

The genomic formation of First American ancestors in East and Northeast Asia

Authors: Chao Ning^{1,2,3,†}, Daniel Fernandes^{4,5,†}, Piya Changmai^{6,†}, Olga Flegontova^{6,7,†}, Eren Yüncü^{6,†}, Robert Maier^{8,9}, N. Ezgi Altınışık^{6,10}, Alexei S. Kassian¹¹, Johannes Krause³, Carles Lalueza-Fox¹², Andrea Manica¹³, Ben A. Potter¹⁴, Martine Robbeets³, Kendra Sirak^{4,15,#}, Veronika Siska¹³, Edward J. Vajda¹⁶, Leonid A. Vyazov¹⁷, Ke Wang³, Lixin Wang², Xiyun Wu¹, Xiaoming Xiao¹⁸, Fan Zhang¹, David Reich^{8,9,19,20}, Stephan Schiffels³, Ron Pinhasi^{4,21,&,*}, Yinqiu Cui^{1,2,&,*}, Pavel Flegontov^{6,8,&,*}

Affiliations:

- 10 ¹ Research Center for Chinese Frontier Archaeology, Jilin University, Changchun, China.
- ² School of Life Sciences, Jilin University, Changchun, China.
- ³ Department of Archaeogenetics and Eurasia3angle research group, Max Planck Institute for the Science of Human History, Jena, Germany.
- ⁴ Department of Anthropology, University of Vienna, Vienna, Austria.
- 15 ⁵ CIAS, Department of Life Sciences, University of Coimbra, Coimbra, Portugal.
- ⁶ Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czechia.
- ⁷ Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czechia.
- 20 ⁸ Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, United States of America.
- ⁹ Broad Institute of MIT and Harvard, Cambridge, MA, United States of America
- ¹⁰ Human-G Laboratory, Department of Anthropology, Hacettepe University, Ankara, Turkey
- ¹¹ Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia.
- 25 ¹² Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain.
- ¹³ Department of Zoology, University of Cambridge, Cambridge, United Kingdom.
- ¹⁴ Arctic Studies Centre, Liaocheng University, Liaocheng, China.
- ¹⁵ Department of Anthropology, Emory University, Atlanta, United States of America.
- 30 ¹⁶ Department of Modern and Classical Languages, Western Washington University, Bellingham, WA, United States of America.
- ¹⁷ Kazan Federal University, Kazan, Tatarstan, Russia.
- ¹⁸ Department of Archaeology, Liaoning University, Shenyang, China.
- ¹⁹ Department of Genetics, Harvard Medical School, Boston, MA, United States of America

²⁰ Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, United States of America

²¹ School of Archaeology, and Earth Institute, University College Dublin, Belfield, Ireland.

5

† These authors contributed equally

‡ Present address: Department Anthropology, Hacettepe University, Ankara, Turkey.

Present address: Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, United States of America.

10

& Co-senior authors

*Correspondence to: Pavel Flegontov (pavel.flegontov@osu.cz, pavel_flegontov@hms.harvard.edu), Yinqiu Cui (cuiyq@jlu.edu.cn), and Ron Pinhasi (ron.pinhasi@univie.ac.at).

15

Abstract: Upward Sun River 1, an individual from a unique burial of the Denali tradition in Alaska (11500 calBP), is considered a type representative of Ancient Beringians who split from other First Americans 22000–18000 calBP in Beringia. Using a new admixture graph model-comparison approach resistant to overfitting, we show that Ancient Beringians do not form the deepest American lineage, but instead harbor ancestry from a lineage more closely related to northern North Americans than to southern North Americans. Ancient Beringians also harbor substantial admixture from a lineage that did not contribute to other Native Americans: Amur River Basin populations represented by a newly reported site in northeastern China. Relying on these results, we propose a new model for the genomic formation of First American ancestors in Asia.

20

25

One Sentence Summary: Ancient Beringians do not form the deepest American lineage, but harbor admixture from Amur River Basin populations.

Main Text:

The peopling of America remains an area of active debate in archaeogenetics (1, 2): the time and place of the "Beringian standstill" (3-6), dates and routes of the first colonization (5, 7-9), and the Asian source for the later migration about 5000 cal yr BP (4, 10, 11) are all contentious. Two recent discoveries are critical for constraining models of American settlement (2). First, the Upward Sun River 1 individual (USR1) from a unique double infant burial of the Denali tradition dated to ca. 11500 cal yr BP (12) was subjected to deep (17×) whole-genome sequencing and was shown to represent a previously unknown genetic lineage (Ancient Beringians) that split from the American stem between 22000 and 18000 cal yr BP (4), much earlier than the two major clades known before, the northern Native Americans (NNA) and southern Native Americans (SNA) that split between 17500 and 14600 cal yr BP (13-19). The region where the divergence of Ancient Beringians took place remains unknown, but Moreno-Mayar *et al.* (4) favored a model where the divergence occurred in Alaska, with the northern and southern American clades diverging south of the ice sheets. Other genomes relevant for reconstructing the earliest history of proto-Americans in Asia belong to an individual from the Duvanny Yar site (9800 cal yr BP), Kolyma River, northeastern Siberia (10) and an older individual from the Ust'-Kiyakhta-3 site (14000 cal yr BP) in the Baikal region (20). Both the Kolyma1 and Ust'-Kiyakhta individuals were modeled as proto-American lineages diverging deeper than USR1 and having Siberian admixture of 21-27% (20). According to Sikora *et al.* (10), this Paleo-Siberian population represented by the Kolyma1 individual was later largely replaced by the Neo-Siberians, with individuals from a 7,600-year-old site at Devil's Gate Cave in the Russian Far East being their type representatives (10, 21). The Kolyma1-related genetic component now is prominent only among Chukotkan, Kamchatkan and some West Siberian indigenous groups (Kets, Selkups, Khanty, Mansi), and in ancient Paleo-Inuit (10).

To date, our knowledge about prehistoric populations from mainland East Asia is relatively limited. Six Early Neolithic individuals (~7600 cal yr BP) from Devil's Gate Cave in the Amur River Basin (ARB) were sequenced at coverage ranging from 0.1 to 6.6× and were shown to be genetically continuous with present-day populations in the ARB and with a majority of present-day Siberians (10, 21). Other genetically similar individuals from the ARB were reported by Ning *et al.* (22) and Wang *et al.* (23): 28 individuals dated between ca. 7400 cal yr BP and 11th-13th centuries CE. To investigate the prehistory of the ARB deeper in time, we generated whole-genome shotgun data from 12 individuals from the Houtaomuga site in Northeast China, who lived between 12000 and 2300 cal yr BP. The oldest individual dated to ~12000 cal yr BP was sequenced to a high coverage (45×) and is the earliest sequenced representative of the ARB genetic cluster.

We developed a new method for admixture inference based on exhaustive "mapping" of a target group on a "skeleton" admixture graph as a mixture of an increasing number of sources and another method that searches the graph space automatically. To allow comparison of less graph models of different complexity, we undertook a study on simulated genetic data (Supplementary Texts 1 and 2) and developed a new model-comparison approach resistant to overfitting. We utilized the new data from the Houtaomuga site for constraining admixture proportions in a graph connecting key Siberian, East Asian, and American lineages. Leveraging the power of that admixture graph and the novel methods, we demonstrate that the USR1 lineage descends from admixture between NNA and ARB populations represented by the Houtaomuga group, questioning the monoclality of USR1 with Native Americans suggested previously (4).

Results

An archaeogenetic transect at the Houtaomuga site

5 An initial sample screening step was performed on 22 ancient individuals from the Houtaomuga site, in the Jilin province of China (Fig. 1, Supplementary Text 3). Based on endogenous ancient DNA yields and authentication (see Supplementary Text 4), the libraries of 12 individuals (two Early Neolithic (EN), four Middle Neolithic (MN) and six Late Bronze Age/Early Iron Age) were sequenced to an adequate coverage (0.1-45x) on the Illumina HiSeq X10 platform (Data S3). The genomic coverages obtained ranged between 0.51–1.40-fold for the Early Neolithic-2, Middle Neolithic, and Iron Age specimens, while a ~12000-year-old Early Neolithic-1 individual was sequenced to a coverage of 45-fold from a UDG-treated library. Deamination frequencies on the 5' and 3' ends of the sequences of these 12 individuals ranged from 6.8 to 40.7%, values typical of aDNA (Data S3).

15 Based on contamination estimates, some samples were not included into the reference Houtaomuga group used for admixture graph analyses, leaving a final set of six (Fig. 1, Data S3). Among these six Houtaomuga individuals, three (M45, M54A, M74B) had no detectable West Eurasian admixture according to an analysis based on a simple admixture graph (Supplementary Text 5). We leveraged this feature of Houtaomuga individuals for constructing an admixture graph uniting diverse Siberian lineages admixed with various West Eurasian groups (Supplementary Text 5). For all the downstream graph-based analyses, we created two alternative Houtaomuga groups with pseudohaploid genotype calls: 1/ M45 and M54A; 2/ those two EN individuals and an Iron Age individual M74B.

20 According to a principal component analysis (PCA), all nine Houtaomuga individuals passing a missing rate cut-off clustered with other Early Neolithic ARB individuals from the Devil's Gate Cave and Boisman sites (10, 23) and were close to a cluster composed of present-day Tungusic and Nivkh speakers in the ARB (Fig. S41). These results suggest long-term genetic continuity in the ARB, in line with previous studies (10, 22, 23).

A new admixture cline between Asians and Americans

30 We assembled a dataset based on a panel of 1.24 million SNPs across the human genome (24, 25) and generated two versions: a version without transition polymorphisms including 208,649 variable autosomal sites and a full version including 1,062,979 sites (Supplementary Text 5). In the PC1/PC2 space both USR1 and Kolyma1 individuals lie between the American and East Asian clusters, with Kolyma1 located much closer to the Asian cluster (Fig. 2). This result remained unchanged when different dataset versions and PCA protocols were used (Supplementary Text 6, Fig. S40, Fig. S51), and the same results were reported before (4, 10, 20). As shown for the simulated USR1 individual, her position in the PC space cannot be interpreted as evidence for or against Siberian admixture since a similar position was reproduced in the absence of Siberian-American gene flow (Fig. 2, Supplementary Text 2). Whole diploid human genomes without missing data were simulated using *msprime* (26) according to the admixture graph shown in Fig. S43c (see also Fig. S21, Table S8, Supplementary Text 2).

40 Both USR1 and Kolyma1 individuals were modelled to date as mixtures of a proto-American lineage (abbreviated as FAM_{EA} below) and a West Eurasian lineage (4, 10). The Kolyma1

individual was alternatively modelled as a mixture of a proto-American, a Siberian, and a West Eurasian lineage, and the same model was proposed for the Ust'-Kyakhta individual (20). To check if these simple models fit the data, we applied the newly developed admixture inference protocol based on admixture graphs to these key individuals. We constructed a skeleton admixture graph for Siberians, East Asians, and Americans, as described in Supplementary Text 5. Building a graph for Siberians is difficult because most groups have a substantial amount of West Eurasian ancestry coming from different sources and probably in multiple waves (27). We decided to simplify the graph and used distal sources to account for almost any type of West Eurasian ancestry: Mal'ta-like, West European hunter-gatherers (WHG), and the basal Eurasian ghost source (28). We added these three gene flows by default to each Siberian branch on the skeleton graph. Two similar versions of the skeleton graph were constructed, including 14 groups on the dataset without transitions (Fig. 3a,b, Fig. S37) or 16 groups on the complete dataset (Fig. S39).

Although the skeleton graph presented here may be a simplification of the real situation, it fits the data (Supplementary Text 7d), and we believe it is reasonably powerful for inferring admixture in Asians and Americans as demonstrated on genetic data simulated using *msprime* (Supplementary Text 2) and on positive controls (ancient and present-day Athabaskan and Eskimo-Aleut speakers, and 10,000 years-old individuals from Brazil) described in Supplementary Text 8. Moreover, the same branching order of Asian lineages (Onge, Tianyuan, Jomon, Tibetans, Southeast Asians, Americans and Siberians, see Table S16) was inferred using a different approach (*TreeMix*) (29).

We mapped USR1 (Supplementary Text 7), Kolyma1, and Ust'-Kyakhta individuals (Supplementary Text 9) on both skeleton graph versions. A sister group of FAM_{EA} and a Mal'ta-related source were assigned as default ancestry sources for the USR1 individual (4), and a Mal'ta-related source was added by default to Kolyma1 and Ust'-Kyakhta following previous results (10, 20). We added one or two additional ancestry sources to each target, testing all possible source combinations within the clade of anatomically modern humans. In the case of USR1 (Data S4a, S4b), only one model in each class was significantly different from others according to the likelihood cut-off of 3 ln-units (11), and each additional gene flow resulted in a log-likelihood (LL) difference and/or difference in the worst model residuals (WR) that was associated with a false positive (FP) rate below 5% on data simulated according to the 14-population skeleton graph (Fig. 3d, Table S18, Supplementary Text 2). The mapping protocol was repeated under eight or four setting combinations as detailed in Supplementary Text 7, showing consistent results.

A final model for USR1 was as follows: FAM_{EA} + a Mal'ta source that for simplicity contributes to all American and Siberian groups (Figs. S37, S39) + a distinct Mal'ta-related lineage (termed "basal Mal'ta" for convenience) + a Houtaomuga-related lineage. Exploration of the likelihood space of admixture proportions has put a lower bound on the Houtaomuga-related ancestry in USR1 at ~6% and the upper bound at 35% (Data S4e, S5). We note that an f_4 -statistic (First Americans, USR1; Southeast Asians, Houtaomuga) deviating from 0 by 3 SE points to attraction between USR1 and Houtaomuga (Table S17). Similar D -statistics (a First American group, USR1; a Siberian group, Han) showed the same signal in the original report, but since absolute Z -scores remained below 3.3 the signal was interpreted as a "slight residual affinity" (4).

