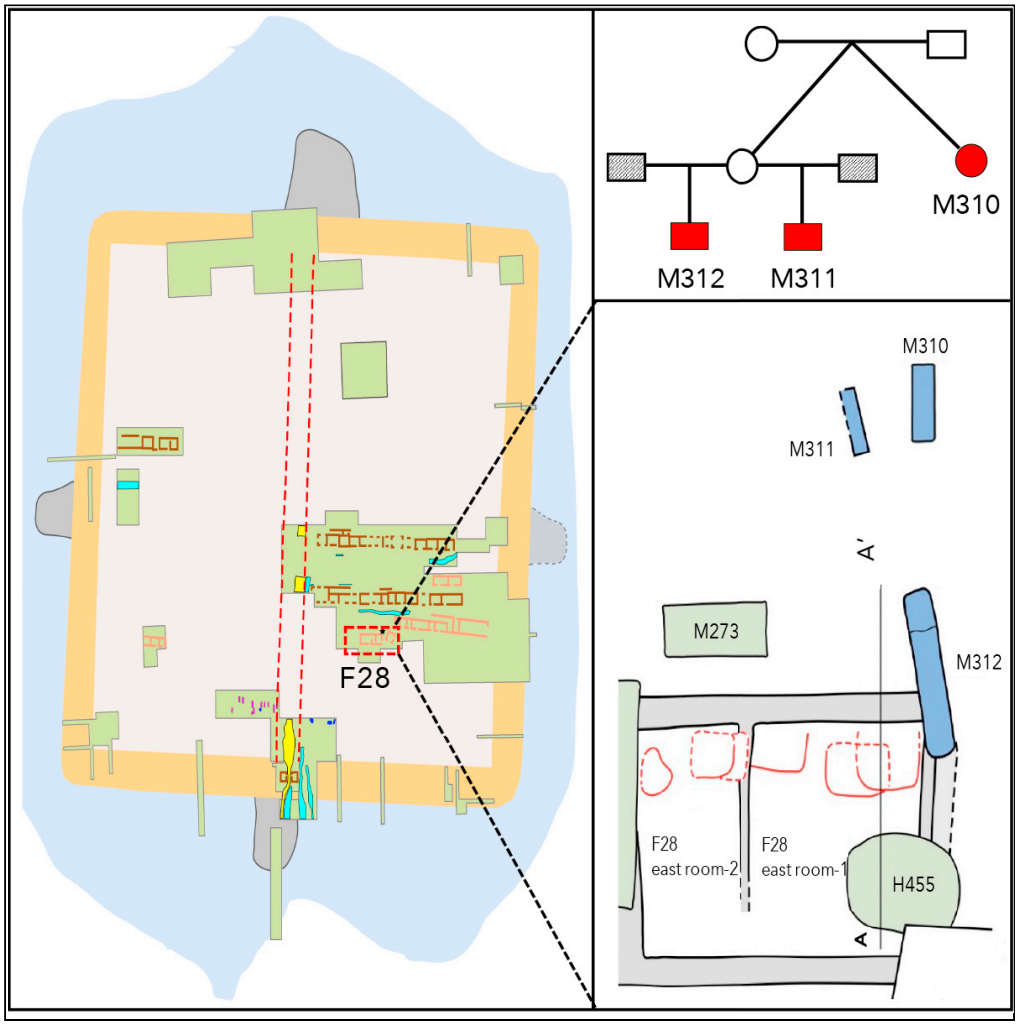
iScience

CellPress
OPEN ACCESS

Article

Ancient genome analyses shed light on kinship organization and mating practice of Late Neolithic society in China



Chao Ning, Fan Zhang, Yanpeng Cao, ..., Martine Robbeets, Hai Zhang, Yinqiu Cui

haizhang@pku.edu.cn (H.Z.)
cuiyq@jlu.edu.cn (Y.C.)

Highlights

Accurate genetic kinship estimation for ancient individuals in Late Neolithic China

Direct evidence of inbreeding in China, 2000 years earlier than historical record

Longshan household was based on extended family beyond the nuclear family

Northward gene flow into Yellow River basin populations since the Yangshao period

Ning et al., iScience 24, 103352
November 19, 2021
https://doi.org/10.1016/j.isci.2021.103352

iScience

CellPress
OPEN ACCESS

Article

Ancient genome analyses shed light on kinship organization and mating practice of Late Neolithic society in China

Chao Ning,[1,2,7] Fan Zhang,[3,7] Yanpeng Cao,[4] Ling Qin,[5] Mark J. Hudson,[2] Shizhu Gao,[6] Pengcheng Ma,[3] Wei Li,[5] Shuzheng Zhu,[4] Chunxia Li,[5] Tianjiao Li,[3] Yang Xu,[3] Chunxiang Li,[3] Martine Robbeets,[2] Hai Zhang,[5,*] and Yinqiu Cui[1,3,8,*]

SUMMARY

Anthropology began in the late nineteenth century with an emphasis on kinship as a key factor in human evolution. From the 1960s, archaeologists attempted increasingly sophisticated ways of reconstructing prehistoric kinship but ancient DNA analysis has transformed the field, making it possible, to directly examine kin relations from human skeletal remains. Here, we retrieved genomic data from four Late Neolithic individuals in central China associated with the Late Neolithic Longshan culture. We provide direct evidence of consanguineous mating in ancient China, revealing inbreeding among the Longshan populations. By combining ancient genomic data with anthropological and archaeological evidence, we further show that Longshan society household was built based on the extended beyond the nuclear family, coinciding with intensified social complexity during the Longshan period, perhaps showing the transformation of large communities through a new role of genetic kinship-based extended family units.

INTRODUCTION

The important role assigned to kinship in human social evolution by Morgan and other early anthropologists had a profound influence on social theory from the late nineteenth century (Morgan, 1877; Engels, 1954). From the 1960s, ethnographic fieldwork among hunter-gatherers found quite flexible post-marital residence arrangements which did not support a historical shift from matrilocality to patrilocality, adding to social anthropology's growing critique of the classical theorists (Marlowe, 2004, 2005; Goody, 1990). During the same period, however, archaeology developed a renewed interest in reconstructing prehistoric social organization in order to deepen its contribution to neo-evolutionary anthropology (Binford, 1962). The techniques of mortuary and ceramic analysis were tested and later supplemented with bioarchaeological methods including craniodental metrics and nonmetrics and stable isotopes (Longacre, 1964; Hill, 1970; Haniharai et al., 1983; Gao and Lee, 1993; Price et al., 1994). Genetics, however, is the only discipline that can demonstrate explicit kin relationships among prehistoric human individuals by analyzing the DNA extracted from ancient human remains. With the application of next-generation sequencing technologies, DNA retrieved from ancient remains can now be sequenced to low (usually < 1x) coverage in a cost-effective way. The rapid growth of this technology means more low-coverage ancient genomes are available each year. However, most such studies have focused on population history, and the great potential of kinship analyses in reconstructing ancient social organization has not received sufficient attention. Although there is a growing trend that the field is moving from a focus on large-scale transregional studies to more local perspectives on socio-economic processes (Schroeder et al., 2019; Mittnik et al., 2019; Amorim et al., 2018; O'Sullivan et al., 2018; Sánchez-Quinto et al., 2019), such studies are still relatively limited. One of the primary reasons for this is that the most commonly used methods, such as PLINK (Lipatov et al., 2015), to identify genetic relatedness among modern individuals would lead to considerable bias when testing ancient individuals. Recently, several programs designed especially to estimate genetic kinship from low-coverage sequencing data have shown great power to determine relatedness among ancient individuals (Lipatov et al., 2015; Kuhn et al., 2018; Ringbauer et al., 2020). Here, we apply ancient DNA analyses to skeletal remains from the Pingliangtai ancient city site (hereafter 'Pingliangtai site'), an archaeological site of the Late Neolithic Longshan culture located in Henan province in the middle Huai River Basin (HR) (Figure 1A).

[1]Research Center for Chinese Frontier Archaeology of Jilin University, Jilin University, Changchun 130012, China

[2]Max Planck Institute for the Science of Human History, 07745 Jena, Germany

[3]School of Life Sciences, Jilin University, Changchun, 130012, China

[4]Henan Provincial Institute of Cultural Heritage and Archaeology, Zhengzhou 450000, China

[5]School of Archaeology and Museology, Peking University, Beijing 100871, China

[6]College of Pharmacia Sciences, Jilin University, Changchun 130021, China

[7]These authors contributed equally

[8]Lead contact

*Correspondence: haizhang@pku.edu.cn (H.Z.), cuiyq@jlu.edu.cn (Y.C.)

