Letter to the Editor

Ancient genome of Empress Ashina reveals the Northeast Asian origin of Göktürk Khanate

Xiao-Min Yang[1,2†], Hai-Liang Meng[1†], Jian-Lin Zhang[3], Yao Yu[4], Edward Allen[4], Zi-Yang Xia[5], Kong-Yang Zhu[6], Pan-Xin Du[1], Xiao-Ying Ren[4], Jian-Xue Xiong[4], Xiao-Yu Lu[4], Yi Ding[4], Sheng Han[7], Wei-Peng Liu[8], Li Jin[1], Chuan-Chao Wang[2,6,9,10]* 🔘, and Shao-Qing Wen[4,11]* 🔘

[1]Ministry of Education Key Laboratory of Contemporary Anthropology, Department of Anthropology and Human Genetics, School of Life Sciences, Fudan University, Shanghai 200433, China

[2]Department of Anthropology and Ethnology, School of Sociology and Anthropology, Institute of Anthropology, Xiamen University, Xiamen 361005, Fujian, China

[3]Shaanxi Academy of Archaeology, Xi'an 710054, China

[4]Institute of Archaeological Science, Fudan University, Shanghai 200433, China

[5]Division of Biosciences, and Department of Genetics, Evolution and Environment, University College London Genetics Institute (UGI), University College London, London WC1E 6BT, UK

[6]State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Xiamen 361005, Fujian, China

[7]Department of History, Fudan University, Shanghai 200433, China

[8]College of Cultural Heritage, Northwest University, Xi'an 710069, China

[9]State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen 361005, Fujian, China

[10]Institute of Artificial Intelligence, Xiamen University, Xiamen 361005, Fujian, China

[11]Center for the Belt and Road Archaeology and Ancient Civilizations (BRAAC), Fudan University, Shanghai 200433, China

[†]These authors contributed equally to this study.

*Authors for correspondence. Chuan-Chao Wang. E-mail: wang@xmu.edu.cn; Shao-Qing Wen. E-mail: wenshaoqing@fudan.edu.cn

The first flourishing of the Türkic peoples took place from the 6th to 8th century with the rise of the powerful pastoral nomadic Göktürk Khanate, which was established by the Ashina clan in the region ranging from the Mongolia Plateau to the Caspian Sea (Grousset, 1970). Arriving in the wake of the Xiongnu, Xianbei and other nomadic confederations, the impact of the Türks was felt across Eurasia and would ultimately impact much of the West Eurasian ethnolinguistic over the following millennium and centuries. Unfortunately, historical records of the Göktürk Khanate are sporadic and frequently inconsistent. While the Türks no longer existed as a unified regime following the Göktürk Khanate's collapse in the mid-8th century, Turkic itself has survived as a linguistic subfamily despite prolonged contact and admixture with other Eurasian languages. Comprised of over 40 languages, Turkic is the largest group in Altaic linguistic family, and 170 million people speak Turkic languages in over 10 countries.

As the ruling tribe of Göktürk Khanate, understanding the Ashina tribe is key to unravelling the mysterious origins of the Türks. Göktürk origins are a contested subject in the relevant Chinese historical records—*Zhoushu* (*Book of Zhou*, 周书), *Beishi* (*History of Northern Dynasties*, 北史), *Suishu* (*Book of Sui*, 隋书), and *Tongdian* (*Comprehensive Manual*, 通典) —but may be summarized in three competing hypotheses: (i) a derivation for Ashina from Xiongnu tribes originating in the Northeast Asian region; (ii) origins in the Pontic-Caspian steppe ("west of the Caspian Sea [Rui, 1991]," 西海之右) or east-central Asia ("Sogdian statelet", 索国) followed by eastward migration; and (iii) multiple origins around Pingliang (平凉) or Gaochang/Turfan (高昌) in northwest China, a process involving both eastern and western Eurasian ethnic groups.

Since cremation was commonly practiced among the Turkic nobility, most known Turkic cemeteries do not leave us with any such skeletal remains. We, fortunately, identified the skeletal remains of Empress Ashina at the Xiaoling Mausoleum (Fig. 1A) in modern-day Dizhang Town, Xianyang City, Shaanxi. As documented in detail in the *Zhoushu* (Volume 9), Empress Ashina (551–582 CE) was the daughter of the Göktürk Muqan Khagan (*Ashina Qijin*, 阿史那俟斤) and married into royalty through her betrothal to Northern Zhou Dynasty Emperor Wu (*Yuwen Yong*, 宇文邕). When the Empress passed away at age 32, Empress Ashina was buried with honors alongside her husband at the Xiaoling Mausoleum. Cultural relics unearthed at this site include the *Xiaolingzhi* (*Epitaph of Xiaoling*, 孝陵志), *Wude Huanghouzhi* (*Epitaph of Empress Wude*, 武德皇后志), a seal belonging to the Empress Dowager Tianyuan (天元皇太后玺)
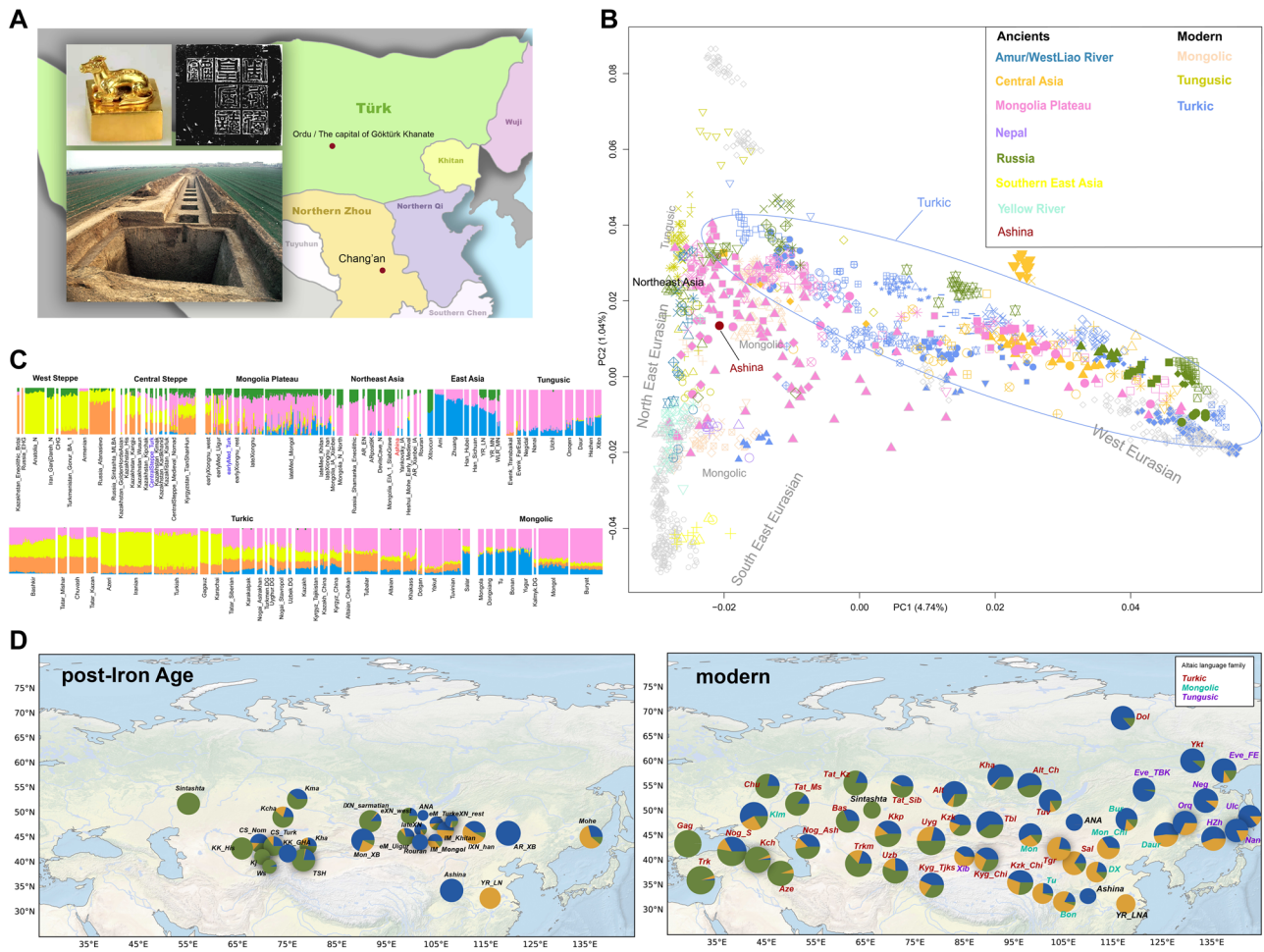
Fig. 1. Burial relics excavated from the Xiaoling Mausoleum and Population structure of newly sampled and published populations in Eurasia. **A**, Maps of ancient regimes in 572 CE (Tan, 1989) and the cemetery of the newly sampled ancient individual, *Epitaph of Empress Wude*, and *Seal of the Empress dowager Tianyuan*. **B**, Principal component analysis (PCA) of ancient individuals projected onto present-day Eurasians. The west–east cline along PC1 and north–south gradient along PC2 were visible (also see Fig. S3). **C**, Unsupervised ADMIXTURE clustering analysis ($K = 5$) was based on HO data set for the ancient individual and selected subset of temporally preceding and later Eurasian ancient and modern populations. **D**, Ancestral composition of post-Iron Age Eastern/Central Steppe pastoralists and modern Altaic-speaking populations based on supervised ADMIXTURE. Mongolia_N_North (labeled as ANA in figure) and Ashina were chosen as ANA ancestry, Russia_Sintashta_MLBA was used as West Eurasian-related ancestry, YR_LN was regarded as the additional millet farmers of Central Plain of China. The diverse ancestral compositions of post-Iron Age Eastern/Central Steppe pastoralists showed a high proportion of East Eurasian ancestry in Eastern Steppe nomadic populations with the exception of earlyXiongnu_west and lateXiongnu_Sarmatian, and West Eurasian ancestry dominant in Central Steppe nomadic populations. Present-day Turkic populations exhibited high genetic heterogeneity with diverse proportions of ANA, and the ANA ancestry decreased almost longitudinally.

