

Supplementary Note 1: Site and specimen description

Supplementary Note 2: Materials and Methods

Supplementary Table:

Table S1. Summary of metadata for each ancient individual sequenced in this study.

Table S2. Individuals with kinship relationships, with the individual retained for downstream population genetic analysis shown under ‘Treatment’.

Table S3. Number of bootstrap replicates (n=1,000) grouping the set of ancient populations indicated by a *.

Table S4. (a) Z-scores for f4 (Bianbian/Boshan/Xiaogao/SD9K/aYR/aLR/aSC, Xiaojingshan; aAR/FE, Mbuti); (b) Z-scores for f4 (Bianbian/Boshan/Xiaogao/SD9K/aYR/aLR, Xiaojingshan; aSC, Mbuti).

Table S5. QpAdm analysis showing results for (a) a two-way source model and (b) a three-way source model.

Table S6. Z-scores for f4 (ShanDong_DWK, ShanDong_LS; YR, Mbuti).

Table S7. Estimated timing of admixture for ancestral source components with DATES.

Table S8. Z-scores for f4(YR, YR; 6,000-1,500 BP ShanDong populations, Mbuti). YR=Yellow River.

Table S9. Z-scores for f4(aJapanese > 3,000 BP, aJapanese < 3,000 BP; 6,000-1,500 BP ShanDong populations, Mbuti).

Table S10. Z-scores for f4(WLR/YR/AR, 6,000-1,500 BP ShanDong populations; post-Yayoi aJapanese, Mbuti).

Supplementary Figures:

Fig S1. Principal component analysis (PCA) of PC1 and PC2 for present-day East Asian populations.

Fig S2A. Dimensionality reduction of PC1-PC10 data using the Umap method.

Fig S2B. Dimensionality reduction of PC1-PC10 data using the t-sne method.

Fig S3. Grouping analysis of individuals within each archaeological site.

Fig S4. Residuals heatmap (m=0 to m=6) generated using Treemix.

Fig S6. CV values from the ADMIXTURE analysis to evaluate the chosen K.

Fig S7. ADMIXTURE results for K = 2 to K = 5.

Fig S8. f4 (X, Nagabaka; SDEN/Xiaojingshan/BQGroup, Mbuti), while X was a simulated population ((1-x)% Jomon+x% Longshan ShanDong populations (YJCGroup, CZYGroup))

Fig S9. f4 (X, Nagabaka; Kofun, Mbuti) on simulated data. In this simulation, X is a simulated population ((1-x)% Kofun+x% ShanDong populations, sample size = 30)

Supplementary Data 1

Supplementary Note 1: Site and specimen description

The ancient individuals in this study were sampled from archaeological sites in ShanDong province (north latitude 34° 22.9' 38° 24.01', east longitude 114° 47.5' 122° 42.3'), which is located in coastal North China. These ancient individuals span a period between 6,000 to 1,500 years before present (BP), which covers the Dawenkou, Longshan, and Yueshi cultural periods, as well as the historical early dynastic period up until Han Dynasty. All new individuals sampled for ancient DNA were curated by the archaeologist co-authors of the study, where full permission for sequencing was obtained from the associated universities or archaeological institutes.

Twenty specimens were selected for accelerator mass spectrometry radiocarbon dating (AMS) (Table S1), and were calibrated using OxCal v4.2.4¹ and the IntCal20 calibration curve². All ages were reported as cal BP (years Before Present) where the present-day is set to AD 1950.

BeiQian: The BeiQian site is located in Jimo District, Qingdao City, ShanDong Province, and was excavated in BeiQian Village, Jinkou Township, with geographic coordinates of 120.44 E, 36.36 N. Material artifacts from the BeiQian site show characteristics of both the late Beixin culture and the Dawenkou culture³. Only a small amount of millet and corn is also observed at this site, indicating that the subsistence strategy used by humans at this site was still primarily dominated by hunting-gathering⁴. 30 individuals belonging to the Dawenkou culture at the BeiQian site were obtained for paleogenomic analysis (Table S1). Two genetic groupings were determined, with 28 individuals showing the main group of the BeiQian site (BQGroup), and two individuals designated as outliers (BQ4610 and BQ4628, see Table S1 and section 7 of Materials and Methods). We radiocarbon dated two random individuals from the BQGroup (BQ4906, BQ4582) and both outlier individuals. BQ4906 and BQ4582 of the BQGroup were dated to 5582 - 5333 cal BP and 5460 - 5370 cal BP, respectively. Of outliers, BQ4610 was dated to 5590-5550 cal BP and BQ4628 was dated to 5470-5320 cal BP.

GangShang: The GangShang site is located in Tengzhou City, Zaozhuang City, ShanDong Province, and was excavated on the bank of Kuohe River, northeast of Chengang Village, Dongshahe Town, with geographic coordinates of approximately 117.26 E, 35.09 N. The material culture from this site are primarily associated with the Dawenkou cultural period, though a small number of artifacts from the Zhou and Han dynasties have also been excavated⁵. Archaeological study of a cemetery at the GangShang site shows that social hierarchical stratification was already present in the middle and late stages of the Dawenkou Culture, where the location of artifacts suggest class differentiation, power structure, and the emergence of a ritual system in the Dawenkou cultural period. Burial objects and the cemetery structure at the GangShang site differ greatly from other cemeteries from the same period in other areas of ShanDong. The site structure and layout show rapid expansion and population growth, with a rapid rise in structures suggesting the town grew rapidly, perhaps due migration from other areas⁵. Three individuals from the GangShang site (Table S1) were obtained for paleogenomic analysis, and all were grouped together, denoted the GSGroup. We randomly selected one GangShang individual for radiocarbon dating, and this individual, GS1693, was radiocarbon dated to 4870 - 4840 cal BP.

FuJia: The FuJia site is located in Dongying City, ShanDong Province, and was excavated in the northern part of FuJia Village, Chengguan Town, Guangrao County, with geographic coordinates of approximately 118.40 E, 37.04 N. The site is associated with the characteristic Dawenkou cultural period widespread in the northern part of ShanDong Province and the excavated funerary objects belong to the middle and late stages of the Dawenkou culture, and they show typical local characteristics⁶. A small number of Zhou, Han, Sui, and Tang dynasty-associated cemeteries are also present within the site locality. Three individuals were sampled at the FuJia site, and they were all grouped into the FJGroup. We randomly selected one individual for radiocarbon dating, FJ4777, who was dated to 4520 - 4260 cal BP (Table S1). Because the upper limit of the dating is close to the lower limit of the Dawenkou cultural period (4,600 BP), and because the associated funerary objects are associated with the Dawenkou cultural period, we categorized the FuJia site as associated with the Dawenkou cultural period.

ChengZiYa: The ChengZiYa site is located in Zhangqiu District, Jinan City, ShanDong Province, and was excavated in the northeast of Longshan Village, Longshan Street, with the coordinates of 117.36 E, 36.74 N. The site is located at an average elevation of about 55 m, and the geomorphology type is a typical tableland⁷. Cultural remains at the ChengZiYa site are associated with the Longshan Culture, Yueshi Culture, and the Zhou Dynasty. Five individuals were sampled for this study, where three individuals are associated with the Longshan cultural period and two individuals are associated with the Yueshi cultural period (Table S1). Based on an intragroup analysis using Pairwise D results (see section 7 of Materials and Methods), there were no significant differences at the genetic level between the five individuals sampled at the ChengZiYa site. Since the Yueshi and Longshan cultures have similar cultural characteristics, are found at a similar time period, and the Yueshi culture is generally considered to have developed from the Longshan culture^{8, 9, 10}, we consider the Longshan and Yueshi culture to be the same, and refer to them both as the Longshan culture in the study. We thus grouped all five individuals for population genetic analyses, forming the CZYGroup. We randomly selected one ChengZiYa individual for radiocarbon dating, CZY4652, whom was dated 4420 - 4240 cal BP.

YinJiaCheng: The YinJiaCheng site is located in Jining City, ShanDong Province, and was excavated in the southern part of YinJiaCheng Village, Sishui County, under the jurisdiction of Sishui County, with the geographic coordinates of 117.19 E, 35.64 N. The YinJiaCheng site underwent five excavations^{8, 11, 12, 13, 14}, and the cultural remains excavated at the site span the Dawenkou, Longshan, and Yueshi cultural periods, as well as the Shang, Zhou, Han, and Song dynasties. However, the main cultural layers are associated with the Longshan culture, and the ceramic quail sculpture unearthed at this site shows the same style as the Longshan cultural artifacts unearthed at ChengZiYa and other sites of the same period. We obtained 6 individuals associated with the Longshan cultural period from the YinJiaCheng site (Table S1) for paleogenomic analysis. Of these, we randomly selected one YinJiaCheng individual for radiocarbon dating, YJC4340, who was dated to 4520 - 4300 cal BP.

TongLin: The TongLin site is located in the Linzi District, Zibo City, ShanDong Province, and was excavated on the terrace between TongLin Village, Zhutai Town, and Tianwang Village, Fenghuang Town, with geographic coordinates of 118.25 E, 36.90 N. Early excavations revealed

that the material culture at this site are primarily associated with the Longshan cultural period, which was around 4000 years old¹⁵, but subsequent archaeological work led to excavation of artifacts from stratigraphic accumulations spanning from the Neolithic to the Tang Dynasty¹⁶. In this study, we obtained 8 individuals associated with early dynasties of China, spanning the Han to Jin Dynasties (Table S1). Of these, we randomly selected one individual for radiocarbon dating, TL4775, who was dated to 1690 - 1520 cal BP.

YiXi: The YiXi site is located in Linzi District, Zibo City, ShanDong Province, excavated in the eastern part of Sunjiazhuang Village, Dawu Township with geographic coordinates of 117.88 E, 36.86 N. The site is a series of graveyards discovered by the Qilu Petrochemical Corporation during construction of the Xindian ethylene two-alcohol site, and the excavation and salvaging process was conducted by the Ethylene Project Archaeological Team¹⁷. The team concluded that the age of the YiXi site spans nearly 600 years, from the late Western Zhou Dynasty to the early Western Han Dynasty. In this study, we obtained 5 individuals from the YiXi site that were associated with the Western Han Dynasty (Table S1) for paleogenomic analysis. We randomly selected one individual from the site for radiocarbon dating, YX5196, who was dated to 2310 - 2150 cal BP.

