bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Inner Asian maternal genetic origin of the Avar period nomadic elite in the 7th century AD Carpathian Basin

Veronika Csáky[1*], Dániel Gerber[1], István Koncz[2], Gergely Csiky[1], Balázs G. Mende[1], Antónia Marcsik[3], Erika Molnár[3], György Pálfi[3], András Gulyás[4], Bernadett Kovacsóczy[5], Gabriella M. Lezsák[6], Gábor Lőrinczy[7], Anna Szécsényi-Nagy[1,*,#], Tivadar Vida[1,2,*,#]

1: Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences, 1097, Budapest, Hungary
2: Institute of Archaeological Sciences, Eötvös Loránd University, 1053, Budapest, Hungary
3: Department of Biological Anthropology, University of Szeged, 6726, Szeged, Hungary
4: Jász Museum, 5100 Jászberény, Hungary
5: Katona József Museum, 6000, Kecskemét, Hungary
6: Institute of History, Research Centre for the Humanities, Hungarian Academy of Sciences, 1097, Budapest, Hungary
7: Móra Ferenc Museum, 6720, Szeged, Hungary

*Corresponding authors: csaky.veronika@btk.mta.hu, szecsenyi-nagy.anna@btk.mta.hu, vida.tivadar@btk.mta.hu
[#]These authors jointly supervised this work.

Abstract

After 568 AD the nomadic Avars settled in the Carpathian Basin and founded their empire, which was an important force in Central Europe until the beginning of the 9th century AD. The Avar elite was probably of Inner Asian origin; its identification with the Rourans (who ruled the region of today's Mongolia and North China in the 4th-6th centuries AD) is widely accepted in the historical research.

Here, we study the whole mitochondrial genomes of twenty-three 7th century and two 8th century AD individuals from a well-characterised Avar elite group of burials excavated in Hungary. Most of them were buried with high value prestige artefacts and their skulls showed Mongoloid morphological traits.

The majority (64%) of the studied samples' mitochondrial DNA variability belongs to Asian haplogroups (C, D, F, M, R, Y and Z). This Avar elite group shows affinities to several ancient and modern Inner Asian populations.

The genetic results verify the historical thesis on the Inner Asian origin of the Avar elite, as not only a military retinue consisting of armed men, but an endogamous group of families migrated. This correlates well with records on historical nomadic societies where maternal lineages were as important as paternal descent.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

The Carpathian Basin in East-Central Europe is generally regarded as the westernmost point of the Eurasian steppe; as such, its history was often influenced by the movements of nomadic people of eastern origin. At the end of the 6th century AD, the nomadic Avars settled in the Carpathian Basin and founded their empire which remained a powerful player in the geopolitical arena of Central and Eastern Europe for a quarter of a millennium[1–3]. The Avar population cannot be regarded as homogeneous, as the Avar occupation of the Carpathian Basin can be best understood if set against the backdrop of how nomadic peoples built their empires[4–7]. The Avars conquered and then united different nomadic people from Eastern Europe and Asia, the Late Antique and Germanic population of the former Roman province Pannonia, Slavs, and other inhabitants of the border provinces of the Byzantine Empire under their rule[8,9].

Historical and archaeological context

The identification of the Avars with the Inner Asian *Rourans* mentioned by several Chinese chronicles appeared as early as the 18th century and it is still the most accepted theories in the historiography of the Avars up until the middle of the 20th century when the rival thesis of their Central Asian Hunnic (*Hephtalite*) descent of the Avars was born. This debate remained unsolved, however a rising number of evidences points towards the Rourans.
The solution of this question is hindered by the paucity and obscurity of sources: according to a written account describing a Byzantine embassy to the Western Turkic Qaganate, the Turkic ruler, Silzibulos, claimed that the Avars were his subjects who escaped from his will and that they unrightfully usurped the title *Qagan* and the name 'Avar' as these fugitives were *Warkhonitai*, i.e. not real Avars in the Turkic *Qagan*'s view[2,10–12]. All of this suggests that at least the Avars' elite was of Inner Asian origin, bringing titles and institutions of a nomadic state to the Carpathian Basin. The Avar ruler's title *Qagan* appeared in Inner Asia among the *Xianbei* people during the 3rd century. The *Rourans*, who founded an empire in today's Mongolia that lasted until the mid-6th century, inherited the title and used it to label their supreme leader[6]. As such, the Avar elite's identification with refugees of the *Rouran* ruling stratum is based on the analysis of written accounts like Theophylact Simocatta's Pseudo-Avar story or by the sudden appearance of titles and personal names of Inner Asian origin (*Qagan, Bayan, Yugurrus, Tarkhan*, etc.)[2,11,12]. The connection to Inner Asia is attested by the clear cultural contacts of their archaeological material as well. The language of the Avars and *Rourans* is unknown. The only evidence that we have is from Inner Asia, which suggests Proto-Mongolian, Proto-Turkic and/or a still not defined Inner Asian language[12,13].

Certain artefacts of the early Avar period material culture, such as elements of the horse harness and weaponry, were parts of the Eurasian nomadic cultures for centuries[3,14,15]. Certain prestigious artefact types have direct analogies from Mongolia and Northern China, while Avars may have contributed in the spread of Asian technological innovations in Europe. The Avar material culture shows, how a nomadic population and its ruling elite remained part of

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the connection network that is the Eurasian steppe, even generations after settling in the Carpathian Basin (Fig. 1).

Background of the bioarchaeological research

The Carpathian Basin witnessed population influxes from the Eurasian Steppes several times, but they are genetically poorly documented yet. The earliest such migration was that of the *Yamnaya* people in the 3rd millennium BC[16]. The second population was the Iron Age Scythians, who settled the Eastern European area and the Carpathian Basin in the 1st millennium BC. Scythians were described by a recent genomic study as genetically highly structured. The Scythian samples from Hungary had relatively increased European farmer ancestry and showed no signs of gene flow from Inner Asian groups[17]. In the 5th century AD, the Carpathian Basin was conquered by the Huns, who are associated with the Inner Asian *Xiongnu* confederation in the historical and archaeological research. Their genetic origin is intensively studied as well[17], but bioarchaeological data on the scarce human remains connected to the Huns from Hungary have not been published yet.
Besides influxes from the east, the Carpathian Basin witnessed the movement of people from the north as well. The Lombards for e.g., who directly preceded the Avars in Transdanubia (today's Western Hungary), migrated across the Carpathian Basin to Italy in the mid-6th century. The population of two analysed Lombard cemeteries in Hungary and Italy showed the Central and North European genomic ancestry of this Germanic tribe[18,19].

Few ancient DNA studies have focused yet on the Avars, and these studies analysed only the control region of the maternally inherited mitochondrial DNA (mtDNA). One research focused on a 7th-9th century Avar group from the micro-region of south-eastern Great Hungarian Plain (Alföld) of the Carpathian Basin[20]. This group buried the deceased in catacomb graves and had mostly Europid, local morphological skull characters. Their maternal gene pool showed predominantly southern and Eastern European composition, with Asian elements presenting only 15.3% of the variation. These Asian haplotypes shared lineages with almost all comparative modern Central and East Asian populations.

Another recent study of a mixed Avar-Slavic population from the 8th-9th centuries from present-day Slovakia showed a mixed Eurasian character too, with a lower frequency (6.52%) of Asian (East Eurasian) elements[21]. Šebest et al. explained the rather European mtDNA composition with high assimilation rate in the investigated population, where the majority of marriages took place between Avar men and Slavic women[21].

The Avar population was uniform in neither cultural nor anthropological characters. More than 100,000 graves from the Avar period were excavated in the territory of present-day Hungary. The osteological record shows heterogeneous anthropological composition with mainly Europid morphological characters, while Asian cranial indices were dominated only in certain regions and periods[22,23] (see Supplementary Information).

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Using whole mitogenome sequencing, our current research focused on the maternal genetic composition and origin of an Avar period group of probably high-ranking individuals dated to the 7th century AD. We also investigated the homogeneity and relationship of the leading stratum of the Avar population, as well as their genetic connection to other ancient and modern Eurasian population. We aimed to obtain information on the organisation of this nomadic group, as well as to track the Avars' genetic admixture with other contemporaneous populations and observe the extent of their genetic legacy in the later Medieval periods.

The studied individuals were excavated at ten different sites (found in small burial groups or single burials). Seven out of ten sites are located in the Danube-Tisza Interfluve (central part of present-day Hungary and the Carpathian Basin) dated to the middle of the 7th century (Fig. 2). This group of lavishly furnished burials—most notable probably the so-called *'Qagan* burial' from Kunbábony—is characterised by high value prestige artefacts such as swords, gold belt sets with pseudo-buckles and certain elements of precious metal tableware[24] (Fig. S1, Supplementary Information). The concentration of burials with these artefacts can, in all likelihood, be linked to leaders of the early Avar polity and the *Qagan*'s military retinue. They can also be seen as a reflection not only of a powerful military centralisation, but also of an Avar power centre in the Danube-Tisza Interfluve[25,26]. The primary focus of the sample selection was to target all available members of this elite group. During the sample collection it became evident that these individuals are also tied together by their physical anthropological characteristics, as the skulls showed mainly Mongoloid morphological traits to a different extent (Supplementary Information). To have a better understanding of this group, we later collected samples from 17 individuals with Mongoloid characteristics, but without any outstanding grave goods unique only to the highest social ranks (Fig. 2, Fig. S2) from the Danube-Tisza Interfluve and the Transtisza region.

