

## ANTHROPOLOGY

## Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads

Maja Krzewińska<sup>1\*†</sup>, Gülşah Merve Kılınc<sup>1\*†</sup>, Anna Juras<sup>2</sup>, Dilek Koptekin<sup>3</sup>, Maciej Chyleński<sup>4</sup>, Alexey G. Nikitin<sup>5</sup>, Nikolai Shcherbakov<sup>6</sup>, Iia Shuteleva<sup>6,7</sup>, Tatiana Leonova<sup>6</sup>, Liudmila Kraeva<sup>8</sup>, Flarit A. Sungatov<sup>9</sup>, Alfija N. Sultanova<sup>9</sup>, Inna Potekhina<sup>10</sup>, Sylwia Łukasik<sup>2</sup>, Marta Krenz-Niedbała<sup>2</sup>, Love Dalén<sup>11</sup>, Vitaly Sinika<sup>12,13</sup>, Mattias Jakobsson<sup>14,15,16</sup>, Jan Storå<sup>17</sup>, Anders Götherström<sup>1†</sup>

For millennia, the Pontic-Caspian steppe was a connector between the Eurasian steppe and Europe. In this scene, multidirectional and sequential movements of different populations may have occurred, including those of the Eurasian steppe nomads. We sequenced 35 genomes (low to medium coverage) of Bronze Age individuals (Srubnaya-Alakulskaya) and Iron Age nomads (Cimmerians, Scythians, and Sarmatians) that represent four distinct cultural entities corresponding to the chronological sequence of cultural complexes in the region. Our results suggest that, despite genetic links among these peoples, no group can be considered a direct ancestor of the subsequent group. The nomadic populations were heterogeneous and carried genetic affinities with populations from several other regions including the Far East and the southern Urals. We found evidence of a stable shared genetic signature, making the eastern Pontic-Caspian steppe a likely source of western nomadic groups.

## INTRODUCTION

The Pontic-Caspian steppe (PCS), stretching from the southern Urals to the western North Pontic lands, was the stage of various demographic changes in the past, and several of those remain unknown. During the Bronze and Iron Age, the area was inhabited by a succession of nomadic populations that had significant impact on the cultural development of both Asia and Europe (1, 2). Possibly the best known of these groups is the Yamnaya. Recent genomic studies have revealed cross-continental Early Bronze Age migrations (~3000 BCE) of the nomadic people associated with the Yamnaya horizon (3, 4). The migration introduced the Caucasus genetic component to the genetic landscape of Europe. In Central Europe, Yamnaya ancestry first appeared among people from the Corded Ware complex and has since been found in many subsequent an-

cient and present-day populations. However, the Pontic-Caspian steppe was critical not only for Early Bronze Age Yamnaya migrations but also because of succeeding movements and population transformations that took place in the developed classical stage of the Late Bronze and Iron Ages between 1800 BCE and 400 CE. This period covered the development of the Srubnaya and Alakulskaya Cultures (~1800–1200 BCE), associated with small settlement sites distributed from the Urals to the Dnieper valley (1). From around 1000 BCE, pre-Scythian nomadic populations started to appear in the western Pontic-Caspian steppe including the Cimmerians known from historical sources (5). Despite regional variation and local peculiarities, the Cimmerians were not associated with any uniform type of archaeological material culture (6). In the seventh century BCE, they were succeeded by the Scythians, who plausibly pushed the Cimmerians into Asia Minor (7). Between 700 and 300 BCE, the Scythians, representing mobile pastoral nomads of a new militaristic type (1), dominated the Pontic-Kazakh steppe, occupying an area from the Altai to the Carpathian Mountains. Their decline began around 300 BCE and was caused by intensifying hostile relations with the Macedonians in the West and the invasion of the Sarmatians from the East. The Sarmatians and the Scythians are thought to have coexisted for a few centuries, but eventually, the former group prevailed (2), resulting in the Scythian downfall. The Sarmatians are believed to comprise a number of groups of similar nomadic background (8), and they became the politically most influential force within the eastern fringes of the Roman Empire at the time. Their decline (~400 CE) was associated with the attack of the Goths and the subsequent invasion of the Huns (8).

The genomic structure of the Bronze and Iron Age (1800 BCE–400 CE) populations in the Pontic-Caspian steppe has not been fully resolved. While earlier genomic studies have suggested close links between the Srubnaya and the central European Late Neolithic and Bronze Age populations (9), our knowledge of the genetic origins of the Cimmerians is limited. Genetic analyses of maternal lineages of Scythians suggest a mixed origin and an east-west admixture gradient across the Eurasian steppe (10–12). The genomics of two early Scythian Aldy-Bel individuals (13) showed genetic affinities to eastern

<sup>1</sup>Archaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, Lilla Frescativägen 7, SE-106 91 Stockholm, Sweden.

<sup>2</sup>Department of Human Evolutionary Biology, Institute of Anthropology, Faculty of Biology, Adam Mickiewicz University in Poznań, Umultowska 89, 61-614 Poznań, Poland. <sup>3</sup>Department of Health Informatics, Middle East Technical University, 06800 Ankara, Turkey. <sup>4</sup>Institute of Archaeology, Faculty of History, Adam Mickiewicz University in Poznań, Umultowska 89D, 61-614 Poznań, Poland. <sup>5</sup>Biology Department, Grand Valley State University, 1 Campus Drive, Allendale, MI 49401, USA. <sup>6</sup>Laboratory of Methodology and Methods of Humanitarian Research, Bashkir State Pedagogical University, Otyabrskoy Revolutsiy, 3A, 450007 Ufa, Russia. <sup>7</sup>Institute of History and State Management, Bashkir State University, Zaki Validy Street 32, 450076 Ufa, Russia. <sup>8</sup>Archaeological Laboratory, Orenburg State Pedagogical University, Orenburg, Russia. <sup>9</sup>Archaeological Laboratory, Bashkir State University, Str. Validi Z. 32, Ufa, Russia.

<sup>10</sup>Institute of Archaeology, National Academy of Sciences of Ukraine, Kiev, Ukraine. <sup>11</sup>Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Post Office Box 50007, 10405 Stockholm, Sweden. <sup>12</sup>Taras Shevchenko University in Tiraspol, October Street 25, 33-00 Tiraspol, Moldova. <sup>13</sup>Nizhnevartovsk State University, Lenin Street, 56, Nizhnevartovsk, 628605, Khanty-Mansi Autonomous District, Yugra, Russia. <sup>14</sup>Subdepartment of Human Evolution, Department of Organismal Biology, Uppsala University, Norbyvägen 18 C, SE-752 36 Uppsala, Sweden. <sup>15</sup>Science for Life Laboratory, Norbyvägen 18 C, SE-752 36 Uppsala, Sweden. <sup>16</sup>Centre for Anthropological Research and Department of Anthropology and Development Studies, University of Johannesburg, Post Office Box 524, Auckland Park 2006, South Africa. <sup>17</sup>Osteoarchaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, Lilla Frescativägen 7, SE-106 91 Stockholm, Sweden.

\*These authors contributed equally to this work.

†Corresponding author. Email: maja.krzewinska@arklab.su.se (M.K.); gulsahhdal@gmail.com (G.M.K.); anders.gotherstrom@arklab.su.se (A.G.)

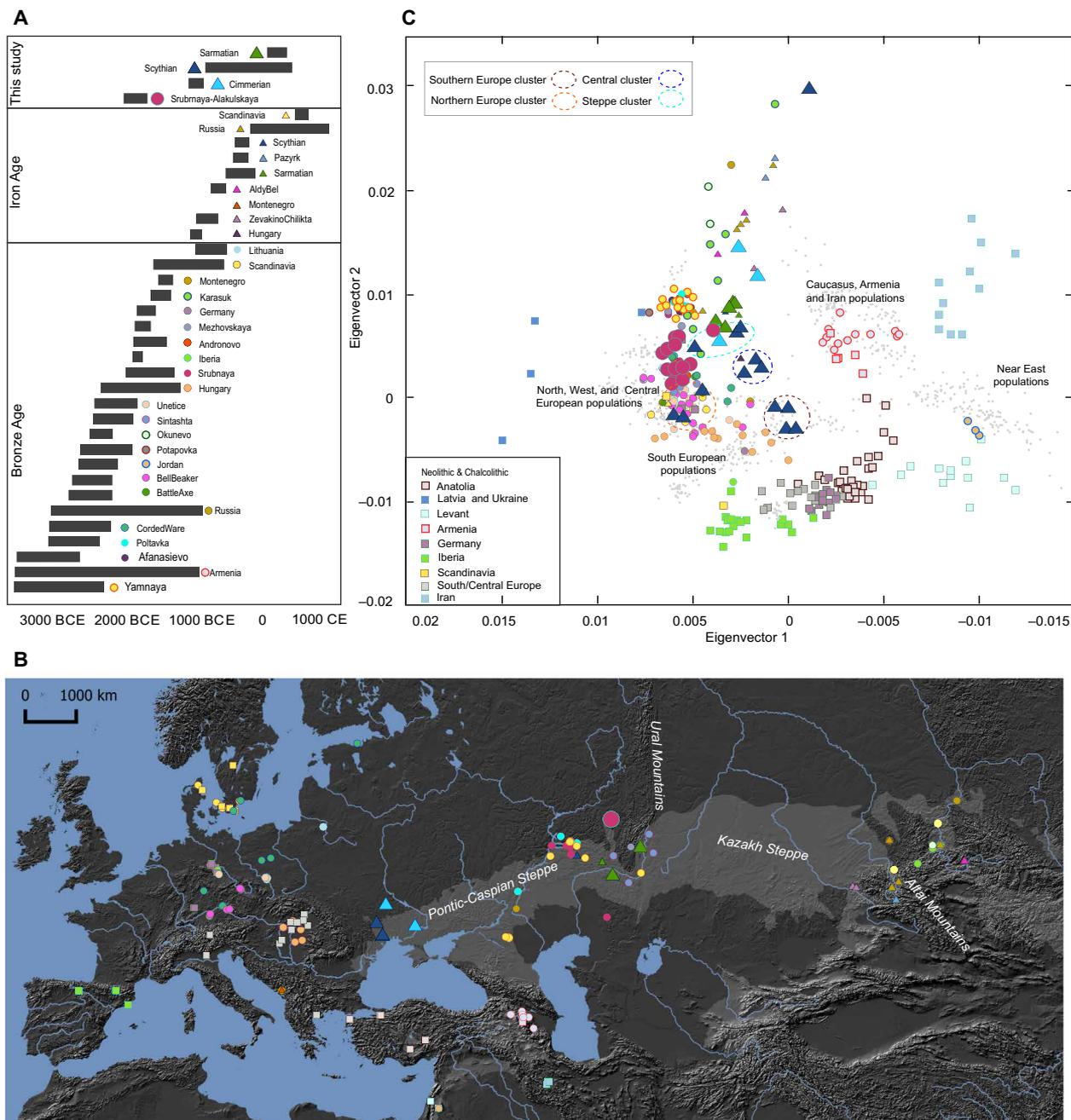
Copyright © 2018  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

Downloaded from <http://advances.sciencemag.org/> on October 10, 2020

populations of Central Asia (12). However, population interactions and the origin of Scythians of the Pontic-Caspian steppe remain poorly understood. Similarly, little is known about the origins and genetic affinities of the Sarmatians. Genomic studies suggest that the latter group may have been genetically similar to the eastern Yamnaya and Poltavka Bronze Age groups (12). To investigate the demographic dynamics in the Pontic-Caspian steppe, we generated and analyzed genomes of the Late Bronze and Iron Age individuals from the region (Fig. 1, A and B).

