

ORIGINAL ARTICLE

Traces of Bronze Age globalization in East Asia: Insights from a revised phylogeography of the Y-chromosome haplogroup Q1a1a-M120

La-Su Mai¹ | Xian-Peng Zhang¹  | Kai-Jun Liu² | Peng-Cheng Ma³ | Hui Li^{4,5} | Jin Sun⁶ | Lan-Hai Wei^{1,5}

¹Institute of Anthropology and Human Sciences, School of Ethnology and Anthropology, Inner Mongolia Normal University, Hohhot, China

²School of International Tourism and culture, Guizhou Normal University, Guiyang, China

³School of Life Sciences, Inner Mongolia University, Hohhot, China

⁴MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China

⁵B&R International Joint Laboratory for Eurasian Anthropology, Fudan University, Shanghai, China

⁶Xingyi Normal University for Nationalities, Xingyi, China

Correspondence

Lan-Hai Wei, School of Ethnology and Anthropology, Inner Mongolia Normal University, Hohhot 010022, China.

Email: ryan.lh.wei@foxmail.com

Jin Sun, Xingyi Normal University for Nationalities, Xingyi 562400, China.

Email: sunjin19851228@gmail.com

Funding information

Scientific and Technology Committee of Shanghai Municipality, Grant/Award Number: 18490750300; National Natural Science Foundation of China, Grant/Award Number: 31900406 to LHW; Key projects of strategic international scientific and technological innovation cooperation of the Chinese Ministry of Science and Technology, Grant/Award Number: 2020YFE0201600

Abstract

Objective: In this study, we aim to explore the genetic imprint of Bronze Age globalization in East Asia from a phylogeographic perspective by examining the Y-chromosome haplogroup Q1a1a-M120, and to identify key demographic processes involved in the formation of early China and the ancient Huaxia people.

Methods: Over the past few decades, we have collected the sequences of 347 Y chromosomes from the haplogroup Q1a1a-M120. These sequences were utilized to analyze and reconstruct a highly revised phylogenetic tree with age estimates. And we analyzed the geographical distribution and spatial autocorrelation of nine major sub-branches of Q1a1a-M120. Finally, we observed the expansion of Q1a1a-M120 from the beginning of the Bronze Age in East Asia, along with the continuous dissemination of its sub-lineages among East Asian populations.

Results: We suggest that certain sub-lineages played a significant role in the formation of states and early civilizations in China, as well as in the development of the ancient Huaxia people, who are the direct ancestors of the Han population. Overall, we propose that haplogroup Q-M120 played a role in the introduction of Bronze Age culture to the central region of East Asia. Therefore, it is haplogroup Q-M120, rather than the Western Eurasian paternal lineage, that expanded and contributed to the gene pool of the East Asian population.

La-Su Mai, Xian-Peng Zhang, and Kai-Jun Liu contributed equally to the study.

© 2024 University College London (UCL) and John Wiley & Sons Ltd.

Conclusion: In summary, the globalization of the Bronze Age led to large-scale population replacement and admixture across various regions of Eurasia; our findings highlight the unique demographic processes that occurred in East Asia during this period.

KEYWORDS

Bronze Age globalization, East Asia, Q1a1a-M120, Y-chromosome

1 | INTRODUCTION

The contributions of ancient pastoralists from the eastern Eurasian Steppe and northwestern China to the formation of early China, early Chinese civilization, and the Huaxia population have been a subject of ongoing debate. Archaeological studies have indicated that pastoralist culture emerged in northwest China during the Chalcolithic and Bronze Ages, approximately 4500 to 2500 years ago (L. Liu, 2004). However, these archaeological sites exhibited mixed characteristics. Cultural elements from pre-indigenous farmers were discovered alongside features associated with steppe pastoralism, such as bronze technology, sheep, wheat, horses, and chariots (L. Liu & Chen, 2012). According to ancient Chinese records, the establishment of the Xia, Shang, and Zhou dynasties was linked to highland pastoralists in northwest China (L. Liu & Chen, 2003). In general, archaeological and historical studies suggest that early China and the Huaxia people emerged from a population admixture of highland pastoralists in northwest China and lowland farmers in the lower and middle Yangtze and Yellow River Basins.

Studies of ancient DNA have shed light on the genetic structure of ancient pastoralists in the northern and western regions of China, including remains from Xinjiang, Ningxia, Shaanxi, Shanxi, and Inner Mongolia (Kumar et al., 2022; Ning et al., 2019, 2020; Robbeets et al., 2021; C.-C. Wang et al., 2021; F. Zhang et al., 2021; Zhao et al., 2014). Ancient populations in northwest China exhibited a high degree of genetic diversity, encompassing paternal lineages C, N, O, Q, and R; ancient DNA studies suggest that western Eurasian genetic lineages spread to the central region of the Mongolian Plateau (Jeong et al., 2018, 2019, 2020; C.-C. Wang et al., 2021); and Xinjiang, located in the southwest of the Mongolia Plateau, serves as the eastern frontier for the spread of western Eurasian genetic lineages (Ning et al., 2019). In China, in the region east of Xinjiang, ancient DNA studies have not detected genetic lineages of western Eurasian populations. By contrast, the haplogroup Q1a1a-M120 is frequently detected at archaeological sites in northern China (Zhao et al., 2014). This haplogroup is the predominant lineage in several archaeological sites, including those of the Yanglang (杨朗) culture

in Ningxia and the cemeteries of the Peng state (棚国) in Shanxi (Zhao et al., 2010, 2014). Scholars have proposed that the Yanglang culture was a remnant of the Rong people (戎) (Y. Zhang, 2018), whereas the Peng state was founded by descendants of the ancient Guifang people (鬼方) (also called “Di people [狄]”) (B.-C. Ma, 2007; Zhao et al., 2014). “Rong-Di people (戎狄人群)” is a general term used by the ancient inhabitants of the central plains of China to refer to the nomadic tribes engaged in animal husbandry in northwest China (L. Liu & Chen, 2003). On the other hand, ancient Q-M120 individuals have also been discovered at archaeological sites in northwestern Mongolia (Jeong et al., 2018) and the northern Black Sea region (Krzewinska et al., 2018). These findings suggest that this paternal lineage was once widespread across the Eurasian Steppe. However, China is the only region where this branch experienced significant expansion.

Previous studies have assessed the genetic impact of Bronze Age globalization on the gene pool of populations in central East Asia. In most parts of China, the frequencies of western Eurasian genetic lineages are low (<7%) (Zhong et al., 2011). Several genetic lineages that may have originated in the region around the eastern Eurasian Steppe are found to be dominant haplogroups within the gene pool of East Asian populations. One of these lineages is N1a2-F1360, and its downstream clade, N1a2b-P43, which is the predominant paternal lineage among Samoyed-speaking populations (Rootsi et al., 2007). Ancient DNA analyses indicate that haplogroup N1a2a-F1101 may be the paternal lineage of the Royal Family of the Zhou dynasty in China (1027 BC to 221 BC) (P. Ma et al., 2021). N. Sun et al. (2019) reported that the expansion of Q-M120 corresponds to the flourishing of Bronze Age culture in northwest China. Ancient populations with the paternal lineage Q-M120 may have played an indispensable role in the formation of both ancient Huaxia and modern Han populations (N. Sun et al., 2019; Zhao et al., 2014). However, the origin, internal phylogenetic structure, diffusion process, and population history of the sub-clades of Q-M120 remain ambiguous.

