Ancient DNA reveals the prehistory of the Uralic and Yeniseian peoples

https://doi.org/10.1038/s41586-025-09189-3

Received: 12 September 2023

Accepted: 23 May 2025

The North Eurasian forest and forest-steppe zones have sustained millennia of sociocultural connections among northern peoples, but much of their history is poorly understood. In particular, the genomic formation of populations that speak Uralic and Yeniseian languages today is unknown. Here, by generating genome-wide data for 180 ancient individuals spanning this region, we show that the Early-to-Mid-Holocene hunter-gatherers harboured a continuous gradient of ancestry from fully European-related in the Baltic, to fully East Asian-related in the Transbaikal. Contemporaneous groups in Northeast Siberia were off-gradient and descended from a population that was the primary source for Native Americans, which then mixed with populations of Inland East Asia and the Amur River Basin to produce two populations whose expansion coincided with the collapse of pre-Bronze Age population structure. Ancestry from the first population, Cis-Baikal Late Neolithic–Bronze Age (Cisbaikal_LNBA), is associated with Yeniseian-speaking groups and those that admixed with them, and ancestry from the second, Yakutia Late Neolithic–Bronze Age (Yakutia_LNBA), is associated with migrations of prehistoric Uralic speakers. We show that Yakutia_LNBA first dispersed westwards from the Lena River Basin around 4,000 years ago into the Altai-Sayan region and into West Siberian communities associated with Seima-Turbino metallurgy—a suite of advanced bronze casting techniques that expanded explosively from the Altai[1]. The 16 Seima-Turbino period individuals were diverse in their ancestry, also harbouring DNA from Indo-Iranian-associated pastoralists and from a range of hunter-gatherer groups. Thus, both cultural transmission and migration were key to the Seima-Turbino phenomenon, which was involved in the initial spread of early Uralic-speaking communities.

Uralic languages are spoken across Northern Eurasia, from Central Europe to Northeastern Siberia, but their homeland has been debated, with theories pointing to the Altai-Sayan mountains, between the Ob' and Yenisei in Siberia, Europe around the confluence of the Volga and Kama rivers, or the East Baltic[2] (Supplementary Information, section 1 gives a guide to geographic terms). Present-day Uralic speakers differ systematically from their Indo-European speaking neighbours in having substantial Siberian ancestry (from around 2% in Estonians to almost all in Nganasans), and a high frequency of Y-chromosome haplogroup N lineages of Siberian origin[3]. Time transects of ancient DNA show that this ancestry arrived in Europe around 3.5 thousand years ago (ka) in Karelia[4] and around 2.6 ka in the East Baltic[5]. In contrast to Indo-European languages, which can be traced by the dispersal of ancestry from the Yamnaya of the European steppe, no genetic 'tracer dye' has been found for the prehistoric dispersals of Uralic-speaking populations.

Yeniseian languages are attested only in populations along the middle and upper Yenisei, and Ket is the sole extant language. However, Yeniseian languages had a broader geographic spread in the past, and are linked in deep time with Na-Dene (Athabaskan–Eyak–Tlingit) languages of North America such as Chipewyan and Navajo, spoken from Alaska to Arizona[6]. Prior studies tried to find a genetic connection between Athabaskans and Kets[7–10], but this has been challenging owing to the genetic similarity of Kets to their non-Yeniseian neighbours[11]. The disruptive effects of migrations associated with the later spread of Indo-European, Turkic and Mongolic languages[12,13] also make it difficult to reconstruct the prehistoric migrations of Uralic and Yeniseian speakers on the basis of genetic variation in present-day people.

We generated genome-wide data for 180 individuals across Northern Eurasia for archaeological cultures from the Mesolithic (approximately 11 ka) to the Bronze Age (approximately 4.0 ka), from the Volga-Ural region to the Lena River Valley of Central Siberia (Fig. 1). Extended Data Fig. 1 provides a map of sites and Extended Data Fig. 2 provides a comprehensive chart showing the placements of the sites within geographic regions and in the archaeological cultures in each region's cultural chronology; archaeological context for each site and culture is provided and organized by region in Supplementary Information, section 3. The Supplementary Information Guide provides the information needed to find the section of the Supplementary Information corresponding to each reference in the text, as well as descriptions Supplementary Tables 1–35 and Supplementary Figs. 1–101, referred to only within the Supplementary Information. We used in-solution enrichment for more than 1.2 million single nucleotide polymorphisms (SNPs) (Methods). We merged with data from 1,312 previously reported ancient individuals from relevant locations and time periods. We also report 88 direct radiocarbon dates, which should be viewed with

---

A list of authors and their affiliations appears at the end of the paper.

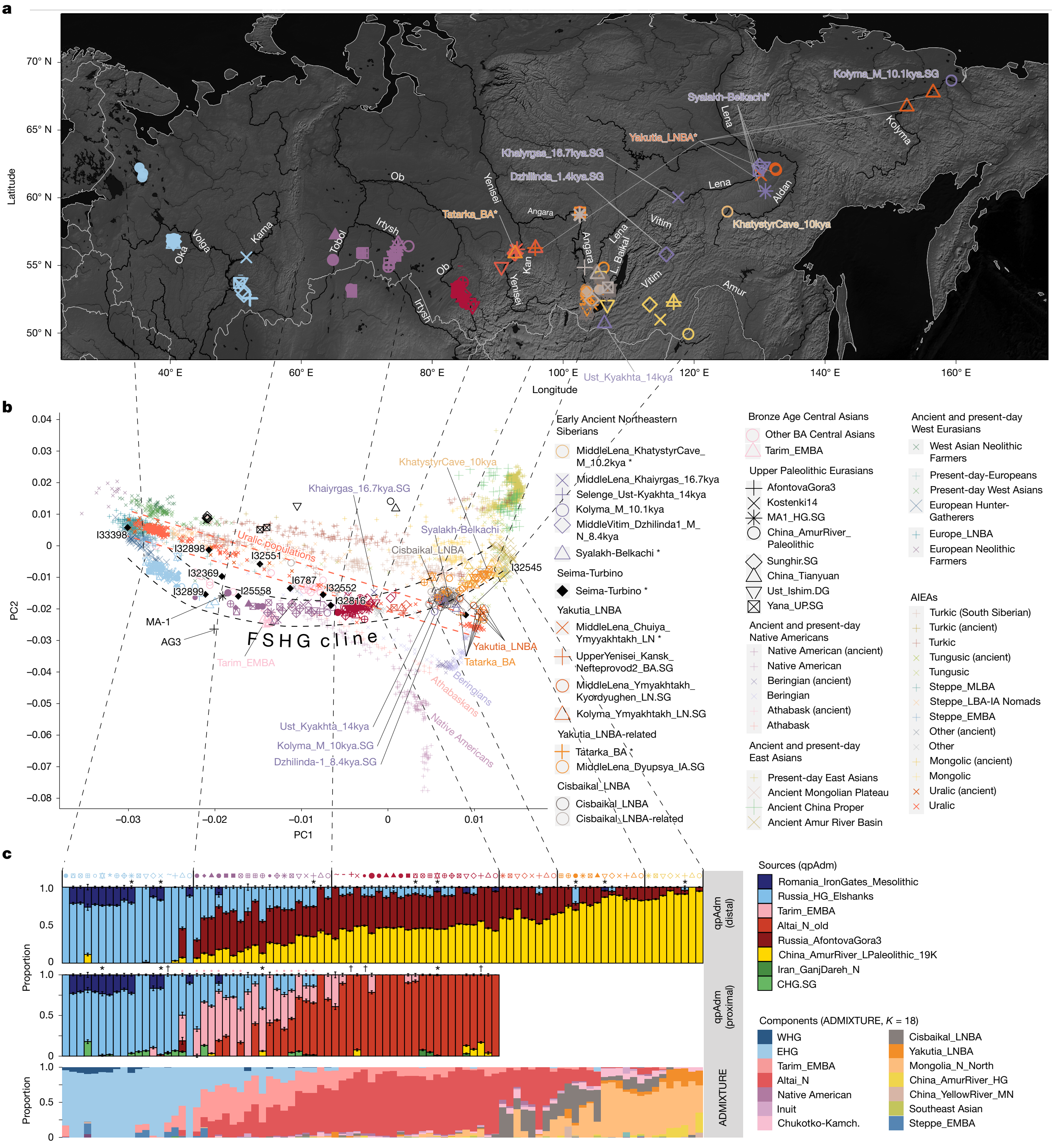


Fig. 1 | See next page for caption.

caution owing to freshwater reservoir effects that can cause overestimates up to a millennium[14] (Supplementary Information, section 2). Our population labels identify genetically homogeneous individuals from a site (Region_Site_ArchaeologicalPeriod_Time), but for some analyses we use more aggregated groupings (glossary in Table 1; see Extended Data Fig. 2 for the geographic and temporal placements of sites and their archaeological cultures).

We performed unsupervised genetic analyses, including principal component analysis (PCA) and ADMIXTURE (Fig. 1 and Supplementary Information, sections 4 and 5), which show that individuals from a belt of pottery-using foraging cultures in the North Eurasian forest-steppe and southern edge of the forest zone around 10–5 ka, form a genetic gradient stretching across approximately 7,000 km (Fig. 1b,c and Extended Data Figs. 3–5) that no longer exists today. We call this the North Eurasian hunter-gatherer (NEAHG) cline. The centre of this cline lies close to the Ancient North Eurasian (ANE) individual Afontova-Gora 3 (AG3), and early Bronze Age people of the Tarim Basin (Tarim_EMBA[15]; Fig. 1b and Extended Data Figs. 3–5). However, many other populations do not

Fig. 1 | The NEAHG cline and its legacy through admixture in ancient northern Eurasia. **a**, Sampling locations of all individuals and selected samples that are mentioned in the text but not on the cline. A higher-resolution version of this image, with all population labels indicated, can be found in Extended Data Fig. 11. **b**, PCA. We project ancient and present-day data onto variation from 122 genotyped present-day Eurasian and Native American populations that were selected to have minimal sub-Saharan African and Oceanian admixture. We observe a continent-spanning NEAHG cline, as well as a cline for Uralic populations stretching from European and Bronze Age Steppe populations to present-day Nganasans, Yakutia_LNBA individuals and the Seima-Turbino period site of Tatarka. In the legend, population labels that include newly sequenced samples are marked with an asterisk. **c**, Admixture proportions for NEAHG cline populations. Top, qpAdm estimates of ancestry related to four sources (Russia_AfontovaGora for ANE, China_AmurRiver_LPaleolithic_19K for East Asian, Russia_HG_Elshanks for EHG, and Romania_IronGatesMesolithic for WHG) for all populations on the NEAHG cline. 84 out of 93 have passing models ($P > 0.01$); populations that do not have an asterisk above the bar plot. In these cases, we show the model with the highest P value. The error bars indicate half a standard error. Middle, estimated admixture proportions for all eight sources in the legend (expanding to include Tarim_EMBA1, Altai_N_9kya, Iran_GanjDareh_N and CHG); a pink dot above the bar plot indicates that all passing qpAdm models have Tarim_EMBA in the sources. A cross indicates a population used as a source (Altai_N_9kya; Russia_MiddleVolga_Elshanka_Chekhalino_4_10kya). For the Elshanka individual, we replaced the EHG source with Russia_Veretye_Mesolithic.SG. Bottom, ADMIXTURE results at $K = 18$ for populations of the NEAHG cline. All populations have their ancestry assigned to components that are maximized among in EHG, Tarim_EMBA, Altai_N, and East Asian populations.

fall on this cline, including Central and Northeast Siberian populations from further north (from deeper in the forest zone or from the Arctic), populations of the Amur Basin, and populations from the Cis-Baikal region after around 5 ka.

To obtain insight into the genetic differences and population changes from around 17 ka to 4.0 ka in this region, we proceeded to group into populations individuals along the NEAHG cline, those from deeper into Northeastern Siberia and those around the Lake Baikal, and then analysed them with a suite of population genetic methods. Extended Data Fig. 6 summarized our key findings graphically. In the following sections, we present our analyses in order: first, on the population histories of Northeastern Siberia; next, on the NEAHG cline; and then, we dive into the connections that link two Bronze Age population groupings: Cisbaikal_LNBA and Yakutia_LNBA, and Yeniseian- and Uralic-speaking populations.

Palaeosiberian legacy in Asia and America

To investigate the population history of this region, we clustered 100 Holocene individuals from Northeastern Siberia and the Cis-Baikal and Transbaikal regions into genetic populations using f_4-statistics. These groupings largely coincide with archaeological cultures (Supplementary Information, section 6 and Supplementary Data 1). We identified seven clusters: five with multiple members and two with single individuals. In chronological order, these are: MiddleLena_KhatystyrCave_M_10.2kya (a newly reported individual, around 10.2 ka from Khatystyr Cave along the Middle Lena), MiddleVitim_Dzhilinda1_M_N_8.4kya (at the Mesolithic-Neolithic boundary, from the Dzhilinda-1 site along the Vitim river from the Ust'-Yumurchen culture[16]), Transbaikal_EMN (8.8–6.2 ka from the Early and Middle Neolithic Kitoi culture east of the Baikal), Cisbaikal_EN (8.0–6.6 ka from the Early Neolithic Kitoi culture west of the Baikal), Syalakh-Belkachi (6.8–6.2 ka from the Early Neolithic Syalakh and Middle Neolithic Belkachi cultures of the Middle Lena Basin), Cisbaikal_LNBA (5.1–3.7 ka from the Late Neolithic and Bronze Age Serovo, Isakovo and Glazkovo cultures west of the Baikal) and Yakutia_LNBA (4.5–3.2 ka, associated chiefly with the Ymyyakhtakh culture). The remaining individuals were genetically intermediate and consistent with being admixtures of other groups that we analysed. In addition to these seven clusters, we added three older individuals: MiddleLena_Khaiyrgas_16.7kya[16], Selenge_Ust-Kyakhta_14kya[17] and Kolyma_M_10.1kya[9], producing a ten-member transect (Extended Data Fig. 7).

We used qpAdm to model each target population as derived from ones preceding or contemporary to them with the 'outgroup rotation' method, which directly competes initially fitting models against each other to find best fits[18] (Supplementary Information, sections 7 and 8 and Methods). For each of the ten populations, we found one or a small number of qualitatively similar passing models ($P > 0.05$). All scans through large numbers of models are expected to allow through some models that are incorrect[18,19], so our protocol should not be viewed as a robust model-selection procedure; instead, it should be viewed as a model-rejection procedure and passing models should be inspected to identify consistent findings, as we do here. The data were also generated through a mixture of wet lab processes (SNP enrichment[8] and shotgun sequencing), which raises concerns about false inferences due to technical biases that have nothing to do with population history[20]. However, our inferences can be replicated in qpAdm setups utilizing sequences generated with a single wet laboratory process (Supplementary Information, section 7). Our qualitative findings are also consistent with simple f_4-statistics that test for affinities with distantly related populations that are plausibly relevant to the peopling of Northeast Siberia (Fig. 2a,b and Supplementary Information, section 8).

The oldest individual in our transect, MiddleLena_Khaiyrgas_16.7kya[16] (from the Middle Lena in Yakutia, of the Dyuktai culture; Supplementary Information, section 3) fits as a sister group of Native Americans, and can be modelled as descending completely from a Native American-related source (Supplementary Information, section 8). The term Ancient Palaeosiberian (APS) was used to designate the ancestry of the third individual in our transect (Kolyma_M_10.1kya)[9], and here we broaden this term to designate the pre-Holocene (meta)population, admixed between ANE and East Asian ancestries that gave rise to Native Americans, of which MiddleLena_Khaiyrgas_16.7kya may be a near-unadmixed representative. APS ancestry had a key role in the genetic formation of all later groups in our Siberian transect and the North American Arctic.

Our findings indicate that APS ancestry may have spread in the Northeast Siberian Upper Palaeolithic with the 'Beringian tradition' of lithics rich in conical and wedge-shaped microcores[21]. It persists in high levels, but admixed with additional East Asian ancestry, in two later individuals, Selenge_Ust-Kyakhta_14kya[17] (south of Lake Baikal on the Selenge River with lithics from this same tradition) and Kolyma_M_10.1kya (close to the Bering Straits[9]; Supplementary Information, section 8). Further west in the Altai, by the early Holocene (around 9 ka), admixture between APS and an ANE-related source formed Altai_N on the NEAHG cline (Supplementary Information, section 8), associated with the Neolithic Kuznetsk-Altai culture of the Upper Ob' and Altai foothills (Supplementary Information, section 3).

Prior work has shown that 'Neosiberian' East Asian ancestry increased while APS ancestry declined in Northeast Siberia throughout the Holocene[9]. We find that this increasing East Asian ancestry can be traced to at least two sources: Inland Northeast Asian-related ancestry, which we proxy by the Inner Mongolian Yumin individual[22] around 8.4 ka (China_NEastAsia_Inland_EN), and Amur Basin-related ancestry, represented by pre-Holocene hunter-gatherers of the Amur Basin[23] around 14 ka (China_AmurRiver_14K). The oldest individual in our Siberian transect with high East Asian and low APS ancestry, MiddleLena_KhatystyrCave_M_10.2kya, had strong affinities to Amur River hunter-gatherers (Fig. 2b and Supplementary Information, section 8), but subsequent populations further south (including the Kitoi-associated Transbaikal_EMN and Cisbaikal_EN at 8.8–6 ka, and

Table 1 | Glossary of acronyms

Term	Usage	Meaning
_M_	Term used to designate the archaeological period in population labels	Mesolithic
_N_	Term used to designate the archaeological period in population labels	Neolithic. Note that in Russian archaeological literature and in the archaeology of much of Northern Eurasia, the Neolithic period is defined by the presence of pottery, and not of agriculture or domesticated animals.
_EN_, _MN_, _EMN_	Term used to designate the archaeological period in population labels	Early Neolithic, Middle Neolithic, Early and Middle Neolithic
_BA_, _EBA_, _LBA_, _MLBA_, _LNBA_	Term used to designate the archaeological period in population labels	Bronze Age, Early Bronze Age, Late Bronze Age, Middle and Late Bronze Age, Late Neolithic and Bronze Age
APS	Acronym used to refer to an ancestry type	Ancient Palaeosiberian ancestry—a term referring to an ancient Siberian population related to the ancestors of Native American populations, who admixed into all later Eastern and Central Siberian populations as well as present-day populations on either side of the Bering Straits
NEAHG	Acronym	North Eurasian hunter-gatherers—a term designating a belt of hunter-gatherer populations spanning Northern Eurasia in the first half of the Holocene.
AIEA	Acronym	Admixed Inner Eurasians—a term designating all populations in Central and Northern Eurasia that are the product of Holocene admixtures between West Eurasian ancestries and East Asian ancestries, including present-day and ancient Mongolic, Turkic, Tungusic and Uralic populations, as well as ancient Scythians, Sarmatians and pre-Scythian nomads of the Iron Age Steppes.

Mongolia_N_North at around 7.5 ka) have increasing affinities to the Inland Northeast Asian source (Fig. 2b). We find that later groups are differentiated by their mix of East Asian ancestries: individuals falling along the NEAHG cline, including Cisbaikal_EN and Transbaikal_EMN, have a characteristic mixture of East Asian ancestries that is intermediate in affinity between the Inland and Amur-related sources, but non-NEAHG populations, such as foragers from the Amur River Basin or Cisbaikal_LNBA, have different ratios (Supplementary Information, section 8).

In the Cis-Baikal region during the mid-Holocene, ancestry from Cisbaikal_EN (8–6.6 ka) was replaced by Cisbaikal_LNBA (5.1–3.7 ka), in a turnover coinciding with the transition from the Early Neolithic Kitoi to the Late Neolithic and Bronze Age Serovo, Isakovo and Glazkovo cultures (Supplementary Information, section 8). Cisbaikal_LNBA is much higher in APS ancestry than its predecessors, which can only be modelled as deriving from an Ust-Kyakhta_14kya-related source, but this result should be viewed with caution owing to the long time gap separating the two populations. Cisbaikal_LNBA is also distinctive in having the most strongly Inland-related East Asian ancestry in our transect (Fig. 2b and Supplementary Information, sections 8 and 11). Despite increased APS ancestry, Cisbaikal_LNBA does not have increased shared drift with Native Americans or Bering Straits populations compared with other groups that are similar mixtures of ANE and East Asian ancestry (such as Ust-Kyakhta_14kya, Khaiyrgas_16.7kya. SG or NEAHG populations from the Upper Yenisei; Figs. 1b and 2a). Instead, it shares high levels of drift with present-day populations from Central Siberia, especially the Yenisei River Basin (Extended Data Fig. 8 and Supplementary Information, section 8). We show below that Cisbaikal_LNBA-related ancestry may be the first of two routes by which APS ancestry persisted into present-day populations, here those of in Central Siberia; that is, it is a 'Route 1' population (Figs. 2c and 3b and Supplementary Information, section 8).

North of the Baikal region along the Lena, the MiddleLena_KhatystyrCave_M_10.2kya individual derived most ancestry from Amur Basin hunter-gatherers, but admixture from a Kolyma_M_10.1kya-related source caused an increase in APS ancestry in the following MiddleVitim_Dzhilinda1_M_N_8.4kya (Supplementary Information, section 8). APS ancestry then declined with admixture from East Asian sources, in a set of population turnovers that seem to coincide with transitions between archaeological cultures. The first saw the transition from MiddleVitim_Dzhilinda1_M_N_8.4kya to the Syalakh-Belkachi population (6.8–6.2 ka), with around 20% admixture from an East Asian source from the Baikal region. The second saw another approximately 50% admixture into Syalakh-Belkachi from Transbaikal_EMN to create the Ymyyakhtakh-associated Yakutia_LNBA population (4.5–3.2 ka).