## RESULTS

We produced genome-wide sequence data with genome coverage between  $0.01\times$  and  $2.9\times$  per individual for 35 Bronze Age and Iron Age individuals from the Pontic-Caspian steppe from four chronologically sequential cultural groups, which comprise Srubnaya-Alakulskaya individuals ( $n = 13$ ), Cimmerians ( $n = 3$ ), Scythians ( $n = 14$ ), and Sarmatians ( $n = 5$ ), with radiocarbon dates between ca. 1900 BCE and 400 CE (Fig. 1, A and B; tables S1 to S3; and fig. S1, A and B). All DNA libraries displayed damage patterns typical of ancient



**Fig. 1. Radiocarbon ages and geographical locations of the ancient samples used in this study.** Figure panels presented counterclockwise: **(A)** Bar plot visualizing approximate timeline of presented and previously published individuals. **(B)** Map showing the locations of ancient individuals sequenced in this study and the locations of previously published ancient individuals used in comparative analyses. **(C)** Principal component analysis (PCA) plot visualizing 35 Bronze Age and Iron Age individuals presented in this study and in published ancient individuals (table S5) in relation to modern reference panel from the Human Origins data set (41).

DNA (fig. S2) (14). To ensure data integrity, we calculated mitochondrial DNA (mtDNA)-based contamination levels using distribution of private polymorphisms in mtDNA (15) and a Bayesian likelihood method (16). The former yielded point estimates of contamination between 0 and 10% [95% confidence intervals (CIs) between 0 and 17%], and the latter method revealed that all individuals carried sequences with >89% probability of being authentic (table S4). Thus, we included all sequenced individuals in the downstream analyses.

Late Bronze Age (LBA) Srubnaya-Alakulskaya individuals carried mtDNA haplogroups associated with Europeans or West Eurasians (17) including H, J1, K1, T2, U2, U4, and U5 (table S3). In contrast, the Iron Age nomads (Cimmerians, Scythians, and Sarmatians) additionally carried mtDNA haplogroups associated with Central Asia and the Far East (A, C, D, and M) (table S3) (11, 18). The absence of East Asian mitochondrial lineages in the more eastern and older Srubnaya-Alakulskaya population suggests that the appearance of East Asian haplogroups in the steppe populations might be associated with the Iron Age nomads, starting with the Cimmerians. The Y chromosome haplogroup variation in 17 of 18 males was limited to two major haplogroup lineages within the macrohaplogroup “R” (table S3). The Srubnaya-Alakulskaya individuals carried the Y haplogroup R1a, which showed a major expansion during the Bronze Age (19). It has previously been found in Bronze Age individuals from the Krasnoyarsk Kurgan in Siberia (20). The Iron Age nomads mostly carried the R1b Y haplogroup, which is characteristic of the Yamnaya of the Russian steppe (4). An exception was a Cimmerian individual (cim358) who carried the Q1\* lineage associated with the east (table S3).

### Genetic relationships between Eurasian steppe nomads and present-day populations

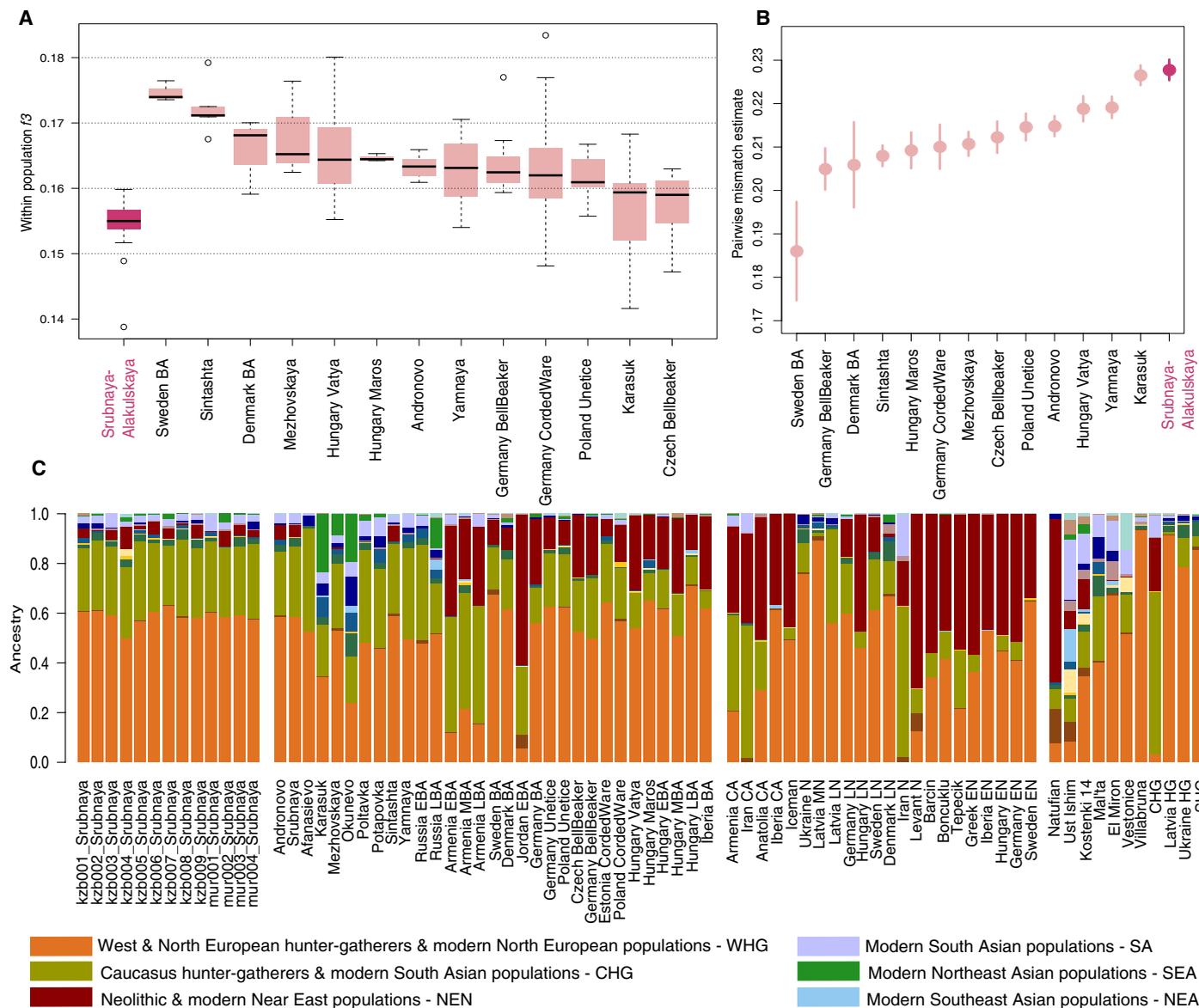
PCA on the autosomal genomic data (Fig. 1C and table S5) revealed the following: (i) Srubnaya-Alakulskaya individuals exhibited genetic affinity to northern and northeastern present-day Europeans (fig. S3), and these results were also consistent with outgroup  $f_3$  statistics (table S6 and fig. S4A). (ii) The Cimmerian individuals, representing the time period of transition from Bronze to Iron Age, were not homogeneous regarding their genetic similarities to present-day populations according to the PCA.  $F_3$  statistics confirmed the heterogeneity of these individuals in comparison with present-day populations (table S6 and figs. S3 and S4C). (iii) The Scythians reported in this study, from the core Scythian territory in the North Pontic steppe (12), showed high intragroup diversity. In the PCA, they are positioned as four visually distinct groups compared to the gradient of present-day populations (Fig. 1C): (i) A group of three individuals (scy009, scy010, and scy303) showed genetic affinity to north European populations, hereafter referred to as a north European (NE) cluster. (ii) A group of four individuals (scy192, scy197, scy300, and scy305) showed genetic similarities to southern European populations, hereafter referred to as a south European (SE) cluster. (iii) A group of three individuals (scy006, scy011, and scy193) located between the genetic variation of Mordovians and populations of the North Caucasus, hereafter referred to as a steppe cluster (SC). In addition, one Srubnaya-Alakulskaya individual (kzb004), the most recent Cimmerian (cim357), and all Sarmatians fell within this cluster. In contrast to the Scythians, and despite being from opposite ends of the Pontic-Caspian steppe, the five Sarmatians grouped close together in this cluster. (iv) A group of three Scythians (scy301, scy304, and

scy311) formed a discrete group between the SC and SE and had genetic affinities to present-day Bulgarian, Greek, Croatian, and Turkish populations, hereafter referred to as a central cluster (CC). All PCA results were consistent with outgroup  $f_3$  statistics (table S6 and figs. S3 and S4, B and D). Finally, one individual from a Scythian cultural context (scy332) is positioned outside of the modern West Eurasian genetic variation (Fig. 1C) but shared genetic drift with East Asian populations (table S6 and fig. S4B).