The best model obtained for the Kolyma1 individual was Houtaomuga + FAM_{EA} + Mal'ta (Data S7a, S7b). The Houtaomuga-like ancestry proportion in Kolyma1 ranged from 37% to 55% (across four setting combinations and both diploid and pseudohaploid genome versions), and the FAM_{EA} ancestry proportion ranged from 19% to 38% (Supplementary Text 9). This result broadly agrees with the previous report (20) where 21% ARB ancestry was inferred in Kolyma1. We support the main result by Yu *et al.* (20), namely that the Ust'-Kyakhta individual represents a divergent proto-American lineage, but we are able to model this individual in a simpler way, without the Siberian gene flow (Supplementary Text 9). A low-level (~6%) gene flow from Kolyma1 or Ust'-Kyakhta to the Saqqaq lineage has received statistical support according to our criteria, although the direction of the gene flow was impossible to resolve (Supplementary Text 9). In contrast, the direction of the Houtaomuga-USR1 gene flow was well-resolved (Supplementary Text 7). The Kolyma1-Saqqaq relationship was reported before (10) and is entirely plausible geographically and chronologically.

Next, we performed an exhaustive exploration of *qpAdm* models for USR1, FAM, and Houtaomuga on the full 1240K-derived dataset, in the context of 12 other groups included into the simpler skeleton graph version (Supplementary Text 10). Using *qpWave*, we tested the cladality of USR1 and various FAM groups: a group of present-day SNA individuals (11) used for most admixture graph analyses in this study, 11 alternative present-day SNA groups, and 45 alternative groups composed of ancient SNA and NNA individuals (13, 14, 18). The cladality of First Americans and USR1 was rejected for 29 to 33 of 57 groups, depending on settings (*p*-values ranged from 6×10^{-7} to 0.01, Data S8a). The FP rate of this test on data simulated under the 14-population skeleton graph was 7% in the presence of ascertainment bias (*p*-value cut-off = 0.01, 100 datasets simulated using *msprime*, 400,000 randomly sampled sites), and <2% in the absence of ascertainment bias. To determine if the FAM or USR1 group is responsible for the cladality violation, we tested all possible 2- and 3-way admixed models for USR1, the composite FAM group, and Houtaomuga as targets. Among 273 2-way models tested for these targets using *qpAdm* with the "rotating" outgroups setup, no model satisfied the following conditions (30): *p*-value > 0.01 and inferred admixture proportions ± 2 SE lie between 0 and 1 for at least one setup. Among 1,092 3-way models tested for the three targets, just one model satisfied these conditions: USR1 = $45 \pm 20\%$ FAM + $35 \pm 13\%$ Houtaomuga + $20 \pm 7\%$ Mal'ta (Data S8b-d). The FP rate of this *qpWave/qpAdm* procedure on data simulated using *msprime* was 5% in the presence of ascertainment bias and <2% in the absence of ascertainment bias. In other words, for 5 out of simulated 100 datasets i/ the cladality of FAM and USR1 was rejected and ii/ only one of three alternative targets was successfully modeled as a 2- or 3-way mixture according to the criteria outlined above.

Finally, we applied a new model-comparison approach resistant to overfitting (see Materials and Methods) to the best graph models selected above for USR1. The baseline model for USR1 (4) was compared to 3-way models "FAM_{EA} + Mal'ta + basal Mal'ta" and "FAM_{EA} + Mal'ta + Houtaomuga", and those were compared to the 4-way model "FAM_{EA} + Mal'ta + basal Mal'ta + Houtaomuga" (Supplementary Text 11). Here we focus on two model comparisons where the Houtaomuga-related ancestry source was added. When comparing distributions of out-of-sample LL differences across 500 bootstrap replicates (Fig. 3c, Fig. S49), the "FAM_{EA} + Mal'ta + Houtaomuga" model fitted the data significantly better (*p*=0.032) than the baseline model for the pseudohaploid USR1 version (Fig. 3c,e). The magnitude of the signal observed on real data in the context of signals on simulated data (Fig. 3e, Table S20) is consistent with the constraints on

the ARB ancestry proportions in USR1 (6 to 35%) found by exploring the admixture graph likelihood space (Data S4e, S5).

5 The gene flow that we revealed in the USR1 lineage challenges previous results which show that the First American source of USR1 forms the deepest lineage in the American clade that bifurcated ca. 20000 years ago in Beringia (4). We tested all possible topologies for USR1, an ancient Athabaskan individual (I5321) admixed with Paleo-Inuit (11) and representing the NNA clade, a Clovis individual (Anzick) (16), and an ancient Peruvian individual (IL7) demonstrating no signal of extra Asian admixture (Data S6). The latter two individuals represented the SNA clade. This American clade with all the necessary Asian and West Eurasian gene flows was grafted on the skeleton graph, and 15 possible topologies were tested under 48
10 setting/dataset/skeleton graph combinations (Supplementary Text 7g, Data S4f, S4g). This model competition has demonstrated that the basal position of the USR1 source was either significantly rejected under all setting combinations or, if it was not rejected, it invariably resulted in a trifurcation in the American clade. The most supported topology that also almost never produced trifurcations was ((I5321, USR1), (Clovis, IL7)). Thus, the American source of USR1 probably belongs to the NNA clade, in agreement with the geographic location in Alaska. A similar topology exploration incorporating the Kolyma1 and Ust'-Kyakhta individuals revealed the following topology as a "winner" having no trifurcations in the proto-American clade: ((Ust'-
15 Kyakhta, Kolyma1), ((I5321, USR1), IL7)) (Supplementary Text 9).

20 When constructing the skeleton graphs (Supplementary Text 5) and conducting mapping and branching order exploration (Supplementary Texts 7, 8, 9), we used certain conventions that helped to standardize the modelling process, but those conventions make the resulting models unrealistic in certain ways. West Eurasian gene flows added to Siberian and American lineages were always right below the leaf nodes, and they were added separately to each Siberian and
25 American branch. For instance, there was no Mal'ta-related gene flow into the common ancestor of First Americans and USR1. To check if the results generated using these skeleton graphs hold for more realistic topologies, we performed another model comparison. We showed that the ARB=>USR1 gene flow is also supported in the context of a graph incorporating separate NNA and SNA lineages and a Mal'ta-related gene flow into the common ancestor of Americans
30 (Supplementary Text 12). When all the necessary extra-American gene flows were incorporated, all three alternative branching orders in the American clade displayed nearly identical model likelihoods, and the branching order (proto-USR1, (proto-Athabaskan, SNA)) again invariably resulted in a trifurcation, in contrast to the other branching orders (Supplementary Text 12). The resulting best model (SNA, (proto-Athabaskan, USR1)) is shown in Fig. 3a.

35 To check if our main result is robust to more profound changes in the analytical setup, we constructed a dataset that is expected to be free of most types of bias, such as: 1/ uneven deamination damage across ancient DNA samples that leads to ancient DNA attraction; 2/
reference bias (Günther and Nettelblad 2019), 3/ ascertainment bias caused by the paucity of rare variants (Supplementary Texts 2 and 13) that is typical for SNP panels (31); 4/ background
40 selection and biased gene conversion which possibly affect 97% of variable sites in the human genome (32); 5/ various levels of missing data across groups. Using the new set of 226,472 G↔C and A↔T polymorphic sites, an entirely different protocol for constructing a skeleton graph and a new set of reference populations (Supplementary Text 13), the ARB=>USR1 gene flow was supported (p -value = 0.05) using the bootstrap model-comparison method. The
45 direction of the gene flow was also supported by a significant log-likelihood difference (Supplementary Text 13).

A complex genetic landscape in the Americas

Strong signals (Supplementary Text 8) of Paleo-Inuit admixture were detected along the Chukotko-Kamchatkan/Eskimo-Aleut/Athabaskan cline (Fig. S42). This finding received attention in previous studies (4, 10, 11, 15, 17, 33), but here the signal was detected in an unsupervised way. These targets served as positive controls, as well as Native Americans having well-known colonial European admixture (Figs. S40, S42, Supplementary Text 8).

A controversial "population Y" signal in Native Americans (13-15, 34, 35) manifests itself as a low-level gene flow from Tianyuan or Onge lineages. We detected the Tianyuan component in one out of five ancient Brazilians from Lapa do Santo (9000-10000 cal yr BP) analyzed here (Fig. S47, Data S6). These individuals were previously analyzed by Posth *et al.* (14), and no signal was found. The Tianyuan admixture proportion inferred by us was low, 6%. Three of four ancient Brazilians from Sumidouro (~10000 cal yr BP) (13) analyzed here showed evidence of gene flows from deeply diverging sources, but several sources were almost equally supported (Data S6). In contrast, for the highest-coverage individual Sumidouro5 only the ancestry sources on the Onge and Tianyuan branches and on the adjacent stem edges were among the winning models (Data S6). The Onge- or Tianyuan-related ancestry proportions in Sumidouro5 were 7% or 8% of the genome (Data S6). A previously unreported Onge/Tianyuan signal (Fig. S47) was found at a low level (2% of the genome) in an ancient individual from Belize (individual ID I3443 from Mayahak Cab Pek) who was dated at 9300 cal yr BP (14). In five out of 12 ancient individuals from Lapa do Santo, Sumidouro and Belize, the Onge/Tianyuan gene flows were supported by the bootstrap model-comparison approach: in CP18, CP22, Sumidouro7, I3443, and I5456 (*p*-values ranged from 0.026 to 0.046). In the case of the Sumidouro5 individual, the "population Y" signal was also significant on the set of neutrally evolving sites derived from shotgun sequencing data (Supplementary Text 13).

Discussion

We question a major result on American pre-history, namely that Ancient Beringians represented by the late Pleistocene/early Holocene USR1 Alaskan individual form the deepest branch in the American clade (4). Here we introduce an alternative interpretation of Ancient Beringian ancestry that is significantly more likely than the one suggested by Moreno-Mayar *et al.* (4). We revealed a new admixture cline that can be represented as a mixture between NNA and a late Pleistocene Siberian population related to contemporary and later groups in the Amur River Basin (ARB). Only two individuals were confidently placed on this cline: USR1 (11500 cal yr BP) in Alaska and Kolyma1 in Chukotka (9800 cal yr BP). These individuals are roughly contemporaneous with the earliest ARB individual reported here and dated to ~12000 cal yr BP (Supplementary Text 3). Ancient Beringians were shown to be genetically unique among Americans (4, 13), and we offer a different explanation for this uniqueness: a unique admixture event in the history of this group instead of their unique phylogenetic position. We summarize various lines of evidence supporting this admixture event in Table S25.

Lexical remnants of a substrate language presumably spoken on the Denali territory before the Athabaskan radiation ca. 2000 – 3000 cal yr BP show similarity with the Chukotko-Kamchatkan language family (Supplementary Text 14a). This suggests that this substrate pre-

Athabaskan language could be a member of the hypothetical Chukotko-Kamchatkan–Nivkh linguistic clade (36). There is also evidence for a distant relationship between the latter clade on the Asian side and Salishan and Algonquian language families in North America (reviewed in Supplementary Text 14b). Since the Nivkh speakers are prominent present-day representatives of the ARB gene pool (Data S10), it is possible that these linguistic traces reflect the Pleistocene gene flow revealed in this study between the ARB cluster and Ancient Beringians.

We believe that the admixture signals in the USR1 and Kolyma1 individuals remained on the verge of statistical significance in previous studies (4, 10) because the dates of divergence of the First American and ARB lineages from the Asian stem were separated by a relatively short period likely coinciding with the Beringian standstill (6, 13), and the ARB lineage(s) contributing to Kolyma1 and USR1 split from the ARB stem shortly after the emergence of the ARB population. In other words, the phylogeny is nearly star-like and thus hardly amenable to standard analytical methods (37). Although without further ancient DNA sampling these split times are hard to estimate, this hypothesis seems probable considering the known divergence dates for First Americans and Siberians (4, 15, 38) and the radiocarbon dates of the individuals mentioned above. Notably, another analysis by Moreno-Mayar *et al.* (4) revealed a statistically significant signal of Asian admixture in USR1 (Table S25). Four simple demographic coalescent models were tested for population pairs USR1/X using *diCal* (39), and a significance threshold for LL difference was optimized on simulated data (4). A clean split model was rejected for all partner populations tested: present-day Aymara, Karitiana, Athabaskans, Han, Koryaks and Nivkhs. Remarkably, only one model was significantly better than the other three models for the Han/USR1 and Nivkh/USR1 pairs, and that was a second contact model (4). And Nivkhs are probably the closest present-day relatives of the ARB cluster according to our graph mapping results, with 81% of their ancestry derived from it (Data S10). We expect that cryptic gene flows similar to the one detected here are common in published admixture graphs, and these undetected gene flows may be important for deciding between competing archaeological interpretations.

The most parsimonious archaeological interpretation of our results, given the geographic anchor points for early NNA, SNA, Ancient Beringians, and ARB individuals between 13000 and 8000 cal yr BP, is that the divergence of the SNA and NNA clades occurred in Asia, immediately after the end of the Beringian standstill (6) that, in fact, could have taken place outside of Beringia (Fig. 4). This model explains the ARB signal in Ancient Beringians without invoking numerous independent crossings of the Bering strait, for which there is no clear archaeological evidence. We hypothesize that SNA ancestors were the first to move to Beringia towards the end of the isolation period, NNA ancestors were the second, and ancestors of Ancient Beringians (an NNA sub-group) were the last to migrate into North America, and thus interacted more with the ARB group ancestors spreading from other refugia at the same time. The hypothesis that the major First American clades diverged in East Siberia, and not in Beringia, also makes the "population Y" signal in present-day and ancient Amazonians (13, 34) less surprising since it provides more opportunities for contacts between recently diverged American groups and various Asian groups.

