Science Bulletin 66 (2021) 1129–1135

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Article

Maternal genetic structure in ancient Shandong between 9500 and 1800 years ago

Juncen Liu [a,b,c,1], Wen Zeng [d,1], Bo Sun [e,1], Xiaowei Mao [a,b,1], Yongsheng Zhao [d], Fen Wang [f], Zhenguang Li [e], Fengshi Luan [f], Junfeng Guo [g], Chao Zhu [e], Zimeng Wang [e], Chengmin Wei [e], Ming Zhang [a,b,c], Peng Cao [a], Feng Liu [a], Qingyan Dai [a], Xiaotian Feng [a], Ruowei Yang [a], Weihong Hou [a,b], Wanjing Ping [a], Xiaohong Wu [h], E. Andrew Bennett [a,b], Yichen Liu [a,b], Qiaomei Fu [a,b,c,*]

[a] Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China
[b] Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing 100044, China
[c] University of Chinese Academy of Sciences, Beijing 100049, China
[d] Institute of Cultural Heritage, Shandong University, Qingdao 266237, China
[e] Shandong Provincial Institute of Cultural Relics and Archaeology, Jinan 250012, China
[f] School of History and Culture, Shandong University, Jinan 250100, China
[g] Jinan Municipal Institute of Archaeology, Jinan 250062, China
[h] School of Archaeology and Museology, Peking University, Beijing 100871, China

ARTICLE INFO

Article history:
Received 24 May 2020
Received in revised form 6 August 2020
Accepted 28 September 2020
Available online 26 January 2021

Keywords:
Ancient DNA
Human
Mitochondrial DNA
East Asia
Neolithic

ABSTRACT

Archaeological and ancient DNA studies revealed that Shandong, a multi-culture center in northern coastal China, was home to ancient populations having ancestry related to both northern and southern East Asian populations. However, the limited temporal and geographical range of previous studies have been insufficient to describe the population history of this region in greater detail. Here, we report the analysis of 86 complete mitochondrial genomes from the remains of 9500 to 1800-year-old humans from 12 archaeological sites across Shandong. For samples older than 4600 years before present (BP), we found haplogroups D4, D5, B4c1, and B5b2, which are observed in present-day northern and southern East Asians. For samples younger than 4600 BP, haplogroups C (C7a1 and C7b), M9 (M9a1), and F (F1a1, F2a, and F4a1) begin to appear, indicating changes in the Shandong maternal genetic structure starting from the beginning of the Longshan cultural period. Within Shandong, the genetic exchange is possible between the coastal and inland regions after 3100 BP. We also discovered the B5b2 lineage in Shandong populations, with the oldest Bianbian individual likely related to the ancestors of some East Asians and North Asians. By reconstructing a maternal genetic structure of Shandong populations, we provide greater resolution of the population dynamics of the northern coastal East Asia over the past nine thousand years.

1. Introduction

Ancient DNA (aDNA) studies have helped clarify human population history and evolution in East Eurasia [1–9], such as revealing a genetic distinction between ancient populations in northern and southern East Asia. Shandong province is situated in the northern coastal region of China, and is geographically connected to both northern and southern China. As a multi-culture center, Shandong was home to the Dawenkou culture (6000–4600 years before present (BP), distributed in the lower Yellow River [10]) and the Shandong Longshan culture (4600–4000 BP, a local manifestation of the Longshan Culture [10]). Previous aDNA studies focused on the short hypervariable region I of the mitochondrial DNA (mtDNA) of several Shandong remains [10–13], but these had limited power identifying haplogroups. Recently, a genome-wide study [6] and a review of the human population history of Eastern Eurasia [9] revealed that ancient Shandong individuals played an important role in defining the population genetic history and past migrations of East Asia: six ancient Shandong individuals (~9500–7700 BP) from four Early Neolithic archaeological sites were shown to have ancestry related to northern East Asian populations, an ancestral component that later spread into southern East Asia [6]. However,

* Corresponding author.
E-mail address: fuqiaomei@ivpp.ac.cn (Q. Fu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.scib.2021.01.029
2095-9273/

the above studies were limited by a scarcity of samples, leaving large gaps in our understanding of the changing genetic relationships over time across Shandong and its neighboring regions.

In this study, we present the analyses of 86 high-quality complete mtDNA sequences of ancient individuals (~9500–1800 BP) collected from the northern Chinese province of Shandong (12 archaeological sites, Fig. 1a). Our study provides a wider and finer temporal and geographical resolution than previous work in describing the population dynamics regarding Shandong's maternal genetic structure, and how these populations connected with those outside of Shandong.

2. Materials and methods

2.1. Ancient DNA extraction and library preparation

DNA was extracted from 88 samples (aged between 9500 and 1800 BP from 12 sites) including temporal bones, limb bones, phalanges, and molars (Table S1 online) using a method described previously [14]. A total of 88 double-stranded libraries were constructed according to previously published protocols [15,16] and all were partially treated with uracil DNA glycosylase ("UDG. half") [17].