This sequence of four populations in Northeast Siberia (Kolyma_M_10.1kya, MiddleVitim_Dzhilinda1_M_N_8.4kya, Syalakh-Belkachi and Yakutia_LNBA) is uniquely shifted towards Native Americans and Bering Straits populations in PCAs (Fig. 1, Extended Data Figs. 5 and 8 and Supplementary Information, section 8). In f_4-statistics, they share more drift with ancient and present-day Bering Straits populations than any groups with similar proportions of ANE and East Asian ancestry (Khaiyrgas_16.7kya, Ust-Kyakhta_14kya, Cisbaikal_LNBA and all NEAHGs east of the Altai; Fig. 2a). Using qpAdm, we confirm that the third member of this sequence—Syalakh-Belkachi—made a major (around 70%) contribution to people of the Arctic Small Tool Tradition in North America (represented by the Palaeo-Eskimo Greenland_Saqqaq.SG and other individuals from the Dorset and related cultures, also reported elsewhere[16]; Supplementary Information, section 9). This Syalakh-Belkachi-related Palaeo-Eskimo ancestry persisted in all later populations around the Bering Straits, including those related to present-day Eskimo-Aleuts, Chukotko-Kamchatkans and Yukaghirs, accounting for the unique trans-Beringian genetic connections of this four-member sequence of populations. We propose that this represents the second major route by which APS ancestry persisted: that is, these are 'Route 2' populations (Fig. 2c and Supplementary Information, section 8).

However, Ancient Athabaskans are an exception in that they do not require this ancestry from Greenland_Saqqaq.SG, corroborated by their behaviour in f_4-statistics (Supplementary Figs. 97 and 98). This suggests that Athabaskans and Palaeo-Eskimos do not derive APS ancestry from the same source—in tension with previous findings by our group[7,8] and confirming suggestions of multiple Holocene migrations from Eurasia into the Americas[9,10,24]. Instead, we find suggestive, but weak, evidence for the involvement of a Route 1 population in the APS admixture into Ancient Athabaskans (Supplementary Information, section 9). Linguists have discovered a connection between Yeniseian languages of Central Siberia and the Na-Dene languages of North America[6], and our results may provide some genetic support for the Dene-Yeniseian hypothesis.

NEAHG cline

Further south, from 10–4 ka, all 150 newly reported and 81 previously published individuals from the North Eurasian forest-steppe and the southern edge of the forest zone fall into a genetic arc—the NEAHG cline—that connects pottery-using Eastern European foragers to their counterparts in the Transbaikal region, visible in ADMIXTURE (Fig. 1c, bottom), and in multiple PCAs (Fig. 1b and Extended Data Figs. 3–5). We grouped NEAHG individuals by site, time and genetic similarity in PCA and ADMIXTURE (Fig. 1a and Extended Data Figs. 3–5 and 9; resulting group labels in Extended Data Fig. 10 and Supplementary Data 1). The great majority can be modelled in qpAdm (restricted to

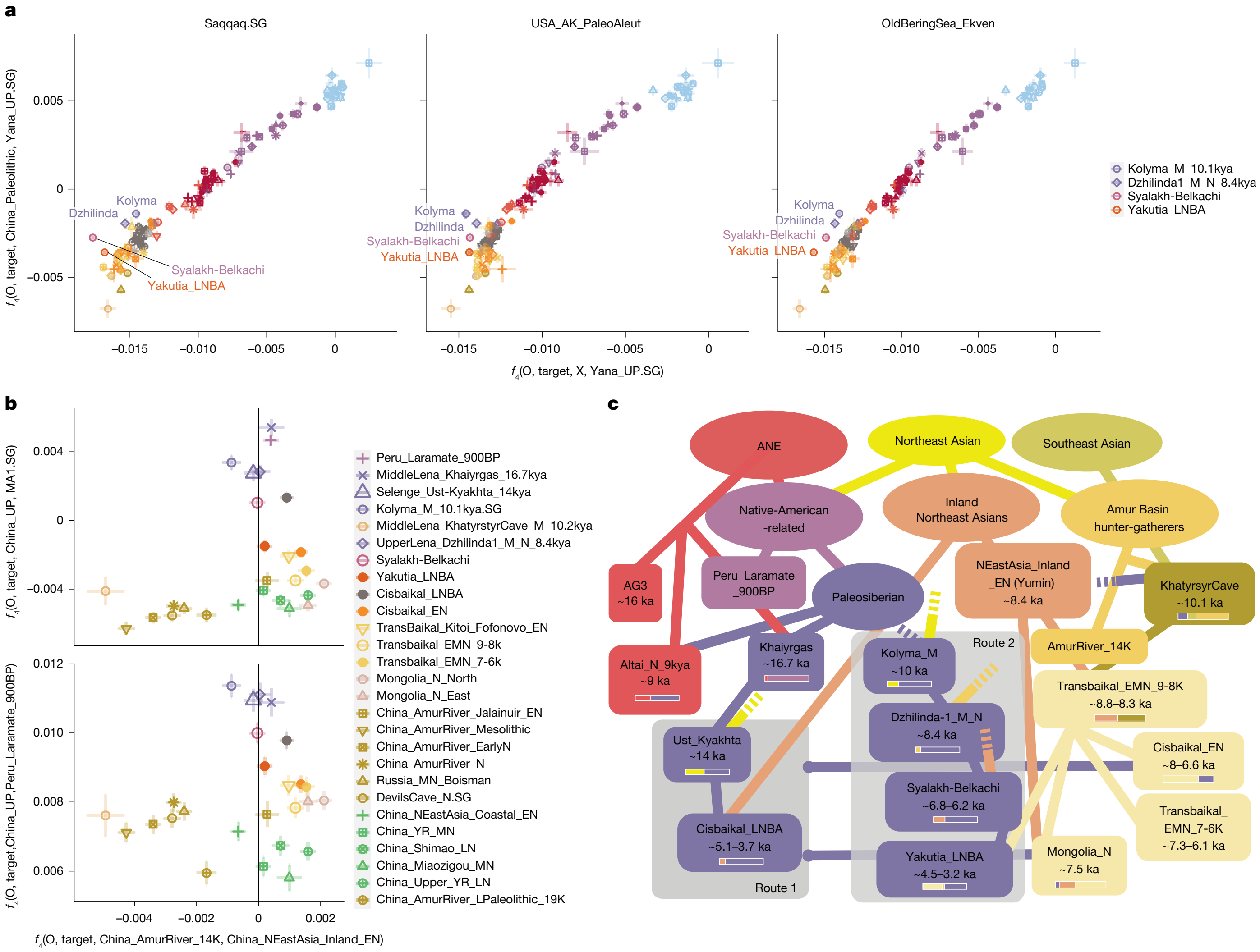


Fig. 2 | Middle Holocene populations and admixture events that formed them. **a**, Statistics of the form f_4(Ethiopia_4500BP, target, China_Paleolithic, Yana_UP) versus f_4(Ethiopia_4500BP, target, X, Yana_UP), where X are ancient Native Americans or populations from the Bering Straits. The position of the target population on the *y* axis is proportional to its ratio of ANE and East Asian ancestry. Kolyma_M_10.1kya, MiddleVitim_Dzhilinda1_M_N_8.4kya, Syalakh-Belkachi and Yakutia_LNBA are shifted left, indicating that they share more drift with ancient Bering Straits groups than other populations with similar ratios of ANE and East Asian ancestry. **b**, Statistics of the form f_4(Ethiopia_4500BP, X, China_NEastAsia_Inland_EN, China_AmurRiver_Mesolithic_14K) versus f_4(Ethiopia_4500BP, X, China_Paleolithic, MA1_HG) (top left) and f_4(Ethiopia_4500BP, X, China_Paleolithic, Peru_Laramate_900BP) (bottom left), where X are ancient populations in Northeast Asia and Siberia. These statistics detect differentiation between an Inland East Asian-related source (proxied by the Yumin hunter-gatherer China_NEastAsia_Inland_EN) and an Amur River-related source (represented by the China_AmurRiver_Mesolithic_14K). Populations from the Amur River region always have high affinity to China_AmurRiver_Mesolithic_14K, whereas those on the Mongolian Plateau and the Baikal area share more affinity with Yumin. The earliest strongly East Asian individual in Siberia, the Mesolithic MiddleLena_KhatystyrCave_M_10.2kya, is extremely Amur River-related; other Northeastern Siberian groups high in APS ancestry, such as MiddleVitim_Dzhilinda1_M_N_8.4kya, Kolyma_M_10.1kya and Syalakh-Belkachi, have both affinities; Cisbaikal_LNBA has extreme Inland Northeast Asian-relatedness. Affinity to China_NEastAsia_Inland_EN increases among agriculturalist populations along the Yellow River Valley. **c**, Schematic of population relationships in Northeast Asia and East Siberia, deduced from qpAdm in a ten-member transect from around 17 ka to around 4 ka. Major findings are: (1) that the MiddleLena_Khaiyrgas_16.7kya population is a near-unadmixed representative of an APS population with Native American affinities; (2) APS ancestry persisted through two routes; and (3) the East Asian ancestry of Siberians derives from an Amur Basin-related source and an Inland East Asian-related source.

1240K data[25]; for analytic details of all qpAdm analyses in this paper, refer to Extended Data Table 1) as mixtures of four ancestries (84 of 93 populations $P > 0.01$): Western hunter-gatherer ancestry (WHG, represented by samples from Serbia[26] after about 10 ka), European hunter-gatherer (EHG) ancestry (by I6413 from the Elshanka culture of the Middle Volga, the oldest pottery-using culture in Eastern Europe, around 8 ka), ANE (by the AG3 individual[27] from about 16 ka) and East Asian (by Amur Basin foragers[23] from about 19 ka; Fig. 1c, bottom, Supplementary Information, section 10 and Supplementary Data 2). In the West, hunter-gatherers from the Baltic to the Urals in such cultures as the Elshanka, Pit-Comb Ware/Lyalovo and Volosovo cultures have mostly EHG with low WHG, consistent with previous findings[28,29]. East of the Urals, in Neolithic populations of the Tobol and Middle Irtysh rivers, and in the circle of Eneolithic West Siberian cultures using Comb-Pit Ware pottery, EHG admixed with ANE and low levels of East Asian ancestry, similar to the Botai population of the Kazakh Steppes[12] (5.4–5.1 ka) and previously described West Siberian Hunter-gatherers[30] (6.6–8.1 ka). Further east, individuals from the Kuznetsk-Altai culture of the upper Ob' and the Altai foothills can be modelled as two-way admixtures of ANE and East Asian ancestry. This continues into individuals from Neolithic sites of the Upper Yenisei and Kansk River Basin, where ANE ancestry declines and East Asian ancestry increases. The gradient

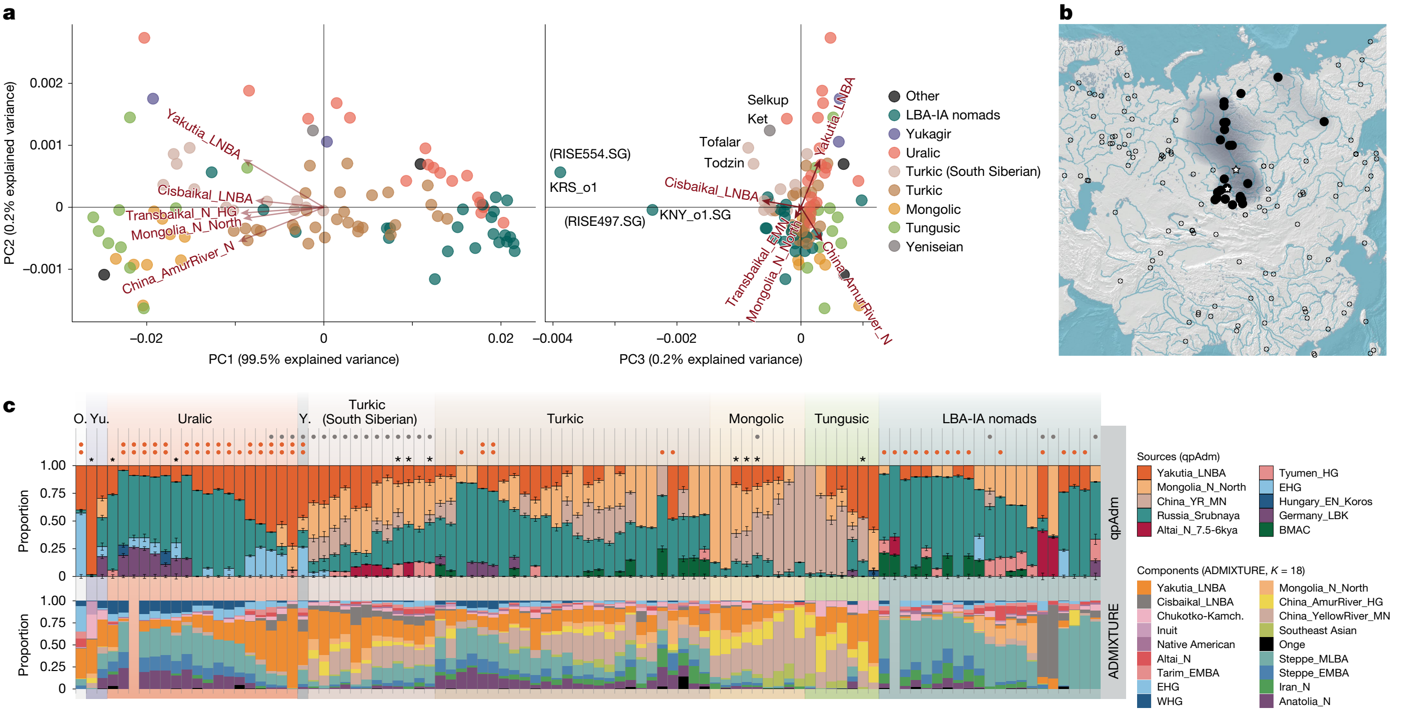


Fig. 3 | Contribution of Yakutia_LNBA and Cisbaikal_LNBA to AIEAs. A version with all population labels indicated is presented in Extended Data Fig. 10. **a**, PCA of f_4-statistics. A version with all population labels indicated is presented in Supplementary Information, section 13. PCA of statistics of the form f_4(Ethiopia_4500BP, AIEA, AG3, East Asian) measure the affinity between the East Asian ancestry of an AIEA population and a panel of tested East Asian populations: China_AmurRiver_N, Mongolia_N_North, Transbaikal_EMN, Cisbaikal_LNBA or Yakutia_LNBA. PC1 is correlated with proportion of any type of East Asian ancestry. At a given proportion of East Asian ancestry, ancient and present-day Uralic-speaking populations shift in PC2 in the direction suggesting disproportionate relatedness to Yakutia_LNBA. PC3 highlights similarity to Cisbaikal_LNBA (right), with most affinity in Yeniseians, South Siberian Turks, Samoyeds and two Upper Yenisei outliers (3.0–2.9 ka, RISE497.SG and RISE554.SG, which our archaeological research suggests are from the Lugavskaya culture). **b**, Cisbaikal_LNBA contribution to present-day populations. Populations with more than 4% Cisbaikal_LNBA ancestry are shown as large black dots. Probable Lugavskaya culture outliers of the Minusinsk Basin are shown as white stars. **c**, Ancestry modelling. Top, qpAdm results for AIEA populations. One orange dot above the bars indicates that all East Asian ancestry can be modelled as Yakutia_LNBA; two orange dots indicates that—additionally—all passing models include Yakutia_LNBA among the sources. We also performed qpAdm with Cisbaikal_LNBA among the references and sources (Supplementary Information, section 11); a grey dot indicates that all passing models include Cisbaikal_LNBA in the sources. Bottom, ADMIXTURE results. Almost all Uralic-speaking populations have East Asian ancestry nearly exclusively assigned to the Yakutia_LNBA component; Yeniseians, South Siberian Turks and Samoyeds are the only populations with appreciable levels of the Cisbaikal_LNBA-related component. The two probable Lugavskaya culture outliers of the Minusinsk Basin are the only individuals with almost all of their ancestry assigned to the Cisbaikal_LNBA component.

extends into the Kitoi culture of the Baikal region through Cisbaikal_EN, to terminate in Transbaikal_EMN.

We sought temporally proximal sources for the ANE ancestry of the NEAHG populations west of the Altai using qpAdm restricted to 1240K data (Extended Data Table 1). Two sources can account for all this ancestry (Fig. 1c, middle and Supplementary Information, section 10): a Tarim_EMBA-like population from Central Asia[15] (around 4 ka) and the population of the Kuznetsk-Altai Neolithic (proxied by Altai_N_9kya). Tarim_EMBA postdates NEAHG populations, but ADMIXTURE and PCA suggest gene flow between a source related to them and NEAHGs in West Siberia (Fig. 1b,c, bottom). West Siberian NEAHGs cannot be modelled without a Tarim_EMBA-related source (Fig. 1c and Supplementary Information, section 10), implying that hunter-gatherer populations related to Tarim_EMBA lived in Central Asia before the Bronze Age[15,25] and contributed to groups living in the north.

The NEAHG cline fragmented in the mid-Holocene following migrations from both West and East (Extended Data Fig. 6). From the West, these brought Steppe_EMBA ancestry with Yamnaya pastoralists, followed by Europe_LNBA and Steppe_LNBA with the expansion of the Fatyanovo, Sintashta and Andronovo cultures[29–31]. In the East, other migrations drove a wedge of Cisbaikal_LNBA ancestry into the Baikal region of the NEAHG cline around 5.4 ka. Subsequently, admixture between Steppe_MLBA and other East Asian ancestries gave rise to admixed groups in multiple genetic clines that connect Turkic-, Mongolic-, Tungusic- and Uralic-speaking populations[12] (Fig. 1c and Extended Data Figs. 3–5). To evaluate the legacy that the NEAHG cline and Central Siberian populations left in later populations across Eurasia, we analysed a set of Admixed Inner Eurasian (AIEA) populations—our term for ancient and present-day Uralic, Turkic, Mongolic, Tungusic and Yeniseian-speaking populations plus pastoralists of the Late Bronze Age and Iron Age such as Scythians, Sarmatians, and Xiongnu[32–35]. We find that NEAHG populations contributed little to these later groups, but two non-NEAHG populations—Cisbaikal_LNBA and Yakutia_LNBA—contributed in important ways.

Cisbaikal_LNBA tracks Yeniseian languages

The Cisbaikal_LNBA group (5.1–3.6 ka; Extended Data Fig. 7) is rich in APS ancestry, occupies a distinct position in PCAs (Extended Data Fig. 6c) and has a uniquely strong affinity to Inland Northeast Asians (Fig. 2b, Extended Data Fig. 8 and Supplementary Information, section 8). Whereas other APS-rich groups from Northeast Siberia (that is, all four Route 2 populations) are more closely related to Bering Straits groups, four lines of analysis show that Cisbaikal_LNBA shares more drift with present-day populations of the Yenisei Basin.

First, ADMIXTURE (Fig. 3 and Extended Data Fig. 11) shows that present-day Yenisei Basin groups such as Kets, Samoyeds and Siberian Turkic speakers are unique in harbouring a Cisbaikal_LNBA-related component (Fig. 3b,c, bottom and Extended Data Fig. 11b,c). Second, qpAdm models for these groups consistently fail when Cisbaikal_LNBA is used as a reference population; Cisbaikal_LNBA is a source in all passing models (bottom rows of Fig. 3c and Extended Data Fig. 11c and Supplementary Information, section 11). Third, in a PCA over f_4-statistics designed to detect differences between AIEA populations in affinities to different East Asian groups, Yeniseian, Samoyedic and South Siberian Turkic speakers are shifted systematically in the direction produced by increased shared drift with Cisbaikal_LNBA (PC3 of Fig. 3a and Extended Data Fig. 11a and Supplementary Information, section 11). Fourth, Y-chromosome sequences related to haplogroup Q-YP1691 found at high frequencies in Kets and at lower frequencies in Samoyedic and Siberian Turkic populations such as Selkups and Tuvinians[28,36–38] have been recovered only from Glazkovo males belonging to Cisbaikal_LNBA (Supplementary Information, section 11).

Ethnolinguistic data and historical records indicate that South Siberian Turks assimilated Yeniseian speakers, beginning with the arrival of the Yenisei Kyrgyz in the sixth century CE and lasting to early modern times. Other Siberian Turkic languages—Yakut and Dolgan—are spoken by populations whose ancestors migrated in the last millennium from the region where South Siberian Turks live today[39]. Further north, ethnographic records indicate that some Samoyedic-speaking groups sustained close relationships with Yeniseian speakers, with much intermarriage (Supplementary Information, section 12).

Unexpectedly, we also found that two published[32] Late Bronze Age (3.0–2.9 ka) East Asian outliers from the Minusinsk Basin along the Upper Yenisei (RISE497.SG, and RISE554.SG) were consistent with having near-complete Cisbaikal_LNBA ancestry (85–95%; Fig. 3b and Supplementary Information, section 11). These individuals had by far the strongest genetic affinity to Cisbaikal_LNBA among all modern or ancient AIEAs (Fig. 3a,c and Extended Data Fig. 11). They were labelled as being from the Karasuk culture in the original publication, but our archaeological investigations indicate instead an alternative assignment to the Lugavskaya culture (Supplementary Information, section 3). Thus, populations with very high Cisbaikal_LNBA were present along the Upper Yenisei, near where Cisbaikal_LNBA is maximized today, by the Late Bronze Age around 3.0 ka (Fig. 3b). Except for Ket, all six other now-extinct Yeniseian languages were spoken in the region where Cisbaikal_LNBA peaks today (Fig. 3b). The Ket themselves reached their current northward location in a recent expansion as late as the seventeenth century (Supplementary Information, section 12).