Results

We investigated 25 Avar period individuals from an archaeologically, geographically and anthropologically well-defined group (Table S1). We managed to sequence the mitochondrial genome using a hybridisation capture method in 2-68x coverage (42x average coverage for the 25 samples, see Table S1). Osteological sex determination was checked by evaluating shallow shotgun sequencing data. The studied Avar group composed of 17 males and 8 females.

The mitochondrial genome sequences can be assigned to a wide range of the Eurasian haplogroups with dominance of the Asian lineages, which represent 64% of the variability: four samples belong to Asian macrohaplogroup C (two C4a1a4, one C4a1a4a and one C4b6); five samples to macrohaplogroup D (one by one D4i2, D4j, D4j12, D4j5a, D5b1), and three individuals to F (two F1b1b and one F1b1f). Each haplogroup M7c1b2b, R2, Y1a1 and Z1a1 is represented by one individual. One further haplogroup, M7 (probably M7c1b2b), was detected (sample AC20); however, the poor quality of its sequence data (2.19x average coverage) did not allow further analysis of this sample.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

European lineages (occurring mainly among females) are represented by the following haplogroups: H (one H5a2 and one H8a1), one J1b1a1, three T1a (two T1a1 and one T1a1b), one U5a1 and one U5b1b (Table S1).

We detected two identical F1b1f haplotypes (AC11 female and AC12 male) and two identical C4a1a4 haplotypes (AC13 and AC15 males) from the same cemetery of Kunszállás; these matches indicate the maternal kinship of these individuals. There is no chronological difference between the female and the male from Grave 30 and 32 (AC11 and AC12), but the two males buried in Grave 28 and 52 (AC13 and AC15) are not contemporaries; they lived at least 2-3 generations apart.

Haplogroup based analyses

We performed Principal Component Analysis (PCA) with the Avar dataset using haplogroup frequencies of another 43 ancient and 64 modern populations (Tables S2 and S3, Figs. S3-4). The Avars show affinities to some Asian populations: they are close to 5th-3rd centuries BC Scythians from the Altai region and 15th-19th centuries Yakuts from East-Siberian Central Yakutia[27] along PC1 and PC2, while along PC3, the Avars are near to South Siberian Bronze Age populations of Andronovo and Krotovo cultures[28] on the Asian part of the plot, which is possibly caused by high loadings of the haplogroup vectors T1 and R (Fig. S3a-b, Table S2). The strict separation of Asian and European populations is also displayed on the Ward-type clustering tree, which is based on the same dataset as the PCA. Here the Avar elite is located on an Asian branch of the tree and clustered together with Central Asian populations dated to the Late Iron Age, Hunnic and Medieval periods[17], and furthermore, with *Xiongnu* period population from present-day Mongolia[29,30] and Scythians from the Altai region[31–34](Fig. 3).

On the PCA plot with modern populations, the Avars are clustered together with Central Asian modern populations from Kazakhstan, Kyrgyzstan and the Altai region as well as with modern Uyghurs living in Northwest-China, East-Siberian Yakuts and Nganasans (Table S3). This phenomenon is pictured on the Ward-type clustering tree too, where the Asian, European and Central-South Asian and Near-Eastern populations are separated on the major branches, and the Avars are situated on the Asian branch, together with modern Altaian, Mongolian and Buryat populations, as well as with Kyrgyz, Uyghurs and Kazakhs (Fig. S5).

Sequence based analyses

The pairwise genetic distances were calculated between ancient and modern Eurasian populations based on 932 ancient and 3,945 modern publicly available whole mitochondrial genome sequences (Tables S4 and S5). The Avar group shows significant genetic distances ($p < 0.05$) from most ancient populations. Only two groups from Central Asia have non-significant differences from the Avar elite: one group containing Late Iron Age samples (nomads from the Late Iron Age and Hun period from the Kazakh Steppe and the Tian Shan) ($F_{ST} = -0.00116$, $p = 0.42382$), and a group of Medieval samples from the Central Asian Steppe and the Tian Shan (Turks, Karluk, Kimak, Karakhanid, Kipchak, Golden Horde and

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

other Medieval nomads)[17] (F_{ST} = 0.00650, p = 0.26839, Table S4). Note, that the reference dataset of the distance calculations is smaller than that used for haplogroup-based analyses, because the whole mitochondrial genome datasets of ancient Asian populations are scarce in the current state of research.

The Multidimensional Scaling (MDS) plots were displayed based on linearised Slatkin F_{ST} values (Tables S4 and S5) that were calculated between the Avars and 22 ancient and 42 modern populations, respectively. The MDS plot of ancient populations does not show a clear chronological or geographical grouping of the populations; however, Asia and Europe are separated. The Avar elite group is close to Central-Asian populations from the Late Iron Age and Medieval period[17] in accordance with the individual F_{ST} results (Fig. 4).

The Avar period elite group exhibits a non-significant difference in genetic distance from the following modern populations: Uyghurs living in Northwest-China (Xinjiang, Turpan prefecture) (F_{ST} = 0.00656, p = 0.14929), Hazara (F_{ST} = 0.007, p = 0.24027) and Burusho (F_{ST} = 0.01748, p = 0.10949) populations of the Central Asian Highlands (Table S5).
The MDS with Avars and modern populations display the differentiation of European, Near-Eastern, Central-Asian and East-Asian populations along coordinate 1. The investigated Avars are located close to Central-Asian and Inner Asian populations (Fig. 5). They cluster with modern Burusho and Hazara populations living in the Central-Asian Highlands, with modern Uyghurs living in Northwest-China, and the Han Chinese.

The maternal genetic lineages of the Avar individuals were further investigated through the Neighbour Joining clustering method (NJ). We separately counted and constructed phylogenetic trees of the 16 mtDNA haplogroups detected (see Table S6, Methods, Supplementary Information). The Neighbour Joining trees of the Asian haplogroups provide evidence of the phylogenetic connection of the 16 Avar individuals with modern Uyghurs from Northwest-China in the first case, but connections are also shown with certain Central- and Inner Asian populations (individuals from Pamir region, Chinese, Yakuts, Buryats, Barguts, Evenks, Evens etc.) (Figs. 6 and S7a-o).

Summary of the Eurasian genetic affinities of the Avars

Connections with modern East Asia are documented on D4j, D5b1 and M7c1b2b phylogenetic trees, where the Avar elite samples shared a common ancestor with several Chinese individuals (Figs. S7c, S7d, S7f). The Han Chinese[35,36] are relatively close to the Avar group on the MDS plot, but their genetic distance is significant. Unfortunately, modern whole mitogenomic data are underrepresented in certain Asian regions (e.g. Mongolia) which probably played an important role in the early Medieval nomadic migrations.

Modern East-Siberian populations, namely Turkic speaking Yakuts and Samoyedic speaking Nganasans, are close to the Avar elite based on their haplogroup composition (Figs. S4-S5, Table S3). Phylogenetic connections of several investigated individuals from the Avar period to the Yakuts and Nganasans as well as to further East-Siberian individuals (Evenks and

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tungusic people) are presented in C4a1a, D4i, D4j, F1b1, Y1a and Z1a NJ trees (Figs. 6 and S7a, S7c, S7e, S7h, S7o).

Some individuals of the Avar elite group show genetic connection with Russian Trans-Baikalian Mongolian-speaking Buryats and Barguts, the affinities of which are displayed on C4a1a, D4i and D4j phylogenetic trees (Figs. 6, S7a, S7c). The Buryats also stay on one branch on the Ward-type clustering tree (Fig. S5). Furthermore, the Buryats appear on C4b, F1b1 and Y1a phylogenetic trees as well (Figs. S7b, S7e, S7h). Derenko et al. recently summarised the genetic research of the Buryats, who show connections to Chinese and Japanese but also to Turkic and Mongolic speaking populations[37]. Yunusbayev et al. concluded based on genome wide genotype data that Tuvinians, Buryats and Mongols are autochthonous to their current southern Siberian and Mongolian residence[38]. The Buryats certainly represent a population that did not migrate much in the last millennia; therefore, they can be a good proxy for the Medieval population of South Siberia.

The genetic connection of Avar elite group with modern Uyghurs from Northwester-China (Xinjiang, Turpan prefecture)[39] is supported by the lowest genetic distance between Avars and Uyghurs compared to the other 42 modern populations studied (Table S5). This connection is displayed on the MDS plot too (Fig. 5). The Uyghurs are relatively near to Avars on the haplogroup frequency-based PCA plots along all three components; the Ward-clustering tree shows the connection as well (Fig. S4-S5, Table S3). The NJ trees of haplogroups C4b, D4i, D4j, D5b, F1b1, M7c1b2, R2, Y1a and Z1a also give evidence of the phylogenetic connection of certain individuals of the Avar elite group to the modern Uyghurs (Figs. 6 and S7b-h, S7o). However, it is important to emphasise that this population is not the descendent of the Medieval *Uighur* Empire, since modern Uyghurs gained their name only during the 20th century.

The genetic distance is small between the investigated Avar elite and some modern-day ethnic groups from the Central-Asian Highlands (lying mostly in the territory of Afghanistan and Pakistan), namely the Hazara and Burusho, Pathan, Balochi and Brahui populations (Table S5)[35], the connections of which are shown on the MDS plot (Fig. 5) and on the haplogroup R2 tree (Fig. S7g) as well. Interestingly, the Hazara population, living mostly in Afghanistan and Pakistan today, probably has a Mongolian origin. Further Central-Asian individuals from the Pamir Mountains show phylogenetic connection with Avars on the D4j, R2 trees, and interestingly also on the European T1a1b tree (Figs. S7c, S7g, S7n).