References and Notes:

1. P. Skoglund, D. Reich, A genomic view of the peopling of the Americas. *Curr. Opin. Genet. Dev.* **41**, 27-35 (2016).
- 5 2. M. R. Waters, Late Pleistocene exploration and settlement of the Americas by modern humans. *Science* **365**, eaat5447 (2019).
3. K. E. Graf, I. Buvit, Human dispersal from Siberia to Beringia: Assessing a Beringian standstill in light of the archaeological evidence. *Curr. Anthropol.* **58** (suppl. 17), S583-S603 (2017).
- 10 4. J. V. Moreno-Mayar, B. A. Potter, L. Vinner, M. Steinrücken, S. Rasmussen, J. Terhorst, J. A. Kamm, A. Albrechtsen, A. S. Malaspinas, M. Sikora, J. D. Reuther, J. D. Irish, R. S. Malhi, L. Orlando, Y. S. Song, R. Nielsen, D. J. Meltzer, E. Willerslev, Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**, 203-207 (2018).
- 5 15 6. B. A. Potter, J. D. Reuther, V. T. Holliday, C. E. Holmes, D. S. Miller, N. Schmuck, Early colonization of Beringia and Northern North America: Chronology, routes, and adaptive strategies. *Quat. Int.* **444**, 36-55 (2017).
- 15 6. E. Tamm, T. Kivisild, M. Reidla, M. Metspalu, D. G. Smith, C. J. Mulligan, C. M. Bravi, O. Rickards, C. Martinez-Labarga, E. K. Khusnutdinova, S. A. Fedorova, M. V. Golubenko, V. A. Stepanov, M. A. Gubina, S. I. Zhadanov, L. P. Ossipova, L. Damba, M. I. Voevoda, J. E. Dipierri, R. Villems, R. S. Malhi, Beringian standstill and spread of Native American founders. *PLoS One* **2**, e829 (2007).
- 20 7. T. J. Braje, T. D. Dillehay, J. M. Erlandson, R. G. Klein, T. C. Rick, Finding the first Americans. *Science* **358**, 592-594 (2017).
8. B. A. Potter, J. F. Baichtal, A. B. Beaudoin, L. Fehren-Schmitz, C. V. Haynes, V. T. Holliday, C. E. Holmes, J. W. Ives, R. L. Kelly, B. Llamas, R. S. Malhi, D. S. Miller, D. Reich, J. D. Reuther, S. Schiffels, T. A. Surovell, Current evidence allows multiple models for the peopling of the Americas. *Sci. Adv.* **4**, eaat5473 (2018).
- 25 9. L. Becerra-Valdivia, T. Higham, The timing and effect of the earliest human arrivals in North America. *Nature* **584**, 93-97 (2020).
10. M. Sikora, V. V. Pitulko, V. C. Sousa, M. E. Allentoft, L. Vinner, S. Rasmussen, A. Margaryan, P. de Barros Damgaard, C. de la Fuente, G. Renaud, M. A. Yang, Q. Fu, I. Dupanloup, K. Giampoudakis, D. Nogués-Bravo, C. Rahbek, G. Kroonen, M. Peyrot, H. McColl, S. V. Vasilyev, E. Veselovskaya, M. Gerasimova, E. Y. Pavlova, V. G. Chasnyk, P. A. Nikolskiy, A. V. Gromov, V. I. Khartanovich, V. Moiseyev, P. S. Grebenyuk, A. Y. Fedorchenko, A. I. Lebedintsev, S. B. Slobodin, B. A. Malyarchuk, R. Martiniano, M. Meldgaard, L. Arppe, J. U. Palo, T. Sundell, K. Mannermaa, M. Putkonen, V. Alexandersen, C. Primeau, N. Baimukhanov, R. S. Malhi, K. G. Sjögren, K. Kristiansen, A. Wessman, A. Sajantila, M. M. Lahr, R. Durbin, R. Nielsen, D. J. Meltzer, L. Excoffier, E. Willerslev, The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182-188 (2019).
- 30 35 40 11. P. Flegontov, N. E. Altınışık, P. Changmai, N. Rohland, S. Mallick, N. Adamski, D. A. Bolnick, N. Broomandkoshbacht, F. Candilio, B. J. Culleton, O. Flegontova, T. M. Friesen, C. Jeong, T. K. Harper, D. Keating, D. J. Kennett, A. M. Kim, T. C. Lamnidis, A. M. Lawson, I. Olalde, J. Oppenheimer, B. A. Potter, J. Raff, R. A. Sattler, P. Skoglund, K. Stewardson, E. J. Vajda, S. Vasilyev, E. Veselovskaya, M. G. Hayes, D. H. O'Rourke, J. Krause, R. Pinhasi, D. Reich, S. Schiffels, Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America. *Nature* **570**, 236-240 (2019).

12. B. A. Potter, J. D. Irish, J. D. Reuther, H. J. McKinney, New insights into Eastern Beringian mortuary behavior: a terminal Pleistocene double infant burial at Upward Sun River. *Proc. Natl. Acad. Sci. U.S.A.* **III**, 17060-17065 (2014).
- 5 13. J. V. Moreno-Mayar, L. Vinner, P. de Barros Damgaard, C. de la Fuente, J. Chan, J. P. Spence, M. E. Allentoft, T. Vimala, F. Racimo, T. Pinotti, S. Rasmussen, A. Margaryan, M. Iraeta Orbegozo, D. Myopotamitaki, M. Wooller, C. Bataille, L. Becerra-Valdivia, D. Chivall, D. Comeskey, T. Devièse, D. K. Grayson, L. George, H. Harry, V. Alexandersen, C. Primeau, J. Erlandson, C. Rodrigues-Carvalho, S. Reis, M. Q. R. Bastos, J. Cybulski, C. Vullo, F. Morello, M. Vilar, S. Wells, K. Gregersen, K. L. Hansen, N. Lynnerup, M. Mirazón Lahr, K. Kjær, A. Strauss, M. Alfonso-Durruty, 10 A. Salas, H. Schroeder, T. Higham, R. S. Malhi, J. T. Rasic, L. Souza, F. R. Santos, A. S. Malaspinas, M. Sikora, R. Nielsen, Y. S. Song, D. J. Meltzer, Willerslev E, Early human dispersals within the Americas. *Science* **362**, eaav2621 (2018).
- 15 14. C. Posth, N. Nakatsuka, I. Lazaridis, P. Skoglund, S. Mallick, T. C. Lamnidis, N. Rohland, K. Nägele, N. Adamski, E. Bertolini, N. Broomandkoshbacht, A. Cooper, B. J. Culleton, T. Ferraz, M. Ferry, A. Furtwängler, W. Haak, K. Harkins, T. K. Harper, T. Hünemeier, A. M. Lawson, B. Llamas, M. Michel, E. Nelson, J. Oppenheimer, N. Patterson, S. Schiffels, J. Sedig, K. Stewardson, S. Talamo, C.-C. Wang, J.-J. Hublin, M. Hubbe, K. Harvati, A. Nuevo Delaunay, J. Beier, M. Francken, P. Kaulicke, H. Reyes-Centeno, K. Rademaker, W. R. Trask, M. Robinson, S. M. Gutierrez, K. M. 20 Prufer, D. C. Salazar-García, E. N. Chim, L. Müller Plumm Gomes, M. L. Alves, A. Liryo, M. Inglez, R. E. Oliveira, D. V. Bernardo, A. Barioni, V. Wesolowski, N. A. Scheifler, M. A. Rivera, C. R. Plens, P. G. Messineo, L. Figuti, D. Corach, C. Scabuzzo, S. Eggers, P. DeBlasis, M. Reindel, C. Méndez, G. Politis, E. Tomasto-Cagigao, D. J. Kennett, A. Strauss, L. Fehren-Schmitz, J. Krause, D. Reich, Reconstructing the Deep Population History of Central and South America. *Cell* **175**, 1185-1197.e22 (2018).
- 25 15. M. Raghavan, M. Steinrücken, K. Harris, S. Schiffels, S. Rasmussen, M. DeGiorgio, A. Albrechtsen, C. Valdiosera, M. C. Ávila-Arcos, A.-S. Malaspinas, A. Eriksson, I. Moltke, M. Metspalu, J. R. Homburger, J. Wall, O. E. Cornejo, J. V. Moreno-Mayar, T. S. Korneliussen, T. Pierre, M. Rasmussen, P. F. Campos, P. de B. Damgaard, M. E. Allentoft, J. Lindo, E. Metspalu, R. Rodríguez-Varela, J. Mansilla, C. Henrickson, A. Seguin-Orlando, H. Malmström, T. Stafford, S. S. 30 Shringarpure, A. Moreno-Estrada, M. Karmin, K. Tambets, A. Bergström, Y. Xue, V. Warmuth, A. D. Friend, J. Singarayer, P. Valdes, F. Balloux, I. Lebreiro, J. L. Vera, H. Rangel-Villalobos, D. Pettener, D. Luiselli, L. G. Davis, E. Heyer, C. P. E. Zollikofer, M. S. Ponce de León, C. I. Smith, V. Grimes, K.-A. Pike, M. Deal, B. T. Fuller, B. Arriaza, V. Standen, M. F. Luz, F. Ricaut, N. Guidon, L. Osipova, M. I. Voevoda, O. L. Posukh, O. Balanovsky, M. Lavryashina, Y. Bogunov, E. 35 Khusnutdinova, M. Gubina, E. Balanovska, S. Fedorova, S. Litvinov, B. Malyarchuk, M. Derenko, M. J. Mosher, D. Archer, J. Cybulski, B. Petzelt, J. Mitchell, R. Worl, P. J. Norman, P. Parham, B. M. Kemp, T. Kivisild, C. Tyler-Smith, M. S. Sandhu, M. Crawford, R. Villems, D. G. Smith, M. R. Waters, T. Goebel, J. R. Johnson, R. S. Malhi, M. Jakobsson, D. J. Meltzer, A. Manica, R. Durbin, C. D. Bustamante, Y. S. Song, R. Nielsen, E. Willerslev, Genomic evidence for the Pleistocene and 40 recent population history of Native Americans. *Science* **349**, aab3884 (2015).
- 45 16. M. Rasmussen, S. L. Anzick, M. R. Waters, P. Skoglund, M. DeGiorgio, T. W. Stafford Jr, S. Rasmussen, I. Moltke, A. Albrechtsen, S. M. Doyle, G. D. Poznik, V. Gudmundsdottir, R. Yadav, A.-S. Malaspinas, S. S. White 5th, M. E. Allentoft, O. E. Cornejo, K. Tambets, A. Eriksson, P. D. Heintzman, M. Karmin, T. S. Korneliussen, D. J. Meltzer, T. L. Pierre, J. Stenderup, L. Saag, V. M. Warmuth, M. C. Lopes, R. S. Malhi, S. Brunak, T. Sicheritz-Ponten, I. Barnes, M. Collins, L. Orlando, F. Balloux, A. Manica, R. Gupta, M. Metspalu, C. D. Bustamante, M. Jakobsson, R. Nielsen, E. Willerslev, The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**, 225-229 (2014).

17. D. Reich, N. Patterson, D. Campbell, A. Tandon, S. Mazieres, N. Ray, M. V. Parra, W. Rojas, C. Duque, N. Mesa, L. F. García, O. Triana, S. Blair, A. Maestre, J. C. Dib, C. M. Bravi, G. Bailliet, D. Corach, T. Hünemeier, M. C. Bortolini, F. M. Salzano, M. L. Petzl-Erler, V. Acuña-Alonzo, C. Aguilar-Salinas, S. Canizales-Quinteros, T. Tusié-Luna, L. Riba, M. Rodríguez-Cruz, M. Lopez-Alarcón, R. Coral-Vazquez, T. Canto-Cetina, I. Silva-Zolezzi, J. C. Fernandez-Lopez, A. V. Contreras, G. Jimenez-Sanchez, M. J. Gómez-Vázquez, J. Molina, A. Carracedo, A. Salas, C. Gallo, G. Poletti, D. B. Witonsky, G. Alkorta-Aranburu, R. I. Sukernik, L. Osipova, S. A. Fedorova, R. Vasquez, M. Villena, C. Moreau, R. Barrantes, D. Pauls, L. Excoffier, G. Bedoya, F. Rothhammer, J.-M. Dugoujon, G. Larrouy, W. Klitz, D. Labuda, J. Kidd, K. Kidd, A. Di Rienzo, N. B. Freimer, A. L. Price, A. Ruiz-Linares, Reconstructing Native American population history. *Nature* **488**, 370-374 (2012).
18. C. L. Scheib, H. Li, T. Desai, V. Link, C. Kendall, G. Dewar, P. W. Griffith, A. Mörseburg, J. R. Johnson, A. Potter, S. L. Kerr, P. Endicott, J. Lindo, M. Haber, Y. Xue, C. Tyler-Smith, M. S. Sandhu, J. G. Lorenz, T. D. Randall, Z. Faltyskova, L. Pagani, P. Danecek, T. C. O’Connell, P. Martz, A. S. Boraas, B. F. Byrd, A. Leventhal, R. Cambra, R. Williamson, L. Lesage, B. Holguin, E. Ygnacio-De Soto, J. Rosas, M. Metspalu, J. T. Stock, A. Manica, A. Scally, D. Wegmann, R. S. Malhi, T. Kivisild, Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **360**, 1024-1027 (2018).
19. P. Verdu, T. J. Pemberton, R. Laurent, B. M. Kemp, A. Gonzalez-Oliver, C. Gorodezky, C. E. Hughes, M. R. Shattuck, B. Petzelt, J. Mitchell, H. Harry, T. William, R. Worl, J. S. Cybulski, N. A. Rosenberg, R. S. Malhi, Patterns of admixture and population structure in native populations of Northwest North America. *PLoS Genet.* **10**, e1004530 (2014).
20. H. Yu, M. A. Spyrou, M. Karapetian, S. Shnaider, R. Radzevičiūtė, K. Nägele, G. U. Neumann, S. Penske, J. Zech, M. Lucas, P. LeRoux, P. Roberts, G. Pavlenok, A. Buzhilova, C. Posth, C. Jeong, J. Krause, Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia. *Cell* **181**, 1232-1245.e20 (2020).
21. V. Siska, E. R. Jones, S. Jeon, Y. Bhak, H.-M. Kim, Y. S. Cho, H. Kim, K. Lee, E. Veselovskaya, T. Balueva, M. Gallego-Llorente, M. Hofreiter, D. G. Bradley, A. Eriksson, R. Pinhasi, J. Bhak, A. Manica, Genome-wide data from two early Neolithic East Asian individuals dating to 7700 years ago. *Sci. Adv.* **3**, e1601877 (2017).
22. C. Ning, T. Li, K. Wang, F. Zhang, T. Li, X. Wu, S. Gao, Q. Zhang, H. Zhang, M. J. Hudson, G. Dong, S. Wu, Y. Fang, C. Liu, C. Feng, W. Li, T. Han, R. Li, J. Wei, Y. Zhu, Y. Zhou, C.-C. Wang, S. Fan, Z. Xiong, Z. Sun, M. Ye, L. Sun, X. Wu, F. Liang, Y. Cao, X. Wei, H. Zhu, H. Zhou, J. Krause, M. Robbeets, C. Jeong, Y. Cui, Ancient genomes from northern China suggest links between subsistence changes and human migration. *Nat. Commun.* **11**, 2700 (2020).
23. C.-C. Wang *et al.*, <https://www.biorxiv.org/content/10.1101/2020.03.25.004606v1> (2020).
24. Q. Fu, M. Hajdinjak, O. T. Moldovan, S. Constantin, S. Mallick, P. Skoglund, N. Patterson, N. Rohland, I. Lazaridis, B. Nickel, B. Viola, K. Prüfer, M. Meyer, J. Kelso, D. Reich, S. Pääbo, An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216-219 (2015).
25. I. Mathieson, I. Lazaridis, N. Rohland, S. Mallick, N. Patterson, S. A. Roodenberg, E. Harney, K. Stewardson, D. Fernandes, M. Novak, K. Sirak, C. Gamba, E. R. Jones, B. Llamas, S. Dryomov, J. Pickrell, J. L. Arsuaga, J. M. B. de Castro, E. Carbonell, F. Gerritsen, A. Khokhlov, P. Kuznetsov, M. Lozano, H. Meller, O. Mochalov, V. Moiseyev, M. A. R. Guerra, J. Roodenberg, J. M. Vergès, J. Krause, A. Cooper, K. W. Alt, D. Brown, D. Anthony, C. Lalueza-Fox, W. Haak, R. Pinhasi, D. Reich, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499-503 (2015).