These findings match reconstructions—based on the distribution of Yeniseian hydronyms—of a Yeniseian homeland between the Cis-Baikal region and the Upper Yenisei (Supplementary Information, section 12). The Cisbaikal_LNBA population first appears genetically 5.4–3.8 ka in the Serovo, Isakovo and Glazkovo cultures (Supplementary Information, section 11). Along the Middle Angara (which drains out of Lake Baikal into the Yenisei), it appears alongside Glazkovo artefacts in samples buried according to Glazkovo traditions (Supplementary Information, section 3). Cisbaikal_LNBA ancestry may thus trace the movements of Yeniseian speakers even further into prehistory.

Yakutia_LNBA tracks Uralic languages

Yakutia_LNBA (Fig. 1 and Extended Data Fig. 7) individuals belong chiefly to the Ymyyakhtakh culture of the Lena River Valley, and are among the Route 2 populations that share distinctive genetic drift with Bering Straits groups (Fig. 2a and Supplementary Information, section 8). They can be modelled as an approximately 50%:50% mixture between the preceding Syalakh-Belkachi population of the Lena Valley and the Transbaikal Kitoi population (Transbaikal_EMN; Supplementary Information, section 8). The connection with Transbaikal_EMN is also supported by shared subclades of Y-chromosome haplogroup N (Supplementary Information, section 13) and is consistent with archaeological reconstructions of Ymyyakhtakh origins (Supplementary Information, section 3). However, an individual recovered from the Krasnoyarsk-Kansk forest-steppe far to the southwest of the Lena River Valley at around 4.2 ka (Kra001.SG from the Nefteprovod-2 site[16]), in a location otherwise occupied by populations from the NEAHG cline, was also genetically Yakutia_LNBA, suggesting that Yakutia_LNBA individuals may have dispersed from Northeast Siberia to the forest-steppes North of the Altai-Sayan shortly before 4.0 ka, which coincides with the spread of Ymyyakhtakh pottery to this region at that time.

Yakutia_LNBA is unambiguously associated with ancient and present-day Uralic-speaking populations. First, in ADMIXTURE at $K = 18$, a component maximized in Yakutia_LNBA appears that peaks today in Nganasans and accounts for almost all East Asian ancestry in Uralic speakers; non-Uralic AIEAs have no Yakutia_LNBA, or other East Asian components in addition to Yakutia_LNBA (Fig. 3c and Extended Data Fig. 11c). Second, in a PCA of f_4-statistics, Uralic speakers are shifted in the direction indicating increased affinity towards Yakutia_LNBA relative to other East Asian ancestries (PC2 in Fig. 3a and Extended Data Fig. 11a and Supplementary Information, section 13). Third, a different set of f_4-statistics indicates that, at any level of East Asian admixture, the AIEA population with the highest affinity to Yakutia_LNBA over other East Asian ancestries is always a Uralic-speaking population (Extended Data Fig. 12 and Supplementary Information, section 13). Fourth, qpAdm models for Uralic speakers always require Yakutia_LNBA as a source, usually accounting for all their East Asian ancestry (top rows of Fig. 3c and Extended Data Fig. 11c and Supplementary Information, section 13), in contrast to other ethnolinguistic groupings of AIEAs who always have other East Asian sources. Finally, Yakutia_LNBA males carry Y-chromosome subclades of haplogroup N that are present at high frequency in present-day speakers of Uralic languages[3] (Supplementary Information, section 13).

Yakutia_LNBA in the Seima-Turbino phenomenon

Populations from Eastern Europe to West Siberia as late as the MLBA (Fatyanovo, Sintashta and Andronovo cultures) do not show any Yakutia_LNBA ancestry[3–5], but present-day Uralic speakers from the same regions do, suggesting a westward spread of Yakutia_LNBA ancestry partially replacing Steppe_MLBA and Europe_LNBA ancestry at about 4 ka at the earliest[29–31]. This transition was potentially accompanied by the dispersal of Uralic-associated Y-haplogroup N, which is absent in Eastern Europe and West Siberia prior to the arrival of Yakutia_LNBA ancestry. Here we show that the earliest stages of this westward dispersal of Yakutia_LNBA ancestry occurred within the Seima-Turbino phenomenon.

The Seima-Turbino phenomenon refers to the sudden appearance of a similar suite of bronze artefacts made with advanced casting techniques that spread across a vast region of Northern Eurasia, from China to the Baltic[1,40], around 4.0 ka. Archaeologists agree that it was responsible for the introduction of metallurgy into East Asia and the dissemination of advanced casting methods for tin bronze into Europe[41,42]. Seima-Turbino items are noted for their sophistication and refinement (Extended Data Fig. 13). Most are weapons, but some are objects of ritual significance. Most Seima-Turbino objects are isolated finds scattered across sites of diverse cultures, but many occur in ceremonial necropoli found across Western Siberia and Eastern Europe, which are large complexes of burials and sometimes empty ritual graves (cenotaphs) with rich collections of Seima-Turbino artefacts and casting moulds. This unusual distribution has fuelled speculation about the social nature of the Seima-Turbino phenomenon, as well as the identity of their bearers[1,40,43,44]. So far, the only material evidence found for the manufacture of Seima-Turbino bronze artefacts were recovered from residential sites of metal-using fisher-foragers of the Ob'-Irtysh basin and the region between the Upper Ob' and Upper Yenisei—an extraordinary cultural association that has generated much comment in the archaeological literature[1,43,44].

We generated genome-wide data from 16 individuals from 4 sites, dated to a tight interval around 4.0 ka (Supplementary Information, section 2 and Supplementary Data 1). Two—Rostovka on the banks of the middle Irtysh in the Ob'-Irtysh Basin with 9 individuals, and Satyga-16, east of the Mid-Ural Mountains with 2 individuals—are Seima-Turbino necropoli. We add to these samples from two Seima-Turbino period sites that have less direct evidence of involvement with the Seima-Turbino phenomenon, but that our genetic analyses suggest may be connected with it: one from Chernoozerye-1, located close to Rostovka, and four males from a previously undescribed site, Tatarka Hill along the Upper Yenisei, on the Krasnoyarsk-Kansk forest-steppe North of the Altai. In our genetic modelling using qpAdm, the four individuals from Tatarka Hill are consistent with being entirely Yakutia_LNBA. By contrast, the individuals from Rostovka, Satyga-16, and Chernoozerye-1 harbour variable proportions of three primary and two minor sources of ancestry (Figs. 1b and 4, Extended Data Figs. 4 and 5 and Supplementary Information, section 15). Based on qpAdm, these ancestries are: (1) Yakutia_LNBA; (2) ANE-rich ancestry from the NEAHG cline; and (3) Steppe_MLBA, occuring in unadmixed individual representatives or intermingled within admixed individuals (two-way: NEAHG ancestry + Steppe_MLBA, or three-way: NEAHG ancestry + Steppe_MLBA + Yakutia_LNBA; Fig. 4b, top). Both individuals from Satyga-16 from further west are admixed (carrying all three ancestry types), contrasting with Rostovka (4 out of 9 single-ancestry individuals: 2 NEAHG, 1 Yakutia_LNBA and 1 Steppe_MLBA; Fig. 4b and Supplementary Information, section 15).

Proximal qpAdm provides insight into the immediate ancestors of Seima-Turbino people. The Yakutia_LNBA ancestry in Rostovka, Satyga-16 and Chernoozerye-1 is related to the people of Tatarka Hill, with no additional mixture from Yakutia_LNBA in Central Siberia (Fig. 4b). This link is reinforced by the presence, at Rostovka, of haplogroup N-L1026 (in the Yakutia_LNBA individual, I32545), also carried by all four males from Tatarka Hill (Supplementary Information, section 13). The subclade in Rostovka (N-L1026>Z1936) is widespread in present-day Uralic populations from West Siberia to the Baltic Sea, attaining maximal frequencies today (up to around 40%) near the Baltic in Finns, Veps and Karelians[45]. The NEAHG ancestry of Seima-Turbino individuals comes in large part from preceding, local Neolithic and Eneolithic populations of West Siberia (Fig. 4e), consistent with an origin in the metallurgical foragers of the nearby Odinovo and Krotovo cultures, who engaged in the systematic casting of Seima-Turbino artefacts[44]. However, some individuals require additional NEAHG ancestry from further afield, from EHG-related or Altai_N-related sources.

The three primary ancestry sources are accompanied by two minor ancestries: non-Yakutia_LNBA East Asian ancestry, and WHG ancestry from as far west as the Baltic region (Fig. 4e). The Seima-Turbino period individual from Chernoozerye-1 (I6787) requires a large fraction of WHG in ADMIXTURE analysis and in all fitting qpAdm models (Fig. 4e and Supplementary Information, section 15). This is a remarkable case of a person whose recent ancestry traces to at least three hunter-gatherer populations from widely separated regions of Eurasia (the Baltic, West Siberia, and the Altai-Sayan). Two individuals from Rostovka, I32816 and I33369, have ancestry from the east—in the former case from a Cisbaikal_LNBA-related source, possibly foragers of the contemporaneous Glazkovo culture.

Our results suggest that Seima-Turbino artefacts were manufactured, exchanged and dispersed in a sociocultural context that integrated people from multiple populations across a continent-spanning network into coherent social groups interred together at single necropolises. Our samples capture a snapshot of this process, indicating a pattern of human mobility that is a genetic correlate to archaeological evidence for similarity in artefacts over vast geographic distances, unusual in cultural groups of the period[1,46].

Discussion

In summary, our study reveals five major ancestry changes in northern Eurasia:

(1) A Pleistocene population related to Native Americans that we call APS ancestry, mixed with two East Asian ancestry sources—Inland Northeast Asian-related and Amur Basin-related—to contribute to later populations throughout Siberia. APS ancestry persisted via two routes, the first of which (Route 1) is responsible for APS ancestry in Central Siberia, and the second of which (Route 2) is responsible for APS ancestry in populations on either side of the Bering Straits.

(2) Early pottery users in a latitudinal belt across Northern Eurasia in the early-to-mid Holocene (10–15 ka) including the forest-steppe and the forest belt immediately adjacent to it, constitute a continent-spanning east-west genetic cline comprising EHG, ANE and East Asian ancestries. This NEAHG cline began to dissolve owing to several major population expansions in the Mid-Holocene (beginning around 5 ka), which either completely displaced or heavily admixed with NEAHG cline groups.

(3) A genetic turnover around 5.4 ka saw the emergence of a population, Cisbaikal_LNBA to the west of Lake Baikal—a Route 1 population. This ancestry spread from the Cis-Baikal region to the Yenisei region by the end of the Late Bronze Age around 3.1 ka. Today, the presence of this ancestry is strongly associated with Yeniseian-speaking populations and those likely to have mixed with them historically. We also discover that the Palaeo-Eskimo Greenland_Saqqaq.SG population contributed high levels of APS ancestry to all later ancient populations on either side of the Bering Straits, explaining their high affinity to Route 2-related populations, but that ancient Athabaskans from Alaska (around 1 ka) are an exception, consistent with them deriving APS ancestry from a separate source, a result that is also buttressed by suggestive evidence that the APS source in Athabaskans have affinity with a Route 1 population, providing the first genetic data in support of the Dene-Yeniseian hypothesis.

(4) A genetic turnover by 4.5 ka saw the emergence of a population in Northeast Siberia, Yakutia_LNBA—one of the populations in Route 2. Today, this ancestry tends to be the only East Asian ancestry present among Uralic-speaking populations, a striking feature not shared by any other ethnolinguistic grouping. This ancestry appears in the Krasnoyarsk region along the Upper Yenisei, far to the Southwest of Yakutia, by 4.2 ka alongside subclades of Y-chromosome haplogroup N found at high frequency among present-day Uralic-speaking males as far as the Baltic Sea. This ancestry was likely dispersed by population movements that spread Uralic languages.

(5) Individuals who lived at the time of the Seima-Turbino phenomenon—an archaeological term for the sudden appearance of a distinct suite of bronze artefacts across an enormous expanse of Northern Eurasia around 4.0 ka—were genetically heterogeneous, but many harbour Yakutia_LNBA ancestry, which occurred in Seima-Turbino sites close to the Urals, far to the west of the original distribution of this ancestry. This geographic distribution supports theories that the Seima-Turbino phenomenon was implicated in the dispersal of early Uralic-speaking communities. The rest of the genetic ancestry of people buried at Seima-Turbino sites was extremely diverse, with ultimate origins from the Baltic to the Baikal. This pattern of genetic ancestry points to a social process that enabled at least several generations of contact and intermarriage between individuals that were genetically and culturally very distant from one another.

Linguistic transmission in large-scale societies need not involve the movement of people, but the same process in smaller-scale societies is likely to require at least some degree of human mobility visible as genetic admixture. One major analytic finding is our identification of Cisbaikal_LNBA as a genetic tracer dye for the spread of early Yeniseian language speakers. We further show that ancient Athabaskans from Alaska around 1.1 ka are unique among Arctic North Americans in lacking ancestry from Palaeo-Eskimo populations but possess tentative

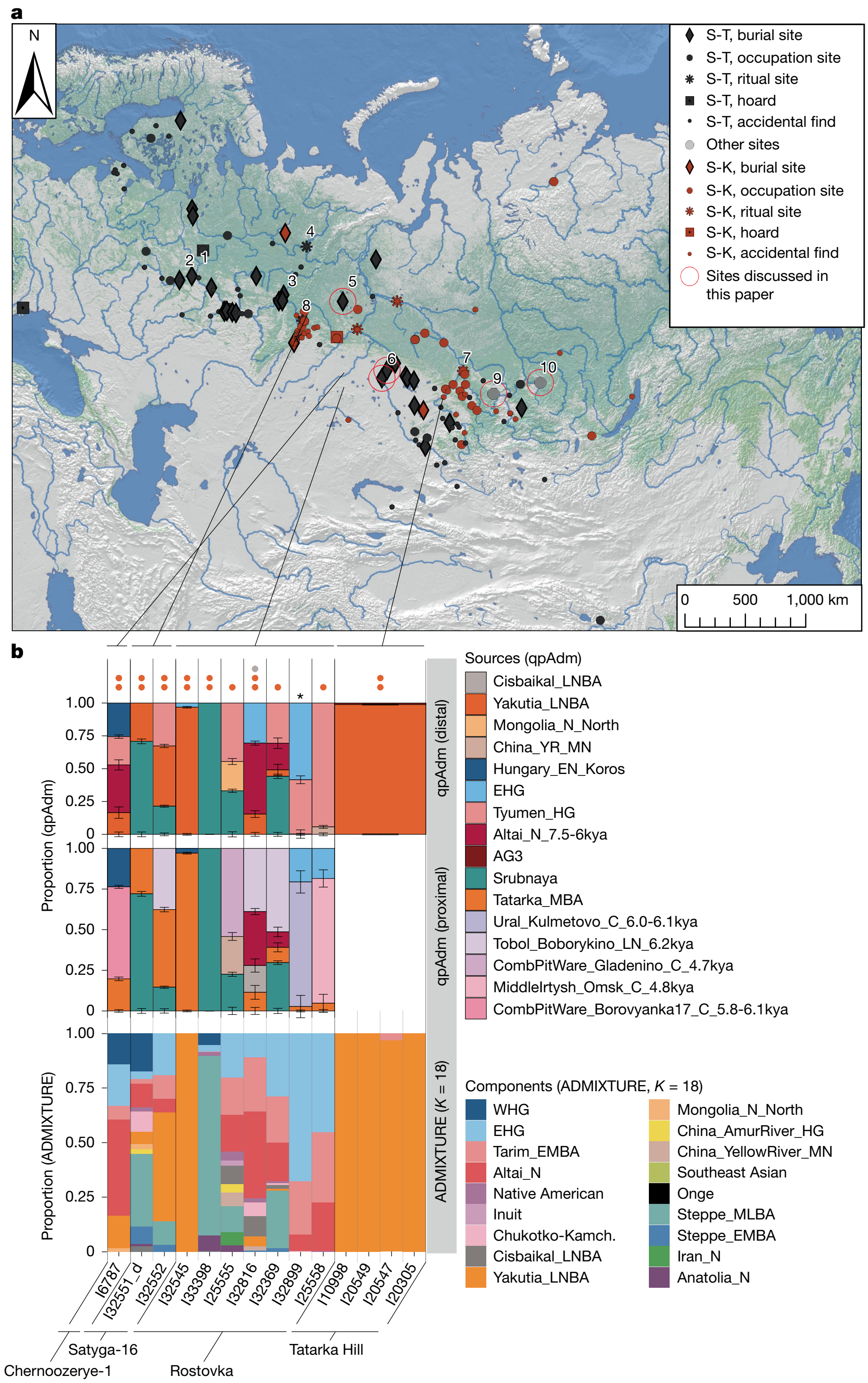


Fig. 4 | Genetics of the Seima-Turbino phenomenon. a, Map of Seima-Turbino sites and finds. Chernoozerye-1, Rostovka, Satyga-16 and Tatarka from which we have genetic data are marked in red circles. other important sites are numbered: (1) Seima; (2) Reshnoe; (3) Turbino; (4) Kaninskaya Cave; (5) Satyga-16; (6) Rostovka; (7) Samus-4; (8) Shaitanskoe Ozero; and (9) Tatarka. Adapted from ref. 61. **b**, ADMIXTURE and qpAdm proportions for 16 Seima-Turbino period individuals: 9 Rostovka, 2 Satyga-16, 1 Chernoozerye-1 and 4 Tatarka. Top, selection of qpAdm models using a distal set of sources, with the simplest passing ($P > 0.01$) qpAdm model with the highest P value always displayed (same modelling as in Fig. 3b). No model passes for I32899 at $P > 0.01$, as indicated by an asterisk above the bar. Middle, qpAdm results for a more proximal set of sources, with individuals from Tatarka used as the source for Yakutia_LNBA and populations from Late Neolithic or Eneolithic Western Siberia (between the Urals and the Altai) as the source for NEAHG; Yakutia_LNBA has been added to the references. Ancestry from the population at Tatarka suffices to account for all the Yakutia_LNBA-related ancestry of the Seima-Turbino individuals even when Yakutia_LNBA is among the references. Samples that can be modelled with their East Asian ancestry derived completely from a Yakutia_LNBA-related source are highlighted with a single orange dot above their qpAdm bar charts, and those that require such a source in every passing model are highlighted with two orange dots. One individual (I32816) requires Cisbaikal_LNBA ancestry among the sources (grey dot); the model displayed for this individual is the simplest passing model that contains both Cisbaikal_LNBA and Yakutia_LNBA ancestry among the sources. In both sets of qpAdm, the individual (I6787) from Chernoozersky-1 requires a contribution from a source from far West of the Urals (WHG ancestry). Bottom, ADMIXTURE proportions at $K = 18$.

signals of ancestry from a Route 1 APS population that also contributed distinctively to Yeniseian-speaking groups. These results help connect the movements of early Yeniseian-speaking groups to the Cis-Baikal region and may also provide tentative genetic support for the linguistic connection between Yeniseian languages of Siberia and Athabaskan languages of North America: the ‘Dene-Yeniseian hypotheses’[5]. In our second major analytic finding, we show that Yakutia_LNBA may serve as an excellent tracer dye for the spread of early Uralic-speaking communities, and that the earliest dispersals of this ancestry west was mediated by people associated with the Seima-Turbino phenomenon.

Archaeologists debate the social processes that drove the rapid spread of Seima-Turbino artefacts across such a wide range of cultures[1,43,44,46]. We find that people buried at Seima-Turbino necropolises were highly genetically variable, contradicting hypotheses of a homogeneous Seima-Turbino people[40,43]. Our results suggest either the one-time amalgamation of individuals from genetically and culturally distinct social groups into a mobile population (an event which may have taken place at a different location and prior to the Seima-Turbino sites themselves), or—based on the multi-way admixtures in Seima-Turbino necropolis individuals—the active, continuous interaction of people from multiple groups in activities that produced the sites over many generations. These findings are consistent with the heterogeneity of other cultural artefacts at Seima-Turbino necropolises, such as pottery (similar to that produced by West Siberian foragers[1,43,46] and people of the Krasnoyarsk-Kansk forest-steppe around Tatarka Hill[47]); artefacts of flint, bone or jade (similar to cultures of far Northeast Siberia and the Baikal); and metal items from non-Seima-Turbino traditions (from the Sintashta and Abashevo cultures)[1,43,46]. The three sources of material culture parallel the three major genetic ancestries at Seima-Turbino sites, also detected in an archaeogenetic study on Rostovka reported simultaneously[48]. Finally, the presence of ancestry from multiple hunter-gatherer populations across vast distances (from the Cis-Baikal to as far West as the Baltic) in Seima-Turbino sites highlights the transformative social impacts of metal exchange networks in the Bronze Age[49,50], and the accumulating but oft-neglected evidence for sociopolitical and economic dynamism in foraging populations[51–53].

