METHODS

Laboratory procedures. Ancient DNA work was conducted in dedicated clean laboratory facilities at the Centre for GeoGenetics, Natural History Museum, University of Copenhagen. We prepared bone powder from remains of the pars petrosa of both USR individuals and extracted DNA following previously published protocols[31]. Double-stranded dual-indexed Illumina libraries were built from uracil-specific excision reagent (USER)- and non-USER-treated extracts and were paired-end sequenced (2×75 bp) on Illumina HiSeq 2500 instruments (Supplementary Information section 2).

Sequence data processing. Raw reads were trimmed for Illumina adaptor sequences and overlapping pairs were collapsed into single reads using AdapterRemoval[32]. Collapsed reads were mapped to the human reference genome build 37 using BWA v.0.6.2-r126[33]; seeding (*-l* parameter) was disabled in order to prevent 5′ terminal substitutions characteristic of ancient DNA to bias the mapping[34]. Reads with mapping quality lower than 30 were discarded, PCR duplicates were removed using MarkDuplicates (http://picard.sourceforge.net) and local realignment was performed using GATK[35]. We called USR1 genotypes using SAMtools mpileup[36] and applied the standard filters described in ref. 2. Called genotypes were phased with shapeit2-r727[37] using the 1,000 genomes phased variant panel (phase 3) as a reference and the HapMap recombination rates as a proxy for the genetic map of the human genome. Sites not included in the 1,000 genomes reference panel were kept as 'unphased' genotypes. Finally, we masked the dataset using a 35-mer 'snpability' mask with a stringency of 0.5 (http://lh3lh3.users.sourceforge.net/snpable.shtml) (Supplementary Information section 3).

Ancient DNA data authentication. We assessed the authenticity of the ancient DNA data by examining the fragment length distributions and the base substitution patterns across non-USER-treated reads using bamdamage[22]. We estimated mtDNA contamination using contamMix[38] on the basis of a majority rule mtDNA consensus sequence and an alignment of 311 worldwide mtDNA sequences[39]. Nuclear contamination was estimated using the two-population model implemented in DICE[40], for which we used the 1,000 Genomes Project 'CEU' population as the putative contaminant and the 'YRI' population as the 'anchor'. Sequencing and genotyping error rates relative to a 'high-quality' sample were obtained following the method described in ref. 41 (Supplementary Information section 4).

Relatedness between USR individuals. We explored the familial relationship between both USR individuals by using NGSrelate[42] and relate[43]. Given the unavailability of allele frequency data for the Ancient Beringian population, we used allele frequencies from the 1,000 Genomes Project 'PEL' population as a proxy, which limited the resolution of these analyses (Supplementary Information section 5).

Reference datasets. We compared the genomes of the USR individuals to a set of 49 worldwide contemporary and ancient genomes and a SNP array dataset comprising 2,537 contemporary individuals from 167 ethnic groups (enriched in Native Americans), genotyped across 199,285 SNP sites. For the latter, European and African ancestry tracts were masked in Native American individuals (Supplementary Information 6).

Population structure analyses. We investigated the relationship between USR1, a set of ancient genomes and the SNP array reference dataset using multi-dimensional scaling as implemented in bammds[22]. Additionally, we explored the genetic ancestry components in the reference panel using ADMIXTURE[20]. We obtained the most likely ancestry proportions in the ancient genomes on the basis of allele frequencies inferred by ADMIXTURE, through the genotype likelihood-based optimization method described in ref. 21 (Supplementary Information sections 7, 8).

***f* statistics.** We computed f_3 statistics to measure the shared drift between by two particular populations or genomes, and used 'basic' and 'enhanced' *D* statistics to formally test hypotheses of treeness and gene-flow. We used admixtools[19] for allele-frequency-based tests and ANGSD[44] for single genome tests. For both tools, standard errors were estimated through a weighted block jackknife approach over approximately 5-Mb blocks (Supplementary Information sections 9–13).

