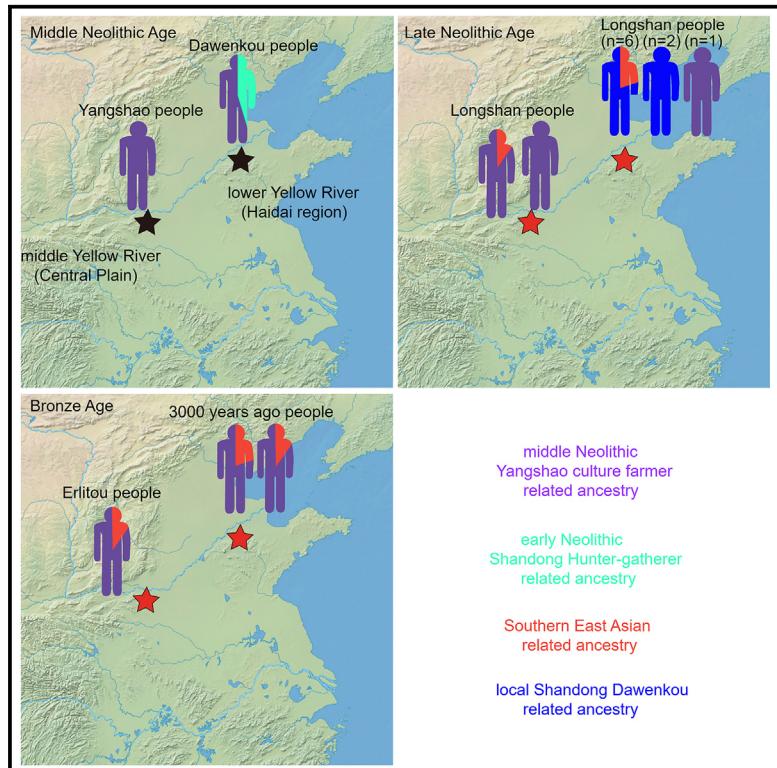


Dynamic history of the Central Plain and Haidai region inferred from Late Neolithic to Iron Age ancient human genomes

Graphical abstract



Authors

Hui Fang, Fawei Liang, Hao Ma, ...,
Yongsheng Zhao, Haiwang Liu,
Chuan-Chao Wang

Correspondence

fangh@sdu.edu.cn (H.F.),
17786126601@163.com (R.W.),
zhaoyongsheng@sdu.edu.cn (Y.Z.),
hnkglh@163.com (H.L.),
chuanchaowang@fudan.edu.cn (C.-C.W.)

In brief

Fang et al. report 31 ancient genomes and describe the dynamic demographic history of the Central Plain and northern Haidai region in the Late Neolithic. The Central Plain and northern Haidai region displayed genetic differentiation in the Late Neolithic but became genetically homogeneous during the Bronze Age.

Highlights

- The genetic differentiation between Late Neolithic Central Plain and northern Haidai
- Late Neolithic northern Haidai and Central Plain experienced rice-farmer-related gene flow
- Previously undescribed population structure in the Late Neolithic Central Plain
- Bronze Age Central Plain were genetically homogeneous with contemporaneous Haidai people



Report

Dynamic history of the Central Plain and Haidai region inferred from Late Neolithic to Iron Age ancient human genomes

Hui Fang,^{1,8,*} Fawei Liang,^{2,8} Hao Ma,^{3,4,8} Rui Wang,^{3,4,8,*} Haifeng He,^{3,4,8} Limin Qiu,³ Le Tao,³ Kongyang Zhu,³ Weihua Wu,² Long Ma,² Huazhen Zhang,² Shuqing Chen,^{1,5} Chao Zhu,⁶ Haodong Chen,³ Yu Xu,³ Yongsheng Zhao,^{1,5,*} Haiwang Liu,^{2,*} and Chuan-Chao Wang^{7,9,*}

¹Institute of Cultural Heritage, Shandong University, Qingdao 266237, China

²Henan Provincial Institute of Cultural Heritage and Archaeology, Zhengzhou 450000, China

³State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Xiamen 361102, China

⁴Department of Anthropology and Ethnology, Institute of Anthropology, Fujian Provincial Key Laboratory of Philosophy and Social Sciences in Bioanthropology, School of Sociology and Anthropology, Xiamen University, Xiamen 361005, China

⁵Ministry of Education Key Laboratory of Archaeological Sciences and Technology, Shandong University, Qingdao 266237, China

⁶Shandong Provincial Institute of Cultural Relics and Archaeology, Jinan 250012, China

⁷Ministry of Education Key Laboratory of Contemporary Anthropology, Department of Anthropology and Human Genetics, School of Life Sciences, Fudan University, Shanghai 200438, China

⁸These authors contributed equally

⁹Lead contact

*Correspondence: fangh@sdu.edu.cn (H.F.), 17786126601@163.com (R.W.), zhaoyongsheng@sdu.edu.cn (Y.Z.), hnglhw@163.com (H.L.), chuanchaowang@fudan.edu.cn (C.-C.W.)

<https://doi.org/10.1016/j.celrep.2025.115262>

SUMMARY

The peopling history of the Yellow River basin (YR) remains largely unexplored due to the limited number of ancient genomes. Our study sheds light on the dynamic demographic history of the YR by co-analyzing previously published genomes and 31 newly generated Late Neolithic to Iron Age genomes from Shandong in the lower YR and the Central Plain in the middle YR. Our analysis reveals the population structure in Shandong and the Central Plain in the Late Neolithic Longshan cultural period. We provide a genetic parallel to the observation of a significant increase in rice farming in the middle and lower YR in the Longshan period. However, the rice-farmer-related gene flow in the Longshan period did not arrive in groups from the Yuzhuang sites in the Central Plain or previously published groups in Shandong. The Bronze Age Erlitou culture genomes validate the genetic stability in the Central Plain and the relative genetic homogeneity between the Central Plain and Shandong.

INTRODUCTION

The Yellow River basin (YR) is the birthplace of ancient Chinese civilization. There is universal agreement among archaeologists that there are two different cultural groups in the Central Plain, in the middle YR and in the Haidai region in the lower reaches of the Yellow River, corresponding to Huaxia (华夏) and Dongyi (东夷), respectively.¹ Geographically, the Central Plain and the Haidai region adjoin each other, lacking significant natural barriers. Archaeologically, the early civilizations in the Central Plain and those in the Haidai region evolved through independent development and interactive integration.^{1,2} The prehistoric cultures in the Haidai region (denoted as Haidai-centered culture hereafter) included the Houli culture (~8.3–7.4 thousand years ago [kya]), Beixin culture (~7.4–6.2 kya), Dawenkou culture (~6.2–4.6 kya), Longshan culture (~4.6–4 kya), and Yueshi culture (~3.9–3.5 kya). The prehistoric cultures in the Central Plain (denoted as Central Plain-centered culture hereafter) included

the Peiligang culture (~9–7 kya), Yangshao culture (~7–5 kya), Longshan culture (~5–4 kya), Erlitou culture (~3.9–3.5 kya), and Shang culture (~3.6–3 kya). The people of the Early Neolithic Houli culture in the Haidai region relied primarily on hunting and fishing.³ The lifestyle of Early Neolithic Beixin culture in the Haidai region and Early Neolithic Peiligang culture in the Central Plain had been transformed from hunting and gathering to millet-based farming.⁴ Crop cultivation was the dominant subsistence strategy in Middle Neolithic Yangshao culture in the Central Plain and Middle Neolithic Dawenkou culture. Yangshao people relied primarily on millet farming, followed by rice as a minor component.⁵ The agricultural pattern of millet and rice mixed farming was formed in the Dawenkou culture, but the levels of reliance on millet and rice were geographically structured.⁶ A sedentary and agricultural way of life led to the massive population growth of the Yangshao people as well as the rapid expansion of populations with painted pottery (a significant feature of the Yangshao culture) to the Haidai region, West Liao River, and southern



China.^{7–11} After the Yangshao and Dawenkou cultural periods, the Late Neolithic Longshan culture occupied the Central Plain and Haidai region. Both unique features that derived from earlier local cultures and common features were observed in Henan Longshan in the Central Plain and Shandong Longshan in the Haidai region.¹² The stable development of the subsistence economy in the Longshan cultural period provided the material basis for intensified social stratification.¹³ There was a significant increase in rice farming in the Central Plain and the Haidai region in the Longshan cultural period.^{10,14} The early Bronze Age Erlitou culture in the Central Plain and the Yueshi culture in the Haidai region were preceded by the Longshan culture. Erlitou was the first state-level society in China.¹⁵ Following the Erlitou cultural period, the Shang Dynasty was the earliest ruling dynasty with records in Chinese history.

The long-term constant interaction between the Central Plain and the Haidai region started during the Peiligang cultural period and flourished from the Yangshao to the Shang period. Interestingly, the dominant role in cultural exchange between the Central Plain and the Haidai region was occasionally shifted.^{2,16} For example, the cultural influence of Yangshao culture in the Central Plain on the Beixin culture in the Haidai region was strong in the early and middle stages of the Yangshao cultural period, while during the Late Yangshao to Longshan culture era, Dawenkou culture in Haidai occupied a dominant role instead. Under the strong eastward expansion of the prosperous Central Plain-centered cultures, the Haidai-centered cultures disappeared and were completely integrated into Central Plain-centered cultures by the middle stage of the Spring and Autumn periods in the Iron Age. Despite this well-developed archaeological narrative, the population history in the Central Plains and Haidai region, especially the extent to which population movements accompanied extensive cultural contacts, remains unclear.

The emergent field of ancient DNA has introduced new means to unravel the population divergence and interaction in East Asia. In northern East Asia, to our knowledge, there are at least four biogeographically structured populations: (1) “Ancient Northeast Asian” (ANA), which was widespread across northeast Asia^{10,17,18}; (2) Yangshao-culture-related ancestry from the middle reaches of the YR that contributed to a vast number of present-day East Asian populations¹⁰; (3) seven Houli-culture-related hunter-gatherers’ genomes from Shandong in the Haidai region, which represent an ancestry referred to as “Shandong_HG”¹¹; and (4) Jomon-culture-related hunter-gatherers from the Japanese archipelago.^{19,20} Focused on the demographic history of the Central Plain, Ning et al.¹⁰ have shown a predominant role for Yangshao-culture-related ancestry (represented by YR_MN, including individuals from the Wanggou and Xiaowu sites) in local Longshan-culture-related (represented by YR_LN, including individuals from Haojiatai, Pingliangtai, and Wadian sites) and Late Bronze and Iron Age (represented by YR_LBIA, including individuals from Luoheguxiang, Jiaozuonie-cun, and Haojiatai sites) people. YR_LN and YR_LBIA were genetically homogeneous and could be characterized by the mixture of Yangshao-culture-related and Southern East Asian-related ancestry. Recent ancient DNA studies supported that the Neolithization of Shandong was related to the demic diffusion of Yangshao culture, as all published Dawenkou people carried

Yangshao-culture-related ancestry.^{8,9} Shandong Dawenkou-culture-related people showed population structure and could be genetically classified into three groups^{8,9}: (1) the direct descendants of Yangshao-culture-related ancestry; (2) a mixture between Yangshao-culture-related ancestry and Shandong_HG-related ancestry, and (3) a mixture between Yangshao-culture-related ancestry and Southern East Asia-related ancestry. Du et al. suggested that Shandong Longshan-culture-related groups from coastal Shandong (represented by Wutai_LS, Sanlihe_LS, and Chengzi_LS) possessed the Shandong Dawenkou-culture-related ancestry.⁸ From the Bronze Age to the historical era, people in Shandong were dominant by Central Plain-related ancestry.