HouLi: The HouLi site is located in the Linzi District, Zibo City, ShanDong Province, and was excavated in the northwestern part of HouLi Guanzhuang Village, Qiling Town, with the geographic coordinates of 118.39 E, 36.83 N. The geographic distribution of the HouLi site is large, and it spans a long time period, where southern localities are associated with the Neolithic, western and northwestern localities are associated with the Eastern Zhou Dynasty, central and northern localities are associated with the Shang, Zhou, and Sui Dynasties, and eastern localities show medium-sized and large-sized burials associated with the Eastern Zhou and Han dynasties, as well as the Song and Yuan dynasties^{18, 19}. In this study, three HouLi individuals were sampled who were associated with the Shang to Zhou dynasties (Tables S1). Two individuals (HL4785 and HL4786) were grouped together as the HLGroup, and one individual (HL4788) showed genetic patterns suggesting that they were a genetic outlier. We randomly selected one individual from the HLGroup for radiocarbon dating, HL4786, who was dated to 3340 - 3180 cal BP. We also radiocarbon dated the outlier HL4788, who was dated to 2760 - 2740 cal BP.

XinZhi: The XinZhi site is located in Changyi City, Weifang City, ShanDong Province, and was excavated in the western part of XinZhi Village, Duchang Street, with the geographic coordinates of 119.39 E, 36.84 N. Many cemeteries have been excavated, and they have been associated with the Zhou, Han, Ming, and Qing dynasties²⁰. In this study, we obtained five XinZhi individuals associated with the Zhou and Han dynasties (Table S1) for paleogenomic analysis. We randomly selected two XinZhi individuals for radiocarbon dating, XZ3468 and XZ3469, who were dated to 2750 - 2720 cal BP and 2320 - 2150 cal BP, respectively.

LiuJiaZhuang: The LiuJiaZhuang site is located in the Gucheng District of Jinan City, ShanDong Province, and was excavated in the village of LiuJiaZhuang, with geographic coordinates of 116.59 E, 36.40 N. Archaeological excavations at the site have revealed material culture associated with the dynastic period, spanning from the Shang Dynasty to the Qing Dynasty^{21, 22}. In

this study, we obtained 15 individuals associated with the Shang and late Han Dynasties (Table S1) for paleogenomic analysis. In a grouping analysis (see section 7 of the Material and Methods), all but one individual showed similar genetic patterns and were grouped as the LJZGroup. One individual showed a different pattern, LZJ4600, and was labeled as an outlier. We randomly selected two individuals from the LZJGroup for radiocarbon dating (see Table S1), LJZ4596 and LJZ4607, who were dated to 2004 - 1880 cal BP and 3080 - 2930 cal BP, respectively. The outlier, LZJ4600, was dated to 3320 - 3080 cal BP.

XiChen: The XiChen site is located in Qufu City, Jining City, ShanDong Province, and was excavated in the western area of XiChen Village, with the geographic coordinates of 116.98 E, 35.55 N. The site is associated with the dynastic period, spanning the Shang Dynasty to the Western Zhou Dynasty, and has been studied extensively for understanding the Shang-Zhou transition²³. In this study, two individuals associated with the Shang dynasty were sampled (Table S1). Of the two, we randomly selected one individual for radiocarbon dating, XC1711, who was dated to 3210 - 3060 cal BP.

Supplementary Note 2: Materials and Methods

(1) Ancient DNA extraction, sequencing, and data processing

For ancient DNA extraction, we primarily selected temporal bone fragments and teeth from human skeletal remains from ancient sites in the ShanDong region for drilling for bone powder. For each specimen, about 100 mg of bone or tooth powder was extracted. In order to avoid inter-sample contamination, use of a disposable drill bit for each specimen was strictly followed during the sampling process. For temporal bone samples, two drilling methods were employed: when we could isolate the temporal bone, we drilled a small hole on the inner side of the temporal bone to obtain bone powder²⁴. When we could not isolate the temporal bone from the intact cranial bone, we drilled from the bottom of the cranial bone²⁵, in order to protect the recognizable morphological features on the surface of the cranial bone.

For DNA library construction, samples from the GangShang and XiChen sites had single-stranded DNA libraries (SS) constructed^{26, 27} and the libraries were not subjected to uracil-DNA glycosylase treatment (non-UDG) (Supplementary Data 1). For samples from the other nine sites, double-stranded DNA libraries (DS) were constructed^{27, 28} and partial uracil-DNA glycosylase treatment was used (half-UDG²⁹). Amplification of DNA libraries was carried out by the AccuPrimepfx DNA enzyme in a polymerase chain reaction (PCR), and libraries were amplified for 35 cycles. The amplification process involved 35 cycles to ensure that enough ancient DNA was available for capture, followed by the addition of P5 and P7 primers to specific libraries. A NanoDrop2000 spectrometer was used to measure the amount of DNA extracted from each sample³⁰.

(2) Sequencing and reads alignment

Oligonucleotide probes designed for ancient nuclear whole genome SNP capture was used, which focused on ~1,240,000 SNPs (1240K SNP array^{31, 32, 33}) (Supplementary Data 1). The enriched

captured DNA fragments were sequenced on Illumina Hiseq2500 and HiSeq X platforms, generating end-paired fragments of 2×100bp and 2×150bp in length. Primer fragments were removed from the original sequences using the leeHom software³⁴, and forward and reverse sequences with at least 11 base pairs of overlap were screened and merged into a single sequence. BWA software³⁵ aligned the merged sequence with the hg19 human reference genome with parameters set to “-n 0.01 -l 16500”. According to the criterion that the mapping quality of the comparison should be greater than or equal to 30, the fragments that did not meet the criterion were filtered. Duplicated sequence fragments, i.e. fragments with the same sequence orientation and same start and end positions, were excluded, where the fragment with the highest quality was retained for further processing.

(3) Test for contamination and genotyping

The C-T substitution rate of each individual terminal nucleotide was calculated. A relatively high C-T substitution rate at the terminal nucleotide is characteristic of ancient DNA³⁶, suggesting that the sequence read represents genetic material from the ancient human sampled. The mitochondrial contamination rate of each individual was assessed by comparing the sequenced fragments with the mitochondrial genomes of 311 present-day humans from around the world using ContamMix software³⁷. For one male individual where mtDNA was not captured, we used an X chromosome contamination test³⁸. Libraries with estimated mitochondrial contamination levels greater than 5% were reprocessed to retain only damaged fragments containing patterns typical of ancient DNA, i.e. they exhibit damage patterns not found in modern DNA³⁹. Damaged fragments were obtained by filtering out fragments containing at least one C-T substitution in the first three positions of the 5' end and in the last three positions of the 3' end by pmdtools0.60 and the “--customterminus” parameter⁴⁰, and the individuals corresponding to these damage-restricted libraries were labeled with a “_d” for subsequent analyses (Table S1). For SNP loci that have reads covered at least once in each individual, a random read was selected to determine the allele for that individual³¹, leading to pseudohaploid genome-wide data for downstream population genetic analyses.

(4) Principal components analysis, Umap dimensionality reduction, t-sne dimensionality reduction

Principal Component Analysis (PCA) was calculated using the smartpca program from the EIGENSOFT package⁴¹, in which we used published present-day humans (34 present-day populations from the HO project⁴², and 17 Tibetan and Han populations differentiated according to region in their published studies⁴³) to determine the principal components (PC1 = 5.46%, PC2 = 3.64%, Fig S1). We then projected ancient ShanDong individuals sampled in this study, as well as previously published ancient individuals^{30, 44, 45, 46, 47, 48, 49} (Fig 1C).

We then assessed PC1 through PC10, collapsing the data through new eigenvalues onto a two-dimensional plane using Umap⁵⁰ and t-sne⁵¹. Compared to the PCA, Umap and t-sne can visualize all 10 PCs on a two-dimensional plane, where Umap (Fig S2) focuses more on global structure and t-sne (Fig S3) focuses more on local structure.

(5) *F3*- and *f4*-analyses

To determine genetic relationships among East Asian populations, the outgroup- f_3 and f_4 analyses found in the software package Admixtools were used⁵². Raghavan et al. first proposed the outgroup f_3 -analysis⁵³, which uses an f -statistic of the form $f_3(\text{Outgroup}; X, Y)$, where the Outgroup is an outgroup population to X and Y . We used the modern Central African population Mbuti as the Outgroup, and ancient East Asians from this and previously published studies as X and Y populations. In practice, we used the qp3Pop software from the AdmixTools⁴² package and plotted heatmaps using the matplotlib package for Python 3.7 (Fig 2A). X and Y populations that share a high f_3 value show high genetic similarity between these two populations. Similarity due to shared ancestry versus admixture can be further differentiated using the f_4 analysis.

We use the qpDstat software in the AdmixTools⁴² package to perform the f_4 analysis and evaluate the relative degree of allele sharing between ancient individuals in East Asia. The f_4 statistic takes the form $f_4(P_1, P_2; P_3, P_4)$, where P_4 is generally fixed as an outgroup to P_1, P_2 , and P_3 . We used the Central African population Mbuti as P_4 , which is outgroup to East Asian populations. In an f_4 analysis, $f_4 > 0$ ($Z > 2.5$ or more strictly $Z > 3$) indicates that the number of alleles shared between the P_1 and P_3 populations is greater than the number of alleles shared between the P_2 and P_3 populations. $f_4 < 0$ ($Z < -2.5$ or more strictly $Z < -3$) indicates that the number of alleles shared between the P_1 and P_3 populations is less than the number of alleles shared between the P_2 and P_3 populations, and $f_4 \sim 0$ ($Z < |2.5|$ or more strictly $Z < |3|$) indicates that the number of alleles shared by P_3 with P_1 and P_2 is approximately equal.

(6) Kinship analyses

In order to exclude the influence of kinship on population genetics analysis, READ (Relationship Estimation from Ancient DNA) software was used to analyze the kinship between human individuals from ancient sites in ShanDong⁵⁴. READ software was specially developed for use with ancient DNA, as the low content of endogenous DNA, fragmentation, terminal damage and other characteristics of aDNA can make estimating kinship difficult. The principle is (1) to divide the genome into non-overlapping windows of 1 Mbp; (2) calculate the proportion of mismatched alleles (P_0) for each window for each pair of individuals; (3) randomly select the expected value of a pair of unrelated individuals in the same population to normalize P_0 ; and finally, (4) classify the kinship between samples according to the threshold value. After processing through the READ analysis, each pair of individuals may be categorized into one of the following four types of kinship: (1) identical individuals/identical twins; (2) first-generation kinship: parents and children, siblings; (3) second-generation kinship: maternal/grandparents and grandchildren, aunts/uncles and nieces/nephews, half-siblings; and (4) unrelated individuals: the distance of kinship is greater than the second-generation range.

