The Genetic Prehistory of the New World Arctic

Maanasa Raghavan¹, Michael DeGiorgio², Anders Albrechtsen¹⁸, Ida Moltke¹³, Pontus Skoglund⁵,⁶,⁹, Thorfinn S. Korneliussen¹, Bjarne Grønnøw⁷, Martin Appelt⁷, Hans Christian Gulløv⁷, T. Max Friesen¹, William Fitzhugh⁴, Helena Malmström¹,⁵, Simon Rasmussen¹⁰, Jesper Olsen¹¹, Linea Melchior¹², Benjamin T. Fuller¹³, Simon M. Fahrni¹³, Thomas Stafford Jr.¹¹,¹⁴, Vaughan Grimes¹⁴,¹⁵, M.A. Priscilla Renouf¹⁴,¹⁵, Jerome Cybulski¹⁶,¹⁷, Niels Lyngren⁷, Marta Mirazon Lahr¹⁸, Kate Britton¹⁸,¹⁹, Rick Knecht¹⁹, Jette Arneborg²⁰, M. Geoffrey Hayes³³,³⁴,³⁵, Joan Coltrain³⁶, Victor A. Spitsyn³⁷, Anders Götherström³⁸, Ludovic Orlando¹, Toomas Kivisild²¹,³⁹, Richard Villems²¹,²², Michael H. Crawford³⁰, Finn C. Nielsen³⁸, Jørgen Dissing¹², Jan Heinemeier¹¹, Morten Meldgaard⁴, Carlos Bustamante²³, Dennis H. O’Rourke³⁶, Matthias Jakobsson³, M. Thomas P. Gilbert¹, Rasmus Nielsen³⁵, and Eske Willerslev¹#.

1. Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen, Denmark.
2. Department of Biology, Pennsylvania State University, 502 Wartik Laboratory, University Park, PA 16802, USA.
3. The Bioinformatics Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark.
4. Department of Statistics, The University of Chicago, 5734 S. University Avenue Chicago, IL 60637, USA.
5. Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, 75233 Uppsala, Sweden.
6. Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA.
7. SILA - Arctic Centre at Ethnographic Collections, The National Museum of Denmark, Frederiksholms Kanal 12, DK-1220 Copenhagen K, Denmark.
8. Department of Anthropology, University of Toronto, 19 Russell St., Toronto, Ontario M5S 2S2, Canada.
9. Arctic Studies Center, P.O. Box 37012, Department of Anthropology, MRC 112, National Museum of Natural History, Smithsonian Institution, Washington, D.C. 20013-7012, USA.
10. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kemitorvet, Building 208, room 006, 2800 Kgs. Lyngby, Denmark.
11. AMS 14C Dating Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark.
12. Anthropological Laboratory, Institute of Forensic Medicine, Faculty of Health Sciences, University of Copenhagen, Frederik V’s Vej 11, 2100 Copenhagen Denmark.
13. Department of Earth System Science, University of California, Irvine, Keck CCAMS Group, B321 Croul Hall, Irvine, California, 92697, USA.
14. Department of Archaeology, Memorial University, Queen’s College, 210 Prince Philip Drive, St. John's, Newfoundland, A1C 5S7, Canada.
15. Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany.
16. Canadian Museum of History, 100 Rue Laurier, Gatineau, Quebec K1A 0M8, Canada.
17. Department of Anthropology, University of Western Ontario, 1151 Richmond St. N., London N6A 5C2, Canada.
18. Leverhulme Centre for Human Evolutionary Studies, Dept. of Archaeology & Anthropology, University of Cambridge, The Henry Wellcome Building, Fitzwilliam Street, Cambridge CB2 1QH, UK.
19. Department of Archaeology, University of Aberdeen, St. Mary’s Building, Elphinstone Road, Aberdeen AB24 3UF, Scotland, UK.