The Central Asian Kazakhs and Kyrgyz cluster together with our Avar group on PCA plots and clustering tree (Figs. S4-5, Table S3). Unfortunately, they cannot be presented on the MDS plot because of the absence of population-level whole mitogenomic data. However, one modern Kazakh individual with the D4i haplogroup shares a common ancestor with an Avar period individual AC6 (Fig. 6).

Caucasian genetic connection is presented only by one sample on the phylogenetic tree of haplogroup H8a, where the AC17 sample from the Avar period is situated on one branch that

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also contains ancient and modern Armenians (Fig. S7j). This result correlates to the first European mention of the Avars in the Caucasus region by Menander Protector[2].

Discussion

In 568 AD the Avars arrived in the Carpathian Basin, which was inhabited in the 6th century by a mixed Germanic, Sarmatian and Late Antique (Roman) populations. Little is known about the maternal genetic composition of these local groups. The inhabitants of the Avar Qaganate had a heterogenic origin, the highest social stratum (presented by the high ranking individuals buried in the region of the Danube-Tisza Interfluve) shows a homogeneous cultural and anthropological character.

We produced valuable information regarding the social organisation of the elite stratum. Among the 25 studied individuals, we detect 23 different maternal lineages. These results show that the group was not organised by matrilineal social rules. Although the maternal genetic data of the investigated Avar group contains both Western and Eastern Eurasian elements, Eastern Eurasian maternal lineages dominate the spectrum in 64%. This suggests that not only a military retinue consisting of males migrated, but also an endogamous group of families.

There is a clear difference between the genetic pattern of the males and females. Of the 25 investigated samples, eight originated from female individuals and four have European haplogroups (that is 50% percent of the European lineages). In the case of the males, the ratio of European lineages is only 17.6% (three out of 17). This correlates well with the picture where maternal lineages were recorded in many historical nomadic societies and were as important as paternal descent[40,41] and where women had the right and the possibility to wield power (see the title of *Qatun* in ref.[2,42]). However, the overrepresentation of European haplogroups in the female group contrasting with the small number of them in the male burials could be the result of archaeological discrepancy. The investigated elite group consists of almost male burials; the women belonging to the same social strata are archaeologically invisible. From the investigated sites in the Danube-Tisza Interfluve, only one female individual was buried with high value artefacts; the other richly furnished female burials are located in the Transtisza region (Fig. 2). Therefore, the males and females investigated here did not necessarily belong to the same social group, although they were part of the same communities (cemeteries). It is important to note that the Avar elite group under investigation does not belong to the first generation of conquerors of the Carpathian Basin. Based on our results, we suggest that the newcomer Avar elite was not admixed with the local 6th century population for ca. a century and remained a consciously maintained closed society; only after that period did the number of intermarriages with local women increase. While certain maternal blood relations can be assumed by our analyses, the socially probably more important paternal contribution of kinship should be further investigated in the male lines.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Interestingly the elite group does not exhibit a genetic connection to the previously investigated small Avar period population from southeast Hungary[20] either, because the latter shows predominantly Eastern European maternal genetic and also different archaeological characters with Eastern European traditions (catacomb graves and partial animal depositions for e.g.). The genetic difference correlates well with the cultural and anthropological differences of this group and shows that the conquering Avar population was heterogeneous and differed in genetic makeup from the local population. We found that the Avar elite group is genetically different from the 6th century Lombard period community of Szólád in Transdanubia[18] (Fig. 2), which has genetic connections to other ancient European populations (Fig. 3), and we can also assume that the other local Germanic populations also had a rather European type mtDNA variability in contrast to the studied Avar group. Comparing the Avar elite with later period datasets from the Carpathian Basin, only a few connections are observable. The mixed Avar-Slavic population from the 8th-9th centuries[21] does not show affinities to the Avar elite. The T1a1b phylogenetic tree contains one individual from the Hungarian conquest period (sample Karos III/14 in Neparáczki et al.[43]) with identical sequence to the Avar HC9, which might indicate the genetic continuity of certain maternal lineages between the 7th and 9th-10th centuries. The overall mtDNA composition of the Avar elite group and the 9-12th century populations of the Carpathian Basin differ significantly, but whole genomic reanalyses of the latter samples are needed to prove these results.

Only loose connections are detectable between the new Avar mtDNA dataset and the available ancient populations in Eurasia. The comparisons are encumbered by the geographically and chronologically scattered nature of the available reference data.
The Avar period elite shows the lowest and non-significant genetic distances to ancient Central Asian populations dated to the Late Iron Age (Hunnic) and to the Medieval period[17], which is displayed on the ancient MDS plot (Fig. 4); these connections are also reflected on the haplogroup based Ward-type clustering tree (Fig. 3). Building of these large Central Asian sample pools is enabled by the small number of samples per cultural/ethnic group.. Further mitogenomic data from Inner Asia are needed to specify the ancient genetic connections; however, genomic analyses are also set back by the state of archaeological research, i.e. the lack of human remains from the 4th-5th century Mongolia, which would be a particularly important region in the study of the Avar elite's origin[44].

The investigated elite group from the Avar period elite also shows low genetic distances and phylogenetic connections to several Central and Inner Asian modern populations. Our results indicate that the source population of the elite group of the Avar *Qaganate* might have existed in Inner Asia (region of today's Mongolia and North China) and the studied stratum of the Avars moved from there westwards towards Europe. Further genetic connections of the Avars to modern populations living to East and North of Inner Asia (Yakuts, Buryats, Tungus) probably indicate common source populations.

The genetic results verify the old historical and archaeological thesis on the Inner Asian origin of the Avar elite and disprove the hypotheses about the Central Asian *Heftalite* origin of the

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Avars. The historical sources suggest that this group was bringing titles and institutions of a Nomadic state to the Carpathian Basin to the discontent of their Asiatic heirs, the Turks. While the already mentioned idea of Avar – *Rouran* identity remained largely unchallenged, the *Rouran* origin of the Avar ruling elite does not automatically mean that the whole population ruled by Avar *Qagans* or even the whole Avar elite was composed exclusively of *Rouran* fugitives. It can rather be imagined as a composition of a second power centre in the steppes against the Western Turkic *Qaganate* resulting in the joining of different groups with no regard for the linguistic or ethnic affiliations who were discontent with the rise of this new nomadic empire in the mid-6th century[2].

The wealth and grave goods found in the single burial of a 50-60 year-old man at Kunbábony (AC2) is outstanding, not only compared to the standards of the early Avar period, but to the group of elite burials located in the Danube-Tisza Interfluve (Kunbábony - Bócsa group). The burial itself is often described as '*Qagan* burial'; while it cannot be clearly proved archaeologically, the 2.34 kilograms of gold buried with him shows that he was indeed a prominent member of society (see Supplementary Information). Out of the 25 investigated samples, eight individuals belonged to this archaeologically defined group of high ranking men buried with similar weaponry covered with precious metal foils, ornamented belt sets and drinking vessels made of gold or silver (Csepel-AC1, Kecskemét-AC23, Kunbábony-AC2, Kunpeszér Grave 3-AC21, 8-AC22, 9-AC20, Petőfiszállás Grave 1-AC19, Szalkszentmárton-AC8). The artefacts found in these burials point to eastern cultural connections, but a more precise definition is hindered by the different distribution pattern of certain artefact types.

The crescent-shaped gold sheet from the Kunbábony burial have direct analogies from 5th-6th century burials attributed to the *Rouran* from Mongolia and Northern China (Inner Mongolia) where these artefacts were used as pectorals based on their documented position in burials (Yihe Nur in China[45] and Talyn Gurvan Kherem in Mongolia[44]), although these items were regarded earlier as head-gears or diadems used as insignia[46] (Fig. 7).

Characteristic ostentatious edged weapons covered with golden or silver sheets were found in early Avar elite burials, and were probably used as prestige goods. These ring-pommel swords have good contemporaneous analogies from the Altai Region, but their distribution reached as far as China, Korea and Japan, while their representations can be found in mural paintings of Old-Samarkand (Afrasiab)[25,47] (Fig. 1).

The strong steppe connection of the early Avar period material culture is present on the level of the common people as well (see the rectangular-mouthed vessels and vessel with peaked and knobbed rim[48]). The technological breakthrough of the iron stirrup appeared in Europe with the Avars; its importance was also showed by the cavalry reform of the Byzantine military after its large-scale distribution[49–51]. However, not only certain artefact types reached Europe during that time, but also some characteristic ways of deposition also spread: sacrificial assemblages so-called '*tainiks*' (cache) composed of weapons and horse harnesses were buried in shallow pits in the Carpathian Basin during this period[8,47,52,53] (Fig. S8). The

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Inner Asian genetic connections of the common people of the Avar society also need to be tested in the future, analysing representative series of individuals.

Conclusion

All kinds of available evidence attest that at least a significant part of the Avar elite was of Inner Asian origin, not only in cultural but also in their genetic heritage. The detected Inner-Asian maternal genetic composition of the elite was preserved through several generations after the conquest of the Carpathian Basin.

This result suggests a consciously maintained closed society, probably through internal marriages or – as opposed to the vivid picture of their flight described by the written sources – intensive contacts with their regions of origin. However, in the Avar population, new elements appeared en route and after the conquest of the Carpathian Basin, showcasing how a successful military elite attracted and assimilated other groups, for example with Caucasian or European origin. The results also hold valuable information regarding the social organisation of the Avar period elite, as it suggests that not only a military retinue consisting of males migrated, but an endogamous group of families.