2.2. In-solution capture and analysis of mitochondrial DNA

Libraries were hybridized with oligonucleotide probes that overlapped with the mitochondrial genome as previously described [18]. From the enriched mitochondrial DNA libraries, we used the Illumina Miseq platform to generate 2 × 76 bp paired-end reads. After sequencing, forward and reverse reads that overlapped by at least 11 bp were merged into a single sequence, with one mismatch allowed [18]. The merged sequences (≥30 bp in length) were mapped to the mtDNA revised Cambridge Reference Sequence using BWA (version 0.5.10evan.91g44db244, https://bitbucket.org/ustenzel/networkawarebwa) [19], after which duplicate reads and those with mapping qualities lower than 30 were removed. The subset of libraries that included the characteristic ancient DNA damage at the first position of the 5′ end and the last position of the 3′ end are noted as "DS.half" [16].

The contamination rate was estimated by calculating the fraction of ancient mtDNA fragments compared to 311 published modern human mitochondrial genomes using ContamMix [20,21]. After manually excluding two individuals with large numbers of missing positions, our final dataset included 86 complete mtDNA genomes (16,569 bp) with high coverage (101- to 591-fold) and low contamination rates (0.01% to 4.52%) (Table S1 online).

2.3. Genetic analysis

To compare with present-day East Asian populations, we prepared mtDNA haplogroup frequencies from published studies [10,22] (Table S3 online) and those of 528 complete mitochondrial sequences (Table S4 online) from populations across East Asia. For our newly sampled individuals, HaploGrep2 and PhyloTree build 17 were used to call the haplogroup of each mtDNA sequence [23,24]. Certain haplotypes were found to be shared by individuals (ranging from 2 to 5) buried in the same graves at the Beiqian site, which may have been due to matrilineal kinship among these individuals (Table S1 online). In these cases, only one individual was used for the analyses. Principal component analysis (PCA) of haplogroup frequencies was carried out to investigate the general population structure. The sequences of the mitochondrial genomes from the ancient samples were aligned using MUSCLE v3.8.3 [25]. Genetic distances (Φst) and corresponding P-values were calculated by Arlequin 3.518 [26]. We also used Arlequin 3.518 [26] to carry out the analysis of molecular variance (AMOVA). To avoid artificially low genetic diversity at certain sites, sites covered by fewer than five individuals were excluded from the PCA and correlation coefficient analyses of haplogroup frequency. The correlation coefficients (R) of the frequencies of different haplogroups in ancient and modern individuals were calculated using the following equation:

$$R(x,y) = \frac{\sum(x-\bar{x})(\hat{y}-\bar{y})}{\sqrt{\sum(x-\bar{x})^2(\hat{y}-\bar{y})^2}},$$

where $0 < R < 1$ means a positive correlation and $-1 < R < 0$ means a negative correlation, x and y are the frequencies of a given haplogroup of the two populations [27]. The network was constructed with complete mtDNA sequences using the median-joining method in the Popart software [28,29]. The Time of the Most Recent Common Ancestor (TMRCA) was estimated by BEAST [30] with the HYK + G + I model because it resulted in the best fit for our data according to jModelTest 2 [31]; we also investigated a model with a Constant Population Size and Strict Clock [20]. For this model, the Markov-Chain Monte Carlo (MCMC) ran for 100,000,000 iterations sampling every 10,000 steps.

3. Results and discussion

3.1. Genetic structure separated by 4600 BP

Complete mtDNA genomes (16,569 bp) were obtained from bone and tooth samples of 86 out of 88 ancient individuals aged between 9500 and 1800 BP from 12 sites in Shandong. From these 86 individuals, we identified a total of 56 haplotypes (Table S1 online). Certain haplotypes (e.g., A5a and B5b2a2) were found to be shared by two to five individuals buried in the same graves at the Beiqian site, which may indicate a matrilineal kinship among them (Table S1 online). For samples having possible matrilineal kinship, only one was kept for the subsequent genetic analyses (10 individuals were removed). The 56 haplotypes could be assigned to 13 basal mtDNA haplogroups [32–37]: A, B, C, D4, D5, F, G, M8, M9, M11, N9a, R, and Z. We used AMOVA to test which population groupings, based on the age of samples, would best represent the genetic structure of all Shandong sites (Table S2 online). We found that the highest variance (variance among groups = 0.17%) was observed when populations were divided into groups older and younger than 4600 BP (group sizes were 27 and 49), which coincides with the interval between the Dawenkou (6000–4600 BP) and Longshan cultural (4600–4000 BP) periods.

3.2. Maternal genetic structure in Shandong before 4600 BP

Twenty-eight individuals were dated to before 4600 BP: a ~9500-year-old individual from Bianbian, an ~8200-year-old Xiaogao individual, an ~8200-year-old Boshan individual [20], three 8000–7700-year-old Xiaojingshan individuals, three 6000 to 4600-year-old individuals from Fujia site, and nineteen 5500 to 5300 BP individuals from Beiqian site (Fig. 1a and Table S1 online). Among them, the samples older than 4600 BP show a high frequency of haplogroups B (B4c1 and B5b2, 24.14%) and D (D4 and D5, 37.93%). The haplogroup B4 is mainly distributed in present-day central-eastern Asians [34], while haplogroup B4c1 is found predominantly in southern East Asians (61.70%), as well as populations from Southeast Asia (Mueang, Tay, PhuLa, Kinh, and LaHu; 25.53%) that are closely related to southern East Asians (Table S6 online). Haplogroup B5b is found widely across East Asia, with its highest diversity in present-day Koreans [34]. B5b2 represents a