In this study, we explored the history of the sub-lineages of Q-M120. We divided Q-M120 into nine major sub-lineages and investigated their phylogenetic relationships, divergence times, geographic distributions, and

population histories. We gained detailed insights into the geographic distribution of Q-M120 sub-lineages, their origins and diffusion, and their role in shaping modern populations in East Asia.

2 | MATERIALS AND METHODS

2.1 | Samples and sequencing

Saliva samples were collected from unrelated healthy males in East Asian populations over the past few decades. All participants provided written informed consent prior to their participation. The study and sample collection processes were reviewed and approved by the Medical Ethics Committee of Fudan University (approval number: BE1945) and Inner Mongolian Normal University (approval number: 20200110), and complied with the ethical principles of the 2013 Helsinki Declaration of the World Medical Association. Genomic DNA was extracted using the DP-318 Kit (Tiangen Biotechnology), following the manufacturer's protocol. First, routine Y-single nucleotide polymorphism (SNP) tests were conducted to determine the general haplogroup of all male samples. Subsequently, DNA specimens extracted from Q1a1a-M120 samples were sent for next-generation sequencing on the Illumina HiSeq2000 platform (Illumina).

2.2 | Data analysis

In this study, a total of 347 sequences of Q-M120 were analyzed (Table S1). We used the procedure that we described previously for the other steps prior to next-generation sequencing, that is, for DNA shearing, adding an adaptor, and gel electrophoresis (J. Sun et al., 2021). Read mapping and SNP calling from next-generation sequencing data were conducted using standard procedures (BWA and SAMtools) and the human reference genome sequence hg38 (Li & Durbin, 2010; Li et al., 2009). To obtain a confident Y-SNP dataset for reconstruction of phylogenetic tree and age estimation, we applied a series of strict filters on the original variants file, including (1) restriction to variants that are SNPs (Y-SNP); (2) removal of all positions with call rate <80% on all samples with sequences; (3) removal of position with heterozygosity call rate >5% on all samples; (4) base coverage ≥ 3 , base quality >20, and distance between SNPs >10 bp; and (5) removal of recurrent or triadic mutations. The regulations proposed by the Y Chromosome Consortium (YCC) were followed to revise the phylogenetic tree with respect to new variants in the non-recombining region of the Y chromosome (Y Chromosome Consortium, 2002). Since there is the phylogenetic

tree of Q-M120 in previous studies, the revised phylogenetic tree was constructed based on previous studies, public resources, and the genetic genealogy community. Bayesian evolutionary analyses were conducted using BEAST (v.2.0.0) (Bouckaert & Drummond, 2017). To calculate divergence times in the phylogenetic tree, a point mutation rate of 0.74×10^{-9} per site per year (Karmin et al., 2015), inferred from the ~12,000-year-old Anzick-1 male infant genome (Rasmussen et al., 2014), was applied. We referred to the regulations established by the Y Chromosome Consortium for the assignment of SNP and revised haplogroup names (Y Chromosome Consortium, 2002). New haplogroup names for sub-branches of Q1a1a-M120 are listed in Table S1.

To analyze the distribution of Q1a1a-M120 sub-lineages, Chinese participants were drawn from our in-house database, which contains various types of information for each sample, including the Y haplogroup, native province, and native city. All steps were followed as described in our previous study (J. Sun et al., 2021). In total, 3581 Q1a1a-M120 individuals were identified among 140,900 Chinese males, resulting in a frequency of 2.54%. Frequency data were utilized to create distribution maps using ArcGIS (version 10.3; Environmental Systems Research Institute Inc.). Please refer to the Supplementary Text of our previous study (J. Sun et al., 2021) for detailed information on the generation of the distribution map. ArcGIS was also used for spatial autocorrelation analysis. The Moran's I Index indicates the degree of correlation between distribution frequency and spatial geographical distribution. The HotSpots plot highlights the regions where hot spots and cold spots cluster. These regions typically correspond to the center of diffusion and its mirror image, as indicated by the overall distribution pattern. The Local Indicators of Spatial Association (Lisa) Cluster shows the clusters and outliers of the general distribution pattern of high and low frequencies.

3 | RESULTS

A schematic representation of the revised phylogeny of haplogroup Q1a1a-M120 with age estimation is shown in Figure 1. The detailed tree is available in Table S1. We identified a complex downstream structure along with numerous new SNPs and sub-branches. The distribution map and the results of the spatial autocorrelation analysis are presented in Figure 2. The Global Moran's I index (0.541720), distribution pattern (Figure 2a), and spatial clustering analysis (Figure 2b, HotSpots plot) suggest strong spatial autocorrelation. This lineage exhibits a high-frequency distribution in the northern regions of China, whereas its frequency is lower in the southern regions. Overall, northern China is the center of