26. J. Kelleher, A. M. Etheridge, G. McVean, Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Comput. Biol.* **12**, e1004842 (2016).
27. C. Jeong, O. Balanovsky, E. Lukianova, N. Kahbatkyzy, P. Flegontov, V. Zaporozhchenko, A. Immel, C.-C. Wang, O. Ixan, E. Khussainova, B. Bekmanov, V. Zaibert, M. Lavryashina, E. Pocheshkhova, Y. Yusupov, A. Agdzhoyan, S. Koshel, A. Bukin, P. Nymadawa, S. Turdikulova, D. Dalimova, M. Churnosov, R. Skhalyakho, D. Daragan, Y. Bogunov, A. Bogunova, A. Shtrunov, N. Dubova, M. Zhabagin, L. Yepiskoposyan, V. Churakov, N. Pislegin, L. Damba, L. Saroyants, K. Dibirova, L. Atramentova, O. Utevska, E. Idrisov, E. Kamenshchikova, I. Evseeva, M. Metspalu, A. K. Outram, M. Robbeets, L. Djansugurova, E. Balanovska, S. Schiffels, W. Haak, D. Reich, J. Krause, The genetic history of admixture across inner Eurasia. *Nat. Ecol. Evol.* **3**, 966-976 (2019).
28. I. Lazaridis, D. Nadel, G. Rollefson, D. C. Merrett, N. Rohland, S. Mallick, D. Fernandes, M. Novak, B. Gamarra, K. Sirak, S. Connell, K. Stewardson, E. Harney, Q. Fu, G. Gonzalez-Fortes, E. R. Jones, S. A. Roodenberg, G. Lengyel, F. Bocquentin, B. Gasparian, J. M. Monge, M. Gregg, V. Eshed, A.-S. Mizrahi, C. Meiklejohn, F. Gerritsen, L. Bejenaru, M. Blüher, A. Campbell, G. Cavalleri, D. Comas, P. Froguel, E. Gilbert, S. M. Kerr, P. Kovacs, J. Krause, D. McGettigan, M. Merrigan, D. A. Merriwether, S. O'Reilly, M. B. Richards, O. Semino, M. Shamoon-Pour, G. Stefanescu, M. Stumvoll, A. Tönjes, A. Torroni, J. F. Wilson, L. Yengo, N. A. Hovhannisyan, N. Patterson, R. Pinhasi, D. Reich, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419-424 (2016).
29. T. Gakuhari, S. Nakagome, S. Rasmussen, M. E. Allentoft, T. Sato, T. Korneliusen, B. N. Chuinneagáin, H. Matsumae, K. Koganebuchi, R. Schmidt, S. Mizushima, O. Kondo, N. Shigehara, M. Yoneda, R. Kimura, H. Ishida, T. Masuyama, Y. Yamada, A. Tajima, H. Shibata, A. Toyoda, T. Tsurumoto, T. Wakebe, H. Shitara, T. Hanihara, E. Willerslev, M. Sikora, H. Oota, Ancient Jomon genome sequence analysis sheds light on migration patterns of early East Asian populations. *Commun. Biol.* **3**, 437 (2020).
30. V. M. Narasimhan, N. Patterson, P. Moorjani, N. Rohland, R. Bernardos, S. Mallick, I. Lazaridis, N. Nakatsuka, I. Olalde, M. Lipson, A. M. Kim, L. M. Olivieri, A. Coppa, M. Vidale, J. Mallory, V. Moiseyev, E. Kitov, J. Monge, N. Adamski, N. Alex, N. Broomandkhoshbacht, F. Candilio, K. Callan, O. Cheronet, B. J. Culleton, M. Ferry, D. Fernandes, S. Freilich, B. Gamarra, D. Gaudio, M. Hajdinjak, É. Harney, T. K. Harper, D. Keating, A. M. Lawson, M. Mah, K. Mandl, M. Michel, M. Novak, J. Oppenheimer, N. Rai, K. Sirak, V. Slon, K. Stewardson, F. Zalzal, Z. Zhang, G. Akhatov, A. N. Bagashev, A. Bagnera, B. Baitanayev, J. Bendezu-Sarmiento, A. A. Bissembaev, G. L. Bonora, T. T. Charynov, T. Chikisheva, P. K. Dashkovskiy, A. Derevianko, M. Dobeš, K. Douka, N. Dubova, M. N. Duisengali, D. Enshin, A. Epimakhov, A. V. Fribus, D. Fuller, A. Goryachev, A. Gromov, S. P. Grushin, B. Hanks, M. Judd, E. Kazizov, A. Khokhlov, A. P. Krygin, E. Kupriyanova, P. Kuznetsov, D. Luiselli, F. Maksudov, A. M. Mamedov, T. B. Mamirov, C. Meiklejohn, D. C. Merrett, R. Micheli, O. Mochalov, S. Mustafokulov, A. Nayak, D. Pettener, R. Potts, D. Razhev, M. Rykun, S. Sarno, T. M. Savenkova, K. Sikhymbaeva, S. M. Slepchenko, O. A. Soltobaev, N. Stepanova, S. Svyatko, K. Tabaldiev, M. Teschler-Nicola, A. A. Tishkin, V. V. Tkachev, S. Vasilyev, P. Velemínský, D. Voyakin, A. Yermolayeva, M. Zahir, V. S. Zubkov, A. Zubova, V. S. Shinde, C. Lalueza-Fox, M. Meyer, D. Anthony, N. Boivin, K. Thangaraj, D. J. Kennett, M. Frachetti, R. Pinhasi, D. Reich, The formation of human populations in South and Central Asia. *Science* **365**, aat7487 (2019).
31. A. Bergström, S. A. McCarthy, R. Hui, M. A. Almarri, Q. Ayub, P. Danecek, Y. Chen, S. Felkel, P. Hallast, J. Kamm, H. Blanché, J.-F. Deleuze, H. Cann, S. Mallick, D. Reich, M. S. Sandhu, P. Skoglund, A. Scally, Y. Xue, R. Durbin, C. Tyler-Smith, Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).

32. F. Pouyet, S. Aeschbacher, A. Thiéry, L. Excoffier, Background selection and biased gene conversion affect more than 95% of the human genome and bias demographic inferences. *eLife* **7**, e36317 (2018).
33. M. Raghavan, M. DeGiorgio, A. Albrechtsen, I. Moltke, P. Skoglund, T. S. Korneliussen, B. Grønnow, M. Appelt, H. C. Gulløv, T. M. Friesen, W. Fitzhugh, H. Malmström, S. Rasmussen, J. Olsen, L. Melchior, B. T. Fuller, S. M. Fahrni, T. Stafford Jr, V. Grimes, M. A. P. Renouf, J. Cybulski, N. Lynnerup, M. M. Lahr, K. Britton, R. Knecht, J. Arneborg, M. Metspalu, O. E. Cornejo, A.-S. Malaspinas, Y. Wang, M. Rasmussen, V. Raghavan, T. V. O. Hansen, E. Khusnutdinova, T. Pierre, K. Dneprovsky, C. Andreasen, H. Lange, M. G. Hayes, J. Coltrain, V. A. Spitsyn, A. Götherström, L. Orlando, T. Kivisild, R. Villems, M. H. Crawford, F. C. Nielsen, J. Dissing, J. Heinemeier, M. Meldgaard, C. Bustamante, D. H. O'Rourke, M. Jakobsson, M. T. P. Gilbert, R. Nielsen, E. Willerslev, The genetic prehistory of the New World Arctic. *Science* **345**, 1255832 (2014).
34. P. Skoglund, S. Mallick, M. C. Bortolini, N. Chennagiri, T. Hünemeier, M. L. Petzl-Erler, F. M. Salzano, N. Patterson, D. Reich, Genetic evidence for two founding populations of the Americas. *Nature* **525**, 104-108 (2015).
35. M. A. Yang, X. Gao, C. Theunert, H. Tong, A. Aximu-Petri, B. Nickel, M. Slatkin, M. Meyer, S. Pääbo, J. Kelso, Q. Fu, 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr. Biol.* **27**, 3202-3208.e9 (2017).
36. M. Fortescue, The relationship of Nivkh to Chukotko-Kamchatkan revisited. *Lingua* **121**, 1359-1376 (2011).
37. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239-2252 (2011).
38. B. Llamas, L. Fehren-Schmitz, G. Valverde, J. Soubrier, S. Mallick, N. Rohland, S. Nordenfelt, C. Valdiosera, S. M. Richards, A. Rohrlach, M. I. B. Romero, I. F. Espinoza, E. T. Cagigao, L. W. Jiménez, K. Makowski, I. S. L. Reyna, J. M. Lory, J. A. B. Torrez, M. A. Rivera, R. L. Burger, M. C. Ceruti, J. Reinhard, R. Spencer Wells, G. Politis, C. M. Santoro, V. G. Standen, C. Smith, D. Reich, S. Y. W. Ho, A. Cooper, W. Haak, Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* **2**, e1501385 (2016).
39. M. Steinrücken, J. Kamm, J. P. Spence, Y. S. Song, Inference of complex population histories using whole-genome sequences from multiple populations. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 17115-17120 (2019).
40. M. Raghavan, P. Skoglund, K. E. Graf, M. Metspalu, A. Albrechtsen, I. Moltke, S. Rasmussen, T. W. Stafford Jr, L. Orlando, E. Metspalu, M. Karmin, K. Tambets, S. Rootsi, R. Mägi, P. F. Campos, E. Balanovska, O. Balanovsky, E. Khusnutdinova, S. Litvinov, L. P. Osipova, S. A. Fedorova, M. I. Voevoda, M. DeGiorgio, T. Sicheritz-Ponten, S. Brunak, S. Demeshchenko, T. Kivisild, R. Villems, R. Nielsen, M. Jakobsson, E. Willerslev, Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87-91 (2014).
41. P. de Barros Damgaard, R. Martiniano, J. Kamm, J. V. Moreno-Mayar, G. Kroonen, M. Peyrot, G. Barjamovic, S. Rasmussen, C. Zacho, N. Baimukhanov, V. Zaibert, V. Merz, A. Biddanda, I. Merz, V. Loman, V. Evdokimov, E. Usmanova, B. Hemphill, A. Seguin-Orlando, F. E. Yediay, I. Ullah, K.-G. Sjögren, K. H. Iversen, J. Choin, C. de la Fuente, M. Ilardo, H. Schroeder, V. Moiseyev, A. Gromov, A. Polyakov, S. Omura, S. Y. Senyurt, H. Ahmad, C. McKenzie, A. Margaryan, A. Hameed, A. Samad, N. Gul, M. H. Khokhar, O. I. Goriunova, V. I. Bazaliiskii, J. Novembre, A. W. Weber, L. Orlando, M. E. Allentoft, R. Nielsen, K. Kristiansen, M. Sikora, A. K. Outram, R. Durbin, E. Willerslev, The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, aar7711 (2018).