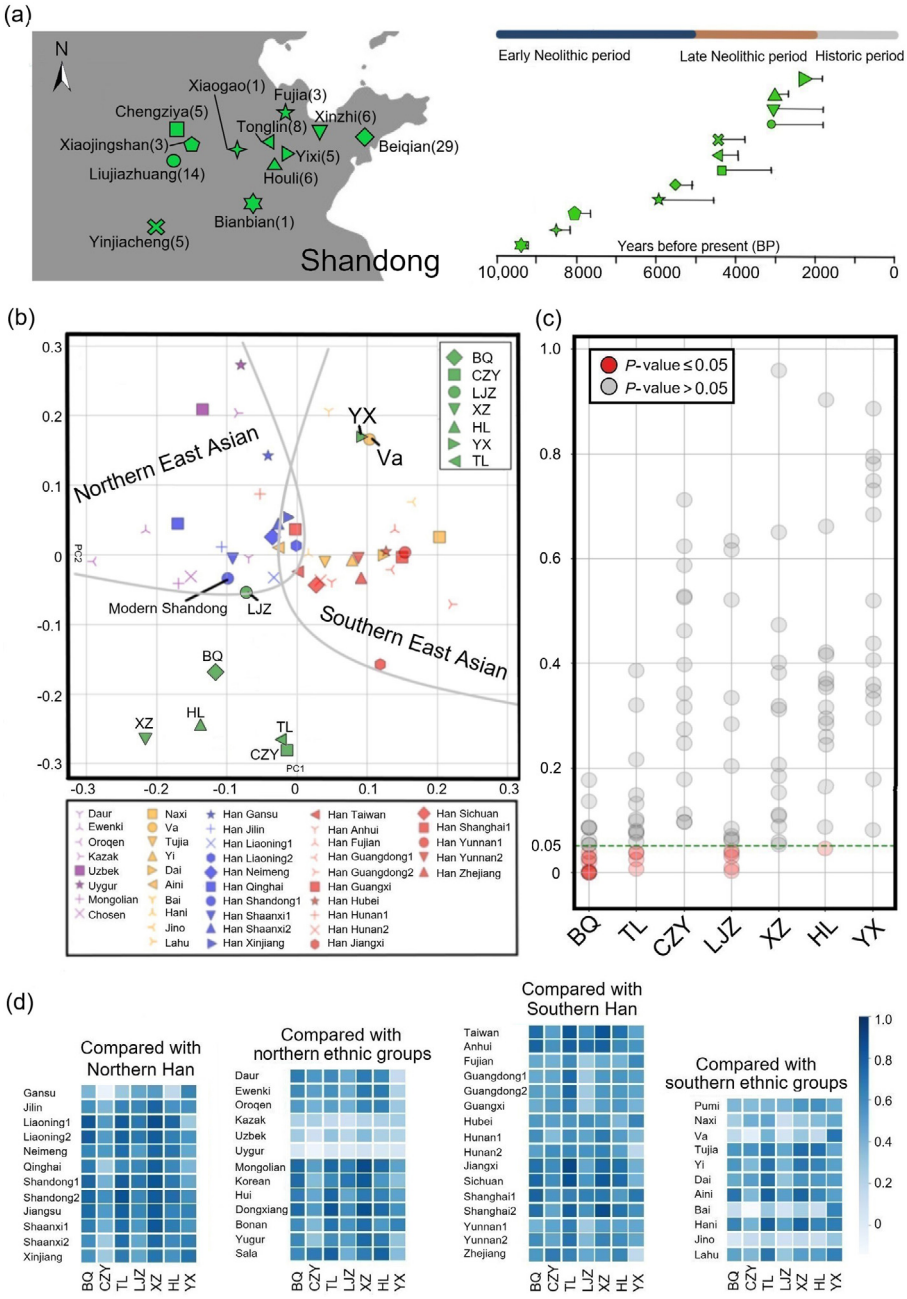


Fig. 1. Genetic analysis of ancient Shandong populations. (a) Geographical and temporal distribution of the ancient humans from 12 archaeological sites across the Shandong Province; (b) PCA plot of the shared mtDNA haplogroup frequencies between the ancient Shandong populations and present-day East Asian populations, showing changes in the maternal genetic structure over time (BQ = Beiqian, CZY = Chengziya, TL = Tonglin, LJZ = Liujiazhuang, XZ = Xinzhi, HL = Houli, and YX = Yixi). Present-day East Asian populations form two major clines, denoted as "northern East Asian" and "southern East Asian"; (c) The *P*-values of genetic distances (Φst) between the ancient Shandong populations and various present-day East Asian populations (Table S4 online) (dots above green dash line have *P*-values > 0.05, representing no significant differentiation, red color means *P*-value $\leq$ 0.05 and color gray means *P*-value > 0.05); (d) Heatmap of mtDNA haplogroup frequency correlations between the ancient Shandong populations and present-day East Asian populations, where values toward 1 indicate a positive correlation (blue), meaning that the two populations share a similar maternal genetic structure, otherwise indicating a negative correlation (white).

more derived haplogroup of B5b that is mainly found in Han, Hezhen, Minnan, and Makatao (63.64%), but is also found in other East Asians, including Japanese, Kinh, Buryat, Khamnigan, and Kizhi (Table S6 online). For haplogroup D, both D4 and D5 have high frequencies in ancient populations of northern East Asia (17.60%–43.75%), but are less frequent in ancient populations of southern East Asia (0–20%) [22,34]. Additionally, N9a was found in one individual aged ~8200 BP from Xiaogao. N9a belongs to the Eastern Eurasian lineage [10], and it is distributed predominantly in East Asia [38]. These results indicate that Shandong populations contained both northern and southern East Asian-associated haplogroups prior to 4600 BP (Fig. 3a).