https://doi.org/10.1016/j.isci.2021.103352

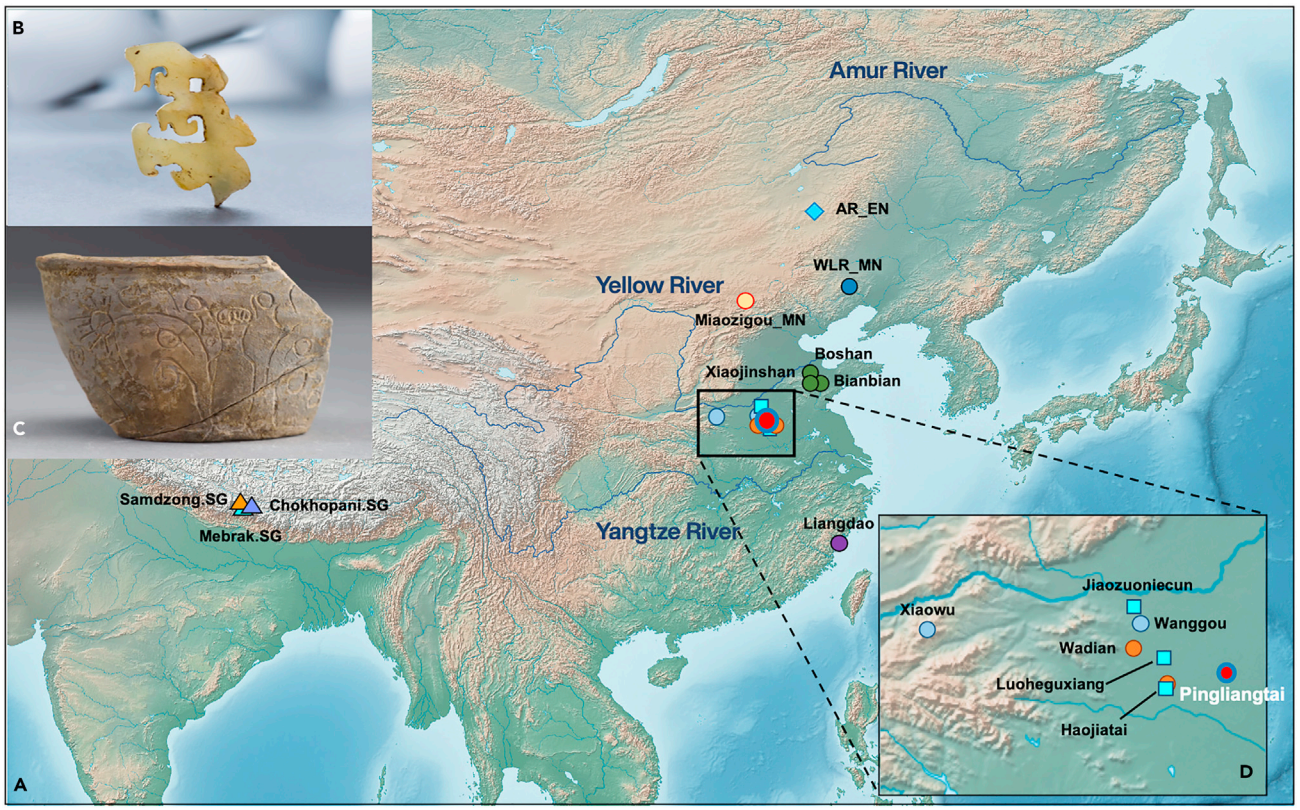


Figure 1. Geographic location of the Pingliangtai site

(A) Location of Pingliangtai and key reference cities in East Asia.

(B) Jade artifact excavated from Pingliangtai.

(C) Longshan pottery excavated from Pingliangtai with incised geometric patterns.

(D) A zoom on Central Plain in China shows the geographical location of the Pingliangtai and other Neolithic archaeological sites.

The Longshan culture (ca. 4,500–3,800 BP) is regarded as an important transitional stage in the development of Chinese civilization from independent communities to dynastic states (Liu, 1996a, 2005; Zhang et al., 2019a). Studies of burial patterns, house distributions, and dietary isotopes confirm major changes in Longshan social organization (Pearson, 1981; Dong et al., 2019; Liu, 1996b). Compared to the Middle Neolithic Yangshao culture, the Longshan witnessed the shrinking of public cemeteries and the disappearance of large tombs such as secondary burials with multiple individuals. At the same time, archaeology indicates enhanced social mobilities in terms of both the integration of multicultural artifacts (Li, 2000) and the gathering of groups of people with different dietary habits (Chen et al., 2016). These changes have suggested to archaeologists that the Longshan witnessed a major transformation in kinship organization, perhaps from large extended to small nuclear families (Pearson, 1981, 1988). However, without accurate identification of relatedness for individuals from the same community or more specifically from the same residence, such a hypothesis cannot be tested empirically. Thus, characterizing the genetic kinship of Longshan individuals is of great importance in understanding the family structure, mating patterns, and underlying the social complexity of prehistoric populations during the formative period of Chinese civilization.

Four individuals excavated from Late Neolithic Pingliangtai were included in this study. Of them, three individuals (M310, M311, M313) are juveniles. M312 was identified as a young male of around 20 years on osteological features (Figures 2C–2E) (Sun et al., 2019). All four individuals were buried nearby the house foundations rather than in the public cemetery, and this type of burial pattern is common in many Longshan societies in the Central Plain area of China (Underhill, 2000; Yang and Xu, 1985; Major and Cook, 2016; Carr, 1995).

It has long been hypothesized that juveniles buried close to a house might live in it before their death and share a close kinship with the owner of the house (Liu, 1996b). To trace the potential genetic kinship

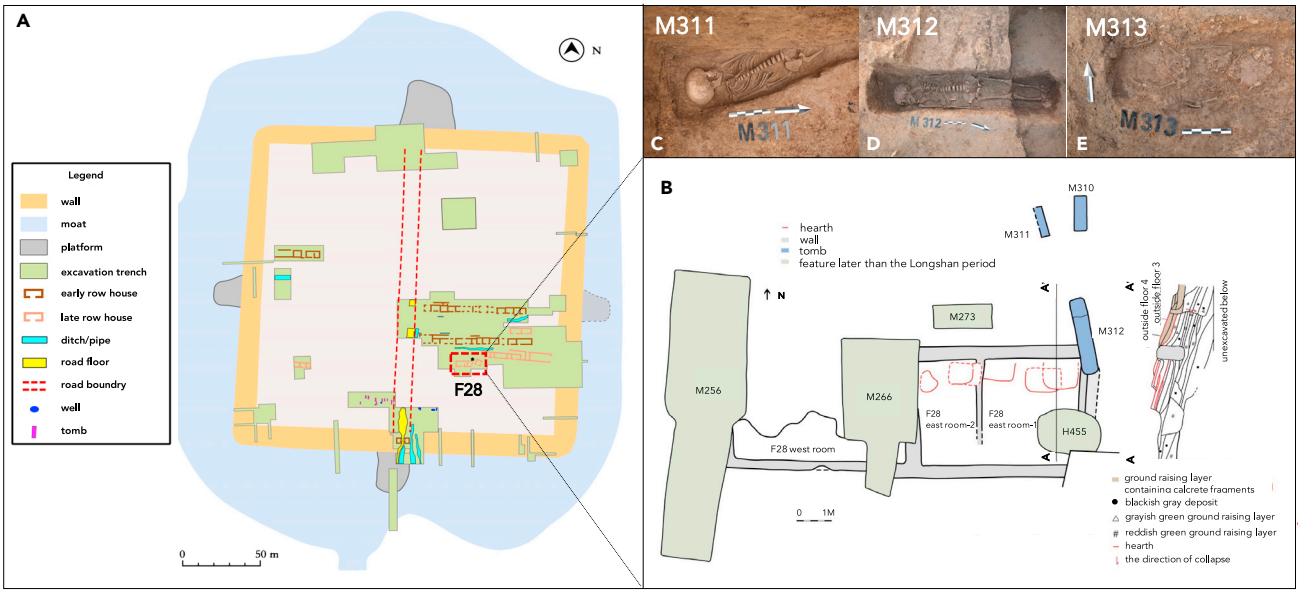


Figure 2. An overview of the Pingliangtai walled town site

(A) The reconstruction of the Pingliangtai walled town based on the archaeological excavations during 2016 and 2019.

(B) A zoom in look at the positional relationship of the three individuals from the north face of house F28: burials of the three individuals are marked in blue.

(C–E) Photos of the three individuals in this study. Burial M310 is not shown as the burial was seriously disturbed.

patterns in the Pingliangtai site, we further shotgun sequenced the four individuals previously reported in (Ning et al., 2020) to a higher mean coverage of 3.2x. Using autosomal, mitochondrial, and Y-chromosomal markers as well as combining the archaeological findings, we identified close genetic kinship among the individuals and provided direct evidence of a genetic kinship-based extended family structure in the Late Neolithic Longshan culture. Subsequently, the runs of homozygosity (ROH) analyses revealed one case of consanguineous union in these samples. By a multidisciplinary approach, including archaeological, anthropological, and archaeogenomics, we gain insight into a complex and transforming society, in which genetic kinship seems to be a focal point of social organization, with evidence that extended family-based household units had already appeared in the central part of China 4,000 years ago.

RESULTS

Archaeological and anthropological insight into the Pingliangtai site

The Pingliangtai site is located southeast of Huaiyang city, Henan province, China. The site was excavated by a joint team of Henan Provincial Institute of Cultural Heritage and the School of Archaeology and Museology of Peking University between 2014 and 2016. A total of 14 tombs dating to the Longshan period were excavated, 8 were aligned neatly and composed a small public cemetery in the southwestern area of the site and were all identified to be adults; the other 6 were scattered nearby the house foundations without any burial goods and were identified as individuals of less than 14 years old with one exception of around 20 years (Henan Provincial Institute of Cultural Heritage and School of Archaeology and Museology of Peking University, 2017). The houses were aligned in three rows, with F26, F22, F23, and F34 to the north, F28 and F36 forming the south row and F30, F40, and F41 in the middle (Figure 2A). Except for M313 who was buried in front of F34, all the other three individuals (M310, M311, M312) were buried in front of F28, which is distant from F36 from the same row and thus are believed to have lived in that house before their death based on archaeological observations. Burying juveniles in front of the house foundation is a tradition that is commonly seen in the eastern part of Henan province during the Yangshao and Longshan period. Both M310 and M311 were identified as from the same stratigraphic layer and were buried after the construction of F28 as the opening of both burials disturbed the early soil cushion of F28 (Figure 2A). M312, however, was buried slightly later than M310 and M311 as the opening of his burial disturbed the later soil cushion of F28, a stratigraphic relationship agreeing with the radiocarbon dates on their skeletons (Figure 2B, Table 1). All the four juveniles were buried in a small tomb without any burial goods, respectively, which was significantly different from other elite burials in the Longshan society. F28 was a two-room house

