LETTER

doi:10.1038/nature25173

Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans

J. Víctor Moreno-Mayar[1]*, Ben A. Potter[2]*, Lasse Vinner[1]*, Matthias Steinrücken[3,4], Simon Rasmussen[5], Jonathan Terhorst[6,7], John A. Kamm[6,8], Anders Albrechtsen[9], Anna-Sapfo Malaspinas[1,10,11], Martin Sikora[1], Joshua D. Reuther[2], Joel D. Irish[12], Ripan S. Malhi[13,14], Ludovic Orlando[1], Yun S. Song[6,15,16], Rasmus Nielsen[1,6,17], David J. Meltzer[1,18] & Eske Willerslev[1,8,19]

Despite broad agreement that the Americas were initially populated via Beringia, the land bridge that connected far northeast Asia with northwestern North America during the Pleistocene epoch, when and how the peopling of the Americas occurred remains unresolved[1–5]. Analyses of human remains from Late Pleistocene Alaska are important to resolving the timing and dispersal of these populations. The remains of two infants were recovered at Upward Sun River (USR), and have been dated to around 11.5 thousand years ago (ka)[6]. Here, by sequencing the USR1 genome to an average coverage of approximately 17 times, we show that USR1 is most closely related to Native Americans, but falls basal to all previously sequenced contemporary and ancient Native Americans[1,7,8]. As such, USR1 represents a distinct Ancient Beringian population. Using demographic modelling, we infer that the Ancient Beringian population and ancestors of other Native Americans descended from a single founding population that initially split from East Asians around 36 ± 1.5 ka, with gene flow persisting until around 25 ± 1.1 ka. Gene flow from ancient north Eurasians into all Native Americans took place 25–20 ka, with Ancient Beringians branching off around 22–18.1 ka. Our findings support a long-term genetic structure in ancestral Native Americans, consistent with the Beringian 'standstill model'[9]. We show that the basal northern and southern Native American branches, to which all other Native Americans belong, diverged around 17.5–14.6 ka, and that this probably occurred south of the North American ice sheets. We also show that after 11.5 ka, some of the northern Native American populations received gene flow from a Siberian population most closely related to Koryaks, but not Palaeo-Eskimos[1], Inuits or Kets[10], and that Native American gene flow into Inuits was through northern and not southern Native American groups[1]. Our findings further suggest that the far-northern North American presence of northern Native Americans is from a back migration that replaced or absorbed the initial founding population of Ancient Beringians.

The details of the peopling of the Americas, and particularly the population history of Beringia, remain unresolved[2,3]. Humans were present in the Americas south of the continental ice sheets by around 14.6 ka[11], indicating that they traversed Beringia earlier. During the Last Glacial Maximum (LGM), this region was marked by harsh climates and glacial barriers[5], which may have led to the isolation of populations for extended periods, and at times complicated dispersal across the region[12]. It remains unknown whether and for how long Native American ancestors were isolated from Asian groups in Beringia before entering the Americas[2,9,13]; whether one or more early migrations gave rise to the founding population of Native Americans[1–4,7,14] (it is commonly agreed that the Palaeo-Eskimos and Inuit populations represent separate and later migrations[1,15,16]); and when and where the basal split between southern and northern Native American (SNA and NNA, respectively) branches occurred. It also remains unresolved whether the genetic affinity between some SNA groups and indigenous Australasians[2,3] reflects migration by non-Native Americans[3,4,14], early population structure within the first Americans[3] or later gene flow[2]. To resolve these uncertainties, a better understanding of the population history of Beringia, the entryway for the Pleistocene peopling of the Americas, is needed.

Genomic insight into that population history has now become available with the recently recovered infant remains (USR1 and USR2) from the Upward Sun River site, Alaska (eastern Beringia), which have been dated to approximately 11.5 ka[6,17]. Mitochondrial DNA sequences (haplogroups C1 and B2, respectively) were previously acquired from these individuals[6,17] (Supplementary Information sections 1, 4.5). We have since obtained whole-genome sequence data, which provide a broader opportunity to investigate the number, source(s) and structure of the initial founding population(s) and the timing and location of their subsequent divergence. We sequenced the genome of USR1 to an average depth of approximately 17×, on the basis of eight sequencing libraries from uracil-specific excision reagent-treated extracts that had previously been confirmed to contain DNA fragments with characteristic ancient DNA misincorporation patterns (Supplementary Information sections 2–4). We estimated modern human contamination to be around 0.14% based on the nuclear genome and about 0.15% based on mitochondrial DNA (Supplementary Information section 4). As expected, the error rate in the uracil-specific excision reagent-treated sequencing data was low (0.09% errors per base), and comparable to other high-coverage contemporary genomes, based on called genotypes (Supplementary Information section 4). Although USR2[6] did not have sufficient endogenous DNA for high-coverage genome sequencing, we found that both individuals were close relatives (Supplementary Information section 5), equally related to worldwide present-day populations (Supplementary Fig. 4g).

We assessed the genetic relationship between USR1, a set of ancient genomes[2,7,8,14,16] and a panel of 167 worldwide populations genotyped for 199,285 single-nucleotide polymorphisms[1,2,18] (Supplementary

[1]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark. [2]Department of Anthropology, University of Alaska, Fairbanks, Alaska 99775, USA. [3]Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, Massachusetts 01003, USA. [4]Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA. [5]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kongens Lyngby, Denmark. [6]Department of Statistics, University of California, Berkeley, California 94720, USA. [7]Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, USA. [8]Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [9]The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark. [10]Department of Computational Biology, University of Lausanne, Lausanne, Switzerland. [11]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [12]Research Centre in Evolutionary Anthropology and Palaeoecology, Liverpool John Moores University, Liverpool L3 3AF, UK. [13]Department of Anthropology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. [14]Carle R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. [15]Computer Science Division, University of California, Berkeley, California 94720, USA. [16]Chan Zuckerberg Biohub, San Francisco, California 94158, USA. [17]Department of Integrative Biology, University of California, Berkeley, California 94720, USA. [18]Department of Anthropology, Southern Methodist University, Dallas, Texas 75275, USA. [19]Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.
*These authors contributed equally to this work.