3.3. Maternal genetic structure in Shandong after 4600 BP

For most of the Shandong sites, haplogroups B and D maintain their high frequencies (>20%) throughout our sampling period (9500–1800 BP) (Table S1 online), while haplogroups C (6.00%, C7a1 and C7b), M9 (6.00%, M9a1), and F (2.00%, F1a1, F2a, and F4a1) are only observed in samples younger than 4600 BP

(Fig. 3b). These newly introduced haplogroups indicate that the maternal genetic structure of Shandong became more diverse beginning with the onset of the Longshan cultural period (4600–4000 BP). In examining the possible sources of these newly introduced haplogroups (Table S3 online), we find haplogroup C is prevalent in northern ethnic groups in East Asia, e.g., the Oroqen population (29.55%, Table S3 online), which is consistent with previous observations that this haplogroup is more diverse among the northern ethnic groups [34]. Haplogroup M9a was suggested to have spread from Southeast Asia northward into the East Asian mainland about 15,000 BP [38], and has a limited distribution among the present-day northern East Asians (0–4.20%) [33]. Haplogroup F1 and F4 are more common in southern East Asia and Southeast Asia than in northern East Asia [34,35], while haplogroup F2 is more frequently distributed among the northern East Asians [10]. As these more recent haplogroups have been found in both southern and northern East Asians, their appearance in Shandong, without replacing those observed before 4600 BP, may represent an influx of populations from outside Shandong after this period.

We further investigated populations from each site between 4600 and 1800 BP using PCA of haplogroup frequencies (Fig. 1b). Present-day northern and southern East Asians can be distinguished along the PC1 axis, driven mainly by the different proportions of haplogroup D, B, and F. Present-day northern East Asians have a higher proportion of haplogroup D (e.g., Mongolian 39.58%, Han-Jilin 35.29%, and Han-Shandong1 36.00%), while haplogroup B and F are more frequent among southern East Asians (e.g., B: Han-Jiangxi 34.78%, Han-Guangdong2 30.43%, F: Va 31.82%, Lahu 33.33%, and Bai 25.00%) (Table S3 online). The individuals from the Chengziya site (CZY, 4400–3300 BP, during and after the Longshan culture) and Tonglin site (TL, 4600–4000 BP, belonging to the Longshan culture) are close to the center of PC1 (Fig. 1b) mainly due to these two sites having high proportions of haplogroups elevated in either southern or northern East Asians. Specifically, the frequencies of haplogroup B of the two sites are higher than 37.50%, and the frequencies of haplogroup D are 20.00% (CZY) and 37.50% (TL), respectively. Haplogroup F was also found in the Tonglin site with a frequency of 12.50%. The Xinzhi (XZ, 3050–1800 BP) and Houli (HL, 3050–2750 BP) are closer to northern East Asians on PC1 (Fig. 1b), which can be explained by their higher proportion of haplogroup D (50.00% in XZ and 33.30% in HL). The population from the Liujiazhuang site (LJZ, 3100–1800 BP) clusters with present-day northern East Asians (Fig. 1b), likely owing to the Liujiazhuang population sharing 59.40% of their haplogroups (A, B, D, G, M11, M8, and N9a) with the present-day Shandong population (Han-Shandong1). On the other hand, the Yixi (YX, 2300–1800 BP) site is closer to southern East Asians on PC1 with a haplogroup F frequency of 40.00% (Fig. 1b).

3.4. Relationship between ancient Shandong populations and present-day East Asians

We then investigated how the ancient Shandong populations are related to present-day East Asians. First, we calculated the Φst between the ancient Shandong populations and present-day East Asians (15 selected populations from Han and other ethnic groups) (Fig. 1c and Table S5 online). Beqian (5500–5300 BP) is found to be more genetically different from present-day East Asians (eight populations show significant Φst with Beiqian), compared with younger Shandong populations which are more similar to present-day East Asians (0–4 populations show significant Φst). We further looked into the haplotype composition (reflected by correlation of haplotype frequencies) between the ancient Shandong populations and present-day East Asians. We observe that ancient Shandong populations are less similar to several northern ethnic groups (Kazak, Uzbek, and Uygur) and southern ethnic groups (Bai and Jino) compared with other present-day East Asians (Fig. 1d).

3.5. Possible genetic exchange between coastal and inland Shandong

We also investigated the population dynamics between the coastal site (Beiqian, 5500–5300 BP) and inland regions across time. In the coastal region, the site contained haplogroups M8 (10.50%) and A (5.30%) (Fig. 3b). However, in the inland sites, we observed these two haplogroups (M8 10% and A 16%) only in sites younger than 3100 BP (Fig. 3b). Furthermore, we found that populations from the coastal site and older inland sites (4600–3100 BP) had a larger value of genetic distance (Φst = 0.037, P = 0.06), compared with that between the coastal site and younger inland sites (3100 to 1800 BP) (Φst = 0.005, P = 0.30). We observed a similar pattern in the AMOVA analysis, where the coastal site and older inland sites had a larger value of variance (5.19%) than that between the coastal site and younger inland sites (1.39%). These results suggest that exchange between populations from the coastal and inland regions was less frequent before 3100 BP, which is consistent with archaeological findings that there were cultural differences between coastal and inland regions before and during the Longshan culture [39–41]. An alternative explanation could be due to the influx of people with ancestry from other parts of East Asia that introduced haplogroups M8 and A [6,13,33].