(Fig. 1A), a bronze bucket and a small number of ornaments. Skeletal remains in this high-rank cemetery were, on the whole, very poorly preserved. We sampled the limb bones from the female individual—Empress Ashina. We have carried out the C14 dating for the Empress Ashina sample. The sample was dated to 564–650 cal AD (1386–1300 cal BP), which is well consistent with the period that Empress Ashina lived (Supplementary Section 2 Material and Methods). The Empress Ashina sample is invaluable for tracing the origins of the Göktürk Khanate.

Previous genetic studies have mostly focused on western Turkic peoples (Yunusbayev et al., 2015; Triska et al., 2017), while core tribes of Göktürk Khanate, such as the Ashina, have been underrepresented. In this study, we attempt to answer three questions through a dissection of Empress Ashina's genetic profile: (i) the ancestral origins of the Göktürk; (ii) genetic relationships between Ashina and other Türkic people and post-Iron Age Central/Eastern Steppe pastoralists; and (iii) the genetic relationship between ancient Göktürk and modern Turkic-speaking populations.

We produced genome-wide sequence data with a coverage of 0.2 x for Empress Ashina (Table S1A). All libraries displayed typical damage patterns of ancient DNA (Renaud et al., 2015) (Fig. S2). The sample featured a low level of modern mitochondrial DNA (mtDNA) contamination (<1%) and no evidence of autosomal contamination (Renaud et al., 2015; Nakatsuka et al., 2020) (Table S1A). The sample was evaluated as female. Her mtDNA belonged to the F1d haplogroup (Weissensteiner et al., 2016), prevalent in Northeast Asia (Table S1B). We merged the ancient data with the published data set for subsequent analyses (Damgaard et al., 2018a, 2018b; Jeong et al., 2019, 2020; Narasimhan et al., 2019; Ning et al., 2020; Yang et al., 2020; Mao et al., 2021; Wang et al., 2021).

1 The Northeast Asia origin of the royal tribe of the Göktürk Khanate

In the principal component analysis (PCA) (Figs. 1B, S3), the Ashina individual clustered with modern Tungusic and Mongolic speakers, ancient populations from Northeast Asia and eastern Mongolia Plateau, and especially with the Northeast Asian hunter-gatherers previously referred to as "Ancient Northeast Asian" (ANA), that is, DevilsCave_N, Mongolia_N_North, Boisman_MN, AR_EN (Jeong et al., 2020; Ning et al., 2020; Wang et al., 2021), as well as post-Iron Age Eastern Steppe nomadic people including Xianbei, Rouran, Khitan, and part of the Mongol population. The shared genetic similarity between Ashina and Northeast Eurasians, especially ANA, was also evident in outgroup-f_3 statistics (Fig. S5A). The pre-Bronze Age populations of the Amur River, hunter-gatherers of the Mongolia Plateau and Lake Baikal, and Late Bronze Age Ulaanzukh culture-related populations of Mongolia Plateau formed a clade together with Ashina, as reflected in f_4 (X, Mbuti; Ashina, pre-Iron Age Northeast Asian)~0 ($|Z| < 3$) (Fig. S6A). The genetic affinity between Ashina and pre-Bronze Age ANA was also displayed in positive f_4 (pre-Iron Age Northeast Asian, Mbuti; Ashina, X), in which X represented East Asians exclusive of millet farmer-related populations such as ancient Yellow River and Sino-Tibetan populations (Fig. S6B).