Our finding that Yakutia_LNBA ancestry first dispersed westwards, almost to Europe, with the Seima-Turbino phenomenon has archaeological and linguistic significance. The Kra001 individual at Nefteprovod-2 around 4.2 ka, close to and just before the Tatarka Hill individuals, shows that Yakutia_LNBA ancestry penetrated onto the Krasnoyarsk-Kansk forest-steppe by 4.2 ka and persisted there before contributing to Seima-Turbino necropolises even further west. Nefteprovod-2 and Tatarka Hill share similar burial rites—suggesting that the genetic population bringing the Yakutia_LNBA ancestry to the Krasnoyarsk-Kansk forest-steppes that impacted Seima-Turbino necropolises, was also culturally cohesive (that we term the Anzhevsky complex; Supplementary Information, section 3). Another material counterpart to the genetic link between Seima-Turbino necropolises and groups of ultimately Northeast Siberian origin can be found in suits of armour made of bone plates, which have been found from the Glazkovo and especially the Ymyyakhtakh cultures. One set was buried with a Yakutia_LNBA male (N4a1.SG from the Kyordyughen site) and others are from the Krasnoyarsk-Kansk forest-steppe around Nefteprovod-2 and Tatarka Hill (Supplementary Information, section 3). Three sets can be found in Rostovka, one associated with a male (I32816 from Grave 33; Supplementary Information, section 3) that bore both Yakutia_LNBA and Cisbaikal_LNBA ancestries.

Linguists have documented hundreds of Indo-Iranian loanwords that present-day Uralic languages have inherited from the Proto-Uralic speech community or from early Uralic communities just after its breakup[54,55]. The Indo-Iranian expansion has been linked to the spread of Steppe_MLBA ancestry from the Sintashta population of the Trans-Ural region into other parts of Central and West Asia (where it persisted into historically attested Iranic speakers[33–35,56–58]), and further into South Asia[30]. Our findings from Rostovka and Satyga-16, showing contact and admixture between a Steppe_MLBA population (which, from archaeological considerations, is plausibly that of the Abashevo culture[1,41,43,59]) and Yakutia_LNBA, provides an attractive context in which this linguistic exchange could have first begun, and offers another line of evidence for Uralic-speaking groups being present at Seima-Turbino sites, in line with prior suggestions[55,59].

Uralic languages, distributed from Western Siberia to Central Europe, are geographically separated from languages of the Eastern Steppes and far Northeast Siberia, but linguists have discovered traces of ancient connections with Yukagiric and Eskimo-Aleut languages on the one hand, and high levels of typological similarity with languages in the ‘Altaic’ language area (Mongolic, Tungusic and Turkic) on the other (Supplementary Information, section 14). To resolve this conundrum, some linguists have suggested a recent eastern origin of the population giving rise to later expansions of Uralic speakers (for example, a “pre-proto-Uralic spoken further east… probably somewhere… near both Mongolia and the watershed area between the Yenisei and the Lena, possibly as recently as 3000 BC”[60])—a scenario compatible with our results. Future ancient DNA sampling from this region would allow for a more precise determination of the archaeological identity of the Proto-Uralic-speaking community, and illuminate the relationship between it and the wider social world of the West Siberian Bronze Age.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-09189-3.

1. Janhunen, J. Proto-Uralic—what, where, and when? *Quasquicentennial Finno Ugrian Soc.* **258**, 57–78 (2009).
2. Tambets, K. et al. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biol.* **19**, 139 (2018).
3. Lamnidis, T. C. et al. Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nat. Commun.* **9**, 5018 (2018).
4. Saag, L. et al. The arrival of Siberian ancestry connecting the Eastern Baltic to Uralic speakers further east. *Curr. Biol.* **29**, 1701–1711.e16 (2019).
5. Vajda, E. Dene-Yeniseian. *Diachronica* **35**, 277–295 (2018).
6. Reich, D. et al. Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).
7. Flegontov, P. et al. Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America. *Nature* **570**, 236–240 (2019).
8. Sikora, M. et al. The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182–188 (2019).
9. Nielsen, S. V. et al. Bayesian inference of admixture graphs on Native American and Arctic populations. *PLoS Genet.* **19**, e1010410 (2023).
10. Flegontov, P. et al. Genomic study of the Ket: a Paleo-Eskimo-related ethnic group with significant ancient North Eurasian ancestry. *Sci. Rep.* **6**, 20768 (2016).
11. Jeong, C. et al. The genetic history of admixture across inner Eurasia. *Nat Ecol. Evol.* **3**, 966–976 (2019).
12. Kidd, K. K. et al. North Asian population relationships in a global context. *Sci. Rep.* **12**, 7214 (2022).
13. Svyatko, S. V. et al. Freshwater reservoir effects in archaeological contexts of Siberia and the Eurasian Steppe. *Radiocarbon* **64**, 377–388 (2022).
14. Zhang, F. et al. The genomic origins of the Bronze Age Tarim Basin mummies. *Nature* **599**, 256–261 (2021).
15. Kılınç, G. M. et al. Human population dynamics and Yersinia pestis in ancient northeast Asia. *Sci. Adv.* **7**, eabc4587 (2021).
16. Yu, H. et al. Paleolithic to Bronze Age Siberians reveal connections with first Americans and across Eurasia. *Cell* **181**, 1232–1245.e20 (2020).
17. Harney, É., Patterson, N., Reich, D. & Wakeley, J. Assessing the performance of qpAdm: a statistical tool for studying population admixture. *Genetics* **217**, iyaa045 (2021).
18. Flegontova, O. et al. Performance of *qpAdm*-based screens for genetic admixture on graph-shaped histories and stepping-stone landscapes. *Genetics* **230**, iyaf047 (2025).
19. Davidson, R. et al. Allelic bias when performing in-solution enrichment of ancient human DNA. *Mol. Ecol. Resour.* **23**, 1823–1840 (2023).
20. Grebenyuk, P. S., Fedorchenko, A. Y., Dyakonov, V. M., Lebedintsev, A. I. & Malyarchuk, B. A. in *Humans in the Siberian Landscapes: Ethnocultural Dynamics and Interaction with Nature and Space* (eds Bocharnikov, V. N. & Steblyanskaya, A. N.) 89–133 (Springer, 2022).
21. Yang, M. A. et al. Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* **369**, 282–288 (2020).

22. Mao, X. et al. The deep population history of northern East Asia from the Late Pleistocene to the Holocene. *Cell* **184**, 3256–3266.e13 (2021).
23. Moreno-Mayar, J. V. et al. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**, 203–207 (2018).
24. Mathieson, I. et al. The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
25. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
26. Raghavan, M. et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
27. de Barros Damgaard, P. et al. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).
28. Saag, L. et al. Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci. Adv.* **7**, eabd6535 (2021).
29. Narasimhan, V. M. et al. The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
30. Posth, C. et al. Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* **615**, 117–126 (2023).
31. Allentoft, M. E. et al. Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
32. de Barros Damgaard, P. et al. 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369–374 (2018).
33. Krzewińska, M. et al. Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Sci. Adv.* **4**, eaat4457 (2018).
34. Järve, M. et al. Shifts in the genetic landscape of the Western Eurasian Steppe associated with the beginning and end of the Scythian dominance. *Curr. Biol.* **29**, 2430–2441.e10 (2019).
35. Wei, L.-H. et al. Paternal origin of Paleo-Indians in Siberia: insights from Y-chromosome sequences. *Eur. J. Hum. Genet.* **26**, 1687–1696 (2018).
36. Karmin, M. et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res.* **25**, 459–466 (2015).
37. YFull. *ISOGG Wiki* https://isogg.org/wiki/YFull (2024).
38. Pakendorf, B. et al. Investigating the effects of prehistoric migrations in Siberia: genetic variation and the origins of Yakuts. *Hum. Genet.* **120**, 334–353 (2006).
39. Chernykh, E. N. & Kuz'minykh, S. V. *Drevnyaya metallurgiya Severnoy Evrazii (Seiminsko-Turbinskiy fenomen)* (Nauka, 1989).
40. Marchenko, Z. V., Svyatko, S. V., Molodin, V. I., Grishin, A. E. & Rykun, M. P. Radiocarbon chronology of complexes with Seima-Turbino type objects (Bronze Age) in Southwestern Siberia. *Radiocarbon* **59**, 1381–1397 (2017).
41. Chernykh, E. N. Formation of the Eurasian 'Steppe Belt' of stockbreeding cultures: viewed through the prism of archaeometallurgy and radiocarbon dating. *Archaeol. Ethnol. Anthropol. Eurasia* **35**, 36–53 (2008).
42. Meicun, L. & Liu, X. The origins of metallurgy in China. *Antiquity* **91**, e6 (2017).
43. Chernykh, E. N. in Nomadic Cultures in the Mega-Structure of the Eurasian World (eds Savinetskaya, I & Hommel, P. N.) 234–249 (Academic Studies, 2017).
44. Molodin, V. I., Durakov, I. A., Mylnikova, L. N. & Nesterova, M. S. The adaptation of the Seima-Turbino tradition to the Bronze Age cultures in the south of the West Siberian plain. *Archaeol. Ethnol. Anthropol. Eurasia* **46**, 49–58 (2018).
45. Ilumäe, A.-M. et al. Human Y chromosome haplogroup N: a non-trivial time-resolved phylogeography that cuts across language families. *Am. J. Hum. Genet.* **99**, 163–173 (2016).
46. Kuzminykh, S. V. Seima-Turbino transcultural phenomenon: migration or diffusion of technology. In *Mobility and Migration: Concepts, Methods, Results: Programme and Abstracts of the V International Scientific Symposium* (eds Molodin, V. I. & Hansen, S.) 52–56 (2019).
47. Makarov, N. P. Khronologiya i periodizatsiya epokhi Neolita i Bronzy Krasnoyarskoy lesostepi [The chronology and periodization of the Neolithic and Bronze Krasnoyarsk forest]. *Izv. Lab. Drevn. Tekhnol.* **1**, 149–171 (2005).
48. Childebayeva, A. et al. Bronze age Northern Eurasian genetics in the context of development of metallurgy and Siberian ancestry. *Commun. Biol.* **7**, 723 (2024).
49. Kristiansen, K. *The Rise of Bronze Age Peripheries and the Expansion of International Trade 1950–1100 BC. Trade and Civilisation* Cambridge (eds Kristiansen, K. et al.) 87–112 (Cambridge Univ. Press, 2018).
50. Powell, W. et al. Tin from Uluburun shipwreck shows small-scale commodity exchange fueled continental tin supply across Late Bronze Age Eurasia. *Sci. Adv.* **8**, eabq3766 (2022).
51. Singh, M. & Glowacki, L. Human social organization during the Late Pleistocene: Beyond the nomadic-egalitarian model. *Evol. Hum. Behav.* **43**, 418–431 (2022).
52. Tarasov, A. & Nordqvist, K. Made for exchange: the Russian Karelian lithic industry and hunter-fisher-gatherer exchange networks in prehistoric north-eastern Europe. *Antiquity* **96**, 34–50 (2022).
53. Piezonka, H. et al. The world's oldest-known promontory fort: Amnya and the acceleration of hunter-gatherer diversity in Siberia 8000 years ago. *Antiquity* **97**, 1381–1401 (2023).
54. Holopainen, S. *Indo-Iranian Borrowings in Uralic: Critical Overview of Sound Substitutions and Distribution Criterion*. Doctoral thesis, Univ. of Helsinki (2019).
55. Grünthal, R. et al. Drastic demographic events triggered the Uralic spread. *Diachronica* **39**, 490–524 (2022).
56. Gnecchi-Ruscone, G. A. et al. Ancient genomic time transect from the Central Asian Steppe unravels the history of the Scythians. *Sci. Adv.* **7**, eabe4414 (2021).
57. Kumar, V. et al. Genetic continuity of Bronze Age ancestry with increased Steppe-related ancestry in Late Iron Age Uzbekistan. *Mol. Biol. Evol.* **38**, 4908–4917 (2021).
58. Guarino-Vignon, P., Marchi, N., Bendezu-Sarmiento, J., Heyer, E. & Bon, C. Genetic continuity of Indo-Iranian speakers since the Iron Age in southern Central Asia. *Sci. Rep.* **12**, 733 (2022).
59. Kovtun, I. V. *Predystoriya Indoariyskoy Mifologii* (Aziya-Print, 2013).
60. Häkkinen, J. in *Iter Polyphonicum Multilinguae* (eds Hyytiäinen, T. et al.) 91–101 (2012).
61. Buchhorn, M. et al. Copernicus Global Land Service: land cover 100m: collection 3: epoch 2019: Globe (V3.0.1) [Data set]. *Zenodo* https://doi.org/10.5281/zenodo.3939050 (2020).

Tian Chen Zeng[1,50]✉, Leonid A. Vyazov[1,2,50]✉, Alexander Kim[3,4,50]✉, Pavel Flegontov[1,2,5], Kendra Sirak[1,3], Robert Maier[1,3], Iosif Lazaridis[1,3], Ali Akbari[1,3], Michael Frachetti[6,7], Alexey A. Tishkin[8], Natalia E. Ryabogina[9], Sergey A. Agapov[10], Danila S. Agapov[10], Anatoliy N. Alekseev[11], Gennady G. Boeskorov[12], Anatoly P. Derevianko[13], Viktor M. Dyakonov[13], Dmitry N. Enshin[14], Alexey V. Fribus[15], Yaroslav V. Frolov[8], Sergey P. Grushin[8], Alexander A. Khokhlov[16], Kirill Yu. Kiryushin[17], Yurii F. Kiryushin[8], Egor P. Kitov[18], Pavel Kosintsev[19,20], Igor V. Kovtun[21], Nikolai P. Makarov[22], Viktor V. Morozov[23], Egor N. Nikolaev[11], Marina P. Rykun[24], Tatyana M. Savenkova[25], Marina V. Shchelchkova[26], Vladimir Shirokov[27], Svetlana N. Skochina[14], Olga S. Sherstobitova[28], Sergey M. Slepchenko[14], Konstantin N. Solodovnikov[14], Elena N. Solovyova[13,29], Aleksandr D. Stepanov[13], Aleksei A. Timoshchenko[13], Aleksandr S. Vdovin[22], Anton V. Vybornov[13], Elena V. Balanovska[30], Stanislav Dryomov[31], Garrett Hellenthal[32], Kenneth Kidd[33], Johannes Krause[34], Elena Starikovskaya[31], Rem Sukernik[31], Tatiana Tatarinova[35], Mark G. Thomas[32], Maxat Zhabagin[36,37], Kim Callan[3,38], Olivia Cheronet[39,40], Daniel Fernandes[39,41], Denise Keating[42], Francesca Candilio[43], Lora Iliev[3,38], Aisling Kearns[1,3], Kadir Toykan Özdoğan[40,44], Matthew Mah[3,38,45], Adam Micco[3,38], Megan Michel[1,3,38], Iñigo Olalde[1,46,47], Fatma Zalzala[3,38], Swapan Mallick[1,3,38,45], Nadin Rohland[1,3,45], Ron Pinhasi[39,40]✉, Vagheesh M. Narasimhan[48,49]✉ & David Reich[1,3,38,45]✉

[1]Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA. [2]Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic. [3]Department of Genetics, Harvard Medical School, Boston, MA, USA. [4]Department of Anthropology, Harvard University, Cambridge, MA, USA. [5]Institute of Parasitology, Biology Centre CAS, České Budějovice, Czech Republic. [6]Department of Anthropology, Washington University in St Louis, St Louis, MO, USA. [7]Spatial Analysis, Interpretation, and Exploration Laboratory, Washington University in St. Louis, St. Louis, MO, USA. [8]Department of Archaeology, Ethnography and Museology, Altai State University, Barnaul, Russia. [9]Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden. [10]Historical, Ecological and Cultural Association 'Povolzhye', Samara, Russia. [11]Institute of Studies the Humanities and Problems of Indigenous People of the North, FIC Yakutsk Scientific Center of the Siberian Branch of the Russian Academy of Sciences, Yakutsk, Russia. [12]Diamond and Precious Metals Geology Institute, Siberian Branch of the Russian Academy of Sciences, Yakutsk, Russia. [13]Institute of Archaeology and Ethnography, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia. [14]Tyumen Scientific Center of the Siberian Branch of Russian Academy of Sciences, Institute of Problems of Northern Development, Tyumen, Russia. [15]Department of Archaeology of Central Asia and the Caucasus, Institute for the History of Material Culture of Russian Academy of Science, Saint Petersburg, Russia. [16]Samara State University of Social Sciences and Education, Samara, Russia. [17]Department of Recreational Geography, Service, Tourism and Hospitality, Institute of Geography, Altai State University, Barnaul, Russia. [18]Center of Human Ecology, Institute of Ethnology and Anthropology, Russian Academy of Science, Moscow, Russia. [19]Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russia. [20]Ural Federal University, Yekaterinburg, Russia. [21]Independent Researcher, Kemerovo, Russia. [22]Krasnoyarsk Regional Museum of Local Lore, Krasnoyarsk, Russia. [23]LLC 'Archaeology of the East European Plain', Moscow, Russia. [24]National Research Tomsk State University, Tomsk, Russia. [25]V. F. Voino-Yasenetsky Krasnoyarsk State Medical University, Krasnoyarsk, Russia. [26]Institute of Natural Sciences, M. K. Ammosov North-Eastern Federal University, Yakutsk, Russia. [27]Institute of History and Archaeology, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russia. [28]Siberian State University of Physical Education and Sport, Omsk, Russia. [29]Arctic Research Center of Sakha Republic, Yakutia, Russia. [30]Human Population Genetics Laboratory, Research Center for Medical Genetics, Moscow, Russia. [31]Laboratory of Human Molecular Genetics, Institute of Molecular and Cellular Biology, Siberian Branch of the Russian Academy of Sciences (SBRAS), Novosibirsk, Russia. [32]Department of Genetics, Evolution and Environment, University College London Genetics Institute (UGI), University College London, London, UK. [33]Department of Genetics, Yale Medical School, New Haven, CT, USA. [34]Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [35]Department of Biology, University of La Verne, La Verne, CA, USA. [36]Laboratory of Archaeogenetics, Nazarbayev University, Astana, Kazakhstan. [37]Laboratory of Human Genetics, National Center for Biotechnology, Astana, Kazakhstan. [38]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA. [39]Human Evolution and Archaeological Sciences, University of Vienna, Vienna, Austria. [40]Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria. [41]CIAS, Department of Life Sciences, University of Coimbra, Coimbra, Portugal. [42]School of Archaeology, University College Dublin, Dublin, Ireland. [43]Servizio di Bioarcheologia, Museo delle Civiltà, Rome, Italy. [44]Utrecht University, Utrecht, Netherlands. [45]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [46]BIOMICs Research Group, Department of Zoology and Animal Cell Biology, University of the Basque Country UPV/EHU, Vitoria-Gasteiz, Spain. [47]Ikerbasque-Basque Foundation of Science, Bilbao, Spain. [48]Department of Integrative Biology, University of Texas, Austin, TX, USA. [49]Department of Statistics and Data Science, University of Texas, Austin, TX, USA. [50]These authors contributed equally: Tian Chen Zeng, Leonid A. Vyazov, Alexander Kim. ✉e-mail: tianchen_zeng@g.harvard.edu; l.a.vyazov@gmail.com; akim@fas.harvard.edu; ron.pinhasi@univie.ac.at; vagheesh@utexas.edu; reich@genetics.med.harvard.edu

Methods

Sampling of ancient individuals

All skeletal samples screened for ancient DNA were analysed with permission from the appropriate authorities including in every case archaeologist or anthropologist custodians of the samples, and/or cultural institutions curating the samples. Descriptions of the archaeological and cultural contexts for all ancient samples analysed, including their grave position within archaeological sites, their grave numbers and burial inventory, as well as references to archaeological publications for the sites themselves (where available), are provided in Supplementary Information, section 3. Contact information for finding out more about the samples we analysed are listed in column G in the sheet labelled 'Ancient individuals' in Supplementary Data 1 (all samples, including previously published samples) and 'Bone samples and libraries' in Supplementary Data 1 (samples analysed for this paper, including material that did not yield enough DNA for analysis). Samples may be identified by their skeletal code listed in the 'Ancient individuals' and 'Bone samples and libraries' sheets of Supplementary Data 1.

Sampling of present-day individuals

We newly genotyped 229 present-day individuals from 10 ethnolinguistic groups using the Affymetrix Human Origins SNP array. All DNA samples were collected with written informed consent for broad studies of population history and full public release of de-identified genetic data, using a protocol approved by the Ethics Committee of the Research Centre for Medical Genetics, Moscow, Russia. All newly reported data are represented either by co-authors of this study or individuals who wished to be mentioned in the Acknowledgments who were involved in sample collection. Details of all present-day genetic samples analysed (all samples, including previously published samples) are given in Table 1 in the sheet 'Present-day individuals' in Supplementary Data 1, while details entirely of newly published samples are provided in the sheet 'Newly-published individuals' in Supplementary Data 1.

Ancient DNA data generation, bioinformatic processing and quality control

We targeted collection of 37 mg of powder from skeleton, after which DNA was extracted using a protocol that retains short and damaged DNA fragments[62,63]. The powder was collected from petrous bones, long bones, teeth and ossicles. Individually barcoded double[64,65] and single-stranded libraries[66] were built after incubation with uracil-DNA-glycosylase (UDG treatment[65], to reduce errors characteristic of ancient DNA damage). We performed in-solution enrichment for ~1.2 million SNPs (1240K enrichment[67]) and also enriched for the mitochondrial genome[68]. Two rounds of enrichment were performed, after which sequencing was performed on the Illumina NextSeq 500 or HiSeq X 10 instruments.