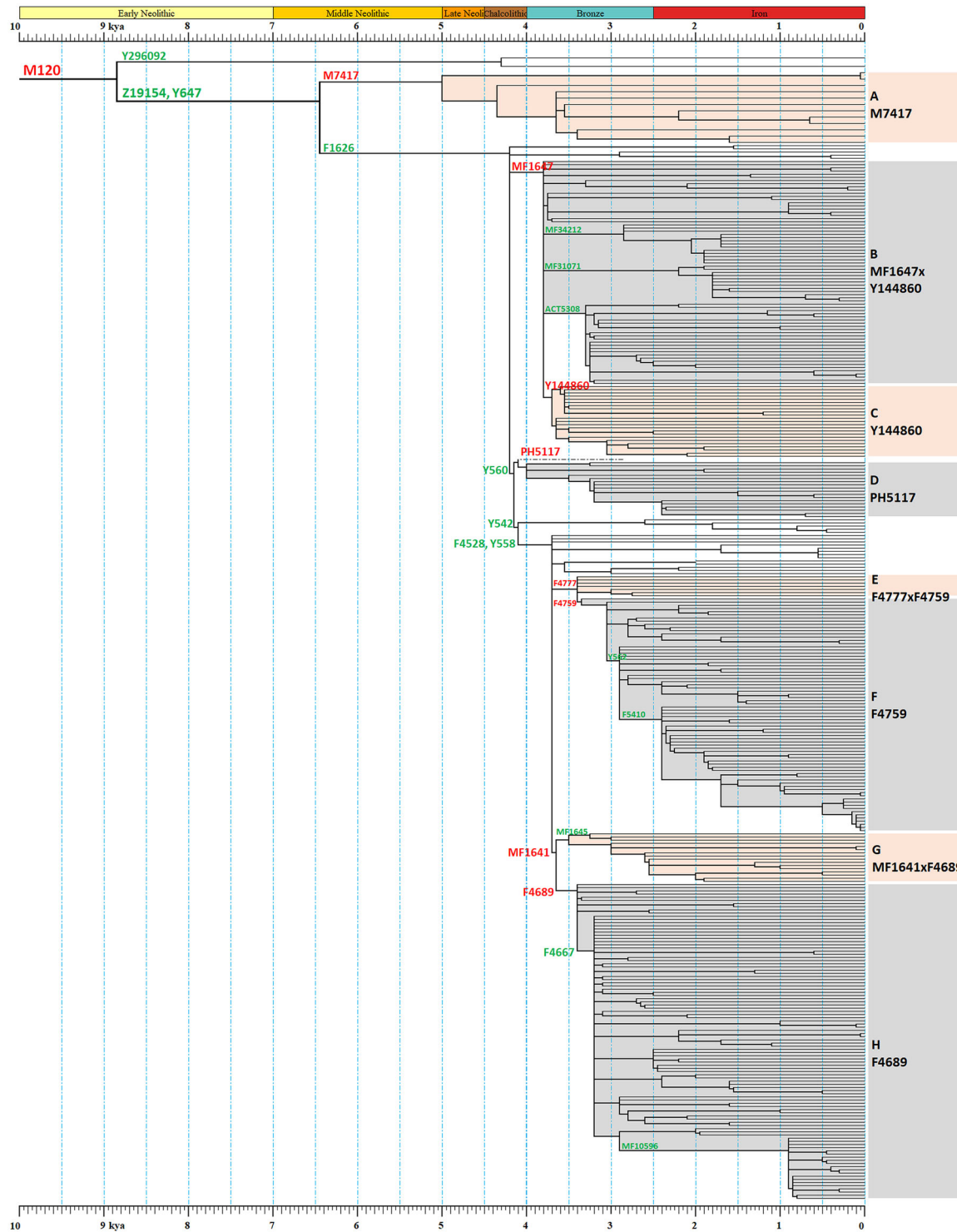


FIGURE 1 Schematic diagram of the revised phylogeny of haplogroup Q1a1a-M120. The red single nucleotide polymorphism (SNP) labels indicated the definition marker of categories of sub-branches in this study. The green SNP labels indicated other important marker of the phylogenetic tree.

diffusion for Q1a1a-M120. A Lisa Cluster map based on Local Moran's I (Figure 2c,d) indicated that Jiangxi province is a spatial outlier; it exhibited a relatively high frequency, deviating from the general pattern of low frequencies in southern China. In this study, we categorized

the sub-lineages of Q1a1a-M120 into nine distinct groups (Figure 1) and explored their geographical distribution across China (Figures 3 and S1).

The sub-lineage Q-M7417 diverged from the major sub-branch of Q-M120 approximately 6500 years ago and has

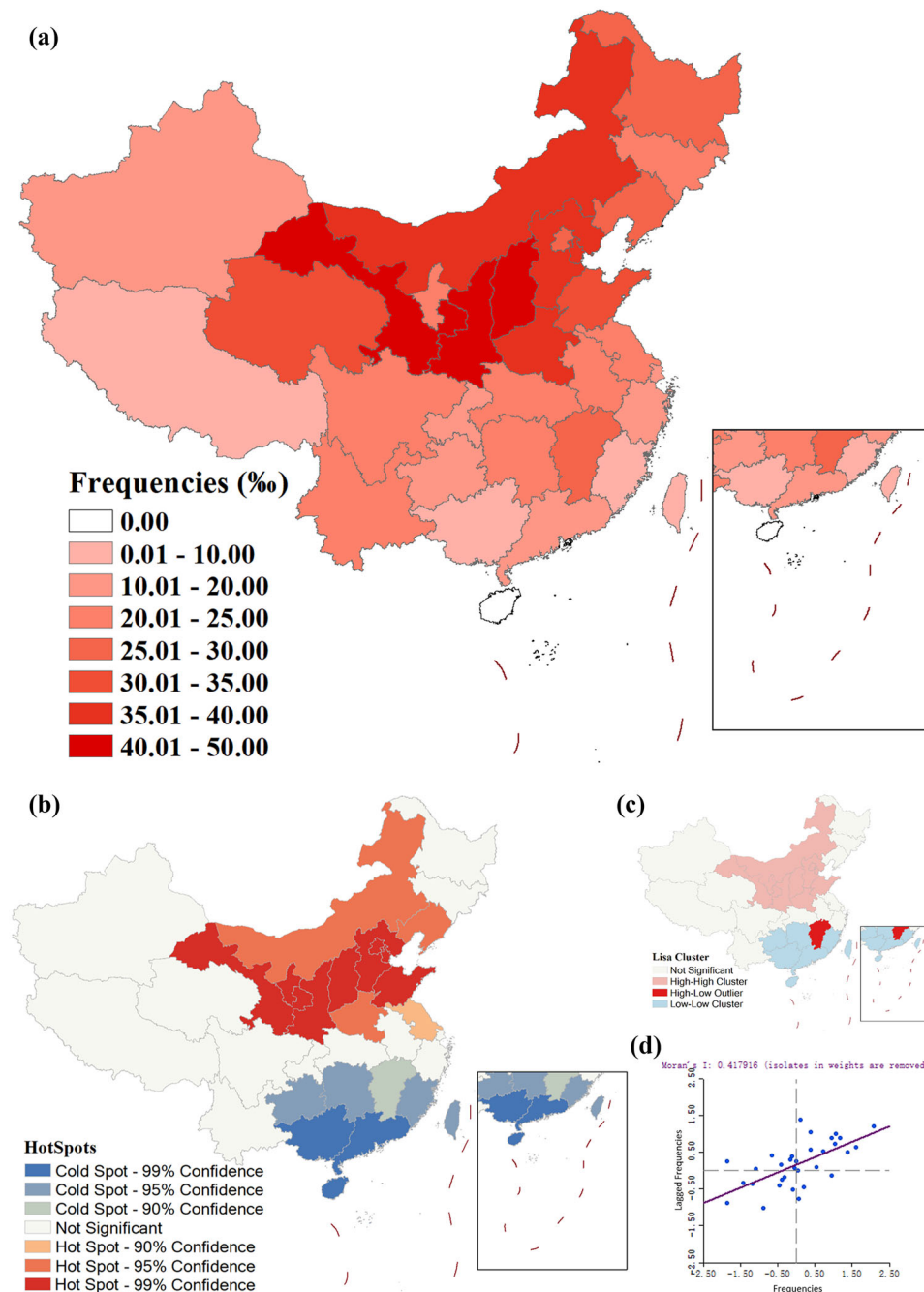


FIGURE 2 Geographic distribution of Q1a1a-M120 and results of a spatial autocorrelation analysis of haplotype frequencies: (a) distribution, (b) hotspots plot, (c) Local Indicators of Spatial Association (Lisa) analysis (local Moran's I index), and (d) Local Indicators of Spatial Association (Lisa) scatter plot frame.

experienced continuous branching over the past 5000 years (Figure 1 and Table S1). A high frequency of Q-M7417 was observed in provinces across both northern and southern China, including Taiwan, Guangxi, and Guangdong (Figure 3a). The low frequency of Q-M7417 in modern populations in central China may be attributed to the expansion of other lineages over the past 2000 years. A 3200-year-old individual with the genetic marker Q-M7417 (ID: ARS015) has been discovered at archaeological sites in

northern Mongolia (Jeong et al., 2018). We speculate that ancient populations with the Q-M7417 haplogroup spread widely throughout the Mongolian Plateau and northern China, playing a crucial role in the diffusion of Bronze Age cultures across the Gobi Desert.