# Corresponding author.
22. Department of Evolutionary Biology, University of Tartu, Tartu 51010, Estonia.
23. Department of Genetics, School of Medicine, Stanford University, 300 Pasteur Dr. Lane Bldg Room L331, Stanford, California 94305, USA.
24. School of Biological Sciences, Washington State University, PO Box 644236, Eastlick Hall 395, Pullman, Washington 99164, USA.
25. Department of Integrative Biology, University of California, Berkeley, CA 94720 USA.
27. Informatics and Bio-computing, Ontario Institute for Cancer Research, 661 University Avenue, Suite 510, Toronto, Ontario, Canada, M5G 0A3, Canada.
28. Center for Genomic Medicine, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark.
29. Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa Russia.
30. Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Bashkortostan, Russia, 450074
31. State Museum for Oriental Art, 12a, Nikitsky Boulevard, Moscow 119019, Russia.
32. Greenland National Museum and Archives, P.O. Box 145, DK-3900 Nuuk, Greenland.
33. Division of Endocrinology, Metabolism and Molecular Medicine, Department of Medicine, Northwestern University Feinberg School of Medicine, 303 East Chicago Avenue, Tarry, 15-759, Chicago, IL 60611, USA.
34. Department of Anthropology, Weinberg College of Arts and Sciences, Northwestern University, 1810 Hinman Avenue, Evanston, IL 60208, USA.
35. Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, 303 E. Superior Street, Lurie 7-125, Chicago, IL 60611, USA.
36. Department of Anthropology, University of Utah, 270 S. 1400 E., Salt Lake City, UT 84112, USA.
37. Medical-Genetic Center of Russian Academy of Sciences, 10 Naberezhnaya Ushaiki, Tomsk 634050, Russia.
38. Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden.
39. Department of Archaeology and Anthropology, University of Cambridge, The Henry Wellcome Building, Fitzwilliam Street, Cambridge CB2 1QH, UK.
40. Laboratory of Biological Anthropology, University of Kansas, 1415 Jayhawk Blvd., 622 Fraser Hall, Lawrence, Kansas 66045, USA.

* Equal contribution
^ Deceased
# Corresponding author
The New World Arctic, the last region of the Americas to be populated by humans, has a relatively well-researched archaeology but is lacking an understanding of its genetic history. We present genome-wide sequence data from ancient and present-day humans from Greenland, Arctic Canada, Alaska, Aleutian Islands and Siberia. We show that Paleo-Eskimos (ca. 3000 BC-1300 AD) represent a migration pulse into the Americas independent of both Native American and Inuit expansions. Furthermore, the genetic continuity characterizing the Paleo-Eskimo period was interrupted by the arrival of the ancestors of present-day Inuit, with evidence of gene flow between these populations. Despite periodic abandonment of major Arctic regions, a single Paleo-Eskimo meta-population likely survived in near-isolation for more than 4,000 years, only to vanish around 700 years ago.
Humans first peopled the North American Arctic (northern Alaska, Canada and Greenland) from the Bering Strait region beginning around 6,000 years before present (1), leaving behind a complex archaeological record [supplementary text S1 (2), Fig. 1A]. Over successive millennia, the pioneering Arctic cultures developed into distinct lifestyles and cultural stages grouped within two broad cultural traditions known as Paleo-Eskimo and Neo-Eskimo. Early Paleo-Eskimo people representing the Denbigh, Pre-Dorset, Independence I and Saqqaq cultures (ca. 3,000-800 BC) lived in tent camps and hunted caribou, musk ox and seals with exquisitely-flaked stone tools similar to those used by northeast Siberian Neolithic cultures (2-5). In northern Alaska, the Denbigh cultural groups were succeeded by the Paleo-Eskimo Choris and Norton cultures starting around 900 BC, with the Norton material culture further developing into the Ipiutak culture around 200 AD (79,83). Simultaneously, during the cold period beginning around 800 BC, innovations in housing and hunting technologies accompanied the formation of the Late Paleo-Eskimo or Dorset culture in eastern Arctic (eastern Canadian Arctic and Greenland), with population growth and more intensive use of marine mammals including walrus (5,6,7). The Dorset culture is divided into three phases: (i) Early Dorset ca. 800 BC – 0 BC/AD, (ii) Middle Dorset ca. 0 BC/AD – 600-800 AD, and, (iii) Late Dorset ca. 600-800 AD – 1300 AD (7). The Paleo-Eskimo tradition in the eastern Arctic ended sometime between 1150-1350 AD; shortly after the sudden appearance of the Neo-Eskimo Thule whale-hunters from the Bering Strait region (7-15).

The Siberian Old Bering Sea culture is the earliest expression of the Neo-Eskimo tradition ca. 2,200 years before present, developing into the Punuk culture around the sixth century AD. Almost concurrently, the Old Bering Sea culture developed into the Birmirk culture in the northern parts of the Bering Strait region. Interactions between people of the Birmirk and Punuk cultures gave rise to the western Thule culture on both sides of the Bering Strait. On the Alaskan side of the Strait, the Ipiutak culture also contributed to the formation of the western Thule culture, (80). By the early second millennium AD, western Thule cultural groups began their movement into the eastern North American Arctic (12,81). With the Thule culture came more effective means of transportation like dog sleds and large skin boats, complex tool kits like sinew-backed bows and harpoon float gear for hunting large whales and sinew-backed bows (8,16), Thule culture spread quickly throughout the eastern Arctic, rapidly replacing Dorset in most, if not all, regions. The decline of whaling during the latter part of the Little Ice Age (sixteenth to nineteenth century AD) resulted in a readjustment to ice-edge and breathing hole hunting of walrus and seal, laying the foundation for modern Inuit cultures (17). Additionally, the Norse (Vikings) formed settlements in Greenland around 985 AD and occupied regions in southern Greenland for about 500 years, contemporaneous with both the Late Dorset and Thule, reaching Newfoundland and Labrador in eastern Canada around 1000 AD (18).