We further observed that Ashina had received additional genetic influence from West Eurasians. First, a model-based ADMIXTURE clustering showed a predominantly East Eurasian-related ancestry and subtle West Eurasian-related strain including Near East farmer and steppe-like ancestry (Figs. 1C, S4). Second, Ashina shared more alleles with West Eurasians than ancient East Eurasians with multiple comparisons showing positive f_4 (West Eurasians, Mbuti; Ashina, ancient East Eurasians) in most cases (Fig. S7). Furthermore, we estimated Ashina had 2.3%–3.9% of West Eurasian-related and 96.1%–97.7% of ANA-related ancestry using qpWave/qpAdm methods (Patterson et al., 2012) when Western Steppe pastoralist Afanasievo and AR_EN or Mongolia_N_North were used as source pairs (Table S2D). The West Eurasian admixture was dated to 1566 ± 396 years ago before the time of the individual when using Mongolia_N_North and Afanasievo as sources in DATES (Table S2F) (Narasimhan et al., 2019). We here used Mongolia_N_North and Afanasievo in DATES because they had larger sample sizes.

We further examined whether there was a gene flow from millet farmers into Ashina. We added Iron Age Yellow River farmers (YR_IA) into the outgroup set of the two-way admixture model with ANA and Afanasievo as sources (Table S2D). YR_IA included samples from the archaeological sites of Luoheguxiang, Jiaozuoniecun, Haojiatai, and Dacaozi, who were the successor of Neolithic Yellow River's millet farmers (Ning et al., 2020). We found that the two-way model still fit Ashina well ($P \geq 0.05$), demonstrating the absence of gene flow from millet farmers. We observed similar results in Eastern Steppe nomadic Rouran and Xianbei.

2 Genetic relationship between Ashina and other post-Iron Age Central/Eastern Steppe pastoralists

From the Iron Age onward, a succession of nomadic regimes rose and fell in the Eastern Steppe: the Xiongnu (the 3rd century BC–the 1st century CE), Xianbei (the 1st century CE—the 6th century CE), Rouran (the 4th century CE–the 6th century CE), Türk (the 6th century CE–the 8th century CE), Uyghur (744–840 CE), Khitan (916–1125 CE), and—most aggressively expansive and successful of any—the Mongol empire (1206–1368 CE). Adjacent to the Eastern Steppe, the Central Steppe also witnessed the shifting fortunes of nomadic regimes, with Hun (the 4th century CE–the 6th century CE), Wusun (the 2nd century BC–the 5th century BC), Kangju (140| BCE), Turkic Karluk (the 7th century CE–13th century CE), Kimak (the 8th century CE–the 11th century CE), Kara-Khanid (the 10th century CE–the 13th century CE), and Kipchak (the 11th century CE). The genetic relationship between the Türkic Khanate and these Eastern/Central Steppe nomadic populations has long been uncertain.

We found that Ashina was genetically different from Central Steppe pastoralists of the Hun, Wusun and Kangju period ($|Z|$ score of f_4(X, Mbuti; Ashina, post-Iron Age Central/Eastern Steppe) >3; Fig. S9). Xiongnu was a genetically heterogeneous population with some samples having received abundant genetic influence from ANA and Yellow River farmers (such as lateXiongnu_han; Table S2E) but others harboring closer connection with West Eurasian (sucha as early Xiongnu, lateXiongnu, and lateXiongnu_sarmatian; Fig. S10) (Jeong et al., 2020). Ashina showed a close relationship with lateXiongnu_han. The Ashina shared the most genetic affinity with Mongolic Rouran and Xianbei, followed by Khitan in outgroup-f_3 statistics (Table S2A). This was consistent with non-significant f_4 (X, Mbuti; Ashina, Rouran/Xianbei /Khitan/Heshui_Mohe) ($|Z| < 2$) and pairwise qpWave ($P > 0.01$) (Figs. S11, S12, S15), which indicated that Ashina formed a genetic clade together with Rouran, Xianbei, Khitan and Mohe. Among nomadic populations after the Türkic Khanate, Eastern Steppe Khitan and Mohe, and Central Steppe Kazakhstan_GoldenHordeAsian (an individual from Jochi Khan's Golden Horde army from the Ulytau mountains) showed genetic similarity with Ashina (non-significant Z-values in f_4 and $P > 0.01$ in pairwise qpWave; Figs. S12, S14, S15). By contrast, Uigur, Karluk, Kimak, Kipchak, Karakhanid and Mongols were genetically quite

Yang et al.

separated from the Ashina sample (significant Z-values in f_4 and $P < 0.01$ in pairwise $qpWave$; Figs. S12B, S13).

Different genetic relationships between Ashina and the post-Iron Age Eastern/Central Steppe nomadic populations were also evident in the ancestral compositions shown in ADMIXTURE, supervised ADMIXTURE and $qpAdm$ (Figs. 1C, 1D; Table S2E): among the Eastern Steppe pastoralists, some of the Xiongnu groups (earlyXiongnu_rest, lateXiongnu_Han, and lateXiongnu), Rouran, Xianbei, Khitan, Mongol and Heishui_Mohe harbored dominating East Eurasian ancestry from 82.9% to 99.8% and additional West Eurasian ancestry. In contrast, the early West Xiongnu (earlyXiongnu_west) and late Sarmatian Xiongnu (lateXiongnu_Sarmatian) derived ancestry mainly from West Eurasian; for example, early West Xiongnu exhibited 68.4% Afanasievo-related ancestry. Among the Central Steppe pastoralists, Wusun, Kangju, and Tianshan Hun derived a majority of their ancestry (62.4%–73%) from Western Steppe nomadic Afanasievo groups with the remainder (37.6%–27%) characterized as BMAC (the Bactria-Margiana Archaeological Complex) and East Eurasian. The Turkic Karluk, Kipchak, and Karakhanid could be modeled derived 35%–50.6% of ancestry from Afanasievo, 10.5%–21.7% from BMAC, and 38.9%–49.4% from YR_IA. The proportion of East Eurasian ancestry increased in the Medieval pastoralists (CentralSteppe_Medieval_Nomad and Kazakhstan_Golden-HordeAsian) from 67.3% to 82.5%, but a historical Kazakhstan individual (Kazakhstan_His) displayed a different genetic profile due to a primary ancestry at 75.5% derived from Western Steppe pastoralists.

The Ashina had not shown close genetic affinity with early Medieval or CentralSteppe Türk as reflected in the lower values of outgroup—f_3(Ashina, early Medieval or Central-Steppe Türk; Mbuti) and the significant values of f_4(X, Mbuti; Ashina, early Medieval or CentralSteppe Türk) ($|Z| > 3$) (Figs. S5, S8). Although earlyMed_Turk showed a genetic affinity with a Rouran individual (Fig. S4), earlyMed_Turk harbored a different genetic profile from the Ashina individual, as shown in the PCA and pairwise $qpWave$. The early Medieval Türk (earlyMed_Turk) derived the major ancestry from ANA at a proportion of 62.2%, the remainder from BMAC (10.7%) and Western Steppe Afanasievo nomad (27.1%) (Figs. 1C, 1D; Table S2E). The geographically remote Central Steppe Türk (Kyrgyzstan_Turk and Kazakhstan_Turk) could be modeled as an admixture of ANA (Mongolia_N_-North), BMAC, and West Steppe pastoralists (Afanasievo) ($P = 0.0196$) (Fig. S5; Table S2E). In contrast to Ashina, Central Steppe and early Medieval Türk exhibited a high but variable degree of West Eurasian ancestry, indicating there was a genetic substructure of the Türkic empire.