Because individuals within two generations of kinship share similar genetic characteristics that can bias population genetics analyses that assume independence of data, we filtered the newly sampled individuals to exclude related individuals in population genetic analyses. That is, for any kinship groups, we retained the individual with the highest data quality. We ultimately excluded 15 individuals (BeiQian = 13, TongLin = 1, GangShang = 1) from downstream population genetics

analyses and marked the excluded individuals with the suffix “_k” (Table S2).

(7) Grouping analysis based on Pairwise D method

To group individuals within sites, we focused on comparing differences between different individuals within the same site using the Pairwise D method. Pairwise D entails using the functionality of the f4 analysis in the AdmixTool software package⁴², with the formula $D(ind1, ind2; Pop, Mbuti)$, where ind1 and ind2 are two different individuals from the same site, and Pop is a published ancient individual or present-day population. A higher number of D-statistics where $|Z| > 3$ indicates genetic differences between ind1 and ind2 that suggest ind1 and ind2 may not share enough genetic similarities to be grouped together. In this study, 107 representative ancient^{30, 33, 45, 46, 48, 49, 55, 56, 57, 58, 59, 60, 61, 62} and present-day^{42, 43} populations were rotated into the Pop position, and the number of D-statistics for each (ind1, ind2) pairing where $|Z| > 3$ was determined. If there were greater than five D-statistics where $|Z| > 3$, then the ancient individuals were divided into subgroups or outliers as appropriate (Fig. S3).

From the grouping analysis using Pairwise D, BQ4628 and BQ4610 were classified as outliers at the BeiQian site, YJC4658 was classified as an outlier at the YinJiaCheng site, XZ3470 was classified as an outlier at the XinZhi site, and HL4788 was classified as an outlier at the HouLi site. Outliers were labeled with the suffix “_o”, and the remaining individuals from that site were grouped together for downstream population genetic analysis (BQGroup of BeiQian site, GSGroup of GangShang site, FJGroup of FuJia site, YJCGroup of YinJiaCheng site, CZYGroup of ChengZiYa site, HLGroup of HouLi site, LJZGroup of LiuJiaZhuang site, XCGroup of XiChen site, XZGroup of XinZhi site, YXGroup of YiXi site, TLGroup of TongLin site).

(8) Phylogeny modeling with Treemix

Treemix v1.13⁶³ was used to determine the phylogenetic relationships of various ancient East Asians^{30, 33, 44, 45, 46, 47, 48}, allowing for admixture events. We rooted the tree using the Central African Mbuti (with the option “-root Mbuti”) and used blocks of 500 SNPs at a time (with the option “-k 500”). We ran 1,000 replicates for each tree, adding the options “-bootstrap -q”. The 1,000 bootstrap trees were assessed in Phylip v3.695 using the “consense” program. With that, we could assess the robustness of each clade in the tree (Table S3). Results for $m = 0$ to $m = 6$, and a heatmap of the residuals were determined (Fig S4-S5), and the tree for $m=3$ is visualized in Fig 2B.

(9) ADMIXTURE analysis

We applied the program ADMIXTURE⁶⁴ to compute stratified components in different East Asian populations based on its likelihood model with a block relaxation algorithm to estimate individual ancestry and cross-validate the population structure. We used PLINK v1.90b3.40⁶⁵ to prune the dataset to minimize linkage disequilibrium, with the parameter “--indep-pairwise 200 25 0.4”. We included present-day and ancient populations used in the PCA analysis. Twenty replicates for each of $K=2$ to $K=9$, using different random seeds. The lowest CV was for $K = 2$ (Fig S6), with

similarly low CVs for $K = 3$ to $K = 5$ (0.4452 - 0.4471). By comparing the results from $K=2$ to $K=5$ (Fig S7), we found that the first separation of components is between North and South East Asians ($K = 2$), the second separation distinguishes continental and island populations amongst northern East Asians ($K = 3$), the third separation distinguishes the Amur River populations (and populations further north) from Yellow River populations ($K = 4$), and the final separation distinguishes Yellow River (inland) and ShanDong (coastal) populations ($K = 5$). These separations across $K=2$ to $K=5$ mirror the population relationships observed in the Treemix analysis (Fig 2B). We visualized the results for $K=5$ in Fig. 2C.

(10) Admixture modeling with qpAdm

To model ancestry proportions for any target population, we used qpAdm³² in AdmixTools with the parameter “allsnps”: default”. We utilized python scripts to implement a rotation strategy to examine the potential ancestral origins of the target population in one-, two-, and three-way mixing scenarios. In the rotated model, a standard outgroup “Yamnaya_Samara”⁵⁹ was added to the “right population”. The possible ancestral source populations are categorized into two groups, “rotating” and “no rotating” (Table S4): (1) Populations in the “rotating” group who are not used as a source in the “left population” are incorporated into the “right population”. The identification of a set of best-fitting mixture models is a major advantage of a rotating qpAdm analysis. Compared to qpAdm without rotation, in a rotating qpAdm analysis, each potential source population is sequentially included as the left_population, where all unused potential sources in that qpAdm analysis are included in the right_population. Finding successful mixture models for one or a few potential sources highlights the optimal sources relative to the other potential sources (where Tail_prob > 0.05, each ancestral mixture proportion > standard error, pnest < 0.05). Using this exhaustive strategy, combined with a smaller number of source populations tends to reveal optimal combinations of sources, because the source populations can only be identified when they outperform the rest of tested potential source populations. Therefore, by this method, we can find the most suitable combination of source populations among all combinations of source populations. (2) Populations in the “non-rotating” group who are not used as a source for the “left population” are not included in the “right population” to avoid situations where the ‘right population’ contains groups that are younger in age than the target population⁶⁶. See Table S4 for details.

The admixture proportions calculated using qpAdm were stratified using the age of the sources, i.e. ShanDong_EN was used as an ancestral source for ShanDong_DWK, ShanDong_DWK was used as an ancestral source for ShanDong_LS, and ShanDong_LS was used as an ancestral source for ShanDong_CD. Thus, to calculate the proportion of the YR component in ShanDong_LS populations, we weighted the ShanDong_EN and YR components based on the ShanDong_DWK proportion observed in each ShanDong_LS population. We used the same method for the ShanDong_CD populations, using the proportions for the ShanDong_LS populations. The results of this re-estimate of proportions was used to visualize the changes in the proportion of YR components in ShanDong populations over the past 6,000 years (Fig 4B).

(11) Estimated mixing time of ancestral source components with DATES.

The timing of admixture events among populations of interest in East Asia was estimated using DATES v4010⁶⁷ (https://github.com/MoorjaniLab/DATES_v4010). The genetic distance was set to 0.45 cM using “lovafit: 0.45”, and the maximum genetic distance was set to a Morgan's maximum using “maxdis: 1” to ensure that it was larger than the confounding LD block. The recommended optimal subgroup size of 0.001 molecules was used (“binsize: 0.001”). Standard errors were estimated by a weighted block jackknife method with the parameter ‘jackknife: Yes’. We considered all results with NRMSD < 0.7, Z > 2, and generations < 200, and assumed that each “generation” corresponds to 28 years⁶⁸ in order to convert generations to years (Table S5).

(12) F4-test based on the simulation method

We included a simulation to test the possible connections between the tested population (Population A) and the possible ancestral component related populations (Population C), which leverages the linear relation between proportion of arbitrary components in the simulated population (Population B) and the value of f4, the line where a series of f4 values are located will pass through the zero point when the ratio of the two components is consistent with that in the population being tested.

Specifically, in f4 (A, B; C, O), (1) Population A contains two population components, i% x and j% y, where i+j=1 and i is a fixed constant; (2) Population B is a series of populations generated by the simulation method and consists of a%x+b%y with a+b=1; Population C is a 100% x component; and (4) O is an outgroup.

We know $f4(A, B; C, O) = (p_A - p_B) \times (p_C - p_O)$, where p_x denotes the frequency of a given allele in population X. So, $p_A = i \times p_x + j \times p_y$; $p_B = a \times p_x + b \times p_y$; $p_C = p_x$; $p_O = 0$.

Then, $f4(A, B; C, O) = (i-a) \times p_x^2 + (a-i) \times p_x p_y$ is a linear equation with respect to the variable a.

Finally, $f4(A, B; C, O) = (p_x \times (i-a) + p_y \times (j-b)) \times p_x = 0$, only if all 3 of the following conditions are met at the same time:

(1) i+j=1; (2) a+b=1; (3) a=i (j-b=(1-i)-(1-a)=a-i).

Also, the conditions that need to be satisfied simultaneously in more complex populations ($A = i\%x + j\%y + k\%z + \dots$) can be further derived from the equation:

(1) i+j+k+...=1; (2) a+b+c+...= 1; (3) a=i, b=j, c=k, ...

So the simulation can be used to demonstrate whether the population components in the tested population A contain only the population components calculated by qpAdm, and to verify if the corresponding component proportions are the same as the proportions of each component calculated by qpAdm when the value of the linear distribution f4 crosses the zero point.