3.6. Discovery of the haplogroup B5b2 lineage in Shandong

A previous study speculated that populations carrying haplogroup B5b migrated from northwestern mainland East Asia into other East Asian areas, bypassing Shandong, as this haplogroup had not been previously observed in ancient Shandong based on ~3000-year-old individuals from northwestern mainland East Asia [10,34]. Here, we identified nine ancient individuals belonging to haplogroup B5b from six sites in Shandong, (9500 to 1800 BP): Bianbian (n = 1), Beiqian (n = 4), Houli (n = 1), Tonglin (n = 1), Xinzhi (n = 1), and Chengziya (n = 1). To investigate this haplogroup in more detail, we obtained 20 complete B5b mitochondrial sequences from present-day individuals (Table S6 online) representing present-day East Asians, including the Northern Han (n = 2), Southern Han (n = 3), Hezhen (n = 1), Minnan (n = 1), Makatao (n = 1), Hakka (n = 1), Kyrgyz (n = 2), and Tingri (n = 1), Buryat (n = 1), Khamnigan (n = 1), Kizhi (n = 1), Japanese (n = 1), Suay (n = 1), Shan (n = 1), and Kinh (n = 2). Mitogenome coalescence times were estimated from the complete data using BEAST [30]. Table S7 (online) shows the B5b lineage after calibrating with the radiocarbon dates of our samples. Under the B5b2 lineage, the B5b2a2 sublineage from the 5500 to 5300-year-old Beiqian (BQ-M2*) shares a common ancestor with the B5b2a2 in present-day Hezhen (HGDP01238), Buryat (JN857016), and Khamnigan (JN857039) (Fig. 2a–c). The estimated Time to the Most Recent Common Ancestor (TMRCA) of B5b2a2 is 8056 BP (95% highest probability density (HPD), 5400–12,192, Table S7 online) with a posterior value of 1.00. Furthermore, two Beiqian individuals (BQ-M31-A and BQ-M24*) are likely to share a common B5b2a ancestor with the Japanese (HGDP00767) around 13,040 BP (95% HPD, 7595–20,329, Table S7 online) with a posterior value of 0.99. Beiqian (BQ-M139-D) and Xinzhi (XZ-M59) individuals are likely to share a common B5b2b ancestor with Han individuals from the Southern Han (HG00690) and Northern Han (NA18643) around 11,513 BP (95% HPD, 6368–18102, Table S7 online) with a posterior value of 0.98. Accordingly, all of them share a common B5b2 ancestor related to the ~9500-year-old Bianbian individual