3 Genetic relationship between ancient Göktürk and modern Turkic-speaking populations

Among the pastoralist Khanates in Eastern Steppe, it has been argued that the spread of Mongolic languages was connected with Xianbei, Rouran, Khitan and Mongol groups, while the Xiongnu, Türk and Uyghur were seen as linked with diffusion of the Turkic languages. In the latter case, two waves of diffusion have been hypothesized: the Bulgharic

Turkic diffusion, beginning in the Hunnic period, instigated by the earlier expansion of the Xiongnu, and followed up by the demic expansion associated with the Türkic Khanate (Nichols, 2011). During the second and third centuries CE, the Central Steppe populated by Iranian-speaking groups was gradually replaced by an increasingly Turkic-speaking population (de la Vaissière, 2005). The genetic relationship between the ancient Göktürk and modern Turkic-speaking populations has remained a matter of controversy. In PCA, Turkic-speaking groups were scattered along the east-west cline in PC2, while Ashina was positioned in the Northeast Asians cluster. Modern populations showing the most similar genetic profile with Ashina were Tungusic, and secondly Mongolic; this relationship was revealed in PCA, unsupervised ADMIXTURE and outgroup-f_3 statistics. Notably, these results did not provide evidence of genetic similarity between ancient Türkic people and present-day Turkic-speaking groups.

To assess the degree of relation to Ashina, we compared Tungusic and Mongolic populations with Turkic populations by formal f_4 statistics (Fig. S16). We found that Tungusic and Mongolic-speaking groups were more related to Ashina when compared with Turkic speakers, that is, f_4 (Mbuti, Ashina; Tungusic/Mongolic, Turkic) <0 ($Z < -3$, except for Yakut and Dolgan), f_4 (X, Mbuti; Ashina, Tungusic/Mongolic) ~0 ($0 < |Z| < 3$). We further observed that Tungusic populations shared the highest genetic similarity with Ashina as reflected in f_4 (Mbuti, Ashina; Tungusic, Mongolic) with the exception of Evenk_FarEast. Turkic-speaking populations harbored significantly divergent genetic profiles when compared with Ashina, that, f_4 (X, Mbuti; Ashina, Turkic) ($Z > 3$ when X included West Eurasians; $Z < -3$ when X included East Eurasians). We observed significant differentiation between Turkic-speaking populations and ancient populations associated with the diffusion of Turkic languages (Fig. S17), indicating the spread of Turkic languages was mainly driven by culture factors rather than demic diffusion and population integration.

Unsupervised ADMIXTURE clustering analysis revealed a west-east admixture pattern among Turkic populations. We also conducted a supervised ADMIXTURE clustering analysis to investigate whether the Ashina-related ancient group had left a genetic legacy in modern Turkic populations. The ancestral proxies selected were Ashina, Mongolia_N_North, YR_LN, and Russia_Sintashta_MLBA based on unsupervised ADMIXTURE. We observed the proportions of ancestry related to ANA (Mongolia_N_North/Ashina) were diverse among Turkic-speaking groups. ANA ancestry was absent among the westernmost Turkic populations (Fig. 1D).

We performed f_4 (Turkic, Mbuti; Ashina, East Eurasian) to further determine whether the Ashina-related lineage is sufficient to explain the East Eurasian ancestry in Turkic populations, where East Eurasian included people who might not be influenced by the West Eurasian-like ancestry. The Z-scores of f_4 (X, Mbuti; Ashina, Turkic) tended towards positive f_4 (most $Z > 0$), showing that the Ashina was genetically closer to East Asians than Turkic groups (Fig. S18). We systematically explored $qpAdm$-based admixture models of Turkic populations and compared them with Tungusic and Mongolic samples exhibiting close proximity to Ashina (Table S3). Ashina was consistent with deriving from

J. Syst. Evol. 00 (0): 1–9, 2023

www.jse.ac.cn

one single source with some Mongolic and Tungusic populations (Bonan, Dongxiang, Evenk_Transbaikal, Oroqen, Tu, and Ulchi) but not with Turkic populations. The two-way and three-way admixture models of West and East Eurasians including Ashina failed for most Turkic populations, but only succeeded in Dolgan, Salar, Tuvinian, Yakut, Chuvash Altaian, Altaian_Chelkan, Khakass, Kazakh_China, Kyrgyz_China and Kyrgyz_Tajikistan, revealing the genetic heterogeneity within Turkic populations. That contrasted with Tungusic and Mongolic speakers that showed a dominating contribution from the Ashina-related lineage in one-way/two-way admixture models. The results of *qpAdm* and supervised ADMIXTURE demonstrated a limited contribution from Ashina in Turkic-speaking populations and the continuity of ANA ancestry showing by Ashina in Northeast Asian Tungusic- and Mongolic-speaking populations.

4 Implications

In summary, we have unveiled the first genomic profile of the ancient Türkic royal family. Our genomic analyses of Empress Ashina revealed Göktürk's Northeast Asian origin (97.7% Northeast Asian ancestry and 2.3% West Eurasian ancestry), refuting the western Eurasian origin and multiple origin hypotheses. We found Ashina shared most genetic affinity with post-Iron Age Tungusic and Mongolic Steppe pastoralists, such as Rouran, Xianbei, Khitan, and Heshui_Mohe, and showed genetic heterogeneity with other ancient Türkic people, suggesting the multiple sources of the Türkic Khanate populations. Furthermore, the limited contribution from ancient Göktürk found in modern Turkic-speaking populations once again validates a cultural diffusion model over a demic diffusion model for the spread of Turkic languages.

Acknowledgements

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The DNA sequences reported in this paper have been deposited in the Genome Sequence Archive in the National Genomics Data Center (GSA) under accession HRA003101.

References

Damgaard PB, Marchi N, Rasmussen S, Peyrot M, Renaud G, Korneliussen T, Moreno-Mayar JV, Pedersen MW, Goldberg A, Usmanova E, Baimukhanov N, Loman V, Hedeager L, Pedersen AG, Nielsen K, Afanasiev G, Akmatov K, Aldashev A, Alpaslan A, Baimbetov G, Bazaliskii VI, Beisenov A, Boldbaatar B, Boldgiv B, Dorzhu C, Ellingvag S, Erdenebaatar D, Dajani R, Dmitriev E, Evdokimov V, Frei KM, Gromov A, Goryachev A, Hakonarson H, Hegay T, Khachatryan Z, Khaskhanov R, Kitov E, Kolbina A, Kubatbek T, Kukushkin A, Kukushkin I, Lau N, Margaryan A, Merkyte I, Mertz IV, Mertz VK, Mijiddorj E, Moiyesev V, Mukhtarova G, Nurmukhanbetov B, Orozbekova Z, Panyushkina I, Pieta K, Smrčka V, Shevnina I, Logvin A, Sjögren KG, Štolcová T, Taravella AM, Tashbaeva K, Tkachev A, Tulegenov T, Voyakin D, Yepiskoposyan L, Undrakhbold S, Varfolomeev V, Weber A, Wilson Sayres MA, Kradin N, Allentoft ME, Orlando L, Nielsen R, Sikora M, Heyer E, Kristiansen K, Willerslev E. 2018a. 137 ancient human genomes from across the Eurasian steppes. *Nature* 557: 369–374.