Table S1. Summary of metadata for each ancient individual sequenced in this study

Sample	Region	Site	Longitude(E)	Latitude(N)	Culture/Dynasty	Age(cal BP)	Grave	Coverage	SNP	Library style	Com	Mt hp	Sex	Y hp
XJS4345_low	Jinan, Zhangqiu	Xiaojingshan	117.55	36.86	HouLi		M5	0.021	24979	DS half UDG	N.D.	N.D.	N.D.	N.D.
XJS4353_low	Jinan, Zhangqiu	Xiaojingshan	117.55	36.86	HouLi		M19	0.028	32771	DS half UDG	N.D.	N.D.	N.D.	N.D.
XJS4354_low	Jinan, Zhangqiu	Xiaojingshan	117.55	36.86	HouLi		M21	0.026	30070	DS half UDG	N.D.	N.D.	N.D.	N.D.
BQ4579	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M35-A	0.631	393794	DS half UDG	0.030	M8a3	M	N.D.
BQ4580_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M8A	0.332	266413	DS half UDG	0.015	D4b2	M	N.D.
BQ4582	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou	5461 - 5374	M2-F	0.627	413310	DS half UDG	0.013	D4h1d	M	N.D.
BQ4583_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M2-J	0.415	316021	DS half UDG	0.024	B5b2a2	F	
BQ4584	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M8B	0.601	386087	DS half UDG	0.009	G1a	M	N.D.
BQ4585	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M4C	0.579	386439	DS half UDG	0.023	M8a3	M	N.D.
BQ4586_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M4B	0.712	454318	DS half UDG	0.033	M8a3	M	N.D.
BQ4587	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M31-B	1.052	489968	DS half UDG	0.019	M11c	M	N.D.
BQ4588	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M31-A	1.181	513165	DS half UDG	0.019	B5b2a	M	O3a2c
BQ4589_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M30-B	1.637	575712	DS half UDG	0.014	Z3	F	
BQ4609	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou	5582 - 5333	M32-11	1.59	541433	DS half UDG	0.005	D4	M	N.D.
BQ4610_o	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou	5593 - 5477	M24-23	1.195	520863	DS half UDG	0.012	B5b2a	M	N.D.
BQ4611_d_low_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M2-H	0.022	26256	DS half UDG	0.044	B5b2a2	M	N.D.
BQ4612_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M2-L	1.034	498079	DS half UDG	0.013	B5b2a2	M	N.D.
BQ4613	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M2-M	0.901	453531	DS half UDG	0.005	D4b1a	M	N.D.
BQ4614	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M81-J	1.232	552418	DS half UDG	0.012	B4c1b2a	M	N.D.
BQ4615	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M139-D	1.07	506078	DS half UDG	0.019	B5b2b	M	N.D.
BQ4618	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M32-23	1.145	493082	DS half UDG	0.009	D4h1d	F	
BQ4619_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M24-4	2.048	599271	DS half UDG	0.013	B5b2a	M	NO
BQ4620_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M2-B	1.481	552423	DS half UDG	0.016	B5b2a2	M	N.D.
BQ4621_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M2-E	0.966	443237	DS half UDG	0.023	B5b2a2	M	N.D.

BQ4622_low_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M35-B	0.003	3410	DS half UDG	0.006	N.D.	M	N.D.
BQ4623	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M35-G	0.962	467181	DS half UDG	0.026	A5a	M	N.D.
BQ4624	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M35-J	1.275	499568	DS half UDG	0.023	M8a3	F	
BQ4625_d_low_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M35-L	0.021	24787	DS half UDG	0.035	A5a	F	
BQ4626	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M54-7	0.686	397087	DS half UDG	0.022	Z3	F	
BQ4627	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M54-B	0.66	391951	DS half UDG	0.013	Z3	M	N.D.
BQ4628_o	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou	5472 - 5319	M57-C	0.826	432686	DS half UDG	0.018	Z3	F	
BQ4648_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M57-F	1.768	614580	DS half UDG	0.012	Z3	F	
BQ4649_k	Qingdao, Jimo	BeiQian	120.44	36.36	Dawenkou		M8-I	1.675	580086	DS half UDG	0.013	D5b1b	M	N.D.
FJ4777	Dongying	FuJia	118.40	37.04	Dawenkou	4518 - 4257	M382	0.046	53882	DS half UDG	0.002	D5b1b	M	N.D.
FJ4778	Dongying	FuJia	118.40	37.04	Dawenkou		M399	0.036	42547	DS half UDG	0.004	D5b1b	M	N.D.
FJ4779	Dongying	FuJia	118.40	37.04	Dawenkou		M489	0.038	45006	DS half UDG	0.003	D5b1b	M	N.D.
GS1689_k	Zaozhuang, Tengzhou	GangShang	117.26	35.09	Dawenkou		M1-2	0.32640	282015	SS	0.048	N.D.	M	N.D.
GS1691	Zaozhuang, Tengzhou	GangShang	117.26	35.09	Dawenkou		M1-4	0.26938	243391	SS	0.018	N.D.	M	N.D.
GS1693	Zaozhuang, Tengzhou	GangShang	117.26	35.09	Dawenkou	4870 - 4835	M2	0.63957	480357	SS	0.019	N.D.	M	N.D.
YJC4337	Jining	YinJiaCheng	117.19	35.64	Longshan		M114	2.291	626184	DS half UDG	0.013	N9a4b	M	O3a
YJC4338_low	Jining	YinJiaCheng	117.19	35.64	Longshan		M118	0.007	8180	DS half UDG	N.D.	N.D.	M	N.D.
YJC4339	Jining	YinJiaCheng	117.19	35.64	Longshan		M119	2.284	637867	DS half UDG	0.016	R+16189	M	NO
YJC4340_low	Jining	YinJiaCheng	117.19	35.64	Longshan	4518 - 4297	M122	0.012	14477	DS half UDG	N.D.	N.D.	M	N.D.
YJC4341	Jining	YinJiaCheng	117.19	35.64	Longshan		F205	2.221	630371	DS half UDG	0.006	N9a8	F	
YJC4659	Jining	YinJiaCheng	117.19	35.64	Longshan		M108	1.451	573217	DS half UDG	0.010	C7b	M	N.D.
CZY4652	Jinan, Zhangqiu	ChengZiYa	117.36	36.74	Longshan	4417 - 4240	M4	1.611	577514	DS half UDG	0.006	N9a1'3	M	N.D.
CZY4653	Jinan, Zhangqiu	ChengZiYa	117.36	36.74	Longshan		M6	0.879	527019	DS half UDG	0.014	C7a1	F	
CZY4654	Jinan, Zhangqiu	ChengZiYa	117.36	36.74	Longshan		M8	0.912	411972	DS half UDG	0.013	B4a1c1	F	
CZY4656	Jinan, Zhangqiu	ChengZiYa	117.36	36.74	Yueshi	3454 - 3345	H293-1	2.639	640347	DS half UDG	0.014	B5b	M	O3a2c
CZY4657	Jinan, Zhangqiu	ChengZiYa	117.36	36.74	Yueshi	3475 - 3366	H393-2	3.532	650858	DS half UDG	0.005	D5a2a1	M	C2e2

XC1711	Jining, Qufu	XiChen	116.98	35.55	Shang	3209 - 3063	TN44E19	0.64899	462497	SS	0.006	N.D.	N.D.
XC1715	Jining, Qufu	XiChen	116.98	35.55	Shang		N34E17	0.54033	414811	SS	0.045	N.D.	M N.D.
HL4785	Zibo, Linzi	HouLi	118.39	36.83	Shang-Zhou		M84	0.06	68798	DS half UDG	0.000	D4a5	F
HL4786	Zibo, Linzi	HouLi	118.39	36.83	Shang-Zhou	3340 - 3177	M108	0.612	347059	DS half UDG	0.002	M9a1a	F
HL4788_o	Zibo, Linzi	HouLi	118.39	36.83	Shang-Zhou	2760 - 2737	J13	0.04	47208	DS half UDG	0.000	C7a1c	M N.D.
LJZ4592	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M78	1.617	596097	DS half UDG	0.015	G3a1'2	M Q1a1
LJZ4593	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M89	1.69	579842	DS half UDG	0.010	B4c1b2a1	F
LJZ4594	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M139	0.423	325209	DS half UDG	0.021	A8a	F
LJZ4595	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M67	1.267	540330	DS half UDG	0.017	M8a2b	F
LJZ4596	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han	2004 - 1880	M46	0.678	441478	DS half UDG	0.025	B4a3	M Q1a1
LJZ4597	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M67	1.506	573854	DS half UDG	0.023	M8a3	M C2e1b1
LJZ4598	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M49	1.775	599532	DS half UDG	0.017	N9a10	F
LJZ4599	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M39	2.253	648253	DS half UDG	0.018	M11a	F
LJZ4600_o	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han	3324 - 3082	M96	1.149	542351	DS half UDG	0.012	D5a2a+16092	F
LJZ4601	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M12	0.231	213967	DS half UDG	0.05*	N.D.	M O3a2c
LJZ4602	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M47	0.267	241961	DS half UDG	0.014	N9a11	M O3a
LJZ4605	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M120	1.628	549870	DS half UDG	0.014	A+152+16362	M N.D.
LJZ4606	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M13	2.026	583807	DS half UDG	0.006	D4c1b1	M O2
LJZ4607	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han	3082 - 2927	M117	1.007	468031	DS half UDG	0.011	Z3	M O2
LJZ4608	Jinan, Gucheng	LiuJiaZhuang	116.59	36.40	Shang-Han		M79	0.678	411997	DS half UDG	0.022	D5b1c	M C2e1b1
XZ3467	Weifang, Changyi	XinZhi	119.39	36.84	Zhou-Han		M74	1.46	612023	DS half UDG	0.018	B4b1a2h	M N.D.
XZ3468	Weifang, Changyi	XinZhi	119.39	36.84	Zhou-Han	2753 - 2717	M149	1.304	535632	DS half UDG	0.007	D4h1d	F
XZ3469	Weifang, Changyi	XinZhi	119.39	36.84	Zhou-Han	2317 - 2151	M60	1.224	580359	DS half UDG	0.002	A15a	M O3a2c
XZ3471	Weifang, Changyi	XinZhi	119.39	36.84	Zhou-Han		M99	1.653	554821	DS half UDG	0.014	D4k	F
XZ5197	Weifang, Changyi	XinZhi	119.39	36.84	Zhou-Han		M59	1.505	903569	DS half UDG	0.009	N.D.	F
YX4790_d_low	Zibo, Linzi	YiXi	117.88	36.86	Han		M15	0.012	14313	DS half UDG	0.037	F2a	M O3a

YX4793	Zibo, Linzi	YiXi	117.88	36.86	Han		M084	1.205	517939	DS half UDG	0.007	N.D.	F	
YX5194	Zibo, Linzi	YiXi	117.88	36.86	Han		M010	1.083	773716	DS half UDG	0.007	N.D.	M	O3a2c
YX5195	Zibo, Linzi	YiXi	117.88	36.86	Han		M021	2.696	1079578	DS half UDG	0.008	N.D.	F	
YX5196	Zibo, Linzi	YiXi	117.88	36.86	Han	2310 - 2149	M023	1.056	760331	DS half UDG	0.006	N.D.	F	
TL4770	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		M4	0.103	113663	DS half UDG	0.008	F4a1a	M	O3a2c
TL4772	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		M13	0.59	394079	DS half UDG	0.020	D5b1b	F	
TL4773_d_k	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		M14	0.08	83250	DS half UDG	0.045	B4a4	M	O3a1c
TL4774_low	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		M17	0.017	19920	DS half UDG	0.000	M11a	M	N.D.
TL4775	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin	1685 - 1521	M23	1.468	563375	DS half UDG	0.005	D5a2a1+@16172	F	
TL4776	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		H299-1	2.138	639993	DS half UDG	0.015	B5b1	F	
TL5192	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		M2	2.493	1052021	DS half UDG	0.006	N.D.	F	
TL5193	Zibo, Linzi	TongLin	118.25	36.90	Han-Jin		M11	1.952	1001366	DS half UDG	0.008	N.D.	F	

(1) The cultural periods or historical dynasties of the sites were determined based on a combination of archaeological studies of burial artifacts, stratigraphic information, and the results of radiocarbon dating of at least one individual from each site.