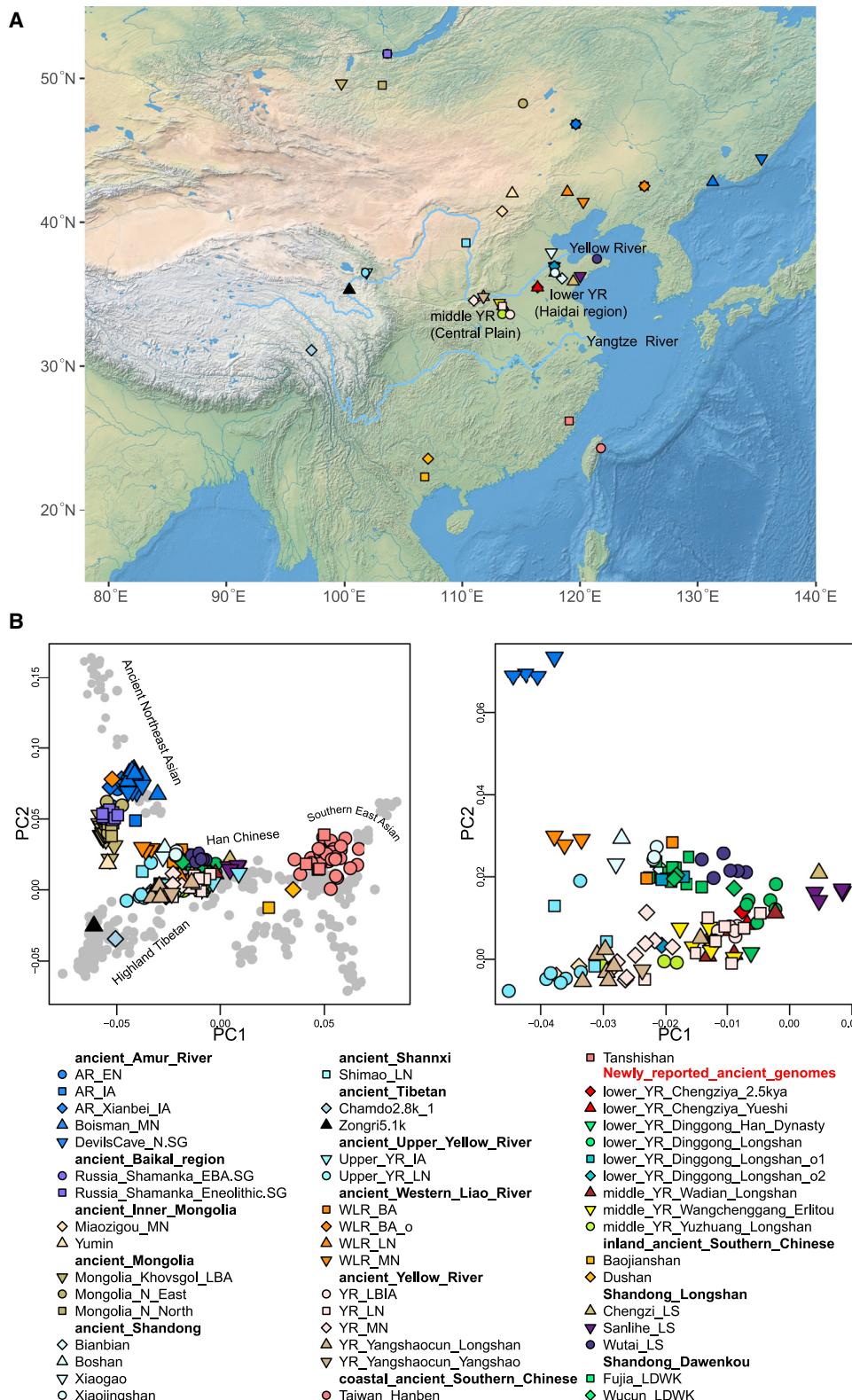
While these studies have revealed the demographic history of the Central Plain and Shandong, several important questions remain to be explored. First, no ancient genomes related to Late Neolithic Longshan culture in northern Shandong have been published. It is difficult to determine whether Longshan-culture-related people in northern Shandong possessed local Dawenkou-culture-related ancestry. Second, in the Central Plain, the extent to which the shared genetic ancestry between Yangshao- and Longshan-culture-related people and whether the genetic diversity in Longshan-culture-related people remain to be examined. Third, the early Bronze Age Erlitou culture in the Central Plain was thought to be the key point of the state formation phase. No ancient genomes related to Erlitou culture have been sampled genetically, making the genetic relationship between Erlitou culture and its preceding Longshan-culture-related people and later populations unclear.

Here, we present the 1240K-captured sequence data of 31 newly generated Late Neolithic to Iron Age genomes from Shandong in the lower reaches of the YR and the Central Plain in the middle reaches of the YR (Table S1). Thirteen individuals were sampled from Shandong, including ten Late Neolithic Longshan-culture-related genomes from the Dinggong sites and three Bronze Age and Iron Age genomes from the Chengziya and Dinggong sites. Eighteen individuals were sampled from the Central Plain, including eight Late Neolithic Longshan-culture-related genomes from the Wadian sites, three Late Neolithic Longshan-culture-related genomes from the Yuzhuang sites, and seven early Bronze Age Erlitou-culture-related genomes from the Wangchenggang sites. An overview of the archaeological sites of the studied samples is shown in Figure 1A. We co-analyzed the whole-genome data of 27 of 31 newly sequenced unrelated individuals alongside published data (Tables S2 and S3) to shed light on how people from the same region genetically connected and how people from the Central Plain and the Haidai region interacted with one another.

RESULTS

Newly generated ancient genome-wide data in this study

The 31 human skeletal remains studied here were recovered from five archaeological sites in the Central Plain and Haidai region (Table S1). We constructed double-strand libraries without



(legend on next page)

uracil-DNA glycosylase (UDG) treatment for each sample and enriched for a panel of around 1.24 million single-nucleotide polymorphisms (SNPs) via in-solution capture. All genomes displayed typical damage patterns of ancient DNA (Figure S1). All individuals displayed low levels of contamination (<5%) (Table S1). Pseudo-haploid genotypes were called on the targeted SNPs by randomly sampling a single allele at each position, resulting in individuals with 36,001 to 1,189,686 SNPs covered on a 1.24 million SNP panel (Table S1). The average autosomal coverage on targeted SNPs ranged from 0.03- to 4.8-fold (mean value 0.93-fold) (Table S1). We removed four individuals determined genetically to be up to second-degree relatives with other higher-coverage individuals in the downstream analyses (Table S4). Runs of homozygosity (ROH) analysis indicated that Longshan-culture individuals M77 and M10 from the Dinggong sites in the Haidai region and one Longshan-culture-related individual, H6, from the Wadian sites in the Central Plain carried large numbers of long ROHs (Figure S2). These results pointed to the recent inbreeding event in Haidai during the Longshan cultural period. Previously published Longshan-culture individuals from Central Plain showed the consanguineous mating pattern.²¹ Our new data also provided direct evidence of consanguineous mating among the Longshan-culture-related individuals in the Haidai region.

Population structure on the Central Plain and Haidai region

To individually describe the genetic profile of newly generated genomes, we first performed principal-component analysis (PCA) by projecting ancient genomes onto the first two dimensions of variation constructed by modern-day East Asians (Figure 1B and Figure S3). We observed a population substructure within Longshan-culture-related individuals from the Central Plain: (1) our newly generated four samples from Wadian (labeled as middle_YR_Wadian_Longshan) clustered with previously published samples from the Wadian, Pingliangtai, and Haojiatai sites¹⁰ (labeled as YR_LN in this study) that were slightly separated from Yangshao-culture-related clusters (labeled as YR_MN in this study) toward the direction of the Southern East Asian-related cluster. (2) Our newly generated three samples from Yuzhuang (labeled as middle_YR_Yuzhuang_Longshan) and recently published seven samples from the Yangshaocun sites in the Central Plain (labeled as YR_Yangshaocun_Longshan in this study) showed closer affinity to the YR_MN group than to YR_LN.

A population substructure was also observed in Longshan-culture-related individuals from the Haidai region. Our newly generated nine Longshan-culture-related genomes from the Dinggong sites were divided into three subgroups in the PCA plot: (1) the major group included six individuals (labeled as low-

er_YR_Dinggong_Longshan). These individuals were plotted closest to the previously published YR_LN-related genetic cluster that was well separated from Early Neolithic Shandong hunter-gatherers (denoted as Shandong_HG hereafter). (2) Two of nine individuals (labeled as lower_YR_Dinggong_Longshan_o1) were positioned outside of the Central Plain-related genetic variation regarding their genetic affinity to Shandong HG. (3) One individual (labeled as lower_YR_Dinggong_Longshan_o2) fell within the genetic variation of Neolithic Central Plain but did not cluster with individuals from the major cluster of lower_YR_Dinggong_Longshan. Our newly reported Longshan-culture-related individuals from Shandong did not overlap with recently published Longshan-culture-related individuals from Shandong (i.e., Chengzi_LS, Sanlihe_LS and Wutai_LS) either in PCA or in geographic locations. From a genetic perspective, Chengzi_LS and Sanlihe_LS individuals showed a close relationship with lower_YR_Dinggong_Longshan and also showed additional affinity to the Southern East Asian-related cluster; Wutai_LS individuals showed a close relationship with lower_YR_Dinggong_Longshan_o1; however, lower_YR_Dinggong_Longshan_o1 showed more affinity with the Shandong_HG-related cluster than Wutai_LS. Recently published Fujia_LDWK and Wucun_LDWK people were geographically close to our newly studied Longshan-culture-related individuals from Dinggong in northern Shandong. In the PCA plot, Longshan-culture-related groups from Dinggong did not completely overlap with the genetic cluster of Fujia_LDWK/Wucun_LDWK. These results mirrored the potential genetic shift between the Dawenkou and the Longshan cultural periods and the genetic diversity in the Longshan cultural period of the Haidai region.

Focused on newly generated Early Bronze Age genomes, we observed that one Yueshi-culture-related individual from Haidai (labeled as lower_YR_Chengziya_Yueshi) and six Erlitou-culture-related genomes from the Central Plain (labeled as middle_YR_Wangchenggang_Erlitou) were projected onto YR_LN-related genetic cluster. Late Bronze and Iron Age genomes, including the published ancient genomes dating to the Shang, Western Zhou, and Western Han dynasties in the Central Plain (labeled as YR_LBIA in this study) and our newly generated one Han-dynasty-related individual from Haidai (labeled as lower_YR_Dinggong_Han_Dynasty) and one 2.5 kya individual from Haidai (labeled as lower_YR_Chengziya_2.5kya), also clustered together with YR_LN.