Damgaard PB, Martiniano R, Kamm J, Moreno-Mayar JV, Kroonen G, Peyrot M, Barjamovic G, Rasmussen S, Zacho C, Baimukhanov N, Zaibert V, Merz V, Biddanda A, Merz I, Loman V, Evdokimov V, Usmanova E, Hemphill B, Seguin-Orlando A, Yediay FE, Ullah I, Sjögren KG, Iversen KH, Choin J, de la Fuente C, Ilardo M, Schroeder H, Moiseyev V, Gromov A, Polyakov A, Omura S, Senyurt SY, Ahmad H, McKenzie C, Margaryan A, Hameed A, Samad A, Gul N, Khokhar MH, Goriunova OI, Bazaliiskii VI, Novembre J, Weber AW, Orlando L, Allentoft ME, Nielsen R, Kristiansen K, Sikora M, Outram AK, Durbin R, Willerslev E. 2018b. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* 360: eaar7711.

de la Vaissière É. 2005. Huns et Xiongnu. *Central Asiatic Journal* 49: 3–26.

Grousset R. 1970. *The empire of the steppes: A history of Central Asia*. New Brunswick: Rutgers University Press.

Jeong C, Balanovsky O, Lukianova E, Kahbatkyzy N, Flegontov P, Zaporozhchenko V, Immel A, Wang CC, Ixan O, Khussainova E, Bekmanov B, Zaibert V, Lavryashina M, Pocheshkhova E, Yusupov Y, Agdzhoyan A, Koshel S, Bukin A, Nymadawa P, Turdikulova S, Dalimova D, Churnosov M, Skhalyakho R, Daragan D, Bogunov Y, Bogunova A, Shtrunov A, Dubova N, Zhabagin M, Yepiskoposyan L, Churakov V, Pislegin N, Damba L, Saroyants L, Dibirova K, Atramentova L, Utevska O, Idrisov E, Kamenshchikova E, Evseeva I, Metspalu M, Outram AK, Robbeets M, Djansugurova L, Balanovska E, Schiffels S, Haak W, Reich D, Krause J. 2019. The genetic history of admixture across inner Eurasia. *Nature Ecology & Evolution* 3: 966–976.

Jeong C, Wang K, Wilkin S, Taylor WTT, Miller BK, Bemmann JH, Stahl R, Chiovelli C, Knolle F, Ulziibayar S, Khatanbaatar D, Erdenebaatar D, Erdenebat U, Ochir A, Ankhsanaa G, Vanchigdash C, Ochir B, Munkhbayar C, Tumen D, Kovalev A, Kradin N, Bazarov BA, Miyagashev DA, Konovalov PB, Zhambaltarova E, Miller AV, Haak W, Schiffels S, Krause J, Boivin N, Erdene M, Hendy J, Warinner C. 2020. A dynamic 6,000-year genetic history of Eurasia's eastern steppe. *Cell* 183: 890–904.

Mao X, Zhang H, Qiao S, Liu Y, Chang F, Xie P, Zhang M, Wang T, Li M, Cao P, Yang R, Liu F, Dai Q, Feng X, Ping W, Lei C, Olsen JW,

Bennett EA, Fu Q. 2021. The deep population history of northern East Asia from the Late Pleistocene to the Holocene. *Cell* 184: 3256–3266.

Nakatsuka N, Harney É, Mallick S, Mah M, Patterson N, Reich D. 2020. ContamLD: Estimation of ancient nuclear DNA contamination using breakdown of linkage disequilibrium. *Genome Biology* 21: 199.

Narasimhan VM, Patterson N, Moorjani P, Rohland N, Bernardos R, Mallick S, Lazaridis I, Nakatsuka N, Olalde I, Lipson M, Kim AM, Olivieri LM, Coppa A, Vidale M, Mallory J, Moiseyev V, Kitov E, Monge J, Adamski N, Alex N, Broomandkhoshbacht N, Candilio F, Callan K, Cheronet O, Culleton BJ, Ferry M, Fernandes D, Freilich S, Gamarra B, Gaudio D, Hajdinjak M, Harney É, Harper TK, Keating D, Lawson AM, Mah M, Mandl K, Michel M, Novak M, Oppenheimer J, Rai N, Sirak K, Slon V, Stewardson K, Zalzala F, Zhang Z, Akhatov G, Bagashev AN, Bagnera A, Baitanayev B, Bendezu-Sarmiento J, Bissembaev AA, Bonora GL, Chargynov TT, Chikisheva T, Dashkovskiy PK, Derevianko A, Dobeš M, Douka K, Dubova N, Duisengali MN, Enshin D, Epimakhov A, Fribus AV, Fuller D, Goryachev A, Gromov A, Grushin SP, Hanks B, Judd M, Kazizov E, Khokhlov A, Krygin AP, Kupriyanova E, Kuznetsov P, Luiselli D, Maksudov F, Mamedov AM, Mamirov TB, Meiklejohn C, Merrett DC, Micheli R, Mochalov O, Mustafokulov S, Nayak A, Pettener D, Potts R, Razhev D, Rykun M, Sarno S, Savenkova TM, Sikhymbaeva K, Slepchenko SM, Soltobaev OA, Stepanova N, Svyatko S, Tabaldiev K, Teschler-Nicola M, Tishkin AA, Tkachev VV, Vasilyev S, Velemínský P, Voyakin D, Yermolayeva A, Zahir M, Zubkov VS, Zubova A, Shinde VS, Lalueza-Fox C, Meyer M, Anthony D, Boivin N, Thangaraj K, Kennett DJ, Frachetti M, Pinhasi R, Reich D. 2019. The formation of human populations in South and Central Asia. *Science* 365: eaat7487.

Nichols J. 2011. Forerunners to globalization: The Eurasian steppe and its periphery. *Studies in Slavic and General Linguistics* 38: 177–195.

Ning C, Li T, Wang K, Zhang F, Li T, Wu X, Gao S, Zhang Q, Zhang H, Hudson MJ, Dong G, Wu S, Fang Y, Liu C, Feng C, Li W, Han T, Li R, Wei J, Zhu Y, Zhou Y, Wang CC, Fan S, Xiong Z, Sun Z, Ye M, Sun L, Wu X, Liang F, Cao Y, Wei X, Zhu H, Zhou H, Krause J, Robbeets M, Jeong C, Cui Y. 2020. Ancient genomes from northern China suggest links between subsistence changes and human migration. *Nature Communications* 11: 2700.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192: 1065–1093.

Renaud G, Slon V, Duggan AT, Kelso J. 2015. Schmutzi: Estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biology* 16: 224.

Rui CM. 1991. The prehistorical residence of Ashina. *North West Ethno-National Studies* 2: 165–176.

Tan QX. 1989. *The historical atlas of China*. Beijing: Sinomap Press.