Supplementary references:

42. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
- 5 43. R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Z. Kucan, I. Gusic, V. B. Doronichev, L. V Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, S. Pääbo, A draft sequence of the Neandertal genome. *Science* **328**, 710-722 (2010).
- 10
44. D. Reich, K. Thangaraj, N. Patterson, A. L. Price, L. Singh, Reconstructing Indian population history. *Nature* **461**, 489-494 (2009).
- 15
45. N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, D. Reich, Ancient admixture in human history. *Genetics* **192**, 1065-1093 (2012)
46. S. Soraggi, C. Wiuf, General theory for stochastic admixture graphs and F-statistics. *Theor. Popul. Biol.* **125**, 56-66 (2019).
- 20
47. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655-1664 (2009).
48. W. Haak, I. Lazaridis, N. Patterson, N. Rohland, S. Mallick, B. Llamas, G. Brandt, S. Nordenfelt, E. Harney, K. Stewardson, Q. Fu, A. Mittnik, E. Bánffy, C. Economou, M. Francken, S. Friederich, R. G. Pena, F. Hallgren, V. Khartanovich, A. Khokhlov, M. Kunst, P. Kuznetsov, H. Meller, O. Mochalov, V. Moiseyev, N. Nicklisch, S. L. Pichler, R. Risch, M. a. Rojo Guerra, C. Roth, A. Szécsényi-Nagy, J. Wahl, M. Meyer, J. Krause, D. Brown, D. Anthony, A. Cooper, K. W. Alt, D. Reich, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* (2015), doi:10.1038/nature14317.
- 25
49. É. Harney et al., <https://www.biorxiv.org/content/10.1101/2020.04.09.032664v1> (2020).
- 30
50. M. E. Allentoft, M. Sikora, K.-G. Sjögren, S. Rasmussen, M. Rasmussen, J. Stenderup, P. B. Damgaard, H. Schroeder, T. Ahlström, L. Vinner, A.-S. Malaspinas, A. Margaryan, T. Higham, D. Chivall, N. Lynnerup, L. Harvig, J. Baron, P. Della Casa, P. Dąbrowski, P. R. Duffy, A. V. Ebel, A. Epimakhov, K. Frei, M. Furmanek, T. Gralak, A. Gromov, S. Gronkiewicz, G. Grupe, T. Hajdu, R. Jarysz, V. Khartanovich, A. Khokhlov, V. Kiss, J. Kolář, A. Kriiska, I. Lasak, C. Longhi, G. McGlynn, A. Merkevicius, I. Merkyte, M. Metspalu, R. Mkrtychyan, V. Moiseyev, L. Paja, G. Pálfi, D. Pokutta, Ł. Pospieszny, T. D. Price, L. Saag, M. Sablin, N. Shishlina, V. Smrčka, V. I. Soenov, V. Szeverényi, G. Tóth, S. V. Trifanova, L. Varul, M. Vicze, L. Yepiskoposyan, V. Zhitenev, L. Orlando, T. Sicheritz-Pontén, S. Brunak, R. Nielsen, K. Kristiansen, E. Willerslev, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-172 (2015).
- 35
- 40
51. I. Lazaridis, A. Mittnik, N. Patterson, S. Mallick, N. Rohland, S. Pfrengle, A. Furtwängler, A. Peltzer, C. Posth, A. Vasilakis, P. J. P. McGeorge, E. Konsolaki-Yannopoulou, G. Korres, H. Martlew, M. Michalodimitrakis, M. Özsait, N. Özsait, A. Papatheanasiou, M. Richards, S. A. Roodenberg, Y. Tzedakis, R. Arnott, D. M. Fernandes, J. R. Hughey, D. M. Lotakis, P. A. Navas, Y. Maniatis, J. A. Stamatoyannopoulos, K. Stewardson, P. Stockhammer, R. Pinhasi, D. Reich, J. Krause, G. Stamatoyannopoulos, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214-218 (2017).
- 45

52. E. R. Jones, G. Zarina, V. Moiseyev, E. Lightfoot, P. R. Nigst, A. Manica, R. Pinhasi, D. G. Bradley, The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr. Biol.* **27**, 576-582 (2017).
53. P. de Barros Damgaard, N. Marchi, S. Rasmussen, M. Peyrot, G. Renaud, T. Korneliussen, J. V. Moreno-Mayar, M. W. Pedersen, A. Goldberg, E. Usmanova, N. Baimukhanov, V. Loman, L. Hedeager, A. G. Pedersen, K. Nielsen, G. Afanasiev, K. Akmatov, A. Aldashev, A. Alpaslan, G. Baimbetov, V. I. Bazaliiskii, A. Beisenov, B. Boldbaatar, B. Boldgiv, C. Dorzhu, S. Ellingvag, D. Erdenebaatar, R. Dajani, E. Dmitriev, V. Evdokimov, K. M. Frei, A. Gromov, A. Goryachev, H. Hakonarson, T. Hegay, Z. Khachatryan, R. Khaskhanov, E. Kitov, A. Kolbina, T. Kubatbek, A. Kukushkin, I. Kukushkin, N. Lau, A. Margaryan, I. Merkyte, I. V. Mertz, V. K. Mertz, E. Mijiddorj, V. Moiyesev, G. Mukhtarova, B. Nurmukhanbetov, Z. Orozbekova, I. Panyushkina, K. Pieta, V. Smrčka, I. Shevnina, A. Logvin, K. G. Sjögren, T. Štolcová, K. Tashbaeva, A. Tkachev, T. Tulegenov, D. Voyakin, L. Yepiskoposyan, S. Undrakhbold, V. Varfolomeev, A. Weber, N. Kradin, M. E. Allentoft, L. Orlando, R. Nielsen, M. Sikora, E. Heyer, K. Kristiansen, E. Willerslev, 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369-374 (2018).
54. É. Harney, H. May, D. Shalem, N. Rohland, S. Mallick, I. Lazaridis, R. Sarig, K. Stewardson, S. Nordenfelt, N. Patterson, I. Hershkovitz, D. Reich, Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nat. Commun.* **9**, 3336 (2018).
55. I. Mathieson, S. Alpaslan-Roodenberg, C. Posth, A. Szécsényi-Nagy, N. Rohland, S. Mallick, I. Olalde, N. Broomandkhoshbacht, F. Candilio, O. Cheronet, D. Fernandes, M. Ferry, B. Gamarra, G. G. Fortes, W. Haak, E. Harney, E. Jones, D. Keating, B. Krause-Kyora, I. Kucukkalipci, M. Michel, A. Mittnik, K. Nägele, M. Novak, J. Oppenheimer, N. Patterson, S. Pfrengle, K. Sirak, K. Stewardson, S. Vai, S. Alexandrov, K. W. Alt, R. Andreescu, D. Antonović, A. Ash, N. Atanassova, K. Bacvarov, M. B. Gusztáv, H. Bocherens, M. Bolus, A. Boroneanț, Y. Boyadzhiev, A. Budnik, J. Burmaz, S. Chohadzhiev, N. J. Conard, R. Cottiaux, M. Čuka, C. Cupillard, D. G. Drucker, N. Elenski, M. Francken, B. Galabova, G. Ganetsovski, B. Gély, T. Hajdu, V. Handzhyiska, K. Harvati, T. Higham, S. Iliev, I. Janković, I. Karavanić, D. J. Kennett, D. Komšo, A. Kozak, D. Labuda, M. Lari, C. Lazar, M. Leppek, K. Leshtakov, D. Lo Vetro, D. Los, I. Lozanov, M. Malina, F. Martini, K. McSweeney, H. Meller, M. Mendušić, P. Mirea, V. Moiseyev, V. Petrova, T. D. Price, A. Simalcsik, L. Sineo, M. Šlaus, V. Slavchev, P. Stanev, A. Starović, T. Szeniczey, S. Talamo, M. Teschler-Nicola, C. Thevenet, I. Valchev, F. Valentin, S. Vasilyev, F. Veljanovska, S. Venelinova, E. Veselovskaya, B. Viola, C. Virag, J. Zaninović, S. Zäuner, P. W. Stockhammer, G. Catalano, R. Krauß, D. Caramelli, G. Zariņa, B. Gaydarska, M. Lillie, A. G. Nikitin, I. Potekhina, A. Papathanasiou, D. Borić, C. Bonsall, J. Krause, R. Pinhasi, D. Reich, The genomic history of southeastern Europe. *Nature* **555**, 197-203 (2018).
56. A. Mittnik, C. C. Wang, S. Pfrengle, M. Daubaras, G. Zariņa, F. Hallgren, R. Allmäe, V. Khartanovich, V. Moiseyev, M. Tõrv, A. Furtwängler, A. Andrades Valtueña, M. Feldman, C. Economou, M. Oinonen, A. Vasks, E. Balanovska, D. Reich, R. Jankauskas, W. Haak, S. Schiffels, J. Krause, The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).
57. I. Olalde, S. Brace, M. E. Allentoft, I. Armit, K. Kristiansen, T. Booth, N. Rohland, S. Mallick, A. Szécsényi-Nagy, A. Mittnik, E. Altena, M. Lipson, I. Lazaridis, T. K. Harper, N. Patterson, N. Broomandkhoshbacht, Y. Diekmann, Z. Faltyskova, D. Fernandes, M. Ferry, E. Harney, P. De Knijff, M. Michel, J. Oppenheimer, K. Stewardson, A. Barclay, K. W. Alt, C. Liesau, P. Rios, C. Blasco, J. V. Miguel, R. M. Garcia, A. A. Fernandez, E. Banffy, M. Bernabo-Brea, D. Billoin, C. Bonsall, L. Bonsall, T. Allen, L. Bustar, S. Carver, L. C. Navarro, O. E. Craig, G. T. Cook, B. Cunliffe, A. Denaire, K. E. Dinwiddy, N. Dodwell, M. Ernee, C. Evans, M. Kucharik, J. F. Farre, C. Fowler, M. Gazenbeek, R. G. Pena, M. Haber-Uriarte, E. Haduch, G. Hey, N. Jowett, T. Knowles, K. Massy, S. Pfrengle, P. Lefranc, O. Lemerrier, A. Lefebvre, C. H. Martinez, V. G. Olmo, A. B. Ramirez, J. L.

- 5 Maurandi, T. Majo, J. I. McKinley, K. McSweeney, B. G. Mende, A. Mod, G. Kulcsar, V. Kiss, A. Czene, R. Patay, A. Endrodi, K. Kohler, T. Hajdu, T. Szeniczey, J. Dani, Z. Bernert, M. Hoole, O. Cheronet, D. Keating, P. Veleminsky, M. Dobe, F. Candilio, F. Brown, R. F. Fernandez, A. M. Herrero-Corral, S. Tusa, E. Carnieri, L. Lentini, A. Valenti, A. Zanini, C. Waddington, G. Delibes, E. Guerra-Doce, B. Neil, M. Brittain, M. Luke, R. Mortimer, J. Desideri, M. Besse, G. Brucken, M. Furmanek, A. Hauszko, M. Mackiewicz, A. Rapinski, S. Leach, I. Soriano, K. T. Lillios, J. L. Cardoso, M. P. Pearson, P. Wodarczak, T. D. Price, P. Prieto, P. J. Rey, R. Risch, M. A. R. Guerra, A. Schmitt, J. Serralongue, A. M. Silva, V. Smrcka, L. Vergnaud, J. Zilhao, D. Caramelli, T. Higham, M. G. Thomas, D. J. Kennett, H. Fokkens, V. Heyd, A. Sheridan, K. G. Sjogren, P. W. Stockhammer, 10 J. Krause, R. Pinhasi, W. Haak, I. Barnes, C. Lalueza-Fox, D. Reich, The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190-196 (2018).
58. I. Olalde, S. Mallick, N. Patterson, N. Rohland, V. Villalba-Mouco, M. Silva, K. Dulias, C. J. Edwards, F. Gandini, M. Pala, P. Soares, M. Ferrando-Bernal, N. Adamski, N. Broomandkoshbacht, O. Cheronet, B. J. Cullen, D. Fernandes, A. M. Lawson, M. Mah, J. Oppenheimer, K. Stewardson, 15 Z. Zhang, J. M. J. Arenas, I. J. T. Moyano, D. C. Salazar-García, P. Castanyer, M. Santos, J. Tremoleda, M. Lozano, P. G. Borja, J. Fernández-Eraso, J. A. Mujika-Alustiza, C. Barroso, F. J. Bermúdez, E. V. Mínguez, J. Burch, N. Coromina, D. Vivó, A. Cebrià, J. M. Fullola, O. García-Puchol, J. I. Morales, F. Xavier Oms, T. Majó, J. M. Vergès, A. Díaz-Carvajal, I. Ollich-Castanyer, F. Javier López-Cachero, A. M. Silva, C. Alonso-Fernández, G. D. De Castro, J. J. Echevarría, A. Moreno-Márquez, G. P. Berlanga, P. Ramos-García, J. Ramos-Muñoz, E. V. Vila, G. A. Arzo, Á. E. Arroyo, K. T. Lillios, J. Mack, J. Velasco-Vázquez, A. Waterman, L. B. De Lugo Enrich, M. B. Sánchez, B. Agustí, F. Codina, G. De Prado, A. Estalrich, Á. F. Flores, C. Finlayson, G. Finlayson, S. Finlayson, F. Giles-Guzmán, A. Rosas, V. B. González, G. G. Atiénzar, M. S. Hernández Pérez, A. Llanos, Y. C. Marco, I. C. Beneyto, D. López-Serrano, M. S. Tormo, A. C. Valera, C. Blasco, C. 20 Liesau, P. Ríos, J. Daura, M. J. De Pedro Michó, A. A. Diez-Castillo, R. F. Fernández, J. F. Farré, R. Garrido-Pena, V. S. Gonçalves, E. Guerra-Doce, A. M. Herrero-Corral, J. Juan-Cabanilles, D. López-Reyes, S. B. McClure, M. M. Pérez, A. O. Foix, M. S. Borràs, A. C. Sousa, J. M. V. Encinas, D. J. Kennett, M. B. Richards, K. W. Alt, W. Haak, R. Pinhasi, C. Lalueza-Fox, D. Reich, The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230-1234 (2019).
- 30 59. M. Feldman, E. Fernández-Domínguez, L. Reynolds, D. Baird, J. Pearson, I. Herskovitz, H. May, N. Goring-Morris, M. Benz, J. Gresky, R. A. Bianco, A. Fairbairn, G. Mustafaoğlu, P. W. Stockhammer, C. Posth, W. Haak, C. Jeong, J. Krause, Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).
- 35 60. M. Feldman, D. M. Master, R. A. Bianco, M. Burri, P. W. Stockhammer, A. Mittnik, A. J. Aja, C. Jeong, J. Krause, Ancient DNA sheds light on the genetic origins of early Iron Age Philistines. *Sci. Adv.* **5**, aax0061 (2019).
61. C. Ning, C.-C. Wang, S. Gao, Y. Yang, X. Zhang, X. Wu, F. Zhang, Z. Nie, Y. Tang, M. Robbeets, J. Ma, J. Krause, Y. Cui, Ancient genomes reveal Yamnaya-related ancestry and a potential source of Indo-European speakers in Iron Age Tianshan. *Curr. Biol.* **29**, 2526-2532.e4 (2019).
- 40 62. V. Villalba-Mouco, M. S. van de Loosdrecht, C. Posth, R. Mora, J. Martínez-Moreno, M. Rojo-Guerra, D. C. Salazar-García, J. I. Royo-Guillén, M. Kunst, H. Rougier, I. Crevecoeur, H. Arcusa-Magallón, C. Tejedor-Rodríguez, I. García-Martínez de Lagrán, R. Garrido-Pena, K. W. Alt, C. Jeong, S. Schiffels, P. Utrilla, J. Krause, W. Haak, Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian peninsula. *Curr. Biol.* **29**, 1169-1177.e7 (2019).
- 45 63. C.-C. Wang, S. Reinhold, A. Kalmykov, A. Wissgott, G. Brandt, C. Jeong, O. Cheronet, M. Ferry, E. Harney, D. Keating, S. Mallick, N. Rohland, K. Stewardson, A. R. Kantorovich, V. E. Maslov, V. G. Petrenko, V. R. Erlikh, B. C. Atabiev, R. G. Magomedov, P. L. Kohl, K. W. Alt, S. L. Pichler, C. Gerling, H. Meller, B. Vardanyan, L. Yeganyan, A. D. Rezepkin, D. Mariaschk, N. Berezina, J.