### Genetic relationships between Srubnaya-Alakulskaya and other ancient populations

The Bronze Age in Eurasia was a dynamic period with several human groups participating in different migratory processes. As a result, there were extensive interactions between European Bronze Age populations (3). The Srubnaya-Alakulskaya individuals, originating from a site of cultural dualism in the forest steppe of the Trans-Volga region, were genetically similar to the previously published Srubnaya and Andronovo individuals from the Pontic-Kazakh steppe (3, 9) and to the European Bronze Age groups, including individuals of the Corded Ware, Unetice, and Bell Beaker complexes (Fig. 1C and fig. S5). Consistent with other Bronze Age populations, the Srubnaya-Alakulskaya individuals were positioned between the genetic variation of the European Mesolithic and the Near East Neolithic populations, being closer to the former and especially to the east European hunter-gatherers (Fig. 1C and figs. S6 and S8). These individuals had higher genetic affinity to Scythians compared to other Iron Age groups (fig. S9). To test whether our Bronze Age sample from a cultural mixing zone between the Srubnaya and the Alakulskaya complexes shared more genetic drift with previously published representatives of Srubnaya, we calculated  $f_4$  statistics of  $f_4(\text{Yoruba}, \text{SrubnayaX}, \text{SrubnayaY}, \text{BronzeX})$ , where “BronzeX” refers to a Bronze Age Russian population. The test revealed that the Srubnaya-Alakulskaya population formed a clade together with Andronovo, Afanasievo, and Sintashta to the exclusion of the previously published data of other Srubnaya individuals (tables S7 and S8). Furthermore, to trace the plausible origin the Caucasus genetic component identified in Srubnaya-Alakulskaya individuals, we adopt the  $f_4$  statistics in the form of  $f_4(\text{Yoruba}, \text{Srubnaya-Alakulskaya}, \text{PopulationX}, \text{Yamnaya})$ , where *PopulationX* was one of the Eurasian Bronze Age populations. The results showed that Srubnaya-Alakulskaya formed a clade together with Yamnaya to the exclusion of other Bronze Age populations from Russia, Armenia, Jordan, and Hungary. This finding indicates that the Caucasus genetic contribution to the Srubnaya-Alakulskaya individuals was mediated by steppe ancestry instead of originating from the Levant (table S9). Both mean  $f_3$  statistics within populations and conditional nucleotide diversity (21) revealed that the genetic diversity was highest in the LBA Srubnaya-Alakulskaya population from the southern Ural region compared to all other Eurasian Bronze Age populations (Fig. 2, A and B, and table S10).

To evaluate the ancestral components of Srubnaya-Alakulskaya, we conducted ADMIXTURE analysis for  $K = 2$  to  $K = 15$  ancestral clusters (fig. S10A).  $K = 15$  clusters revealed that Srubnaya-Alakulskaya individuals consisted of two major ancestral components; first, an “orange” component predominantly found in west and NE hunter-gatherers (WHG) and in present-day NE populations, and second, a “light green” component typical of Caucasus hunter-gatherers (CHG) found in south Asian (SA) modern populations. The component associated with Neolithic groups and present-day Near



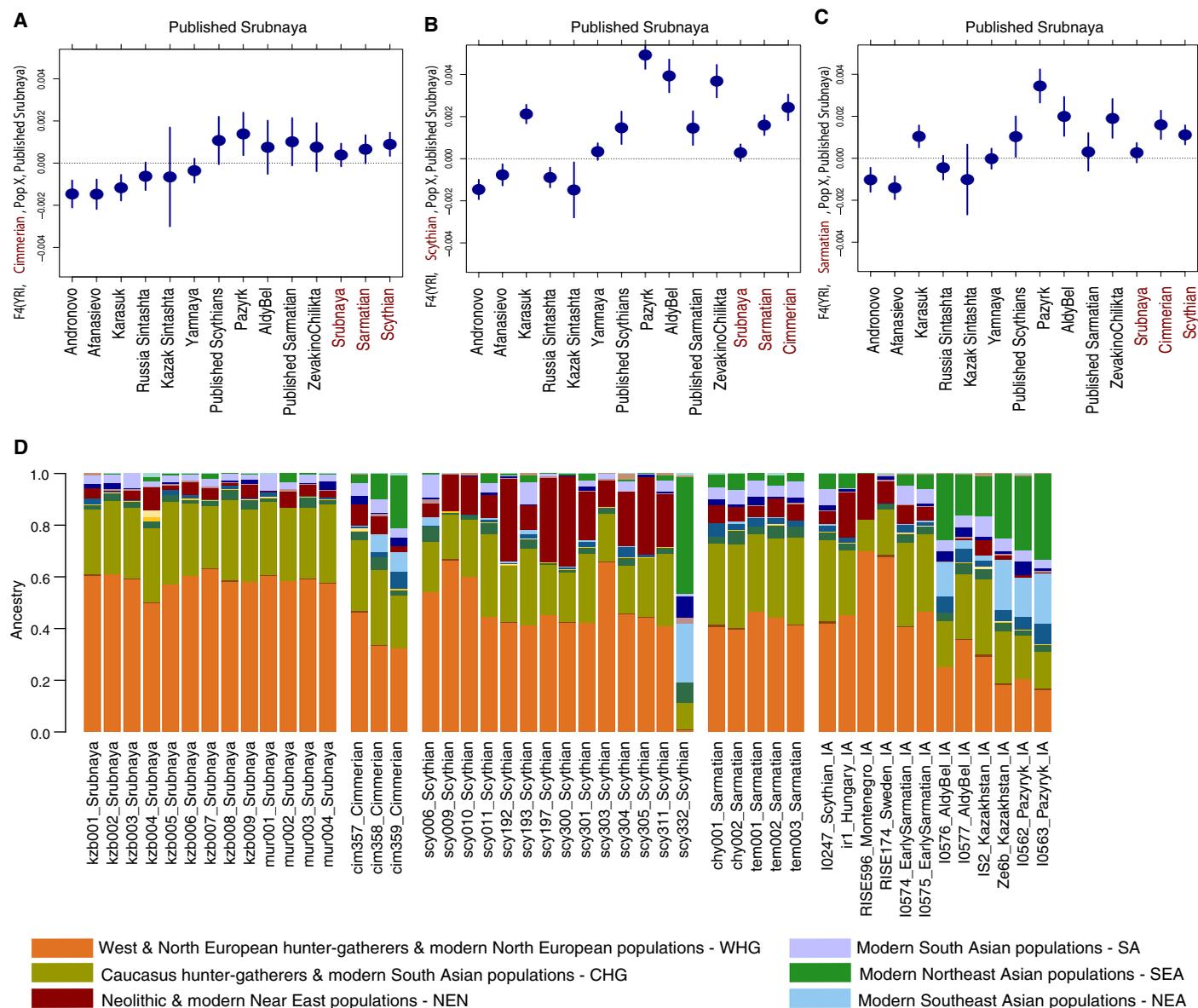
**Fig. 2. Genetic diversity and ancestral components of Srubnaya-Alakulskaya population.** Diversity and ancestral components of Srubnaya-Alakulskaya population (here called “Srubnaya”): (A) Mean  $f_3$  statistics for Srubnaya and other Bronze Age populations. Srubnaya group was color-coded the same as with PCA. (B) Pairwise mismatch estimates for Bronze Age populations. (C) ADMIXTURE results for  $K = 15$ .  $K = 15$  was chosen to display since it shows SA component (lilac) and Northeast Asian (NEA, “dark green”) components in addition to the other components. BA, Bronze Age; EBA, Early Bronze Age; MBA, Middle Bronze Age; CA, Chalcolithic; N, Neolithic; EN, Early Neolithic; LN, Late Neolithic; HG, hunter-gatherers; SHG, Scandinavian hunter-gatherers (fig. S10).

Eastern populations (NEN, “dark red”) contributed less to our Srubnaya-Alakulskaya individuals compared to the European Bronze Age populations (Fig. 2C and fig. S10B).

**Genetic relationships between the Iron Age nomads and other ancient populations**

In the 10th century BCE, during the transition from the Bronze Age to the Iron Age, the Cimmerians appear in the Pontic-Caspian steppe. In the PCA, the chronologically youngest Cimmerian individual (cim357) grouped within the SC including all Sarmatians, one Srubnaya-Alakulskaya individual, and three Scythian individuals. The other Cimmerian individuals were positioned in close proximity to a number of eastern Iron Age individuals from the Altai region (3), as well as

Aldy-Bel and Zevakino-Chillikta individuals of the Altai and Western Mongolia regions (Fig. 1C and fig. S9) (12). We tested whether Cimmerians formed a clade with the subsequent Iron Age populations including Scythians and Sarmatians, to the exclusion of Bronze Age populations by calculating  $f_4$  statistics in the form of  $f_4(\text{Yoruba}, \text{Cimmerians}; \text{Srubnaya} \& \text{Scythians} \& \text{Sarmatians}, \text{BronzeAgePopulation})$ , showing that the Cimmerians shared more drift with the Bronze Age populations of Russia compared to the Srubnaya/-Alakulskaya but not compared to Scythians or Sarmatians (Fig. 3A and tables S11, S12, and S16). The Cimmerians shared more drift with the far eastern Karasuk population compared to the geographically more closely located Srubnaya/-Alakulskaya population (table S16), corroborating the existence of the “Karasuk-Cimmerian cultural-historical



**Fig. 3. Genetic relationship between Srubnaya-Alakulskaia population and Iron Age nomads.** Visual summary of  $f_4$  statistics of a form  $f_4(YRI, TestPop, Pop X, Srubnaya)$ , where *TestPop* is (A) Cimmerian; (B) Scythian and (C) Sarmatian. (D) Admixture selection: Our ancient individuals and Iron Age individuals from closely related populations ( $K = 15$ ).  $K = 15$  was chosen to display since it shows SA (lilac) and NEA (dark green) components in addition to the other components. Srubnaya-Alakulskaia population is named as Srubnaya in the panels. IA, Iron Age (fig. S10).

community” (22). Pairwise mismatch estimates revealed a slightly higher genetic diversity in Cimmerians compared to that of the Srubnaya-Alakulskaia population (table S10). ADMIXTURE analysis ( $K = 15$ ) revealed that Cimmerian individuals carried predominantly the WHG and CHG components similar to the Srubnaya individuals, but they also carried other ancestral components including NEN (dark red), Northeast Asian (NEA) (“dark green”), and Southeast Asian (SEA, “light blue”) components. From the oldest to the most recent sample, the amount of NEN component increased, while the NEA and SEA components decreased through time in the Cimmerians (Fig. 3D and fig. S10). The admixture results were further tested with  $f_3$  statistics, which were consistent with the observed pattern by means of shared genetic drift between the Cimmerians

and the tested populations used as proxies for different genetic components identified in admixture analyses (table S13). These results implicate a more eastern origin of the Cimmerians compared to that of the Srubnaya-Alakulskaia population and an increased amount of NEN-originated admixture in these individuals through time.

In the seventh century BCE, the Scythians appeared in the North Pontic region. Iron Age western Scythians displayed slightly higher intragroup diversity compared to that of the Bronze Age groups and formed four discrete clusters including NE, SE, SC, and CC clusters (Fig. 1C and table S10). Scythians belonging to the SE cluster were closer to Hungarian Bronze Age and Iron Age individuals including Vatya and Maros. The NE Scythian cluster fell close to the Iron Age individuals from modern-day Montenegro and Sweden (fig. S9).