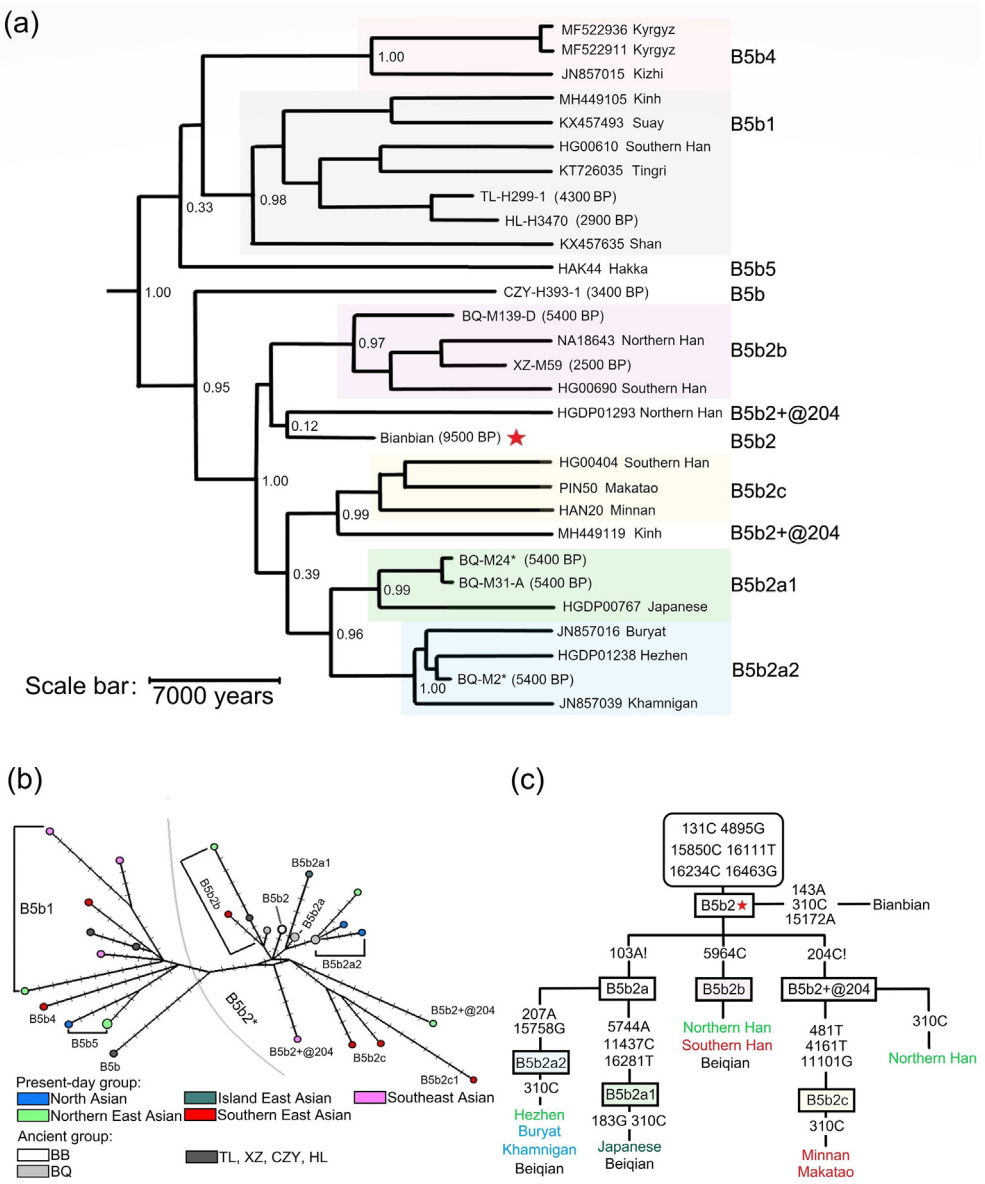


Fig. 2. Genetic analysis of populations in the B5b lineage. (a) Bayesian coalescence tree of the B5b lineage with the basal position of Bianbian indicated (red star) (BQ-M2-B, BQ-M2-E, BQ-M2-H, BQ-M2-L, and BQ-M2-J, these 5 individuals in one matrilineal kinship were regarded as BQ-M2*, BQ-M24-4 and BQ-M24-23 in one matrilineal kinship were regarded as BQ-M24*), all the individuals labeled with the sample ID and dates (without a population name) are from ancient Shandong sites, the whole tree can be seen in Fig. S3 (online); (b) Network of the B5b lineage (B5b2 lineage is in the right part of the grey line) using complete mitochondrial genomes of ancient Shandong and present-day Asians showing that the Bianbian individual and Beiqian population are closest to the root of haplogroup B5b2, and in particular, some Beiqian sequences are ancestral to present-day East Asians and North Asians; (c) Private mutations under the B5b2 lineage showing that Bianbian is three mutations (143A, 310C, and 15172A) away from the root of the haplogroup B5b2, and that Bianbian and Beiqian are closely related to their common ancestor with present-day East Asians and North Asians (@ = assumed back mutation or missing mutation).

around 17,293 years ago (95% HPD, 10,503–26,726 BP, Table S7 online). These results suggest that the B5b2 lineage found in the Bianbian individual (9500 BP) is likely ancestral to East Asians and North Asians (Fig. 2a–c and Fig. 3a). This coincides with the population dynamics in Shandong proposed in previous studies, with an outward spreading to the west and south of East Asia and interactions with other cultures (e.g., the Yangshao culture) [41–43].

4. Conclusion

In this study, we reconstructed the maternal genetic structure of Shandong populations over the past nine thousand years (Fig. 3). Since 9500 BP, the main haplogroups of present-day northern and southern East Asia could be observed in the ancient populations of Shandong. After 4600 BP, additional connections were identified between populations inside and outside of Shandong. After 3100 BP, a possible genetic exchange between the coastal and inland regions within Shandong was observed. Additionally, in the B5b2 lineage newly discovered in Shandong populations, we observed (1) the Bianbian individual likely belongs to a lineage ancestral to East Asians and North Asians; (2) a continuous maternal genetic structure from the ~9500-year-old Bianbian individual to the 5500–5300-year-old Beiqian population. A wider range of temporal and geographical Y-chromosome and nuclear genomic data could further build on these insights and provide a more comprehensive picture of the genetic history and population movements of the people of ancient Shandong.