- Gresky, K. Fuchs, C. Knipper, S. Schiffels, E. Balanovska, O. Balanovsky, I. Mathieson, T. Higham, Y. B. Berezin, A. Buzhilova, V. Trifonov, R. Pinhasi, A. B. Belinskij, D. Reich, S. Hansen, J. Krause, W. Haak, Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nat. Commun.* **10**, 590 (2019).
- 5 64. D. M. Fernandes, A. Mittnik, I. Olalde, I. Lazaridis, O. Cheronet, N. Rohland, S. Mallick, R. Bernardos, N. Broomandkoshbacht, J. Carlsson, B. J. Culleton, M. Ferry, B. Gamarra, M. Lari, M. Mah, M. Michel, A. Modi, M. Novak, J. Oppenheimer, K. A. Sirak, K. Stewardson, K. Mandl, C. Schattke, K. T. Özdoğan, M. Lucci, G. Gasperetti, F. Candilio, G. Salis, S. Vai, E. Camarós, C. Calò, G. Catalano, M. Cueto, V. Forgia, M. Lozano, E. Marini, M. Micheletti, R. M. Micciché, M. R. Palombo, D. Ramis, V. Schimmenti, P. Sureda, L. Teira, M. Teschler-Nicola, D. J. Kennett, C. Lalueza-Fox, N. Patterson, L. Sineo, A. Coppa, D. Caramelli, R. Pinhasi, D. Reich, The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nat. Ecol. Evol.* **4**, 334-345 (2020).
- 10 65. A. Furtwängler, A. B. Rohrlach, T. C. Lamnidis, L. Papac, G. U. Neumann, I. Siebke, E. Reiter, N. Steuri, J. Hald, A. Denaire, B. Schnitzler, J. Wahl, M. Ramstein, V. J. Schuenemann, P. W. Stockhammer, A. Hafner, S. Lössch, W. Haak, S. Schiffels, J. Krause, Ancient genomes reveal social and genetic structure of Late Neolithic Switzerland. *Nat. Commun.* **11**, 1915 (2020).
- 15 66. J. H. Marcus, C. Posth, H. Ringbauer, L. Lai, R. Skeates, C. Sidore, J. Beckett, A. Furtwängler, A. Olivieri, C. W. K. Chiang, H. Al-Asadi, K. Dey, T. A. Joseph, C. C. Liu, C. Der Sarkissian, R. Radzevičiūtė, M. Michel, M. G. Gradoli, P. Marongiu, S. Rubino, V. Mazzarello, D. Rovina, A. La Fragola, R. M. Serra, P. Bandiera, R. Bianucci, E. Pompianu, C. Murgia, M. Guirguis, R. P. Orquin, N. Tuross, P. van Dommelen, W. Haak, D. Reich, D. Schlessinger, F. Cucca, J. Krause, J. Novembre, Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nat. Commun.* **11**, 939 (2020).
- 20 67. J. Kamm, J. Terhorst, R. Durbin, Y. S. Song, Efficiently inferring the demographic history of many populations with allele count data. *J. Am. Stat. Assoc.* **115**, 1472-1487 (2020).
- 25 68. S. Schiffels, W. Haak, P. Paajanen, B. Llamas, E. Popescu, L. Loe, R. Clarke, A. Lyons, R. Mortimer, D. Sayer, C. Tyler-Smith, A. Cooper, R. Durbin, Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat. Commun.* **7**, 10408 (2016).
- 30 69. L. Excoffier, I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, M. Foll, Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9**, e1003905 (2013).
70. M. Lipson, Applying f_4 -statistics and admixture graphs: Theory and examples. *Mol. Ecol. Resour.* 10.1111/1755-0998.13230 (2020).
- 35 71. C. Jeong, S. Wilkin, T. Amgalantugs, A. S. Bouwman, W. T. T. Taylor, R. W. Hagan, S. Bromage, S. Tsolmon, C. Trachsel, J. Grossmann, J. Littleton, C. A. Makarewicz, J. Krigbaum, M. Burri, A. Scott, G. Davaasambuu, J. Wright, F. Irmer, E. Myagmar, N. Boivin, M. Robbeets, F. J. Rühli, J. Krause, B. Frohlich, J. Hendy, C. Warinner, Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11248–E11255 (2018).
- 40 72. S. Mallick, H. Li, M. Lipson, I. Mathieson, M. Gymrek, F. Racimo, M. Zhao, N. Chennagiri, S. Nordenfelt, A. Tandon, P. Skoglund, I. Lazaridis, S. Sankararaman, Q. Fu, N. Rohland, G. Renaud, Y. Erlich, T. Willems, C. Gallo, J. P. Spence, Y. S. Song, G. Poletti, F. Balloux, G. Van Driem, P. De Knijff, I. G. Romero, A. R. Jha, D. M. Behar, C. M. Bravi, C. Capelli, T. Hervig, A. Moreno-Estrada, O. L. Posukh, E. Balanovska, O. Balanovsky, S. Karachanak-Yankova, H. Sahakyan, D. Toncheva, L. Yepiskoposyan, C. Tyler-Smith, Y. Xue, M. S. Abdullah, A. Ruiz-Linares, C. M. Beall, A. Di Rienzo, C. Jeong, E. B. Starikovskaya, E. Metspalu, J. Parik, R. Villems, B. M. Henn, U. Hodoglugil, R. Mahley, A. Sajantila, G. Stamatoyannopoulos, J. T. S. Wee, R. Khusainova, E. Khusnutdinova, S.
- 45

- Litvinov, G. Ayodo, D. Comas, M. F. Hammer, T. Kivisild, W. Klitz, C. A. Winkler, D. Labuda, M. Bamshad, L. B. Jorde, S. A. Tishkoff, W. S. Watkins, M. Metspalu, S. Dryomov, R. Sukernik, L. Singh, K. Thangaraj, S. Paäbo, J. Kelso, N. Patterson, D. Reich, The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201-206 (2016).
- 5 73. L. Pagani, D. J. Lawson, E. Jagoda, A. Mörseburg, A. Eriksson, M. Mitt, F. Clemente, G. Hudjashov, M. Degiorgio, L. Saag, J. D. Wall, A. Cardona, R. Mägi, M. A. W. Sayres, S. Kaewert, C. Inchley, C. L. Scheib, M. Järve, M. Karmin, G. S. Jacobs, T. Antao, F. M. Iliescu, A. Kushniarevich, Q. Ayub, C. Tyler-Smith, Y. Xue, B. Yunusbayev, K. Tambets, C. B. Mallick, E. Pocheshkhova, G. Andriadze, C. Muller, M. C. Westaway, D. M. Lambert, G. Zoraqi, S. Turdikulova, D. Dalimova, Z. Sabitov, G. N. 10 N. Sultana, J. Lachance, S. Tishkoff, K. Momynaliev, J. Isakova, L. D. Damba, M. Gubina, P. Nymadawa, I. Evseeva, L. Atramentova, O. Utevska, F. X. Ricaut, N. Brucato, H. Sudoyo, T. Letellier, M. P. Cox, N. A. Barashkov, V. Škaro, L. Mulahasanović, D. Primorac, H. Sahakyan, M. Mormina, C. A. Eichstaedt, D. V. Lichman, S. Abdullah, G. Chaubey, J. T. S. Wee, E. Mihailov, A. Karunas, S. Litvinov, R. Khusainova, N. Ekomasova, V. Akhmetova, I. Khidiyatova, D. MarjanoviÄ, 15 L. Yepiskoposyan, D. M. Behar, E. Balanovska, A. Metspalu, M. Derenko, B. Malyarchuk, M. Voevoda, M. Voevoda, L. P. Osipova, M. M. Lahr, P. Gerbault, M. Leavesley, A. B. Migliano, M. Petraglia, O. Balanovsky, E. K. Khusnutdinova, E. Metspalu, M. G. Thomas, A. Manica, R. Nielsen, R. Villems, E. Willerslev, T. Kivisild, M. Metspalu, Genomic analyses inform on migration events during the peopling of Eurasia. *Nature* **538**, 238-242 (2016).
- 20 74. A. Scally, R. Durbin, Revising the human mutation rate: implications for understanding human evolution. *Nat. Rev. Genet.* **13**, 745-753 (2012).
75. J. N. Fenner, Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128**, 415-423 (2005).
- 25 76. M. Lipson, I. Ribot, S. Mallick, N. Rohland, I. Olalde, N. Adamski, N. Broomandkhoshbacht, A. M. Lawson, S. López, J. Oppenheimer, K. Stewardson, R. N. Asombang, H. Bocherens, N. Bradman, B. J. Culleton, E. Cornelissen, I. Crevecoeur, P. de Maret, F. L. M. Fomine, P. Lavachery, C. M. Mindzie, R. Orban, E. Sawchuk, P. Semal, M. G. Thomas, W. Van Neer, K. R. Veeramah, D. J. Kennett, N. Patterson, G. Hellenthal, C. Lalueza-Fox, S. MacEachern, M. E. Prendergast, D. Reich, 30 Ancient West African foragers in the context of African population history. *Nature* **577**, 665-670 (2020).
77. S. H. Martin, J. W. Davey, C. D. Jiggins, Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* **32**, 244-257 (2014).
78. E. Frichot, S. Schoville, G. Bouchard, O. François, Correcting principal component maps for effects of spatial autocorrelation in population genetic data. *Front. Genet.* 10.3389/fgene.2012.00254 (2012).
- 35 79. G. McVean, A genealogical interpretation of principal components analysis. *PLoS Genet.* **5**, 1000686 (2009).
80. D. Reich, A. L. Price, N. Patterson, Principal component analysis of genetic data. *Nat. Genet.* **40**, 491-492 (2008).
- 40 81. J. Novembre, M. Stephens, Interpreting principal component analyses of spatial population genetic variation. *Nat. Genet.* **40**, 646-649 (2008).
82. H. B. Hansen, P. B. Damgaard, A. Margaryan, J. Stenderup, N. Lynnerup, E. Willerslev, M. E. Allentoft, Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS One* **12**, e0170940 (2017).
- 45 83. M. T. P. Gilbert, H.-J. Bandelt, M. Hofreiter, I. Barnes, Assessing ancient DNA studies. *Trends Ecol. Evol.* **20**, 541-544 (2005).

84. E. Willerslev, A. Cooper, Ancient DNA. *Proc. Biol. Sci.* **272**, 3-16 (2005).
85. K. A. Sirak, D. M. Fernandes, O. Cheronet, M. Novak, B. Gamarra, T. Balassa, Z. Bernert, A. Cséki, J. Dani, J. Z. Gallina, G. Kocsis-Buruzs, I. Kővári, O. László, I. Pap, R. Patay, Z. Petkes, G. Szenthe, T. Szeniczey, T. Hajdu, R. Pinhasi, A minimally-invasive method for sampling human petrous bones from the cranial base for ancient DNA analysis. *BioTechniques* **62**, 283-289 (2017).
86. Meyer *et al.* (2010).
87. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448 (2010).
88. S. Andrews, A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
89. M. Schubert, S. Lindgreen, L. Orlando, AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
90. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
91. A. Peltzer, G. Jäger, A. Herbig, A. Seitz, C. Kniep, J. Krause, K. Nieselt, EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).
92. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
93. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682-1684 (2013).
94. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
95. D. J. Kennett, S. Plog, R. J. George, B. J. Culleton, A. S. Watson, P. Skoglund, N. Rohland, S. Mallick, K. Stewardson, L. Kistler, S. A. Leblanc, P. M. Whiteley, D. Reich, G. H. Perry, Archaeogenomic evidence reveals prehistoric matrilineal dynasty. *Nat. Commun.* **8**, 14115 (2017).
96. Q. Fu, H. Li, P. Moorjani, F. Jay, S. M. Slepchenko, A. A. Bondarev, P. L. F. Johnson, A. Aximu-Petri, K. Prüfer, C. de Filippo, M. Meyer, N. Zwyns, D. C. Salazar-García, Y. V. Kuzmin, S. G. Keates, P. A. Kosintsev, D. I. Razhev, M. P. Richards, N. V. Peristov, M. Lachmann, K. Douka, T. F. G. Higham, M. Slatkin, J.-J. Hublin, D. Reich, J. Kelso, T. B. Viola, S. Pääbo, Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449 (2014).
97. M. Lipson, D. Reich, A working model of the deep relationships of diverse modern human genetic lineages outside of Africa. *Mol. Biol. Evol.* **34**, 889-902 (2017).
98. H. McColl, F. Racimo, L. Vinner, F. Demeter, T. Gakuhari, J. Víctor Moreno-Mayar, G. Van Driem, U. G. Wilken, A. Seguin-Orlando, C. De la Fuente Castro, S. Wasef, R. Shoocongdej, V. Souksavatdy, T. Sayavongkhamdy, M. M. Saidin, M. E. Allentoft, T. Sato, A. S. Malaspinas, F. A. Aghakhanian, T. Korneliussen, A. Prohaska, A. Margaryan, P. de Barros Damgaard, S. Kaewsutthi, P. Lertrit, T. M. H. Nguyen, H. chun Hung, T. M. Tran, H. N. Truong, G. H. Nguyen, S. Shahidan, K. Wiradnyana, H. Matsumae, N. Shigehara, M. Yoneda, H. Ishida, T. Masuyama, Y. Yamada, A. Tajima, H. Shibata, A. Toyoda, T. Hanihara, S. Nakagome, T. Deviese, A. M. Bacon, P. Durringer, J. L. Ponche, L. Shackelford, E. Patole-Edoumba, A. T. Nguyen, B. Bellina-Pryce, J. C. Galipaud, R. Kinaston, H. Buckley, C. Pottier, S. Rasmussen, T. Higham, R. A. Foley, M. M. Lahr, L. Orlando, M.