(2) “Com” indicates the contamination rate, where the mitochondrial contamination rate was estimated for most individuals. For one male individual, LJZ4601, mitochondrial data were not captured, so the X-chromosome contamination rate was used. We added a “*” after the contamination rate value to indicate the different methodology used to estimate the contamination rate.

(3) In the column labeled “Sample”, we added the suffixes “_d”, “_low”, “_k”, “_o”:

“_d” indicates where only damage-restricted fragments were included in the libraries; “_low” denotes individuals with SNP counts below 40,000. “_low” individuals where a mitochondrial or X-chromosome contamination rate could not be determined were grouped with unrelated high-coverage individuals from the same site to retain more data and genetic diversity. “_k” denotes lower coverage individuals that were related to another individual at that site. These individuals were excluded from downstream population genetic analyses. “_o” indicates individuals that were categorized as outliers in the grouping analysis for each site. For more details on the suffixes, see the Material and Methods section.

(4) Nuclear genomic data from the Xiaojingshan site have been published in other studies³⁰, and nuclear genomic data from three new individuals were added to the current study and grouped together with the previously published individuals.

- (5) The upper bound of radiocarbon dating from the FuJia site is ~4,500 BP, about 100 years lower than the conventional lower limit of the Dawenkou cultural period. However, archaeological evidence suggests these individuals are best associated with the middle and late Dawenkou culture, so we still regard the FuJia site as a Dawenkou site in the present study (see SI 1 for more explanation).
- (6) Data for which no estimates were determined are shown as “N.D.”

Table S2. Individuals with kinship relationships, with the individual retained for downstream population genetic analysis shown under ‘Treatment’

IND1	IND2	Kinship	Z_upper	Z_lower	Treatment
BQ4579	BQ4623	second-generation kinship	1.93	-10.67	Using BQ4623
BQ4579	BQ4624	first-generation kinship	8.34	-22.03	Using BQ4624
BQ4625_d	BQ4579	second-generation kinship	0.86	-3.87	Using BQ4624
BQ4580	BQ4584	second-generation kinship	3.6	-7.6	Using BQ4584
BQ4580	BQ4649	first-generation kinship	7.83	-18	Using BQ4584
BQ4582	BQ4612	first-generation kinship	6.04	-21.87	Using BQ4582
BQ4582	BQ4620	first-generation kinship	6.97	-22.44	Using BQ4582
BQ4582	BQ4621	first-generation kinship	6.37	-21.33	Using BQ4582
BQ4582	BQ4649	second-generation kinship	0.23	-12.66	Using BQ4582
BQ4582	BQ4611_d	first-generation kinship	2.02	-7.98	Using BQ4582
BQ4611_d	BQ4583	first-generation kinship	3.85	-6.11	Using BQ4582
BQ4583	BQ4612	first-generation kinship	6.94	-21.24	Using BQ4582
BQ4583	BQ4620	first-generation kinship	7.88	-21.36	Using BQ4582
BQ4583	BQ4621	first-generation kinship	7.54	-20.49	Using BQ4582
BQ4585	BQ4586	first-generation kinship	5.79	-19.19	Using BQ4585
BQ4589	BQ4624	second-generation kinship	1.29	-13.12	Using BQ4624
BQ4610	BQ4619	second-generation kinship	1.55	-11.85	Using BQ4610
BQ4611_d	BQ4612	first-generation kinship	1.58	-9.06	Using BQ4582
BQ4611_d	BQ4620	first-generation kinship	3.7	-7.39	Using BQ4582
BQ4611_d	BQ4621	first-generation kinship	3.81	-7.09	Using BQ4582
BQ4612	BQ4620	first-generation kinship	3.4	-23.02	Using BQ4582
BQ4612	BQ4621	first-generation kinship	9.06	-18.21	Using BQ4582
BQ4620	BQ4621	first-generation kinship	12.88	-13.74	Using BQ4582
BQ4622	BQ4623	second-generation kinship	0.88	-1.51	Using BQ4623
BQ4623	BQ4625_d	first-generation kinship	3.71	-6.83	Using BQ4623
BQ4628	BQ4648	first-generation kinship	7.47	-23.61	Using BQ4628
TL4773_d	TL4774	first-generation kinship	0.34	-5.17	Using TL4774
GS1689	GS1693	second-generation kinship	4.35	-5.85	Using GS1693

Table S3. Number of bootstrap replicates (n=1,000) grouping the set of ancient populations indicated by a *.

counts Number	Tianyuan ³³	aFulian EN ³⁰	EarlyJomon ⁴⁴	Nagabaka historic ⁴⁷	Yumin ³⁰	WLR MN ⁴⁶	ARpost14K ⁴⁵	ShanDong BIA	ShanDong EN ³⁰	Xiaojingshan ³⁰	ShanDong MLN	ShanDong LN	YR LN ⁴⁶	YR MN ⁴⁶	Mbuti
1000	.	*	*	*	*	*	*	*	*	*	*	*	*	*	.
1000	.	.	.	.	.	.	.	.	.	.	.	.	*	*	.
999	.	.	.	.	*	*	*	*	*	*	*	*	*	*	.
998	.	.	*	*	*	*	*	*	*	*	*	*	*	*	.
998	.	.	*	*	.	.	.	.	.	.	.	.	.	.	.
996	.	.	.	.	.	.	.	.	*	*	*	.	.	.	.
635	.	.	.	.	*	*	*	.	*	*	*	.	.	.	.
635	.	.	.	.	.	*	*	*	*	*	*	.	.	.	.
592	.	.	.	.	.	.	.	*	*	*	*	*	*	*	.
592	.	.	.	.	.	.	.	.	*	*	*	*	*	*	.
592	.	.	.	.	.	.	.	*	*	*	*	*	*	*	.
575	.	.	.	.	*	*	.	.	.	.	.	.	.	.	.
423	.	.	.	.	*	*	*	*	*	*	*	.	.	.	.
406	.	.	.	.	*	*	*	*	*	*	*	*	.	.	.
361	.	.	.	.	*	*	*	.	*	*	*	.	.	.	.
361	.	.	.	.	*	*	*	.	*	*	*	*	.	.	.
361	.	.	.	.	*	*	*	*	*	*	*	*	.	.	.
361	.	.	.	.	*	*	*	.	*	*	*	.	.	.	.
46	.	.	.	.	*	*	*	.	*	*	*	*	.	.	.
46	.	.	.	.	*	*	*	.	*	*	*	*	.	.	.
4	.	.	.	.	.	.	.	.	.	*	*	*	.	.	.
4	.	.	.	.	*	*	*	.	*	.	.	.	.	.	.
3	.	.	.	.	*	*	.	.	*	.	.	.	.	.	.
2	.	*	.	*	*	*	*	*	*	*	*	*	*	*	.
2	.	.	.	.	.	*	.	.	*	.	.	.	.	.	.
2	.	.	.	*	*	*	*	*	*	*	*	*	*	*	.
1	.	.	.	.	*	.	*	.	*	.	.	.	.	.	.
1	.	.	.	.	.	.	.	*	.	*	*	*	*	*	.
1	.	.	.	*	*	*	*	*	*	*	*	*	.	.	.
1	.	.	.	.	.	.	.	.	*	*	*	*	.	.	.
1	.	.	.	*	*	*	*	.	*	*	*	*	.	.	.
1	.	.	.	.	.	.	.	*	.	.	.	*	.	.	.

Table S4. (a) Z-scores for f4 (Bianbian/Boshan/Xiaogao/SD9K/aYR/aLR/aSC, Xiaojingshan; aAR/FE, Mbuti);

		aAmurRiver					
		AR19K	AR14K	AR13-10K	ARpost9K	AR9.2K_o	Yankovsky IA
aYellowRiver	YR_MN_WGM	-3.2	-7.8	-7.9	-8.4	-5.1	-3.7
	YR_MN_XW	-1.0	-7.4	-6.6	-7.8	-4.4	-3.3
	YR_LN	-3.6	-8.6	-9.9	-10.1	-5.4	-4.1
	YR_LBIA	-2.6	-9.3	-10.4	-10.9	-6.0	-4.3
aShandong	Bianbian	0.8	-2.0	-0.2	-0.8	0.2	0.3
	Xiaogao	-0.4	-3.3	-3.3	-3.7	-3.9	-0.9
	Boshan	1.7	-2.2	-3.0	-1.8	-2.6	0.4
	SD9K	0.6	-3.8	-3.8	-3.3	-3.4	-0.4
aLiaoRiver	WLR_MN	-1.5	-3.5	-3.5	-3.8	-4.2	-0.8
	WLR_LN	-2.6	-7.9	-8.4	-8.1	-4.8	-2.8
	WLR_BA	-1.5	-4.7	-3.9	-3.6	-2.7	-0.3
	WLR_BA_o	0.7	-0.2	1.5	3.6	-2.3	3.5
	HMMH_MN	-2.5	-3.6	-3.2	-2.4	-3.0	-0.8
aSouthChina	Qihe	-4.1	-6.3	-8.3	-8.5	-2.3	-6.1
	Qihe3	-4.4	-10.9	-9.6	-11.5	-5.0	-7.9
	Liangdao1	-2.2	-7.7	-8.8	-10.3	-3.8	-6.3
	Liangdao2	-3.2	-10.0	-10.5	-12.2	-7.3	-7.4
	Suogang	-1.5	-2.1	-3.4	-3.4	-0.7	-2.5
	Xitoucun	-1.1	-4.0	-5.8	-6.4	-2.6	-4.0
	Tanshishan	-1.0	-5.0	-4.4	-6.5	-3.3	-5.5
	Chuanyun	0.5	-0.8	-1.6	-1.6	-0.9	-0.3
	Taiwan_Hanben	-3.6	-11.5	-13.2	-14.8	-5.6	-6.7
	Taiwan_Gongguan	-1.3	-5.7	-5.9	-6.3	-3.0	-4.3
	Longlin	-6.3	-8.4	-10.5	-12.5	-3.4	-7.8
	Baojianshan	-8.2	-12.5	-14.7	-16.1	-7.8	-10.7
	Dushan	-3.9	-9.9	-11.3	-11.7	-3.9	-8.5
	LaCen	-1.3	-7.1	-7.9	-7.9	-4.3	-5.7
	GaoHuaHua	-2.8	-6.9	-7.0	-7.3	-2.6	-3.9
	BaBanQinCen	-2.5	-6.2	-9.8	-8.8	-2.9	-4.0
	Shenxian	-1.6	-6.9	-6.5	-8.7	-3.9	-4.8
	Layi	-1.8	-4.9	-6.5	-6.0	-3.6	-3.5
	Yiyang	-1.1	-3.7	-3.1	-2.7	-0.7	-1.9