The resulting read pairs were separated using library-specific barcode pairs or index pairs (for double-stranded and single-stranded libraries respectively) and merged prior to alignment. Read pairs were merged if: (1) 15 or more base pairs (bp) overlap; (2) at most one mismatch occurred and base quality was at least 20; (3) at most three mismatches occurred and base quality was lower than 20. The resulting sequences were aligned to the human genome reference sequence (hg19)[69] and the mitochondrial RSRS genome using samse from bwa-v.0.6.1[70,71]. Duplicated sequences were removed if they shared start and stop positions, orientation, and (for double-stranded libraries) barcode pairs. Analysis was performed on sequences at least 30 bp in length. We trimmed 2 bp from the ends of each read to reduce deamination errors. For each sample, we merged the sequences from all libraries. Most of the datasets used for population genetic analysis were generated by randomly sampling at each SNP on chromosomes 1–22 and X, with a mapping threshold of 10 and base quality 20.

We flagged as 'questionable' libraries that had evidence of contamination based on the upper bound of the match rate to the mitochondrial consensus sequence (assessed using contamMix v1.0-10,[72]) being less than 95%; we also flagged as 'critical' libraries if this value was less than 90% (sheet labelled 'Ancient individuals' in Supplementary Data 1). We flagged as 'questionable' males with evidence of high polymorphism on the X chromosome (lower bound of the 95% confidence interval for mismatch rate >1%), or as 'critical' (if >5%), estimated using ANGSD v0.923[73]. For high-coverage contaminated individuals, we generated alternative sequences restricting to molecules showing signs of characteristic ancient DNA damage (designated by a suffix '_d' in the Genetic ID of the sample in the 'Ancient individuals' sheet of Supplementary Data 1).

For a subset of 15 individuals with high percentages of human DNA, we generated shotgun sequences (designated by the suffix '.SG' in the 'Ancient individuals' sheet of Supplementary Data 1) using the pre-enrichment libraries. We carried out sequencing on an Illumina HiSeq X Ten instrument. These shotgun sequences were used for analysis only in PCAs (Supplementary Information, section 4).

Uniparental analysis

Mitochondrial haplogroups were determined with Haplogrep v2.1.1[74]. Y-chromosome haplogroups were evaluated using the methodology described in[75], section S5, using both targeted and off-target SNPs. Allelic status was determined by majority rule.

ADMIXTURE and PCA

All relatives and shotgun sequences were excluded from ADMIXTURE analysis. For relative pairs or groups, the lower-coverage individual was excluded.

We used ADMIXTURE v.1.3.0[76] after pruning SNPs with high missingness in plink v.1.0.7 (using option–geno 0.5[77]), after which 597,573 autosomal SNPs were retained. We used $K = 18$ as the first K value where Yakutia_LNBA and Cisbaikal_LNBA were separated from East Asian components characteristic of NEAHG populations (for example, the components maximized in Mongolia_N_North and AmurRiver_14K). Further details of our application of ADMIXTURE can be found in Supplementary Information, section 5, including our ADMIXTURE cross-validation error (Supplementary Information, section 5).

We pruned individuals from PCA analysis if they were found to be a first-degree relative of another individual in the dataset with high coverage. PCA was performed using smartPCA in the EIGENSOFT package[78], using numoutlier: 0 and lsqproject: YES for three out of four PCAs. Further details on our PCAs can be found in Supplementary Information, section 4.

qpAdm analyses and f_4-statistics

All f_4-statistics were calculated using the qpDstat package of Admixtools v.7.0[79] with the f4mode: YES parameter. Further details of each set of f_4-statistic calculations can be found where they are presented, in Supplementary Information, sections 8, 9, 11 and 13.

All qpAdm analyses were run using the R package Admixtools2[80]. Precalculated f_2-statistics, used to speed up the process of f_4-ratio estimation central to qpAdm, were performed allowing for maximal missingness = 0.99 over multiple datasets. Further details for each set of qpAdm can be found in Supplementary Information, sections 8–11 and 13. Additionally, details on all our sets of qpAdm analyses can be found in Extended Data Table 1.

The results of these qpAdm analyses are found in Supplementary Data 2–7. In these files, the tables listing qpAdm results are sorted first by target; then for each target, models with all positive coefficients are listed first, ahead of the rest. The all-positive-coefficient models for each target are themselves sorted, first by simplicity (that is, one-source all-positive models listed first, then two-source all-positive models, then three-source, etc.), and then (among the all-positive models with the same number of sources) ranked by P value. This same ordering is used for models with negative coefficients for each target (that is, they are listed first by simplicity, then by *P* value). Results for each target population are easily accessible by filtering on the 'Target' column, and then by the threshold *P* value one picks, which would automatically list all passing models starting with the simplest all-positive models with the highest *P* value.

We made sure whenever possible that the populations included in left/sources and right/references in our qpAdm sets were always processed through only one set of wet laboratory procedures: through 1240K enrichment. For analyses where population groups in the left/sources and right/references included both 1240K and shotgun sequences, wherever possible, we performed replicate analyses where shotgun individuals were purged from all the group labels in the left/sources and right/references. Our replicate analyses show that our main conclusions in qpAdm are relatively robust to the effects of allelic bias (Supplementary Information, sections VI.C.ii.a and VI.D.ii.a).

Relatedness and runs of homozygosity

We looked for kinship relationships between the individuals included in our study. We computed pairwise allelic mismatch rates in the autosomes by randomly sampling one DNA sequence at each 1240K polymorphic position, following the same strategy as in refs. 81–83, which is similar to that in ref. 84. We then estimated relatedness coefficients *r* for each pair as in ref. 81:

$$r = 1 - ((x - b)/b)$$

with x being the mismatch rate of the pair under analysis and b the base mismatch rate expected for two genetically identical individuals from the population under analysis, which we estimated by computing intra-individual mismatch rates. We also computed 95% confidence intervals using block jackknife standard errors over 5-megabase (Mb) blocks.

Ethics statement

This study prioritized minimizing damage to skeletal remains during analysis. Authorization was obtained from local authorities in each region of origin. Sample stewards are local archaeologists or museum curators, who either contributed as authors or are acknowledged in this work. To uphold open science principles, we have made both the electronic sequence data and the physical ancient DNA libraries publicly accessible, ensuring comprehensive data availability. Researchers interested in conducting additional sequencing of these should contact the corresponding author, D.R. Requests will be accommodated as long as the libraries remain intact in our care, with no obligation to include us as collaborators or co-authors on subsequent publications.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The newly reported data in this study can be obtained from the European Nucleotide Archive under accession number PRJEB86428. Bam files of aligned reads for the 180 newly published ancient individuals and 15 newly reported whole-genome sequences from a subset of these individuals can be found at secondary accession ERP169776, and the genotypes that we used for analysis can be found at secondary accession ERZ25719453. Genotype files in PLINK format for the 229 modern individuals for whom we newly report SNP array can be found at secondary accession ERZ26790638. All maps in the main text and in the Supplementary Information were created using ArcGIS 10.6.1 and QGIS 3.40.6. Figures presenting genetic data were created using Rstudio running R version 4.4.1, and further edited in Adobe Illustrator version 28. Archaeological images in Supplementary Information, section 3 were edited in Adobe Photoshop 25.12.2 and Adobe Acrobat 2025.001.20458.

62. Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* **13**, 2447–2461 (2018).
63. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
64. Briggs, A. W. & Heyn, P. in *Ancient DNA. Methods in Mol. Biol.* (eds Shapiro, B. & Hofreiter, M.) https://doi.org/10.1007/978-1-61779-516-9_18 (2012).
65. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* **370**, 20130624 (2015).
66. Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & M MEYER, Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020).
67. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
68. Maricic, T., Whitten, M. & Pääbo, S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS ONE* **5**, e14004 (2010).
69. Speir, M. L. et al. The UCSC Genome Browser Database: 2016 update. *Nucleic Acids Res.* **44**, D717–D725 (2016).
70. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
71. Behar, D. M. et al. A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90**, 675–684 (2012).
72. Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
73. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
74. Weissensteiner, H. et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).
75. Lazaridis, I. et al. The genetic history of the Southern Arc: a bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
76. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
77. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
78. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
79. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
80. Maier, R., Flegontov, P., Flegontova, O., Işıldak, U., Changmai, P. & Reich, D. On the limits of fitting complex models of population history to *f*-statistics. *eLife* **12**, e85492 (2023).
81. Kennett, D. J. et al. Archaeogenomic evidence reveals prehistoric matrilineal dynasty. *Nat. Commun.* **8**, 14115 (2017).
82. Van de Loosdrecht, M. et al. Pleistocene North African genomes link near Eastern and sub-Saharan African human populations. *Science* **360**, 548–552 (2018).
83. Olalde, I. et al. The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).
84. Monroy Kuhn, J. M., Jakobsson, M. & Günther, T. Estimating genetic kin relationships in prehistoric populations. *PLoS ONE* **13**, e0195491 (2018).

Acknowledgements We dedicate this paper to Oleg Balanovsky, who had a leading role in the collection of present-day samples newly reported in this study, and who would have been an author had he not died in 2021. The authors thank N. Adamski, R. Bernardos, N. Bradman, A. Chizhevsky, M. Ferry, E. Idrisov, J. Kidd, E. Kostyleva, S. Kuz'minykh, K. Mandl, P. Nymadawa, O. Poshekhonova, H. Ringbauer, L. Saroyants, K. Stewardson, S. Tur, Y. Yusupov and Z. Zhang for wet laboratory or bioinformatic support, providing permission to analyse samples that they shared or critical comments. We acknowledge E. Besprozvanny, T. Chikisheva, A. Chizhevskiy, O. Goryunova, E. Kostyleva, N. Kungurova, D. Maslyuzhenko, A. Polevodov, A. Shalapinin, G. Sinitsyna, Z. Trufanova and V. Zakh for providing permissions to use their previously published figures in the supplementary information. A.A.T. acknowledges support from the Russian Science Foundation (project 22-18-00470). M.Z. acknowledges support from the Collaborative Research Grants Program 091019CRP2119 to Nazarbayev University. The research of G.G.B. was conducted within the framework of the scientific research programme of the Diamond and Precious Metals Geology Institute, Siberian Branch of the Russian Academy of Science (project FUFG-2024-0005). Research by A.D.S., E.N.S. and V.M.D. was carried out within the research programme of the Institute of Archeology and Ethnography of the Siberian Branch of the Russian Academy of Sciences 'The Stone Age of Northern Asia: Cultural and Ecological context (FWZG-2025-0010). Research by D.N.E., S.N.S., S.M.S. and K.N.S. was carried out within the State Assignment FWRZ-2021-0006. A.V.F. was supported by the IHMC RAS research programme (FMZF-2025-0008). M.G.T. is supported by ERC Horizon 2020 research and innovation programme grant agreements: 951385 (COREX), 865515 (SUSTAIN), 324202 (NeoMilk), 788616 (YMPACT), and by Wellcome Senior Research Fellowship Grant 100719/Z/12/Z. G.H. was supported by BBSRC (BB/L009382/1), Wellcome Trust and Royal Society (098386/Z/12/Z, 224575/Z/21/Z). We thank the Museum of the Institute of Plant and Animal Ecology UB RAS for sharing samples. P.F. was supported by the Czech Science Foundation (project 21-27624S) and the EU Operational Program Just Transition ('LERCO—Life Environment Research Center Ostrava', project CZ.10.03.01/00/22_003/0000003). L.A.V. was supported by the Czech Ministry of Education, Youth and Sports (programme ERC CZ, project LL2103). P.F., R.P. and D.R. were supported by John Templeton Foundation grant 61220. P.F. and D.R. were supported by gifts from Jean-Francois Clin. D.R. was supported by National Institutes of Health grant HG012287 and by the Allen Discovery Center programme, a Paul G. Allen Frontiers Group advised programme of the Paul G. Allen Family Foundation, and is an Investigator of the Howard Hughes Medical Institute.

Author contributions T.C.Z., L.A.V., A.K. and D.R. wrote the manuscript and supplementary materials with input from all co-authors. S.M., N.R., R.P., V.M.N. and D.R. supervised different aspects of the study. T.C.Z. carried out the main genetic analyses under the supervision of P.F., R.M. and V.M.N. I.O. and I.L. contributed additional genetic analyses. M.F., P.F., V.M.N. and A.A.T. contributed to the framing and interpretation of results. L.A.V. edited archaeological information with input from D.R. and P.F. K.S. and L.A.V. contributed extensively to sample procurement. A.A.T, N.E.R., S.A.A., D.S.A., A.N.A., G.G.B., A.P.D., V.M.D., D.N.E., A.V.F., Y.V.F., S.P.G., A.A.K., K.Y.K., Y.F.K., E.P.K., P.K., I.V.K., N.P.M., V.V.M., E.N.N., M.P.R., T.M.S., M.V.S., V.S., S.N.S., O.S.S., S.M.S., K.N.S., E.N.S., A.D.S., A.A.T., A.S.V., A.V.V. and R.P. contributed anthropological remains and/or contributed to the creation of the archaeological supplement. N.E.R. wrote the 'geophysical context' section of the Supplementary Data. E.V.B., S.D., G.H., K.K., J.K., E.S., R.S., T.T., M.G.T. and M.Z. contributed genetic data from present-day people. A.A., M. Mah, A.M. and S.M. carried out bioinformatic data processing. K.C., O.C., D.F., D.K., F.C., L.I., A.K., K.T.Ö., F.Z. and M. Michel carried out wet laboratory work.

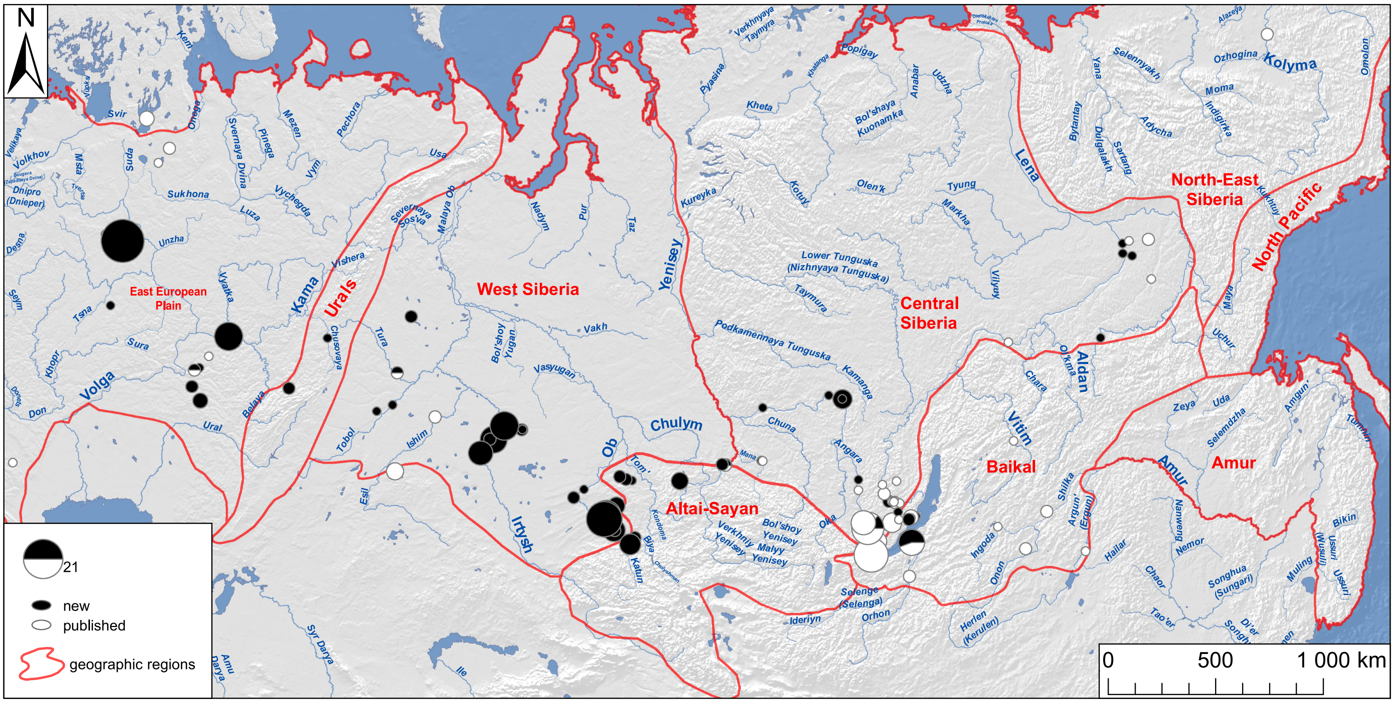
Competing interests The authors declare no competing interests.

Additional information

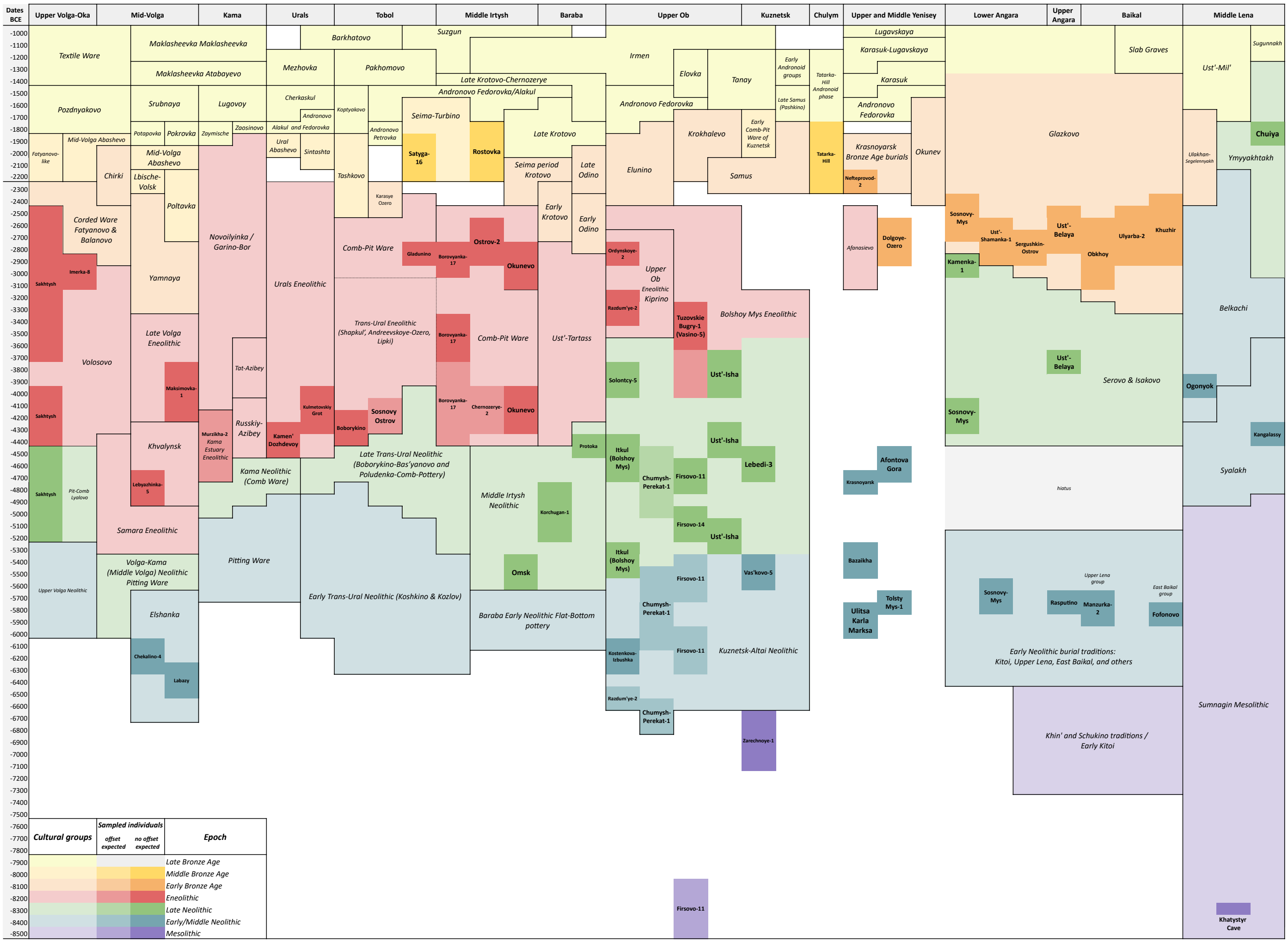
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-09189-3.

Correspondence and requests for materials should be addressed to Tian Chen Zeng, Leonid A. Vyazov, Alexander Kim, Ron Pinhasi, Vagheesh M. Narasimhan or David Reich.