- Sikora, M. E. Phipps, H. Oota, C. Higham, D. M. Lambert, E. Willerslev, The prehistoric peopling of Southeast Asia. *Science* **361**, 88-92 (2018).
- 5 99. M. Lipson, O. Cheronet, S. Mallick, N. Rohland, M. Oxenham, M. Pietrusewsky, T. O. Pryce, A. Willis, H. Matsumura, H. Buckley, K. Domett, G. H. Nguyen, H. H. Trinh, A. A. Kyaw, T. T. Win, B. Pradier, N. Broomandkhosbacht, F. Candilio, P. Changmai, D. Fernandes, M. Ferry, B. Gamarra, E. Harney, J. Kampuansai, W. Kutanan, M. Michel, M. Novak, J. Oppenheimer, K. Sirak, K. Stewardson, Z. Zhang, P. Flegontov, R. Pinhasi, D. Reich, Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science* **361**, 92-95 (2018).
- 10 100. C. Jeong, A. T. Ozga, D. B. Witonsky, H. Malmström, H. Edlund, C. A. Hofman, R. W. Hagan, M. Jakobsson, C. M. Lewis, M. S. Aldenderfer, A. Di Rienzo, C. Warinner, Long-term genetic stability and a high-altitude East Asian origin for the peoples of the high valleys of the Himalayan arc. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7485-7490 (2016).
- 15 101. M. Rasmussen, Y. Li, S. Lindgreen, J. S. Pedersen, A. Albrechtsen, I. Moltke, M. Metspalu, E. Metspalu, T. Kivisild, R. Gupta, M. Bertalan, K. Nielsen, M. T. P. Gilbert, Y. Wang, M. Raghavan, P. F. Campos, H. M. Kamp, A. S. Wilson, A. Gledhill, S. Tridico, M. Bunce, E. D. Lorenzen, J. Binladen, X. Guo, J. Zhao, X. Zhang, H. Zhang, Z. Li, M. Chen, L. Orlando, K. Kristiansen, M. Bak, N. Tommerup, C. Bendixen, T. L. Pierre, B. Grønnow, M. Meldgaard, C. Andreasen, S. a Fedorova, L. P. Osipova, T. F. G. Higham, C. B. Ramsey, T. V. O. Hansen, F. C. Nielsen, M. H. Crawford, S. Brunak, T. Sicheritz-Pontén, R. Villems, R. Nielsen, A. Krogh, J. Wang, E. Willerslev, Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757-762 (2010).
- 20 102. M. A. Yang, X. Fan, B. Sun, C. Chen, J. Lang, Y. C. Ko, C. H. Tsang, H. Chiu, T. Wang, Q. Bao, X. Wu, M. Hajdinjak, A. M. S. Ko, M. Ding, P. Cao, R. Yang, F. Liu, B. Nickel, Q. Dai, X. Feng, L. Zhang, C. Sun, C. Ning, W. Zeng, Y. Zhao, M. Zhang, X. Gao, Y. Cui, D. Reich, M. Stoneking, Q. Fu, Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* **369**, 282-288 (2020).
- 25 103. C. Osgood, *The ethnography of the Tanaina* (Yale University Publications in Anthropology, Yale University, New Haven, CT, 1937).
104. J. Kari, *Dena'ina topical dictionary* (Alaska Native Language Center, Fairbanks, AK, 2007).
- 30 105. J. Kari, "Some linguistic insights into Dena'ina prehistory" in *Athapaskan linguistics: current perspectives on a language family (Trends in Linguistics. State-of-the-Art Reports)*, E.-D. Cook, K. D. Rice, Eds. (Mouton de Gruyter, 1989), pp. 533-574.
106. J. Kari, "Linguistic traces of Dena'ina strategy at the archaic periphery" in *Adventures through time: readings in the anthropology of Cook Inlet, Alaska: proceedings of a symposium*, N. Yaw Davis, W. E. Davis, Eds. (Cook Inlet Historical Society, Anchorage, AK, 1996), pp. 51-63.
- 35 107. P. Y. Skorik, К вопросу о классификации чукотско-камчатских языков [On the classification of the Chukotko-Kamchatkan languages]. *Вопросы языкознания [Voprosy yazykozvaniya]* **1**, 21-35 (1958).
108. M. Fortescue, *Language relations across Bering Strait: reappraising the archaeological and linguistic evidence* (Open Linguistics Series, Cassell, London, 1998).
- 40 109. M. Fortescue, Diachronic typology and the genealogical unity of Chukotko-Kamchatkan. *Linguist. Typology* **7**, 51-88 (2003).
110. M. Fortescue, *Comparative Chukotko-Kamchatkan dictionary* (De Gruyter, Berlin, 2005).
- 45 111. T. Kurebito, M. Kurebito, Y. Nagayama, C. Ono, M. Yazu, *Comparative basic vocabulary of the Chukchee-Kamchatkan language family* (Endangered Languages of the Pacific Rim, Faculty of Informatics, Osaka Gakuin University, Suita, 2001), vol. 1.

112. O. A. Mudrak, *Этимологический словарь чукотско-камчатских языков [Etymological dictionary of the Chukotko-Kamchatkan languages]* (Языки русской культуры [Yazyki russkoy kultury], Moscow, 2000).
113. A. P. Volodin, *Ительменский язык [Itelmen language]* (Prosveshchenie, Leningrad, 1976).
- 5 114. D. S. Worth, La place du kamtchadal parmi les langues soi-disant paleosiberiennes. *Orbis* **11**, 579-599 (1962).
115. A. P. Volodin, "Чукотско-камчатские языки [Chukotko-Kamchatkan languages]" in *Языки мира: палеоазиатские языки [Languages of the World: Paleoasiatic languages]*, A. P. Volodin, Ed. (Indrik, Moscow, 1997), pp. 12-22.
- 10 116. S. Georg, A. P. Volodin, *Die itelmenische Sprache: Grammatik und Texte* (Tunguso-Sibirica, Harrassowitz, Wiesbaden, 1999), vol. 5.
117. S. G. Thomason, T. Kaufman, *Language contact, creolization, and genetic linguistics* (University of California Press, Berkeley, 1988).
118. O. A. Mudrak, "Itelmen etymology" in *The Tower of Babel. An etymological database project*, S. A. Starostin, Ed. (Moscow, 2005), <http://starling.rinet.ru/cgi-bin/main.cgi>.
- 15 119. O. A. Mudrak, *Свод камчатской лексики по памятникам XVIII века [Kamchatkan lexicon according to the 18th century sources]* (Восточная литература [Vostochnaya Literatura], Moscow, 2008).
120. A. S. Kastian *et al.*, <https://www.academia.edu/39903804/> (2019).
- 20 121. A. S. Kastian, G. S. Starostin, M. A. Zhivlov, Proto-Indo-European-Uralic comparison from the probabilistic point of view. *J. Indo-Eur. Stud.* **43**, 301-347 (2015).
122. M. Fortescue, *Comparative Nivkh dictionary* (Languages of the World. Dictionaries, LINCOM GmbH, München, 2016), vol. 62.
- 25 123. D. Anthony, *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World* (Princeton University Press, Princeton, NJ, 2007).
124. J. Nichols, *Linguistic diversity in space and time* (University of Chicago Press, Chicago, 1992).
125. Y. Berezkin, *Мифы заселяют Америку [Myths Populate America]* (OGI, Moscow, 2007).
126. M. Fortescue, E. Vajda, *Mid-Holocene Language Connections Across Bering Strait* (Brill, Leiden, in press).
- 30 127. O. Mudrak, S. Nikolaev, *Gilyak and Chukchi-Kamchatkan as Almosan-Keresiouan languages: Lexical evidence (Preliminary report)* (Explorations in Language Macrofamilies, Universitätsverlag Dr. Norbert Brockmeyer, Bochum, 1989), pp. 67-87.
128. S. Nikolaev, Toward the reconstruction of Proto-Algonquian-Wakashan. Part 1: Proof of the Algonquian-Wakashan relationship. *J. Language Relationship* **13**, 23-61 (2015).
- 35 129. S. Nikolaev, Toward the reconstruction of Proto-Algonquian-Wakashan. Part 2: Algonquian-Wakashan sound correspondences. *J. Language Relationship* **13**, 289-328 (2015).
130. L. R. Morgan, *Kootenay-Salishan linguistic comparison. A preliminary study* (M.A. thesis, University of British Columbia, Vancouver, 1982).
- 40 131. M. Dryer, "Kutenai, Algonquian and the Pacific Northwest from an areal perspective" in *Papers of the 38th Algonquian Conference*, H. C. Wolfart, Ed. (University of Manitoba, Winnipeg, 2007), pp. 155-206.

132. J. H. Greenberg, *Language in the Americas* (Stanford University Press, Stanford, CA, 1987).
133. I. Goddard, "Comparative Algonquian" in *The languages of Native America: Historical and comparative assessment*. L. Campbell, M. Mithun, Eds. (University of Texas Press, Austin, TX, 1979), pp. 70-132.
- 5 134. L. Campbell, W. J. Poser, *Language Classification: History and Method* (Cambridge University Press, Cambridge, 2008).
135. H. Matsumae *et al.*, <https://www.biorxiv.org/content/10.1101/513929v1> (2019).
136. E. Vajda, "Patterns of innovation and retention in templatic polysynthesis" in *Handbook of Polysynthesis*, M. Fortescue, M. Mithun, N. Evans, Eds. (Oxford University Press, Oxford, 2017), pp. 363-391.
- 10 137. P. Bakker, "Algonquian-Ritwan (Kutenai) and Salish: Proving a distant genetic relationship" in *Papers for the 41st International Conference on Salish and neighbouring languages* (University of British Columbia Press, Vancouver, 2006), pp. 1-32.
138. D. Beck, Grammatical convergence and the genesis of diversity in the Northwest Coast Sprachbund. *Anthropol. Linguistics* **42**, 1-67 (2000).
- 15 139. E. Sapir, "Central and North American languages" in *Encyclopedia Britannica* (A & C Black, London, ed. 14, 1929), vol. 5, pp. 128-140.
140. I. Lazaridis, N. Patterson, A. Mitnik, G. Renaud, S. Mallick, K. Kirsanow, P. H. Sudmant, J. G. Schraiber, S. Castellano, M. Lipson, B. Berger, C. Economou, R. Bollongino, Q. Fu, K. I. Bos, S. Nordenfelt, H. Li, C. de Filippo, K. Prüfer, S. Sawyer, C. Posth, W. Haak, F. Hallgren, E. Fornander, N. Rohland, D. Delsate, M. Francken, J.-M. Guinet, J. Wahl, G. Ayodo, H. a. Babiker, G. Bailliet, E. Balanovska, O. Balanovsky, R. Barrantes, G. Bedoya, H. Ben-Ami, J. Bene, F. Berrada, C. M. Bravi, F. Brisighelli, G. B. J. Busby, F. Cali, M. Churnosov, D. E. C. Cole, D. Corach, L. Damba, G. van Driem, S. Dryomov, J.-M. Dugoujon, S. A. Fedorova, I. Gallego Romero, M. Gubina, M. Hammer, B. M. Henn, T. Hervig, U. Hodoglugil, A. R. Jha, S. Karachanak-Yankova, R. Khusainova, E. Khusnutdinova, R. Kittles, T. Kivisild, W. Klitz, V. Kučinskás, A. Kushniarevich, L. Laredj, S. Litvinov, T. Loukidis, R. W. Mahley, B. Melegh, E. Metspalu, J. Molina, J. Mountain, K. Näkkäläjärvi, D. Nesheva, T. Nyambo, L. Osipova, J. Parik, F. Platonov, O. Posukh, V. Romano, F. Rothhammer, I. Rudan, R. Ruizbakiev, H. Sahakyan, A. Sajantila, A. Salas, E. B. Starikovskaya, A. Tarekegn, D. Toncheva, S. Turdikulova, I. Uktveryte, O. Utevska, R. Vasquez, M. Villena, M. Voevoda, C. a. Winkler, L. Yepiskoposyan, P. Zalloua, T. Zemunik, A. Cooper, C. Capelli, M. G. Thomas, A. Ruiz-Linares, S. a. Tishkoff, L. Singh, K. Thangaraj, R. Villems, D. Comas, R. Sukernik, M. Metspalu, M. Meyer, E. E. Eichler, J. Burger, M. Slatkin, S. Pääbo, J. Kelso, D. Reich, J. Krause, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 399-413 (2014).
- 20
25
30
35

Acknowledgments:

Funding: This work was supported by the Major project of Humanities and Social Sciences Key Research Base of the Ministry of Education of China (16JJD780005), the Major project of the National Social Science Fund of China (15ZDB055), the Czech Ministry of Education, Youth and Sports from the Large Infrastructures for Research, Experimental Development and Innovations project „IT4Innovations National Supercomputing Center – LM2015070". C.N. was supported by the European Research Council (646612) grant to M.R. and the Max Planck Society. E.Y., O.F., P.C., and P.F., were supported by the Institutional Development Program of the University of Ostrava. O.F. was supported by the program PPLZ of the Czech Academy of

40
45

Sciences. D.R. was supported by the NIH (NIGMS) grant GM100233, the Paul Allen Foundation, and the John Templeton Foundation grant 61220, and is an Investigator of the Howard Hughes Medical Institute. **Author contributions:** P.F., R.P., and Y.C. supervised the study. L.W., X.X., R.P., and Y.C. assembled the collection of archaeological samples. L.W. was responsible for radiocarbon dating and calibration. C.N., D.F., K.S., R.P., and Y.C. performed laboratory work and supervised ancient DNA sequencing. P.F., C.N., P.C., D.F., O.F., E.Y., N.E.A., C.L.-F., K.W., and S.S., analyzed genetic data, R.M. developed novel software, and B.A.P. prepared archaeological maps. A.S.K. and E.J.V. wrote the supplemental sections on linguistics. P.F., C.N., and R.P. wrote the manuscript with additional input from all other co-authors. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** Raw sequence data (bam files) from the 12 newly reported ancient individuals have been deposited at the Genome Sequence Archive at the China National Genomics Data Center under accession number HRA000268. Custom code used in this manuscript is available at a dedicated github repository: <https://github.com/uqrmaie1/admixtools>.