Table 1. A summary of Pinglinagtai samples reported in this study

Sample ID	Age at death	Date (cal BCE)	Genetic sex	Autosomal coverage	MtDNA coverage	MtDNA haplogroup	Y haplogroup	X-Contam	mt-Contam (95% CI)	1240k SNPs
M310	10 ± 3	2275-2045	Female	3.38	115.84	D4b1a	–	–	0.01 (0.01–0.02)	972,422
M311	8 ± 2	2201-1844	Male	4.36	203.06	D4b1a	N1b2*	0.0082 ± 0.0009	0.01 (0.01–0.02)	1,063,341
M312	20 ± 3	2135-1939	Male	0.74	23.11	D4b1a	N1b2*	0.0315 ± 0.0192	0.03 (0.02–0.04)	454,223
M313	Infant[a]	2118-1894	Female	4.29	274.6	pre-F2h	–	–	0.01 (0.01–0.02)	1,068,952

[a]The skeleton is too degraded and the exact age is hard to determine.

that was no obvious difference from the other house foundations. This provides an ideal case to test the basic household structure of the Longshan society by characterizing the genetic kinship of individuals who lived in the same household. The separation of the juvenile burials from the public cemetery where adults were buried was a custom broadly practiced in the Yellow River region during the Neolithic. Some archaeologists believe this may relate to some kind of ancestor worship wherein individuals who died at a young age or prior to mating were considered to be inauspicious and were not allowed to be buried in the public cemetery (Dong et al., 2019).

Ancient DNA authentication and uniparental genetic analyses

Four individuals from nearby house foundations at the Pingliangtai site were shotgun sequenced to medium coverage, ranging from 0.74x to 4.36x. Radiocarbon dates place the four individuals within a time frame of 2,275–1,844 cal BCE (Table 1). The authenticity of the ancient DNA was verified with multiple methods. All individuals showed damage patterns characteristic of ancient DNA (see Figure S1; Data S1A) and low levels of modern human DNA contamination (Table 1). We determined the biological sex of the four Pingliangtai individuals by comparing the ratio of X and Y chromosome coverages to autosomes (Fu et al., 2016). As a result, M310 and M313 were assigned as females with X rates of 0.783 and 0.862, as well as negligible Y rates (0.003 and 0.01, respectively). M311 and M312 were identified as males with approximately similar X and Y ratios (X and Y ratio: 0.412, 0.392; 0.44, 0.28, respectively) (see Figure S2; Data S1A). We retrieved complete mtDNA sequences for all four individuals with coverage varying from 23 to 274 fold, and they were further assigned to an explicit haplogroup (Table 1). Burials M310, M311, and M312 from the same layer share the same mitochondrial haplogroup belonging to D4b1a and all of them hold identical mtDNA consensus. The later burial M313, however, carries a different haplogroup that was assigned to pre-F2h (Table 1). Both D4b1a and pre-F2h are most prevalent in present-day East Asian populations (Dryomov et al., 2015; Kutanan et al., 2017). Haplogroup D4b1a has its highest frequency in Northeast Asian populations such as Han Chinese, Japanese, and Koreans, as well as populations in the Amur Basin and the Russian Far East (Dryomov et al., 2015). In contrast, pre-F2h is prevalent in considerable frequency in Southeast Asian populations such as in Taiwan and Thailand (Kutanan et al., 2017). The two male individuals (M311 and M312) were assigned to Y chromosome haplogroup N1b2*, which is a widespread lineage in modern Northeastern Asians such as Sino-Tibetan speaking populations (Wen et al., 2019). In short, from uniparental genetic analyses, we found that three of four individuals (M310, M311, and M312) from the Pingliangtai site were potentially maternally related and the two male individuals (M311 and M312) were probably paternally related.

Neolithic period genetic contribution into the Yellow River basin from southern China

To investigate the genetic ancestry and to determine the relationship of the four Neolithic Pingliangtai individuals with preceding and present-day populations in China, we prepared a dataset by merging our data with the "Human Origins" panel (Mathieson et al., 2015) as well as 96 published ancient populations

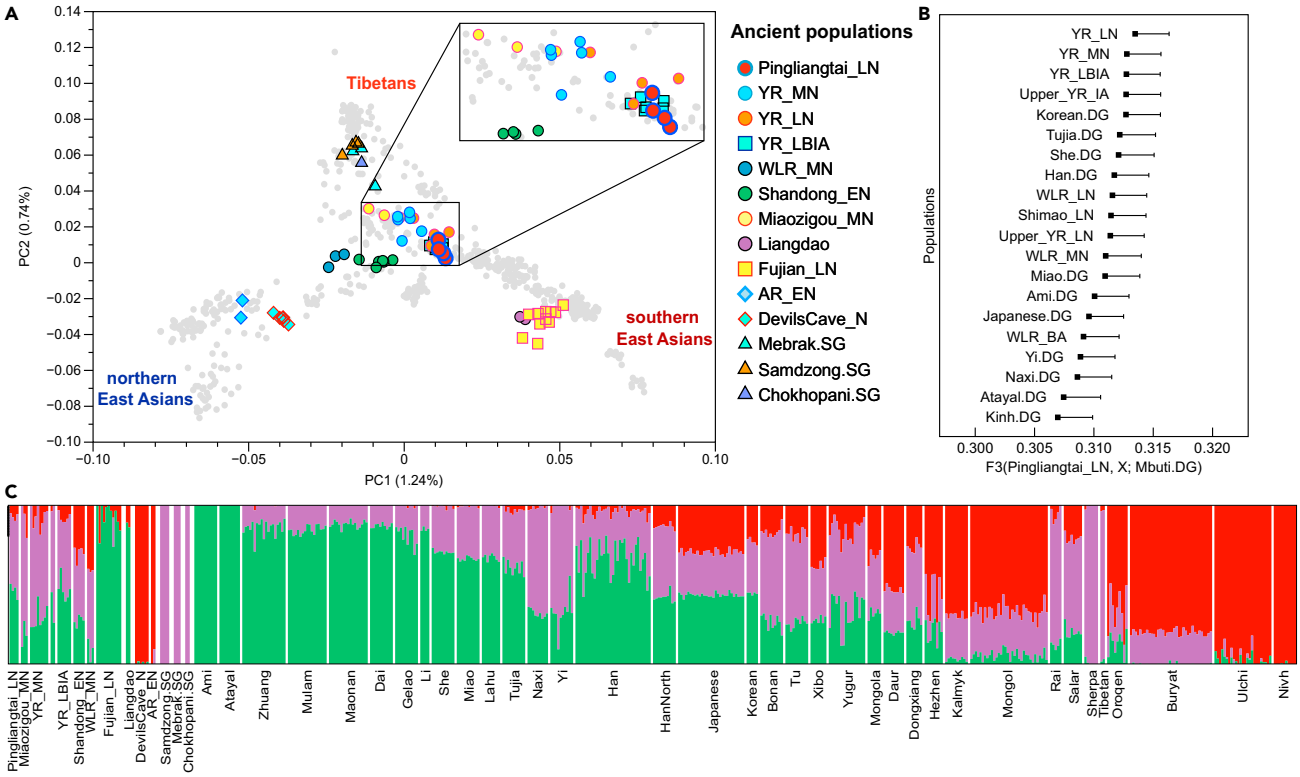


Figure 3. Genetic structure of Pingliangtai individuals

(A) The PCA was constructed from present-day East Asian populations in the "HumanOrigins" dataset, the ancient individuals are projected onto the top PCs.

(B and C) (B) The top 20 populations who share the closest relationship with Pingliangtai individuals across 257 world-wide populations in the "1240k" panel with the 1 standard errors marked in black, horizontal bars represent ±1 standard error (SE) calculated by 5 cM block jackknifing (C) Plot of ADMIXTURE (K = 3) results containing various modern and ancient populations in East Asia.

YR: Yellow River basin; WLR: West Liao River basin; AR: Amur River basin; EN: Early Neolithic; MN: Middle Neolithic; LN: Late Neolithic; LBIA: Late Bronze and Iron Age; Liangdao: Early Neolithic individuals from the Liang island; Mebrak.SG: Ancient individuals from Nepal (2,400-1,850 BP); Samdzong.SG: Ancient individuals from Nepal (1,750-1,250 BP); Chokhopani.SG: Ancient individuals from Nepal (3,150-2,400 BP). Detailed descriptions and references of comparative populations are provided in (see Data S2).

(see Data S2A). We then performed principal component analysis (PCA) with a set of modern East Asian individuals (see Data S2B). We found that all four Pingliangtai individuals (Pingliangtai_LN) fall into the East Asian gene pool and clustered with other Late Neolithic Longshan culture individuals in the Yellow River basin (YR_LN) in the literature (Figure 3A) (Ning et al., 2020; Yang et al., 2020), consistent with our observation in the outgroup *f*3-statistics that Pingliangtai individuals share the highest affinity with YR_LN (Figure 3B), suggesting genetic homogeneity for the regional Late Neolithic Longshan populations. Compared with the preceding Middle Neolithic Yangshao individuals from the Yellow River basin (YR_MN), the Pingliangtai individuals are shifted toward populations from southern China in the PCA plot (Figure 3A). Similar genetic patterns were also observed in the unsupervised model-based ADMIXTURE analysis that Pingliangtai individuals share a similar genetic profile with the YR_LN and harbor a larger proportion of the green component than YR_MN that is maximized in the indigenous populations in Taiwan (e.g., Ami and Atayal) and in other populations from southern China (Figure 3C). This result is further supported by the symmetric *f*4-statistics of the form *f*4(Mbuti, world-wide; YR_LN/Pingliangtai_LN, YR_MN), when compared to the YR_MN groups, both later YR_LN and Pingliangtai_LN individuals showed a significant genetic affinity with populations from south China and Southeast Asia, such as Ami, Dai, etc. ($|Z| > 3$) (see Figure S3), documenting extensive genetic contribution into the YR from further south, echoing the previous finding in (Ning et al., 2020).