Triska P, Chekanov N, Stepanov V, Khusnutdinova EK, Kumar GPA, Akhmetova V, Babalyan K, Boulygina E, Kharkov V, Gubina M, Khidiyatova I, Khitrinskaya I, Khrameeva EE, Khusainova R, Konovalova N, Litvinov S, Marusin A, Mazur AM, Puzyrev V, Ivanoshchuk D, Spiridonova M, Teslyuk A, Tsygankova S, Triska M, Trofimova N, Vajda E, Balanovsky O, Baranova A, Skryabin K, Tatarinova TV, Prokhortchouk E. 2017. Between Lake Baikal and the Baltic Sea: Genomic history of the gateway to Europe. *BMC Genetics* 18: 110.

Wang CC, Yeh HY, Popov AN, Zhang HQ, Matsumura H, Sirak K, Cheronet O, Kovalev A, Rohland N, Kim AM, Mallick S, Bernardos R, Tumen D, Zhao J, Liu YC, Liu JY, Mah M, Wang K, Zhang Z, Adamski N, Broomandkhoshbacht N, Callan K, Candilio F, Carlson KSD, Culleton BJ, Eccles L, Freilich S, Keating D, Lawson AM, Mandl K, Michel M, Oppenheimer J, Özdoğan KT, Stewardson K, Wen S, Yan S, Zalzala F, Chuang R, Huang CJ, Looh H, Shiung CC, Nikitin YG, Tabarev AV, Tishkin AA, Lin S, Sun ZY, Wu XM, Yang TL, Hu X, Chen L, Du H, Bayarsaikhan J, Mijiddorj E, Erdenebaatar D, Iderkhangai TO, Myagmar E, Kanzawa-Kiriyama H, Nishino M, Shinoda KI, Shubina OA, Guo J, Cai W, Deng Q, Kang L, Li D, Li D, Lin R, Nini, Shrestha R, Wang LX, Wei L, Xie G, Yao H, Zhang M, He G, Yang X, Hu R, Robbeets M, Schiffels S, Kennett DJ, Jin L, Li H, Krause J, Pinhasi R, Reich D. 2021. Genomic insights into the formation of human populations in East Asia. *Nature* 591: 413–419.

Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt HJ, Kronenberg F, Salas A, Schönherr S. 2016. HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Research* 44: W58–W63.

Yang MA, Fan X, Sun B, Chen C, Lang J, Ko YC, Tsang CH, Chiu H, Wang T, Bao Q, Wu X, Hajdinjak M, Ko AM, Ding M, Cao P, Yang R, Liu F, Nickel B, Dai Q, Feng X, Zhang L, Sun C, Ning C, Zeng W, Zhao Y, Zhang M, Gao X, Cui Y, Reich D, Stoneking M, Fu Q. 2020. Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* 369: 282–288.

Yunusbayev B, Metspalu M, Metspalu E, Valeev A, Litvinov S, Valiev R, Akhmetova V, Balanovska E, Balanovsky O, Turdikulova S, Dalimova D, Nymadawa P, Bahmanimehr A, Sahakyan H, Tambets K, Fedorova S, Barashkov N, Khidiyatova I, Mihailov E, Khusainova R, Damba L, Derenko M, Malyarchuk B, Osipova L, Voevoda M, Yepiskoposyan L, Kivisild T, Khusnutdinova E, Villems R. 2015. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genetics* 11: e1005068.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12938/suppinfo:

Table S1. A, The genome coverage, contamination based on mtDNA and Linkage Disequilibrium and damage pattern of the sampled ancient individual. **B,** Genetic gender and mitochrondrial haplotype of ancient individual. **C,** Populations used in supervised ADMIXTURE.

Table S2. A, The outgroup-f_3 of Ashina including 210 modern/ancient Eurasians showed Ashia shared genetic similarity with Northeast Asians including ANA, Xianbei, Rouran, present-day Tungusic. **B,** The outgroup-f_3 of earlyMed_Turk including 210 modern/ancient Eurasians showed the most close genetic affinity with Rouran in earlyMed_Turk. **C,** The outgroup-f_3 of CentralSteppe_Turk including 210 modern/ancient Eurasians. **D,** The admixture modelings of Ashina. We use ANA (AR_EN, ARpost9K, DevilsCave_N, Boisman_MN, Mongolia_N_North, Mongolia_N_East, Russia_Shamanka_E-neolithic), Afanasievo and YR_IA as sources to model the admixture of Ashina and CentralSteppe_Turk, earlyMed_-Turk. Models with *P*-value > 0.05 provided an adequate model fit and were highlighted in light blue, and the models with *P* > 0.05 but the proportion of source or standard error > CoefRef were highlighted in gray. We used a base set of eight outgroups ("Base"): Mbuti.DG, Indian_GreatAnda-man_100BP.SG, Yana_UP.SG, Iran_Wezmeh_N.SG, Turkme-nistan_Gonur_BA_1, Loschbour.DG, Anatolia_N, Kazakhstan_Eneolithic_Botai.SG. Afanasievo and YR_IA were used as extract outgroups and additional sources to test the robust admixture modeling of target populations. Rotated *qpadm* of Ashina and earlyMed_Turk, CentralSteppe_Turk showed ANA

ancestry continued and the detected additional gene flow from Afanasiveo pastoralists in Ashina. The admixture model of Ashina failed in earlyMed_Turk and CentralSteppe_Turk, indicating they harbored different genetic profiles with Ashina. **E**, The admixture modelings of post-Iron Age (IA) East/Central Steppe pastoralists. For modeling, we used Mongolia_N_North/Chemurchek_northAlaiMongolia_Khovsgol_LBA/Ulaanzuukh_SlabGrave, Botai, BMAC (Turkmenistan_Gonur_BA_1), Afanasievo and YR_IA as sources to model the admixture of post-IA East/Central Steppe pastoralists. Models with P-value > 0.05 provided an adequate model fit and were highlighted in blue, models with $P > 0.5$ but standard error > CoefRef or negative proportion were highlighted in gray. We used a base set of seven outgroups: Mbuti.DG, Indian_GreatAndaman_100BP.SG, Yana_UP.SG, Iran_Wezmeh_N.SG, Loschbour.DG, Anatolia_N, AR19K. **F**, Genetic admixture dates for Ashina based on DATES including Mongolia_N_North ($n = 7$), Ulaanzuukh_SlabGrav ($n = 16$),DevilsCave_N ($n = 5$), YR_IA ($n = 10$), CHB.SG ($n = 103$), Russia_Afanasievo ($n = 31$),Russia_Alan.SG ($n = 5$), Russia_Sarmatian.SG ($n = 8$).CEU.SG ($n = 99$), Russia_MLBA_Sintashta ($n = 50$). The Ashina showed west Eurasian admixture dating to about 2000 years ago (~1330–1550 years, assuming 29 years/generation. Mongolia_N_North + Sarmatian/Mongolia_N_North + Afanasievo), while earlyMed_Turk and CentralSteppe_Turk showed more recently mixing date.

Table S3. Rotated *qpadm* of present-day Altaic-speaking populations. **A**, Rotated *qpadm* of Turkic populations showed genetic divergency within population with different Ashina-related ancestry proportions, indicated the limitedly genetical contribution from Ashina in Turkic-speaking populations. **B**, Rotated *qpadm* of Tungusic and Mongolic populations showed genetic continuity with Ashina-related ancestry and the additional gene flow from millet farmer in Yellow River and West Steppe nomad ancestry in Tungusic/Mongolic populations.