Continuities in chipped stone bifaces and blade and burin technology point to Paleo-Eskimo origins among Siberian Neolithic cultures (5,8,19,20). Genetic evidence suggests that the earliest eastern Arctic Paleo-Eskimo people represented an independent Siberian migration into the New World (21,22) (Fig. 1C). However, to date we have been unable to identify the likely Siberian ancestral population. Some have argued for origins from an ill-defined 8,000-year-old South Alaska Eskimo-Aleut or Na Dene blade and burin complex (23), but these remain controversial. A recent genetic study is in support of the Early Paleo-Eskimos, specifically Saqqaq, sharing ancestry with Na Dene Native Americans (Fig. 1C), as part of a three-wave peopling model of the Americas consisting of (a) Amerindians, (b) Eskimo-Aleuts and, (c) Saqqaq and Na Dene (24). Alternate hypotheses on Dorset origins include the Aleutian Islanders (27) and earlier theories of Amerindian cultures in eastern Canada and even further south (17,28-30) (Fig. 1C). Current views favor an in situ origin of Dorset from Canadian Pre-Dorset in northern Hudson Bay (8,11,31,32) (Fig. 1B).

Additionally, whether the individual Early, Middle and Late Dorset phases represent genetic continuity of the same peoples or not remains unresolved (Fig. 1B). The Dorset chronological sequence from the T1 site on Southampton Island, Igloolik, South Baffin, and Labrador in Canada provide evidence of cultural continuity through Early to Late Dorset (16, 17, 18, 32), although regional differences and settlement discontinuities are also common (34, 35, 36). Significant regional discontinuities and occupation gaps also occur in Greenland (7). It is also debated whether abandoned areas were re-occupied by people from a different genetic background and whether this occurred from a Central Arctic ecological ‘core area’ or
regional core areas into which human populations retreated and restructured before expanding again into periodically refurbished marginal zones (16,31). The resolution of these controversies has been hindered by the limited amount of Paleo-Eskimo human material, difficulties in assigning cultural affiliation of some finds (36), and dating uncertainties resulting from the strong marine component in the Arctic diet (7).

**Samples and sequence data**

We collected bone, teeth, and hair samples from the field and museums representing 169 ancient human remains from Arctic Siberia, Alaska, Canada and Greenland (Fig. S2, Table S1). These remains have been assigned to one of several ancient Arctic cultures on the basis of typological and/or stratigraphic evidence and, in some cases, radiocarbon dating [supplementary text S1 (2)]. To circumvent drawing conclusions from single genomes (40), we generated mitochondrial DNA (mtDNA) data from 158 and low-coverage whole genome datasets from 26 of the ancient samples (up to 0.3X depth) [supplementary text S3 (2), Tables S7, S9A]. Despite colder temperatures in the Arctic, DNA survival in the ancient samples was surprisingly low, ranging from ~0 to 3.2% endogenous content based on the genome sequencing data (Table S7). This low endogenous content may be explained by the remains being largely surface burials that suffered from fluctuating temperatures and humidity, and to subsequent storage conditions at museums.

We also sequenced two high-coverage genomes from present-day North American Native Americans belonging to the Na Dene family (the Dakelh of British Columbia, hereafter referred to as Athabascans), and five unrelated, present-day Greenlandic Inuit (n = 2), Aleutian Islander (n = 1) and Siberian Nivkhs (n = 2) to average depths of 20-40X [supplementary text S3 (2), Table S8]. Only the Aleutian Islander showed evidence of recent European admixture, and was masked for non-Native American ancestry tracts prior to analyses [supplementary text S5 (2)]. Additionally, we radiocarbon dated 27 ancient samples and corrected 25 of the dates for marine reservoir effect to account for the dominant marine component in these individuals’ diets (15 of these samples are represented in the aforementioned genomic dataset) [supplementary text S2 (2)]. This is critical in the accurate cultural assignment of these individuals, especially in cases where stratigraphic information is inconclusive or contentious [supplementary text S1 (2)].