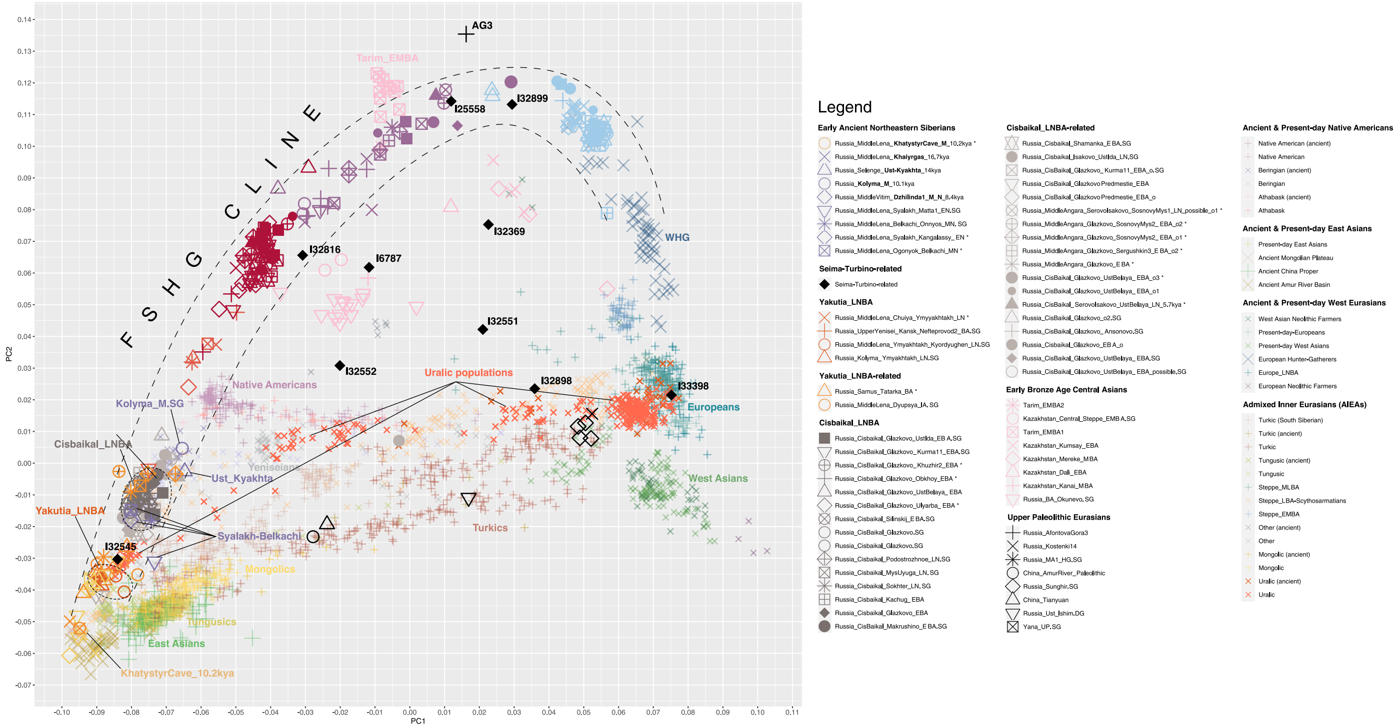
Peer review information *Nature* thanks Henny Piezonka, Edward Vajda and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.



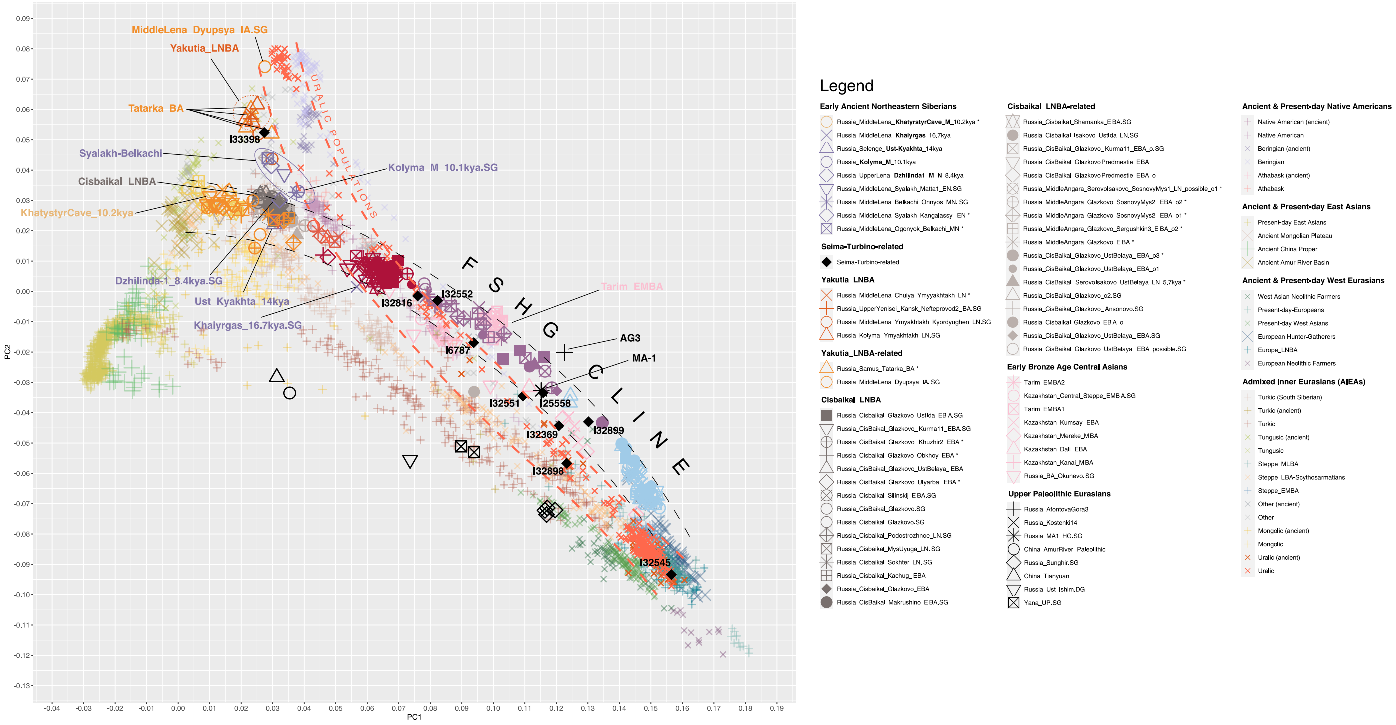
Extended Data Fig. 1 | Sites with newly-reported samples. This map displays all the sites which are the sources of the samples in the major populations that are the focus in this paper. These include all sites 1) whose samples fall on the NEAHG cline, 2) whose samples fall in the Cisbaikal_LNBA cluster or are admixed with it, 3) whose samples fall in the Yakutia_LNBA cluster or are admixed with it, 4) whose samples are a part of the ten-population East Siberian transect described in our qpAdm modelling, and 5) whose samples are from Seima-Turbino period individuals. Each site is represented by a pie chart, whose size is proportional to the number of individuals from that site; the white fraction represents previously-published samples, and the black newly-published samples. Our sampling fills geographic and temporal lacunae.



Extended Data Fig. 2 | Chronology of sites and cultures in each geographic region. Temporal and geographic disposition of cultures from the Mesolithic to the Late Bronze and Iron Ages across Northern Eurasia. Sites whose samples are analyzed in our paper are highlighted in darker boxes, within containing boxes indicating archaeological cultures. Sites whose colors are darker are those that we believe are most securely dated (based on radiocarbon, isotopic, and archaeological evidence).

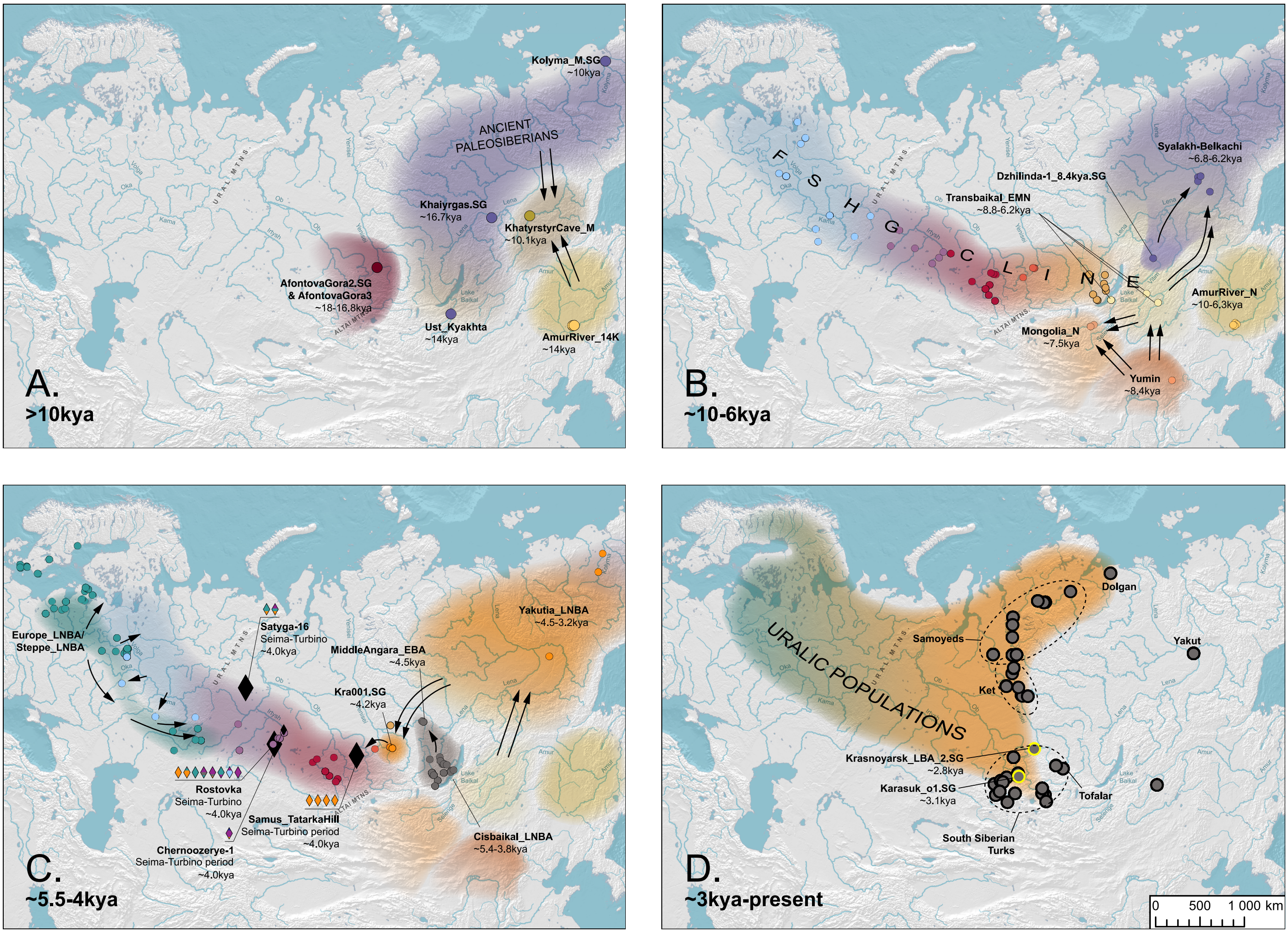


Extended Data Fig. 3 | PCA with target populations projected onto ancient populations with an especially high fraction of ANE ancestry. To illuminate the role that levels of *ANE* ancestry plays in generating variation among the populations we analyze, we use as a basis for another projection 71 shotgun-sequenced ancient individuals from across Eurasia, of which a large proportion are enriched in *ANE* ancestry and fall outside the range of present-day variation (e.g. individuals from populations like *Tyumen_HG.SG* or *Kazakhstan_Botai.SG*; for full list, see Supplementary Information section 4). The North Eurasian Hunter-Gatherer cline forms a curved arc stretching from *EHG* populations to present-day East Asians; the center of the arc dominated by populations rich in *ANE* ancestry is moved toward the positive direction in PC2. The individual furthest along the positive direction in PC2 is AG3. Clines formed by later Inner Asian populations, such as present-day Uralic, Turkic, and Mongolic speakers, as well as Late Bronze Age and Iron Age steppe populations such as Scythians and Sarmatians, are distinguished from the *NEAHG* cline by their much lower values along PC2, suggesting a much lower level of ANE ancestry. This PCA shows that populations along the NEAHG cline, remaining stable for many millennia, were substantially outside the range of present-day genetic variation in Northern Eurasia.

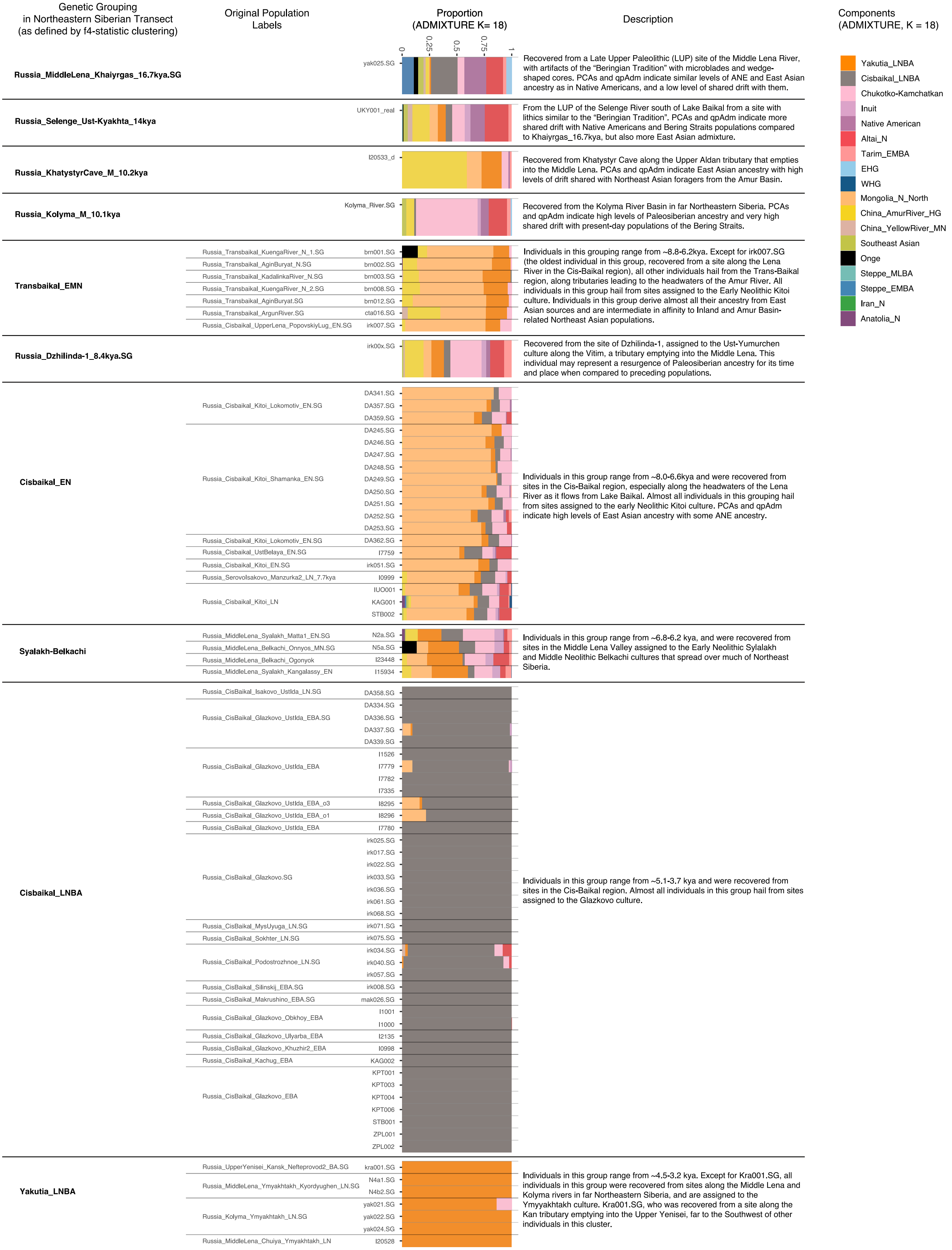


Extended Data Fig. 4 | PCA focusing on East Eurasian populations. To further uncover possible structure among the East Asian ancestries within the populations that we analyze, we constructed a third PCA, using as a basis 37 East Asian present-day populations that have minimal West Eurasian admixture, and a single West Eurasian population (Norwegian), all genotyped on the Affymetrix Human Origins array (for a full list of populations analyzed, refer to Supplementary Information section 4). We projected all other shotgun-sequenced and hybridization-captured ancient and present-day individuals onto this basis. Once again, the North Eurasian Hunter-Gatherer cline forms a curved arc stretching from West Eurasian populations to present-day East Asians, with the center of the arc deflected toward the AG3 individual. East Asian populations are now differentiated along PC2, with Southeast Asians and East Asian agriculturalists taking on especially negative values along that dimension; populations from the Amur River Basin taking on intermediate values; then populations on the Mongolian Plateau and surrounding areas. A large gap separates these populations from *Yakutia_LNBA* and Russia_Tatarka_BA, which take on very positive values along PC2, close to present-day Nganasans and a genetically very similar Iron-Age individual from Yakutia who clusters with Nganasans in the previous two PCAs (Yakutia_IA.SG; also see Extended Data Fig. 9). As one moves East along the *NEAHG* cline, their positions along PC2 tend to converge to the values found among populations of the Mongolian Plateau. In contrast, the Dzhilinda1_M_N_8.4 kya and Kolyma_M_10.1 kya individuals, and the Syalakh_Belkachi, *Yakutia_LNBA* and Russia_Tatarka_BA populations do not fall on the *NEAHG* cline and are shifted in the positive direction on PC2, toward the positions occupied by Nganasans, Beringian populations, and Native Americans. Lastly, Uralic populations possess the most positive values among PC2 when compared to Turkic, Mongolic and Tungusic populations.

Extended Data Fig. 5 | PCA focusing on ancient individuals from Northern Eurasia and the Americas. To understand structure among *NEAHG* populations and non-*NEAHG* Siberians, we constructed two PCAs with ancient individuals including all individuals from the *NEAHG* cline, ancient non-*NEAHG* Siberians, and a selection of ancient Beringians and Native Americans. Notably, all these populations possess combinations of only *WHG*, *EHG*, *ANE* and East Asian ancestries. No individuals were projected in these PCAs. The first PCA (Extended Data Fig. 6a) includes all individuals in the set, and the second (Extended Data Fig. 6b) includes only individuals East of the Altai mountains. (**A**) In the first PCA we highlight several patterns. 1) the North Eurasian Hunter-Gatherer cline forms a curved arc stretching from West Eurasian populations to East Asian populations along PC1 and PC2. Populations rich in East Asian ancestry are differentiated along PC3, with individuals and populations within or closely related to the *Cisbaikal_LNBA* cluster having the most positive values, followed by those in the *Transbaikal_EMN* cluster and populations of the Mongolian Plateau, followed by individuals and populations in the *Yakutia_LNBA* cluster, followed by those from the Amur River Basin, followed by populations from the Bering Straits and the Americas. Notably, all individuals along the *NEAHG* cline, including individuals rich in East Asian ancestry (e.g. *Cisbaikal_EN*, *Transbaikal_EMN*, and all *NEAHG* individuals from the Krasnoyarsk region) form a straight line in PC3, suggesting a constant source of East Asian ancestry at the East Asian terminus of the *NEAHG* cline. 2) *Khaiyrgas_16.7 kya* occupies a central position among the other groups rich in East Asian ancestry in East Siberia, Beringia and the Americas, suggesting a lack of shared drift with later populations of the Bering region or the Americas. The situation is different for later populations: *Kolyma_M_10.1 kya* falls among ancient Beringian populations, while the more East Asian-admixed *Ust-Kyakhta_14 kya* and *Dzhilinda1_M_N_8.4kya* occupy a position in between *Syalakh-Belkachi* and ancient Bering Straits populations, with the even more East Asian-admixed *Syalakh-Belkachi* population showing even less of this displacement towards ancient Bering Straits populations. (**B**) We find a similar pattern in the second PCA, except with an opposite ordering of the clusters along PC3. Our results suggest that the distinctions we discover between groupings produced by the clustering analyses in Supplementary Information Section 6 can be recovered in PCA analyses aimed at recovering fine-scale structure, despite underlying similarities in deep ancestry in populations in East Siberia, Beringia, and the Americas—all the products of admixture between ANE and East Asian ancestry.