We also compared different ancient YR individuals as well as other present-day populations in China to test whether they share a common genetic substratum. As expected, we observed significant genetic affinity of

present-day populations in China such as Han, Naxi, Lahu, Yi, Tibetan, and Tujia who are all Sino-Tibetan speakers as well as ancient populations from the northeastern part of the Tibetan Plateau (Qijia individuals) (Ning et al., 2020) and from Nepal (Jeong et al., 2016) to the Neolithic YR populations (see Figure S4). As a result, all the present-day Sino-Tibetan speakers can be modeled as possessing a majority of ancestry (35.1%–86.7%) from Neolithic YR populations (see Data S1F), which is compatible with the Northern China origin of Sino-Tibetan languages coinciding with the expansion of Neolithic millet farmers from the YR (Zhang et al., 2019a, 2019b; Sagart et al., 2019).

Three Pingliangtai individuals share second-degree relatedness (SDR) to each other

To estimate the genetic relatedness of the four Pingliangtai individuals at a finer scale, we determined the degree of genetic relatedness between individuals from autosomes. We first calculated pairwise mismatch rates (PMRs) from haploid genotypes of the "1240k" panel (Mathieson et al., 2015), but further excluded the SNPs on the X and Y chromosomes. This approach calculated the rates of alleles mismatching for each individual pair to infer the degree of relatedness between individual pairs (see STAR Methods for details). All six pairs in our study had more than 40,000 overlapping SNPs on the autosomes (see Data S1B), and the sufficient data make the results of pair-based PMR quite accurate. As a result, the PMR values for the four individuals range from 0.19 to 0.24 (see Data S1B), and two major clusters are observed (see Figure S5A). The first cluster includes the comparison of M313 to the other three Pingliangtai individuals (M313-M310, M313-M311, M313-M312), for which a high PMR value (0.237–0.241) was obtained, suggesting that M313 shares no close relatedness with the other individuals, a finding in agreement with the observation that the mtDNA of M313 belonged to a different haplogroup from the other Pingliangtai individuals. The second group includes three individual pairs (M310-M311, M310-M312, M312-M313), and the PMR values range from 0.19268 to 0.2126, roughly 7/8 of the baseline value (unrelated), suggesting that the three individuals share SDR to each other (Jeong et al., 2018; Ning et al., 2019).

Second, we implemented READ to further affirm the genetic kinship between the four Pinglinagtai individuals. Since this method includes a step to normalize the data, which requires additional data for individuals with explicit relationships to normalize the P_0 (the proportion of mismatching alleles in each non-overlapping 1Mbps window), we included all published ancient genomes from five distinct Neolithic sites in the same region into the analysis, and a total of 120 pairs of comparisons were characterized. The results show that the normalized P_0 between M310, M311, M312 ranges from 0.8242 to 0.9011, and those of the unrelated individuals are from 1.0084 to 1.0286, with a small standard error of ~0.005 (see Figure S5B; Data S1C). Such a result echoes the PMR analysis in that M310, M311, M312 are estimated as harboring a SDR to each other. Third, we employed lcMLkin, a method that uses information from genotype likelihoods rather than observed genotypes in a maximum likelihood framework to estimate the overall coefficient of relatedness as well as individual genetic kinship components between pairs of individuals. To have an adequate amount of individuals to estimate allele frequencies, we combined the Pingliangtai samples with other available ancient genomic data from Central and North China (Ning et al., 2020). For absolutely unrelated individuals, we would expect $k_0 = 1$ (the probability that two diploid individuals share 0 alleles). Our results show that the k_0 values for M310, M311, and M312 range from 0.497 to 0.673 (see Data S1D), representing second-degree relatives between each other. For the other individuals, including M313 from this study, we observed a high k_0 value (close to 1), documenting that they are genetically unrelated to each other. By plotting the coefficient of relatedness (r) against k_0 of up to the third degree of relatedness within the resolution of our data, we were able to directly visualize the relatedness between the Pingliangtai individuals. M310, M311, and M312 fall within the range of second-degree relationships, and M313 (the red dot at the bottom) and the other three individuals are identified as unrelated to the others (Figure 4), such a conclusion was further confirmed by NgsRelate2 (10.1093/gigascience/giz034) (Hanghøj et al., 2019) (see Data S1E), a method that follows the same logic as LcMLkin. In conclusion, all the above analyses consistently support that M310, M311, and M312 share a SDR with each other.

Parental relatedness in the Pingliangtai individuals

ROH are contiguous regions lacking variation in the genome, the length of these long stretches of DNA segments potentially reflecting the pedigree inbreeding. When long ROH arise, the only plausible explanation is that the person inherited two copies of the genome from their parents who are closely genetically related. In order to identify if the Pingliangtai individuals were descended from some degree of recent inbreeding, we estimated the ROH of selected genomes in our study following an approach implemented in Ringbauer et al. (Ringbauer et al., 2020)

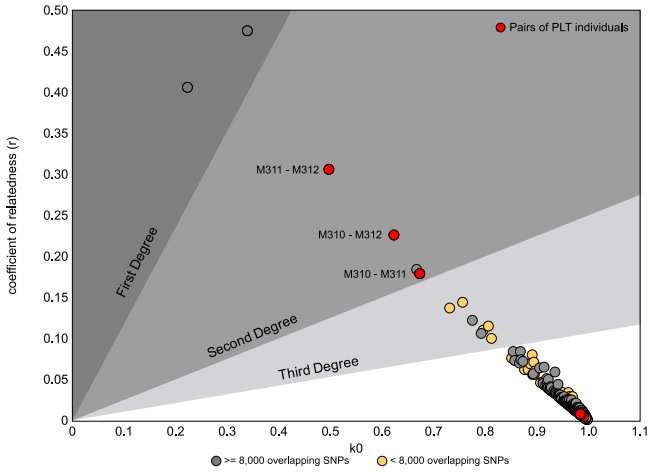


Figure 4. Pingliangtai genetic kinship analysis based on lcMLkin

K_0 represents the probability of two individuals sharing zero alleles by IBD, and the coefficient of relatedness (r) is related to $k_1/2 + k_2$. A smaller k_0 and larger R value indicate a close relatedness. Only red plots represent genetic kinship statistics of Pingliangtai individual pairs, the gray and light orange plots represent genetic kinship statistics among other published ancient Chinese individuals.

Our results show that three out of four Pingliangtai individuals (M311, M312, M313) do not have any detected ROH segments longer than 4 centimorgans (cM), but M310 who is SDR with both M311 and M312, has a sum of 109cM long ROH segments (>20cM) in her genome, approximal to a degree of offspring of second cousins (~135cM is the average value for blocks of the expected offspring from first cousins) (Figure 5B). We next applied this method to all the available ancient genomic data from the Central Plain area of China (Ning et al., 2020; Yang et al., 2020) to further explore whether the inbreeding event was detected at Pingliangtai is a common pattern in ancient China. We found that 20 out of 33 individuals screened did not even have any detected ROH longer than 4 cM (see Data S1G) pointing to a large local population size persisting over a range of 3000 years. Only one individual dating to the Han period (ca. 2000 BP) from central China was characterized as an offspring of a pair of parents who shared a third degree or more distant relatedness (1/33) (see Figure 5A; Data S1G), suggesting that endogamy was limited in prehistoric China and providing direct evidence that singular consanguineous mating events could be detected in both Neolithic and Iron Age China.

DISCUSSION

Our study utilizes ancient genomic data to reconstruct mating strategy and underlying social organizations for a prehistoric community in China. Three types of genetic markers were utilized in this study, namely the Y chromosome and mtDNA, which are uniparentally inherited and are widely used to track paternal and maternal lines, as well as autosome DNA that can reflect a more detailed population admixture history, genetic kinship relations, and parental relatedness (Ringbauer et al., 2020). Three (M310, M311, and M312) out of four individuals from the Pingliangtai site were identified as second-degree relatives. Genetically speaking, SDR is individuals who share a quarter of their genes with each other. This includes relationships such as grandparent–grandchild, uncle/aunt–nephew/niece, half-siblings, and double cousins. Specifically, M310, M311, and M312 share identical mitochondrial sequences and belong to the mtDNA haplogroup D4b1a, which has not been observed in other ancient individuals from China (Ning et al., 2020; Yang et al., 2020; Wu et al., 2019; Zhao et al., 2010), but has a wide distribution in present-day East Asians (Derenko et al., 2010). This suggests that the three individuals were probably descended from the same maternal line, and even from a common mother. The Y chromosome evidence shows that M311 and M312 share the same set of specific novel SNPs under the Y chromosomal haplogroup N1b2* showing that they are paternally related, although we cannot formally claim that they were descended from the same paternal line because of the low coverage and limited informative markers available (1.92x and 0.3x for M311 and M312, respectively). According to the above relatedness and the inbreeding signals, we detected in the M310, several most probable parsimony family trees were reconstructed (see Figures 5C and S6). However, if we assume that M311 and M312 share the same paternal line and they both share the same maternal line with M310, then the only family tree topologies that can explain the characterized