The SC Scythians further grouped with Early Sarmatians (12) and the Iron Age Scythian from modern-day Hungary (23). It has been hypothesized by Terenozhkin that Scythians reached the Pontic-Caspian steppe region from Central Asia (for example, from Andronovo) (7). To formally test this hypothesis, we calculated  $f_4$  statistics in the form of  $f_4(\text{Yoruba}, \text{Scythian}, \text{Srubnaya}, \text{BronzeAgeX})$ , where “BronzeAgeX” refers to either Andronovo, Sintashta, or Afanasievo. This analysis revealed that the western Scythians (tested either as a single population or four different clusters) formed a clade together with Andronovo and two other Andronovo-originated populations to the exclusion of Srubnaya-Alakulskaya from the southern Urals (Fig. 3B and table S14), supporting the Central Asian origin of the western Scythians. Furthermore, the western Scythians shared more drift with Andronovo, Afanasievo, Sintashta, and Mezhovskaya to the exclusion of Yamnaya (table S15). While the eastern Scythians from an earlier study (12) formed a clade together with Srubnaya-Alakulskaya to the exclusion of Yamnaya, the western Scythians of the present study did not show this pattern. Only when compared with Karasuk and Okunevo, Yamnaya seemed closer (although scy332 is an exception), which is in line with a shared steppe origin of the Yamnaya and the western nomads (table S15). ADMIXTURE analysis at  $K = 15$  revealed that the Scythians belonging to the NE cluster had a more visible West Eurasian component that is highest in the northern populations, while Scythians that belong to the SE cluster had a more visible NEN genetic component, most commonly found in Neolithic Anatolian individuals. The NEN component was also found in individuals falling into the CC, who additionally carried an SA component (“lilac”) shared with the Cimmerians, Sarmatians, and two more Scythian individuals (scy193 and scy006). Scy332 had a SEA (“light blue”) component and was most similar to the Karasuk individuals (Fig. 3 and fig. S10, A and B).

The five Sarmatians grouped close together in the SC (Fig. 1C), but they also had the highest pairwise mismatch estimates compared to other Iron Age nomads, suggesting a larger effective population size (table S10).  $F_4(\text{Yoruba}, \text{Sarmatians}, \text{PopulationX}, \text{Cimmerian})$ , where the *PopulationX* is Yamnaya, Srubnaya, or Karasuk, revealed that Sarmatians formed a clade together with all these three populations to the exclusion of Cimmerians (table S18). Multiway  $f_4$  statistics testing the relationships between Srubnaya/-Alakulskaya, Cimmerians, and Scythians revealed that both Scythians and Cimmerians formed a clade together with Srubnaya/-Alakulskaya to the exclusion of the other population and that Srubnaya/-Alakulskaya was closer to Scythians among the two (Fig. 3C and table S17). The results point to the presence of a deep shared ancestry of all Iron Age nomadic groups associated with Bronze Age populations of the steppe, which, however, is not equivalent with a direct genetic continuity between Srubnaya-Alakulskaya and the western Scythians.

## DISCUSSION

### The origins of the four steppe populations

The Bronze Age Srubnaya-Alakulskaya individuals from Kazburun 1/Muradym 8 presented genetic similarities to the previously published Srubnaya individuals. However, in  $f_4$  statistics, they shared more drift with representatives of the Andronovo and Afanasievo populations compared to the published Srubnaya individuals. Those apparently West Eurasian people lacked significant Siberian components (NEA and SEA) in ADMIXTURE analyses but carried traces of the SA component that could represent an earlier connection to

ancient Bactria. The presence of an SA component (as well as finding of metals imported from Tien Shan Mountains in Muradym 8) could therefore reflect a connection to the complex networks of the nomadic transmigration patterns characteristic of seasonal steppe population movements [see (2), figure 6.1, p. 205]. These movements, although dictated by the needs of the nomads and their animals, shaped the economic and social networks linking the outskirts of the steppe and facilitated the flow of goods between settled, semi-nomadic, and nomadic peoples. In contrast, all Cimmerians carried the Siberian genetic component. Both the PCA and  $f_4$  statistics supported their closer affinities to the Bronze Age western Siberian populations (including Karasuk) than to Srubnaya. It is noteworthy that the oldest of the Cimmerians studied here (cim357) carried almost equal proportions of Asian and West Eurasian components, resembling the Pazyryks, Aldy-Bel, and Iron Age individuals from Russia and Kazakhstan (12). The second oldest Cimmerian (cim358) was also the only one with both uniparental markers pointing toward East Asia. The Q1\* Y chromosome sublineage of Q-M242 is widespread among Asians and Native Americans and is thought to have originated in the Altai Mountains (24). It has previously been identified in numerous ancient samples from Siberia, the Americas, and in representatives of the Siberian Bronze Age and nomadic populations (4, 24). This is the first indication that Cimmerians did not originate in the PCS region but were nomads tracing their origin to the Far East.

The origins of Scythians of the western Pontic-Caspian steppe are difficult to resolve. We identified four different clusters within our geographically continuous sample set, which likely represent a varying gradient of different genetic components: the Northern cluster, SC, CC, and SE cluster. The latter was characterized by the presence of the NEN component representing local semi-nomadic Scythians with clear genetic uptake from the locals and possibly from other settlers such as the Greeks around the Black Sea region. Finally, the Sarmatians fell between all other nomads that form the bulk of the SC, suggesting that southern Urals is where the continuity of western nomads was sustained.

### Intragroup genetic diversity of the steppe populations

The diversity of Srubnaya-Alakulskaya individuals was on par with that of Sarmatians and Scythians (table S10) and was also the highest among all published Bronze Age individuals. The Muradym 8/Kazburun 1 site is a unique cultural mixing zone with an unusual number of culturally distinct burials of two different traditions (Srubnaya-Alakulskaya), but the individuals are genetically uniform across the time span of the sites, suggesting that diffusion may have been the main mode of cultural dispersal in the LBA.

Cimmerians resemble individuals within the SC and had higher diversity compared to that of the Srubnaya-Alakulskaya individuals. Despite clustering tightly on the PCA plot, the Cimmerians and Sarmatians had the highest pairwise mismatch estimates among the nomads studied (Fig. 1C and table S10). Although the Scythians showed high variation based on the PCA, their intragroup genetic diversity was comparatively lower than that of the Cimmerians and Sarmatians. This observation is best explained by a smaller ancestral Scythian population base in western Pontic-Caspian steppe than in Sarmatians of the southern Urals. When compared with published Early Sarmatians from the same region (12), the Sarmatians studied here seem to have remained genetically uniform through an approximately 300- to 500-year time span. This might

suggest the genetic continuity of Sarmatians in the southern Urals despite a distinct cultural shift and a suspected earlier population replacement between Early and Middle/Late Sarmatians in the region. Thus, the observed high genetic diversity in the Sarmatians could result from a large effective population size rather than gene flow into the region.

### Mutual relations and shared ancestry between steppe populations

Our results suggest that the Cimmerians were largely similar to the more eastern Sarmatians (cim357) albeit with an increased amount of a Siberian (NEA) component (cim358 and cim359), thus representing a genetic link between the Iron Age people of the Kazakh steppe region. Similar to the cis-Uralic Srubnaya-Alakulskaya, Cimmerians were not direct ancestors of the Scythians in the Pontic-Caspian steppe; however, all those populations shared a common ancestral gene pool.

Scythians who were thought to have displaced the Cimmerians from the Pontic-Caspian steppe region differed from both Cimmerians and the later Sarmatians. However, it is unclear whether the observed pattern resulted from replacement, since the genetic composition of Cimmerians may have undergone a temporal change, as witnessed by the observed temporal reduction of the NEA and the SEA components in more recent individuals. It has previously been noted that Scythians were not uniform across Eurasia, exhibiting distinct differences between Eastern/Western Eurasian groups (12). In contrast to the eastern steppe Scythians (Pazyryks and Aldy-Bel) that were closely related to Yamnaya, the western North Pontic Scythians were instead more closely related to individuals from Afanasievo and Andronovo groups. Some of the Scythians of the western Pontic-Caspian steppe lacked the SA and the East Eurasian components altogether and instead were more similar to a Montenegro Iron Age individual (3), possibly indicating assimilation of the earlier local groups by the Scythians. Finally, one individual from Nesterivka (scy011), which was the most recent of the Scythians, and possibly representing a transition between different groups, carried a distinct resemblance to the later steppe Sarmatians, corroborating the idea of a more “stable”/“common” steppe signature shared with all Sarmatians or reflecting physical/social connection to the southern Ural steppe zone maintained within the western nomadic world. Toward the end of the Scythian period (fourth century CE), a possible direct influx from the southern Ural steppe zone took place, as indicated by scy332. However, it is possible that this individual might have originated in a different nomadic group despite being found in a Scythian cultural context. This individual instead resembles, genetically, the early eastern Altai Scythians, albeit with even higher contribution of the NEA and the SEA genetic components. This result suggests the presence of a continuous connection between the Western fringes of the Scythian empire and central parts of Eurasia or maybe even the source region in the Altai Mountains. Such a connection could reflect communication associated with population exchange and internal high mobility among the Scythians.

In conclusion, our genomic analyses revealed that both the Bronze Age and Iron Age were highly dynamic periods in the Pontic-Caspian steppe. The time span between 1800 BCE and 400 CE was characterized by mobility, population movements, and replacements, which shaped the complex demography of the region through time. Our results showed that the Western Eurasian steppe nomads were not direct descendants of the Bronze Age Srubnaya-Alakulskaya individuals but shared elements of common ancestry with contribution

from different peoples. The early nomads could thus be referred to as a “cultural and chronological horizon” represented by various cultures of the Scythian-Siberian world that was not composed of a genetically homogeneous and/or isolated group. Quite the contrary is observed. We observe little evidence of mobility from the Far East, suggesting that the main source of most Western nomads is likely found in eastern Pontic-Caspian steppe and southern Urals. Thus, we propose that the region, similar to the so-called Mongolian steppe generator of peoples during the Middle Ages, served as the generator of the west nomadic peoples that sustained the western nomadic horizon in the Iron Age.

## MATERIALS AND METHODS

### Ancient DNA extraction

DNA extraction of the Srubnaya ( $n = 13$ ), Sarmatian ( $n = 5$ ), and four Scythian individuals was done at the Archaeological Research Laboratory at Stockholm University, specially equipped for ancient DNA work. The remaining Scythians from Moldova ( $n = 10$ ) and Cimmerians ( $n = 3$ ) underwent DNA extraction in ancient DNA facilities at the Adam Mickiewicz University in Poland. The surface of the material was cleaned mechanically and was further subjected to ultraviolet light ( $265 \text{ nm J/cm}^2$ ) on each side before sampling. The teeth were additionally pretreated with bleach solution (0.5 to 1% NaOCl) and purified double-distilled  $\text{H}_2\text{O}$  (dd $\text{H}_2\text{O}$ ). The amount of bone powder used varied between 80 and 150 mg. DNA extraction and library preparation was performed in a separate ancient DNA laboratory in laminar flow hoods. The extraction was done following a silica column-based extraction procedure (25) with adaptations (26). The extraction buffer that we used contained 1000  $\mu\text{l}$  of 0.5 M EDTA (pH 8.0), 1 M urea, and 10  $\mu\text{l}$  of proteinase K (10 mg/ml). After digestion, 1 ml of extract was concentrated to 100  $\mu\text{l}$  using Amicon Ultra centrifugal filters and purified to 110  $\mu\text{l}$  of cleaned product using a MinElute kit (Qiagen).