Q-F1626 is a major sub-branch of Q-M120 and encompasses several significant sub-lineages, including Q-MF1647, Q-Y144860, Q-PH5117, Q-F4777, and Q-MF1641. The revised phylogeny revealed a bottle-

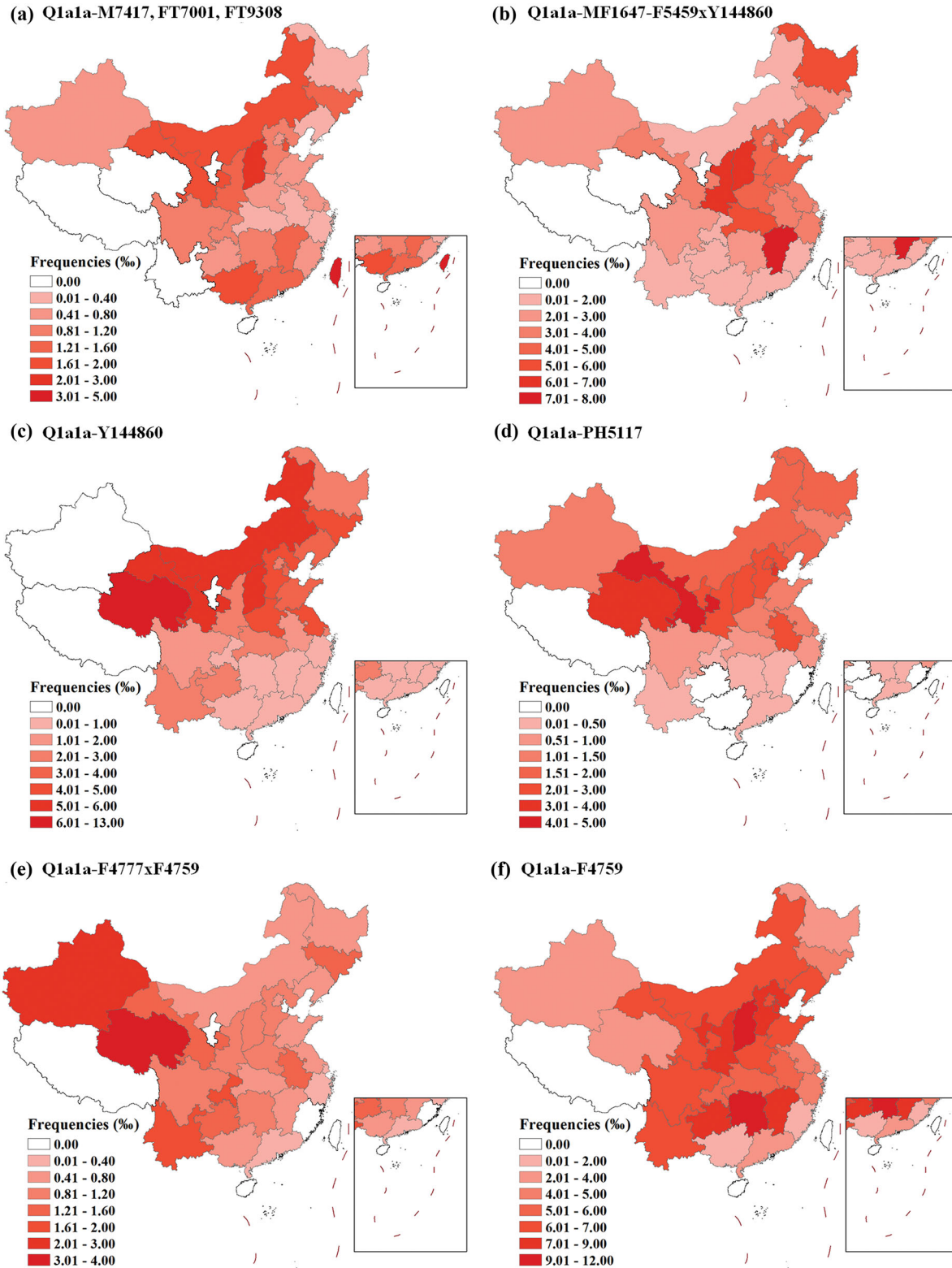


FIGURE 3 Geographic distribution of sub-lineages of Q1a1a-M120: (a) Q1a1a-M7417, FT7001, FT9308, (b) Q1a1a-MF1647-F5459xY144860, (c) Q1a1a-Y144860, (d) Q1a1a-PH5117, (e) Q1a1a-F4777xF4759, (f) Q1a1a-F4759, (g) Q1a1a-MF1641xF1827, (h) Q1a1a-F4689, F1827.