A dynamic demographic history in the Haidai region between the Middle Neolithic to the historical period

To explore whether new arrivals of the Central Plain-related gene flow into Haidai Longshan-culture-related people compared to Dawenkou-culture-related people, we performed f_4 statistics

Figure 1. Geographic locations and genetic profiles of ancient Central Plain and Shandong individuals

(A) The geographic locations of newly sampled and previously published representative populations in East Asia. Symbols as shown in (B). “EN” refers to the Early Neolithic Age, “MN” refers to the Middle Neolithic Age, “LN” refers to the Late Neolithic Age, “LB” refers to the Late Bronze Age, and “IA” refers to the Iron Age. “LS” refers to Longshan culture and “DWK” refers to Dawenkou culture. A detailed archaeological description of our newly generated samples is given in Table S1. This map was made with Natural Earth. Free vector and raster map data are at naturalearthdata.com.

(B) Principal-component analysis (PCA). The ancient individuals are projected on the PCs constructed by modern East Asian individuals using the “Isqproject: YES” option. The modern individuals are shown in light gray circles. See also Figure S3 for further details.

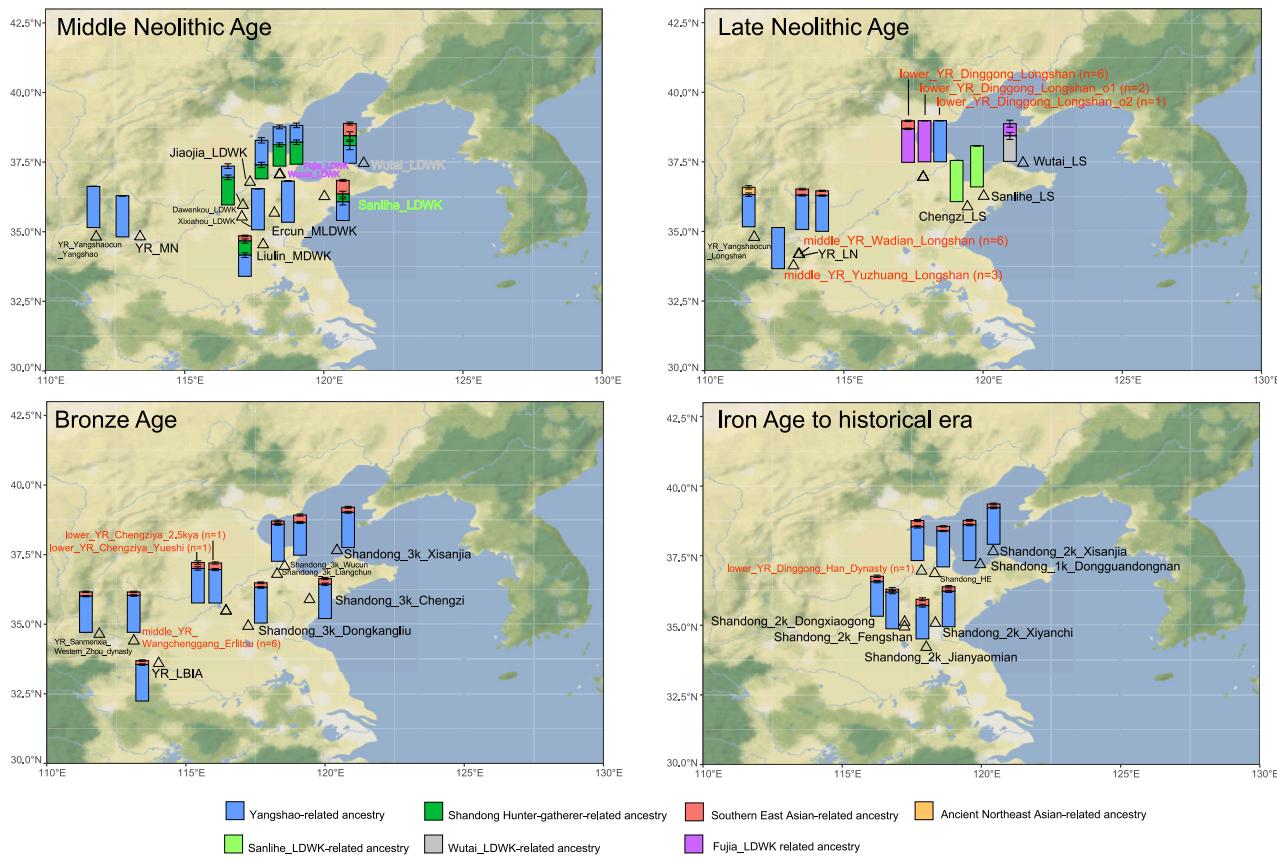


Figure 2. Ancestry modeling for Neolithic to historical era populations from the Central Plain and Haidai region

We used YR_MN (Yangshao culture-related ancestry) to represent Central Plain-related ancestry (blue component). We used Boshan to represent Shandong hunter-gatherer-related ancestry (dark green component). We used Ami from coastal Southern East Asia as the representative of Southern East Asian-related ancestry (red component). We used Devils' Cave from Russia Far East to represent Ancient Northeast Asian (ANA)-related ancestry (orange component). We used Fujia_LDWK (purple component) as the Dawenkou cultural period ancestry for Longshan-culture-related Dinggong groups (lower_YR_Dinggong_Longshan, lower_YR_Dinggong_Longshan_01, and lower_YR_Dinggong_Longshan_02). Sanlihe_LDWK (light green component) was the Dawenkou cultural period group in Sanlihe. We used Sanlihe_LDWK (light green component) as the source for the Longshan-culture-related Sanlihe group (Sanlihe_LS) and neighboring Chengzi group (Chengzi_LS). Wutai_LDWK (gray component) was the Dawenkou cultural period group in Wutai. We used Wutai_LDWK as the source for the Longshan-culture-related Wutai group (Wutai_LS). Horizontal bars represent ± 1 standard error. See also Tables S7 and S8.

(Table S5) in the form of f_4 (Yoruba, YR_MN; Wutai_LS, Wutai_LDWK), f_4 (Yoruba, YR_MN; Sanlihe_LS, Sanlihe_LDWK), and f_4 (Yoruba, YR_MN; lower_YR_Dinggong_Longshan, Fujia_LDWK/Wucun_LDWK) (Fujia_LDWK and Wucun_LDWK were geographically close to our newly studied Dinggong sites; we here used Fujia_LDWK and Wucun_LDWK to represent the Dawenkou-culture-related ancestry for Dinggong). The non-significant f_4 values ($-0.901 < Z \text{ score} < 0.696$) suggested that Longshan-culture-related groups did not receive additional Central Plain-related gene flow compared with local Dawenkou-culture-related groups. Next, we quantitatively explored the shared genetic shift between the Dawenkou and the Longshan cultural periods in the region of the Dinggong sites, one of the central settlements during the Longshan cultural periods in the Haidai region.²² In the outgroup f_3 statistics (Table S6 and Figure S4), lower_YR_Dinggong_Longshan shared a high genetic drift with Fujia_LDWK and Wucun_LDWK. Compared with Fujia_LDWK and Wucun_LDWK, lower_YR_Dinggong_Longshan showed genetic

affinity with Southern East Asian-related ancestry, as shown by f_4 (Yoruba, Taiwan_Hanben; lower_YR_Dinggong_Longshan, Fujia_LDWK/Wucun_LDWK) (Z score = -3.529 and -5.155) (Table S5). $qpAdm$ modeling suggested that lower_YR_Dinggong_Longshan could be modeled as the mixture of $\sim 80\%$ Fujia_LDWK/Wucun_LDWK and $\sim 20\%$ Southern East Asian (represented by Amis) (Table S7 and Figure 2). lower_YR_Dinggong_Longshan could also be modeled as the mixture among Shandong_HG- (represented by Boshan, 30.5%), Central Plain- (represented by YR_MN, 51.2%), and Southern East Asian-related ancestry (represented by Amis, 18.3%) (Table S7 and Figure S5). The non-significant values in all f_4 (Yoruba, X; lower_YR_Dinggong_Longshan_01, Fujia_LDWK/Wutai_LDWK) suggested that lower_YR_Dinggong_Longshan_01 (including two individuals) were the direct descendants of Fujia_LDWK/Wutai_LDWK (Table S5 and Figure 2). This result was supported by $qpAdm$ analysis (Table S7). Another individual, lower_YR_Dinggong_Longshan_02, was the direct descendant of Yangshao-culture-related

ancestry with no need for Southern East Asian- and Fujia_LDWK/Wucun_LDWK-related genetic contribution, as shown by the non-significant values in all f_4 statistics (Yoruba, X; lower_YR_Dinggong_Longshan_o2, YR_MN) and $qpAdm$ analysis (Tables S5 and S7 and Figure 2). lower_YR_Dinggong_Longshan_o2 was also genetically homogeneous with Xixiahou_LDWK and Ercun_MLDWK from Shandong, which were genetically homogeneous with Yangshao-culture-related ancestry (Table S7).

Previously published Shandong_3k_Wucun and Shandong_3k_Liangchun could be used to represent the 3 kya people in the Dinggong sites as their close geographic locations. Previously published Shandong_HE individuals from Shen et al.²³ and our newly generated one Han dynasty individual from Dinggong (lower_YR_Dinggong_Han_dynasty) could be used to represent the historical era people in the Dinggong sites as their close geographic locations. The outgroup f_3 statistics (Table S6 and Figure S4) supported that Shandong_3k_Wucun shared high genetic drift with Central Plain-related ancestry. The significant negative values in f_4 (Yoruba, YR_MN; Shandong_3k_Wucun, Wucun_LDWK/lower_YR_Dinggong_Longshan) (Z score = -3.205 and -3.350) (Table S5) also suggested Central Plain-related gene flow into the region covered Dinggong/Wucun sites between the Longshan cultural period (~4 kya) and 3 kya. The significant negative value in f_4 (Yoruba, Taiwan_Hanben; Shandong_3k_Wucun, Wucun_LDWK/YR_MN/YR_LN) (-4.053 < Z score < -3.240) (Table S5) suggested that Shandong_3k_Wucun shared extra affinity with Southern East Asian-related ancestry compared to Wucun_LDWK/YR_MN/YR_LN, while this Southern East Asian-related gene flow was not observed compared to lower_YR_Dinggong_Longshan, i.e., f_4 (Yoruba, Taiwan_Hanben; Shandong_3k_Wucun, lower_YR_Dinggong_Longshan) ~0 (Z score = -1.129) (Table S5). The non-significant values in f_4 (Yoruba, Shandong_HG; Shandong_3k_Wucun, YR_MN/YR_LN) (Table S5) suggested no additional Shandong_HG-related signal in Shandong_3k_Wucun compared to Central Plain-related ancestry (Z score = -1.056 and -0.447). The marginally significant negative value in f_4 (Yoruba, lower_YR_Dinggong_Longshan/Wucun_LDWK; Shandong_3k_Wucun, YR_MN/YR_LN) (-2.931 < Z score < -2.143) (Table S5) suggested that Shandong_3k_Wucun might have extra affinity with local Dawenkou- and Longshan-related ancestry compared with YR_MN/YR_LN. We therefore used YR_MN/YR_LN, Southern East Asian, lower_YR_Dinggong_Longshan/Wucun_LDWK as the potential source for Shandong_3k_Wucun and conducted one-way, two-way, and three-way $qpAdm$ modeling (Table S7). We found that Shandong_3k_Wucun could be modeled only as the mixture of ~92% YR_LN and ~8% Southern East Asian (Figure 2). This two-way modeling fitted well even when local Dawenkou and Longshan groups (i.e., Wucun_LDWK and lower_YR_Dinggong_Longshan) were added to the outgroup set (Table S7), while Shandong_3k_Liangchun and our newly reported Bronze Age and Iron Age individuals (lower_YR_Chengziya_Yueshi and lower_YR_Chengziya_2.5kya), as well as Shandong_HE and lower_YR_Dinggong_Han_Dynasty, were the direct descendants of YR_LN, as suggested by non-significant values in all f_4 and $qpAdm$ analysis (Tables S5 and S7 and Figure 2).