Supplementary Materials:

Materials and Methods

Supplementary Text, sections 1-14

Figures S1-S55

Tables S1-S25

References (42-140)

Data S1-S10



Fig. 1. Geographic location of the Houtaomuga site analyzed in this study. The table in the inset includes chronological information and genomic coverage for all screened individuals. Bold and italicized individual codes denote individuals used for population genetic analysis; asterisks denote individuals used for *qpGraph* modeling. Dates are based on confidence intervals from radiocarbon dating results (Data S3). (A, B) Ceramic shards and vessels from Phases II and IV, respectively. (C) The burial of individual M45. The Amur river basin is marked by color.

5

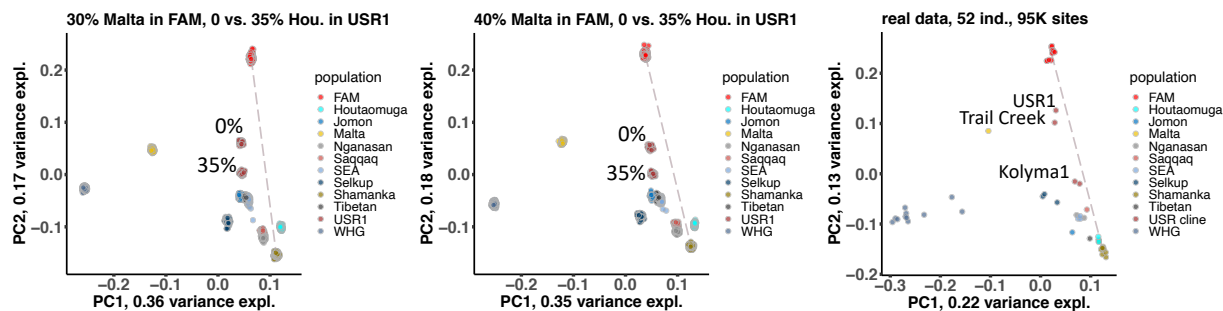


Fig. 2. PC1/PC2 plots constructed using PLINK on simulated (A,B) and real data (C). PCs were calculated using 50 simulated ind. and the same number of real counterparts. In the case of the real data, 3 USR1-related individuals were added: Trail Creek and two versions of the Kolyma 1 genome (labelled in panel c). Dataset sizes in terms of sites were also nearly identical, ca. 95,000. Centroids of the FAM and Shamanka EN clusters are connected with a dashed line. Plots for 100 simulation iterations were overlapped. Panels A and B show results for different ratios of Mal'ta ancestry in the simulated FAM and USR1 groups: 30/30% and 40/30%. Clusters of USR1 points are labelled according to the simulated proportion of Houtaomuga ancestry.

5

10

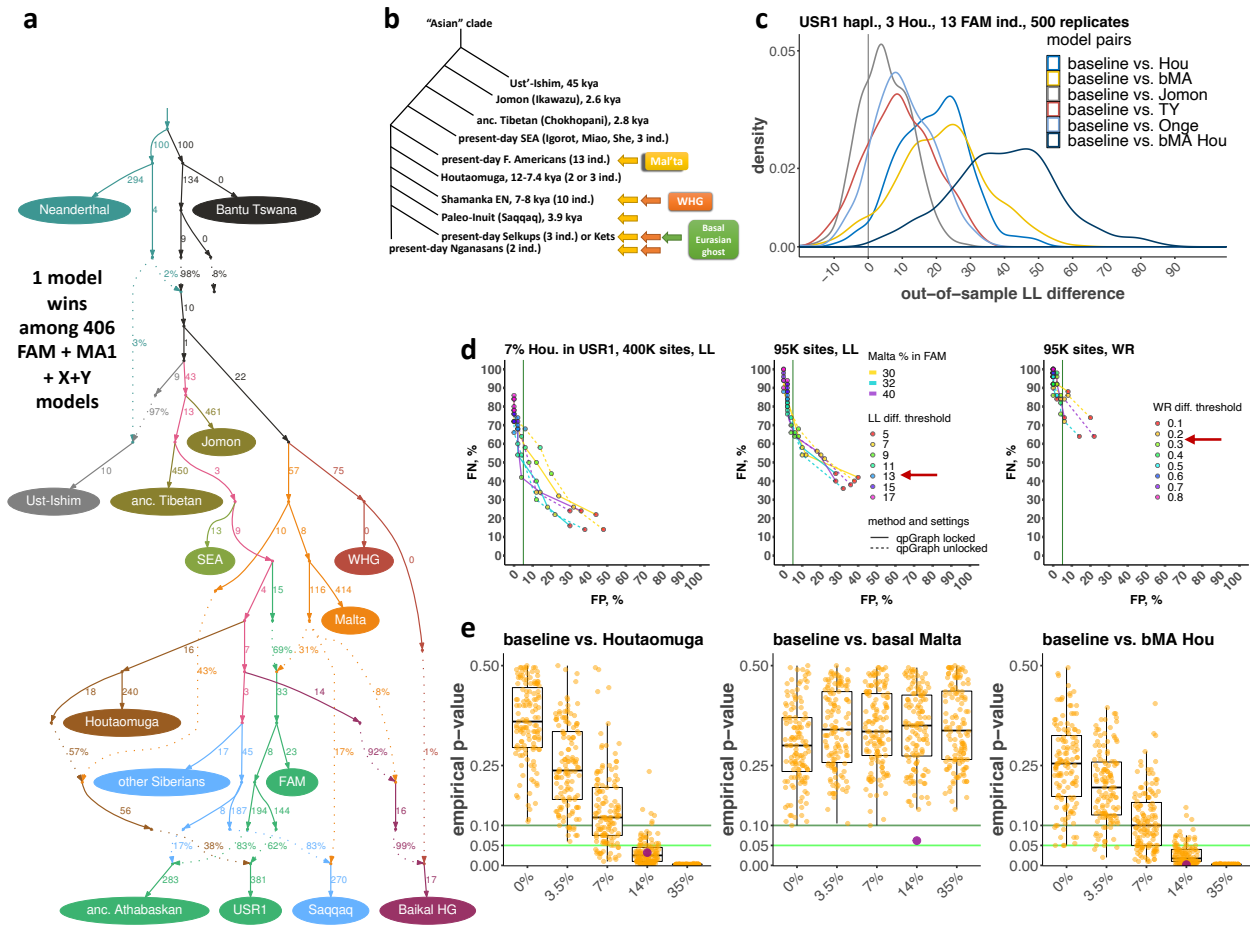


Fig. 3. Modelling the genetic history of the USR1 individual. The graph in (A) presents a model in the 4-way admixed class that is significantly different from the other 405 models of that class. Panel (B) displays the branching order in the Asian clade and the West Eurasian gene flows incorporated into the skeleton model. In panel (D) false positive and false negative rates of the graph mapping protocol on simulated genetic data are shown (baseline and "FAM_{EA} + Mal'ta + X" models compared). LL and WR difference thresholds are coded by point colour, and the optimized thresholds are marked by arrows. Panel (E) shows results of out-of-sample testing of admixture graph models on resampled replicates of simulated data. Each dot shows a one-tailed empirical p -value. The analysis was repeated for 130 replicates of the history and for different proportions of *Houtaomuga* ancestry in USR1. The bright green line marks the significance threshold of 0.05. Model comparison p -values on real data (ca. 407,000 sites across all groups) are shown as magenta dots in panel E, and the distributions of out-of-sample LL differences across 500 bootstrap replicates are shown in panel C for six model pairs labelled according to the legend.

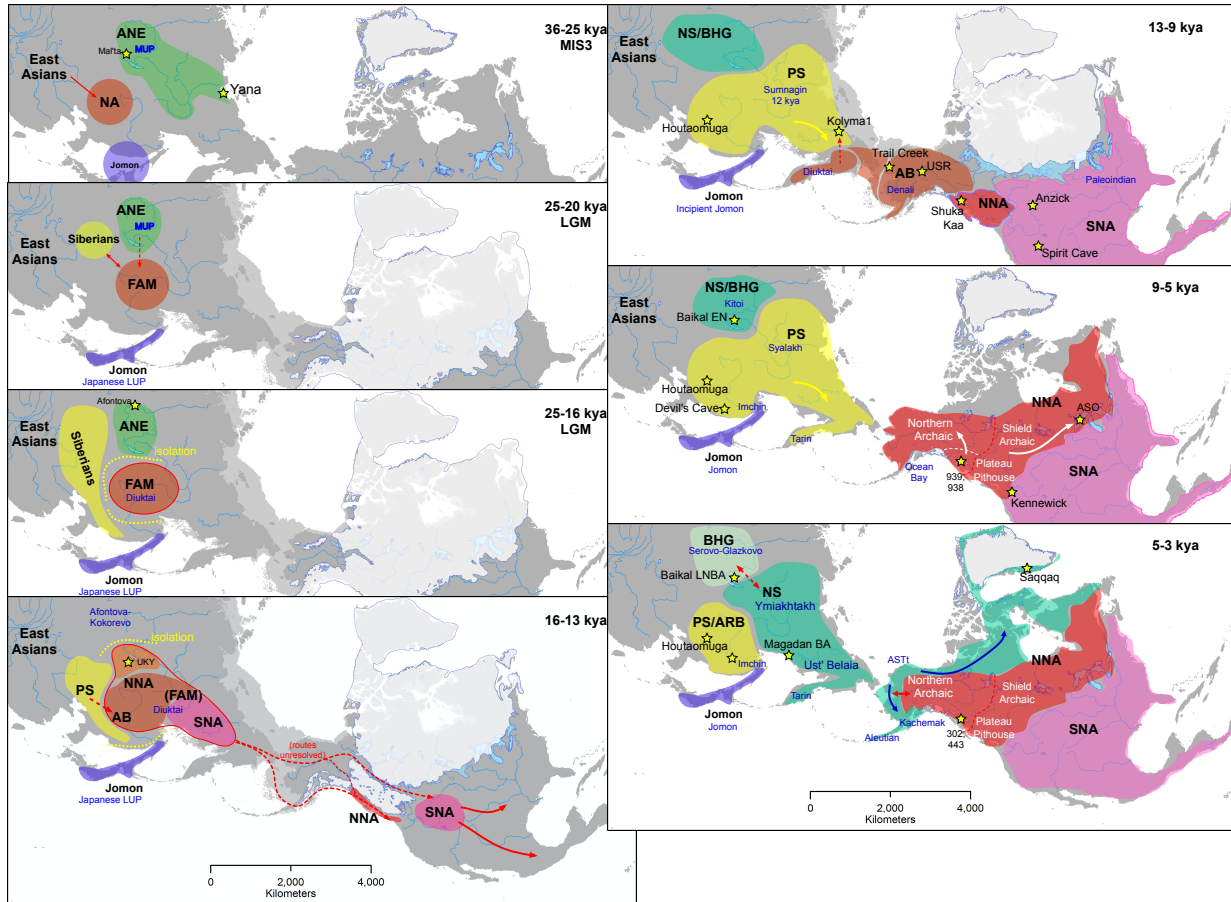


Fig. 4. Maps of archaeological cultures and migrations in Northeast Asia, North and Central America. Seven "time slices" are shown. Key archaeological sites and ancient genomes are marked with yellow stars, archaeological cultures with blue text. Past shorelines and glaciers are also shown. The following abbreviations are used: AB, Ancient Beringians *sensu* Moreno-Mayar *et al.* (4), ANE, ancient North Eurasians *sensu* Raghavan *et al.* (40); BHG, Baikal hunter-gatherers *sensu* Damgaard *et al.* (41), FAM, the First American lineage in Asia and America; LGM, last glacial maximum, LUP, Late Upper Paleolithic; MUP, Middle Upper Paleolithic; NNA and SNA, northern and southern Native American lineages *sensu* Raghavan *et al.* (15); NS and PS, Neo- and Paleo-Siberians (the meaning of the terms is different from that used by Sikora *et al.* (10)); UKY, the Ust'-Kyakhta-3 site.