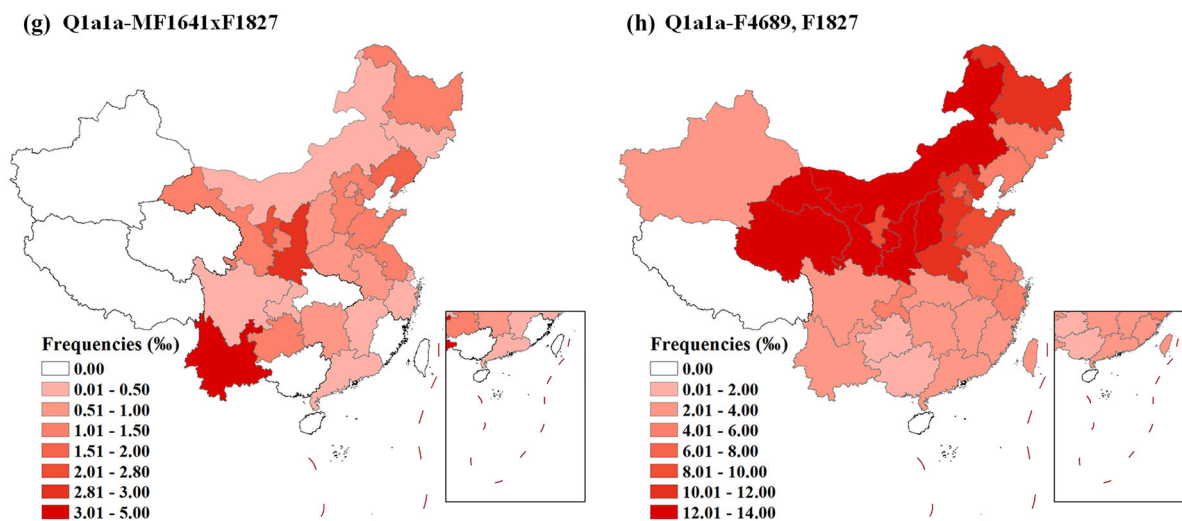


FIGURE 3 Continued

neck event lasting 2500 years for Q-F1626, followed by a significant expansion beginning 4400 years ago. The frequencies of Q*-MF1647(xY144860) (Figure 3b) and its sub-lineage Q-Y144860 (Figure 3c) were both high in the northwestern provinces of China. The high frequency of Q*-MF1647(xY144860) in Jiangxi province, located in southern China, can be attributed to the recent flourishing of a particular family. The initial expansion of sub-lineage Q-MF1647 occurred approximately 3800 years ago (Figure 1 and Table S1), coinciding with the rise of the Xia Dynasty in ancient China. This lineage could represent a major paternal lineage of either the Xia people or the barbarian tribes (Rong-Di tribes) in northwest China during the same historical period. Ancient DNA evidence is required to test this hypothesis.

Sub-lineage Q-PH5117 is a minor sub-branch of Q-M120. Beginning 4000 years ago, this haplogroup experienced a brief period of continuous differentiation (Figure 1 and Table S1). The two regions with the highest frequencies of this sub-lineage were Qinghai and Gansu in northwest China (Figure 3d). This minor haplogroup has been found among the Bhutanese (Hallast et al., 2015), the Murut people of Brunei (Karmin et al., 2015), and ancient Cimmerians in Eastern Europe (Krzewinska et al., 2018). In this study, we also identified a Q-PH5117 individual from Lhasa city in Tibet. This lineage may correspond to the Q-M242 haplogroup found among Tibetans in previous research (Qi et al., 2013). Based on these data, we propose that haplogroup Q-PH5117 represents a paternal lineage of the ancient population that established the Qijia culture (4200–3800 Years Before Present (YBP)), the first culture with bronze tools and pastoral subsistence in northwest China (F. Liu et al., 2010). The descendants of

the ancient Qijia culture may have contributed to the rise of pastoralism on the northeastern edge of the Qinghai–Tibet Plateau (M. Ma et al., 2016). Subsequently, these individuals contributed to the Tibeto–Burman-speaking populations, who later spread across a vast geographic region. Further studies are needed to uncover the detailed demographic history. Furthermore, the lineage Q-F30415 identified in our previous study (N. Sun et al., 2019) corresponds to Q-MF82700 in this study (upstream of Q-PH5117, Figure 3d). Lineage Q-FGC14753 from our previous study corresponds to Q-MF1647/F5459 in this study (Figure 3b,c). Lineage Q-SK1928 from our previous study is identical to Q-MF1642/F4719 in this study (upstream of Q-MF1645, as shown in Figure 3g).

The distributions of lineages Q*-F4777(xF4759) and Q-F4759 are shown in Figure 3e,f. The frequency of Q*-F4777(xF4759) was high in several provinces in western China. The demographic history of this lineage remains unclear. It may be related to the westward diffusion of ancient pastoral tribes from northwest China. Q-F4759 is a major sub-branch of Q-M120. This lineage began expanding continuously around 2900 years ago, coinciding with the start of the Zhou dynasty (1027 to 221 BC). Q-F4759 exhibits a high frequency in southern China (Figure 3 f). In general, we propose that Q-F4759 is the predominant paternal lineage of various clans of the Huai people (怀姓九宗), who became part of the ancient Huaxia civilization during the Zhou dynasty. Interestingly, we found that a sub-branch of Q-F4759 belongs to the “Zhang (张)” family, and this lineage can be traced back 2400 years. Ancient legends claim that the ancestor of the Zhang family, “Hui (挥),” is a descendant of the “Yellow Emperor” and the creator of the bow and arrow (L.-Q. Wang, 1981). These legends suggest that the Zhang family may have originated

from an ancient pastoralist tribe in northwest China. However, the paternal lineage Q is relatively rare in the current Chinese population with the surname Zhang (<4%). Further research is needed to explore the origin and evolution of the Zhang surname.

Sub-lineage Q-MF1645 (Figure 3g, equivalent to Q*-MF1641(xF1827)) diverged 3500 to 2000 years ago and was widely distributed in East Asia, with a high frequency in Shaanxi province. According to the raw data on clan names, these results can be attributed to the expansion of several families over the past 2000 years. The reason for the high frequency of this lineage in Yunnan province remains unclear (Figure 3g).

Sub-lineage Q-F4689 is one of the main sub-branches of Q-M120 and has undergone continuous expansion beginning 3200 years ago. As shown in Figure 3h, the haplogroup frequency is high in many provinces in northern China. The raw data show that this lineage is predominantly found among individuals with the surnames Lu and Jia. The origin of these two surnames can be traced directly to the ancient Rong-Di people (戎狄) in northern and northwestern China. Some of the Rong-Di people were also called Guifang (鬼方). In addition, there are other branches of the Rong-Di people, such as the “Red Di (赤狄)” and “White Di (白狄).” Historical analyses suggest that the ancient Rong-Di people contributed to the formation of the Han population and many other minor ethnic groups in China (L. Liu & Chen, 2003). The current distribution of the Q-F4689 lineage aligns with the demographic history of these ancient people.

4 | DISCUSSION

4.1 | Migration to the central plain of China from the Eurasian Steppe

Previous studies have demonstrated that Bronze Age globalization contributed to large-scale human migration and intense admixture of populations across nearly all peripheral regions of Eurasia (Haak et al., 2015; Lazaridis et al., 2016; Narasimhan et al., 2019). The Bronze Age culture played an essential role in the early state formation and civilization of the central plain of China; elements related to steppe culture have been discovered at archaeological sites from early China. However, demographic processes in East Asia during this period remain unclear. According to available ancient DNA data mentioned in the sections above, western Eurasian genetic lineages were nearly absent in the northern boundary region of China during the Bronze Age. This suggests that alternative demographic processes need to be identified (Kumar et al., 2022; Ning et al., 2019,

2020; Robbeets et al., 2021; C.-C. Wang et al., 2021; F. Zhang et al., 2021; Zhao et al., 2014).

To date, paternal lineages C2a-L1373, C2b-F1067, N1a2a-F1101, N1a1a-M178, O2a2b1a1a-M117, and Q1a1a-M120 have been found in archeological sites along the northern boundary of China between 5000 and 2000 years ago (Cui et al., 2020; Jeong et al., 2020; P. Ma et al., 2021; Ning et al., 2020; Robbeets et al., 2021; C.-C. Wang et al., 2021; Yang et al., 2020; F. Zhang et al., 2021). Among these lineages, only Q1a1a-M120 and N1a2a-F1101 have connections with ancient people in the eastern Eurasian Steppe and experienced significant expansions in northern China during the Bronze Age. The other haplogroups are likely lineages of indigenous populations from the central plain of China. In addition, we speculated that the ancient R1a1a-M17 lineage might have reached the northern boundary of China; however, there is no evidence to support this hypothesis.