The cultural transition between Yangshao and Longshan was not necessarily accompanied by the population turnover in the Central Plain

With the widespread genetic influence of Central Plain-related ancestry on the Haidai region, we next examined the extent to which Haidai-specific Shandong_HG-related lineage impacted the Central Plain. No gene flow related to Shandong_HG was detected in the Longshan cultural period compared to Yangshao-culture-related populations, as Shandong_HG shared equal amounts of alleles with YR_MN and all Longshan-culture-related groups in the Central Plain, i.e., f_4 (Yoruba, Shandong_HG; YR_LN/middle_YR_Wadian_Longshan/middle_YR_Yuzhuang_Longshan, YR_MN) ~0 (-0.244 < Z score < 0.874) (Table S5). Given that some Dawenkou-culture-related groups in Shandong, such as Ercun_LDWK and Xixiahou_LDWK, were already the direct descendants of YR_MN, we could not exclude the scenario of population migration of the Dawenkou groups with no trace of Haidai-specific Shandong_HG-related ancestry to the Central Plain accompanied by the cultural interactions.

The newly generated Longshan culture population from the Wadian archaeological site was genetically homogeneous with previously published YR_LN genomes on the basis that all f_4 statistics of the form f_4 (Yoruba, X; middle_YR_Wadian_Longshan, YR_LN) ~0 (i.e., all $|Z$ scores| < 3) (Table S5). The genetic homogeneity between middle_YR_Wadian_Longshan and YR_LN can be further supported by the $qpAdm$ analysis (Table S8 and Figure 2). Another newly reported Longshan-culture-related group in the Central Plain was sampled from the Yuzhuang site (i.e., middle_YR_Yuzhuang_Longshan). Yuzhuang sites are currently the largest Longshan-era settlement in the Central Plain region. Yuzhuang had cultural exchanges with surrounding areas. The red pottery cup, painted pottery, and elaborate funerals in Yuzhuang are believed to have been influenced by the Qujialing and Shijiahe cultures in the Jianghan region. The unearthed eggshell pottery and the use of roe deer teeth as burial objects are similar to factors of the Dawenkou and Longshan cultures in Shandong. It was unclear whether gene flows were accompanied by the cultural influence from the Shandong and Jianghan regions into Yuzhuang. We found that Yuzhuang shared the most genetic drift with YR_MN-related ancestry in the outgroup f_3 statistics (Table S6 and Figure S4). f_4 statistics in the form of f_4 (Yoruba, X; middle_YR_Yuzhuang_Longshan, YR_MN) (Table S5) did not produce significant values, even when X = Southern East Asian and Shandong ancients. The $qpAdm$ analysis also supported the genetic homogeneity between YR_MN and middle_YR_Yuzhuang_Longshan (Table S8 and Figure 2). Therefore, from a genetic perspective, Longshan-culture-related Yuzhuang people did not receive gene flow from Shandong- and Southern East Asian-related ancestry compared with preceding Yangshao-culture-related ancestry.

The early Bronze Age Erlitou-culture-related people were the direct descendants of YR_LN-related ancestry

Here, we report the first batch of Erlitou-culture-related genomes from the Wangchenggang sites in the Central Plain (i.e., middle_YR_Wangchenggang_Erlitou). No significant negative f_4 values were produced in f_4 (Yoruba, X;

middle_YR_Wangchenggang_Erlitou, YR_MN/YR_LN), even when $X = \text{Shandong_HG/Shandong Dawenkou/Shandong Longshan-related groups}$ (Table S5). The *qpAdm* analysis (Table S8 and Figure 2) rejected the model in which middle_YR_Wangchenggang_Erlitou were unadmixed descendants of YR_MN. middle_YR_Wangchenggang_Erlitou people could be one-way modeled by YR_LN. This suggests the genetic stability of Late Neolithic YR_LN-related ancestry in the Central Plain, with no trace of Neolithic Shandong-related ancestry.

DISCUSSION

The relationship between the Central Plain and the Haidai-centered Chinese civilizations has been of great interest over the past decades. Previous ancient genome studies suggested that, in contrast to the oldest genomes from the Central Plain (represented by Middle Neolithic Yangshao-culture-related ancients) and the Haidai region (represented by Early Neolithic Houli-culture-related ancients), later populations from the Central Plain and Haidai showed less genetic differentiation from each other, which could be largely attributed to an extensive process of population migration from the Central Plain to the Haidai region, which led to all Middle Neolithic Late Dawenkou-culture-related groups carrying Central Plain-related ancestry.^{8,9} In this study, we offer new observations on the demographic history of population groups in the Central Plain and northern Shandong from the Late Neolithic to the Bronze Age.

After the Yangshao and Dawenkou cultural periods, Longshan culture occupied the Central Plain and Haidai region. Both unique features that derived from earlier local cultures and common features were observed in Henan Longshan in the Central Plain and Shandong Longshan in the Haidai region. From the genetics perspective, we observed the genetic legacy of Dawenkou-culture-related ancestry and Yangshao-culture-related ancestry in Longshan-culture-related people from Shandong and the Central Plain, respectively. We also observed a genetic shift between the Dawenkou/Yangshao and the Longshan cultural periods in both Shandong and the Central Plain. However, these genetic shifts were not related to the migration of the Central Plain and Shandong hunter-gatherer-related ancestry. In Shandong, previously published Longshan-culture-related groups (Wutai_LS, Sanlihe_LS, and Chengzi_LS) did not show extra affinity with Central Plain-related ancestry compared with the Dawenkou-culture-related group from the same archaeological site. Our newly reported Longshan-culture-related genomes from the Dinggong sites display a population substructure: two individuals (labeled as “lower_YR_Dinggong_Longshan_o1”) were the direct descendants of the local Dawenkou-culture-related genetic profile (represented by Fujia_LDWK and Wucun_LDWK) (Figure 2); six individuals (labeled as “lower_YR_Dinggong_Longshan”) maintained the local Dawenkou-culture-related genetic profile and received additional Southern East Asian-related ancestry that could be linked to the rice farmers from Southern East Asia (Figure 2). This result provided a genetic parallel to the observation of a significant increase in rice farming in the lower reaches of the YR in the Longshan period.¹⁴ Given that some Shandong groups were the direct descendants of YR_MN in the Dawenkou cultural period (repre-

sented by Xixiahou_LDWK), it is more likely that a single individual with 100% YR_MN-related ancestry in Dinggong (labeled as “lower_YR_Dinggong_Longshan_o2”) was a migrant from other regions of Shandong. In the Central Plain, neither previously published¹⁰ nor newly generated Longshan-culture-related ancients received the Shandong HG-related genetic impact compared to Yangshao-culture-related ancestry. Our newly generated Longshan-culture-related individuals from the Yuzhuang sites were genetically homogeneous with Yangshao-culture-related people (YR_MN). They were genetically heterogeneous with YR_LN and our newly generated individuals from the Wadian sites (Figure 2). This could be attributed to the additional Southern East Asian-related ancestry (~10%) in YR_LN and newly generated Wadian individuals compared with YR_MN and Longshan-culture-related Yuzhuang people. Six of nine Longshan-culture-related individuals from Dinggong also received additional Southern East Asian-related ancestry compared to Dawenkou-culture-related people. These results reflected that Southern East Asians separately migrated to Haidai and the Central Plain during the Longshan cultural period. Yuzhuang was the direct descendant of YR_MN. The case of Yuzhuang suggests that the cultural transition between Yangshao and Longshan was not necessarily accompanied by the population turnover in the Central Plain. Moreover, based on the archaeological information, the female individual M64XR from the Longshan period in the Yuzhuang site was a human sacrifice (the grave owner did not produce sufficient genomic data). However, there was no significant genetic difference between this individual and the other two Yuzhuang individuals nor was there a Yangshao-culture-related ancestry.

The genetic turnover in Shandong from the Longshan cultural period to the Bronze Age was linked to the strong expansion of Central Plain-related ancestry. The Bronze Age people in northern Shandong were represented by previously published Shandong_3k_Wucun and Shandong_3k_Liangchun. The genetic profile of these two groups was adequately explained by YR_LN and/or Southern East Asia, with no need for Shandong_HG-related ancestry as local Dawenkou- and Longshan-culture-related people (Figure 2). In the central part of the Central Plain, the early Bronze Age people represented by our newly generated Erlitou-culture-related people were genetically homogeneous with YR_LN, suggesting genetic stability in the central area of Central Plain (Figure 2). In the western Central Plain, there was a population turnover caused by the migration of YR_LN-related ancestry between the Longshan cultural period and the Bronze Age (Figure 2), as shown by recently published Longshan-culture-related genomes from the Yangshaocun sites (represented by YR_Yangshaocun_Longshan²⁴) and Bronze Age genomes from the Lusixi sites (represented by YR_Western_Zhou_Dynasty²⁵). YR_Yangshaocun_Longshan was genetically heterogeneous with YR_LN, while at least in the Western Zhou dynasty, people (represented by YR_Western_Zhou_Dynasty) were genetically homogeneous with YR_LN. There are also divergent views regarding the extent of the cultural influence of Yueshi-related culture on Erlitou culture in the Central Plain from the archaeological perspective.² Erlitou-culture-related people lacked Shandong HG/Shandong Dawenkou/Shandong Longshan-related ancestry compared

with YR_LN and YR_MN. These results suggest that the cultural diffusion of Haidai since the Middle Neolithic might not have accompanied massive population migrations with significant Shandong_HG-related ancestry into the Central Plain.

Limitations of the study

It should be noted that our limited sampling may not fully capture the genetic diversity of the Central Plain and Haidai regions. Further sampling from various archaeological sites from the Peiligang cultural period to the present, particularly in search of high-coverage ancient genomes, will be needed to obtain a more comprehensive understanding of the genetic interaction between Haidai and the Central Plain.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Chuan-Chao Wang (chuanchaowang@fudan.edu.cn).