### Library preparation and sequencing

Blunt-end, double-stranded libraries were built from 20  $\mu\text{l}$  of DNA extract following Meyer and Kircher (27). The libraries were assessed using real-time quantitative polymerase chain reaction (qPCR). With 1  $\mu\text{l}$  of DNA library, 12.5  $\mu\text{l}$  of 1 $\times$  Maxima SYBR Green Master Mix, 10.5  $\mu\text{l}$  of d $\text{H}_2\text{O}$ , and 1  $\mu\text{l}$  of each of the 10  $\mu\text{M}$  primers (IS7 and IS8 primers) were ran in duplicates. The libraries were thereafter amplified in five separate PCR reactions with AmpliTaq Gold (Applied Biosystems). PCR products were pooled and cleaned with magnetic beads (Agencourt AMPure XP, Beckman Coulter). We received estimates of the concentration and molarity of the DNA in the cleaned DNA libraries using the 2100 Bioanalyzer Instrument (Agilent Technologies). Consequently, libraries were shotgun-sequenced on Illumina HiSeq 2500 [ $2 \times 125$  base pairs (bp), pair-end (PE)] and HiSeq X ( $2 \times 150$  bp, PE) platforms at the Science for Life Laboratory (SciLifeLab) facilities in Uppsala (SNP&SEQ Technology Platform, Uppsala University) and Stockholm [National Genomics Infrastructure (NGI) Stockholm]. The base calling was performed using an Illumina CASAVA software, and reads were demultiplexed on the basis of six nucleotide index sequences used in library preparation (27).

### UDG-treated libraries

On the basis of screening results of conventional blunt-end libraries, one individual from Kazburun 1 (kzb002) was identified as having

exceptional preservation. To use the individual in a future high-coverage genome database, we commenced production of libraries with repaired damages using uracil-DNA-glycosylase (UDG) treatment to remove deaminated cytosine sites (28). Two new libraries were prepared using the standard blunt-end protocol listed above (27) albeit with an additional 3-hour incubation step at 37°C with 3  $\mu$ l of Uracil-Specific Excision Reagent (USER) enzyme (New England Biolabs) instead of T4 DNA polymerase. After the incubation, 1  $\mu$ l of T4 DNA Polymerase (Thermo Fisher Scientific) was added to the mix, and the mix was further incubated at 25°C for 15 min and at 12°C for 5 min. The libraries were tested using qPCR (as above) and subsequently amplified in five separate reactions of 25  $\mu$ l with 1  $\mu$ l of each 10  $\mu$ M PCR primer (index primer P7 and IS4), 2.5  $\mu$ l of an AccuPrime Pfx buffer, 0.5  $\mu$ l of Herculase II Fusion DNA Polymerase (Agilent Technologies), 16 to 17  $\mu$ l of ddH<sub>2</sub>O, and 4  $\mu$ l of repaired DNA library. The thermocycling conditions consisted of an initial denaturation at 95°C for 2 min, 10 to 14 cycles of 95°C for 15 s, 60°C for 30 s, 68°C for 1 min, and a final extension at 68°C for 5 min. The libraries were pooled, cleaned with magnetic beads (Agencourt AMPure XP, Beckman Coulter), and quantified using the 2100 Bioanalyzer Instrument (Agilent Technologies).

### Whole-genome capture

One individual, scy006, underwent whole-genome capture using 7  $\mu$ l of previously screened library. Hybridization capture reaction was performed using commercially biotinylated probes for human from MYcroarray ([www.mycroarray.com](http://www.mycroarray.com)). The reactions conducted at 60°C for 40 hours in a final volume of 30  $\mu$ l. Captured library was purified with Dynabeads MyOne Streptavidin C1 magnetic beads (Invitrogen) following the manufacturer's protocol (version 3.1) and was subsequently eluted in 30  $\mu$ l of 1 $\times$  tris-EDTA buffer with 0.05% Tween 20 (TET) buffer. Postcapture amplification was performed in five separate reactions with 1  $\mu$ l of each 10  $\mu$ M PCR primer (IS5 and IS6), 10  $\mu$ l of 5 $\times$  Herculase II amplification buffer, 1  $\mu$ l of Herculase II Fusion DNA Polymerase (Agilent Technologies), 0.5  $\mu$ l of deoxynucleotide triphosphates (dNTPs) (25 mM), 32.5  $\mu$ l of ddH<sub>2</sub>O, and 5  $\mu$ l of captured DNA library. The thermocycling conditions consisted of an initial denaturation at 98°C for 30 s, 16 cycles of 98°C for 20 s, 60°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 5 min. After postcapture amplification, the five amplified samples per reaction were pooled and purified with the use of MinElute spin columns (Qiagen) following the manufacturer's instructions. Despite successful library preparation, we did not observe much improvement in terms of single-nucleotide polymorphism (SNP) yield in the capture library. The screening results from captured library were merged with previous results obtained from conventional blunt-end libraries to increase the final coverage to the level used in this paper.

### Sequence data processing

Using MergeReadsFastQ\_cc.py, paired-end reads were merged by requiring at least an 11-bp overlap, and residual adapter sequences were trimmed (29). Merged reads were aligned to the human reference genome (version hs37d5) in single-end mode using Burrows-Wheeler Aligner (BWA) "aln" (version 0.7.12) (30) with parameters "-l 16500 -n 0.01 -o 2" and using BWA "samse." For individuals with more than one library sequenced, Binary Alignment Map (BAM) files from different libraries were merged with SAMtools "merge" (version 1.3) (30). PCR duplicates were removed using a modified version of FilterUniqSAMCons\_cc.py as in (31). Using SAMtools "calmd" (version 1.3) (30), MD tags in

BAM files were generated. Finally, sequences that are shorter than 35 bp and that contain more than 10% mismatches to the reference genome were filtered using a custom script. Ancient genome sequence data of previously published individuals in table S5 were processed following the same procedure to prevent bias in comparative analysis. Genomic coverage was calculated using "genome-CoverageBed" module of BEDTools (32).

### Molecular sex estimation

Molecular sex of each individual was estimated using the Ry method reported in (33). Briefly, the method is based on dividing the number of sequences mapped to the Y chromosome to the number of those mapped to X and Y chromosomes. For this calculation, only sequences with mapping quality of minimum 30 were considered.

### Authenticity testing

Authenticity of the sequences was firstly evaluated on the basis of the presence of ancient DNA-specific damage patterns including increased frequency of C to T type transitions at 5' end of the sequences due to deamination and based on short mean sequence length due to fragmentation (14, 34). PMDtools (35) was used to assess the nucleotide misincorporation patterns at the 5' ends of the reads. All sequence data showed increased C to T frequencies at 5' ends. The sequencing statistics of all libraries are listed in table S3. mtDNA contamination was estimated using two different methods (15, 16). In the first method (15), private or near-private alleles (less than 5% in 311 modern mtDNA sequences) with a minimum sequencing depth of 10 and with a base quality of 30 were considered. Transition-type substitutions were filtered, and the point estimate of contamination was calculated by summing the counts of consensus and alternative alleles across all sites. For the second method (16), to calculate probabilities of being authentic based on mtDNA, SAMtools "mpileup" (30) and vcftools (36) were used. Probabilities of being authentic were calculated using "contamMix" library in R.

### Uniparental markers

Mitochondrial consensus sequence per sample was called using SAMtools "mpileup" (30) and vcftools (36). Haplogroups were determined using Haplofind (37), and final assignment of haplogroups was performed on the basis of visual inspection of each consensus against PhyloTree mtDNA tree (build 17) (38). Variable site calls were prepared using MITOMASTER (39). Mitochondrial DNA genomes of a number of individuals published here were discussed in detail elsewhere (11).

Similarly, Y chromosome sequences were filtered out from BAM files aligned to the human reference genome 19 (hs37d5) using SAMtools version 0.1.19 (30). The Y chromosome haplogroups were assigned using haplogroup identifying SNPs included in minimal PhyloTree<sub>Y</sub> (version 30-Nov-2014) (40) and ISOGG Y-DNA haplogroup tree 2017 version 12.29 (International Society of Genetic Genealogy). Only SNPs of base and mapping qualities of 30 were used. Whenever possible, the haplogroup assignment was restricted to derived transversions to avoid misclassification resulting from damage. All uniparental markers are listed in table S3.

### Data sets

Population genetics analyses were restricted with known present-day genetic variation to minimize false positives. Two different present-day reference panels including (i) the Human Origins genotype data

set (41, 42) and (ii) the 1000 Genomes whole-genome sequence data (43) were used to call the SNPs of ancient individuals sequenced in this study and in published individuals (table S5).

#### Human Origins data set

A curated version of the Human Origins data set consisting of 594,924 autosomal SNP genotype calls for 2730 present-day individuals was obtained from (41). Using SAMtools (version 1.3) “mpileup” (30), SNPs of the ancient individuals overlapping with this data set were discovered. Genotypes of the ancient samples for positions of C/T or G/A type substitutions were encoded as missing data. For overlapping SNPs, genotypes of ancient individuals were merged with reference data set, and a full data set was haploidized by randomly picking one allele per position using a custom script. This data set consisting of 2730 present-day individuals together with a total of 364 ancient samples (table S5) was used for ADMIXTURE and PCA.

#### 1000 Genomes project data set

For statistical analysis, we used an SNP data set ascertained in African Yoruba population as in (26). For this, a data set of 1,979,918 transversions with minor allele frequencies of greater than 10% in African Yoruba individuals from 1000 Genomes Project phase 3 (43) was obtained from (26). SNPs of the ancient individuals overlapping with those were discovered using SAMtools (version 1.3) “mpileup” (30) and merged with this data set using PLINK (44). A full data set was haploidized, as described above.

### Population genetics analysis

#### Principal component and ADMIXTURE analyses

To investigate the genetic affinities of ancient individuals to present-day populations and their affinities with ancient groups in the context of modern genetic variation, we performed PCA. Principal components were calculated using 53 present-day West Eurasian populations (728 individuals) genotyped for 594,924 autosomal SNPs, and 364 ancient individuals were projected (table S5) on an inferred PCA space using the smartpca program of EIGENSOFT (version 7.2) (45). The program was run with “shrinkmode: YES” and “lsqproject: YES” parameters. The result was plotted using “Ploteig” code provided in EIGENSOFT (version 7.2) (45).