Our first genetic results on the leader class of the Avar society provides new evidence of the history of an important early Medieval Empire. Nevertheless, further genetic data from ancient and modern Inner Asian populations are needed to more exactly describe the genetic relations and the different waves of nomadic migration through the Eurasian steppe.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Ancient DNA work

Twenty-five samples were collected from ten different cemeteries dated to the Avar period (7th-8th centuries) according to their geographical position, grave goods, funerary custom and anthropological characteristics (see Table S1 and the site and grave descriptions in the Supplementary Information).

All stages of the work were performed under sterile conditions in a dedicated ancient DNA laboratory (Laboratory of Archaeogenetics in the Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences) following well-established ancient DNA workflow protocols[20,54]. The laboratory work was carried out wearing clean overalls, facemasks and face-shields, gloves and over-shoes. All appliances, containers and work areas were cleaned with DNA-ExitusPlus™ (AppliChem) and/or bleach and irradiated with UV-C light. All steps were carried out in separate rooms. In order to detect possible contamination by exogenous DNA, one extraction and library blank were used as a negative control for every batch of five/seven samples. Haplotypes of all persons involved in the ancient DNA work were determined and compared with the results obtained from the ancient bone samples.

Usually, pars petrous bone fragments were used for analyses, except for three individuals where teeth and long bone fragments were collected because the skulls were not preserved (Table S1).

The DNA extraction was performed based on the protocol of Dabney et al.[55] with some modifications described also by Lipson et al.[54]. DNA libraries were prepared using UDG-half treatment methods[56]. We included library negative controls and/or extraction negative controls in every batch. Unique P5 and P7 adapter combinations were used for every library[56,57]. Barcode adaptor-ligated libraries were then amplified with TwistAmp Basic (Twist DX Ltd), purified with AMPure XP beads (Agilent) and checked on a 3% agarose gel. The DNA concentration of each library was measured on a Qubit 2.0 fluorometer. In solution, the hybridisation method was used to capture the target short sequences that covered the whole mitochondrial genome, as described by Haak et al. and Lipson et al.[16,54]. Captured samples as well as raw libraries for shotgun sequencing were indexed using universal iP5 and unique iP7 indexes[57]. NGS sequencing was performed on an Illumina MiSeq System using the Illumina MiSeq Reagent Kit v3 (150-cycles).

Bioinformatics analyses

The final BAM files were obtained by a custom pipeline for both shotgun and capture datasets. The paired-end reads were merged using SeqPrep master (https://github.com/jstjohn/SeqPrep), allowing a minimum overlap of 5 bp and minimum length of 15 bp. Then, the reads were filtered by size and barcode content using *cutadapt* version 1.9.1[58], allowing no barcode mismatch, and a minimum length of 15 bp. BWA version

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

0.7.12-r1039[59] was used to map the capture sequencing reads to the Cambridge Reference Sequence (rCRS) and the shotgun sequencing reads to hg38 human genome assembly allowing a 3 bp difference in seed sequence, 3 bp gap extension and 2 gap opening per reads. The downstream analyses including SAM-BAM conversion, sorting, indexing and PCR duplicate removal was performed by *samtools* version 1.6[60].

For capture data, indel realignment was performed using Picard tools version 2.5.0 (https://github.com/broadinstitute/picard) and GATK version 3.6[61–63]. The presence of a deamination pattern was estimated by MapDamage version 2.0.8 (https://ginolhac.github.io/mapDamage/) and summarised in Table S1. Due to the relatively young age and half-UDG treatment of the samples required, the deamination frequency did not reach the minimum limit for software *schmutzi* in most cases; therefore, the final validation of the sequences was performed by eye on the final bam files. The shotgun sequencing provided a raw estimate of the endogenous content and genetic sex determination according to Haak et al.[16]. These data are summarised in Table S1.

The consensus sequences (with a minimum coverage of 3x) and SNPs according to rCRS[64] and RSRS[65] (with a minimum variant frequency of 0.7 and minimum coverage of 5x) were evaluated by Geneious 8.1.7 software (https://www.geneious.com/). The haplogroups were determined using HaploGrep (v2.1.1) (https://haplogrep.uibk.ac.at/) based on *phylotree* 17[66].

Population genetic analyses

Standard statistical methods were used for the calculation of genetic distances between the investigated Avar elite population and Eurasian ancient and modern populations.

We excluded sample AC20 from any statistical and phylogenetic analyses because of the large number of haplogroup-diagnostic positions missing. Furthermore, we excluded sample HC9 from population-genetic statistical analyses because it belongs to a later period (end of 7th – early 9th centuries), and also excluded sample RC26 from sequence-based analyses because of a large number of unreadable and missing parts of the mitochondrial sequence, which inhibit the haplotype-based calculation of genetic distances.

The whole mitochondrial genomes of the samples were aligned in SeaView[67] by ClustalO with default options. Positions with poor alignment quality were discarded in the case of ancient and modern sequences as well.

Population pairwise F_{ST} values were calculated based on 3,945 modern-day and 932 ancient whole mitochondrial sequences using Arlequin 3.5.2.2[68]. The Tamura and Nei substitution model was used[69] with a gamma value of 0.62, 10,000 permutations and significance level of 0.05 in case of comparison between the investigated Avar elite population and 43 modern-day Eurasian populations (for the references see Table S5). For the comparison of 23 ancient populations, the Tamura and Nei model was performed with a gamma value of 0.599, 10,000 permutations and significance level of 0.05. The number of usable loci for distance

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

computation in this case was 13,634 because 2,934 np had too much missing data (for the references see Table S4). The genetic distances of linearised Slatkin F_{ST} values[70] were used for Multidimensional scaling (MDS) and visualised on two-dimensional plots (Figs. 4-5) using the metaMDS function based on Euclidean distances implemented in the vegan library of R 3.4.1[71].

Principal component analyses (PCA) were performed based on mtDNA haplogroup frequencies of 64 modern and 44 ancient populations. Thirty-two mitochondrial haplogroups were considered in the PCA of ancient populations, while 36 mitochondrial haplogroups in the PCA of modern populations were considered (Tables S2 and S3). The PCAs were carried out using the prcomp function in R 3.4.1 and visualised in two-dimensional plots displaying the first two (PC1 and PC2) or the first and third principal components (PC1 and PC3) (Figs. S3a-b and S4a-b).

For hierarchical clustering, Ward type algorithm[72] and Euclidean distance measurement method were used based on haplogroup frequencies of ancient and modern populations and displayed as a dendrogram in R3.4.1 (Figs 3 and S5). The same population-pools were used for this clustering as those used in the two PCAs.

Phylogenetic analysis

Phylogenetic analyses aimed to detect close maternal relative lineages within the group of samples belonging to a certain haplogroup. All available human mitochondrial genome sequences in NCBI (more than 33,500) were downloaded and sorted according to their haplogroup assignments. Multiple sequence alignment was performed for each sample set using the same procedure mentioned in the Population genetic analyses section, with an exception that only the 303-318 sites were discarded on this highly repetitive and indel prone region due to poor alignment quality. Then neighbour joining trees were calculated using the *dnadist* and *neighbor* subprograms of Phylip version 3.696[73] with default options. The Median Joining Network, which is a favoured method for analysing haplotype data, was rejected due to unresolvable ties. The trees were drawn in Figtree version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree). We did not use bootstrap analyses due to the low quantity of informative positions, which highly biases the supporting values.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Golden, P. B. *An Introduction to the History of the Turkic Peoples: Ethnogenesis and State Formation in Medieval and Early Modern Eurasia and the Middle East. Turcologica 9* (Harrassowitz, O., 1992).

2. Pohl, W. *Die Awaren. Ein Steppenvolk in Mitteleuropa 567-822 n. Chr.* (2002).

3. Daim, F. in *Regna and Gentes: The Relationship between Late Antique and Early Medieval Peoples and Kingdoms in the Transformation of the Roman World* (eds. Goetz, H.-W., Jarnut, J. & Pohl, W.) 463–570 (2003).

4. Barfield, T. J. in *Empires* (ed. Susan E. Alcock, Terence N. D'Altroy, Kathleen D. Morrison, C. M. S.) 11–41. (Cambridge University Press, 2001).

5. Di Cosmo, N. & Wyatt, D. J. *Political Frontiers, Ethnic Boundaries and Human Geographies in Chinese History*. (Routledge Curzon, 2003).

6. Kradin, N. N. From tribal confederation to empire: the evolution of the Rouran society. *Acta Orient. Acad. Sci. Hungariae* **58,** 149–169 (2005).

7. Крадин, Н. Н. *Кочевники Евразии.* (Дайк-пресс, 2007).

8. Tomka, P. Die Lehre der Bestattungsbräuche. *Antaeus Commun. ex Inst. Archaeol. Acad. Sci. Hungaricae* **29/30,** 233–263 (2008).

9. Vida, T. in *The Other Europe in the Middle Ages: Avars, Bulgars, Khazars and Cumans. East Central and Eastern Europe in the Middle Ages, 450-1450, 2.* (ed. Curta, F.) 13–46 (Brill, 2008).

10. Sinor, D. *The Cambridge History of Early Inner Asia.* (Cambridge University Press, 1990).

11. Dobrovits, M. in *Ērān ud Anērān. Studies presented to Boris Ilich Marshak on the Occasion of His 70th Birthday.* (Universidad del Salvador, 2003).

12. Golden, P. B. in *The Steppe Lands and the World beyond them. Studies in honor of Victor Spinei on his 70th birthday.* (eds. Curta, F. & Maleon, B.-P.) 43–66. (2013).