**Origins of Paleo-Eskimos**

Diagnostic mtDNA coding region markers were targeted in the ancient samples to determine their mtDNA haplogroup (hg) affinities. While hgs A, B, C, D and X are among the five founding haplogroups in the Americas, previous studies have shown the near-absence of hgs B, C and X in Paleo-Eskimos and Thule as well as among present-day Inuit (Saillard et al., 2000; Hayes et al., 2005; Helgason et al. 2006; Gilbert et al. 2007; Gilbert et al. 2008, Raff et al. 2011). We observe mtDNA hg D, specifically the lineage hg D2a, in both Early and Late Paleo-Eskimos, with the majority of the Pre-Dorset/Saqqaq and Middle Dorset samples further classified as hg D2a1 [supplementary text S4 (2), Table S9A]. The absence of biological remains affiliated to the Early Dorset phase precludes genetic testing for this period. Hgs D2a and D2a1 are found in present-day Aleutian Islanders and Siberian Eskimos (42), who are genetically among the closest living populations to the previously sequenced Greenlandic Paleo-Eskimo (Saqqaq) individual that also belonged to hg D2a1 (hereafter, high-coverage Saqqaq) (22). The single Canadian Pre-Dorset sample (XIV-H:168, Rocky Point) was typed to hg D4e (referred to as hg D2 in (42)), which is ancestral to hg D2a, but lacks further phylogenetic resolution due to the low coverage nature of the shotgun data [supplementary text S4 (2)].

Maximum likelihood trees based on nuclear DNA variation place the Middle Dorset, Late Dorset and the Canadian Pre-Dorset individuals as sister groups to the high-coverage Saqqaq individual (Figs. 2A, S9B-D), separately from contemporary Greenlandic Inuit or Native Americans [represented by the South American
Karitiana (44) and the ancient Clovis Anzick-1 individual (45). Similar results were obtained when allowing for admixture between populations with TreeMix (43) (Figs. S9C-D). When the analysis was repeated including the Aleutian Islander, which was masked for European ancestry over ca. 80% of its genome [supplementary text S5 (2)], the Canadian Pre-Dorset, Middle Dorset and the Late Dorset individuals grouped with the high-coverage Saqqaq; as before (Fig. S10).

Pairwise outgroup f2-statistics (46,47) and D-statistics (46,48) confirm Early and Late Paleo-Eskimos as being significantly (no overlap at three standard errors for the SNP chip data and at one standard error for the sequencing data) closer to one another than to any of the sampled present-day populations, including those from the Americas and Siberia (Figs. 2B, S11A-D). Furthermore, admixture clustering profiles (49) of the high-coverage Saqqaq and a Dorset individual are near-identical (Fig. S8), with both sharing components with present-day Siberian Chukchi and Greenlandic Inuit and, to a lesser extent, with other Siberians and East Asians (Han). Variations in TreeMix graph topologies occurred with the inclusion of the different ancient samples [supplementary text S5 (2)]; however, the resulting conclusions from these trees are consistent with other analyses. Hence, evidence from mitochondrial as well as nuclear markers suggests that all Paleo-Eskimos, from both Canada and Greenland, represent a continuum of the same single ancestral population (Figs. 1B, C). Present-day populations that are genetically closely related to the Paleo-Eskimos include the Greenlandic Inuit, Aleutian Islanders and far-east Siberians (Figs. 2C, S11E).

We additionally tested the claim that Saqqaq and Na Dene were part of the same expansion into the Americas (26) by including present-day Athabascans, who represent distinct early branches of Native Americans (45), in the TreeMix (43) analysis. The maximum likelihood tree places the Athabascans as a sister clade to Karitiana and Anzick-1 (Fig. S12A). Similarly, using single nucleotide polymorphism (SNP) chip data from select Old and New World populations, masked for European admixture, we observe that the high-coverage Saqqaq individual forms a clade with the far-east Siberian Koryaks instead of with the Chipewyan, another Na Dene population (26) which groups with the South American Karitiana (Fig. S12B). Outgroup f2-statistics (46,47) and D-statistics (46,48) show that the high-coverage Saqqaq individual is closer to Greenlandic Inuit than to the Na Dene (Figs. S13A-C). Furthermore, the lack of support placing the Saqqaq closer to the Athabascans than to Karitiana (Fig. S13B) is incompatible with a scenario where Saqqaq and Na Dene share a fraction of their ancestries through a secondary Asian stream (26). Overall, our results support the Paleo-Eskimo migration into the Americas as being separate from that of the Na Dene Native Americans (Fig. 1C).

**Genetic affiliations of Neo-Eskimos**

Greenlandic and Canadian Neo-Eskimo Thule, and present-day Greenlandic Inuit form a clade in the maximum likelihood trees (Fig. 3A), even under admixture scenarios generated with TreeMix (43) (Fig. S14A-B). This supports genetic continuity over the last ca. 1,000 years between these populations, which is also evident by the shared mtDNA haplogroups (hgs A2a, A2b and D3a2a) between them [supplementary text S4 (2), Table S9A]. Furthermore, outgroup f2-statistics (46,47) and D-statistics (46,48) demonstrate that both Greenlandic and Thule are closer to present-day Inuit than to other sampled present-day populations or the high-coverage Saqqaq individual (Figs. 3B-C, S15A-C). These results are in agreement with the archaeological literature which suggests that present-day Greenlandic Inuit are direct descendants of the Thule (2,7,50-52) (Fig. 1C).