To investigate ancestry components in our ancient individuals and to compare them with other ancient and present day individuals, we followed an unsupervised clustering approach implemented in ADMIXTURE (version 1.3) (46). ADMIXTURE analysis was performed following the same procedure in (47). Briefly, the Human Origins data set described in the “Human Origins data set” section for linkage disequilibrium was filtered using PLINK with “--indep-pairwise 200 25 0.4” option (44). This retained a total of 575,658 SNPs. Ancestry components of 2345 individuals from 206 present-day populations were calculated using ADMIXTURE (version 1.3) (46) software. Using these ancestral allele frequencies, cluster memberships of each ancient sample were estimated, as described in (47). Clustering was performed for  $K = 2$  to  $K = 15$ . For each  $K$ , 30 replicate runs with different random seeds were performed. LargeKGreedy algorithm of CLUMPP (48) was used to determine the common signals between each independent runs. Results were plotted using “ggplot2,” “rworldmap,” “maps,” “SDMTools,” and “RColorBrewer” libraries of R.

#### $f_3$ and $f_4$ statistics

To formally test genetic affinity of 35 newly reported ancient individuals to 127 modern populations (1800 modern individuals) from the Human Origins data set, shared genetic drift between them

since their divergence from an outgroup population was estimated using  $f_3$  statistics implemented as “qp3pop” in ADMIXTOOLS (42). The African Yoruba population in the data set was used as an outgroup, and a test was performed for the topology of  $f_3(\text{Outgroup}, \text{ancient individual}, \text{present day population})$ . Results were projected on a geographical map using “ggplot2,” “rworldmap,” “maps,” “SDMTools,” and “RColorBrewer” libraries of R. To measure shared drift between pairs of ancient individuals, admixture  $f_3$  statistics in the form of  $f_3(\text{Outgroup}, X\text{Pop}_1, X\text{Pop}_2)$  was calculated using “qp3pop,” where  $X\text{Pop}_1$  and  $X\text{Pop}_2$  are the  $n_1$  and  $n_2$  members of  $\text{Pop}_1$  and  $\text{Pop}_2$ . Means of these  $n_1 * n_2 f_3$  values were used as genetic similarity between  $\text{Pop}_1$  and  $\text{Pop}_2$ , as well as within population diversity in the presence of  $f_3(\text{Outgroup}, X\text{Pop}_1, X\text{Pop}_1)$ . A 1000 Genomes data set containing 1,979,918 transversions was used for this analysis. The African Yoruba population was used as an outgroup. To test deviations from tree-like topology,  $f_4$  statistics was calculated using “qpDstat” tool of ADMIXTOOLS (42) with “f4mode: YES” option. Significance was assessed by calculating SEs using a block jackknife of 0.5 Mbp. African Yoruba population was used as an outgroup.

#### Diversity estimation

Population diversity was estimated, as described in (26). Briefly, two contemporary individuals with highest number of SNPs in each group were used to represent each population (table S10), and the average number of mismatches between two individuals for all sites in 1000 Genomes data set ascertained in African Yoruba individuals was estimated. Significance was estimated by calculating SEs using block jackknife over blocks of 500 SNPs.

#### Radiocarbon dating and stable isotope data

A total of 29 bone or tooth samples were submitted to Accelerator Mass Spectrometry (AMS) radiocarbon analysis at BETA Analytic Inc. and at Poznan Radiocarbon Laboratory (table S1). Methods for extraction and analyses followed the standard protocols of the two laboratories. Calibration of the conventional radiocarbon ages were performed using Oxcal version 4.3.2 (online) (49) and the IntCal13 atmospheric curve (50). Calibrated results were reported with 95% CIs ( $2\sigma$ ), with a few exceptions. Calibrated ages were reported as years cal BCE or CE. For all but two samples (scy192 and scy197), reliable results were obtained. The two dates fall in a considerably older period, which is not possible when the archaeological find context is considered. The four undated samples (mur001, mur002, scy006, and scy193) all have what may be regarded secure archaeological find contexts to securely be grouped in broad chronological intervals.

The radiocarbon datings showed chronological differences that fall well within earlier estimations of their cultural and archaeological find contexts. Two individuals (scy009 and scy010) from different kurgans at Starosillya have slightly older dates than most of the others. The evaluation of these dates is problematic as they fall within the s.c. Hallstatt radiocarbon calibration plateau (ca. 800–400 cal BC) (51) where the chronological resolution is poor. This needs to be considered, but it has no critical impact on the interpretations presented here.

The samples form five rather clear chronological clusters: Srubnaya-Alakulskaia ( $n = 11$ ), Cimmerians ( $n = 3$ ), Scythians ( $n = 10$ ), and Sarmatians ( $n = 5$ ) (fig. S1A). The oldest group is composed of the Srubno-Alakulskaia individuals, all falling roughly between 1900 and 1550 cal BC. The three Cimmerian samples are slightly younger and all fall between 1000 and 800 cal BC. The dates for the Scythian samples fall between 800 and 100 cal BC, except for one sample that fall between 248 and 391 cal AD (scy332). The last date was younger

than the earlier estimated chronological range of the cemetery ending in the second century (52), as well as the end date of the Scythians. This individual is apparently from a secondary burial. Two individuals from different kurgans at Starosillya are slightly older than most of the others (scy009 and scy010). Finally, the dates for the Huns-Sarmatian samples fall between roughly 0 to 400 AD, however, with some variation within this interval. The studied samples thus represent defined chronological and spatial samples of ancient individuals, well suited for palaeogenomic analyses.

Isotope-Ratio Mass Spectrometry (IRMS) stable isotope analyses were performed on 19 teeth and bone samples in connection to the radiocarbon datings (at Beta Analytic Inc.) (fig. S1B). Dietary patterns are not the main focus here, but the isotopic variation and thus protein input in the diet may affect the integrity of the dates due to reservoir effects. Earlier analyses of stable isotopes have shown that the inclusion of fresh water fish in many areas, and an associated fresh water reservoir effect (FRE) may affect the age estimations, as would also geographical and thus ecological differences with associated trophic-level variations (53, 54). The isotopic values fall in three separate clusters that correlate with the archaeological context of the samples and thus probably also with geographical (and chronological) factors. The variation in stable isotope values in human bone and teeth is probably related to general regional dietary differences between the cultural groups and to the high levels of mobility [see, for example, (55)]. Two Scythian samples, from the Glinoe cemetery, were analyzed, and they exhibited  $\delta^{13}\text{C}$  values that indicate an input of marine protein in the diet. Considering the proximity to the coast of the Black Sea, this is not unexpected. Inclusion of proteins with a marine origin may be expected on this site studied. However, these results cannot be considered as representative for the dietary variation among the buried Scythians at Glinoe. The Srubno-Alakulskaya individuals exhibited a rather narrow variation in  $\delta^{13}\text{C}$  values, while the variation was slightly larger for  $\delta^{15}\text{N}$ , however, within a range below 3‰.

Finally, the individuals from the Late Sarmatian contexts also formed a distinct group. However, two clusters are visible in the scatterplot that are separated by the  $\delta^{15}\text{N}$  values. The rather high  $\delta^{15}\text{N}$  indicated the consumption of proteins probably originating from freshwater fish and probably also that there were some dietary differences between the studied individuals. Also, this may have some consequences for the calibrated age ranges. The chronologically older individuals actually exhibit slightly higher  $\delta^{15}\text{N}$  values than the younger individuals. This may not be the result of a changing dietary pattern through time but rather be interpreted as a sign of FRE. This, however, does not have major consequences for the present study.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/4/10/eaat4457/DC1>

Archaeological context information

Fig. S1. Radiocarbon dating and diet.

Fig. S2. Nucleotide misincorporation patterns at last 30-bp sequences.

Fig. S3. PCA with modern populations.

Fig. S4. Outgroup  $f_3$  statistics.

Fig. S5. PCA with Bronze Age individuals.

Fig. S6. PCA with Paleolithic and Mesolithic individuals.

Fig. S7. PCA with Neolithic individuals.

Fig. S8. PCA with Chalcolithic individuals.

Fig. S9. PCA with Iron Age individuals.

Fig. S10. ADMIXTURE analysis.

Table S1. Archaeological information for individuals used in this study.

Table S2. Stable isotope and radiocarbon dating results information for individuals used in this study.

Table S3. Sequencing statistics and mitochondrial variants for individuals sequenced in this study.

Table S4. Mitochondrial contamination estimates for individuals sequenced in this study.

Table S5. Ancient sample data set details including the number of SNPs overlapping with modern reference.

Table S6. Outgroup  $f_3$  statistics for individuals sequenced in this study.

Table S7. Summary  $f_4$  statistics between Srubnaya and Andronovo, Afanasievo, and Sintashta.

Table S8. Summary  $f_4$  statistics between Srubnaya-Alakulskaya tested as a single population or individuals and Karasuk and other Bronze Age populations.

Table S9. Summary  $f_4$  statistics between Srubnaya-Alakulskaya tested as a single population or individuals and Yamnaya and other Bronze Age populations.

Table S10. Diversity estimates in Bronze Age populations.

Table S11. Summary  $f_4$  statistics between Cimmerians, Scythians, and other ancient populations.

Table S12. Summary  $f_4$  statistics between Cimmerians, Sarmatians, and other ancient populations.

Table S13.  $F_3$  support for different components changing with time in Cimmerians as observed in ADMIXTURE.

Table S14. Summary  $f_4$  statistics between Srubnaya/Srubnaya-Alakulskaya and Andronovo, Afanasievo, and Sintashta.

Table S15. Summary  $f_4$  statistics between Scythians tested as population, Yamnaya, and other Bronze Age populations.

Table S16. Summary  $f_4$  statistics between Cimmerians tested as a population, Srubnaya/Srubnaya-Alakulskaya, and other Bronze Age populations.

Table S17. Summary  $f_4$  statistics of multiway comparisons of individuals from the study together with Srubnaya.