We have previously found that haplogroup N1a2a-F1101 may be the predominant paternal lineage in the royal family of the Zhou dynasty (P. Ma et al., 2021). Early sub-lineages of N1a2a-F1101 have been found from northeast China to western Siberia. The central plain of China is likely the center of diffusion for haplogroup N1a2a-F1101; however, its origin may lie farther north. In this study, we further explored the divergence time and geographic distribution of the Q-M120. Sub-lineage Q-F7417, the earliest sub-branch of Q-M120, was found in the Mongolian Plateau, whereas other sub-branches were primarily distributed in central China. Many sub-branches of Q-M242 were distributed across Eurasia, with an early center of dispersal in southern Siberia. Haplogroup Q1a1b-B143, which is closely related to Q1a1a-M120, is one of the main paternal lineages of the Chukotko–Kamchatkan-speaking population in the Arctic region. Overall, a similar pattern of divergence was observed for N1a2a-F1101 and Q1a1a-M120. These haplogroups were primarily distributed in China, whereas closely related branches were mainly found in North Asia.

In general, we propose that ancient populations, in which haplogroups N1a2a-F1101 and Q1a1a-M120 were the main paternal lineages, played an important role in introducing Bronze Age culture to the central plains of China. During the Chalcolithic and Bronze Age, these populations might have been distributed both to the north and south of the Gobi Desert between China and Mongolia. They had a wide range of activities as hunter-gatherers and were well acquainted with routes through the Gobi Desert. The southward expansion of pastoralism and Bronze Age culture from the Mongolian Plateau, dominated by haplogroups N1a2a-F1101 and Q1a1a-M120, was driven by global cooling 4200 to 4000 years ago. Subsequently, haplogroups N1a2a-F1101 and Q1a1a-M120 experienced

significant expansions in northern China. Other paternal lineages may be distributed far from the Gobi Desert, suggesting that they were not involved in the final stage of introducing Bronze Age culture to the central plains of China.

4.2 | Rise of pastoralists in northwest China contributed to the formation of early China

According to archaeological and historical research, early pastoralism in northern and northwestern China emerged between 4500 and 2000 years ago (L. Liu, 2004; L. Liu & Chen, 2003). These pastoralists were referred to as “Rong-Di (戎狄)” in ancient historical records and were distinctly characterized by their use of a wide variety of metal weapons and their high mobility. The Rong-Di people were believed to have a different ancestry from the ancient populations of the central plain of China. They gradually mixed with the Huaxia people and other East Asian populations around 2000 years ago.

Different cultural and lifestyle landscapes of the Rong-Di people have been revealed through archaeological and historical research. Archeological evidence has shown that the Rong-Di people are descendants of local Neolithic farmers and migrants from across the Gobi Desert. They exhibited a mixed system of pastoralism and millet civilization, and formed a city-like settlement. The similarity between the cultural relics of the Rong-Di people and those found in the Eurasian Steppe may be attributed to the routes through the Gobi Desert, which facilitated the long-distance exchange of materials across the Eurasian Steppe. Overall, the long-distance exchange of materials and cultural elements was the primary connection between North China and the Eurasian Steppe, whereas gene flow remained limited during the Bronze Age. Ancient DNA data and the distributions of haplogroups N1a2a-F1101 and Q1a1a-M120 detected in this study support this scenario.

Archaeological and historical studies have indicated that pastoralists in northwest China played an important role in the early state formation and civilization of the central plain of China. Previous studies combined with our analyses support this view. The Huaxia people were formed through the population admixture of farmers from the Yellow River Basin and Yangtze River Basin with pastoralists from northern and northwestern China. During the formation of the ancient Huaxia people, N1a2a-F1101 and Q1a1a-M120 became the dominant paternal lineages.

4.3 | History of Han people since the Qin and Han dynasties

Approximately 4500 to 2000 years ago, admixture between the Huaxia people and ancient populations from peripheral regions of East Asia, including the Rong-Di people, contributed to the formation of the Han populations during the Han dynasty. In this study, we found that Q-MF1647/F5459 experienced significant expansion in North China during the Xia dynasty (2070 to 1600 BC). Ancient individuals with Q-MF1647/F5459 may have contributed substantially to populations in the central plain during this period. However, relevant ancient DNA evidence is currently lacking. Q-F1827 is a primary sub-clade of Q-M120. Modern individuals with Q-F1827, often bearing the surnames Lu and Jia, may be descendants of the Rong-Di people from northern and northwestern China during the Zhou dynasty. The surname Zhang is relatively common among modern individuals with the Q-F4759 haplogroup and can be traced back to the Zhou dynasty. In general, the Rong-Di people and their descendants played an essential role in the formation of China, Chinese culture, and the Han people.

In conclusion, we reconstructed a high-resolution phylogenetic tree and determined the divergence times and geographic distribution of sub-clades based on 347 newly generated Q1a1a-M120 sequences. We detected a complex substructure of haplogroup Q-M120, which is primarily distributed in northern and northwestern China. We propose that haplogroup Q-M120 was the dominant paternal lineage of the Rong-Di people in northern and northwestern China 4000 years ago. These people were significantly influenced by the eastward expansion of Eurasian Steppe pastoralists during the Bronze Age and exhibited a mixed subsistence pattern of agriculture and pastoralism. The ancient pastoralists of northern and northwestern China facilitated the exchange of materials and culture between eastern and western Eurasia. They played an indispensable role in the formation of early China, Chinese culture, and the Huaxia people, the ancestors of the Han people.

AUTHOR CONTRIBUTIONS

Lan-Hai Wei and Jin Sun designed this study and revised the manuscript. La-Su Mai wrote this manuscript; Xian-Peng Zhang and Kai-Jun Liu analyzed the data and visualized results. Material preparation and data collection were performed by Hui Li and Peng-Cheng Ma. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank all donors for providing DNA samples and/or DNA sequences. This study was supported by the key projects of strategic international scientific and technological innovation cooperation of the Chinese Ministry of Science and Technology (2020YFE0201600), the National Natural Science Foundation of China (31900406 to LHW), and the Scientific and Technology Committee of Shanghai Municipality (18490750300). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.


CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Following the regulations of the Human Genetic Resources Administration of China (HGRAC), the raw sequence data reported in this paper are available on request from the corresponding author.