We additionally analyzed five ancient individuals dated to the 6th-7th century AD and associated with the Siberian Birnirk culture, which is part of the Neo-Eskimo tradition and may be one of the cultural ancestors of the Thule (53-55). Evidence from both mitochondrial typing (hg A2a) (Table S9A) and nuclear markers (Figs. S16A-D) reveal that these individuals are genetically closely related to present-day Greenlandic Inuit, providing the first genetic evidence of an Old World population that was not only a cultural precursor of the Thule, but also either closely related to or a component of the ancestral Inuit gene pool (Fig. 1C). We also reassessed claims of the Sadlermuit population from Southampton Island in the Hudson Bay region being
remnants of the Dorset culture on the basis of cultural similarities and mtDNA markers (56, 57). We typed ten Sadlermiut individuals, dating to the 15th-19th century, to mtDNA hgs A2b and D3a2a [supplementary text S4 (2), Tables S9A, S10], which are characteristic of the Thule/Inuit. Also, from their nuclear genome sequences, two of these individuals form a clade with present-day Greenlandic Inuit and not the high-coverage Saqqaq individual (Figs. 3A, S14A). This implies that the Sadlermiut were either genetically derived from or closely related to the Thule, rather than the Paleo-Eskimos (Fig. 1C).

Additionally, whether the Norse admixed with neighbouring Dorset or Thule groups remains debated, even though no anthropological evidence supporting such admixture has been found (58). In order to address this hypothesis, we tested 34 Norse individuals from southern Greenland, across the time span of their occupation in the region, for matrilineal admixture with neighboring Paleo- and Neo-Eskimo populations [supplementary text S1 (2)]. We found no New World-specific mitochondrial markers in these Norse samples, especially the Arctic-specific hgs A and D, suggesting that no detectable matrilineal gene flow occurred from the Dorset or Thule into the sampled Greenlandic Norse [supplementary text S4 (2), Tables S9B-C].

Admixture signals in Paleo- and Neo-Eskimos

We further investigated the observed genetic affinity between present-day Greenlandic Inuit and Paleo-Eskimos. D-statistics tests (46,48) and outgroup f-stats (46) support the Paleo-Eskimos as closer to Greenlandic Inuit than to other present-day New World populations (Figs. S13A, S17A-B), with the exception of the Aleutian Islanders whose genetic proximity to the Inuit is evident in analyses from both sequencing data (Fig. S10) and SNP chip data (26) and, Naukan who have been shown to possess Eskimo-Aleut ancestry as a result of back-migration of a related population (26). Maximum likelihood trees with migration edges inferred with TreeMix show evidence for admixture between Paleo-Eskimos and Greenlandic Inuit (Fig. 4A), mediated by the ancient Neo-Eskimos which include the Canadian Thule, the Greenlandic Thule and the Siberian Birnirk (Figs. S9A, S9C-D, S14A-B, S16A). In all the cases, this gene flow event is among the first few migration edges to be inferred by TreeMix (number of migration edges, m = 1-4). Bootstrap support suggested that gene flow likely occurred in both directions between the various Paleo- and Neo-Eskimo groups (Table S13).

While evidence for gene flow events was also observed in the high-coverage Saqqaq individual, we wanted to test whether this was due to TreeMix being unable to distinguish between Saqqaq and Dorset due to their genetic similarity and, if admixture might have actually occurred between Dorset and Neo-Eskimos. Therefore, we generated simulated datasets with varying levels of admixture (10% and 25%) between Dorset and Inuit in both directions, and determined whether a migration edge was observed in the TreeMix admixture graphs between the high-coverage Saqqaq individual and the Inuit in the absence of the Dorset [supplementary text S5 (2)]. Simulated admixture from Dorset to Inuit was detected by the analysis in the form of gene flow between Saqqaq and Inuit, but the reverse, that is, from Inuit to Dorset through gene flow involving Saqqaq and Inuit, was less likely (Figs. S18A-B, Table S13). Since our data show support for gene flow from the various Neo-Eskimo groups, including present-day Inuit, into the high-coverage Saqqaq individual, it is likely that the admixture involving Saqqaq is not an artifact and occurred from the Neo-Eskimo lineage into Saqqaq. We cannot, however, exclude the possibility of gene flow also from Paleo-Eskimos into Neo-Eskimos, or that subsequent gene flow did not occur between the later Dorset phases and the Neo-Eskimo lineage (Table S13).