Fig. S1. The geographical location of this individual and Records of Empress Wude.

Fig. S2. Proportion of C > T and G > A substitutions in human DNA across DNA fragments in the Ashina individual. The red curve represented C > T substitutions that increased at the 5′-end and the blue curve represented G > T substitutions that increased at the 3′-end, as expected for authentic ancient DNA.

Fig. S3. Detailed principal component analysis (PCA) of 1667 Eurasian exhibited the west–east Eurasian cline and north–south Eastern Eurasian cline. The Ashina individual fell within present-day and ancient Eastern Eurasians along PC2, and clustered with modern Tungusic and Mongolic speakers, ancient populations in Northeast Asia and eastern Mongolia Plateau, especially the populations previously referred to as "Ancient Northeast Asian" (ANA) that are genetically homogeneous hunter-gatherers from Northeast Asia ("Baikal_EN", 5200–4200 BCE, "Mongolia_N", 6000–4400 BCE, "AR_EN", 5500–5300 BCE, and "DevilsCave_N", 5700 BCE) and post-Iron Age East Steppe nomadic people including Xianbei, Rouran, Khitan and part of Mongols.

Fig. S4. ADMIXTURE results for selected Eurasians for K = 2–6. **A,** Cross-validation (CV) errors for the ADMIXTURE

analysis based on 1240 K data set. **B**, CV errors for the ADMIXTURE analysis based on HO data set. **C**, ADMIXTURE results for K = 2–6 based on 1240 K data set. After pruning for linkage disequilibrium, the number of SNPs included in this analysis was 871 858, the included modern populations were from HGDP. **D**, ADMIXTURE results included Turkic populations for K = 2–6 based on HO data set. The number of included SNPs was 279 737. The ADMIXTURE analyses base on 1240 K and HO data set both revealed two summarized ancestries in Ashina: one related to West Eurasian ancestry, the dominating one associated with East Eurasians including ANA and southern ancestry. The results of ADMIXTURE also presented the genetic discrepancy among Ashina, earlyMed_Turk and CentralSteppe_Turk and the genetic heterogeneity among Ashina and present-day Turkic-speaking populations.

Fig. S5. Close genetic relationship between ancient and present-day Eurasian populations and geographically different individuals of the Türkic Khaganate. We presented top 60 outgroup-f_3 signal for each different group of the Türkic Khaganate. Horizontal bars represent the point estimate ±3 standard error. The outgroup-f_3 of Ashina showed the close genetic affinity with East Eurasians than West Eurasians, especially northern East Eurasian. earlyMed_Turk who sampled in Mongolia Plateau presented close genetic relationship with East Eurasians, especially Rouran.The CentralSteppe_Turk harbored similar genetic relationship with Eurasians.

Fig. S6. f Statistics in forms of f_4 (X, Mbuti; Ashina, pre-Iron Age Northeast Asian populations) and f_4 (pre-Iron Age Northeast Asian populations, Mbuti; Ashina, X) test whether pre-Iron Age Northeast Asian populations make contribution to the formation of Ashina. Darker orange and blue squares showed sharing at |Z-score| > 6. Blue and orange squares show sharing at 3 < |Z-score| <6. Gray squares indicate significantly less sharing at the same threshold. **A**, f_4 (X, Mbuti; Ashina, pre-Iron Age Northeast Asian populations) showed Ashina formed a genetic clade together with Neolithic hunter-gatherers in Northeast Asia including Amuer River, Mongolia Plateau and DevilsCave, Baikal Lake (Russia_Shamanka_Eneolithic). In addition, Ashina showed close genetic affinity with population related to Bronze Age SlabGrave and Ulaanzuukh culture in Mongolia Plateauwho harbors dominating ANA ancestry. **B**, f_4 (pre-Iron Age Northeast Asian populations, Mbuti; Ashina, X) further provided evidence that ancestry related to hunter-gathers in Northeast Asia dedicated to the gene pool of Ashina as reflected in significantly positive f_4 with the exception of Sino-Tibetan populations.

Fig. S7. The analysis of f_4 (West Eurasian, Mbuti; Ashina, East Eurasian) detected whether gene influx from West Eurasian flowed into the gene pool of Ashina basing on 1240 K data set. **A–L**, West Eurasians included West Steppe nomadic populations from Bronze Age to Iron Age, Anatolian and Iranian farmers, East hunter-gatherers (EHG), ANE (MA1, Ancient North European) and Botai and the Bactria-Margiana Archaeological Complex (BMAC) of Central Asia. East Eurasian included reported population without gene flow from West Eurasian in previous studies. Different shapes represented different Z-scores of f_4. The results showed the gene pool of Ashina was influenced by the eastward

migration of West Steppe pastoralists and the population related to BMAC who was attested to migrated along Inner Asia Mountains Corridor into the Mongolia Plateau and influenced ancient Tianshan and Mongolian people ($f_4 > 0$). The detected Anatolia farmers, EHG and Iranian farmers, Botai-related ancestry in Ashina might be mediated by the eastward migration of West Steppe pastoralists and BMAC, respectively.

Fig. S8. The significantly genetic heterogeneity within the Türkic Khaganate. **A, B,** The genetic divergency in individuals of the Türkic Khaganate revealed by f_4(X, Mbuti; Ashina, earlyMed_Turk/CentralSteppe_Turk) showing earlyMed_Turk/CentralSteppe_Turk harbored more allele with West Eurasian than Ashina. **A,** earlyMed_Turk. **B,** CentralSteppe_Turk.

Fig. S9. The genetic relationship of Ashina with Iron-Age Tianshan nomadic populations. **A–C,** The results of f_4 in form of f_4 (X, Mbuti; Ashina, Kazakhstan_Kanju/Kazakhstan_Wusun/Kyrgzstan_TianshanHun) all showed Iron-Age Tianshan nomadic people harbored diverse genetic profile with Ashina with more genetic affinity with West Eurasian and West Siberian hunter-gatherers (WSHG). In contrast, Ashina shared more allele with East Eurasians than Iron-Age Tianshan nomadic populations. Different shapes represented different Z-score.

Fig. S10. The genetic relationship of Ashina with Xiongnu populations in different periods. Jeong et al. (2020) split early Xiongnu into two subgroups, earlyXiongnu_west (SKT010, SKT001, SKT003, SKT009, SKT008, AST001) and earlyXiongnu_rest (JAG001, SKT002, SKT004, SKT005, SKT006, SKT012), based on their individual genetic modeling results. **A, B,** The results of f_4 in form of f_4 (X, Mbuti; Ashina, earlyXiongnu_rest/earlyXiongnu_west) showed Ashina shared more genetic affinity with East Eurasians than early Xiongnu. The genetic discrepancy with Ashina was different in earlyXiongnu_rest and earlyXiongnu_west, showing the genetic affinity with ancient populations in Northeast Asia and Mongolia Plateau was similar in earlyXiongnu_rest and Ashina (no significant Z-score). While that genetic profile did not exist in earlyXiongnu_west who harbored more West Eurasian affinity than Ashina. **C–E,** The results of f_4 (X, Mbuti; Ashina, lateXiongnu/lateXiongnu_han/lateXiongnu_sarmatian) presented the difference of harboring genetic relationships with Asina among late Xiongnu: late Xiongnu shared similar genetic affinity to ancient Northeast Asians with Ashina (non-significant Z-score when X included Northeast Asian populations), lateXiongnu_han possessed semblable genetic profile with Ashina (non-significant Z-score, except for AR_EN), that contrasted with lateXiongnu_sarmatian who shared more genetic affinity with West Eurasians than Ashina (significantly negative Z-score when X was West Eurasian populations).

