodium acetate (3M)	Fisher Scientific	Cat. No. 56-742-2100 ML
Tris EDTA Buffer, Molecular Biology Grade	Fisher Scientific	Cat. No. AAJ75893AP
MinElute PCR Purification Kit	Qiagen	Cat. No. 28006
Deposited data		
TYYS Shotgun BAM file	Available upon request from lead contact	TYYS_shotgun.final.bam
TYYS Shotgun fastq files	Available upon request from lead contact	TYYS_Shotgun_HYC3LDSXX_s4_1_ MY7351-MY5349_SL416825.fastq.gz TYYS_Shotgun_HYC3LDSXX_s4_2_ MY7351-MY5349_SL416825.fastq.gz
TYYS Mitochondrial Enrichment fastq files	Available upon request from lead contact	TYYS_Mito_Enrichment_HYC3LDSXX_ s2_1_MY7351-MY5349_SL416731.fastq.gz TYYS_Mito_Enrichment_HYC3LDSXX_ s2_2_MY7351-MY5349_SL416731.fastq.gz
TYYS mitochondrial fasta file	Available upon request from lead contact	TYYS_Mitogenome.fa
Oligonucleotides		
5′ 8bp i5 indexed adapter: 5′AATGATACGGCGACCA CCGAGATCTACAC[i5 index] ACACTCTTTCCCTACACGAC GCTCTTCCGATCT	Illumina	5′ 8bp i5 indexed adapter
3′ 8bp i7 indexed adapter: 3′AGATCGGAAGAGCACAC GTCTGAACTCCAGTCAC[i7 index]ATCTCGTATGCCGTCTTCTGCTT	Illumina	3′ 8bp i7 indexed adapter
Software and algorithms		
AdapterRemoval (version 2.3.1)	Schubert et al.[64]	https://github.com/MikkelSchubert/adapterremoval
BWA (version 0.7.13)	Li and Durbin[65]	http://bio-bwa.sourceforge.net/
Bowtie2 (version 2.2.8)	Langmead and Salzberg[69]	https://bowtie-bio.sourceforge.net/bowtie2/
MapDamage (version 2.0.8)	Jónsson et al.[70]	https://github.com/ginolhac/mapDamage
Schmutzi (version 0.0.1)	Renaud et al.[68]	https://github.com/grenaud/schmutzi
Samtools (version 1.9)	Li et al.[66]	https://github.com/samtools/samtools

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HaploGrep2 (version 2.4)	Weissensteiner et al.[41]	https://haplogrep.uibk.ac.at/
BEAST (version 1.10.4)	Suchard et al.[73]	https://github.com/beast-dev/beast-mcmc
EIGENSOFT (version 6.1.4)	Patterson et al.[80,81]	https://github.com/DReichLab/EIG
AdmixTools (version 7.0.1)	Patterson et al.[79]	https://github.com/DReichLab/AdmixTools
Admixtools2	Maier et al.[86]	https://github.com/uqrmaie1/admixtools
ADMIXTURE (version 1.3.0)	Alexander et al.[48]	https://dalexander.github.io/admixture/
PLINK (version 1.9)	Chang et al.[85]	https://www.cog-genomics.org/plink/1.9/
PhyloTree (build 17)	van Oven and Kayser[39]	http://www.phylotree.org/
GATK (version 4.1.6)	Poplin et al.[77]	https://gatk.broadinstitute.org/hc/en-us/sections/360008481611-4-1-6-0
RFMix (version 2)	Maples et al.[84]	https://github.com/slowkoni/rfmix
Shapeit (version 2)	Delaneau et al.[83]	https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to the lead contact, Charlotte Lindqvist (cl243@buffalo.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- To promote responsible access to genomic data from Indigenous communities, including Ancestors, the raw sequence and mapped data are available upon request from the corresponding author and review of its use by the Wrangell Cooperative Association tribal council and completion of a Data Use Agreement.
- No new code was generated in this study.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The human remains represent a small skeletal fragment from a humerus with a diameter of about 3 cm excavated from Lawyer's Cave (also called Phalanges Phreatic Tube), which is located in the Alexander Archipelago of Southeast Alaska.