Table S4. (b) Z-scores for f4 (Bianbian/Boshan/Xiaogao/SD9K/aYR/aLR, Xiaojingshan; aSC, Mbuti)

		aSouthChina																		
		Qihc	Qihc3	Liangdao1	Liangdao2	Siogang	Xitoucun	Tanshishan	Chuanynun	Taiwan_Hanben	Taiwan_Gongguan	Longlin	Baojianshan	Dushan	LaCen	GaoHuaHua	BaBanQinCen	Shenxian	Layi	Yiyang
aYellowRiver	YR_MN_WGM	-3.6	-2.7	-3.5	-2.1	-2.7	-4.3	-3.4	-1.3	-2.5	0.4	-1.8	-2.1	-2.9	-1.9	-3.1	-5.0	-2.5	-1.1	-1.7
	YR_MN_XW	-3.2	-1.4	-2.1	-2.0	-2.6	-3.0	-1.8	-1.9	-0.7	0.4	-0.5	-3.0	-2.5	-1.3	-2.3	-3.2	-1.5	-0.9	-3.4
	YR_LN	-4.5	-3.2	-3.3	-2.6	-2.8	-4.3	-2.6	-0.3	0.3	1.4	-1.4	-2.3	-3.8	-0.9	-3.0	-3.4	-0.9	-0.8	-1.7
	YR_LBIA	-4.6	-2.5	-3.8	-2.1	-2.5	-4.0	-3.2	-0.3	-0.7	1.6	-1.1	-2.6	-3.8	-0.5	-2.1	-3.2	-1.0	-0.7	-1.6
aShandong	Bianbian	-0.2	0.5	0.6	-0.5	0.1	1.1	0.9	-0.2	0.3	1.3	1.5	-0.8	-0.7	0.7	-1.7	-0.5	1.2	-0.7	-1.5
	Xiaogao	-2.5	-1.2	-1.3	-2.1	-2.5	-2.1	-1.9	-1.0	-2.5	-1.1	-0.6	-2.6	-1.7	-2.5	-3.0	-3.4	-2.6	-1.9	-0.5
	Boshan	-3.2	-1.8	-0.9	-2.0	-1.9	-3.7	-1.8	-0.6	-0.7	0.2	-1.0	-0.7	-1.2	-0.5	-2.8	-2.1	-0.9	-0.9	0.3
	SD9K	-3.6	-1.7	-1.4	-2.4	-2.1	-3.2	-1.7	-0.8	-1.8	-0.4	-0.6	-2.2	-1.8	-1.5	-3.5	-3.4	-1.8	-1.8	-0.5
aLiaoRiver	WLR_MN	-4.4	-4.1	-3.6	-3.6	-3.4	-5.5	-3.8	-1.6	-3.5	-0.5	-1.8	-2.6	-3.1	-3.3	-5.2	-5.1	-3.7	-3.8	-2.8
	WLR_LN	-4.6	-2.9	-3.2	-2.5	-2.2	-4.2	-3.3	-1.2	-1.7	0.3	-1.1	-2.4	-2.4	-1.5	-3.1	-2.3	-1.8	-1.3	-1.5
	WLR_BA	-3.9	-1.7	-2.8	-2.5	-1.8	-3.1	-3.0	-0.7	-1.4	-0.4	-2.7	-2.5	-3.2	-1.0	-2.0	-2.4	-1.1	-0.2	0.6
	WLR_BA_o	-4.0	-2.7	-4.1	-2.8	-3.0	-6.2	-5.1	-1.6	-4.6	-1.7	-1.2	-1.5	-2.2	-4.0	-4.2	-3.6	-2.1	-2.1	-2.8
	HMMH_MN	-4.4	-3.7	-3.9	-4.2	-3.5	-6.9	-5.9	-4.1	-7.2	-4.2	-2.6	-3.9	-4.0	-3.6	-6.2	-8.1	-3.8	-2.7	-4.3
aAmurRiver	AR19K	-3.6	-1.7	-1.3	-0.8	-2.5	-3.1	-0.4	-2.6	-3.8	-1.4	-0.8	-2.1	-1.6	-1.4	-4.6	-4.0	-2.3	-2.0	-3.2
	Boisman_MN	-4.1	-4.2	-3.8	-2.8	-3.2	-5.4	-4.5	-3.0	-4.9	-1.7	-1.9	-2.4	-4.3	-4.1	-5.8	-5.9	-3.5	-1.8	-3.2
	DevilsCave_N	-4.4	-4.6	-5.5	-3.6	-2.3	-7.2	-5.5	-3.0	-5.8	-2.4	-0.7	-3.9	-5.0	-3.5	-6.8	-7.9	-3.5	-4.0	-4.1
	AR14K	-1.4	-2.9	-2.5	-2.4	-0.3	-0.3	-1.0	-1.0	-4.0	-0.8	0.1	0.2	-1.8	-2.1	-3.4	-2.5	-2.6	-1.5	-1.5
	AR13-10K	-2.0	-1.1	-2.5	-1.5	-3.0	-1.5	-0.1	-1.5	-4.6	0.0	0.9	-0.1	-2.0	-1.0	-3.3	-4.5	-0.9	-2.4	-0.2
	ARpost9K	-1.6	-1.7	-2.9	-2.1	-1.1	-1.7	-1.2	-1.4	-5.5	-0.7	-0.6	-0.9	-1.2	-1.6	-3.7	-3.8	-2.1	-1.7	-1.3
	AR9.2K_o	0.3	0.2	0.1	-2.3	0.2	0.0	-1.4	-1.7	-1.8	-0.4	2.8	-1.4	-0.4	-0.8	-1.7	-1.0	-2.0	-2.0	0.8
	Yankovsky_IA	-1.7	-1.9	-1.7	-0.6	-1.8	-1.1	-2.5	-0.6	-0.4	-0.4	0.0	0.1	-1.4	-1.0	-1.1	-1.2	-1.0	-1.0	-0.5

Table S5. QpAdm analysis showing results for (a) a two-way source model and (b) a three-way source model (with “Tail prob” > 0.05 and “pro” > onefold standard error (errx, x = 1, 2, 3)).

(a) Two-way source model

	Target	Source1		Source2		Tail_prob	STD.ERR		
		Pop Name	pro	Pop Name	pro		err1	err2	pnest
DWK	BQGroup	Xiaojingshan	0.87	YR_LN	0.13	0.09	0.05	0.05	0.01
	GSGGroup	Xiaojingshan	0.38	YR_MN	0.62	0.44	0.06	0.06	0.00
	GSGGroup	Xiaojingshan	0.29	YR_LN	0.71	0.34	0.07	0.07	0.00
	GSGGroup	SD9K	0.40	YR_LN	0.60	0.14	0.09	0.09	0.00
	GSGGroup	SD9K	0.50	YR_MN	0.50	0.09	0.08	0.08	0.00
	FJGroup	BQGroup	1.00			0.47			
	FJGroup	SD9K	1.00			0.08			
LS	CZYGroup	GSGGroup	0.87	ARpost14K	0.14	0.87	0.08	0.26	0.13
	CZYGroup	GSGGroup	0.74	BQGroup	0.26	0.06	0.10	0.10	0.02
	YJCGGroup	GSGGroup	0.68	BQGroup	0.32	0.68	0.09	0.09	0.00
	YJCGGroup	GSGGroup	0.79	Xiaojingshan	0.21	0.56	0.08	0.08	0.02
	YJCGGroup	GSGGroup	0.79	SD9K	0.22	0.13	0.14	0.14	0.14
early dynastic period	HLGroup	CZYGroup	1.00			0.09			
	LJZGroup	CZYGroup	1.00			0.47			
	TLGroup	GSGGroup	1.00			0.17			
	XCGGroup	GSGGroup	1.00			0.72			
	XCGGroup	CZYGroup	1.00			0.15			
	XZGroup	GSGGroup	1.00			0.24			
	YXGroup	YR_MN	0.87	ARpost14K	0.13	0.62	0.05	0.05	0.01
	YXGroup	YR_MN	0.75	CZYGroup	0.25	0.28	0.06	0.06	0.00
	YXGroup	YR_MN	0.84	WLR_BA	0.16	0.22	0.04	0.04	0.00
	YXGroup	YR_LN	0.82	CZYGroup	0.18	0.17	0.06	0.06	0.00
	YXGroup	YR_LN	0.90	ARpost14K	0.10	0.15	0.05	0.05	0.05
	YXGroup	YR_MN	0.80	YJCGGroup	0.20	0.15	0.06	0.06	0.00
	YXGroup	YR_MN	0.87	Yumin	0.13	0.11	0.04	0.04	0.00
	YXGroup	YR_MN	0.85	SD9K	0.15	0.10	0.05	0.05	0.01
	YXGroup	YR_LN	0.91	SD9K	0.09	0.09	0.05	0.05	0.11
	YXGroup	YR_LN	0.92	aFujian_EN	0.08	0.07	0.04	0.04	0.04
	YXGroup	YR_LN	0.90	WLR_BA	0.10	0.07	0.04	0.04	0.02
	YXGroup	YR_LN	0.85	YJCGGroup	0.15	0.07	0.06	0.06	0.01
	YXGroup	YR_MN	0.87	BQGroup	0.13	0.05	0.04	0.04	0.00
aJP	Nagabaka_2800BP	EarlyJomon	1.00			0.17			
	Nagabaka_2800BP	InitialJomon	1.00			0.42			
	Nagabaka_2800BP	LateJomon	1.00			0.62			
	Nagabaka_historic	YJCGGroup	0.75	LateJomon	0.25	0.06	0.01	0.01	0.00
	Nagabaka_historic	CZYGroup	0.75	LateJomon	0.25	0.37	0.01	0.01	0.00