13. Vovin, A. in *Ötüken'den İstanbul'a Türkçenin 1290 Yılı (720-2010) Sempozyumu/From Ötüken to Istanbul, 1290 Years of Turkish (720-2010) 3-5 Aralık 2010 İstanbul / 3th-5th December 2010, İ* (eds. Ölmez, M., Aydın, E., Zieme, P. & Kaçalin, M. S.) 1–10 (2011).

14. Bálint, C. *Die Archäologie der Steppe. Steppenvölker zwischen Volga und Donau vom 6. bis zum 10. Jahrhundert*. (1989).

15. Curta, F. in *The Other Europe in the Middle Ages. Avars, Bulgars, Khazars and Cumans. ECEE 2.* (ed. Curta, F.) 297–325. (Brill, 2008).

16. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522,** 207–211 (2015).

17. Damgaard, P. de B. *et al.* 137 ancient human genomes from across the Eurasian steppes. *Nature* 1 (2018). doi:10.1038/s41586-018-0094-2

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

18. Alt, K. W. *et al.* Lombards on the move - An Integrative study of the migration period cemetery at Szólád, Hungary. *PLoS One* **9,** e110793 (2014).

19. Amorim, C. E. G. *et al.* Understanding 6th-century barbarian social organization and migration through paleogenomics. *bioRxiv* (2018). at <http://biorxiv.org/content/early/2018/02/20/268250.abstract>

20. Csősz, A. *et al.* Maternal genetic ancestry and legacy of 10th century AD Hungarians. *Sci. Rep.* **6,** 33446 (2016).

21. Šebest, L. *et al.* Detection of mitochondrial haplogroups in a small Avar-Slavic population from the eigth-ninth century AD. *Am. J. Phys. Anthropol.* **165,** 536–553 (2018).

22. Lipták, P. *Avars and Ancient Hungarians.* (1983).

23. Fóthi, E. Összehasonlító antropológiai vizsgálat a Kárpát-medence népeinek etnogeneziséhez. Embertani kapcsolatok a 6.-8. században az eurázsiai steppe és a Kárpát-medence között. *Móra Ferenc Múzeum Évkönyve* **IV,** 497– 515. (1998).

24. H. Tóth, E. & Horváth, A. *Kunbábony. Das Grab eines Awarenkhagans*. (Museumsdirektion der Selbstverwaltung des Komitats Bács-Kiskun, 1992).

25. Csiky, G. in *East-Central Europe in the Middle Ages, 450-1450 34. Leiden – Boston 2015.* (2015).

26. Vida, T. in *Between Byzantium and the Steppe: Archaeological and Historical Studies in Honour of Csanád Bálint on the Occasion of His 70th Birthday* (eds. Bollók, Á., Csiky, G. & Vida, T.) 51–70 (MTA BTK Régészeti Intézet, 2016).

27. Crubézy, E. *et al.* Human evolution in Siberia: from frozen bodies to ancient DNA. *BMC Evol. Biol.* **10,** 25 (2010).

28. Molodin, V. I. *et al.* in *Population Dynamics in Prehistory and Early History: New Approaches Using Stable Isotopes and Genetics* 93–112 (2012). doi:10.1515/9783110266306.93

29. Keyser-Tracqui, C., Crubézy, E. & Ludes, B. Nuclear and mitochondrial DNA analysis of a 2,000-year-old necropolis in the Egyin Gol Valley of Mongolia. *Am. J. Hum. Genet.* **73,** 247–60 (2003).

30. Kim, K. *et al.* A western Eurasian male is found in 2000-year-old elite Xiongnu cemetery in Northeast Mongolia. *Am. J. Phys. Anthropol.* **142,** 429–440 (2010).

31. Pilipenko, A., Romaschenko, A., Molodin, V., Parzinger, H. & Kobzev, V. Mitochondrial DNA studies of the Pazyryk people (4th to 3rd centuries BC) from northwestern Mongolia. *Archaeol Anthr. Sci* **2,** 231–236 (2010).

32. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522,** 167–172 (2015).

33. González-Ruiz, M. *et al.* Tracing the origin of the east-west population admixture in the Altai region (Central Asia). *PLoS One* **7,** e48904 (2012).

34. Ricaut, F. X., Keyser-Tracqui, C., Bourgeois, J., Crubézy, E. & Ludes, B. Genetic

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

analysis of a Scytho-Siberian skeleton and its implications for ancient Central Asian migrations. *Hum. Biol.* **76,** 109–25 (2004).

35. Lippold, S. *et al.* Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. *Investig. Genet.* **5,** 13 (2014).

36. Zheng, H. X. *et al.* Major population expansion of East Asians began before Neolithic time: Evidence of mtDNA genomes. *PLoS One* **6,** (2011).

37. Derenko, M., Denisova, G., Malyarchuk, B., Dambueva, I. & Bazarov, B. Mitogenomic diversity and differentiation of the Buryats. *J. Hum. Genet.* **63,** 71–81 (2017).

38. Yunusbayev, B. *et al.* The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLOS Genet.* **11,** e1005068 (2015).

39. Zheng, H.-X. *et al.* MtDNA genomes reveal a relaxation of selective constraints in low-BMI individuals in a Uyghur population. *Hum. Genet.* **136,** 1353–1362 (2017).

40. De Nicola, B. *Women in Mongol Iran. The Khatuns, 1206-1335.* (Edinburgh University Press, 2017).

41. Kradin, N. The social system of the nomads: discussion and problems. *Quest. Hist.* **4,** 21–32 (2001).

42. Drompp, M. From Qatun to Refugee. The Taihe Princess among the Uighurs. in *The Role of Women in the Altaic World Permanent International Altaistic Conference 44th Meeting, Walberberg, 26–31 August 2001. Asiatische Forschungen 152.* (ed. Veit, V.) (Harrassowitz Verlag, 2007).

43. Neparáczki, E. *et al.* Revising mtDNA haplotypes of the ancient Hungarian conquerors with next generation sequencing. *PLoS One* **12,** 1–11 (2017).

44. Ochir, A. & Ankhbayar, B. in *Монголын эртний булш оршуулга. – Ancient Funeral Monuments of Mongolia III. Ulaanbaatar* (eds. Eregzen, G. & Эрэгзэн, Г.) 190–195 (2016).

45. Yongzhi, C., Guodong, S. & Yan, M. The Results of the Excavation of the Yihe-Nur Cemetery in Zhengxiangbai Banner (2012-2014). *Silk Road* **14,** 42–57. (2016).

46. Stark, S. Central and Inner Asian parallels to a find from Kunszentmiklós-Bábony (Kunbábony): Some thoughts on the Early Avar headdress. *Anc. Civilisations from Scythia to Sib.* **15,** 287–305. (2009).

47. Мамадаков, Ю. Т., Кунгуров, А. Л., Чекурашев, А. Г. & Барзынов, Я. Р. Воинский схрон тюркского времени на горе Ак-Кая (Онгудайский район, Республика Алтай). *Теория и Практика археологических исследований* **18,** 30–40. (2017).

48. Vida, T. in *Awarenforschungen 1.* (ed. Daim, F.) 517–557 (1992).

49. Серегин, Н. Н. Появление и первоначальное распространение петельчатых стремян (хронологический и этнокультурный аспекты). *Известия Алтайского Государственного Университета* **80,** 194–198. (2013).

50. Азбелев, П. П. Ещё раз о ранних стременах. *Российский Археологический*

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ежегодник **4,** (2004).

51. Graff, D. *The Eurasian way of war: Military practice in seventh-century China and Byzantium.* (Routledge, 2016).

52. Евтюхова, Л. А. in *По следам древних культур. От Волги до Тихого океана.* 195–224 (1954).

53. Киселев, С. В. *Древняя история Южной Сибири*. (1951).

54. Lipson, M. *et al.* Parallel paleogenomic transects reveal complex genetic history of early European farmers. *Nature* **551,** 368–372 (2017).

55. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A.* **110,** 15758–63 (2013).

56. 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).

57. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* (2010). doi:10.1101/pdb.prot5448

58. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10 (2011).

59. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26,** 589–595 (2010).

60. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

61. McKenna, A. *et al.* The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20,** 1297–1303 (2010).

62. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43,** 491–498 (2011).

63. Van der Auwera, G. A. *et al.* in *Current Protocols in Bioinformatics* **43,** 11.10.1–11.10.33 (John Wiley & Sons, Inc., 2013).

64. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23,** 147 (1999).

65. Behar, D. M. *et al.* A 'Copernican' reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90,** 675–84 (2012).

66. van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30,** E386–94 (2009).

67. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27,** 221–224 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

68. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10,** 564–567 (2010).

69. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10,** 512–26 (1993).

70. Slatkin, M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139,** 457–62 (1995).

71. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.r–project.org/ (2017).

72. Ward, J. H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58,** 236–244 (1963).

73. Felsenstein, J. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* **5,** 164–166. (1989).

74. Koch, A. Überlegungen zum Transfer von Schwerttrag- und -kampfesweise im frühen Mittelalter am Beispiel chinesischer Schwerter mit p-förmigen Tragriemenhaltern aus dem 6.-8. Jahrhundert n. Chr. *Jahrb. des Römisch-Germanischen Zentralmuseums Mainz* **45,** 571–598 (1999).

75. Ito, A. *Zur Chronologie der frühsillazeitlichen Gräber in Südkorea*. *Abhandlungen der Bayerischen Akademie der Wissenschaften, Philosophisch-historische Klasse; Neue Folge* **71,** (1971).