METHOD DETAILS

Archaeological context and sex determination

The human remains, given the field number PP-00294 and analyzed here, were found in Lawyer's Cave also called Phalanges Phreatic Tube by cavers, referring to the cave's shape and the toe bones of a bear found inside. The cave, which has two entrances and consists of an approximately 20 m long, nonbranching crawlway from end to end, is located along Blake Channel on the southeast Alaskan mainland east of Wrangell Island in the Alexander Archipelago. The cave is rich in postglacial remains that were discovered during two excavations in 1998 and 2003,[56,57] including bones of various mammals, birds, and fish. Notably, the cave also harbors human and dog remains,[58] suggesting human occupation of the cave. Several artifacts, including a bone spear point, bone awl, shell beads, partial obsidian biface, obsidian microblade, and obsidian flakes further support the human occupation of the cave. The bone awl was dated to 3,050 $\pm$ 40 ^{14}C years BP.[56,57]

The human remains is a small skeletal fragment from a humerus with a diameter of about 3 cm. Radiocarbon dating of this sample was performed in 2004 by the geochronology laboratory at the University of Arizona

(AA-57000; see Heaton & Grady[56] for details) and returned a date of 3,425 ± 50 ^{14}C years BP. Because the stable isotope results from this individual suggest a fully marine diet, the ^{14}C date was calibrated using the Marine20 calibration curve[59] in Calib 8.2,[60] with marine Delta R following[61] and 2-sigma dates reported. The calibrated median date was estimated to 2,950 cal BP ranging from 2,500 to 3,379 cal BP. The δ^{13}C value of was −12.5‰.

We identified the remains as coming from a female using a method that identifies sex based on the number of reads from shotgun sequencing that map to the X chromosome versus those that map to the Y chromosome.[62]

Community engagement

Cooperation and common goals between Alaska Native tribes and the scientific community have aided in significant finds including those presented in this paper. Forest Service archaeologists and Wrangell District rangers worked closely with the Wrangell Cooperative Association (WCA), the tribal group associated with the area, regarding the archaeological component at Lawyer's Cave. Our collaboration, with outreach and respectful interest from the scientific community, resulted in the repatriation of human remains recovered from the cave and scientific analysis reported here, including radiometric analysis for age determination and paleogenomic analyses for genetic information. The WCA named the ancient individual analyzed in this study as *Tatóok yík yées sháawat* ("Young lady in cave"). This name (as well as the acronym *TYYS* for the purpose of this paper) is used here.

DNA extraction and sequencing

Genomic DNA was extracted in a dedicated cleanroom facility appropriate for ancient DNA research, physically separated from any handling of modern samples and post-PCR procedures. The ancient DNA extraction followed the protocol described in Dabney et al.,[63] with some modifications. Using a dentist drill, approximately 100 mg of fine bone powder was obtained. Following overnight digestion with proteinase K and addition of the binding buffer, the mixture was purified and concentrated with a Qiagen MinElute PCR Purification Kit (Qiagen, USA). The final elution step was performed twice with 25 μL of TE buffer for a total DNA volume of 50 μL. A negative control was prepared alongside the extraction. Ancient DNA single-stranded Illumina library preparation (ssDNA 2.0), mitochondrial DNA target enrichment, and whole genome shotgun sequencing was performed by Daicel Arbor Biosciences. The library was target enriched using Daicel Arbor Biosciences' predesigned "Human, Modern Global" mitogenome bait panel and single-plex dual-round captures. The enriched library, as well as shotgun sequencing, was sequenced with Illumina NovaSeq PE150.

Trimming and mapping

Adapters were removed from the Illumina reads using AdapterRemoval 2.3.1.[64] After adapters were removed, trailing and leading stretches of Ns and/or low-quality bases were trimmed. Additionally, both the 3′ and 5′ ends of the reads were trimmed by 2 bases. After trimming, reads shorter than 20 bp were discarded. Finally, overlapping reads were collapsed into single reads.

The trimmed shotgun and mitochondrial enrichment reads were separately aligned to the Cambridge Reference for the human mitochondrial genome using the same pipeline. In particular, we used BWA 0.7.13[65] with the aln algorithm and seed value set to 1024, to map the collapsed reads as well as the uncollapsed read pairs separately to the mitochondrial reference. The resulting BAM files were merged using Samtools' merge option. Unmapped reads were obtained from the merged file using Samtools,[66] and mapped using BWA-mem.[67] Resulting files from BWA-aln and BWA-mem were then merged. PCR duplicates were removed using the MarkDuplicates tool in Picard 1.119 (http://broadinstitute.github.io/picard/). To obtain better coverage, files containing mapped reads from shotgun sequencing and mitochondrial enrichment were merged using the MergeSamFiles option in Picard. The combined width of coverage was 100% and the depth was 112X. The resulting file was used as input for Schmutzi[68] to produce a consensus sequence of the endogenous mitochondrial genome. In particular, we first used the contDeam.pl module to obtain an initial estimate of the present-day human contamination using C to T misincorporations at the ends of the reads. Next, we used the wrapper script schmutzi.pl, which takes as input a set of potential mitochondrial contaminants. Based on the misincorporation patterns in the focal reads and presence sequences from potential contaminants, schmutzi.pl produces a consensus sequence for the endogenous mitochondrial genome, along with a final estimate of present-day contamination. Five

percent of our mapped reads were estimated to be contaminants. The endogenous mitochondrial genome obtained here was used for all downstream analysis.