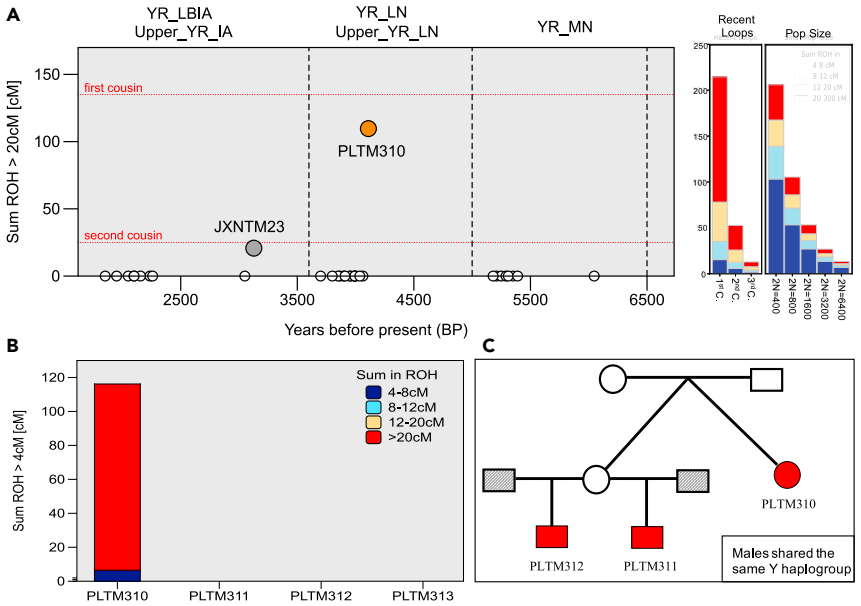


Figure 5. Run of homozygosity of ancient individuals from Central Plain of China

(A) ROH of 33 ancient genomes dating between the Middle Neolithic and the Late Iron Age of the Yellow River region. Each dot represents one individual; the orange dot with elevated ROH represents PLTM310 in this study, the gray dot displays the JXNTM23 individual from the Bronze Age Jiaozuoniecun site.

(B) ROH results of four Late Neolithic Pingliangtai individuals.

(C) The most plausible family tree of the three Pingliangtai individuals. In this scenario, PLTM311, PLTM312, and PLTM310 share the same mtDNA haplogroup, PLTM311, and PLTM312 carry the same Y haplogroup, and share relatively higher relatedness than with M310.

relatedness is that at least two of the three individuals were descended from a sororate/levirate mating. Levirate/sororate mating is a custom in which a brother or sister inherits a sibling's mating, sometimes in order to produce descendants and to continue a family line (Ebrey and Walthall, 2013). In China, this custom was first attested in historical documents (e.g. the *Tso Chuan* 左传, a commentary on the fifth century BCE Spring and Autumn Annals) in the Late Bronze Age Eastern Zhou dynasty as *Ying* (媵), involving the mating of two sisters to one noble in upper-class society. In such cases, M311 and M312 do not show a strict SDR since either their mothers or fathers are genetically related to each other. Thus, M311 and M312 should share relatively higher relatedness than with M310. This agrees with and is confirmed by the observation that M311 and M312 share the highest genetic relatedness with each other than with the other combinations, although they both share a SDR with M310 (see Figures 4, S5, and S6A).

Taking the shared second-degree genetic kinship for the Longshan-period juveniles who came from the same household before their death, we provide clear DNA-based evidence that extended family beyond the nuclear family as proposed by (Pearson, 1981, 1988), served as the basic household unit in Longshan society. This observation is in agreement with the archaeological findings that Longshan is a crucial Neolithic culture in East Asia, which was proceeded by the Yangshao culture and followed by the early Bronze Age Erlitou culture. The Erlitou culture was usually considered as the basis of the first Chinese dynasty, referred to as Xia (Liu, 2005). Compared with the preceding Yangshao culture, the Longshan culture manifests a process of social change from a more egalitarian to a stratified society where town walls were built, violence, and warfare were widespread, and there was a transition from regional units to settlements with defined social or political hierarchies (Liu, 1996a; Underhill, 1994). Our results suggested that the social units of Longshan society are different from either the clan-based Yangshao culture or the patriarchal family in the latter Chinese dynasties known from historical records, and are in a transitional stage from a clan-based to a family-based society.

Starting from at least 6,000 BCE, the Central Plains of China was one of the earliest centers in the world where millet (foxtail millet, *Setaria italica* and broomcorn millet, *Panicum miliaceum*) were first cultivated and domesticated (Lu et al., 2009; Yang et al., 2012). The shift from hunting and gathering subsistence

to cereal agriculture enabled a rapid growth of population size (Yan et al., 2014). By the Longshan period, intensified millet and rice agriculture had developed, and compared to the preceding Yangshao culture in the Central Plains, the population size had increased significantly, as shown by the higher density of Longshan sites (Liu, 2005). This is consistent with our genetic results, which point to the ancient YR people as a genetically stable community that could have dense populations (Figure 5B). Although increasing local population size tends to decrease the length of ROH, one individual from the Pingliangtai site harbors long ROH fragments that were similar in degree to that expected for a descendant of first cousins, providing direct evidence of consanguineous mating in Longshan society roughly 4,000 years ago. Mating between close relatives is documented in many societies and was prevalent among European royal/elite families (Cassidy et al., 2020; Bittles and Black, 2010). To keep high social status and to establish strong political alliances, royal family members cannot marry commoners and the only viable option is to marry their relatives. For example, the kings of the Spanish Habsburg dynasty (1516–1700 CE) frequently married close relatives producing numerous uncle-niece, first cousin, and other close consanguineous matings (Ceballos et al., 2018; Alvarez et al., 2009). A similar mating strategy is also well-documented in Chinese historical records from the Han Dynasty (200 BCE–200 CE), when according to the Shiji and Hanshu (*史记; 汉书*) royal family members married their relatives to strengthen the political power of the royal line. Here we show that consanguineous mating was already taking place in Longshan society 4,000 years ago.

The Longshan period (ca. 2,500–1,800 BCE) was a time of major cultural and demographic change in prehistoric China, and the kinship, social organization, and mating practice of this time have been a major interest for historians, archaeologists, and anthropologists. Through multidisciplinary research, we are able to reconstruct aspects of the genetic kinship, mating strategy, and underlying social organization for this prehistoric society. Here we provide direct evidence that consanguineous mating had been practiced in Late Neolithic Longshan society, around 2,000 years earlier than its attestation in the historical record in China. In addition, by characterizing the genetic kin relations of individuals from the same household, we provide an explicit indication that extended family beyond the nuclear family served as a basic household in Longshan society and that genetic kinship still acted as a major focus of social organization during the Longshan period. It must be emphasized that the individuals in this study came from a single site; further such studies on a larger sample size from various regions and cemeteries will provide more detailed knowledge about the mating customs, burial patterns, and social organization of Longshan society.

Limitations of study

In our study, we explored genome-wide data of four ancient individuals from the Pingliangtai site in China, the sample size is relatively limited, and the current analyses are only limited to the "1240K" touchdown alleles that only allows for accurately estimating the degree of genetic kinship up to the third-degree relatives. We note that further studies with more samples from various archaeological sites during the Late Neolithic and especially high-coverage ancient genomes to obtain a more comprehensive understanding the social organization of the central plain area and the whole of China.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHODS DETAILS
  - Ancient DNA extraction and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Genomic data processing
  - Ancient DNA authentication
  - Genetic sexing and uniparental haplogroup assignment
  - Related databases and population genetic structure
  - Run of homozygosity analysis
  - Genetic relatedness analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103352.

ACKNOWLEDGMENTS

We are grateful to Henan Provincial Institute of Cultural Heritage and Archaeology for providing materials for analysis. This work was supported by the major project of Humanities and Social Sciences Key Research Base of the Ministry of Education (Grant No.16JJD780005), the National Natural Science Foundation of China (Grant No. 42072018 and 41971251); Science and Technology Development Program of Jilin Province (20200901030SF); British Academy, Newton Advanced Fellowship (Grant No. NAFR2180079); C.N. and Mark Hudson were supported by the European Research Council (ERC) under the European Union's Horizon 2010 research and innovation program (grant agreement No 646612) granted to M.R..

AUTHOR CONTRIBUTIONS

Y.C. conceived and supervised this study. F.Z., T.L., S.G., Y.X., and C.L. performed the laboratory works. H.Z., Y.C., L.Q., W.L., S.Z, and C.L. provided archaeological materials and associated information. F.Z., C.N., and P.M. analyzed data. C.N., F.Z., M.H., and H.Z. wrote the manuscript with input from all co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 12, 2021
Revised: August 12, 2021
Accepted: October 22, 2021
Published: November 19, 2021

REFERENCES

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109.

Alvarez, G., Ceballos, F.C., and Quinteiro, C. (2009). The role of inbreeding in the extinction of a European royal dynasty. PLoS One *4*, e5174. https://doi.org/10.1371/journal.pone.0005174.

Amorim, C.E.G., Vai, S., Posth, C., Modi, A., Koncz, I., Hakenbeck, S., La Rocca, M.C., Mende, B., Bobo, D., Pohl, W., et al. (2018). Understanding 6th-century barbarian social organization and migration through paleogenomics. Nat. Commun. *9*, 3547. https://doi.org/10.1038/s41467-018-06024-4.

Binford, L.R. (1962). Archaeology as anthropology. Am. Antiq. *28*, 217–225. https://doi.org/10.2307/278380.

Bittles, A.H., and Black, M.L. (2010). Consanguinity, human evolution, and complex diseases. PNAS *107*, 1779–1786. https://doi.org/10.1073/pnas.0906079106.

Carr, C. (1995). Mortuary practices: Their social, philosophical-religious, circumstantial, and physical determinants. J. Archaeol Method Theor. *2*, 105–200. https://doi.org/10.1007/BF02228990.