Admixture graph fitting using TreeMix. We used the heuristic approach in TreeMix[23] to assess the phylogenetic placement of USR1 in the broader context of Eurasian and Native American populations and to explore the origin of the Na-Dene and Inuit (see 'Admixture graph fitting using qpGraph'). We restricted the analysis to transversion sites where all considered populations have at least one individual with a non-missing genotype call. We grouped the resulting number of SNPs into approximately 5-Mb blocks to account for linkage disequilibrium, and for each number of migrations we ran 1,000 replicates with random seeds and kept the run with the highest likelihood. We estimated the support for internal nodes and migration edges through a bootstrap procedure (Supplementary Information sections 14, 17).

Pairwise branch lengths and genomic divergence. We used the method from ref. 7 to measure the amount of drift leading to different pairs of genomes after their split. We restricted this analysis to sites that are variable in five African genomes and obtained the counts for each of the five possible genotype configurations between a given pair of genomes, after which we used numerical optimization to infer maximum likelihood parameters (Supplementary Information section 15). We computed the average DNA divergence between pairs of genomes using the triangulation method from ref. 45, and estimated standard errors using a weighted block jackknife approach over 5-Mb blocks (Supplementary Information section 16).

Admixture graph fitting using qpGraph. We used a two-step approach to assess the origin of the Na-Dene and Inuit. First, we found the most likely Eurasian ancestry sources for these groups by using TreeMix. We then fitted *f*-statistics-based admixture graphs[1,19] incrementally, such that for each new 'admixed leaf' we enumerated all possible pairs of edges using ref. 46 and kept the admixture event that produced the graph with the best maximum $|Z|$ and fitting scores. We assessed the robustness of this model and its predictions using pooled *D* statistics and by fitting the model using alternative datasets (Supplementary Information section 17).

Demographic inference using the sequentially Markov coalescent. We used diCal2[27] to estimate the key demographic parameters relating pairs of genomes including USR1 (sample dated to 11.5 ka) and a set of present-day Asian and Native American genomes. We analysed these pairs under different models, including a clean split, isolation with migration until the present, isolation with migration with a stopping time and isolation with migration with a stopping time and a second contact. We tested competing models through a simulation study and obtained confidence intervals for the inferred parameters through a parametric bootstrap strategy (Supplementary Information section 18).

Demographic inference using the site frequency spectrum. We used a combination of SMC++[29] and momi2[28] to infer demographic parameters for USR1 and a set of present-day genomes. We estimated the marginal sizes over time for each population using SMC++. We used these demographic histories as a basis for fitting a joint 'backbone demography' for the present-day populations using momi2. We then inferred the most likely join-on point for USR1 onto the backbone demography using momi2. Confidence intervals were obtained through a parametric bootstrap strategy (Supplementary Information section 19).

Data availability. Sequence data were deposited in the ENA under accession: PRJEB20398.

31. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522,** 167–172 (2015).
32. Lindgreen, S. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res. Notes* **5,** 337 (2012).
33. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
34. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13,** 178 (2012).
35. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43,** 491–498 (2011).
36. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
37. Delaneau, O., Zagury, J.-F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10,** 5–6 (2013).
38. Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23,** 553–559 (2013).
39. Green, R. E. *et al.* A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134,** 416–426 (2008).
40. Racimo, F., Renaud, G. & Slatkin, M. Joint estimation of contamination, error and demography for nuclear DNA from ancient humans. *PLoS Genet.* **12,** e1005972 (2016).
41. Orlando, L. *et al.* Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499,** 74–78 (2013).
42. Korneliussen, T. S. & Moltke, I. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics* **31,** 4009–4011 (2015).
43. Albrechtsen, A. *et al.* Relatedness mapping and tracts of relatedness for genome-wide data in the presence of linkage disequilibrium. *Genet. Epidemiol.* **33,** 266–274 (2009).
44. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15,** 356 (2014).
45. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328,** 710–722 (2010).
46. Leppälä, K., Nielsen, S. V. & Mailund, T. admixturegraph: an R package for admixture graph manipulation and fitting. *Bioinformatics* **33,** 1738–1740 (2017)