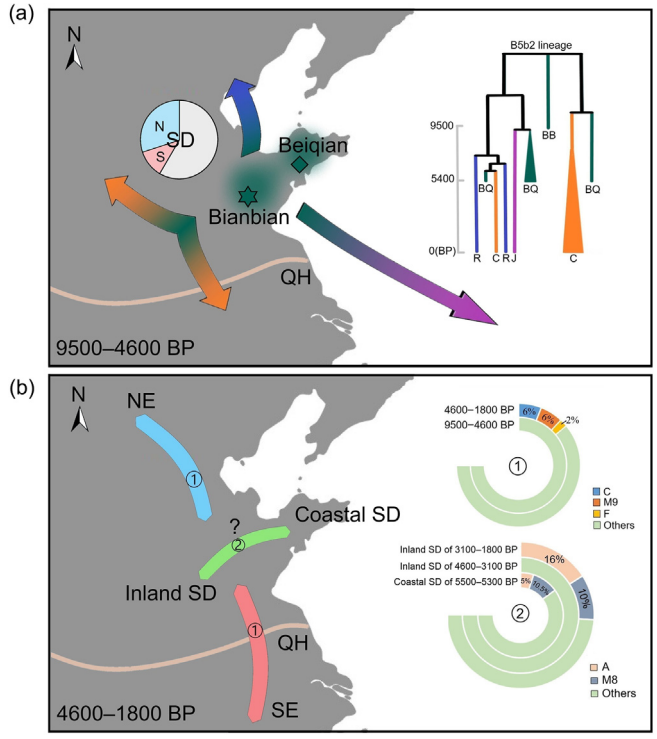


Fig. 3. The maternal genetic history of ancient Shandong population from 9500 to 1800 BP. (a) The pie chart shows that Shandong populations from 9500 to 4600 BP contained both northern and southern East Asian associated haplogroups (N = northern East Asian associated haplogroups; S = southern East Asian associated haplogroups); arrows with different colors indicate that the populations with B5b2 lineage in the Shandong region (the ~9500-year-old Bianbian individual and the 5500- to 5300-year-old Beiqian population) were ancestral to present-day East Asians and North Asians. (SD = Shandong populations in color green; C = Han, Hezhen, Minnan, and Makatao in color orange; R = Buryat and Khamnigan (located in North Asia) in color blue; J = Japanese in color purple; QH and the brown line = Qinling-Huaihe); (b) The maternal genetic sources in Shandong became more diverse since the Longshan cultural period (4600–4000 BP). New haplogroups (C, M9, and F) appeared since 4600 BP in the ancient Shandong population, indicating the appearance of northern (blue arrow) and southern (red arrow) maternal lineages in Shandong; The possible exchange with haplogroups M8 and A between the populations from the coastal and inland regions was less frequent before 3100 BP but frequent after 3100 BP (green arrow with the question mark) (NE = northern East Asian, SE = southern East Asian).

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (41672021, 41925009, 91731303, and 41630102), "Research on the Roots of Chinese Civilization" of Zhengzhou University (XKZDJC202006), the Chinese Academy of Sciences (XDB26000000, XDA1905010, and QYZDB-SSW-DQC003), the Tencent Foundation through the XPLORER PRIZE, the Howard Hughes Medical Institute (55008731), the National Social Science Foundation of China (15CKG013), the Shandong University Multidisciplinary Research and Innovation Team of Young Scholars (2020QNQT018). We would like to thank Albert Min-Shan Ko, Jacob Gardner, Mian Li, Tianyi Wang, and Wenjun Wang from the Institute of Vertebrate Paleontology and Paleoanthropology of Chinese Academy of Sciences, and Chao Ning from the Max Planck Institute for the Science of Human History for their comments, as well as archaeological teams from Shandong. Maps in this paper were reviewed by the Ministry of Natural Resources of the People's Republic of China (GS(2021)184).

Author contributions

Qiaomei Fu designed the research project. Qiaomei Fu, Juncen Liu, Wen Zeng, Bo Sun, and Xiaowei Mao managed the project. Wen Zeng, Bo Sun, Yongsheng Zhao, Fen Wang, Zhenguang Li, Fengshi Luan, Junfeng Guo, Chao Zhu, Zimeng Wang, Chengmin Wei, Ming Zhang, Weihong Hou, Wanjing Ping, and Xiaohong Wu collected archaeological samples and dating. Qiaomei Fu, Peng Cao, Ruowei Yang, Feng Liu, Xiaotian Feng, and Qinyan Dai, performed or supervised wet laboratory work. Qiaomei Fu and Xiaotian Feng did the data processing and quality control. Qiaomei Fu, Ming Zhang, Juncen Liu, and Xiaowei Mao supervised or performed the analyses. Juncen Liu, Wen Zeng, Xiaowei Mao, and Qiaomei Fu wrote the manuscript. E. Andrew Bennett, Yichen Liu, and all other authors discussed, critically revised, and approved the final version of the manuscript.

Appendix A. Supplementary materials

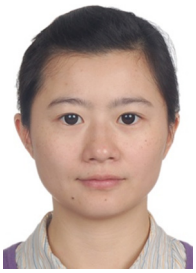
Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2021.01.029.