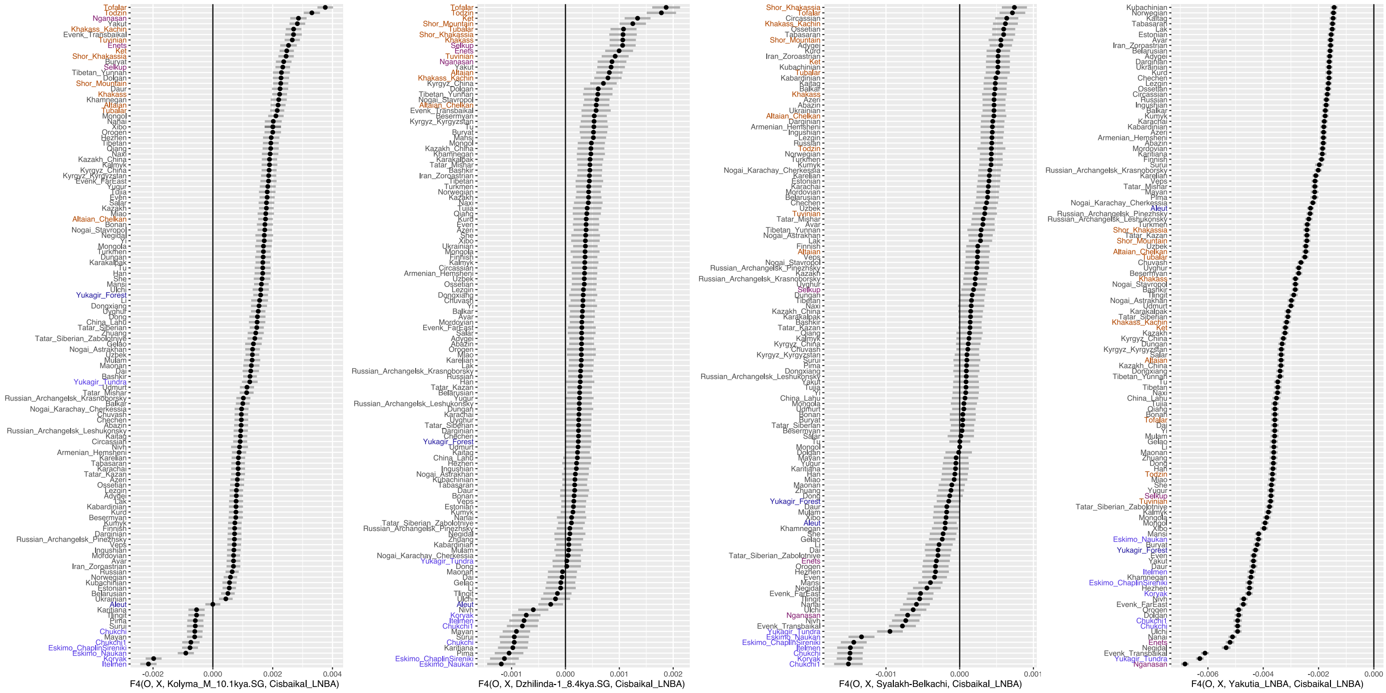


Extended Data Fig. 6 | Graphical Summary of Genetic Changes Taking Place in Northern Eurasia. Panel A shows the widespread distribution of individuals with Ancient Paleosiberian (APS) ancestry in Siberia before the Holocene, >10 kya. Panel B shows the formation of the NEAHG cline by ~10 kya, and the formation of the population on its eastern terminus (Transbaikal_EMN) through admixture between Amur River and Inland East Asian ancestries. Panel C shows the emergence of Cisbaikal_LNBA and Yakutia_LNBA in genetic turnovers in the Cis-Baikal and Northeastern Siberian regions in the Mid-Holocene, and the genetic diversity of Seima-Turbino period individuals ~4.0 kya. Panel D shows the genetic gradient between West Eurasian ancestry and Yakutia_LNBA formed by present-day Uralic populations, along with all locations from which present-day populations with Cisbaikal_LNBA ancestry were sampled (grey dots ringed with black), alongside the geographic locations of two late Bronze Age/early Iron Age individuals (grey dots ringed with yellow) with >90% Cisbaikal_LNBA ancestry.

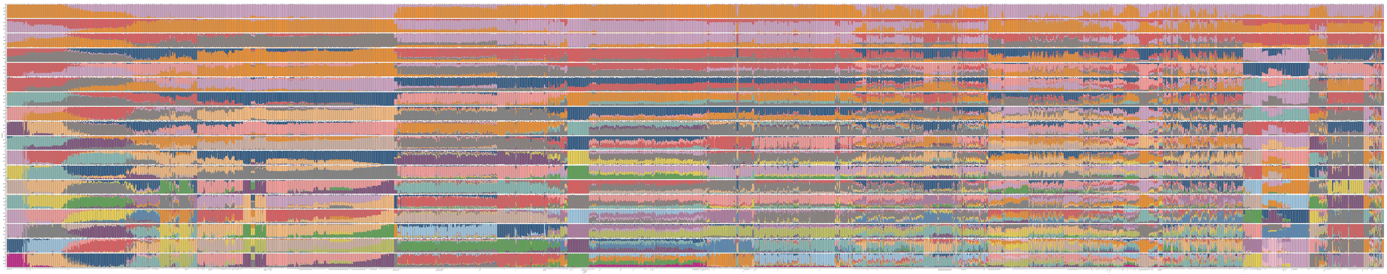


Genetic Grouping in Northeastern Siberian Transect (as defined by f4-statistic clustering)	Original Population Labels	Proportion (ADMIXTURE K= 18)	Description
Russia_MiddleLena_Khaiyrgas_16.7kya.SG		yak025.SG	Recovered from a Late Upper Paleolithic (LUP) site of the Middle Lena River, with artifacts of the "Beringian Tradition" with microblades and wedge-shaped cores. PCAs and qpAdm indicate similar levels of ANE and East Asian ancestry as in Native Americans, and a low level of shared drift with them.
Russia_Selenge_Ust-Kyakhta_14kya		UKY001_real	From the LUP of the Selenge River south of Lake Baikal from a site with lithics similar to the "Beringian Tradition". PCAs and qpAdm indicate more shared drift with Native Americans and Bering Straits populations compared to Khaiyrgas_16.7kya, but also more East Asian admixture.
Russia_KhatystyrCave_M_10.2kya		I20533_d	Recovered from Khatystyr Cave along the Upper Aldan tributary that empties into the Middle Lena. PCAs and qpAdm indicate East Asian ancestry with high levels of drift shared with Northeast Asian foragers from the Amur Basin.
Russia_Kolyma_M_10.1kya		Kolyma_River.SG	Recovered from the Kolyma River Basin in far Northeastern Siberia. PCAs and qpAdm indicate high levels of Paleosiberian ancestry and very high shared drift with present-day populations of the Bering Straits.
Transbaikal_EMN	Russia_Transbaikal_KuengaRiver_N_1.SG	brn001.SG	Individuals in this grouping range from ~8.8-6.2kya. Except for irk007.SG (the oldest individual in this group, recovered from a site along the Lena River in the Cis-Baikal region), all other individuals hail from the Trans-Baikal region, along tributaries leading to the headwaters of the Amur River. All individuals in this group hail from sites assigned to the Early Neolithic Kitoi culture. Individuals in this group derive almost all their ancestry from East Asian sources and are intermediate in affinity to Inland and Amur Basin-related Northeast Asian populations.
	Russia_Transbaikal_AginBuryat_N.SG	brn002.SG	
	Russia_Transbaikal_KadalinkaRiver_N.SG	brn003.SG	
	Russia_Transbaikal_KuengaRiver_N_2.SG	brn008.SG	
	Russia_Transbaikal_AginBuryat.SG	brn012.SG	
	Russia_Transbaikal_ArgunRiver.SG	cta016.SG	
	Russia_Cisbaikal_UpperLena_PopovskiyLug_EN.SG	irk007.SG	
Russia_Dzhilinda-1_8.4kya.SG		irk00x.SG	Recovered from the site of Dzhilinda-1, assigned to the Ust-Yumurchen culture along the Vitim, a tributary emptying into the Middle Lena. This individual may represent a resurgence of Paleosiberian ancestry for its time and place when compared to preceding populations.
Cisbaikal_EN	Russia_Cisbaikal_Kitoi_Lokomotiv_EN.SG	DA341.SG	Individuals in this group range from ~8.0-6.6kya and were recovered from sites in the Cis-Baikal region, especially along the Lena River as it flows from Lake Baikal. Almost all individuals in this grouping hail from sites assigned to the early Neolithic Kitoi culture. PCAs and qpAdm indicate high levels of East Asian ancestry with some ANE ancestry.
		DA357.SG	
		DA359.SG	
	Russia_Cisbaikal_Kitoi_Shamanka_EN.SG	DA245.SG	
		DA246.SG	
		DA247.SG	
		DA248.SG	
		DA249.SG	
		DA250.SG	
		DA251.SG	
		DA252.SG	
		DA253.SG	
	Russia_Cisbaikal_Kitoi_Lokomotiv_EN.SG	DA362.SG	
	Russia_Cisbaikal_UstBelaya_EN.SG	I7759	
	Russia_Cisbaikal_Kitoi_EN.SG	irk051.SG	
	Russia_SerovoIsakovo_Manzurka2_LN_7.7kya	I0999	
	Russia_Cisbaikal_Kitoi_LN	IUO001	
		KAG001	
		STB002	
Syalakh-Belkachi	Russia_MiddleLena_Syalakh_Matta1_EN.SG	N2a.SG	Individuals in this group range from ~6.8-6.2 kya, and were recovered from sites in the Middle Lena Valley assigned to the Early Neolithic Sylalakh and Middle Neolithic Belkachi cultures that spread over much of Northeast Siberia.
	Russia_MiddleLena_Belkachi_Onnyos_MN.SG	N5a.SG	
	Russia_MiddleLena_Belkachi_Ogonyok	I23448	
	Russia_MiddleLena_Syalakh_Kangalassy_EN	I15934	
Cisbaikal_LNBA	Russia_CisBaikal_Isakovo_UstIda_LN.SG	DA358.SG	Individuals in this group range from ~5.1-3.7 kya and were recovered from sites in the Cis-Baikal region. Almost all individuals in this group hail from sites assigned to the Glazkovo culture.
	Russia_CisBaikal_Glazkovo_UstIda_EBA.SG	DA334.SG	
		DA336.SG	
		DA337.SG	
		DA339.SG	
	Russia_CisBaikal_Glazkovo_UstIda_EBA	I1526	
		I7779	
		I7782	
		I7335	
	Russia_CisBaikal_Glazkovo_UstIda_EBA_o3	I8295	
	Russia_CisBaikal_Glazkovo_UstIda_EBA_o1	I8296	
	Russia_CisBaikal_Glazkovo_UstIda_EBA	I7780	
	Russia_CisBaikal_Glazkovo.SG	irk025.SG	
		irk017.SG	
		irk022.SG	
		irk033.SG	
		irk036.SG	
		irk061.SG	
		irk068.SG	
	Russia_CisBaikal_MysUyuga_LN.SG	irk071.SG	
	Russia_CisBaikal_Sokhter_LN.SG	irk075.SG	
	Russia_CisBaikal_Podostrozhnoe_LN.SG	irk034.SG	
		irk040.SG	
		irk057.SG	
	Russia_CisBaikal_Silinskij_EBA.SG	irk008.SG	
	Russia_CisBaikal_Makrushino_EBA.SG	mak026.SG	
	Russia_CisBaikal_Glazkovo_Obkhoy_EBA	I1001	
		I1000	
	Russia_CisBaikal_Glazkovo_Ulyarba_EBA	I2135	
	Russia_CisBaikal_Glazkovo_Khuzhir2_EBA	I0998	
	Russia_CisBaikal_Kachug_EBA	KAG002	
	Russia_CisBaikal_Glazkovo_EBA	KPT001	
		KPT003	
		KPT004	
		KPT006	
		STB001	
		ZPL001	
		ZPL002	
Yakutia_LNBA	Russia_UpperYenisei_Kansk_Nefteprovod2_BA.SG	kra001.SG	Individuals in this group range from ~4.5-3.2 kya. Except for Kra001.SG, all individuals in this group were recovered from sites along the Middle Lena and Kolyma rivers in far Northeastern Siberia, and are assigned to the Ymyyakhtakh culture. Kra001.SG, who was recovered from a site along the Kan tributary emptying into the Upper Yenisei, far to the Southwest of other individuals in this cluster.
	Russia_MiddleLena_Ymyakhtakh_Kyordyughen_LN.SG	N4a1.SG	
		N4b2.SG	
	Russia_Kolyma_Ymyakhtakh_LN.SG	yak021.SG	
		yak022.SG	
		yak024.SG	
	Russia_MiddleLena_Chuiya_Ymyakhtakh_LN	I20528	

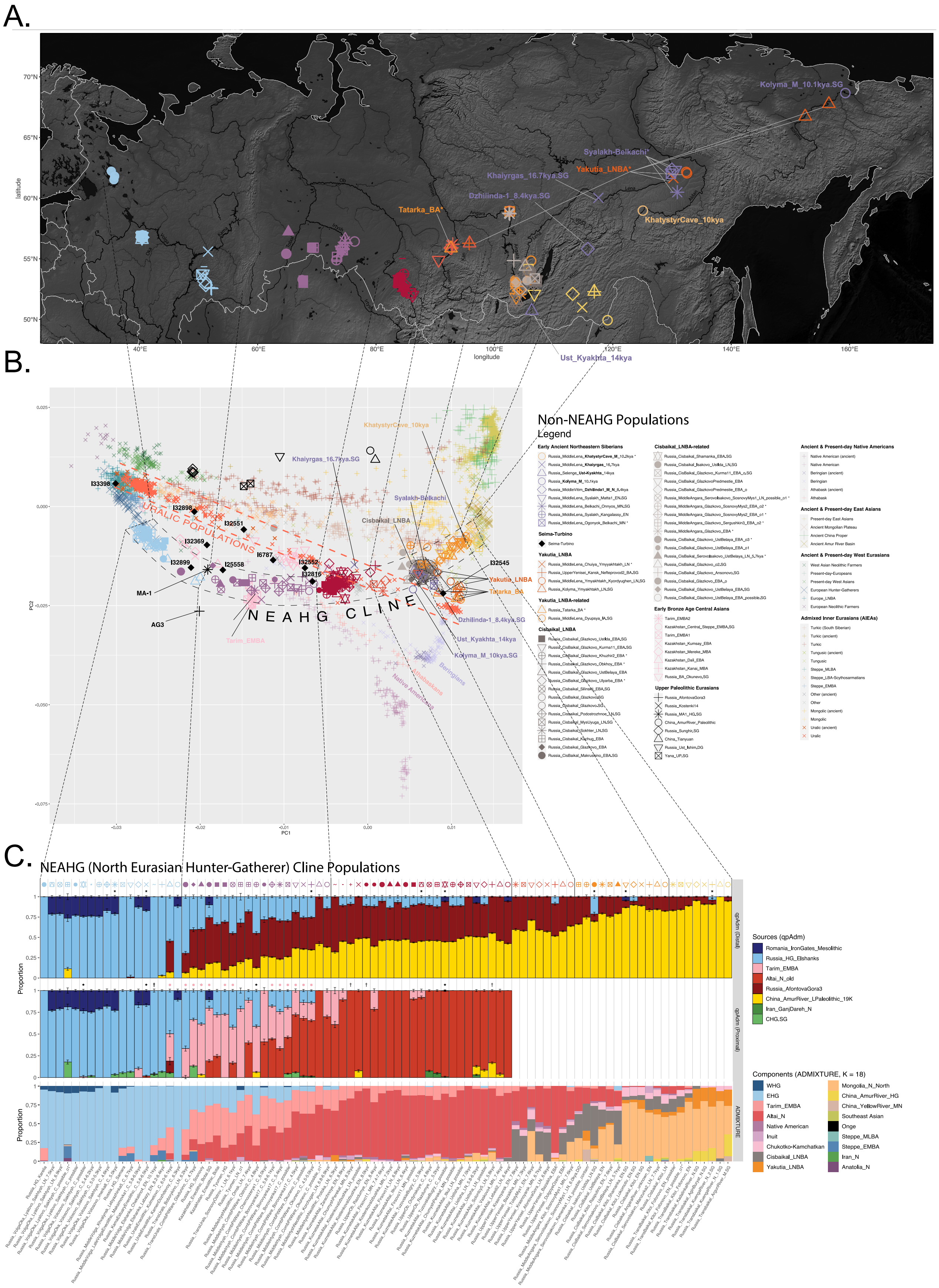
Extended Data Fig. 7 | Populations created by genetic grouping procedure applied over Northeast Siberians. Details of populations created by the grouping procedure applied to individuals in Northeastern Siberia.



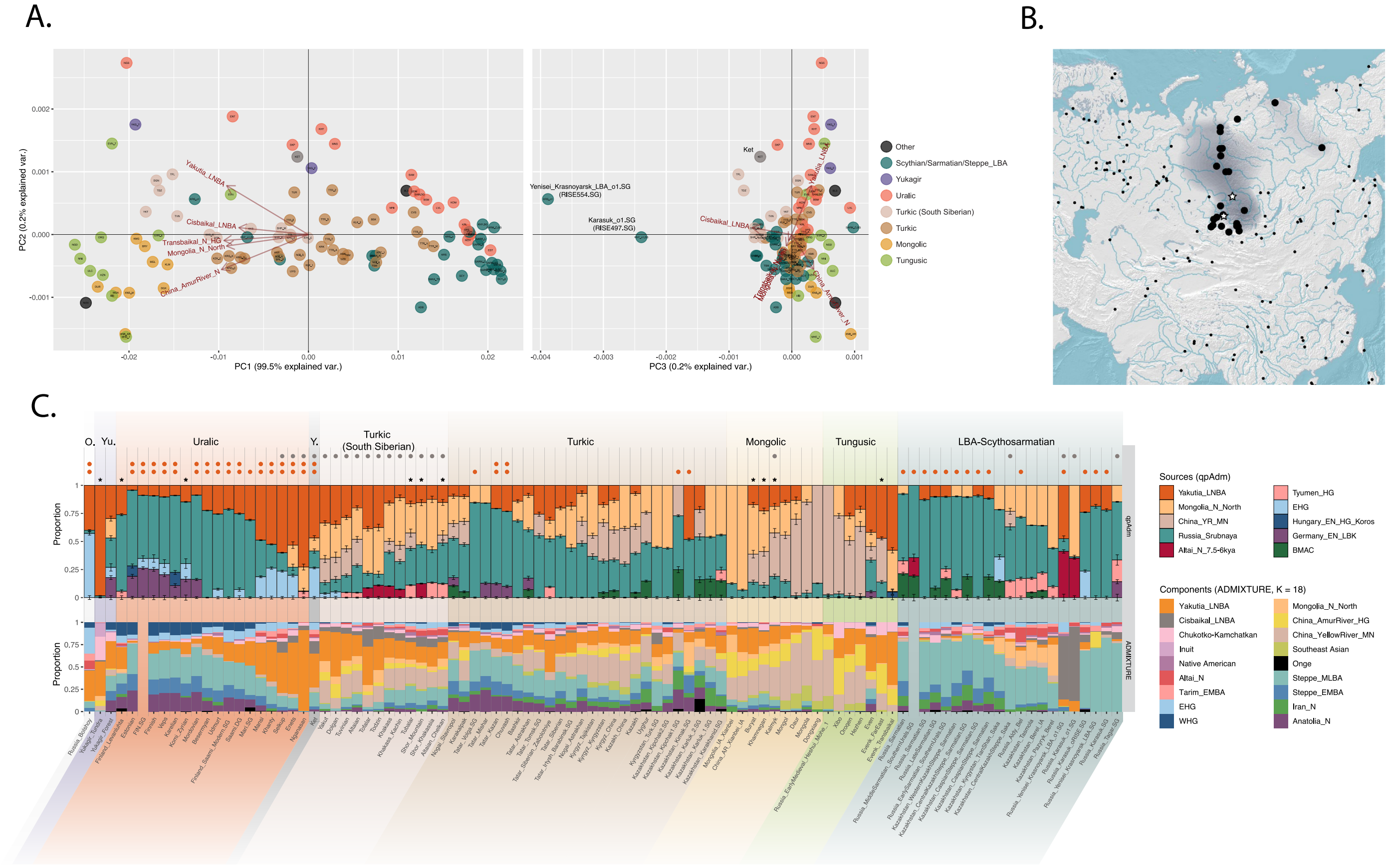
Extended Data Fig. 8 | Statistics of the form f4(Ethiopia_4500BP.SG, Target, "Route 2" population, Cisbaikal_LNBA). Central Siberian populations from the Yenisei Basin (including Kets and South Siberian Turks) are highlighted in brown, while Arctic North American and Asian populations on either side of the Bering Straits populations are highlighted in blue. Bering Straits populations that are heavily European-admixed (Aleut and Yukagir_forest) are colored dark blue, while Samoyedic populations (Enets, Selkup, and Nganasan) are colored violet. Despite the similarity of the APS-rich populations in this comparison (all being admixtures between APS ancestry and East Asian ancestry), present-day groups of the Bering Straits are always closer to groups with "Route 2" APS ancestry (i.e., Kolyma_M_10.1 kya → Dzhilinda1_8.4 kya → Syalakh-Belkachi → Yakutia_LNBA), while Central Siberian populations of the Yenisei Basin are always closer to Cisbaikal_LNBA. For the version including a comparison with Ust-Kyakhta, refer to Supplementary Information Section 8; Figs. S94 & S95.