Cassidy, L.M., Maoldúin, R.Ó., Kador, T., Lynch, A., Jones, C., Woodman, P.C., Murphy, E., Ramsey, G., Dowd, M., Noonan, A., et al. (2020). A dynatic elite in monumental Neolothic society. Nature *582*, 384–388. https://doi.org/10.1038/s41586-020-2378-6.

Ceballos, F.C., Joshi, P.K., Clark, D.W., Ramsay, M., and Wilson, J.F. (2018). Runs of homozygosity: Windows into population history and trait architecture. Nat. Rev. Genet. *19*, 220–234. https://doi.org/10.1038/nrg.2017.109.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.

Chang, K. (1968). Archeology of ancient China. Science *162*, 519–526. https://doi.org/10.1126/science.162.3853.519.

Chen, X.-L., Fang, Y.-M., Hu, Y.-W., Hou, Y.-F., Lü, P., Yuan, J., Song, G.-D., Fuller, B.T., and Richards, M.P. (2016). Isotopic reconstruction of the late Longshan period (ca. 4200–3900 BP) dietary complexity before the onset of state-level societies at the Wadian site in the Ying River Valley, Central Plains, China. Int. J. Osteoarchaeol. *26*, 808–817. https://doi.org/10.1002/oa.2482.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., et al. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. PNAS *110*, 15758–15763. https://doi.org/10.1073/pnas.1314445110.

Derenko, M., Malyarchuk, B., Grzybowski, T., Denisova, G., Rogalla, U., Perkova, M., Dambueva, I., and Zakharov, I. (2010). Origin and post-glacial dispersal of mitochondrial DNA haplogroups C and D in northern Asia. PLoS One *5*. https://doi.org/10.1371/journal.pone.0015214.

Dong, Y., Lin, L., Zhu, X., Luan, F., and Underhill, A.P. (2019). Mortuary ritual and social identities during the late Dawenkou period in China. Antiquity *93*, 378–392. https://doi.org/10.15184/aqy.2019.34.

Dryomov, S.V., Nazhmidenova, A.M., Shalaurova, S.A., Morozov, I.V., Tabarev, A.V., Starikovskaya, E.B., and Sukernik, R.I. (2015). Mitochondrial genome diversity at the Bering Strait area highlights prehistoric human migrations from Siberia to northern North America. Eur. J. Hum. Genet. *23*, 1399–1404. https://doi.org/10.1038/ejhg.2014.286.

Ebrey, P.B., and Walthall, A. (2013). East Asia: A Cultural, Social, and Political History (Cengage Learning).

Engels, F. (1954). Origin of the Family, Private Property, and the State (Foreign Languages Publishing House).

Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., et al. (2016). The genetic history of Ice Age Europe. Nature *534*, 200–205. https://doi.org/10.1038/nature17993.

Gao, Q., and Lee, Y.K. (1993). A biological perspective on Yangshao Kinship. J. Anthropol. Archaeol. *12*, 266–298. https://doi.org/10.1006/jaar.1993.1008.

Goody, J. (1990). The Oriental, the Ancient and the Primitive: Systems of Marriage and the Family in the Pre-industrial Societies of Eurasia, Studies

in Literacy, the Family, Culture and the State (Cambridge University Press). https://doi.org/10.1017/CBO9780511621703.

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature *522*, 207–211. https://doi.org/10.1038/nature14317.

Hanghøj, K., Moltke, I., Andersen, P.A., Manica, A., and Korneliussen, T.S. (2019). Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. GigaScience *8*. https://doi.org/10.1093/gigascience/giz034.

Haniharai, K., Yamauchi, A., and Mizoguchi, Y. (1983). Statistical analysis on kinship among skeletal remains excavated from a neolithic site at Uwasato, Iwate Prefecture. J. Anthropol. Soc. Nippon *91*, 49–68. https://doi.org/10.1537/ase1911.91.49.

Henan Provincial Institute of Cultural Heritage, School of Archaeology and Museology of Peking University (2017). Excavation report for the Pingliangtai site of the Longshan period in Huaiyang, Henan province (in Chinese). Huaxia Kaogu, 3–15. https://doi.org/10.16143/j.cnki.1001-9928.2017.03.001.

Hill, J.N. (1970). Broken K Pueblo: Prehistoric Social Organization in the American Southwest (University of Arizona Press).

Jeong, C., Balanovsky, O., Lukianova, E., Kahbatkyzy, N., Flegontov, P., Zaporozhchenko, V., Immel, A., Wang, C.-C., Ixan, O., Khussainova, E., et al. (2019). The genetic history of admixture across inner Eurasia. Nat. Ecol. Evol. *3*, 966–976. https://doi.org/10.1038/s41559-019-0878-2.

Jeong, C., Ozga, A.T., Witonsky, D.B., Malmström, H., Edlund, H., Hofman, C.A., Hagan, R.W., Jakobsson, M., Lewis, C.M., Aldenderfer, M.S., et al. (2016). Long-term genetic stability and a high-altitude East Asian origin for the peoples of the high valleys of the Himalayan arc. Proc. Natl. Acad. Sci. U S A *113*, 7485–7490. https://doi.org/10.1073/pnas.1520844113.

Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A.S., Taylor, W.T.T., Hagan, R.W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., et al. (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. Proc. Natl. Acad. Sci. U S A *115*, E11248–E11255. https://doi.org/10.1073/pnas.1813608115.

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.

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.

Kennett, D.J., Plog, S., George, R.J., Culleton, B.J., Watson, A.S., Skoglund, P., Rohland, N., Mallick, S., Stewardson, K., Kistler, L., et al. (2017). Archaeogenomic evidence reveals prehistoric matrilineal dynasty. Nat. Commun. *8*, 1–9. https://doi.org/10.1038/ncomms14115.

Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., and Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques *59*, 87–93. https://doi.org/10.2144/000114320.

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. BMC Bioinformatics *15*, 356. https://doi.org/10.1186/s12859-014-0356-4.

Kuhn, J.M.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.

Kutanan, W., Kampuansai, J., Srikummool, M., Kangwanpong, D., Ghirotto, S., Brunelli, A., and Stoneking, M. (2017). Complete mitochondrial genomes of Thai and Lao populations indicate an ancient origin of Austroasiatic groups and demic diffusion in the spread of Tai–Kadai languages. Hum. Genet. *136*, 85–98. https://doi.org/10.1007/s00439-016-1742-y.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

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.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 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.

Li, L. (2000). Ancestor Worship: an archaeological investigation of ritual activities in Neolithic north China. J. East Asian Archaeol. *2*, 129–164. https://doi.org/10.1163/156852300509826.

Lipatov, M., Sanjeev, K., Patro, R., and Veeramah, K.R. (2015). Maximum likelihood estimation of biological relatedness from low coverage sequencing data. bioRxiv. https://doi.org/10.1101/023374.

Liu, L. (2005). The Chinese Neolithic: Trajectories to Early States, New Studies in Archaeology (Cambridge University Press).

Liu, L. (1996a). Settlement patterns, chiefdom variability, and the development of early states in North China. J. Anthropol. Archaeol. *15*, 237–288. https://doi.org/10.1006/jaar.1996.0010.

Liu, L. (1996b). Morturary ritual and social hierarchy in the Longshan culture. Early China *21*, 1–46.

Longacre, W.A. (1964). Archeology as anthropology: A case study. Science *144*, 1454–1455. https://doi.org/10.1126/science.144.3625.1454.

Lu, H., Zhang, J., Liu, K., Wu, N., Li, Y., Zhou, K., Ye, M., Zhang, T., Zhang, H., Yang, X., et al. (2009). Earliest domestication of common millet (Panicum miliaceum) in East Asia extended to 10,000 years ago. PNAS *106*, 7367–7372. https://doi.org/10.1073/pnas.0900158106.

Major, J.S., and Cook, C.A. (2016). Ancient China: A History (Taylor & Francis).

Marlowe, F.W. (2005). Hunter-gatherers and human evolution. Evol. Anthropol. *14*, 54–67. https://doi.org/10.1002/evan.20046.

Marlowe, F.W. (2004). Marital residence among foragers. Curr. Anthropol. *45*, 277–284. https://doi.org/10.1086/382256.

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503. https://doi.org/10.1038/nature16152.

Mittnik, A., Massy, K., Knipper, C., Wittenborn, F., Friedrich, R., Pfrengle, S., Burri, M., Carlichi-Witjes, N., Deeg, H., Furtwängler, A., et al. (2019). Kinship-based social inequality in Bronze age Europe. Science, eaax6219. https://doi.org/10.1126/science.aax6219.

Morgan, L.H. (1877). Ancient Society (World Publishing).

Ning, C., Li, T.J., Wang, K., Zhang, F., Li, T., Wu, X.Y., 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.

Ning, C., Wang, C.-C., Gao, S., Yang, Y., Zhang, X., Wu, X., Zhang, F., Nie, Z., Tang, Y., Robbeets, M., et al. (2019). Ancient genomes reveal Yamnaya-related ancestry and a potential source of Indo-European speakers in Iron age Tianshan. Curr. Biol. *29*, 2526–2532.e4. https://doi.org/10.1016/j.cub.2019.06.044.

O'Sullivan, N., Posth, C., Coia, V., Schuenemann, V.J., Price, T.D., Wahl, J., Pinhasi, R., Zink, A., Krause, J., and Maixner, F. (2018). Ancient genome-wide analyses infer kinship structure in an Early Medieval Alemannic graveyard. Sci. Adv. *4*, eaao1262. https://doi.org/10.1126/sciadv.aao1262.

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.

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.

Pearson, R. (1988). Chinese neolithic burial patterns: Problems of method and interpretation. Early China *13*, 1–45. https://doi.org/10.1017/S0362502800005198.