Fig. S11. The genetic relationship of Ashina with contemporary East Steppe nomadic populations including Xianbei and Rouran. **A–C,** The result of f_4 in form of f_4(X, Mbuti; Ashina, AR_Xianbei_IA/Mongolia_Xianbei/Rouran) revealed the genetic homogeneity of Ashina with Xianbei and Rouran (no significant Z score).

Fig. S12. The genetic relationship of Ashina with Mongolic and Tungusic East Steppe pastoralists after the Türkic Khanate. (**A–C**) The result of f_4 in form of f_4(X, Mbuti;

Ashina, lateMed_Khitan/lateMed_Mongol/Russia_Heishui_Mohe_early_Medieval) demonstrated the discrepancy of genetic relationship with Ashina in Khitan and Mongol empire; Mongolic Khitan and Tungusic Heshui_Mohe had genetic similarity with Ashina, while Mongols in Mongol empir period showed genetic differentiation with Ashina.

Fig. S13. The genetic relationship of Ashina with historically Turkic-speaking pastoralists of Central Steppe after the collapse of the Türkic Khaganate. **A–E,** The result of f_4 in form of f_4 (X, Mbuti; Ashina, Kimak/KaraKhanid/Uigur/Karluk/Kipchak) indicated the high genetic differentiation between Ashina and later Turkic-speaking pastoralists of Central Steppe including Kimak, KaraKhanid, Uigur, Karluk and Kipchak Khaganate. The Central Steppe pastoralists had evident genetic affinity with West Eurasians (significantly negative Z-score when X included West Eurasians), indicating the prominent ancestry from West Eurasian in the Central Steppe pastoralists, that contrasted with Ashina who harbored dominating ancestry from Northeast Asian.

Fig. S14. The genetic relationship of Ashina with Central Steppe pastoralists in Medieval period. **A–C,** The result of f_4 in form of f_4(X, Mbuti; Ashina, Kazakhstan_His/Kazakhstan_GoldenHordeAsian/CentralSteppe_Medieval_Nomad) showed diverse genetic relationship with Ashina among Central Steppe pastoralists in Medieval period. Kazakhstan_GoldenHordeAsian formed a genetic clade together with Ashina, which indicating the similar genetic profile between them. However, Kazakhstan_His and CentralSteppe_Medieval_Nomad showed closer genetic affinity with West Eurasians than Ashina.

Fig. S15. Genetic heterogeneity of the Türkic Khaganate. Heatmap of P-value of pairwise qpWave among post-Iron Age Central/East Steppe pastoralists. "++" represented values greater than 0.05, and "+" represented values <0.05 and >0.01. Computations were based on the outgroup set (Mbuti.DG + Onge.DG + Russia_MA1_HG.SG + Russia_Kostenki14.SG + Iran_GanjDareh_N + Kazakhstan_Eneolithic_Botai.SG + Russia_Sintashta_MLBA.SG + AR19K + Mongolia_N_North+UpperMid_YR_LN).

Fig. S16. The genetic relationship with Ashina of present-day Altaic-speaking populations. **A,** We conducted f_4(X, Mbuti; Ashina, Tungusic/Mongolic/Turkic) to investigate the genetic relationship of present-day Altaic populations with Ashina, and found that Mongolic and Tungusic speakers formed genetic clade with Ashina, and influenced by additional gene flow from millet farmers in Yellow River or West Eurasian in some groups (negative Z-score). By contrast, Turkic populations showed extremely disparate genetic profile with Ashina with the exception of Yakut, Dolgan, Tuvinian and Salar who presented some extent of genetic affinity with Ashina. **B,** We further performed f_4 (Mbuti, Ashina; Tungusic/Mongolic, Turkic) to provide robust evidence of supporting the close genetic affinity with Ashina in Tungusic and Mongolic populations rather than in Turkic populations. Tungusic and Mongolic speakers did have closer genetic relationship with Ashina than Turkic populations ($Z < -3$). "-"represented $-6 < Z$-score < -3; "–" represented Z-score < -6; "+" represented $3 < Z$-score < 6; "++" represented Z-score >6. **C,** The result of f_4 (Mbuti, Ashina; Tungusic, Mongolic) further revealed Tungusic population shared more alleles with Ashina than Mongolic populations as reflected in significantly negative f_4 values with the exception of Eavek_FastEast.

Fig. S17. The genetic affinity between Turkic-speaking populations and ancient populations related with the diffusion of Turkic language. Heatmap of P-value of pairwise $qpWave$ among Turkic-speaking populations and ancient populations related with the diffusion of Turkic languages. The genetic relationships between Turkic-speaking population and ancient populations associated with the diffusion of Turkic language revealed the significantly genetic differentiation of Ashina/earlyMed_Turk/earlyXiongnu/lateXiongnu/earlyMed_Uigur and Turkic-speaking populations ($P < 0.01$), CentralSteppe_Turk formed pairwise clade with Uzbek and Dolgan, the formed pairwise clades existed in other pairwise combination, including pairwise combination of Kimak with Tatar_Siberian, Uyghur, Karakalpak, Nogai_Astrakhan, Nogai_Stavropol and Uzbek, pairwise combination between lateXiongnu to Karakalpak ($P > 0.05$). "++" represented values >0.05, and "+" represented values <0.05 and >0.01. Computations were based on the outgroup set (Mbuti.DG + Onge.DG + Russia_MA1_HG.SG + Russia_Kostenki14.SG + Iran_GanjDareh_N + Anatolia_N + CHG + AR19K + Mongolia_N_North+UpperMid_YR_LN).

Fig. S18. The potential East Eurasian ancestral source of Turkic population. We performed f_4(X, Mbuti; Pop, Turkic) to further determine whether other East Eurasians are more suited to explain the East Eurasian ancestry of Turkic than Ashina. X was chosen from representative populations including Chuanyun (the Southern East Asian), YR_LN (millet farmers in Yellow River), Russia_Sintashta_MLBA and Botai. Pop included other potential East Eurasian ancestral source of Turkic-speakers. We observed that Ashina was more suited to approximate East Eurasian ancestry of Turkic populations where Z-scores of f_4 (X, Mbuti; Ashina, Turkic) tended to approach to zero.

Supplementary Document. Section 1 Archaeological Site and Sample Description of the Xiaoling Mausoleum. **Section 2** Material and Methods. **Section 3** Ancient Data Analyses and Genetic Characterization of Ancient Individual.