ORCID

Xian-Peng Zhang  <https://orcid.org/0000-0003-2145-4223>

REFERENCES

- Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, 17(1), 42. <https://doi.org/10.1186/s12862-017-0890-6>
- Cui, Y., Zhang, F., Ma, P., Fan, L., Ning, C., Zhang, Q., Zhang, W., Wang, L., & Robbeets, M. (2020). Bioarchaeological perspective on the expansion of Transeurasian languages in Neolithic Amur River basin. *Evolutionary Human Sciences*, 2, E15. <https://doi.org/10.1017/ehs.2020.16>
- Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Bánffy, E., Economou, C., Francken, M., Friederich, S., Pena, R. G., Hallgren, F., Khartanovich, V., ... Reich, D. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), 207–211. <https://doi.org/10.1038/nature14317>
- Hallast, P., Batini, C., Zadik, D., Maisano Delser, P., Wetton, J. H., Arroyo-Pardo, E., Cavalleri, G. L., De Knijff, P., Destro Bisol, G., Dupuy, B. M., Eriksen, H. A., Jorde, L. B., King, T. E., Larmuseau, M. H., Lopez De Munain, A., Lopez-Parra, A. M., Loutradis, A., Milasin, J., Novelletto, A., ... Jobling, M. A. (2015). The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. *Molecular Biology and Evolution*, 32(3), 661–673. <https://doi.org/10.1093/molbev/msu327>
- Jeong, C., Balanovsky, O., Lukianova, E., Kahbatkyzy, N., Flegontov, P., Zaporozhchenko, V., Immel, A., Wang, C.-C., Ixan, O., Khussainova, E., Bekmanov, B., Zaibert, V., Lavryashina, M., Pocheshkhova, E., Yusupov, Y., Agdzhoian, A., Koshel, S., Bukin, A., Nymadawa, P., ... Krause, J. (2019). The genetic history of admixture across inner Eurasia. *Nature Ecology & Evolution*, 3(6), 966–976. <https://doi.org/10.1038/s41559-019-0878-2>
- Jeong, C., Wang, K. E., Wilkin, S., Taylor, W. T. T., Miller, B. K., Bemmman, J. H., 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., ... Warinner, C. (2020). A dynamic 6,000-year genetic history of Eurasia's Eastern Steppe. *Cell*, 183(4), 890–904.e829. <https://doi.org/10.1016/j.cell.2020.10.015>
- Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A. S., Taylor, W. T. T., Hagan, R. W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., Littleton, J., Makarewicz, C. A., Krigbaum, J., Burri, M., Scott, A., Davaasambuu, G., Wright, J., Irmer, F., Myagmar, E., ... Warinner, C. (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *PNAS*, 115(48), E11248–e11255. <https://doi.org/10.1073/pnas.1813608115>
- Karmin, M., Saag, L., Vicente, M., Sayres, M. A. W., Järve, M., Talas, U. G., Rootsi, S., Ilumäe, A.-M., Mägi, R., Mitt, M., Pagani, L., Puurand, T., Faltyskova, Z., Clemente, F., Cardona, A., Metspalu, E., Sahakyan, H., Yunusbayev, B., Hudjashov, G., ... Kivisild, T. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research*, 25(4), 459–466. <https://doi.org/10.1101/gr.186684.114>
- Krzewińska, M., Kılınc, G. M., Juras, A., Koptekin, D., Chyleński, M., Nikitin, A. G., Shcherbakov, N., Shuteleva, I., Leonova, T., Kraeva, L., Sungatov, F. A., Sultanova, A. N., Potekhina, I., Łukasik, S., Krenz-Niedbała, M., Dalén, L., Sinika, V., Jakobsson, M., Storå, J., & Götherström, A. (2018). Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Science Advances*, 4(10), eaat4457. <https://doi.org/10.1126/sciadv.aat4457>
- Kumar, V., Wang, W., Zhang, J., Wang, Y., Ruan, Q., Yu, J., Wu, X., Hu, X., Wu, X., Guo, W., Wang, B., Niyazi, A., Lv, E., Tang, Z., Cao, P., Liu, F., Dai, Q., Yang, R., Feng, X., ... Fu, Q. (2022). Bronze and Iron Age population movements underlie Xinjiang population history. *Science*, 376(6588), 62–69. <https://doi.org/10.1126/science.abk1534>
- Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D. C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., Connell, S., Stewardson, K., Harney, E., Fu, Q., Gonzalez-Forbes, G., Jones, E. R., Roodenberg, S. A., Lengyel, G., Bocquentin, F., ... Reich, D. (2016). Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), 419–424. <https://doi.org/10.1038/nature19310>
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liu, F., Zhang, Y., Feng, Z., Hou, G., Zhou, Q., & Zhang, H. (2010). The impacts of climate change on the Neolithic cultures of Gansu-