Pearson, R. (1981). Social complexity in Chinese coastal Neolithic sites. Science *213*, 1078–1086. https://doi.org/10.1126/science.213.4512.1078.

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., and Nieselt, K. (2016). EAGER:

Efficient ancient genome reconstruction. Genome Biol. *17*, 60. https://doi.org/10.1186/s13059-016-0918-z.

Price, T.D., Johnson, C.M., Ezzo, J.A., Ericson, J., and Burton, J.H. (1994). Residential mobility in the prehistoric southwest United States: A preliminary study using strontium isotope analysis. J. Archaeol. Sci. *21*, 315–330. https://doi.org/10.1006/jasc.1994.1031.

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.

Ringbauer, H., Novembre, J., and Steinrücken, M. (2020). Human parental relatedness through time - detecting runs of homozygosity in ancient DNA. bioRxiv. https://doi.org/10.1101/2020.05.31.126912.

Sagart, L., Jacques, G., Lai, Y., Ryder, R.J., Thouzeau, V., Greenhill, S.J., and List, J.-M. (2019). Dated language phylogenies shed light on the ancestry of Sino-Tibetan. PNAS *116*, 10317–10322. https://doi.org/10.1073/pnas.1817972116.

Sánchez-Quinto, F., Malmström, H., Fraser, M., Girdland-Flink, L., Svensson, E.M., Simões, L.G., George, R., Hollfelder, N., Burenhult, G., Noble, G., et al. (2019). Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. PNAS *116*, 9469–9474. https://doi.org/10.1073/pnas.1818037116.

Schroeder, H., Margaryan, A., Szmyt, M., Theulot, B., Włodarczak, P., Rasmussen, S., Gopalakrishnan, S., Szczepanek, A., Konopka, T., Jensen, T.Z.T., et al. (2019). Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. PNAS *116*, 10705–10710. https://doi.org/10.1073/pnas.1820210116.

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.

Sun, L., Pechenkina, K., Cao, Y., Zhang, H., and Qi, X. (2019). Cases of endocranial lesions on juvenile skeletons from Longshan cultural sites in Henan Province, China. Int. J. Paleopathol. *26*, 61–74. https://doi.org/10.1016/j.ijpp.2019.06.001.

Underhill, A. (2000). An analysis of mortuary ritual at the Dawenkou site, Shandong, China. J. East Asian Archaeol. *2*, 93–127. https://doi.org/10.1163/156852300509817.

Underhill, A.P. (1994). Variation in settlements during the Longshan period of northern China. Asian Perspect. *33*, 197–228.

Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönherr, 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.

Wen, S.-Q., Yao, H.-B., Du, P.-X., Wei, L.-H., Tong, X.-Z., Wang, L.-X., Wang, C.-C., Zhou, B.-Y., Shi, M.-S., Zhabagin, M., et al. (2019). Molecular genealogy of Tusi Lu's family reveals their paternal relationship with Jochi, Genghis Khan's eldest son. J. Hum. Genet. *64*, 815–820. https://doi.org/10.1038/s10038-019-0618-0.

Wu, X., Ning, C., Bao, Q., Gao, S., Zhang, F., Wu, S., Li, T., Fan, L., Li, T., Yang, X., et al. (2019). Mitochondrial genome of an 8,400-year-old individual from northern China reveals a novel subclade under C5d. Hum. Biol. *91*, 21–30. https://doi.org/10.13110/humanbiology.91.1.04.

Yan, S., Wang, C.-C., Zheng, H.-X., Wang, W., Qin, Z.-D., Wei, L.-H., Wang, Y., Pan, X.-D., Fu, W.-Q., He, Y.-G., et al. (2014). Y chromosomes of 40% Chinese descend from three neolithic super-grandfathers. PLoS One *9*, e105691. https://doi.org/10.1371/journal.pone.0105691.

Yang, B., and Xu, G. (1985). Archaeological excavation report on the Hougang site in Anyang, 1979 (in Chinese). Acta Archaeol. Sin. *1*, 137–148.

Yang, M.A., Fan, X., Sun, B., Chen, C., Lang, J., Ko, Y.-C., Tsang, C., Chiu, H., Wang, T., Bao, Q., et al. (2020). Ancient DNA indicates human population shifts and admixture in Northern and Southern China. Science. https://doi.org/10.1126/science.aba0909.

Yang, X., Wan, Z., Perry, L., Lu, H., Wang, Q., Zhao, C., Li, J., Xie, F., Yu, J., Cui, T., et al. (2012). Early millet use in Northern China. PNAS *109*, 3726–3730. https://doi.org/10.1073/pnas.1115430109.

Zhang, C., Pollard, A.M., Rawson, J., Huan, L., Liu, R., and Tang, X. (2019a). China's major Late Neolithic centres and the rise of Erlitou. Antiquity *93*, 588–603. https://doi.org/10.15184/aqy.2019.63.

Zhang, M., Yan, S., Pan, W., and Jin, L. (2019b). Phylogenetic evidence for Sino-Tibetan origin in Northern China in the late neolithic. Nature *569*, 112–115. https://doi.org/10.1038/s41586-019-1153-z.

Zhao, Y.-B., Li, H.-J., Cai, D.-W., Li, C.-X., Zhang, Q.-C., Zhu, H., and Zhou, H. (2010). Ancient DNA from nomads in 2500-year-old archeological sites of Pengyang, China. J. Hum. Genet. *55*, 215–218. https://doi.org/10.1038/jhg.2010.8.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Ancient human remains	This paper	PLTM310
Ancient human remains	This paper	PLTM311
Ancient human remains	This paper	PLTM312
Ancient human remains	This paper	PLTM313
Chemicals, peptides, and recombinant proteins		
T4 Polynucleotide Kinase	New England Biolabs	M0236L
T4 DNA Polymerase	New England Biolabs	M0203L
Bst DNA Polymerase	New England Biolabs	M0275S
Q5 High-Fidelity DNA Polymerase	New England Biolabs	M0491S
BSA 20mg/mL	New England Biolabs	B9000S
1x Tris-EDTA pH 8.0	AppliChem	A8569
0.5M EDTA PH 8.0	AppliChem	A4892
20% SDS Solution	Serva	39575.01
5 M NaCl	Sigma Aldrich	S5150
Guanidine hydrochloride	Sigma Aldrich	50933
Proteinase K	Sigma Aldrich	P6556
Isopropanol	Sigma Aldrich	67-63-0
Sodium acetate	Sigma Aldrich	S2889
1M Tris-HCL, PH 8.0	Sigma Aldrich	AM9856
Tween-20	Sigma Aldrich	P9416
1M NaOH	Sigma Aldrich	71463
dNTP Mix	Thermo	R1121
ATP	Thermo	R0441
T4 DNA Ligase	Thermo	EL0011
Agencourt AMPure XP beads (60ml)	Beckman Coulter	A63881
Critical commercial assays		
Min Elute PCR Purification Kit	QIAGEN	28006
Quick Ligation ™ Kit	New England Biolabs	M2200
Deposited data		
BAM files are available in the BIG Data Center Genome Sequence Archive (http://bigd.big.ac.cn/gsa-human/)	This paper	HRA000253
Software and algorithms		
AdapterRemoval v2.2.0	Schubert et al., 2016	https://github.com/MikkelSchubert/adapterremoval; RRID:SCR_011834
BWA v0.7.12	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/; RRID: SCR_010910
samtools v1.3.165	Li and Durbin, 2010	http://samtools.sourceforge.net/; RRID:SCR_002105
DeDup v0.12.2	Peltzer et al., 2016	https://github.com/ apeltzer/DeDup
pileupCaller	https://github.com/stschiff/sequenceTools	https://github.com/stschiff/sequenceTools

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
mapDamage v2.0.6	Jónsson et al., 2013	https://ginolhac.github.io/mapDamage/; RRID:SCR_001240
ANGSD v0.910	Korneliussen et al., 2014	http://www.popgen.dk/angsd/index.php/ANGSD
Schmutzi	Renaud et al., 2015	https://github.com/grenaud/schmutzi
Geneious v11.1.3	https://www.geneious.com/	https://www.geneious.com/
HaploGrep2	Weissensteiner et al., 2016	https://haplogrep.uibk.ac.at/index.html
EIGENSOFT v6.1.4	Patterson et al., 2006	https://github.com/DReichLab/EIG; RRID:SCR_004965
ADMIXTURE	Alexander et al., 2009	http://dalexander.github.io/admixture/download.html; RRID:SCR_001263
hapROH	Ringbauer et al., 2020	https://pypi.org/project/hapROH/
READ	Kuhn et al., 2018	https://bitbucket.org/tguenther/read/src/master/
lcMLkin	Lipatov et al., 2015	https://github.com/COMBINE-lab/maximum-likelihoodrelatedness-estimation
PMR	Kennett et al., 2017	Kennett et al., 2017
NgsRelate	Hanghøj et al., 2019	https://github.com/ANGSD/NgsRelate

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yinqiu Cui (cuiyq@jlu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The accession number for all newly reported sequencing data reported in this paper are available from the BIG Data Center Genome Sequence Archive for Human: HRA000253. The basemap used in Figure 1 is in the public domain and accessible through the Natural Earth website (https://www.naturalearthdata.com/downloads/10m-raster-data/).

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The Pingliangtai site is a square-walled town first discovered and excavated in the late 1970s (Chang, 1968). The skeletal samples analysed here were discovered during excavations carried out by the School of Archaeology and Museology, Peking University and Henan Provincial Institute of Cultural Heritage and Archaeology in 2016 and 2019. No skeletal trauma is visible on the four analysed individuals. Scientific investigations of these remains were approved by the School of Archaeology and Museology, Peking University and Henan Provincial Institute of Cultural Heritage and Archaeology, which holds the custodianship of the studied remains. This study is based on previously excavated archaeological remains and included no new excavation effort not study of live human or animal subjects.