References

[1] Ko AS, Chen CY, Fu Q, et al. Early austronesians: into and out of Taiwan. Am J Hum Genet 2014;94:426–36.
[2] Lipson M, Cheronet O, Mallick S, Rohland N, Oxenham M, Pietrusewsky M, et al. Ancient genomes document multiple waves of migration in Southeast Asian prehistory. Science 2018;361:92–5.
[3] McColl H, Racimo F, Vinner L, et al. The prehistoric peopling of Southeast Asia. Science 2018;361:88–92.
[4] Ning C, Wang CC, Gao S, et al. Ancient genomes reveal Yamnaya-related ancestry and a potential source of Indo-European speakers in Iron Age Tianshan. Curr Biol 2019;29:2526–2532.e4.
[5] Siska V, Jones ER, Jeon S, et al. Genome-wide data from two early Neolithic East Asian individuals dating to 7700 years ago. Sci Adv 2017;3:e1601877.
[6] Yang MA, Fan X, Sun B, et al. Ancient DNA indicates human population shifts and admixture in northern and southern China. Science 2020;369:282–8.
[7] Yang MA, Fu Q. Insights into modern human prehistory using ancient genomes. Trends Genet 2018;34:184–96.
[8] Yang MA, Gao X, Theunert C, et al. 40,000-Year-Old individual from Asia provides insight into early population structure in Eurasia. Curr Biol 2017;27:3202–3208.e9.
[9] Zhang M, Fu Q. Human evolutionary history in eastern Eurasia using insights from ancient DNA. Curr Opin Genet Dev 2020;62:78–84.
[10] Zhang F. Study on mtDNA polymorphism of ancient Chinese population. Thesis. Fudan University 2005 (in Chinese).
[11] Dong Y, Li C, Luan F, et al. Low mitochondrial DNA diversity in an ancient population from China: Insight into social organization at the Fujia site. Hum Biol 2015;87:71–84.
[12] Wang Li, Oota H, Saitou N, et al. Genetic structure of a 2,500-year-old human population in China and its spatiotemporal changes. Mol Biol Evol 2000;17:1396–400.
[13] Yao Y, Kong Q, Man X, et al. Reconstructing the evolutionary history of China: a caveat about inferences drawn from ancient DNA. Mol Biol Evol 2003;20:214–9.
[14] Dabney J, Knapp M, Glocke I, et al. Complete mitochondrial genome sequence of a middle pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc Natl Acad Sci USA 2013;110:15758–63.
[15] Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res 2012;40:e3.
[16] Meyer M, Kircher M, Gansauge MT, et al. A high-coverage genome sequence from an archaic Denisovan individual. Science 2012;338:222–6.
[17] Rohland N, Harney E, Mallick S, et al. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos Trans R Soc Lond B Biol Sci 2015;370:20130624.
[18] Fu Q, Meyer M, Gao X, et al. DNA analysis of an early modern human from Tianyuan Cave, China. Proc Natl Acad Sci USA 2013;110:2223–7.
[19] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60.
[20] Fu Q, Mittnik A, Johnson PF, et al. A revised timescale for human evolution based on ancient mitochondrial genomes. Curr Biol 2013;23:553–9.
[21] Reich D, Patterson N, Campbell D, et al. Reconstructing native American population history. Nature 2012;488:370–4.

[22] Li J, Zeng W, Zhang Ye, et al. Ancient DNA reveals genetic connections between early Di-Qiang and Han Chinese. BMC Evol Biol 2017;17:239.
[23] van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 2009;30:E386–94.
[24] Weissensteiner H, Pacher D, Kloss-Brandstätter A, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res 2016;44:W58–63.
[25] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32:1792–7.
[26] Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010;10:564–7.
[27] Jia J, He X, Jin Y. Statistics. 6th edition. Beijing: China Renmin University Press; 2015 (in Chinese).
[28] Bandelt HJ, Forster P, Rohl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 1999;16:37–48.
[29] Leigh JW, Bryant D. PopART: Full-feature software for haplotype network construction. Methods Ecol Evol 2015; 6: 1110-6
[30] Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 2007;7:214.
[31] Darriba D, Taboada GL, Doallo R, et al. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 2012;9:772.
[32] Kivisild T, Tolk HV, Parik J, et al. The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol 2002;19:1737–51.
[33] Kong QP, Yao YG, Liu M, et al. Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. Hum Genet 2003;113:391–405.
[34] Tanaka M, Cabrera VM, González AM, et al. Mitochondrial genome variation in eastern Asia and the peopling of Japan. Genome Res 2004;14:1832–50.
[35] Yao YG, Kong QP, Bandelt HJ, et al. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 2002;70:635–51.
[36] Yao Y, Kong Q, Wang C, et al. Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in China. Mol Biol Evol 2004;21:2265–80.
[37] Yao YG, Nie L, Harpending H, et al. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. Am J Phys Anthropol 2002;118:63–76.
[38] Derenko M, Malyarchuk B, Denisova G, et al. Complete mitochondrial DNA analysis of eastern Eurasian haplogroups rarely found in populations of northern Asia and eastern Europe. PLoS One 2012;7:e32179.
[39] Gao G, Shao W. One of the birthplaces of Chinese civilization-Haidai historical and cultural district. Prehistory 1984;1:7–26 (in Chinese).
[40] Guo N. Research on the historical evolution of Shandong culture and the division of Shandong culture. Thesis. Anhui Normal University 2006 (in Chinese).
[41] He D. Study on environmental archaeology of Neolithic age in Shandong. Cultural Relics of the East 2004;2:26–40 (in Chinese).
[42] He D. A new study of the origin of civilization between Haidai area and central plains. Cultural Relics of Central China 2007;6:22-7+38 (in Chinese).
[43] Zhao Y, Xiao Y, Zheng W. Talking about the westward migration of Dawenkou cultural residents from human bone materials. Southeast Culture 2019;5:56–65 (in Chinese).

Juncen Liu is studying in the ancient DNA laboratory at the Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences as a graduate student, and he graduated from Shandong University with a major in Life Sciences in 2018. His research interest focuses on the population genetics of East Asia, specifically in the Shandong area.



Wen Zeng is the director of the Molecular Archaeology Laboratory at the Joint International Research Laboratory of Environmental and Social Archaeology, Shandong University. Her research involves fields like archaeology, physical anthropology, and ecology, seeking to understand the population histories through the study of ancient and contemporary genomics.

Bo Sun is a professor at the Shandong Provincial Institute of Cultural Relics and Archaeology. His research area is Neolithic Archaeology of China. He successively led the archaeological investigation and excavation of Chengziya, Dawenkou, Bianbian Cave, and other sites.

Xiaowei Mao is an associate professor at the Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences. His research focuses on exploring the ancient human demography, animal domestication, and evolution of Mendelian and complex traits for both human and animals.

Qiaomei Fu is a professor and the director of the Ancient DNA Laboratory at the Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences. Her research area is evolutionary genetics and population genetics.