Since the observed admixture involves the 4,000-year-old Saqqaq individual, it implies that the meeting and intermixing of Paleo-Eskimo and Neo-Eskimo ancestors predates the first archaeological observations of their co-existence in the eastern Arctic by ca. 3,000 years or earlier, since Thule groups entered eastern Canada and Greenland, previously occupied by the Late Dorset, only in the 12th or 13th century AD. This also suggests that the admixture is more likely to have occurred either in the Old World prior to the entry of the Neo-Eskimos into the Americas, or in Beringia, but not further east in the New World since no
archaeological evidence of a parallel existence of Paleo- and Neo-Eskimos around 4,000 years ago or earlier has been documented in this region (Fig. 1C).

We also investigated the level of genetic contribution from the ~24,000-year-old Siberian boy from Mal’ta (MA-1) (47) into Paleo-Eskimos and Greenlandic Inuit. Pairwise outgroup f$_{ST}$-statistics tests indicate that sampled western Eurasians and MA-1 are slightly closer to the high-coverage Saqqaq than to the Han Chinese (47) (Figs. 4B, S19). To confirm that this genetic affinity between Saqqaq and MA-1 was not due to ancient DNA bias, we also included the ~7,000-year-old La Brana 1 sample from Spain (59) in our analysis and found that it clustered close to present-day Europeans, hence confirming that our results were not skewed by ancient DNA attraction (Fig. 4B). TreeMix predicts a gene flow event from the high-coverage Saqqaq into MA-1 and vice-versa, consistent with (60), but with low support to substantiate both this signal and the inferred direction [Fig. 4C, Table S13, supplementary text S5 (2)].

In contrast, gene flow from MA-1 into the root of the clade comprising Native American populations (Karitiana and Anzick-1) and including the Greenlandic Inuit is detected with high bootstrap support, in agreement with (47) (Figs. 1B, 4C, Table S13). Additionally, D-statistics (46) and outgroup f$_{ST}$-statistics (46,47) from SNP chip data show that MA-1 is significantly (no overlap at three standard errors) closer to the Chipewyan than to the high-coverage Saqqaq (Fig. S20), which is compatible with other results presented in this study that also reject a single wave model for Saqqaq and Na Dene. Overall, while there is evidence of genetic affinity between MA-1 and Saqqaq, we are unable to ascertain whether this is due to gene flow or shared ancestry between the two lineages. It is also possible that this affinity is a consequence of the aforementioned gene flow from Neo-Eskimos, who received MA-1 gene flow, into the Saqqaq lineage.

Discussion

We overcome the difficulties of studying the peopling of the Arctic by including an extensive collection of Paleo-Eskimo remains for both ancient DNA and radiocarbon analyses. We have shown that Paleo-Eskimos likely represent a single migration pulse into North America from Siberia; separate from the migration events giving rise to Native Americans and Inuit. However, while being genetically distinct from other New World populations, Paleo-Eskimos are still more closely related to these populations than to non-New World populations, which supports the Beringian Standstill model (Tamm et al.) stating that a single ancestral population gave rise to many sub-populations, and possibly many migration pulses. Moreover, while our data are in agreement with Reich et al. (26), we find no support for Saqqaq or the rest of the Paleo-Eskimo tradition being a part of one of the two waves of Native American ancestors entering the more southern regions of the Americas. Therefore, an additional Paleo-Eskimo migration wave should be added to the three-wave hypothesis in explaining the peopling of the Americas (26).