- (1) Left population: YR_MN⁴⁶ (YR_MN_WGM, YR_MN_XW), YR_LN⁴⁶, ShanDong_EN³⁰ (Bianbian, Xiaogao, Boshan), Xiaojingshan³⁰, aFuJian_EN³⁰ (Qihe2, Liangdao2), WLR_MN⁴⁶, WLR_BA (including WLR_BA_o)⁴⁶, Yumin³⁰, ARpost14K⁴⁵ (AR14K, AR13-10K, ARpost9K), BQGroup, GSGGroup, YJCGGroup, CZYGroup, Japan_Honshu_EarlyJomon⁴⁴, Japan_Shikoku_InitialJomon⁴⁴, Japan_Shikoku_LateJomon⁴⁴, Kofun⁴⁷;
- (2) Not rotated in the left population: BQGroup, GSGGroup, CZYGroup, YJCGGroup, Japan_Shikoku_LateJomon, Kofun.
- (3) Right population (Not rotated): Yamnaya_Samara⁵⁹

(b) Three-way source model

Target	Source1		Source2		Source3		STD.ERR				pnest	
	Pop Name	pro	Pop Name	pro	Pop Name	pro	Tail prob	err1	err2	err3	pnest	2pnest_1
Xiaojingshan	SD9K	0.74	aFujian_EN	0.10	ARpost14K	0.16	0.99	0.10	0.05	0.08	0.07	0.10
Nagabaka_historic	YR_LN	0.20	YJCGroup	0.56	LateJomon	0.25	0.12	0.10	0.09	0.01	0.04	0.00
Nagabaka_historic	YR_LN	0.17	CZYGroup	0.58	LateJomon	0.24	0.42	0.10	0.10	0.01	0.11	0.00
Nagabaka_historic	WLR_MN	0.30	YJCGroup	0.46	LateJomon	0.25	0.10	0.16	0.16	0.01	0.05	0.00
Nagabaka_historic	WLR_BA	0.13	YJCGroup	0.63	LateJomon	0.25	0.31	0.09	0.09	0.01	0.20	0.00
Nagabaka_historic	WLR_BA	0.11	CZYGroup	0.65	EarlyJomon	0.24	0.06	0.07	0.07	0.01	0.09	0.00
Nagabaka_historic	WLR_BA	0.12	CZYGroup	0.63	InitialJomon	0.26	0.17	0.07	0.07	0.01	0.09	0.00
Kofun	YR_MN	0.54	WLR_BA	0.34	EarlyJomon	0.11	0.14	0.04	0.04	0.01	0.00	0.00
Kofun	YR_MN	0.54	WLR_BA	0.34	InitialJomon	0.13	0.11	0.04	0.04	0.01	0.00	0.00
Kofun	YR_MN	0.56	WLR_BA	0.32	LateJomon	0.12	0.38	0.06	0.06	0.01	0.00	0.00
Kofun	YR_MN	0.70	ARpost14K	0.18	EarlyJomon	0.12	0.28	0.05	0.05	0.01	0.00	0.00
Kofun	YR_MN	0.69	ARpost14K	0.18	InitialJomon	0.13	0.35	0.05	0.05	0.01	0.00	0.00
Kofun	YR_MN	0.72	ARpost14K	0.15	LateJomon	0.13	0.44	0.07	0.06	0.01	0.03	0.00
Kofun	YR_LN	0.41	WLR_MN	0.46	LateJomon	0.13	0.12	0.13	0.13	0.01	0.00	0.00
Kofun	YR_LN	0.59	WLR_BA	0.29	LateJomon	0.13	0.42	0.06	0.06	0.01	0.00	0.00
Kofun	YR_LN	0.49	Yumin	0.38	EarlyJomon	0.13	0.08	0.05	0.05	0.01	0.00	0.00
Kofun	YR_LN	0.48	Yumin	0.38	InitialJomon	0.14	0.07	0.06	0.05	0.01	0.00	0.00
Kofun	YR_LN	0.56	Yumin	0.30	LateJomon	0.14	0.14	0.07	0.07	0.01	0.00	0.00
Kofun	YR_LN	0.62	ARpost14K	0.25	EarlyJomon	0.13	0.12	0.05	0.05	0.01	0.00	0.00
Kofun	YR_LN	0.62	ARpost14K	0.25	InitialJomon	0.13	0.18	0.05	0.05	0.01	0.00	0.00
Kofun	YR_LN	0.66	ARpost14K	0.20	LateJomon	0.13	0.65	0.07	0.07	0.01	0.00	0.00
Kofun	GSGGroup	0.71	WLR_BA	0.16	LateJomon	0.13	0.08	0.12	0.12	0.01	0.37	0.00

- (1) Left population: YR_MN⁴⁶ (YR_MN_WGM, YR_MN_XW), YR_LN⁴⁶, ShanDong_EN³⁰ (Bianbian, Xiaogao, Boshan), Xiaojingshan³⁰, aFuJian_EN³⁰ (Qihe2, Liangdao2), WLR_MN⁴⁶, WLR_BA (including WLR_BA_o)⁴⁶, Yumin³⁰, ARpost14K⁴⁵ (AR14K, AR13-10K, ARpost9K), BQGroup, GSGGroup, YJCGroup, CZYGroup, Japan_Honshu_EarlyJomon⁴⁴, Japan_Shikoku_InitialJomon⁴⁴, Japan_Shikoku_LateJomon⁴⁴, Kofun⁴⁷;
- (2) Not rotated in the left population: BQGroup, GSGGroup, CZYGroup, YJCGroup, Japan_Shikoku_LateJomon, Kofun.
- (3) Right population (Not rotated): Yamnaya_Samara⁵⁹

Table S6. Z-scores for f4 (ShanDong_DWK, ShanDong_LS; YR, Mbuti)

		aYellowRiver			
		YR_MN_WGM	YR_MN_XW	YR_LN	YR_LBIA
P2: CZY	BQGroup	0.1	0.5	1.3	1.6
	GSGroup	3.3	1.0	1.0	2.1
	YJCGroup	2.2	0.7	2.3	1.7
P2: YJC	FJGroup	-0.2	-0.8	0.0	0.8
	BQGroup	-2.3	-0.8	-1.8	-0.8
	GSGroup	1.2	-0.5	-0.9	0.0

Table S7. Estimated timing of admixture for ancestral source components with DATES.

S1	S2	Target	nrmsd	mean	Z
BQGroup	YR_LN	TLGroup	0.371	18.915	2.089
BQGroup	YR_MN	TLGroup	0.452	11.5	3.054
BQGroup	YR_MN	YXGroup	0.457	5.154	2.04
BQGroup	YR_MN	XZGroup	0.52	4.662	2.272
YJGroup	YR_LN	TLGroup	0.568	31.105	2.646
BQGroup	YR_LBIA	TLGroup	0.588	20.766	3.185
CZYGroup	YR_MN	YXGroup	0.654	10.193	2.501
CZYGroup	YR_LN	TLGroup	0.694	15.809	2.13
LJZGroup	Ngbk_2.8kBP	Ngbk_his	0.18	43.342	2.935
CZYGroup	Ngbk_2.8kBP	Ngbk_his	0.238	66.916	2.127
XCGroup	Ngbk_2.8kBP	Ngbk_his	0.287	102.426	3.704
WLR_LN	Ngbk_2.8kBP	Ngbk_his	0.33	55.36	2.702
WLR_BA	Ngbk_2.8kBP	Ngbk_his	0.354	35.156	2.412
HLGroup	Ngbk_2.8kBP	Ngbk_his	0.617	62.091	2.103

(1) "nrmsd" is a parameter that assesses the goodness of fit of the model and is usually required to be less than 0.7⁶⁷.

(2) "mean" is calculated as the average of the number of generations since the mixing event occurred, and one generation is usually considered to be about 28 years⁶⁸.

(3) We also tried to compute the ancestral components of the Kofun population via qpAdm using the same rotation scheme, but even with the addition of the ShanDong population (6,000-1,500 BP) as a possible ancestral population, only the 3-way model passed (Table S5b). Among the 3-way models that passed, we did not observe the ShanDong population as an ancestral model except for the GS population, a population with a large YR component), and except for the components of Jomon and YR, which are stable, the components of another northern East Asian population can be replaced by northern populations from several different regions.

Table S8. Z-scores for f4(YR, YR; 6,000-1,500 BP ShanDong populations, Mbuti). YR=Yellow River.