METHODS DETAILS

Ancient DNA extraction and sequencing

We first extracted ancient DNA from the petrous bones in a dedicated clean facility specially designed for ancient DNA studies at Jilin University, following a strict protocols. In brief, this method involves the following steps: 1) 50 mg of bone powder were incubated for 12-16 h (37°C) in 1ml of extraction buffer consisting of 0.5M EDTA (pH 8.0) and 0.25 mg/ml proteinase K; 2) DNA released from powder is bound to silica, which is added as a suspension together with a binding buffer containing guanidine hydrochloride, Sodium acetate (pH 5.2), and isopropanol, in combination with the Min Elute PCR Purification Kit; 3) the DNA is eluted into a 100ml low-salt TET buffer containing 1M Tris-HCl, pH 8.0, 0.5mM EDTA, pH8.0, and Tween-20 (Dabney et al., 2013; Korlević et al., 2015). Then, a double-stranded library was prepared per extract (Dabney et al., 2013) and followed by a purification step using AMPure XP bead (*Beckman Coulter Ltd*). The libraries were then sequenced on an Illumina HiSeq X10 platform.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genomic data processing

The raw sequencing outputs were demultiplexed by allowing one mismatch in each of the two 8-bp index sequences. Data came from the same libraries, but sequences in different batches were first merged into one. Then we clipped the Illumina sequencing adaptors by AdapterRemoval v2.2.0 (Schubert et al., 2016) and mapped the reads to the human reference genome (hs37d5) using BWA v0.7.12 (Li and Durbin, 2010). Next, we removed the PCR duplications by DeDup v0.12.2 (Peltzer et al., 2016). To minimize the post-mortem DNA damage bias, we trimmed the aligned reads by soft-masking ten bp on both 5′ and 3′ read ends implemented by bamUtils v1.0.13 (Jun et al., 2015). For the SNPs in the "1240K" panel (Haak et al., 2015), we randomly sampled a single high-quality base (Phred-scaled base quality score 30 or higher) as pseudo-haploid genotypes using the pileupCaller program (https://github.com/stschiff/sequenceTools).

Ancient DNA authentication

We used multiple methods to assess the quality and authentication of our ancient genomic data. First, we tabulated patterns of post-mortem chemical modification expected for ancient DNA using mapDamage v2.0.6 (Jónsson et al., 2013). Second, we estimated mitochondrial contamination rates for all individuals using Schmutzi (Renaud et al., 2015), an approach that calls endogenous DNA based on the deamination patterns and computes the contamination rate by comparison to a set of known contaminants. Third, we measured the nuclear genome contamination rate in males based on X chromosome data, as implemented in ANGSD v0.910 (Korneliussen et al., 2014). Since males have only a single copy of the X chromosome, mismatches between bases aligned to the same polymorphic position that are beyond the sequencing errors are considered as evidence of contamination. First, all ancient samples showed damage patterns that are characteristic of ancient DNA (see Figure S1). Second, all individuals yield low mitochondrial contamination of less than 4% (Table 1). Third, the X chromosome contaminations of the two males (M311 and M312) were also minimal (< 3%). All the above analyses show minimum modern human contamination and confirm the authentication of our ancient DNA data.

Genetic sexing and uniparental haplogroup assignment

We determined the genetic sex of the newly reported samples in this study by counting the number of reads overlapping with the '1240k' panel (Mathieson et al., 2015; Haak et al., 2015). We extracted high plausibly bases, and mapping quality sequencing reads (samtools depth -q30 -Q37) using samtools v1.3.165 (Li et al., 2009). We calculated the ratios of the numbers of reads mapped on to the X or Y chromosome compared with that mapped to the autosomes (X-rate and Y-rate, respectively). Samples with an X-rate < 0.42 and a Y-rate > 0.26 were assigned as males and those with an X-rate > 0.68 and a Y-rate < 0.02 were assigned as females. To determine the mtDNA haplogroup, we first aligned the adapter trimmed reads to the revised Cambridge Reference Sequence (rCRS; NC_012920.1) and removed low-quality sequences (-q30). Next, we generated the mtDNA consensus sequences of our ancient individuals using the Geneious v11.1.3 (https://www.geneious.com/) and then assigned their mtDNA haplogroups using HaploGrep2 (Weissensteiner et al., 2016). We determined the male Y chromosome haplogroup by examining a set of positions on the 25,660 diagnostic positions on the ISOGG database (https://isogg.org/) and assigned the final haplogroups by the most downstream derived SNPs.

Related databases and population genetic structure

We merged our ancient individuals to two sets of world-wide genotype panels that are widely used in ancient genomic studies, one based on the Affymetrix HumanOrigins Axiom Genome-wide Human Origins 1 array ('HumanOrigins'; 593,124 autosomal SNPs) (Jeong et al., 2019) and the other '1240k' panel covering 1,233,013 autosomal SNPs (Jeong et al., 2018; Ning et al., 2019). A Principal Component Analysis (PCA) was implemented in the smartpca program in Eigensoft v6.1.4 package (Patterson et al., 2006) with default parameters, and Shrinkmode: YES and lsqproject: YES options. The ADMIXTURE analysis was performed from K=2 to K=20 with Linkage Disequilibrium (LD) pruned by using PLINK with the parameter indep-pairwise 20 5 0.2 (Alexander et al., 2009; Chang et al., 2015) , and we chose K = 3 as it gave the lowest cross-validation (CV) error (see Figure S7). The outgroup *f*3-statistics were calculated using the qp3Pop (v435) programs in the ADMIXTOOLS v5.1 package using default parameters (Patterson et al., 2012). We carried out PCA and ADMIXTURE analyses against the 'HumanOrigins' dataset mainly because it harbors more present-day worldwide populations than the '1240k' panel. We used *qpAdm* v.810 from ADMIXTOOLS with the option 'allsnps: NO' to identify the most likely sources of ancestry and proportions of ancestry for the ancient populations from East Asia, as well as present-day Sino-Tibetan speakers groups (Haak et al., 2015).

Run of homozygosity analysis

We detected Runs of homozygosity (ROH) blocks (Ringbauer et al., 2020) of our ancient individuals using the python package *hapROH* (https://pypi.org/project/hapROH/) with default settings, and the hapROH method was specially designed for low-coverage ancient genomes (>0.5x). First, we prepared pseudo-haploid data on 1240k SNPs from the Eigenstrat files generated by pileupCaller program. Next, we ran hapROH and extracted results from the output CSV files. Finally, we visualized the results using DataGraph v4.5.1 and plot functions in the hapROH package. Here, we reported the total sum of ROH > 4, >8, >12, >20 cM for Pingliangtai individuals and also presented the sum of long ROH segments (more than 20cM) of 33 ancient genomes from Central Plain.

Genetic relatedness analysis

Because of the high missing rate of ancient DNA data, we applied four different methods to calculate the genetic kinship of the Pingliangtai individuals. 1) Pairwise mismatch rate (Kennett et al., 2017) (PMR) : the PMR approach estimates genetic kinship by calculating the pairwise mismatch rate of haploid genotypes across autosomal SNPs. For each pair of individuals, we defined the PMR value by dividing the number of SNP sites for which two individuals have different alleles sampled by the total number of sites covered in both individuals. In general, the PMR of the identical individuals (r = 1) should be half of that between the unrelated individuals (r = 0, identified as the population baseline, no inbreeding). Likewise, the PMR for first- (r = 0.5) and second-degree relatives (r = 0.25) should be 3/4 and 7/8 of the baseline. 2) READ (Kuhn et al., 2018) (Relationship Estimation from Ancient DNA): READ calculates the nonidentical allele ratio between the samples in each non-overlapping 1 Mb segments in the genome (P_0), moreover, the P_0 need to be normalized by the expected value for a randomly chosen pair of unrelated individuals from a similar population to exclude the interference of the diversity within the target population. Lower P_0 values mean more shared chromosomal segments. 3) lcMLkin (Lipatov et al., 2015): the method utilizes genotype likelihoods through a maximum likelihood approach and calculates the coefficient of relatedness (r), which is defined by the proportion of the genome common to two individuals due to direct genetic kinship. The method defines k_0, k_1, k_2 as the probability that no, one or two alleles were shared by identical by descent (IBD) between the pair of individuals. The coefficient of relatedness (r) can be related to IBD through $r = k_1/2 + k_2$. By comparing r and k_0, we can infer the relatedness to at least third degree for low coverage genome. The approach can clearly distinguish between sibling-sibling and parent-offspring relationships. For example, for a parent-offspring pair, k_0 is expected to be 0 (with $k_1 = 1$ and $k_2 = 0$), while k_0 should be 1/4 (with $k_1 = 1/2$ and $k_2 = 1/4$) between siblings depending on the rate of recombination, although their relatedness (r) are all 0.5. 4) NgsRelate: similar to lcMLkin, the NgsRelate software (version 2) infers the genetic relatedness on low-coverage genomic data using the genotype likelihood estimates than genotype calls. Besides, this method can accurately estimates genetic relatedness between pairs of inbred-individuals as well as the inbreeding coefficient. As opposed to NgsRelate, lcMLkin assumes that the individuals themselves are not inbred. In order to distinguish different pedigree relationships among the putative related individual pairs, we estimated the genetic kinship from both the autosomal and X chromosome data, however, due to the low coverage on the X chromosomes, all the four methods described in the above section failed to estimate the X chromosome genetic kinship. Overall, the PMR and READ methods were implemented on the pseudo-haploid genotype data, while, the lcMLkin and NgsRelate2 methods were implemented on the VCF files from indexed BAM files.