Furthermore, Paleo-Eskimo and Inuit peoples appear to have occupied the New World Arctic for more than 4,000 years, with only a single population replacement (Thule) less than 700 years ago. In contrast with the dynamic responses of the Thule people to climate change (19,61,62), Siberian iron trade (64) and Norse contact (14), the 4,000-year Paleo-Eskimo period presents a single tradition of continuous technological and social development, including geographical dislocations and periods of relative stability punctuated by episodes of rapid change (16). The long-term continuity of the Paleo-Eskimo population and its culture is especially striking given the climatic and ecological cycles over 4,000 years - from warm late hypsithermal to cold sub-boreal through the early medieval warm event, as well as changes in sea ice distribution, changes in animal population cycles and distribution (11). In light of this, Paleo-Eskimo survival must be due to a remarkable resilience among small, dispersed local groups with the ability to shift their small population units to new areas when their homes became untenable; like the complete de-population of Greenland and the abandonment of the Canadian High Arctic between ca. 1 – 700 AD (7). Such events argue for the re-formulation of the original ‘core area’ hypothesis to a network of regional ‘core areas’ that served as demographic reservoirs for repopulating areas abandoned due to climate change, animal population crashes, or human over-hunting.
This paper also contributes to the long-standing debate about the Dorset-Thule transition. Dorset harpoons, snow knives, snow houses, soapstone vessels and Dorset art in Thule sites are considered evidence of sustained contact and acculturation (McGhee 2000, 2009; Fitzhugh 2004; Friesen 2004; Friesen and Arnold 2008, Sutherland 2009; Appelt and Gulløv 2009). The re-dating of the Thule migration to ca. 1300 AD, coincident with the latest Dorset radiocarbon dates, reduces the possibility for contact and exchange and strengthens the case for isolation (Park 1993, 2000). However, despite our results showing evidence of early contact and admixture between the Paleo-Eskimo and Neo-Eskimo lineages, dating to at least ca. 4,000 years ago, we cannot preclude a scenario where subsequent gene flow might also have occurred between Late Dorset and Thule groups. A related result of this study is the Thule affiliation of the Sadlermiut individuals. This culture that went extinct in 1903 from European disease has long been considered Thule-acculturated Dorset people (Collins 1958; deLaguna 1947; Thomson 1988); genetic evidence now suggests they were Thule people who had somehow acquired Dorset stone technology (Park 1993; Rowley 1994). If Dorset and Thule overlapped chronologically and geographically, as it appears they did at least in some areas, what social mechanisms resulted in genetic isolation? So far, there is no evidence of massacres, although the spread of Norse diseases have been suggested (Agger and Maschner 2009). Similar questions can be raised with regard to lack of matrilineal gene flow between the Thule or Dorset and the Greenland Norse (Appelt and Gulløv 2009).

Our study contrasts with previous population-level genetic studies, such as those focusing on the introduction of agriculture in Europe (Neolithization), which found that population movements were the instigators of changes in culture and subsistence strategies (60,65-67). Paleo-Eskimo technological innovations seem to have occurred solely by movement of ideas within a single resident population. Hence, our findings suggest that caution is required when using cultural similarities and differences as proxies for population movements and migrations into new and dramatically different environments (36).

Methods

DNA from 169 ancient human bone, teeth and hair samples from Arctic Siberia, Alaska, Canada and Greenland was extracted and targeted for haplogroup diagnostic mitochondrial DNA markers, while a subset of 26 samples was converted into Illumina libraries and sequenced, using standard laboratory procedures [supplementary text S3 (2)]. Two present-day Greenlandic Inuit, two Nivkh, one Aleutian Islander and two Athabascans were genome sequenced with no objections from The National Committee on Health Research Ethics, Denmark (H-3-2012-FSP21) [supplementary text S2 (2)]. 27 samples were radiocarbon dated and corrected for marine reservoir offset [supplementary text S2 (2)]. Mitochondrial DNA contamination estimates were computed as noted in supplementary text S4 (2). Error rate analysis, ancient DNA damage analysis, multidimensional scaling analysis on SNP chip and sequencing data, NGSadmix analysis, ABBA-BABA tests on sequencing data, D-statistics and f$_r$-statistics tests on SNP chip and sequencing data, TreeMix analysis on SNP chip and sequencing data, neighbor-joining analysis, and, ancestry painting of the Aleutian genome were performed as described in supplementary text S5 (2).

Acknowledgements

Figure legends

Figure 1. Chronology, origins and continuity of Paleo-Eskimos and Neo-Eskimos. Support from genetic results presented in this study is indicated by ‘(S)’ and rejection by ‘(R)’. A) A chronological framework for the prehistoric cultures in the New World Arctic and northeast Siberia, based on a combination of screened radiocarbon dates of associated terrestrial materials, typological studies, and contexts (eg. 7,31,32,35,79,84,85). Fading colors symbolize uncertainties concerning the beginnings or ends of the archaeological cultures, due to plateaus or wiggles in the radiocarbon calibration curve or lack of data. Defined archaeological phases within a culture are separated by a white line. Dark reddish-brown towards the top of the figure indicates historical times. Cultural contexts from which samples included in this study arise are highlighted in yellow. B) A two-wave migration model into the New World Arctic, with continuity throughout the Paleo-Eskimo tradition, followed by the Neo-Eskimo migration, is supported. Black thunderbolt symbols represent genetic discontinuity. C) This schematic summarizes the proposed origins of Paleo- and Neo-Eskimos in the archaeological and genetic literature and their relationships with other ancient populations in the North American Arctic. Our data shows support for Paleo-Eskimos constituting a migration pulse into the Americas independent to those of other New World populations (scenario 1) (2,21-23). For reference, we show the maximal geographical distribution of the Paleo-Eskimos and Neo-Eskimos in the New World Arctic and far-east Siberia (7). Additionally, plotted are Paleo-Eskimo (Pre-Dorset, Saqqaq, Dorset), Thule, Birnirk and Norse sites from which samples in this study derive; for further information see Fig. S2 and Table S1.