76. Al'Baum, L. I. *Zhivopis' Afrasiaba*. (1975).

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

This study was funded by a Hungarian research grant of T.V. (NKFIH-OTKA NN 113157) based on Eötvös Loránd University, Budapest. We thank Ildikó Pap, Ágnes Kustár (Hungarian Natural History Museum, Budapest), the co-workers of the Tessedik Samuel Museum (Szarvas) and the Katona József Museum (Kecskemét) for providing DNA samples. We are grateful to Csilla Balogh (Istanbul Medeniyet University, Istambul), and Erika Wiker (Katona József Museum, Kecskemét) for providing archaeological description and information about the studied sites. We thank Levente Samu and Viktor Szinyei for preparing the base maps.

Author contributions

I. K, B.G.M, A.S-N, T.V. designed the study. V.Cs. and G.D. performed the ancient DNA analyses. V.Cs. G.D. and A.S-.N, performed population genetic analyses. I.K., G.Cs., A. G., B. K. G. M. L., G. L. and T.V. performed the archaeological analysis and provided the historical background and interpretation. A. M, B.G.M., E. M. and Gy. P. summarised the anthropological data of the human remains. V. Cs., D.G., I.K, G.Cs., B.G.M, A. S-N, T.V wrote the paper. All authors read and discussed the manuscript.

Competing Interests

The authors declare no competing financial interests

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures

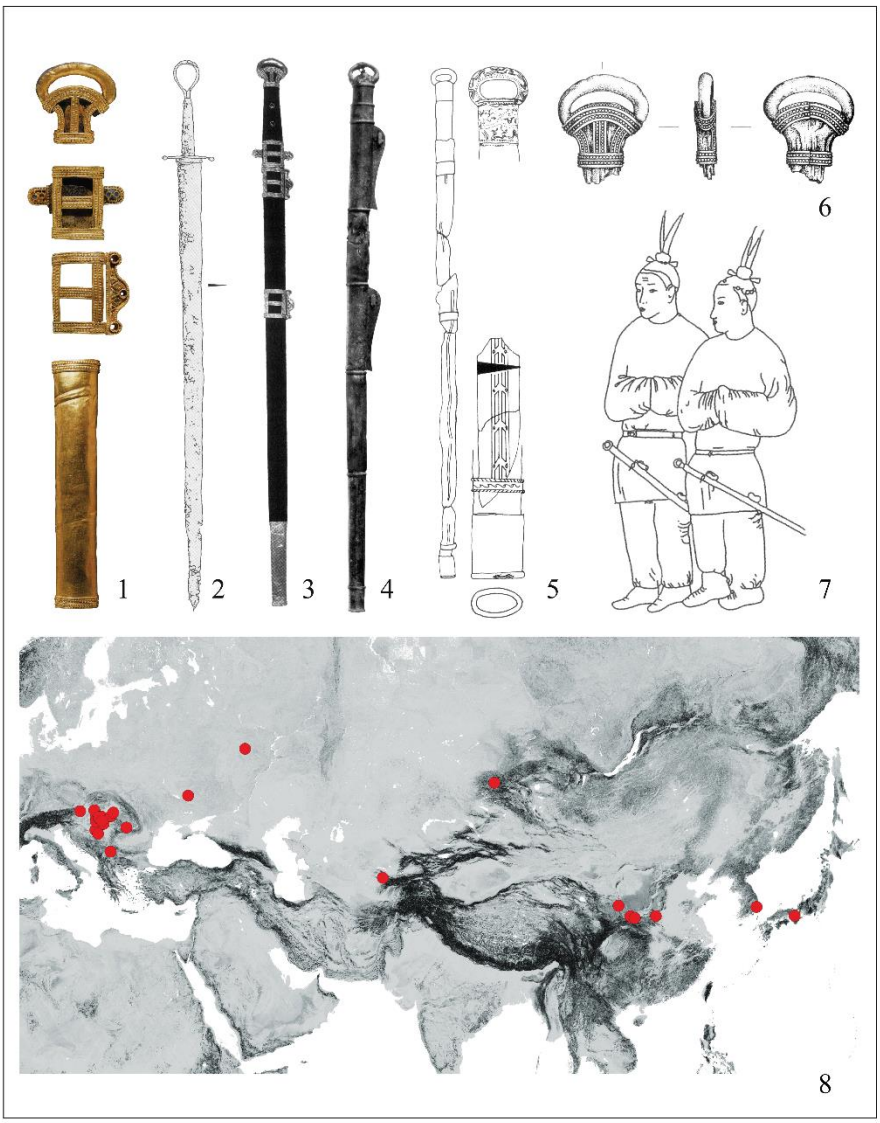

Figure 1. Ring-pommel swords.
1. Kunbábony[24], 2. Sopron-téglagyári agyagbánya[25], 3. Bócsa[25], 4. Miniature sword from the burial of Emperor Wu of Northern Zhou, Xian-yang, Xi'an, Shaanxi province[74], 5. Gyeongju, South Korea, grave No. 126[75] 6. Bócsa[25], 7. Afrasiab 1. building: representation of Korean envoys (after Al'baum 1975[76] and Csiky 2015[25]), 8. Distribution of the ring-pommel swords[25]. Map is based on the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). Data are available from the U.S. Geological Survey (https://lta.cr.usgs.gov/GMTED2010) and downloaded on 11.05.2018 by L. Samu.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

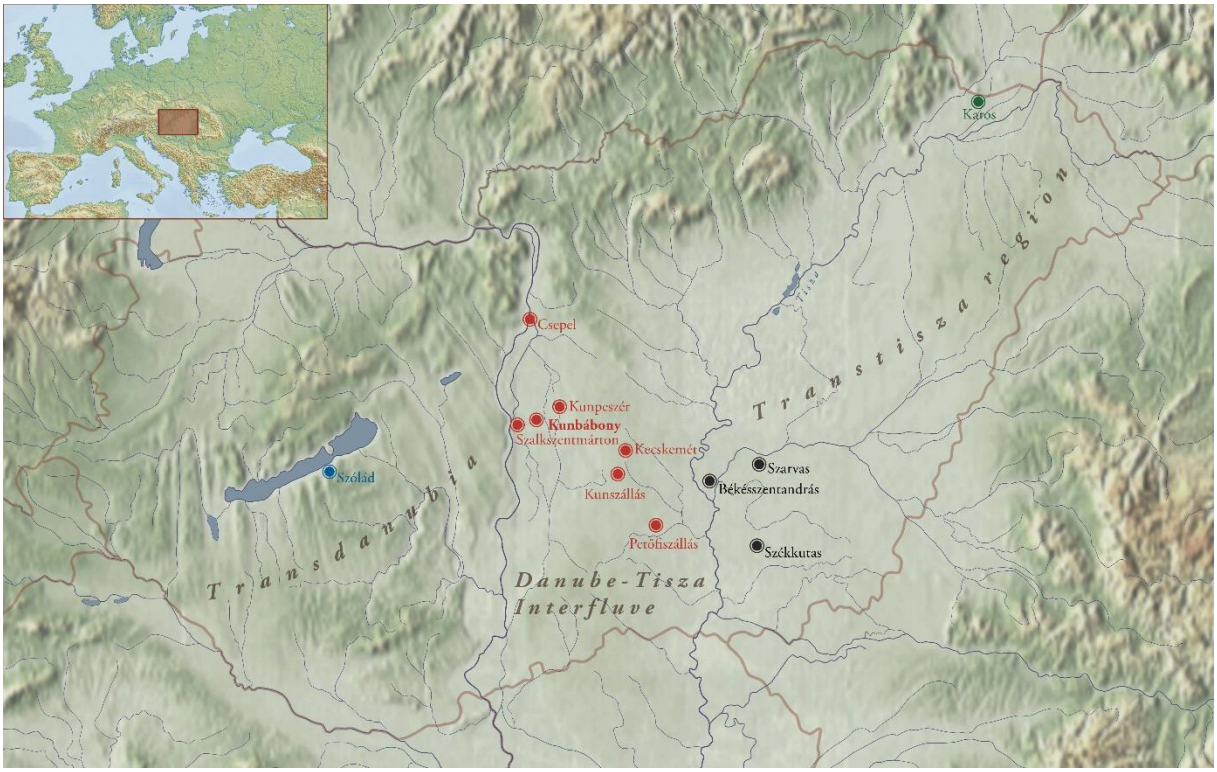

Figure 2. Location of the samples and investigated sites in the Carpathian Basin (Hungary).

The investigated sites of the Kunbábony group (7th century) are marked with red, 7th century supplementary sites are marked with black. Szólád (blue) and Karos (green) are comparative materials from the 6th and 9th-10th centuries. (Map was made by V. Szinyei IA RCH HAS. The map of Europe was downloaded from MAPSWIRE https://mapswire.com/europe/physical-maps/).