Table S18. Summary  $f_4$  statistics between Sarmatians and other Bronze Age populations. References (56–60)

## REFERENCES AND NOTES

1. D. W. Anthony, *The Horse, The Wheel, and Language : How Bronze-Age Riders From the Eurasian Steppes Shaped the Modern World* (Princeton Univ. Press, 2007).
2. L. N. Koryakova, A. V. Epimakhov, *The Urals and Western Siberia in the Bronze and Iron Ages* (Cambridge Univ. Press, 2007).
3. M. E. Allentoft, M. Sikora, K. G. Sjögren, S. Rasmussen, M. Rasmussen, J. Stenderup, P. B. Damgaard, H. Schroeder, T. Ahlström, L. Vinner, A. S. Malaspinas, A. Margaryan, T. Higham, D. Chival, N. Lynnerup, L. Harvig, J. Baron, P. Della Casa, P. Dąbrowski, P. R. Duffy, A. V. Ebel, A. Epimakhov, K. Frei, M. Furmanek, T. Gralak, A. Gromov, S. Gronkiewicz, G. Grupe, T. Hajdu, R. Jarysz, V. Khartanovich, A. Khokhlov, V. Kiss, J. Kolář, A. Kriiska, I. Lasak, C. Longhi, G. McGlynn, A. Merkevicius, I. Merkyte, M. Metspalu, R. Mkrtychyan, V. Moiseyev, L. Paja, G. Pálfi, D. Pokutta, E. Pospieszny, T. D. Price, L. Saag, M. Sablin, N. Shishlina, V. Smrčka, V. I. Soenov, V. Szeverényi, G. Tóth, S. V. Trifanov, L. Varul, M. Vízce, L. Yepiskoposyan, V. Zhitenov, L. Orlando, T. Sicheritz-Pontén, S. Brunak, R. Nielsen, K. Kristiansen, E. Willerslev, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
4. W. Haak, I. Lazaridis, N. Patterson, N. Rohland, S. Mallick, B. Llamas, G. Brandt, S. Nordenfelt, E. Harney, K. Stewardson, Q. Fu, A. Mittnik, E. Bánffy, C. Economou, M. Francken, S. Friederich, R. G. Pena, F. Hallgren, V. Khartanovich, A. Khokhlov, M. Kunst, P. Kuznetsov, H. Meller, O. Mochalov, V. Moiseyev, N. Nicklisch, S. L. Pichler, R. Risch, M. A. Rojo Guerra, C. Roth, A. Szécsényi-Nagy, J. Wahl, M. Meyer, J. Krause, D. Brown, D. Anthony, A. Cooper, K. W. Alt, D. Reich, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
5. S. F. Adali, in *Eurasian Empires in Antiquity and the Early Middle Ages: Contact and Exchange between the Graeco-Roman World, Inner Asia and China*, F. J. Vervaeke, H. J. Kim, S. F. Adali, Eds. (Cambridge University Press, 2017), pp. 34–59.
6. S. V. Makhortkyh, The North Black Sea Steppes in the Cimmerian Epoch, in *Impact of the Environment on Human Migration in Eurasia*, E. M. Scott, A. Y. Alekseev, G. Zaitseva, Eds. (Springer, 2005), pp. 35–44.
7. J. Davis-Kimball, V. A. Bashilov, L. T. Yablonski, *Nomads of the Eurasian Steppes in the Early Iron Age* (Zinat Press, 1995).
8. R. Brzeziński, M. Mielczarek, *The Sarmatians 600 BC - AD 450* (Osprey Publishing, 2002).
9. I. Mathieson, I. Lazaridis, N. Rohland, S. Mallick, N. Patterson, S. A. Roodenberg, E. Harney, K. Stewardson, D. Fernandes, M. Novak, K. Sirak, C. Gamba, E. R. Jones, B. Llamas, S. Dryomov, J. Pickrell, J. L. Arsuaga, J. M. B. de Castro, E. Carbonell, F. Gerritsen, A. Khokhlov, P. Kuznetsov, M. Lozano, H. Meller, O. Mochalov, V. Moiseyev, M. A. R. Guerra, J. Roodenberg, J. M. Vergès, J. Krause, A. Cooper, K. W. Alt, D. Brown, D. Anthony, C. Laluzza-Fox, W. Haak, R. Pinhasi, D. Reich, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).

10. C. Der Sarkissian, O. Balanovsky, G. Brandt, V. Khartanovich, A. Buzhilova, S. Koshel, V. Zaporozhchenko, D. Gronenborn, V. Moiseyev, E. Kolpakov, V. Shumkin, K. W. Alt, E. Balanovska, A. Cooper, W. Haak; Genographic Consortium, Ancient DNA evidence prehistoric gene-flow from Siberia in the complex human population history of North East Europe. *PLoS Genet.* **9**, e1003296 (2013).
11. A. Juras, M. Krzewińska, A. G. Nikitin, E. Ehler, M. Chyleński, S. Łukasik, M. Krenz-Niedbala, V. Sinika, J. Piontek, S. Ivanova, M. Dabert, A. Götherström, Diverse origin of mitochondrial lineages in Iron Age Black Sea Scythians. *Sci. Rep.* **7**, 43950 (2017).
12. M. Unterländer, F. Palstra, I. Lazaridis, A. Piliipenko, Z. Hofmanová, M. Groß, C. Sell, J. Blöcher, K. Kirsanow, N. Rohland, B. Rieger, E. Kaiser, W. Schier, D. Pozdniakov, A. Khokhlov, M. Georges, S. Wilde, A. Powell, E. Heyer, M. Currat, D. Reich, Z. Samashev, H. Parzinger, V. I. Molodin, J. Burger, Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nat. Commun.* **8**, 14615 (2017).
13. A. Y. Alekseev, N. A. Bokovenko, Y. Boltrik, K. A. Chugunov, G. Cook, V. A. Dergachev, N. Kovalyukh, G. Possnert, J. van der Plicht, E. M. Scott, A. Sementsov, V. Skripkin, S. Vasiliev, G. Zaitseva, A chronology of the Scythian antiquities of Eurasia based on new archaeological and <sup>14</sup>C data. *Radiocarbon* **43**, 1085–1107 (2001).
14. M. Hofreiter, V. Jaenicke, D. Serre, A. von Haeseler, S. Pääbo, DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res.* **29**, 4793–4799 (2001).
15. R. E. Green, A.-S. Malaspina, J. Krause, A. W. Briggs, P. L. F. Johnson, C. Uhler, M. Meyer, J. M. Good, T. Maricic, U. Stenzel, K. Prüfer, M. Siebauer, H. A. Burbano, M. Ronan, J. M. Rothberg, M. Egholm, P. Rudan, D. Brajković, Ž. Kučan, I. Gušić, M. Wikström, L. Laakkonen, J. Kelso, M. Slatkin, S. Pääbo, A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134**, 416–426 (2008).
16. Q. Fu, A. Mittnik, P. L. F. Johnson, K. Bos, M. Lari, R. Bollongino, C. Sun, L. Giemisch, R. Schmitz, J. Burger, A. M. Ronchitelli, F. Martini, R. G. Cremonesi, J. Svoboda, P. Bauer, D. Caramelli, S. Castellano, D. Reich, S. Pääbo, J. Krause, A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
17. D. M. Behar, M. van Oven, S. Rosset, M. Metspalu, E.-L. Loogväli, N. M. Silva, T. Kivisild, A. Torroni, R. Villems, A “Copernican” reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90**, 675–684 (2012).
18. V. I. Molodin, A. S. Piliipenko, A. G. Romaschenko, A. A. Zhuravlev, R. O. Trapezov, T. A. Chikisheva, D. V. Pozdnyakov, Human migrations in the southern region of the West Siberian Plain during the Bronze Age: Archaeological, palaeogenetic and anthropological data, in *Population Dynamics in Prehistory and Early History. New Approaches by Using Stable Isotopes and Genetics*, E. Kaiser, J. Burger, W. Schier, Eds. (De Gruyter, 2012), pp. 93–112.
19. P. Balaresque, N. Poulet, S. Cussat-Blanc, P. Gerard, L. Quintana-Murci, E. Heyer, M. A. Jobling, Y-chromosome descent clusters and male differential reproductive success: Young lineage expansions dominate Asian pastoral nomadic populations. *Eur. J. Hum. Genet.* **23**, 1413–1422 (2015).
20. C. Keyser, C. Bouakaze, E. Crubézy, V. G. Nikolaev, D. Montagnon, T. Reis, B. Ludes, Ancient DNA provides new insights into the history of south Siberian Kurgan people. *Hum. Genet.* **126**, 395–410 (2009).
21. P. Skoglund, H. Malmström, M. Raghavan, J. Storå, P. Hall, E. Willerslev, M. T. P. Gilbert, A. Götherström, M. Jakobsson, Origins and genetic legacy of neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 446–469 (2012).
22. N. L. Chlenova, *Chronology of the Monuments of the Belonging to the Karasuk Epoch* (Moscow Institute of Archaeology Press, 1972).
23. C. Gamba, E. R. Jones, M. D. Teasdale, R. L. McLaughlin, G. Gonzalez-Fortes, V. Mattiangeli, L. Domboróczki, I. Kóvári, I. Pap, A. Anders, A. Whittle, J. Dani, P. Raczky, T. F. G. Higham, M. Hofreiter, D. G. Bradley, R. Pinhasi, Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
24. S. L. Zegura, T. M. Karafet, L. A. Zhivotovskiy, M. F. Hammer, High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol. Biol. Evol.* **21**, 164–175 (2004).
25. D. Y. Yang, B. Eng, J. S. Wayne, J. C. Dudar, S. R. Saunders, Improved DNA extraction from ancient bones using silica-based spin columns. *Am. J. Phys. Anthropol.* **105**, 539–543 (1998).
26. T. Günther, C. Valdiosera, H. Malmström, I. Ureña, R. Rodriguez-Varela, Ó. O. Sverrisdóttir, E. A. Daskalaki, P. Skoglund, T. Naidoo, E. M. Svensson, J. M. B. de Castro, E. Carbonell, M. Dunn, J. Storå, E. Iriarte, J. L. Arsuaga, J.-M. Carretero, A. Götherström, M. Jakobsson, Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11917–11922 (2015).
27. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **6**, pdb.prot5448 (2010).
28. A. W. Briggs, P. Heyn, Preparation of next-generation sequencing libraries from damaged DNA. *Methods Mol. Biol.* **840**, 143–154 (2012).
29. M. Kircher, Analysis of high-throughput ancient DNA sequencing data, in *Ancient DNA. Methods in Molecular Biology*, B. Shapiro, M. Hofreiter, Eds. (Humana Press, 2012), vol. 804, pp. 197–228.
30. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
31. T. Günther, H. Malmström, E. M. Svensson, A. Omrak, F. Sánchez-Quinto, G. M. Kılınc, M. Krzewińska, G. Eriksson, M. Fraser, H. Edlund, A. R. Munters, A. Coutinho, L. G. Simões, M. Vicente, A. Sjölander, B. J. Sellevold, R. Jørgensen, P. Claes, M. D. Shriver, C. Valdiosera, M. G. Netea, J. Apel, K. Lidén, B. Skar, J. Storå, A. Götherström, M. Jakobsson, Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
32. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
33. P. Skoglund, J. Storå, A. Götherström, M. Jakobsson, Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
34. A. J. Hansen, E. Willerslev, C. Wiuf, T. Mourier, P. Arctander, Statistical evidence for miscoding lesions in ancient DNA templates. *Mol. Biol. Evol.* **18**, 262–265 (2001).
35. P. Skoglund, B. H. Northoff, M. V. Shunkov, A. P. Derevianko, S. Pääbo, J. Krause, M. Jakobsson, Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2229–2234 (2014).
36. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
37. D. Vianello, F. Sevini, G. Castellani, L. Lomartire, M. Capri, C. Franceschi, HAPLOFIND: A new method for high-throughput mtDNA haplogroup assignment. *Hum. Mutat.* **34**, 1189–1194 (2013).
38. M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, E386–E394 (2009).
39. M. T. Lott, J. N. Leipzig, O. Derbeneva, H. M. Xie, D. Chalkia, M. Sarmady, V. Procaccio, D. C. Wallace, mtDNA Variation and analysis using Mitomap and Mitomaster. *Curr. Protoc. Bioinformatics* **44**, 1.23.1–1.23.26 (2013).
40. M. van Oven, A. Van Geystelen, M. Kayser, R. Decorte, M. H. D. Larmuseau, Seeing the wood for the trees: A minimal reference phylogeny for the human Y chromosome. *Hum. Mutat.* **35**, 187–191 (2014).
41. I. Lazaridis, N. Patterson, A. Mittnik, G. Renaud, S. Mallick, K. Kirsanow, P. H. Sudmant, J. G. Schraiber, S. Castellano, M. Lipson, B. Berger, C. Economou, R. Bollongino, Q. Fu, K. I. Bos, S. Nordenfält, H. Li, C. de Filippo, K. Prüfer, S. Sawyer, C. Posth, W. Haak, F. Hallgren, E. Fornander, N. Rohland, D. Delsate, M. Francken, J.-M. Guinet, J. Wahl, G. Ayodo, H. A. Babiker, G. Bailliet, E. Balanovska, O. Balanovsky, R. Barrantes, G. Bedoya, H. Ben-Ami, J. Bene, F. Berrada, C. M. Bravi, F. Brisighelli, G. B. J. Busby, F. Cali, M. Churnosov, D. E. C. Cole, D. Corach, L. Damba, G. van Driem, S. Dryomov, J.-M. Dugoujon, S. A. Fedorova, I. G. Romero, M. Gubina, M. Hammer, B. M. Henn, T. Hervig, U. Hodoglugil, A. R. Jha, S. Karachanak-Yankova, R. Khushainova, E. Khusnutdinova, R. Kittles, T. Kivisild, W. Klitz, V. Kučinskas, A. Kushniarevich, L. Laredj, S. Litvinov, T. Loukidis, R. W. Mahley, B. Melegh, E. Metspalu, J. Molina, J. Mountain, K. Näkkäläjärvi, D. Nesheva, T. Nyambo, L. Osipova, J. Parik, F. Platonov, O. Posukh, V. Romano, F. Rothhammer, I. Rudan, R. Ruizbakiev, H. Sahakyan, A. Sajantila, A. Salas, E. B. Starikovskaya, A. Tarekgn, D. Toncheva, S. Turdikulova, I. Uktveryte, O. Utevska, R. Vasquez, M. Villena, C. Voevoda, C. A. Winkler, L. Yepiskoposyan, P. Zalloua, T. Zemunik, A. Cooper, C. Capelli, M. G. Thomas, A. Ruiz-Linares, S. A. Tishkoff, L. Singh, K. Thangaraj, R. Villems, D. Comas, R. Sukernik, M. Metspalu, M. Meyer, E. E. Eichler, J. Burger, M. Slatkin, S. Pääbo, J. Kelso, D. Reich, J. Krause, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
42. N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschorek, T. Webster, D. Reich, Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
43. The 1000 Genomes Project Consortium, An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
44. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
45. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
46. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
47. G. M. Kılınc, A. Omrak, F. Özer, T. Günther, A. M. Büyükkarakaya, E. Biçakçı, D. Baird, H. M. Dönertaş, A. Ghalichi, R. Yaka, D. Koptekin, S. C. A. Çan, P. Parvizi, M. Krzewińska, E. A. Daskalaki, E. Yüncü, N. D. Dağtaş, A. Fairbairn, J. Pearson, G. Mustafaoğlu, Y. S. Erdal, Y. G. Çakan, I. Togan, M. Somel, J. Storå, M. Jakobsson, A. Götherström, The demographic development of the first farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).