We also mapped the trimmed shotgun reads to human reference hg19 using the following pipeline. We used Bowtie2 version 2.2.8,[69] with options -L 10, -N 1, and –local, to map collapsed reads as single reads (using the -U option), and uncollapsed reads as paired reads (using −1 and −2 options). Unmapped reads were extracted using Samtools and remapped using BWA-aln. Unmapped reads were extracted again and then mapped using BWA-mem. Files containing mapped reads from Bowtie2, BWA-aln, and BWA-mem were then merged using Samtools. Only reads with a mapping quality of 30 or above were retained. Finally, PCR duplicates were removed using the MarkDuplicates tool in Picard 1.119. After mapping the reads with the aforementioned quality control, we obtained a coverage width of 0.14 and a depth of 1.1X. The mapping statistics are summarized in Table S1. Additionally, we obtained the fragment misincorporation plots for mapped reads (both, with and without end-trimming) using MapDamage2 (Figure S1).[70]

Mitochondrial genome analysis

To investigate whether the ~2,950 yo *Tatóok yík yées sháawat* (*TYYS*) from Southeast Alaska exhibits a mitochondrial lineage that is also observed in present-day Indigenous Americans from the PNWC, we constructed a dataset containing 523 complete mitochondrial genomes, including sequences from *TYYS*, *Anzick-1*,[37] and *Shuká Káa*.[23] The remainder of the mitogenomes were obtained from GenBank. Figure 2A includes the NCBI accession codes for samples that are not from the Pacific Northwest. The accession codes for *939, 938, 152, Tsimshian 018, 160a, Nisga'a B009,* and *Tsimshian 069* are KC998701-KC998707, respectively.

We aligned these sequences using MAFFT version 7.[71,72] For phylogenetic analysis of these sequences, we used BEAST 1.10.4,[73] with the following parameters: 1) gamma site model with 6 gamma categories and no invariant sites, 2) the generalized time-reversible substitution model, 3) coalescent constant population, and 4) a strict constant clock model with a normal prior with $\mu = 1.665 \times 10^{-8}$ and $\sigma = 1.479 \times 10^{-9}$.[74,75] We ran three BEAST replicates, each with a Markov chain length of 10 million. We used Tracer v. 1.7.2[38] to conclude that each replicate had converged. LogCombiner was used to combine the results from the three replicates, with 10% of the sampled posterior trees being discarded for each replicate in the burn-in process. Finally, we used TreeAnnotator to summarize the sampled trees into a single Maximum Clade Credibility (MCC) tree. We then visualized the MCC tree using EMBL's Interactive Tree of Life (iTOL).[76]

Variant calling

We used *GATK* best practices workflow (https://gatk.broadinstitute.org/hc/en-us/articles/360035535932-Germline-short-variant-discovery-SNPs-Indels-) to call genotypes for *TYYS*. In particular, we assembled a set of BAM files including that for *TYYS*, and those for a number of previously published ancient and modern genomes for which BAM files were publicly available (see Table S7 for a complete list). First, for each of these BAM files, we separately called variants using GATK's HaplotypeCaller tool in GVCF mode.[77] Next, we use the *CombineGVCFs* tool to combine the per-sample gVCF files into a multi-sample gVCF file. Using the *GenotypeGVCFs* tool on the multi-sample gVCF file to perform joint genotyping on the samples. The *VariantAnnotato*r tool was used to annotate the VCF file for FisherStrand, MappingQualityRankSumTest, ReadPosRankSumTest, RMSMappingQuality, and QualByDepth. Then, we used VariantRecalibrator, which uses machine learning to identify annotation profiles for variants that are likely to be true based on a training dataset, and accordingly assigns a score to each variant in the target VCF; for this step, we used HapMap 3.3, OmniChip 2.5 and dbSNP138 resources with priors 15, 12, and 2 respectively as the training datasets. Lastly, we used *ApplyVQSR* to filter variants using a truth-sensitivity threshold of 99.0.