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

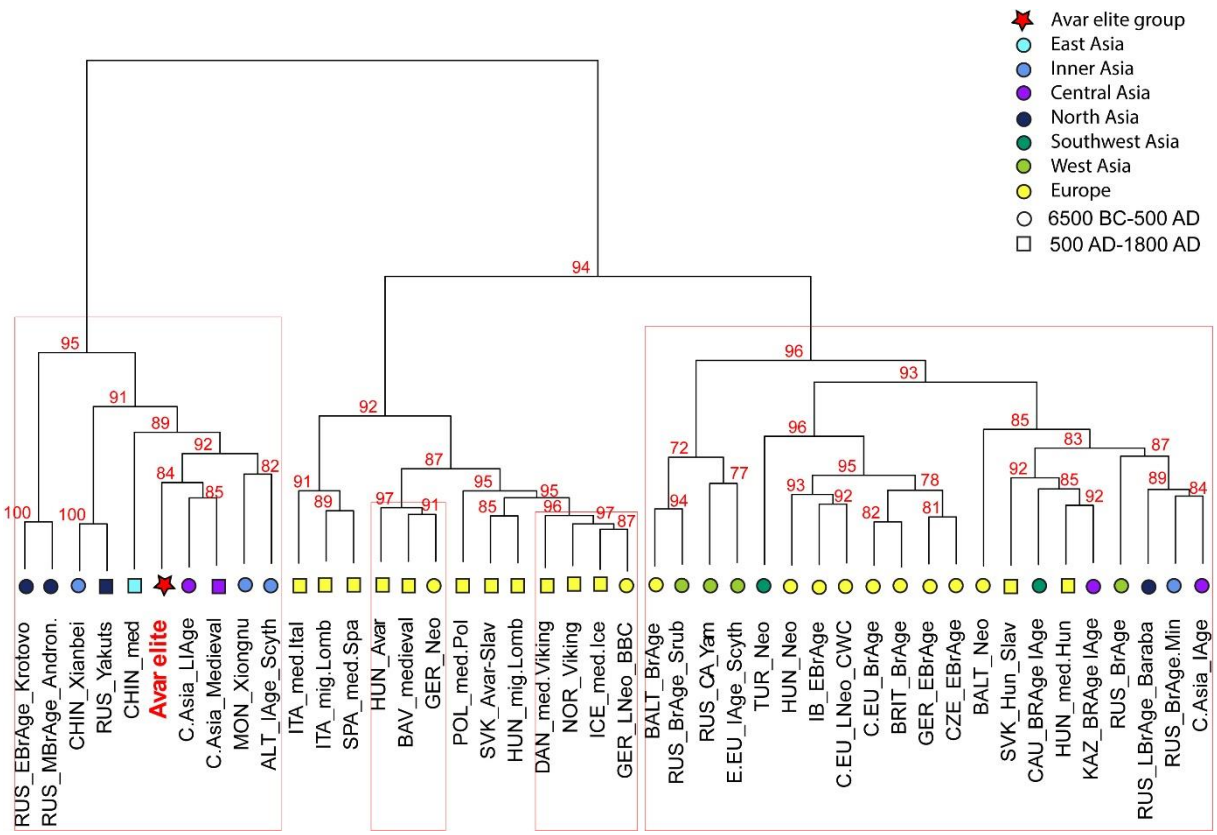

Figure 3. Ward type clustering of 44 ancient populations.

The Ward type clustering shows separation of Asian and European populations. The Avar elite group (AVAR) is situated on an Asian branch and clustered together with Central Asian populations from Late Iron Age (C-ASIA_LIAge) and Medieval period (C-ASIA_Medieval), furthermore with Xiongnu period population from Mongolia (MON_Xiongnu) and Scythians from the Altai region (E-EU_IAge_Scyth). P values are given in percent as red numbers on the dendogram, where red rectangles indicate clusters with significant p values. The abbreviations and references are presented in Table S2.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

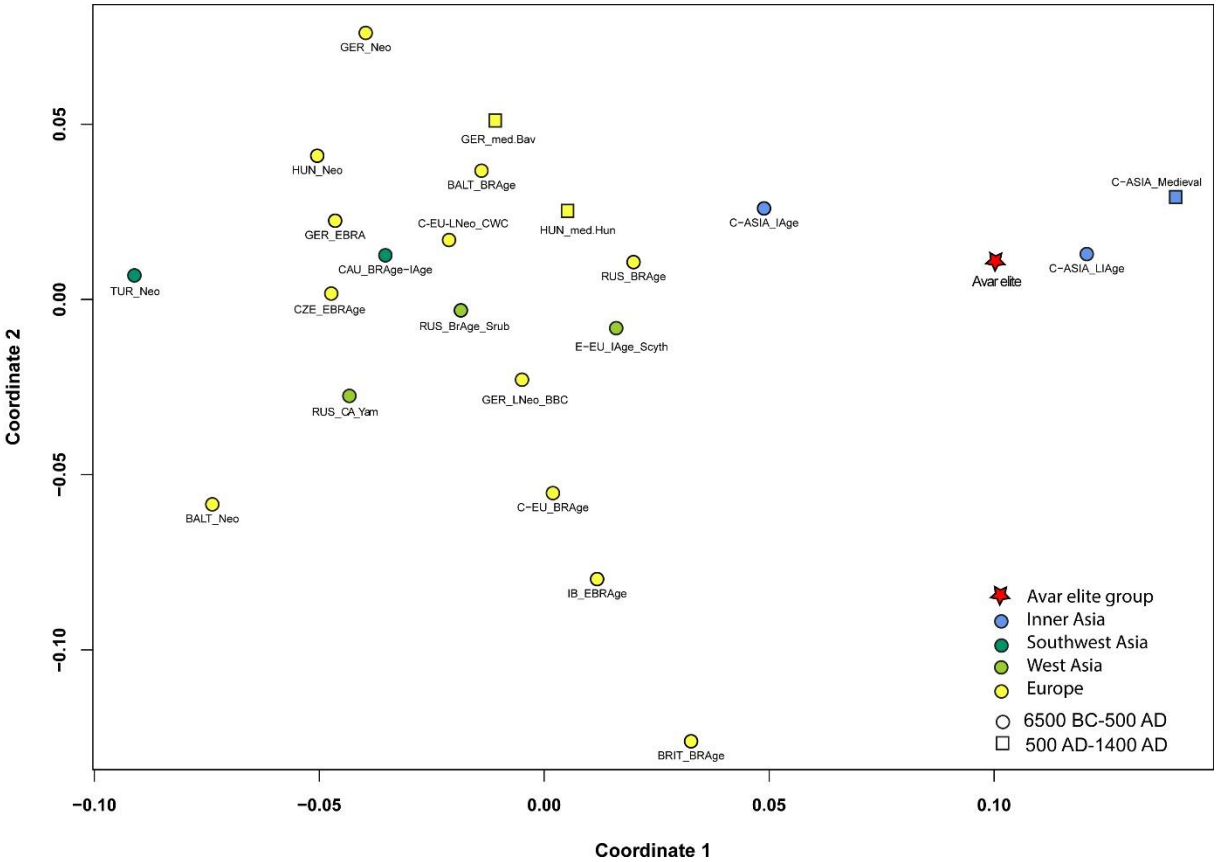

Figure 4. MDS with 23 ancient populations.

The Multidimensional Scaling plot is based on linearised Slatkin F_{ST} values that were calculated based on whole mitochondrial sequences (stress value is 0.1581). The MDS plot shows the connection of the Avars (AVAR) to the Central-Asian populations of the Late Iron Age (C-ASIA_LIAge) and Medieval period (C-ASIA_Medieval) along coordinate 1 and coordinate 2, which is caused by non-significant genetic distances between these populations. The European ancient populations are situated on the left part of the plot, where the Iberian (IB_EBRAge), Central-European (C-EU_BRAge) and British (BRIT_BRAge) populations from Early Bronze Age and Bronze Age are clustered along coordinate 2, while the Neolithic populations from Germany (GER_Neo), Hungary (HUN_Neo), Near-East (TUR_ _Neo) and Baltic region (BALT_Neo) are located on the skirt of the plot along coordinate 1. The linearised Slatkin F_{ST} values, abbreviations and references are presented in Table S4.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