48. M. Jakobsson, N. A. Rosenberg, CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806 (2007).
49. C. B. Ramsey, Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360 (2009).
50. P. J. Reimer, E. Bard, A. Bayliss, J. W. Beck, P. G. Blackwell, C. B. Ramsey, C. E. Buck, H. Cheng, R. L. Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, H. Haflidason, I. Hajdas, C. Hatté, T. J. Heaton, D. L. Hoffmann, A. G. Hogg, K. A. Hughen, K. F. Kaiser, B. Kromer, S. W. Manning, M. Niu, R. W. Reimer, D. A. Richards, E. M. Scott, J. R. Southon, R. A. Staff, C. S. M. Turney, J. van der Plicht, IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55**, 1869–1887 (2013).
51. J. van der Plicht, Radiocarbon, the calibration curve and scythian chronology, in *Impact of the Environment on Human Migration in Eurasia*, E. M. Scott, A. Y. Alekseev, G. Zaitseva, Eds. (Springer, 2005), pp. 45–61.
52. S. Łukasik, J. Bijak, M. Krenz-Niedbala, G. Liczbińska, V. Sinika, J. Piontek, Warriors die young: Increased mortality in early adulthood of Scythians from Glinoe, Moldova, fourth through second Centuries BC. *J. Anthropol. Res.* **73**, 584–616 (2017).
53. N. Shishlina, V. Sevastyanov, E. Zazovskaya, J. van der Plicht, Reservoir effect of archaeological samples from Steppe Bronze Age cultures in Southern Russia. *Radiocarbon* **56**, 767–778 (2014).
54. S. V. Svyatko, P. J. Reimer, R. Schulting, Modern freshwater reservoir offsets in the Eurasian Steppe: Implications for archaeology. *Radiocarbon* **59**, 1597–1607 (2017).
55. C. Gerling, A multi-isotopic approach to the reconstruction of prehistoric mobility and economic patterns in the West Eurasian steppes 3500 to 300 BC. *eTopoi J. Ancient Stud.* **3**, 1–21 (2014).
56. I. Shuteleva, N. Sherbakov, K. Gorshkov, The archaeological study of Kazburun Barrow burial ground: Initial results reports. *Eur. Archaeol.*, 21–23 (2010).
57. N. Sherbakov, I. Shuteleva, G. Obydenovaa, M. Balonovaa, O. Khohlovab, A. Golyevac, Some results of the application of a complex approach to the research of the Late Bronze Age Muradymovo settlement in the Volgo-Ural region. *Nat. Sci. Arc.* **1**, 29–36 (2010).
58. V. S. Sinika, S. N. Razumov, S. D. Lysenko, H. P. Telnov, *Report on the Activities of Dniester Archaeological Expedition on the Investigation of Barrows near Glinoe Village of Slobozia District in 2016* (Archive of the Dniester Area Archaeology Museum of Taras Shevchenko Transnistria State University, 2016), p. 95.
59. O. V. Nazarov, in *New research of the Scythian burial mounds in the river Ros' Basin // "Museum readings" Materials of an international scientific conference devoted to the 90th anniversary of the prominent Ukrainian archaeologist O.I.Terenozhkin* (1998), pp. 77–79.
60. L. Kraeva, E. Kryukova, I. Shuteleva, N. Sherbakov, in *Archaeological microdistricts of northern Eurasia* (Omsk, 2009), pp. 78–82.

**Acknowledgments:** We thank A. R. Munters from the Department of Organismal Biology, Uppsala University, Uppsala, Sweden for data management; S. V. Levykin from the Institute of the Steppes of the Ural Branch of the Russian Academy of Sciences in Orenburg for geographical consultation; R. Rodríguez-Varela, N. Kashuba, and V. Sobrado from the Archaeological Research Laboratory, Stockholm University; and M. Somel from the Department of Biological Sciences, Middle East Technical University in Ankara, Turkey. Finally, we thank two anonymous reviewers whose comments and suggestions were invaluable.

**Funding:** The project was supported by the Knut and Alice Wallenberg Foundation (1000 Ancient Genome Project). M.K. was partially supported by Riksbankens Jubileumsfond (P16-0553:1). Sequencing was conducted at the Uppsala University SNP&SEQ Technology Platform. Computations were performed at UPPMAX resources (Uppsala Multidisciplinary Centre for Advanced Computational Science) under the projects b2013203, b2013240, b2015307, b2015056, and b2017011. **Author contributions:** M.K., G.M.K., and A.G. conceived and designed the study. M.K. and A.J. performed the experiments. M.K., G.M.K., D.K., and J.S. processed and analyzed the data. G.M.K., D.K., and M.C. prepared the figures. A.G.N., N.S., I.S., T.L., L.K., F.A.S., A.N.S., I.P., S.Ł., and V.S. contributed samples and provided input about archaeological and anthropological context. M.K.-N., L.D., M.J., J.S., and A.G. coordinated the study. M.K., G.M.K., A.J., J.S., and A.G. wrote the manuscript with input from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The sequence data can be accessed and downloaded from the European Nucleotide Archive under the following study accession number: PRJEB27628. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 27 February 2018  
Accepted 23 August 2018  
Published 3 October 2018  
10.1126/sciadv.aat4457

**Citation:** M. Krzewińska, G. M. Kılınc, A. Juras, D. Koptekin, M. Chyleński, A. G. Nikitin, N. Shcherbakov, I. Shuteleva, T. Leonova, L. Kraeva, F. A. Sungatov, A. N. Sultanova, I. Potekhina, S. Łukasik, M. Krenz-Niedbala, L. Dalén, V. Sinika, M. Jakobsson, J. Storå, A. Götherström, Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Sci. Adv.* **4**, eaat4457 (2018).

## Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads

Maja Krzewinska, Gülsah Merve Kiliç, Anna Juras, Dilek Koptekin, Maciej Chylenski, Alexey G. Nikitin, Nikolai Shcherbakov, Iia Shuteleva, Tatiana Leonova, Liudmila Kraeva, Flarit A. Sungatov, Alfija N. Sultanova, Inna Potekhina, Sylwia Lukasiak, Marta Krenz-Niedbala, Love Dalén, Vitaly Sinika, Mattias Jakobsson, Jan Storå and Anders Götherström

*Sci Adv* 4 (10), eaat4457.  
DOI: 10.1126/sciadv.aat4457

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/4/10/eaat4457>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2018/10/01/4.10.eaat4457.DC1>

### REFERENCES

This article cites 46 articles, 4 of which you can access for free  
<http://advances.sciencemag.org/content/4/10/eaat4457#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2018 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).