Figure 2. Origins of Paleo-Eskimos and genetic continuity. A) Sequencing data-based maximum likelihood trees constructed with TreeMix (43) with the high-coverage Saqqaq (23), Middle Dorset and Late Dorset datasets, and 17 present-day and two ancient populations; for Greenlandic Inuit see supplementary text S5 (2). The scale bar represents ten times the average standard error (s.e.) of the values in the covariance matrix.
Residual matrices are shown in Figs. S9C-D. B) SNP chip data-based $D$-statistic tests of the form $D$(Yoruba, Dorset; X, Saqqaq), where X represents present-day American and Siberian populations from (26) (coloured centers represent populations divided by linguistic affiliation, see legend at bottom). Thick and thin lines represent one and three standard errors of the $D$-statistics, respectively. Middle and Late Dorset (left and right panels, respectively) are significantly (no overlap at three standard errors) closer to Saqqaq than to the sampled present-day populations. C) Heat map of the SNP chip data-based statistic $f_3$(Yoruba; Saqqaq, X), where X represents present-day worldwide non-African populations. The graded heat key (to the right of the image) represents the magnitude of the computed $f_3$-statistics.

Figure 3. Genetic affinities of the Neo-Eskimo Thule. A) Sequencing data-based maximum likelihood trees constructed with TreeMix (43) with the high-coverage Saqqaq (23), Canadian Thule and Greenlandic Thule datasets, and 17 present-day and two ancient populations. The scale bar represents ten times the average standard error (s.e.) of the values in the covariance matrix. Residual matrices are shown in Figure S14A-B. B) SNP chip data-based $D$-statistic tests of the form $D$(Yoruba, Thule; X, West Greenlanders), as in Fig. 2. Canadian and Greenlandic Thule (left and right panels, respectively) are significantly (no overlap at three standard errors) closer to present-day West Greenlandic Inuit than to other worldwide present-day populations. Similar results are observed with East Greenlandic Inuit in place of West Greenlandic Inuit (Fig. S15C). C) Heat maps of the SNP chip data-based statistic $f_3$(Yoruba; Thule, X), as in Fig. 2.

Figure 4. Admixture signals in Paleo-Eskimos and Greenlandic Inuit. A) Admixture graph (number of migration edges, $m = 3$) constructed with TreeMix (43), with the high-coverage Saqqaq (23), 17 present-day and two other ancient individuals. The scale bar represents ten times the average standard error (s.e.) of the values in the covariance matrix, and the migration weight represents the fraction of ancestry derived from the migration edge. Migration edges are observed between the Paleo-Eskimos and Greenlandic Inuit, shown here as gene flow from the high-coverage Saqqaq to the root of East and West Greenlandic Inuit, although gene flow in the other direction is also observed (Table S13). Admixture graphs with other Paleo-Eskimo and Neo-Eskimo populations, different numbers of migration edges and, residual matrices are shown in Figs. S9A-D, S14A-B and S16A. B) Biplot of SNP chip data-based statistic $f_3$(Yoruba; Saqqaq, X) versus $f_3$(Yoruba; Han, X), where X represents present-day worldwide non-African populations. Thick and thin errors bars represent one and three standard errors of the $f_3$-statistics, respectively. Western Eurasian populations (Europeans and, South and Central Asians) are shifted towards the high-coverage Saqqaq compared to the Han Chinese. La Brana was included in the analysis to evaluate potential ancient DNA bias between the Saqqaq and MA-1 datasets. C) Sequencing data-based admixture graph with 17 present-day and four ancient individuals. A known migration edge is inferred from MA-1 to the root of Native Americans and Inuit, but this gene flow event excludes the high-coverage Saqqaq individual. Admixture graphs with different number of migration edges and residual matrices are shown in Fig. S21, and bootstrapping results are shown in Table S12. See supplementary text S5 (2) regarding the migration edge from Saqqaq into MA-1.
References


60. P. Skoglund et al., Science, (10.1126/science.1253448).


72.


221. A. Albrechtsen et al., Relatedness mapping and tracts of relatedness for genome-wide data in the presence of linkage disequilibrium, Genetic Epidemiology 33, 266-274 (2009).

222. S. Purcell et al., PLINK: a tool set for whole-genome association and population-based linkage analyses, Am J Hum Genet 81, 559-575 (2007).


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


234. D. Reich et al., Genetic history of an archaic hominin group from Denisova Cave in Siberia, Nature 468, 1053-1060 (2010).