Materials availability

This study did not generate new reagents.

Data and code availability

- The BAM files reported in this paper have been deposited in the Genome Sequence Archive in the National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (<https://hgdc.cncb.ac.cn/gsa-human>) and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#).
- This paper does not report the original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

The work was funded by the National Natural Science Foundation of China (T2425014 and 32270667), the Natural Science Foundation of Fujian Province of China (2023J06013), the Major Project of the National Social Science Foundation of China (21&ZD285), the National Key Research and Development Program (2023YFC3303701-02), the Open Research Fund of State Key Laboratory of Genetic Engineering at Fudan University (SKLGE-2310), the Open Research Fund of Forensic Genetics Key Laboratory of the Ministry of Public Security (2023FGKFKT07), and the Major Special Project of Philosophy and Social Sciences Research of the Ministry of Education (2022JZDZ023).

AUTHOR CONTRIBUTIONS

H.F. and C.-C.W. conceived and supervised the project. H.F., H.L., F.L., W.W., L.M., H.Z., S.C., C.Z., and Y.Z. provided the materials. H.M., H.H., L.Q., and L.T. performed the wet laboratory work. R.W. and K.Z. performed the genetic data analysis and prepared the figures. R.W. and C.-C.W. wrote and edited the manuscript. All authors contributed to the article and approved the final version for submission.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)

- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
 - Description of archaeological background to human remains
 - Method details
 - Quantification and statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2025.115262>.

Received: July 10, 2024

Revised: December 4, 2024

Accepted: January 13, 2025

Published: February 1, 2025

REFERENCES

1. Han, J. (2023). *The Making of the Chinese Civilization* (Springer Nature Singapore). <https://doi.org/10.1007/978-981-99-4213-8>.
2. Jin, S. (2005). Exchange and Integration of Archaeological Culture in the Heluo and Haidai Regions (In Chinese) (Zhengzhou University). PhD dissertation.
3. Jin, G. (2012). Preliminary Study on the Economic Development of the Houli Culture Industry (In Chinese). *Oriental Archaeology* 00, 579–594.
4. Lu, T. The Transition from Foraging to Farming and the Origin of Agriculture in China. The Australian National University; 1999. PhD dissertation.
5. Zhao, Z. (2011). New Archaeobotanic Data for the Study of the Origins of Agriculture in China. *Curr. Anthropol.* 52, S295–S306. <https://doi.org/10.1086/659308>.
6. Wu, R. (2018). Research on the Cultural and Industrial Economy of Dawenkou (In Chinese) (Shandong University). Master dissertation.
7. Luan, F. (1997). Archaeological Research in Haidai Area (In Chinese) (Shandong University Press).
8. Du, P., Zhu, K., Wang, M., Sun, Z., Tan, J., Sun, B., Sun, B., Wang, P., He, G., Xiong, J., et al. (2024). Genomic dynamics of the Lower Yellow River Valley since the Early Neolithic. *Curr. Biol.* 34, 3996–4006.e11. <https://doi.org/10.1016/j.cub.2024.07.063>.
9. Wang, F., Wang, R., Ma, H., Zeng, W., Zhao, Y., Wu, H., Tang, Z., He, H., Fang, H., and Wang, C.-C. (2024). Neolithization of Dawenkou culture in the lower Yellow River involved the demic diffusion from the Central Plain. *Sci. Bull.* 69, 3677–3681. <https://doi.org/10.1016/j.scib.2024.08.016>.
10. Ning, C., Li, T., Wang, K., Zhang, F., Li, T., Wu, X., Gao, S., Zhang, Q., Zhang, H., Hudson, M.J., et al. (2020). Ancient genomes from northern China suggest links between subsistence changes and human migration. *Nat. Commun.* 11, 2700. <https://doi.org/10.1038/s41467-020-16557-2>.
11. Yang, M.A., Fan, X., Sun, B., Chen, C., Lang, J., Ko, Y.-C., Tsang, C.H., Chiu, H., Wang, T., Bao, Q., et al. (2020). Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* 369, 282–288. <https://doi.org/10.1126/science.aba0909>.
12. Zhang, C. (2021). Longshanization, Longshan Period and Longshan Period - Rereading "Longshan Culture and Longshan Period. Southern Cultural Relics 1, 62–69.
13. Cheng, B. (2022). Research on the Complexity Process of Prehistoric Society in Zhengluo Region (In Chinese) (Shandong University). PhD dissertation.
14. Jin, G., and Luan, F. (2006). Progress and Issues in Research on Rice Agriculture during the Longshan Era in Haidai Prefecture (in Chinese). *Agricultural Archaeology* 1, 46–55.
15. Yuan, G. (2005). Research on Erlitou Culture (In Chinese) (Zhengzhou University). PhD dissertation.
16. Zhang, X. (2015). The Research of Dawenkou Culture (In Chinese) (Jilin University). PhD dissertation.

17. Mao, X., Zhang, H., Qiao, S., Liu, Y., Chang, F., Xie, P., Zhang, M., Wang, T., Li, M., Cao, P., et al. (2021). The deep population history of northern East Asia from the Late Pleistocene to the Holocene. *Cell* 184, 3256–3266.e13. <https://doi.org/10.1016/j.cell.2021.04.040>.
18. 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., et al. (2021). Genomic insights into the formation of human populations in East Asia. *Nature* 591, 413–419. <https://doi.org/10.1038/s41586-021-03336-2>.
19. Gakuhari, T., Nakagome, S., Rasmussen, S., Allentoft, M.E., Sato, T., Korneliussen, T., Chuinneagáin, B.N., Matsumae, H., Koganebuchi, K., Schmidt, R., et al. (2020). Ancient Jomon genome sequence analysis sheds light on migration patterns of early East Asian populations. *Commun. Biol.* 3, 437. <https://doi.org/10.1038/s42003-020-01162-2>.
20. McColl, H., Racimo, F., Vinner, L., Demeter, F., Gakuhari, T., Moreno-Mayar, J.V., van Driem, G., Gram Wilken, U., Seguin-Orlando, A., de la Fuente Castro, C., et al. (2018). The prehistoric peopling of Southeast Asia. *Science* 361, 88–92. <https://doi.org/10.1126/science.aat3628>.
21. Ning, C., Zhang, F., Cao, Y., Qin, L., Hudson, M.J., Gao, S., Ma, P., Li, W., Zhu, S., Li, C., et al. (2021). Ancient genome analyses shed light on kinship organization and mating practice of Late Neolithic society in China. *iScience* 24, 103352. <https://doi.org/10.1016/j.isci.2021.103352>.
22. Luan, F., Fang, H., and Xu, H. (1993). Summary of the fourth and fifth excavations at Dinggong Site in Zouping, Shandong Province (in Chinese). *Archaeology* 4, 295–299+385–386.
23. Shen, Q., Wu, Z., Zan, J., Yang, X., Guo, J., Ji, Z., Wang, B., Liu, Y., Mao, X., Wang, X., et al. (2024). Ancient genomes illuminate the demographic history of Shandong over the past two millennia. *Journal of Genetics and Genomics*. <https://doi.org/10.1016/j.jgg.2024.07.008>.
24. Li, S., Wang, R., Ma, H., Tu, Z., Qiu, L., Chen, H., Jiang, L., Geng, Y., Liu, H., Wang, J., et al. (2024). Ancient genomic time transect unravels the population dynamics of Neolithic middle Yellow River farmers. *Sci. Bull.* 69, 3365–3370. <https://doi.org/10.1016/j.scib.2024.09.002>.
25. Ma, H., Zhou, Y., Wang, R., Yan, F., Chen, H., Qiu, L., Zhao, J., Jin, L., and Wang, C.-C. (2024). Ancient genomes shed light on the long-term genetic stability in the Central Plain of China. *Sci. Bull.* 7, e24. <https://doi.org/10.1016/j.scib.2024.07.024>.
26. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* 9, 88. <https://doi.org/10.1186/s13104-016-1900-2>.
27. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
28. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
29. Jun, G., Wing, M.K., Abecasis, G.R., and Kang, H.M. (2015). An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* 25, 918–925. <https://doi.org/10.1101/gr.176552.114>.
30. Peitzner, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., and Nievelt, K. (2016). EAGER: efficient ancient genome reconstruction. *Genome Biol.* 17, 60. <https://doi.org/10.1186/s13059-016-0918-z>.
31. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684. <https://doi.org/10.1093/bioinformatics/btt193>.
32. Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* 16, 224. <https://doi.org/10.1186/s13059-015-0776-0>.
33. Fu, Q., Mitnik, A., Johnson, P.L.F., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., et al. (2013). A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes. *Curr. Biol.* 23, 553–559. <https://doi.org/10.1016/j.cub.2013.02.044>.
34. Katoh, K., Misawa, K., Kuma, K.I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. <https://doi.org/10.1093/nar/gkf436>.
35. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinf.* 15, 356. <https://doi.org/10.1186/s12859-014-0356-4>.
36. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönher, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* 44, W58–W63. <https://doi.org/10.1093/nar/gkw233>.
37. Ralf, A., Montiel González, D., Zhong, K., and Kayser, M. (2018). Yleaf: Software for Human Y-Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. *Mol. Biol. Evol.* 35, 1291–1294. <https://doi.org/10.1093/molbev/msy032>.
38. Patterson, N., Price, A.L., and Reich, D. (2006). Population Structure and Eigenanalysis. *PLoS Genet.* 2, e190. <https://doi.org/10.1371/journal.pgen.0020190>.
39. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient Admixture in Human History. *Genetics* 192, 1065–1093. <https://doi.org/10.1534/genetics.112.145037>.
40. Ringbauer, H., Novembre, J., and Steinrücken, M. (2021). Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* 12, 5425. <https://doi.org/10.1038/s41467-021-25289-w>.
41. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. <https://doi.org/10.1086/519795>.
42. Monroy Kuhn, J.M., Jakobsson, M., and Günther, T. (2018). Estimating genetic kin relationships in prehistoric populations. *PLoS One* 13, e0195491. <https://doi.org/10.1371/journal.pone.0195491>.
43. Zhu, C., Sun, B., Lv, K., and Zhang, Z. (2019). Excavation Report of Chengziya Site in Zhangqiu District, Jinan City from 2013 to 2015 (in Chinese). *Archaeology* 4, 3–24+2.
44. Jia, Z., Kuang, Y., and Jiang, T. (1983). Brief Report on Excavation of Wadian Site in Yuxian County (in Chinese). *Cultural Relics* 3, 37–48+100.
45. Fang, Y., and Liu, X. (2006). Excavation Report of Wangchenggang Site in Dengfeng City, Henan Province in 2002 and 2004 (in Chinese). *Archaeology* 9, 3–15+97.
46. Wu, W., and Zhang, W. (2022). Brief Report on M10 Excavation at Yuzhuang Site in Yexian County, Henan Province (in Chinese). *Huaxia Archaeology* 4, 3–6+37+2.
47. Van Klinken, G.J. (1999). Bone Collagen Quality Indicators for Palaeodietary and Radiocarbon Measurements. *J. Archaeol. Sci.* 26, 687–695. <https://doi.org/10.1006/jasc.1998.0385>.
48. Zhu, S., Ding, P., Wang, N., Shen, C., Jia, G., and Zhang, G. (2015). The compact AMS facility at Guangzhou Institute of Geochemistry, Chinese Academy of Sciences. *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. Atoms* 361, 72–75. <https://doi.org/10.1016/j.nimb.2015.06.040>.
49. Ramsey, C.B., and Lee, S. (2013). Recent and Planned Developments of the Program OxCal. *Radiocarbon* 55, 720–730. <https://doi.org/10.1017/S003382200057878>.
50. Reimer, P.J., Austin, W.E.N., Bard, E., Bayliss, A., Blackwell, P.G., Bronk Ramsey, C., Butzin, M., Cheng, H., Edwards, R.L., Friedrich, M., et al. (2020). The IntCal20 Northern Hemisphere Radiocarbon Age Calibration