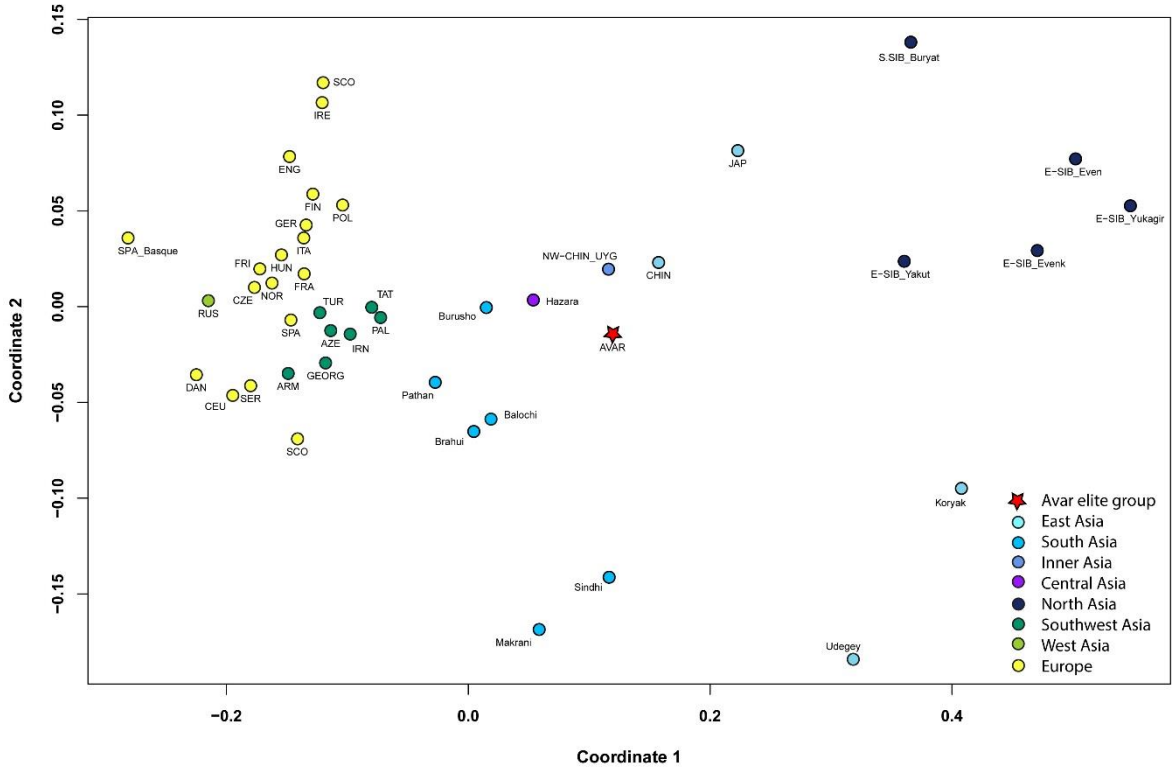

Figure 5. MDS with the 44 modern populations and the Avar elite group

The Multidimensional Scaling plot is displayed based on linearised Slatkin F_{ST} values calculated based on whole mitochondrial sequences (stress value is 0.0677). The MDS plot shows differentiation of European, Near-Eastern, Central- and East-Asian populations along coordinates 1 and 2. The Avar elite (AVAR) is located on the Asian part of plot and clustered with Uyghurs from Northwest-China (NW-CHIN_UYG) and Han Chinese (CHIN), as well as with Burusho and Hazara populations from the Central-Asian Highland (Pakistan). The linearised Slatkin F_{ST} values, abbreviations and references are presented in Table S5.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

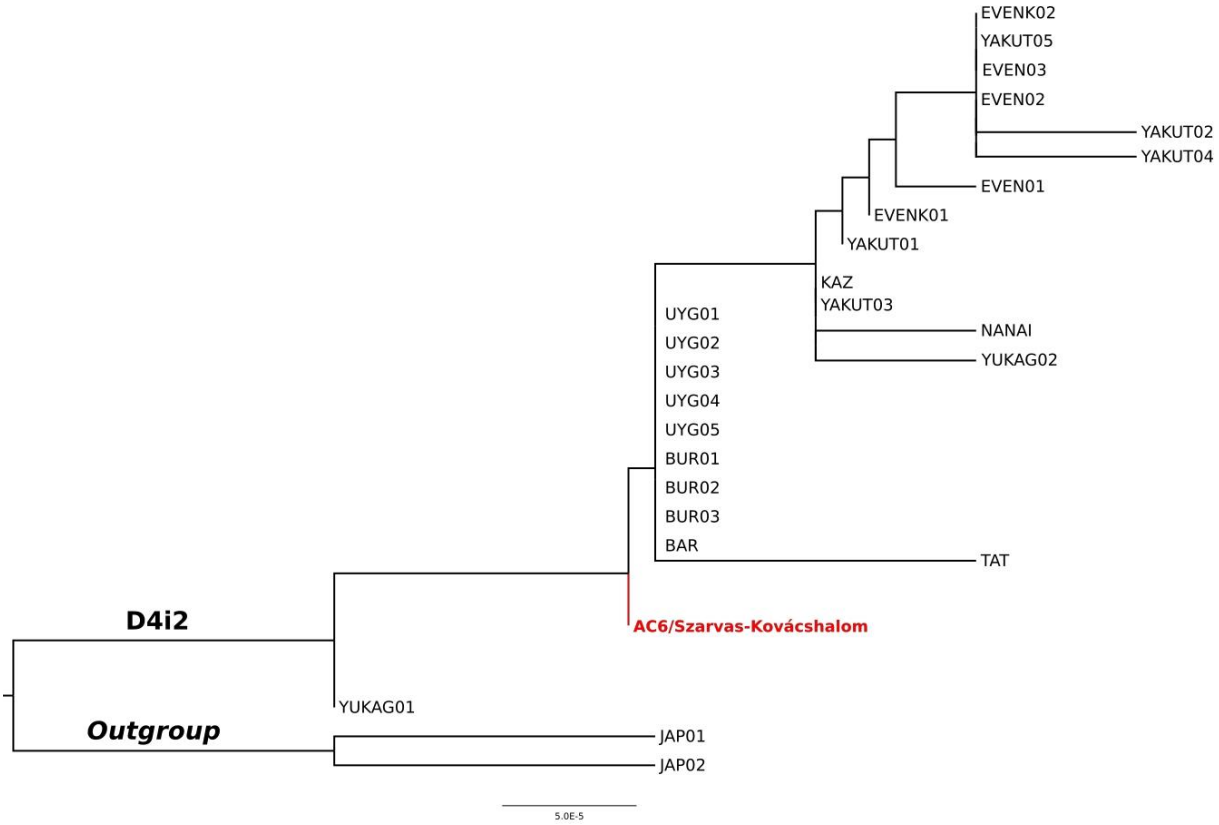

Figure 6. Phylogenetic tree of D4i2 sub-haplogroup.

Phylogenetic tree of D4i2 sub-haplogroup shows AC6 to be the mitochondrial founder of most of the other D4i2 lineages from in Inner Asia and North Asia, which indicates a close shared maternal ancestry between the populations represented by these individuals.

The references of individuals displayed on the tree are presented in Table S6.

bioRxiv preprint doi: https://doi.org/10.1101/415760; this version posted September 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

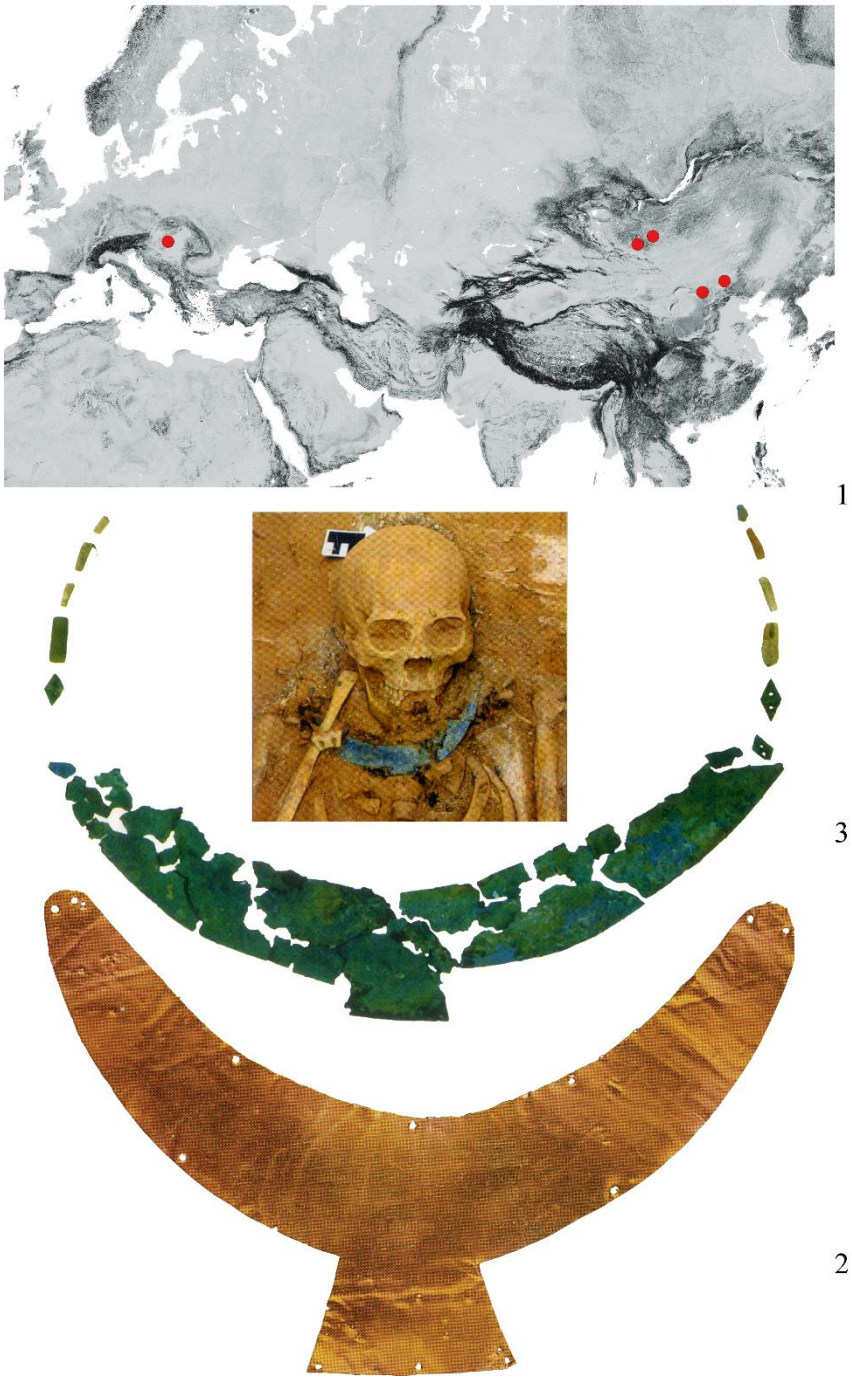

Figure 7. Crescent-shaped gold sheets used as pectorals in graves of the nomadic elite.
1. Distribution of the crescent-shaped gold sheets used as pectorals from East to West: Yihe Nur[45], Bikeqi[46]; Talyn Gurvan Kherem (with photos in figure part 3)[44], Galuut sum[46], Kunbábony (pictured in figure part 2)[24]. Basic map is based on the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). Data is available from the U.S. Geological Survey (https://lta.cr.usgs.gov/GMTED2010) and downloaded at 11.05.2018 by L. Samu.