Dataset for nuclear DNA analysis

The dataset we used for nuclear DNA analysis was constructed by merging the variant call file we created for *TYYS* and a number of other individuals (described above) with five previously compiled variant call datasets: 1) whole-genome sequences from 929 individuals in the Human Genome Diversity Project[78]; 2) (Single Nucleotide Polymorphism) SNP chip data from 167 East Asian, Siberian, and Greenlandic individuals[4]; 3) SNP chip data from 21 Dakelh Athabascans[4]; 4) SNP chip data from 19 Greenlandic individuals[3]; and 5) SNP chip data from 82 Pacific Northwest coastal and inland individuals.[20] The merging was performed using the *mergeit* tool in ADMIXTOOLS.[79] Thus, only the union of individuals and the intersection of SNPs was retained in the resulting merged dataset. Table S8 contains a list of all individuals in this dataset.

QUANTIFICATION AND STATISTICAL ANALYSIS

PCA

We performed two rounds of PCA using the *smartpca* program in the *EIGENSOFT 6.1.4* package[80,81]: one using the global panel (excluding Africans and Oceanians), and the other using modern and ancient individuals from the Pacific Northwest. In both cases, the PCs were calculated using modern samples, and the ancient samples were projected on top of PCs using the *lsqproject* option in the *smartpca* program. 483,090 SNPs were used for calculating the principal components. For projected samples, the number of intersected SNPs were as follows: 41,163 for *TYYS*, 112,222 for *939*, 119,835 for *302*, 174,697 for *443*, 6,284 for *Shuká Káa*, 188,852 for *The Ancient One*, 472,090 for *Anzick-1*, and 420,433 for *Saqqaq*.

ADMIXTURE

To gauge the levels of European admixture in the Indigenous individuals in our dataset, we performed model-based clustering of the individuals in our dataset using *ADMIXTURE*[48]; we use only European, Siberian, and Indigenous American – including Inuit – individuals for this analysis. In particular, we ran 10 replicates of *ADMIXTURE* for $2 \leq K \leq 12$, retaining results with the largest log likelihood value for each K. We visualized the results using *pong*.[82]

Masking for alleles found in Europeans

For all downstream analysis, we removed Pacific Northwest individuals with less than 30% NNA ancestry based on *ADMIXTURE* results with K = 7. The remaining Pacific Northwest individuals were masked for European admixture based on a three-step process: first, we phased variant calls for individuals in our dataset; secondly, for each Pacific Northwest individual, we classified regions of the genome as belonging to either European, East Asian, African, or Indigenous American ancestry; and lastly, SNPs falling in the European, East Asian, or African regions were removed for each Pacific Northwest Individual. The phasing was performed using *shapeit2*,[83] using the HapMap2 genetic map. We performed local ancestry estimation using *RFMix*.[2,84] To this end, we used a reference panel composed of 48 Europeans, 48 East Asians, 48 Africans, and 48 Indigenous Americans: including 8 Athabascans with an NNA ancestry $\geq$95% based on *ADMIXTURE* results. For each Pacific Northwest individual, regions were thus identified as coming from European, East Asian, African, or Indigenous American ancestry, based on Viterbi calls. For each individual, SNPs within the regions identified as European were set to "missing" using the *–zero-cluster* option in *PLINK*.[85]