- Curve (0–55 cal kBP). Radiocarbon 62, 725–757. <https://doi.org/10.1017/RDC.2020.41>.
51. Zhu, K., He, H., Tao, L., Ma, H., Yang, X., Wang, R., Guo, J., and Wang, C.-C. (2024). Protocol for a comprehensive pipeline to study ancient human genomes. STAR Protoc. 5, 102985. <https://doi.org/10.1016/j.xpro.2024.102985>.
52. Knapp, M., Clarke, A.C., Horsburgh, K.A., and Matisoo-Smith, E.A. (2012). Setting the stage – Building and working in an ancient DNA laboratory. Annals of Anatomy - Anatomischer Anzeiger 194, 3–6. <https://doi.org/10.1016/j.aanat.2011.03.008>.
53. Llamas, B., Valverde, G., Fehren-Schmitz, L., Weyrich, L.S., Cooper, A., and Haak, W. (2017). From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. Star: Science & Technology of Archaeological Research 3, 1–14. <https://doi.org/10.1080/20548923.2016.1258824>.
54. Rohland, N., Mallick, S., Mah, M., Maier, R., Patterson, N., and Reich, D. (2022). Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. Genome Res. 32, 2068–2078. <https://doi.org/10.1101/gr.276728.122>.
55. Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mitnik, A., et al. (2016). The genetic history of Ice Age Europe. Nature 534, 200–205. <https://doi.org/10.1038/nature17993>.
56. Lee, J., Miller, B.K., Bayarsaikhan, J., Johannesson, E., Ventresca Miller, A., Warinner, C., and Jeong, C. (2023). Genetic population structure of the Xiongnu Empire at imperial and local scales. Sci. Adv. 9, eadf3904. <https://doi.org/10.1126/sciadv.adf3904>.
57. Mallick, S., Micco, A., Mah, M., Ringbauer, H., Lazaridis, I., Olalde, I., Patterson, N., and Reich, D. (2024). The Allen Ancient DNA Resource (AADR) a curated compendium of ancient human genomes. Sci. Data 11, 182. <https://doi.org/10.1038/s41597-024-03031-7>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Ancient human remains	This paper	H6
Ancient human remains	This paper	M13_WD
Ancient human remains	This paper	M16
Ancient human remains	This paper	G11-1
Ancient human remains	This paper	M1
Ancient human remains	This paper	M32-1
Ancient human remains	This paper	M8_WD
Ancient human remains	This paper	M21
Ancient human remains	This paper	M63_WCG
Ancient human remains	This paper	M67-1
Ancient human remains	This paper	M10
Ancient human remains	This paper	M31
Ancient human remains	This paper	M5
Ancient human remains	This paper	M63_DG
Ancient human remains	This paper	M67
Ancient human remains	This paper	M77
Ancient human remains	This paper	M8_DG
Ancient human remains	This paper	M110
Ancient human remains	This paper	M53
Ancient human remains	This paper	M61-11
Ancient human remains	This paper	M13_DG
Ancient human remains	This paper	M3
Ancient human remains	This paper	H393-2
Ancient human remains	This paper	M24_Wd
Ancient human remains	This paper	M31_Yz
Ancient human remains	This paper	M91_Yz
Ancient human remains	This paper	M77_wcg
Ancient human remains	This paper	M79_wcg
Ancient human remains	This paper	M21_Wd
Ancient human remains	This paper	M71_wcg
Ancient human remains	This paper	M64XR_merge
Chemicals, peptides, and recombinant proteins		
Ethanol	Sinopharm	100092008
NaClO	Sinopharm	80010428
0.5 M EDTA, pH 8.0	Thermo Fisher Scientific	AM9262
Proteinase K	Beyotime	ST533
Guanidine hydrochloride	Sigma	G3272
Isopropanol	Sigma	34863
Acetic acid	Sigma	695092
Sodium acetate	Sigma	S7899
Tween 20	Sigma	P7947
Isothermal amplification buffer	NEB	B0537S
Deoxynucleotide (dNTP) solution mix	NEB	N0447L

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bst 2.0 DNA polymerase	NEB	M0537L
AMPure XP Beads	Beckman	A63881
Agarose	Biowest	111860
Tris-EDTA buffer solution (1003)	Sigma	T9285
Critical commercial assays		
MinElute PCR Purification Kit	QIAGEN	28006
NEBNext Ultra II DNA Library Prep Kit	NEB	E7645
Twist Mitochondrial Panel Kit	Twist	102040
Twist Binding and Purification Beads	Twist	100984
Twist Universal Blockers	Twist	100767
Twist Hybridization Reagents	Twist	100982
Twist Wash Buffers	Twist	100846
Twist Ancient Human DNA Panel	Twist	106658
Deposited data		
BAM files reported in this paper have been deposited in the GSA-Human (https://ngdc.cncb.ac.cn/gsa-human/) and are publicly available as of the date of publication.	This paper	HRA009449
Software and algorithms		
AdapterRemoval v2.3.3	Schubert et al. ²⁶	https://github.com/MikkelSchubert/adapterremoval ; RRID:SCR_011834
BWA v0.7.17	Li et al. ²⁷	https://bio-bwa.sourceforge.net/ ; RRID:SCR_010910
SAMtools v1.18	Li et al. ²⁸	http://samtools.sourceforge.net ; RRID:SCR_002105
bamUtil v1.0.15	Jun et al. ²⁹	https://github.com/statgen/bamUtil
DeDup v0.12.8	Peltzer et al. ³⁰	https://github.com/apeltzer/DeDup
pileupCaller	https://github.com/stschiff/sequenceTools	https://github.com/stschiff/sequenceTools
mapDamage v2.2.2	Jónsson et al. ³¹	https://ginolhac.github.io/mapDamage/ ; RRID:SCR_001240
Schmutzi	Renaud et al. ³²	https://github.com/grenaud/schmutzi
ContamMix	Fu et al. ³³	https://github.com/plfjohnson/contamMix
MAFFT	Katoh et al. ³⁴	https://mafft.cbrc.jp/alignment/software/ ; RRID:SCR_011811
ANGSD v0.910	Korneliussen et al. ³⁵	http://www.popgen.dk/angsd/index.php/ANGSD ; RRID:SCR_021865
HaploGrep2 v2.4.0	Weissensteiner et al. ³⁶	https://haplogrep.uibk.ac.at
Yleaf v2.2	Ralf et al. ³⁷	https://github.com/genid/Yleaf
EIGENSOFT	Patterson et al. ³⁸	https://github.com/DReichLab/EIG ; RRID:SCR_004965
ADMIXTOOLS (<i>qp3Pop</i> , <i>qpDstat</i> , <i>qpAdm</i>)	Patterson et al. ³⁹	https://github.com/DReichLab/AdmixTools/ ; RRID:SCR_018495
hapROH	Ringbauer et al. ⁴⁰	https://pypi.org/project/hapROH
PLINK v1.9	Purcell et al. ⁴¹	https://www.cog-genomics.org/plink/1.9 ; RRID:SCR_001757
READ	Monroy Kuhn et al. ⁴²	https://bitbucket.org/tguenther/read/src/master/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Description of archaeological background to human remains

In this study, 31 ancient samples from 5 archaeological sites were successfully sequenced (Figure 1A and Table S1). These samples were collected from the related provincial archaeology institutes and research universities, with their appropriate permissions. Our study was reviewed and approved by the Medical Ethics Committee of Xiamen University (XDYX202412K88).

Dinggong sites (丁公遗址)

The Dinggong site is located in the east of Dinggong Village, Changshan Town, Zouping County, Shandong Province, China²². Many relics of Longshan and Yueshi culture have been discovered at the site. The discovery of the prehistoric city site indicated that the Dinggong site was one of the central settlements during the Longshan cultural period in the Haidai area²². This study analyzed 10 Longshan culture-related individuals and 1 Han dynasty-related individual from Dinggong sites (Table S1). One Longshan culture-related sample, M77, was directly dated by radiocarbon dating to 4,396-4,093 cal BP.

Chengziya sites (城子崖遗址)

The Chengziya site is located in Zhangqiu District, Jinan City, Shandong Province, China⁴³. It is the first Neolithic site discovered and excavated by Chinese archaeologists. It is also the discovery and naming place of Longshan culture. This study analyzed 1 Longshan culture-related individual and 1 Yueshi culture-related individual from Chengziya sites (Table S1). Sample M3 with the archaeological background of Longshan culture was directly dated by radiocarbon dating to 2,756-2,542 cal BP. The radiocarbon dating for sample M3 did not overlap with the timespan of Longshan culture. We therefore labeled sample M3 as “lower_YR_Chengziya_2.5kya” in our study according to the radiocarbon dating result. Sample H393-2, with the archaeological background of Yueshi culture, was directly dated by radiocarbon dating to 3,550-3,367 cal BP.

Wadian sites (瓦店遗址)

The Wadian site is located in Yuzhou City, Henan province, China⁴⁴. It mainly contains early, middle, and late remains of the Longshan Culture. It is one of the largest human settlement sites in the late Longshan Culture period in China. This study analyzed 8 Longshan culture-related individuals from Wadian sites (Table S1). One Longshan culture-related sample, M21, was directly dated by radiocarbon dating to 3,982-3,780 cal BP.

Wangchenggang sites (王城岗遗址)

The Wangchenggang site is in Bafang Village, Gaocheng Town, Dengfeng, Henan Province, China⁴⁵. It is an ancient cultural site mainly composed of the middle and late stages of the Longshan Culture, the early Neolithic Peiligang Culture, and the Shang and Zhou Dynasties. 7 Erlitou culture-related individuals from Wangchenggang sites were analyzed in this study (Table S1). One Erlitou culture-related sample, M71_wcg, was directly dated by radiocarbon dating to 3,636-3,466 cal BP.