- Qinghai region during the late Holocene Megathermal. *Journal of Geographical Sciences*, 20(3), 417–430. <https://doi.org/10.1007/s11442-010-0417-1>
- Liu, L. (2004). *The Chinese Neolithic: Trajectories to early states*. Cambridge University Press.
- Liu, L., & Chen, X.-C. (2003). *State formation in early China*. Duckworth.
- Liu, L., & Chen, X.-C. (2012). *The archaeology of China: From the late Paleolithic to the early Bronze Age*. Cambridge University Press.
- Ma, B.-C. (2007). *Kao Gu Yu WenWu (Archaeology and Cultural Relics)*, 6(6), 37–43.
- Ma, M., Dong, G., Jia, X., Wang, H., Cui, Y., & Chen, F. (2016). Dietary shift after 3600 cal yr BP and its influencing factors in northwestern China: Evidence from stable isotopes. *Quaternary Science Reviews*, 145, (Supplement C), 57–70. <https://doi.org/10.1016/j.quascirev.2016.05.041>
- Ma, P., Yang, X., Yan, S., Li, C., Gao, S., Han, B., Hou, K., Robbeets, M., Wei, L.-H., & Cui, Y. (2021). Ancient Y-DNA with reconstructed phylogeny provides insights into the demographic history of paternal haplogroup N1a2-F1360. *J Genet Genomics*, 48(12), 1130–1133. <https://doi.org/10.1016/j.jgg.2021.07.018>
- Narasimhan, V. M., Patterson, N., Moorjani, P., Rohland, N., Bernardos, R., Mallick, S., Lazaridis, I., Nakatsuka, N., Olalde, I., Lipson, M., Kim, A. M., Olivieri, L. M., Coppa, A., Vidale, M., Mallory, J., Moiseyev, V., Kitov, E., Monge, J., Adamski, N., ... Reich, D. (2019). The formation of human populations in South and Central Asia. *Science*, 365(6457), eaat7487. <https://doi.org/10.1126/science.aat7487>
- Ning, C., Li, T., Wang, K., Zhang, F., Li, T., Wu, X., Gao, S., Zhang, Q., Zhang, H., Hudson, M. J., Dong, G., Wu, S., Fang, Y., Liu, C., Feng, C., Li, W., Han, T., Li, R., Wei, J., ... Cui, Y. (2020). Ancient genomes from northern China suggest links between subsistence changes and human migration. *Nature Communications*, 11(1), 2700. <https://doi.org/10.1038/s41467-020-16557-2>
- Ning, C., Wang, C.-C., Gao, S., Yang, Y., Zhang, X., Wu, X., Zhang, F., Nie, Z., Tang, Y., Robbeets, M., Ma, J., Krause, J., & Cui, Y. (2019). Ancient genomes reveal Yamnaya-related ancestry and a potential source of Indo-European speakers in Iron Age Tianshan. *Current Biology*, 29(15), 2526–2532.e2524. <https://doi.org/10.1016/j.cub.2019.06.044>
- Qi, X., Cui, C., Peng, Y. I., Zhang, X., Yang, Z., Zhong, H., Zhang, H., Xiang, K., Cao, X., Wang, Y. I., Ouzhuluobu, Basang, Ciwangsangbu, Bianba, Gonggalanzi, Wu, T., Chen, H., Shi, H., & Su, B. (2013). Genetic evidence of Paleolithic colonization and Neolithic expansion of modern humans on the Tibetan Plateau. *Molecular Biology and Evolution*, 30(8), 1761–1778. <https://doi.org/10.1093/molbev/mst093>
- Rasmussen, M., Anzick, S. L., Waters, M. R., Skoglund, P., Degiorgio, M., Stafford, T. W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S. M., Poznik, G. D., Gudmundsdottir, V., Yadav, R., Malaspina, A.-S., White, S. S. 5th, Allentoft, M. E., Cornejo, O. E., Tambets, K., Eriksson, A., ... Willerslev, E. (2014). The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature*, 506(7487), 225–229. <https://doi.org/10.1038/nature13025>
- Robbeets, M., Bouckaert, R., Conte, M., Savelyev, A., Li, T., An, D.-I., Shinoda, K.-I., Cui, Y., Kawashima, T., Kim, G., Uchiyama, J., Dolińska, J., Oskolskaya, S., Yamano, K.-Y., Seguchi, N., Tomita, H., Takamiya, H., Kanzawa-Kiriyama, H., Oota, H., ... Ning, C. (2021). Triangulation supports agricultural spread of the Transeurasian languages. *Nature*, 599(7886), 616–621. <https://doi.org/10.1038/s41586-021-04108-8>
- Rootsi, S., Zhivotovsky, L. A., Baldovici, M., Kayser, M., Kutuev, I. A., Khusainova, R., Bermisheva, M. A., Gubina, M., Fedorova, S. A., Ilumäe, A.-M., Khusnutdinova, E. K., Voevoda, M. I., Osipova, L. P., Stoneking, M., Lin, A. A., Ferak, V., Parik, J., Kivisild, T., Underhill, P. A., & Villems, R. (2007). A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics*, 15(2), 204–211. <https://doi.org/10.1038/sj.ejhg.5201748>
- Sun, J., Li, Y.-X., Ma, P.-C., Yan, S., Cheng, H.-Z., Fan, Z.-Q., Deng, X.-H., Ru, K., Wang, C.-C., Chen, G., & Wei, L.-H. (2021). Shared paternal ancestry of Han, Tai-Kadai-speaking, and Austronesian-speaking populations as revealed by the high resolution phylogeny of O1a-M119 and distribution of its sub-lineages within China. *American Journal of Physical Anthropology*, 174(4), 686–700. <https://doi.org/10.1002/ajpa.24240>
- Sun, N., Ma, P.-C., Yan, S., Wen, S.-Q., Sun, C., Du, P.-X., Cheng, H.-Z., Deng, X.-H., Wang, C.-C., & Wei, L.-H. (2019). Phylogeography of Y-chromosome haplogroup Q1a1a-M120, a paternal lineage connecting populations in Siberia and East Asia. *Annals of Human Biology*, 46(3), 261–266. <https://doi.org/10.1080/03014460.2019.1632930>
- Wang, C.-C., Yeh, H.-Y., Popov, A. N., Zhang, H.-Q., Matsumura, H., Sirak, K., Cheronet, O., Kovalev, A., Rohland, N., Kim, A. M., Mallick, S., Bernardos, R., Tumen, D., Zhao, J., Liu, Y.-C., Liu, J.-Y., Mah, M., Wang, K., Zhang, Z., ... Reich, D. (2021). Genomic insights into the formation of human populations in East Asia. *Nature*, 591(7850), 413–419. <https://doi.org/10.1038/s41586-021-03336-2>
- Wang, L.-Q. (1981). *Annotation of Feng Shu Tong Yi (in Chinese)*. Zhonghua Book Company.
- Y Chromosome Consortium. (2002). A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research*, 12(2), 339–348. <https://doi.org/10.1101/gr.217602>
- Yang, M. A., Fan, X., Sun, B., Chen, C., Lang, J., Ko, Y.-C., Tsang, C.-H., Chiu, H., Wang, T., Bao, Q., Wu, X., Hajdinjak, M., Ko, A. M.-S., Ding, M., Cao, P., Yang, R., Liu, F., Nickel, B., Dai, Q., ... Fu, Q. (2020). Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science*, 369(6501), 282–288. <https://doi.org/10.1126/science.aba0909>
- Zhang, F., Ning, C., Scott, A., Fu, Q., Bjørn, R., Li, W., Wei, D., Wang, W., Fan, L., Abuduresule, I., Hu, X., Ruan, Q., Niyazi, A., Dong, G., Cao, P., Liu, F., Dai, Q., Feng, X., Yang, R., ... Cui, Y. (2021). The genomic origins of the Bronze Age Tarim Basin mummies. *Nature*, 599(7884), 256–261. <https://doi.org/10.1038/s41586-021-04052-7>
- Zhang, Y. (2018). Origin of Yanglang type of Xirong culture of the eastern Zhou Dynasty. *Journal of National Museum of China*, 29(3), 29–35.
- Zhao, Y.-B., Li, H.-J., Cai, D.-W., Li, C.-X., Zhang, Q.-C., Zhu, H., & Zhou, H. (2010). Ancient DNA from nomads in 2500-year-old archeological sites of Pengyang, China. *Journal of Human Genetics*, 55(4), 215–218. <https://doi.org/10.1038/jhg.2010.8>

- Zhao, Y.-B., Zhang, Y., Li, H. J., Cui, Y. Q., Zhu, H., & Zhou, H. (2014). Ancient DNA evidence reveals that the Y chromosome haplogroup Q1a1 admixed into the Han Chinese 3,000 years ago. *American Journal of Human Biology*, 26(6), 813–821. <https://doi.org/10.1002/ajhb.22604>
- Zhong, H., Shi, H., Qi, X.-B., Duan, Z.-Y., Tan, P.-P., Jin, L., Su, B., & Ma, R. Z. (2011). Extended Y chromosome investigation suggests post-glacial migrations of modern humans into East Asia via the northern route. *Molecular Biology and Evolution*, 28(1), 717–727. <https://doi.org/10.1093/molbev/msq247>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mai, L.-S., Zhang, X.-P., Liu, K.-J., Ma, P.-C., Li, H., Sun, J., & Wei, L.-H. (2025). Traces of Bronze Age globalization in East Asia: Insights from a revised phylogeography of the Y-chromosome haplogroup Q1a1a-M120. *Annals of Human Genetics*, 89, 12–23. <https://doi.org/10.1111/ahg.12580>