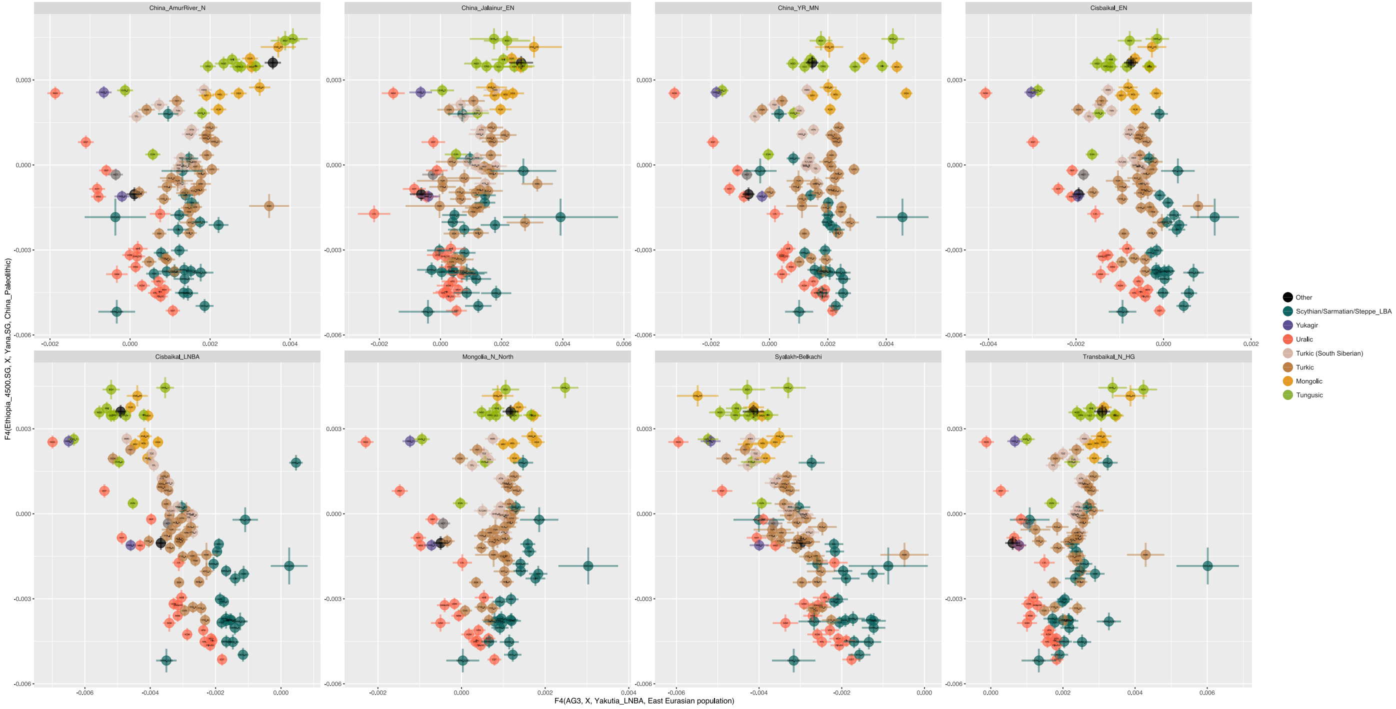
Extended Data Fig. 9 | ADMIXTURE results. For details, refer to Supplementary Information Section 5.



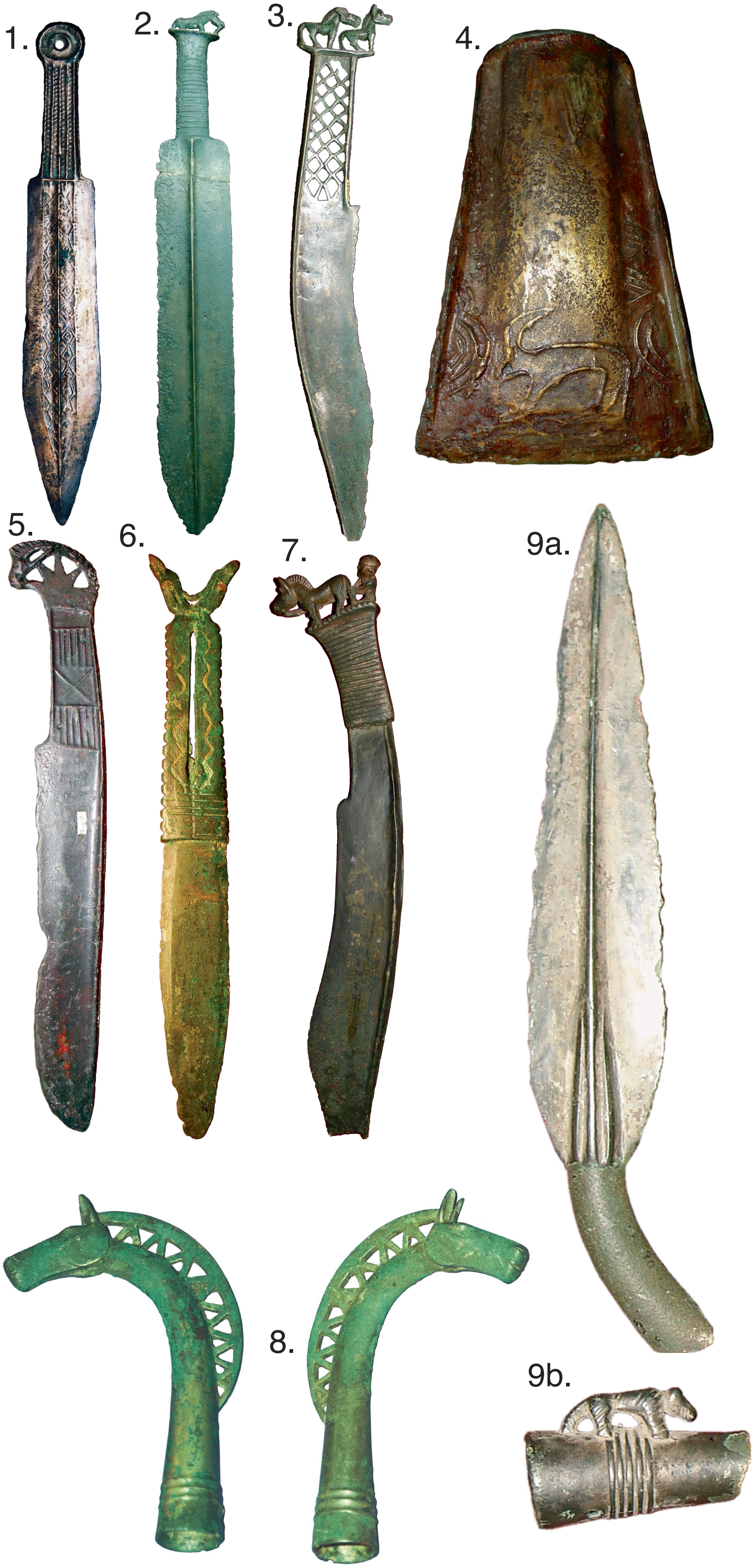
Extended Data Fig. 10 | The North Eurasian Hunter-Gatherer (NEAHG) Cline and its legacy through admixture in ancient northern Eurasia. Higher-resolution version of Fig. 1, containing the group/population labels of Fig. 1c.



Extended Data Fig. 11 | Contribution of Yakutia_LNBA and Cisbaikal_LNBA to Admixed Inner Eurasians (AIEA). Higher-resolution version of Fig. 3, containing the group/population labels. The codes are: ATN, Altaian; ATN_C, Altaian_Chelkan; BSK, Bashkir; BSM, Besermyan; BRY, Buryat; XNB_AR, China_AR_Xianbei_IA; CVS, Chuvash; DUR, Daur; DGN, Dolgan; DGX, Dongxiang; ENT, Enets; EST, Estonian; EVN, Even; EVN_E, Evenk_FarEast; EVN_T, Evenk_Transbaikal; FIN.SG, FIN.SG; LVL, Finland_Levanluhta; SAM, Finland_Saami_Modern.SG; FIN, Finnish; HZN, Hezhen; KLM, Kalmyk; KKP, Karakalpak; KRL, Karelian; KZK, Kazakh; KZK_C, Kazakh_China; BRL, Kazakhstan_Berel_IA; SARM_C, Kazakhstan_CaspianSteppe_Sarmatian; SARM_C.SG, Kazakhstan_CaspianSteppe_Sarmatian.SG; SAKA_K, Kazakhstan_CentralKazakhSteppe_Saka; SARM_K, Kazakhstan_CentralKazakhSteppe_Sarmatian.SG; KRK, Kazakhstan_Karakhanid.SG; KLK_1, Kazakhstan_Karluk_1.SG; KLK_2, Kazakhstan_Karluk_2.SG; KMK, Kazakhstan_Kimak.SG; KPC_1, Kazakhstan_Kipchak1.SG; KPC_2, Kazakhstan_Kipchak2.SG; SAKA_TS, Kazakhstan_Kyrgystan_TianShan_Saka; BRL_P, Kazakhstan_Pazyryk_Berel; TSM, Kazakhstan_Tasmola; SARM_W, Kazakhstan_WesternKazakhSteppe_Sarmatian; KET, Ket; KKS, Khakass; KKS_K, Khakass_Kachin; KMG, Khamnegan; KHT, Khanty; KOM, Komi_Zyrian; KRG_C, Kyrgyz_China; KRG_K, Kyrgyz_Kyrgyzstan; KRG_T, Kyrgyz_Tajikistan; TUR, Kyrgyzstan_Turk.SG; MNS, Mansi; MRI, Mari.SG; SCY, Moldova_Scythian; MGL, Mongol; MGA, Mongola; XNB_M, Mongolia_IA_Xianbei; MDV, Mordovian; NNI, Nanai; NGD, Negidal; NGS, Nganasan; NVH, Nivh; NGI_A, Nogai_Astrakhan; NGI_K, Nogai_Karachay_Cherkessia; NGI_S, Nogai_Stavropol; ORQ, Oroqen; ADB, Russia_Aldy_Bel; BLS, Russia_Bolshoy; MHE_1, Russia_EarlyMedieval_Heshui_Mohe_1; MHE_2, Russia_EarlyMedieval_Heshui_Mohe_2; SARM_S, Russia_EarlySarmatian_SouthernUrals.SG; KRS_o1, Russia_Karasuk_o1.SG; KRS_o, Russia_Karasuk_oRISE.SG; KRS, Russia_Karasuk.SG; SARM_L, Russia_LateSarmatian.SG; SARM_S.SG, Russia_MiddleSarmatian_SouthernUrals.SG; SARM, Russia_Sarmatian; SARM.SG, Russia_Sarmatian.SG; TGR, Russia_Tagar.SG; SAM.DG, Saami.DG; SKP, Selkup; SHR_K, Shor_Khakassia; SHR_M, Shor_Mountain; TTR_A, Tatar_Astrakhan; TTR_I, Tatar_Irtysh_Barabinsk.SG; TTR_K, Tatar_Kazan; TTR_M, Tatar_Mishar; TTR_S, Tatar_Siberian; TTR_Z, Tatar_Siberian_Zabolotniye; TTR_T, Tatar_Tomsk.SG; TTR_V, Tatar_Volga.SG; TDZ, Todzin; TFL, Tofalar; TBL, Tubalar; TKM, Turkmen; TVN, Tuvinian; UDM, Udmurt; SCY_U, Ukraine_Scythian; ULC, Ulchi; UYG, Uyghur; UZB, Uzbek; VPS, Veps; XIB, Xibo; YKT, Yakut; YKG_F, Yukagir_Forest; YKG_T, Yukagir_Tundra; KNY.SG, Russia_Yenisei_Krasnoyarsk_LBA.SG; KNY_o1.SG, Russia_Yenisei_Krasnoyarsk_LBA_o1.SG.



Extended Data Fig. 12 | f4 statistics of the form f4(*Ethiopia_4500BP.SG, X, Yana.SG, China_Paleolithic*) plotted against f4(*AG3, X, Yakutia_LNBA, East Eurasian Population*). *China_Paleolithic* includes the Tianyuan and Amur_River_33K genomes, "East Eurasian Population" is some population grouping in Siberia or Northeast Asia other than *Yakutia_LNBA*, and X are Admixed Inner Eurasian populations (AIEA populations) including ancient Central Asian nomads from the Late Bronze to Iron Age down to the Scytho-Sarmatian period, as well as modern or ancient populations that speak languages from the Yukaghiric, Yeniseian (Kets), Uralic, Turkic, Mongolic, Tungusic, and Nivkh language families. Modern Uralic-speaking populations, and ancient putatively Uralic-speaking populations uniformly prefer *Yakutia_LNBA* to other East Asian ancestries no matter the other population used in the comparison. Furthermore, at any level of admixture between East and West Eurasian ancestries, the population with the greatest affinity to *Yakutia_LNBA* is always a Uralic-speaking population. f4-statistics therefore highlight the connection between Uralic populations and Yakutia_LNBA ancestry over other sources of East Asian ancestry.



Extended Data Fig. 13 | Characteristic Seima-Turbino artifacts. 1. Double-bladed dagger with a ring-shaped pommel, robbery find, unknown provenance (probable Omsk region or Rostovka). 2. Double-bladed dagger with a horse figurine on the pommel, an accidental find near Shemonaikha, East Kazakhstan. 3., 5., 7. Crook-backed knives with figurines on pommels: 3. from Seyma; 5. from Elunino-1, burial 1, 7. from Rostovka, burial 2. 4. Scapula-shaped celt with goat image, Rostovka, cluster of finds near burial 21. 6. double-bladed plate dagger with a double elk-head figurine pommel, an accidental find near Perm' (probably associated with the Turbino site). 8. Top of staff with a horse figurine, an accidental find near Omsk. 9a. & 9b. Single-ear long spearhead with a relief figurine of a *Felidae* predator (tiger or mountain leopard) on the socket (9a. the spear tip,10b. the detail of the socket), an accidental find near Omsk.

Extended Data Table 1 | Summary of qpAdm analyses

qpAdm analyses	Purpose	Associated Section of Supplementary Material	Associated Figures displaying results	Individuals included in f2-statistic calculation and their population labels for that set of qpAdm	Number of SNPs retained after f2blocks calculation	Supplementary Data file with full results	Wet laboratory data type
qpAdms targeting NEAHG populations (proximal)	Investigating the origin of the ANE ancestry in NEAHG populations	Supplementary Information section 10	Figure 1C, bottom row	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients” Column D	1196712	Supplementary Data 2, in the sheet “Proximal NEAHG qpAdms”	References: 1240K only Sources: 1240K only Targets: 1240K & shotgun
qpAdms targeting NEAHG populations (distal)	Investigating the deep genetic affinities of NEAHG populations	Supplementary Information section 10	Figure 1C, Middle Row	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column D	1135474	Supplementary Data 2, in the sheet “Distal NEAH qpAdms”	References: 1240K only Sources: 1240K only Targets: 1240K and shotgun
qpAdms targeting Seima-Turbino individuals (proximal models)	Investigating the proximal ancestry sources of Seima-Turbino individuals	Supplementary Information section 15	Figure 4B, bottom row	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column O	1135472	Supplementary Data 7, in the sheet “Proximal ST qpAdms”	References: 1240K & shotgun Sources: 1240K & shotgun Targets: 1240K & shotgun
qpAdms targeting Seima-Turbino individuals (distal models)	Investigating the distal ancestry sources of Seima-Turbino individuals	Supplementary Information section 15	Figure 4B, Middle Row	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column N	1196712	Supplementary Data 7, in the sheet “Distal ST qpAdms”	References: 1240K & shotgun Sources: 1240K & shotgun Targets: 1240K & shotgun
qpAdms targeting AIEA populations, and Seima-Turbino individuals (distal models)	Investigating the ancestry sources of AIEA populations (with China_AmurRiver_N and without China_AmurRiver_14K in the sources and references)	Supplementary Information sections 11,13,15	Figure 3C	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column M	593124	Supplementary Data 6, in the paired sheets “AIEA Yakutia_LNBA qpAdms" with passing models in "Results (Yakutia_LNBA)”; and in the paired sheets “AIEA Cisbaikal_LNBA qpAdms” with passing models in “Results (Cisbaikal_LNBA, 1240k)”	References: 1240K & shotgun Sources: 1240K & shotgun Targets: 1240K & shotgun
qpAdms targeting AIEA populations (1240K populations used in sources and references only)	Investigating the ancestry sources of AIEA populations, with only 1240K sequences in both sources and references	Supplementary Information sections 11,13,15	-	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column P	546292	Supplementary Data 6, in the paired sheets “AIEA Yakutia_LNBA qpAdms 1240k" with passing models in "Results (Yakutia_LNBA 1240k)”; and in the paired sheets “AIEA Cisbaikal_LNBA qpAdms1240k” with passing models in “Results (Cisbaikal_LNBA)”	References: 1240K only Sources: 1240K only Targets: 1240K & shotgun
qpAdms targeting populations in 10-member Siberian transect	Investigating the genetic origins of a ten-member transect of Siberian population history	Supplementary Information section 8	Figure 2C	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column Q	1196712	Supplementary Data 6, in the sheet “AIEA Yakutia_LNBA + Amur”	References: 1240K & shotgun Sources: 1240K & shotgun Targets: 1240K & shotgun
qpAdms targeting Bering Straits Populations	Investigating the genetic relationship between Bering Straits populations and the populations in the ten-member transect	Supplementary Information section 9	-	Listed in Supplementary Data 1, in the sheet “Population Labels of Ancients”, Column R	1135467	Supplementary Data 5, Table, in the sheets “Saqqaq qpAdms” and “Beringian qpAdms (NNA + SNA)”	References: 1240K & shotgun Sources: 1240K & shotgun Targets: 1240K & shotgun

This table provides information about the Supplementary Data files that should be referred to in order to understand the details of each qpAdm analysis. There are four major rounds of qpAdm analysis (NEAHG, Seima-Turbino, AIEA populations, and 10-member Siberian transect plus Bering Straits populations). These qpAdm analyses vary in their goals. They may also involve different data types (such as 1240k data only, or mixture of 1240k and shotgun), and their analytic setup.

Corresponding author(s): Tian Chen Zeng, Leonid Vyazov, Alexander Kim, David Reich. 2023-08-15223C

Last updated by author(s): Oct 2, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Software: BWA-v.0.6.1a, contamMix v1.0-10,
Data analysis	HaploGrep2, hapROH, ADMIXTURE v1.3.0, smartpca v18160, PLINK1.9, ADMIXTOOLS v6.0, ADMIXTOOLS2 2.0.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The aligned sequences will be available through the European Nucleotide Archive under a new accession number. Genotype data used in analysis will be available at https://reich.hms.harvard.edu/datasets. Any other relevant data are available from the corresponding authors upon reasonable request.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	We newly genotyped 229 present-day individuals from 10 ethnolinguistic groups using the Affymetrix Human Origins SNP array. All DNA samples were collected with informed consent for broad studies of population history and full public release of de-identified genetic data. All newly reported data are represented either by co-authors of this study or individuals who wished to be mentioned in the Acknowledgments who were involved in sample collection. Sex and gender were not relevant to data collection procedures from present-day individuals. We make a careful distinction between sex and gender in the main text; because ancient gender identities cannot be known with certainty, we only ever refer to bioloigical sex.
Reporting on race, ethnicity, or other socially relevant groupings	We make a careful point to refer to linguistic identities and to use only such terms to refer to populations, so as not to reify ethnic distinctions, while still drawing attention to the correlation that might sometimes obtain between genetic variation and language due to culture-gene co-transmission.
Population characteristics	Samples were collected without regard to phenotypic information and solely based on ethnolinguistic identity.
Recruitment	See above.
Ethics oversight	We are analyzing previously collected DNA samples, which were collected under formal ethical review board supervision at the relevant institutions. Aliquots of these samples, which were de-identified samples so that the researchers involved in the study are not able to connect back to the volunteers who provided the samples, were sent for genotyping for this study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Genetic analyses were performed on DNA data generated from ancient human skeletons and from samples from present-day populations. Population genetic statistics, primarily testing historical relationships by measuring allele-sharing patterns across populations, were computed using genome-wide SNP genotypes.
Research sample	201 samples of skeletal material were screened for ancient DNA data. The resulting genetic sequences from 181 ancient individuals who passed screening and quality controls were co-analyzed with 229 newly-published present-day genetic sequences. These are co-analyzed in turn with 2313 other previously-published present-day samples and 1331 other previously-published ancient samples.
Sampling strategy	We sampled available bones from 201 ancient individuals from across Northern Eurasia and obtained working data from 181. We targeted approximately 1.2 million genome-wide SNPs, which effectively cover almost all independent loci (due to linkage disequilibrium) and provide good power in population history analyses.
Data collection	DNA from the ancient remains was extracted, sequenced, and processed into SNP genotype calls.
Timing and spatial scale	Ancient individuals were sampled from across the forest and forest-steppe zones of Northern Eurasia. Ancient individuals lived between ~10,000-3500 calibrated years before the present.
Data exclusions	20 of the sampled skeletons did not yield working data as assessed by pre-established ancient DNA quality criteria.
Reproducibility	All attempts to reproduce were successful.
Randomization	Samples were grouped based on a multi-step process. Samples were first divided into forest-steppe-hunter-gatherer and non-forest-steppe-hunter-gatherer populations based on unsupervised analyses in ADMIXTURE and PCA. Non-forest-steppe hunter-gatherer populations where then grouped using f4-statistics, while forest-steppe hunter-gatherer populations were grouped using ADMIXTURE and PCA and time-stratified according to C14 dates, where available.
Blinding	Analyses were performed either for all individuals separately, or for groupings of individuals produced by f4-statistics or based on the results of unsupervised analyses and C14 dates ('clusters'); other sample-specific features were not relevant to results.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Palaeontology and Archaeology

Specimen provenance	Samples analyzed in this study came from museum collections stewarded by our co-authors. Descriptions of the archaeological and cultural contexts for all ancient samples analyzed in this study, including their grave position within archaeological sites, their grave numbers and burial inventory, archaeological publications describing the sites themselves (where available), are provided in Supplementary information section I. For specimen source and deposition details, see below.
Specimen deposition	The skeletal samples are under the stewardship of the co-authors or museum collections that are listed in SI Data Table 2, column G. Samples may be accessed by their skeletal code listed in Supplementary Data S1 Table 2, along with the contact details of the persons stewarding the samples in our author list.
Dating methods	We report 76 new radiocarbon dates using standard methods from the Pennsylvania State University Radiocarbon Laboratory.

☒ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	Every sample is represented by a co-author who was involved in sample collection with permission from all relevant local authorities and institutions.

Note that full information on the approval of the study protocol must also be provided in the manuscript.