Yuzhuang sites (余庄遗址)

The Yuzhuang site is located in Ye County, Pingdingshan City, Henan Province, China⁴⁶. It is a large-scale settlement site from the Longshan period, with an area of approximately 500,000 square meters. More than 50 types of relics, such as tombs, house sites, and cellars, have been discovered. This study analyzed 3 Longshan culture-related individuals from Yuzhuang (Table S1). One Longshan culture-related sample, M64XR, was directly dated by radiocarbon dating to 4,406-4,153 cal BP.

Method details**Radiocarbon dating**

The preparation of bone/tooth samples involves standard acid-base-acid (ABA) procedure and collagen extraction. Firstly, bone/tooth samples are ultrasonicated in ultrapure water, dried, grinded and sieved to get the appropriately sized sample fraction (0.5–1 mm). Then samples are treated with 0.5M hydrochloric acid (~18 hrs), 0.1–0.2M sodium hydroxide (30 min-1h), and 0.5M hydrochloric acid (1 hr). Bone/tooth collagen gelatinization is performed in pH 3 solution at 70°C for 20 hrs. Gelatine solution is filtered using a cleaned filter and freeze-dried. The quality of collagen is monitored by carbon and nitrogen content in collagen, atomic C/N ratio determination and collagen yield. Samples are dated if the collagen yield is above 1% and the C:N ratio of the collagen is between 2.9 and 3.5. The samples that deviate from these ratios are deemed unsuitable for dating⁴⁷. The pure collagen samples were combusted with CuO and silver in vacuo to CO₂ at 900°C for 2 hrs. The CO₂ was purified, graphitized and measured on a compact 0.5 MeV NEC Accelerator Mass Spectrometry in the Guangzhou Institute of Geochemistry, Chinese Academy of Sciences⁴⁸. All four radiocarbon dates were modelled in OxCal⁴⁹ using the IntCal20⁵⁰ calibration curve.

Ancient DNA extraction and sequencing

We initially screened ancient samples in a controlled environment at Xiamen University’s Ancient DNA Laboratory, China, using established measures for handling ancient human DNA^{51–53}. We washed the bone samples with 75% ethanol to remove surface dirt and attachments. We carefully removed the surface layer using an electric drill to expose a fresh surface. Subsequently, we applied a 10% sodium hypochlorite solution on the newly exposed bone surface for 10 minutes to disinfect and eliminate the remaining contaminants. After rinsing the samples with absolute ethanol, we air-dried them under ultraviolet light for a minimum duration of 30 minutes to minimize microbial contamination before milling. Approximately 80–180 mg of bone powder was obtained through a methodical process, which entailed drilling into cementum, petrous parts of temporal bones, or bone sections of limb bones while carefully avoiding penetration into the marrow cavity. The bone powder was subjected to lysis by overnight incubation. Each sample received 1 ml of 0.5 M EDTA and 1 µl of 20 mg/ml Proteinase K, with adjustments made for samples containing more than 120 mg of bone powder. The samples were agitated at 300 rpm and incubated at 37°C overnight. We used a silica membrane-based MinElute Kit (Qiagen, Germany) to conduct DNA purification. The laboratory-prepared binding buffer containing 5M guanidine hydrochloride, 40% isopropanol, 25 mM sodium acetate, and 0.05% Tween-20 (Sigma Aldrich, Germany) at pH 5.5, which provides the correct salt concentration and pH for adsorption of DNA to the MinElute membrane. The clarified supernatant obtained after overnight digestion for each sample was combined with approximately 13 ml of binding buffer. The purified DNA was ultimately eluted using distilled water.

We prepared double-stranded libraries with no uracil-DNA-glycosylase (UDG) treatment to preserve the typical terminal damage of ancient DNA using the NEBNext Ultra II DNA Library Prep Kit (New England BioLabs) in conjunction with customized sticky-ends adaptors. We followed the protocol from David Reich's lab, employing Twist capture reagents for hybridization capture to enhance the acquisition of additional SNP loci⁵⁴. Post-PCR clean-up throughout the entire experimental procedure was achieved using 1.8X AMPure XP beads (Beckman Coulter, USA) to eliminate non-target DNA fragments below 100 bp. We then sequenced the libraries on the DNBSEQ-T7 platform.

Quantification and statistical analysis

DNA sequence data processing

We trimmed adapters from both read pairs and trimmed bases at 5'/3' termini with quality scores ≤ 20 and the ambiguous bases (N) (-trimns -trimqualities -minquality 20) and collapsed forward and reverse reads (-collapse) and discarded reads shorter than 30 bp (-minlength 30) using AdapterRemoval (version 2.3.3)²⁶. Collapsed reads were aligned against the human reference genome hs37d5 (GRCh37 with decoy sequences) using the aln and samse modules in the Burrows-Wheeler Aligner (BWA) program (version 0.7.17)²⁷ with options for turning off seeding (-l 1024) and allowing additional mismatches (-n 0.01). BAM files were sorted and indexed using Samtools (version 1.18)²⁸ and then used to remove PCR duplicates using dedup (version 0.12.8)³⁰. BAM files were filtered for a minimal Phred-scaled mapping quality score of 30 using Samtools (version 1.18)²⁸.

Data quality control

We assessed the authenticity of ancient DNA using multiple measures. First, we calculated the postmortem damage pattern using mapDamage (version 2.2.2)³¹. We checked whether each library showed increased C-to-T misincorporations at 5' termini and G-to-A misincorporations at 3' termini, as expected for double-stranded libraries. Second, we estimated the contamination rate. We used the ContamMix³³ R-script "estimate.R" to calculate the amount of authentically mapping mitochondrial reads at 311 modern diagnostic marker positions. We run ContamMix³³ with default options and the following two input-files we generated from all reads in the BAM file that mapped against the MT-genome: (1) An alignment of MT reads against their own consensus (-samFn). This consensus was constructed by iteratively mapping the reads with MIA (mia) and mapping the last iteration with MIA (mia) to obtain one FASTA-sequence (<https://github.com/mpieva/mapping-iterative-assembler>). The MT reads were mapped against their consensus using bwa aln and samse²⁷ and filtered for a minimum mapping quality of 30 with Samtools (version 1.18)²⁸. (2) A multiple alignment of the consensus genome and a fasta-file containing the diagnostic marker positions against each other (-malnFn) obtained with mafft³⁴. We used one minus the value of MAP authentic (the probability of being authentic) reported by ContamMix software to represent the mtDNA contamination rate.

The mtDNA contamination rate was also assessed using the schmutzi.pl module in schmutzi³² based on the contaminant database designed for Eurasian samples as the authors recommended (share/schmutzi/alleleFreqMT/eurasian/freqs). We used values reported by schmutzi_pl to represent the mtDNA contamination rate if ContamMix software failed to output results.

The X chromosomal contamination rate was estimated for individuals with assigned male biological sex using the contamination module of the ANGSD software (version 0.910)³⁵. We used the HapMap resources for CHB (Han Chinese in Beijing), which were provided with the ANGSD software to define polymorphic sites. We restricted the analysis to the non-recombining portion of the X chromosome (X:5,000,000-154,900,000). We used "Method2: new_llh: Methods of Moments" (MoM) and its standard error SE(MoM) reported by ANGSD software to represent the X chromosomal contamination rate. If a male individual had low contamination of the nuclear DNA as determined by the X chromosome but slightly higher contamination of the mtDNA, we used all fragments of this male individual for further analyses.

Biological sex determination

We used "male" to indicate an individual with one X chromosome and one Y chromosome and "female" to indicate an individual with two X chromosomes. The biological sex of samples based on 1240K capture data was determined by the method described in Reference⁵⁵. This method calculated the ratio of the alignments to chromosome Y to chromosome autosomes, divided by the expected value of the quantity based on the number of SNPs in the relevant target set. The depth of coverage was calculated by the depth module of Samtools (version 1.18)²⁸ on 1240K SNPs with mapping quality ≥ 30 ("–Q 30") and base quality ≥ 30 ("–q 30"). The expected value of $\frac{Y_{Cov}}{\text{autoCov}}$ is ~ 0.5 for males and ~ 0 for females. The individuals with the observed value of $\frac{Y_{Cov}}{\text{autoCov}} > 0.3$ are reported as male, and individuals with the observed value of $\frac{Y_{Cov}}{\text{autoCov}} < 0.1$ are reported as female, as described in Reference⁵⁶.

Genotyping

To minimize the bias due to ancient DNA deamination, we masked 9 bp from both ends using the trimBam module in BamUtil (version 1.0.15)²⁹. We randomly called the genotype for the Twist panel of SNPs with high-quality base and mapping quality (–q 30 and –Q30) using pileupcaller (<https://github.com/stschiff/sequenceTools>).

Estimation of genetic relatedness

We first applied PLINK software (version 1.9) to exclude non-polymorphic and low-frequency variants (–maf 0.01). We then used the READ software⁴² with default parameters based on pseudo-haploid genotypes to estimate the degrees of kinship between every pair of newly generated 31 individuals.

Uniparental haplogroup assignment

We generated mitochondrial consensus sequences of quality ≥ 30 using the log2fasta program in the Schmutz³². Mitochondrial haplogroups were then inferred with Haplogrep (version2)³⁶. For the males individuals, we used Yleaf (version2.2)³⁷ program with option “-r 1-q 30 -b 90” to infer Y chromosomal haplogroups.

We here described the manual check for the Y chromosomal haplogroup assignment for 12 males:

- (1) middle_YR_Wadian_Longshan sample H6 could be assigned as O2a1a1a1 based on the mutation O2a1a1a1-CTS8587:18093660G->A, but we caution this might be caused by ancient DNA damage. This sample also had upstream-derived mutations for haplogroup O2a1a-F3143:21178130T, O2a1a1-F2726:18160098T, O2a1a1a-F1867:14928001A. This sample also showed ancestral alleles at O2a1a1a2-MF14363:7328526C, O2a1a1b-CTS7667:17602041G, O2a1a1b-F3182:21365517T-, O2a1a1b-F3417:23195006G, O2a1a1b-F2633:17911458C, O2a1a1b-F1675:14033835G.
- (2) middle_YR_Wadian_Longshan sample M13_WD could be assigned as O2a2b1a2a1 based on the mutation O2a2b1a2a1-CTS2643:14402768G->C. However, we note this assignment may not be correct due to the low coverage of the data and only a few observed mutations. This sample also had upstream-derived mutations for haplogroup O2a2b1a2a-F79:7540917C->T and O2a2b1a2a-F314:15947171G->T. This sample also showed ancestral alleles at O2a2b1a2a1a1-F523:19137297A and O2a2b1a2a1a1-F163:8725496C.
- (3) The Y haplogroup assignment estimated by Yleaf software for middle_YR_Wadian_Longshan sample M21 was not reliable. The software assigned sample M21 as R1b1a1b1a1a1a2b, however, all derived mutations for haplogroup R (R1a1a1-b1a3a2b2b-AM00559:18175815G->A, R1b1a1b-CTS894:7073423G->A, R1b1a1b-L753:18865298C->T, R1b1a1b1a1a1a2b-A561:8554311G->), were related to G->A or C->T the mutation type related to ancient DNA damage for double-strand libraries.
- (4) middle_YR_Wangchenggang_Erlitou sample M63_WCG could be assigned as O2a2b1a1a2a based on the mutation O2a2b1a1a2a-F310:15929506A->C. However, we note this assignment may not be correct due to the low coverage of the data and only a few observed mutations. This sample also had upstream derived mutations for haplogroup O2a2b1a1-CTS5128:16028396T->C, O2a2b1a1-CTS3251:14768861G->A, O2a2b1a1-F476:18209413G->C, O2a2b1a1-CTS6623:17014495G->T, O2a2b1a1F649:23578463T->C, O2a2b1a1-F342:16472742A->G. This sample also showed ancestral alleles at O2a2b1a1a2a1a-F1531:9387194G and O2a2b1a1a2a1a-F2781:18584672A.
- (5) lower_YR_Dinggong_Longshan_o1 sample M10 could be assigned as N1b1a2b~ based on the mutation N1b1a2b~Y13913:2797465T->G, N1b1a2b~Y13362:16324441T->C, N1b1a2b~Y13910:23999094T->C, N1b1a2b~Y13361:15111956G->A, N1b1a2b~Y13911:15005817A->G, N1b1a2b~Y13914:7800785C->A, N1b1a2b~Y13916:7870645C->T, N1b1a2b~Y13363:17241808A->G, N1b1a2b~Y13922:17197864C->A, N1b1a2b~Y13912:15312203C->A, N1b1a2b~Y13360:7915150T->A, N1b1a2b~Y13919:9872564A->G. This sample also had upstream-derived mutations for haplogroup N1b1a2~L727:6909100G->A. This sample also showed ancestral alleles at N1b1a2b~Y16099:17339823T, N1b1a2b~Y15972:19153013C, N1b1a2b~Y16109:18560399A, N1b1a2b~Y16088:8670198G, N1b1a2b~Y16104:22824093C, N1b1a2b~Y16098:16271412T. The derived mutation for N1b1a2~L727:6909100G->A might be caused by ancient DNA damage.
- (6) lower_YR_Dinggong_Longshan sample M31 could be assigned as N1b1a2b~ based on the mutation N1b1a2b~Y13366:22797209C->G, N1b1a2b~Y13361:15111956G->A, N1b1a2b~Y13362:16324441T->C, N1b1a2b~Y13908:15547580G->A, N1b1a2b~Y13915:7870528T->C, N1b1a2b~Y13914:7800785C->A, N1b1a2b~Y13922:17197864C->A. This sample also had upstream-derived mutations for haplogroup N1b1-CTS8796:18208410T->C, N1b1-Z4936:21282669C->T, N1b1-F4250:17881667C->T, N1b1-Z19696:19301928C->T. This sample also showed ancestral alleles at N1b1a2b~Y16104:22824093C, N1b1a2b~Y16098:16271412T, N1b1a2b~Y16099:17339823T.
- (7) lower_YR_Dinggong_Longshan_o2 sample M77 could be assigned as C2b1a2b2 based on the mutation C2b1a2b2-FGC45562:14734460A->G, C2b1a2b2-FGC45566:22007680G->C, C2b1a2b2-FGC45551:7739238G->A, C2b1a2b2-Z31671:28778595G->A, C2b1a2b2-FGC45560:19545799G->A, C2b1a2b2-PH1109:14312438C->T, C2b1a2b2-FGC45548:8131816T->C, C2b1a2b2-Z45207:9051877C->G, C2b1a2b2-PH1906:15940435T->G, C2b1a2b2-PH2194:16337272T->A. This sample also had upstream derived mutations for haplogroup C2b1a2b-F13141:22091718A->G, C2b1a2b-F12415:18103043T->C, C2b1a2b-F12881:21353808A->G, C2b1a2b-F14035:28768557C->T, C2b1a2b-CTS10989:22874654G->A, C2b1a2b-F12528:18713612C->T, C2b1a2b-F9951:7742460G->T, C2b1a2b-F12583:18943811G->T, C2b1a2b-F10136:8545183C->T, C2b1a2b-CTS2457.1:14313081C->T, C2b1a2b-F12408:18060147C->T, C2b1a2b-F10004:7961051C->A, C2b1a2b-F13062:21893184T->G. This sample also showed ancestral alleles at C2b1a2b2a~Z31672:2815306G, C2b1a2b2a~Z31669:8399738C, C2b1a2b2a~PH404:7056501C, C2b1a2b2b~Z45210:7352364A, C2b1a2b2b~Z45209:7267405T, C2b1a2b2b~Z45208:7126834G, C2b1a2b2b~Z45211:7616819C, C2b1a2b2b~Z45212:7734496C.
- (8) lower_YR_Dinggong_Longshan sample M110 could be assigned as O1b1a1b1b based on the mutation O1b1a1b1b-F15640:7151735C->T (coverage=3). This sample also had upstream-derived mutations for O1b1a1b1-Z23602:9927052G->A and O1b1a1b1-CTS3857:15228137G->C. This sample also showed ancestral alleles at O1b1a1b1b1a-Y160

504:6428783A and O1b1a1b1b1a-F21414:17137456A. The derived mutation for C2a1b-BY74835:8640307C->T might be caused by ancient DNA damage.

- (9) lower_YR_Chengziya_Yueshi sample H393-2 could be assigned as C2b1b based on the mutation C2b1b-Z12264:8770161G->A, but we caution this might be caused by ancient DNA damage. This sample also had upstream derived mutations for C2b1-CTS6374:16859254A->G, C2b1-Z4149:24001146G->A, C2b1-Z1313:7788727C->T, C2b1-F1223:8463332T->C, C2b1-Z3967:8195260G->A, C2b1-F3789:17792483T->G, C2b1-CTS7414:17468144A->G, C2b1-F978:7527930G->A, C2b1-K493:9051841A->C, C2b1-CTS11292:23059222C->T, C2b1-F2385:17113314G->T, C2b1-F3854:7557268C->T, C2b1~CTS6866:17157702G->T, C2b1~F1453:8895766G->A, C2b1~CTS5399:16251407C->T, C2b1~CTS12964:28766196G->A, C2b1~CTS10238:19340515T->C. The derived mutation for C2a1b-BY74835:8640307C->T might be caused by ancient DNA damage.
- (10) middle_YR_Wangchenggang_Erlitou sample M77_wcg could be assigned as O2a1b1a1a1a based on the mutation O2a1b1a1a1a-F11:2815303C->G, O2a1b1a1a1a-F60:7158696A->G, O2a1b1a1a1a-F119:8231717T->C. This sample also had upstream derived mutations for O2a1b1a1a1-F379:16946526G->A, O2a1b1a1a1-FGC12474:21569893G->T, O2a1b1a1a1-F62:7177161T->C, O2a1b1a1a1-F68:7204313A->G. This sample also showed ancestral alleles at O2a1b1a1a1a1-F646:23482948C and O2a1b1a1a1a1-Z25065:21616774T.
- (11) middle_YR_Wangchenggang_Erlitou sample M79_wcg could be assigned as O2a based on the mutation O2a-P199:18647042A->G. This sample also had upstream-derived mutations for O-F600:22166455G->T and O2-M122:21764674A->G. This sample also showed derived alleles at O2a2b1a2b1a-A16615:2764643T->C and O2a2b1a2b1a-CTS4325:15579254G->A. However, the genotypes of upstream mutations for O2a2b1a2b1a were unknown.
- (12) middle_YR_Wadian_Longshan sample M21_Wd could be assigned as C2b1a1a1a based on the mutation C2b1a1a1a-F6968:14843790C->G and C2b1a1a1a-Y12946:22576525T->C. This sample also had upstream-derived mutations for C2b1a1a-CTS11990:23406293C->A and C2b1a1a-F3949:16904813A->T. This sample also showed ancestral alleles at C2b1a1a1a1a4~Z31866:14652065A, C2b1a1a1a1a5b~-B107:15984162A, and C2b1a1a1a2~Z45403:8273488C. The derived mutation for C2b1b2-BY106576:16274452C->T might be caused by ancient DNA damage.

Data merging

We have provided the co-analyzed ancient and modern genome information and the corresponding original reference paper in **Tables S2** and **Table S3**. The *convertf* program in EIGENSOFT³⁸ software was used to extract the populations listed in **Tables S1B** and **S1C** from the 1240K and Human Origin (HO) datasets curated by The Allen Ancient DNA Resource (AADR)⁵⁷, respectively. We then used the *mergeit* program in EIGENSOFT software³⁸ to merge the 1240K and HO datasets with our newly generated dataset, respectively.

Runs of homozygosity

ROH refers to segments of the genome where the two chromosomes in an individual are identical to each other owing to recent common ancestry. Therefore, long ROH segments strongly suggest that an individual's parents are related. We applied the hapROH method using the Python library hapROH⁴⁰ with default parameters. The method was developed to identify ROH from low-coverage genotype data typical of ancient DNA and is still robust enough to identify ROH for individuals with coverage down to 0.5×. We reported the total sum of ROH longer than 4, 8, 12, and 20 cM and visualized the results using Datagraph.

Principal component analysis (PCA)

We performed PCA on the merged HO dataset using the smartpca (version 16000) program in EIGENSOFT software³⁸ with the options “Isqproject: YES” and “numoutlieriter:0.” The option “Isqproject: YES” was used to project all ancient genomes on the modern background. The option “numoutlieriter:0” was used to turn off outlier removal.

F statistics

We used *qp3pop* (version 651) from the ADMIXTOOLS software³⁹ with the option “inbreed: YES” to calculate outgroup f_3 . Outgroup- f_3 statistics has the form $f_3(A, B; C)$ where C is an outgroup (i.e., equally closely related to all left populations) and is used to calculate the shared genetic drift between populations A and B after divergence from the outgroup C. Yoruba from sub-Saharan Africa is often chosen as plausible outgroups with human populations from Eurasia. The larger the outgroup f_3 value, the more genetic drift between population A and population B; in other words, the closer the genetic relationship between population A and population B. The standard error (Std. err) was calculated with 5 centiMorgan (cM) block jackknifing implemented in the ADMIXTOOLS software³⁹.

We used *qpDstat* (version 980) from the ADMIXTOOLS software³⁹ with the option “f4mode: YES” to calculate f_4 statistics in the form of f_4 (Yoruba, X; Y, Z). The standard error (Std. err) was calculated with 5 centiMorgan (cM) block jackknifing implemented in the ADMIXTOOLS software³⁹. The null hypothesis for f_4 statistics is Y and Z formed a sister clade relative to X. f_4 statistics computed a Z-score for the deviation of f_4 -statistics from zero based on a block jackknife standard error. The absolute value of a Z-score larger than 3 suggests a significant rejection of the null hypothesis. Z-score>3 indicated X shared more derived alleles with Z than with Y. Z-score<-3 indicated X shared more derived alleles with Y than with Z.

Admixture modeling using qpAdm

We used *qpAdm* (version 1000) from the ADMIXTOOLS software³⁹ to model our genomes as an admixture of potential ancestral source populations and estimate ancestry proportions, with the options “details: YES” and “allsnps: YES”. The outgroup set applied in each *qpAdm* result was listed in the legend of Supplementary Tables.