f- and D-statistics

For f_2, f_3, f_4, and D-statistics[2,50] we used Pacific Northwest individuals that were masked for European, Asian, and African alleles. We calculated the outgroup f_3 statistic of the form (X, *TYYS*; Han), where X is an American or Siberian population. This was done using the *qp3pop* tool, with the *outgroupmode* option, in *ADMIXTOOLS*.[79] We also calculated the f_4 statistics[50] of the form (Han, *TYYS*; X, Y), where X and Y are North American populations using the *qpDstat* tool, with the *f4mode* option enabled, in *ADMIXTOOLS*. The D-statistics, too, were calculated using the *qpDstat* tool, but with *f4mode* option disabled. The f_2 results were obtain using the f2 function in *ADMIXTOOLS 2*.[86] For *qpAdm* modeling, we used *ADMIXTOOLS*. We used the following set of populations as the "right" populations (outgroups): {Mbuti, Papuan, Buryats, Evenkis, Dolgans, Nganassans, Yakut, Tuvinians, Altaians, Selkups, Dai}. We used the set {X, Y, Z} as the "left" populations (target and sources): where X is the target population or individual, and Y and Z are the sources. A model was rejected if the tail probability associated with the model was less than 0.05, or if the ancestry contribution associated with any source was outside the [0, 1] interval. Following the tradition in the literature,[87] if a target could be modeled as both a combination of two sources and only one of the two sources, we preferred the model which required only one source. To test for "proto-Paleo-Eskimo" (PPE) geneflow into present-day PNW populations, we followed two different strategies for target populations: 1) for each PNW population, only using unmasked genomes with less than 1% European ancestry based on *ADMIXTURE* results (Table S3); and 2) using masked genomes for each PNW population (Table S4). In both cases, Y was either a South American population or an ancient individual; Z was either *Saqqaq* or speakers of Chukotka-Kamchatkan languages (only individuals with less than 1% European ancestry based on *ADMIXTURE* results were obtained). For testing PPE gene flow into ancient individuals (Table S5), we let X be *443, 302, TYYS, 939, the Ancient One,* or *Anzick-1*. Y could be either *Saqqaq* or the speakers of Chukotka-Kamchakan languages (only individuals with less than 1% European ancestry based on *ADMIXTURE* results were retained). We let Z be Tlingit, Nisga'a, Haida, Tsimshian, Splatsin, Stswecem'c, Karitiana, Surui, Pima, or any ancient individual other than X for a given test. To test for First People gene

flow into Saqqaq, we let X be Saqqaq; Y was Chukotka-Kamchatkan Siberians; and Z was Tlingit, Nisga'a, Haida, Tsimshian, Splatsin, Stswecem'c, *302, 443, TYYS, 939, the Ancient One, Anzick-1*, Karitiana, Surui, or Pima (Table S6).

Conceptual concerns in testing for admixture between PPE and first people

This analysis was performed using *qpAdm* modeling.[52,53] Testing whether 1) present-day populations from the Pacific Northwest (PNW), and 2) ancient individuals from the PNW harbor PPE ancestry is a complicated task. To test whether present-day populations can be modeled as a combination of PPE and ancient individuals from the region, we must assume that ancient individuals from the region are free of PPE ancestry, and to model ancient individuals from the region as a combination of PPE and present-day populations, we have to assume that present-day population are free from PPE ancestry. Consequently, the reasoning becomes circular. However, using a wide range of populations with First People ancestry as one of the sources may help making sense of the data. Moreover, since *The Ancient One* (*Kennewick man;* 8,500 ka), and *Anzick-1* (~13,000 ka) are dated to times prior to the estimated time of arrival (~6 ka) of the Paleo-Inuit into the Americas, they can be used as sources free of PPE ancestry.

To test for PPE ancestry in present-day PNW populations, we used two different strategies: 1) using unmasked genomes for each population, retaining only those individuals who had less than 1% European ancestry based on *ADMIXTURE* results; and 2) using masked genomes. In both cases, source-1 is a representative of First People ancestry, and source-2 is a representative of PPE ancestry. Using first strategy (Table S3), we find that the Tlingit, Haida, and Nisga'a can be modeled as a combination of *Anzick-1* and PPE. Using the second strategy (Table S4), we find that only Nisga'a can be modeled as a combination of *Anzick-1* and PPE. Curiously, the results do not hold if source-1 is represented by *The Ancient One* or any other representative of the First People ancestry. We suspect we get these results because *Anzick-1* is a very old sample. The older the sample, the greater the genetic distance between it and present-day PNW populations, and the smaller the genetic distance between PPE and that sample. It is possible that *qpAdm* is unable to distinguish between the two sources in such a case. We think that an argument made for the presence of PPE ancestry in present-day PNW tribes based on our results would be tenuous at best.

To test for PPE ancestry in ancient genomes, we sought to model them as combinations of PPE and various representative of the First People ancestry (Table S5). Here we find instances wherein the target ancient individual with First People ancestry can be modeled as a combination of PPE and another ancient individual with First People ancestry. However, in all these cases, the target individual can also be modeled with 0% PPE ancestry. In these cases, we use parsimony to reject dual-source models. We find no evidence suggesting that *302, 443, TYYS, 939, The Ancient One,* or *Anzick-1* harbor PPE ancestry.

Lastly, we test for the First People ancestry in *Saqqaq* genome (Table S6). We find that while *Saqqaq* cannot be modeled as a combination of Southern Native Americans (SNA) and PPE, it can be modeled as combination of NNA and PPE ancestries. The latter result holds regardless of whether the NNA representative is an ancient individual or a present-day population. Here too, we take the results where *Anzick-1* is souce-2 with a grain of salt.