P1	P3	P2				P3	P2				P3	P2			
		YR_MN_WGM	YR_MN_XW	YR_LN	YR_LBIA		YR_MN_WGM	YR_MN_XW	YR_LN	YR_LBIA		YR_MN_WGM	YR_MN_XW	YR_LN	YR_LBIA
YR_MN_WGM	BQ	0	1.3	-1.1	-1.3	FJ	0	0.2	-0.5	-0.5	QS	0	2	1.9	1.2
YR_MN_XW		-1.3	0	-2.9	-3		-0.2	0	-0.8	-1.2		-2	0	-0.9	-1.5
YR_LN		1.1	2.9	0	-0.4		0.5	0.8	0	-0.4		-1.9	0.9	0	-1.2
YR_LBIA		1.3	3	0.4	0		0.5	1.2	0.4	0		-1.2	1.5	1.2	0
YR_MN_WGM	CZY	0	1.1	-0.1	0.1	YJC	0	2.3	0.3	0.8					
YR_MN_XW		-1.1	0	-1.7	-1.5		-2.3	0	-2.4	-2.1					
YR_LN		0.1	1.7	0	0.2		-0.3	2.4	0	0.3					
YR_LBIA		-0.1	1.5	-0.2	0		-0.8	2.1	-0.3	0					
YR_MN_WGM	TL	0	0.5	-0.5	-0.5	HL	0	-0.2	-0.7	-1.4	LJZ	0	0.4	0.3	0.2
YR_MN_XW		-0.5	0	-1.3	-1.5		0.2	0	-0.7	-1.4		-0.4	0	-0.7	-0.9
YR_LN		0.5	1.3	0	-0.5		0.7	0.7	0	-1.1		-0.3	0.7	0	-0.5
YR_LBIA		0.5	1.5	0.5	0		1.4	1.4	1.1	0		-0.2	0.9	0.5	0
YR_MN_WGM	XZ	0	1	-0.6	-0.7	YX	0	0.2	0	0.2	XC	0	2.5	1	0.6
YR_MN_XW		-1	0	-1.7	-1.7		-0.2	0	-0.3	-0.3		-2.5	0	-2.5	-2.3
YR_LN		0.6	1.7	0	-0.2		0	0.3	0	0		-1	2.5	0	0
YR_LBIA		0.7	1.7	0.2	0		-0.2	0.3	0	0		-0.6	2.3	0	0

Table S9. Z-scores for f4(aJapanese > 3,000 BP, aJapanese < 3,000 BP; 6,000-1,500 BP ShanDong populations, Mbuti). JM=Jomon, Ngbk=Nagabaka.

P3	P1	P2					P3	P1	P2				
		Ngbk_28k	JM_28k	Yayoi	Kofun	Ngbk_his			Ngbk_28k	JM_28k	Yayoi	Kofun	Ngbk_his
BQ	InitialJM	-1.9	-1.1	-4.2	-16.5	-15.5	HL	InitialJM	-0.8	-0.1	-1.3	-10.7	-9.2
	EarlyJM	-2.1	-0.3	-5.8	-20.6	-18.6		EarlyJM	-2.2	-1	-2.6	-14	-12.2
	MidLateJM	-2.1	-0.5	-5.8	-18.5	-16.8		MidLateJM	-1.6	-0.4	-1.5	-12.2	-10.6
	Ngbk_4k	-2.7	-1.7	-1.4	-7.6	-7.1		Ngbk_4k	-1.7	-0.9	-0.8	-4.6	-5.5
	LateJM	-0.1	2.3	-3.6	-12.9	-10.3		LateJM	0.5	1.3	-0.8	-6.9	-5.9
FJ	InitialJM	-0.6	0.9	-0.9	-10.1	-8.2	LJZ	InitialJM	-2.1	-1.6	-3.3	-18.3	-15.9
	EarlyJM	-1.3	0.7	-1.5	-12.8	-10.2		EarlyJM	-1.3	0	-4	-20.9	-17.6
	MidLateJM	-1.5	0.2	-1.5	-12	-9.5		MidLateJM	-1.3	-0.3	-4.3	-20	-16.2
	Ngbk_4k	-0.5	1.5	0.2	-2.5	-2.6		Ngbk_4k	-1.5	-0.8	-2.2	-7.6	-7
	LateJM	-0.9	1.3	-1	-7.4	-4.3		LateJM	0.5	1.8	-2	-13.8	-10.4
GS	InitialJM	-1.8	-0.7	-2.4	-13.3	-10.7	XZ	InitialJM	-2	-2.5	-3.3	-17.4	-13.8
	EarlyJM	-2.1	-0.4	-2.6	-14.9	-13		EarlyJM	-1.5	-0.9	-4.6	-20.7	-15.3
	MidLateJM	-1.5	0.1	-3	-13.5	-11.4		MidLateJM	-0.7	-0.4	-5.1	-18.5	-14.1
	Ngbk_4k	-1.8	1.3	-2.2	-3.8	-3		Ngbk_4k	-3.4	-2.1	-0.8	-8.8	-7.4
	LateJM	-0.3	1	-1.3	-9	-6.9		LateJM	1.3	2.2	-1.2	-13.1	-8.6
CZY	InitialJM	-1.7	-1.5	-3.8	-16.9	-14.8	YX	InitialJM	-1.1	-1.2	-3.6	-16.3	-13.9
	EarlyJM	-1.6	-0.8	-4.6	-20.7	-17.2		EarlyJM	-1.1	-0.9	-3.8	-20.2	-16.4
	MidLateJM	-1.5	-1.4	-5.2	-18.6	-15.8		MidLateJM	-1.3	-1.1	-4.4	-18.4	-15.4
	Ngbk_4k	-1.8	-0.1	-2.3	-7	-6.6		Ngbk_4k	-2.2	-2.1	-1.6	-8.5	-7.3
	LateJM	0.6	2	-1.5	-12.9	-9.8		LateJM	0.7	1.6	-1.3	-13.3	-9.8
YJC	InitialJM	-0.9	-0.8	-2.6	-16.1	-14	XC	InitialJM	-1.4	-1.4	-1.4	-12.3	-10.6
	EarlyJM	-2.1	-0.8	-4.5	-20.3	-17.6		EarlyJM	-0.9	-0.8	-2.4	-14.4	-12.3
	MidLateJM	-1.9	-1.1	-3.9	-19.4	-16.2		MidLateJM	-1.3	-1.4	-2.4	-13.9	-11.9
	Ngbk_4k	-2.2	-0.7	-1.8	-6.8	-6.5		Ngbk_4k	-0.6	-0.7	-1.5	-4.1	-3.8
	LateJM	1.6	3.1	-2	-11.2	-9.2		LateJM	-0.2	1.1	-1.3	-8.9	-7.1
TL	InitialJM	-0.6	-0.5	-3.6	-16.5	-13.3							
	EarlyJM	-1.6	-0.9	-4.4	-21.7	-17.3							
	MidLateJM	-1.4	-0.8	-4.5	-19.8	-15.2							
	Ngbk_4k	-2.6	-1.1	-2.1	-7.8	-7							
	LateJM	1	1.9	-1.7	-13.2	-9							

Table S10. Z-scores for f4(WLR/YR⁴⁶/AR, 6,000-1,500 BP ShanDong populations; post-Yayoi aJapanese, Mbuti). WLR=West Liao River, YR=Yellow River, AR = Amur River.

P1	P2	P3			P2	P3			P2	P3			P2	P3		
		Nagabaka_28k	Kofun	Nagabaka_his		Nagabaka_28k	Kofun	Nagabaka_his		Nagabaka_28k	Kofun	Nagabaka_his		Nagabaka_28k	Kofun	Nagabaka_his
YR_MN_WGM	BQ	-3.7	-1.1	-4.7	CZY	-1.5	0.5	-1	HL	-2.3	-0.2	-1.5	YX	-2.1	-1.8	-3.1
YR_MN_XW		-1.8	-0.7	-4.5		-0.1	0.5	-2.1		-1.1	-0.7	-1.5		-0.5	-1.1	-3.5
YR_LN		-4.3	0	-5.7		-0.9	1.9	-1.6		-2.3	0.5	-2.1		-1.8	-0.7	-3
YR_LBIA		-3.5	-0.1	-4.7		-0.3	1.8	-0.7		-2.1	0.2	-1.2		-0.9	-1.2	-2.5
WLR_MN		-2.1	0.1	-4.8		-0.2	1.2	-2		-1.4	0.1	-2		-0.6	-0.3	-3.1
WLR_LN		-4	0.2	-3.9		-1.9	1.7	-1.2		-2.2	0.5	-1.6		-2.2	-0.1	-2.4
WLR_BA		-2.4	-0.1	-3.4		-1.6	0.4	-1.5		-1	0.9	-0.7		-1.1	-0.6	-2.2
WLR_BA_o		-0.3	-0.3	-2		1.2	1	0.1		-0.2	0.2	-0.4		0.4	-0.7	-0.6
HMMH_MN		-1.9	-2.4	-4.8		0.1	-1.3	-2.5		-0.8	-1.7	-2.8		-0.5	-2.6	-3.4
AR19K		0.5	-3.8	-4.7		1.8	-2.2	-2.5		-1.3	-3.7	-3.9		1.3	-3.8	-4
AR_EN		-2.2	-3.3	-4.9		-1.6	-2.5	-2.6		-1.7	-1.1	-2.6		-0.9	-2.6	-2.5
Boisman_MN		4.2	0.7	-1.6		5.7	2.7	2.1		2.4	1.1	0.5		5.3	-0.4	0.7
DevilsCave_N		-0.3	0.5	-1.9		1.7	2.1	0.7		-0.2	1.1	-0.1		1	-0.7	-0.9
AR14K		-1.9	-4.2	-3.9		0.2	-2.1	-1		-1.3	-2.4	-1.6		-0.2	-4.1	-2.3
AR13-10K		-0.8	-0.9	-2.5		1.8	0.5	0.1		-0.5	-0.7	-0.3		0.9	-0.8	-1.2
ARpost9K		-1.8	-2.2	-3.3		1.1	0.3	0.9		-0.7	-0.2	-0.5		0	-2.3	-1
AR9.2K_o		-1.2	-3	-0.8		0.6	-3	0.6		-0.5	-1.6	-0.8		-0.3	-2.4	0
Yankovsky_IA		-0.1	1.1	0.8		1.3	2.2	2.7		-0.6	1.2	0.7		1	0.6	1.4
AR_Xianbei_IA		-1.2	-2.3	-3.4		-0.4	-1.7	-0.6		-1	-1.1	-1.2		-1.2	-2.9	-2.5
Heshui_Mohe		-1.2	-2.1	-3		0.1	-0.8	-2		-1.3	-0.9	-1.9		-0.3	-1.8	-2.4
YR_MN_WGM	FJ	-0.8	-1.1	-2.2	YJC	-1.9	-0.3	-3.7	LJZ	-2	0	-2.1	XC	-1.8	-0.4	-2
YR_MN_XW		0.4	0	-1.1		-1.1	-0.4	-3.3		-0.5	0.6	-2.6		-0.4	0.6	-1.7
YR_LN		-1.5	-1	-1.6		-2.6	0.2	-3.6		-2.5	1.6	-2.6		-1.2	1.3	-1.3
YR_LBIA		-0.8	-0.9	-1.9		-1.5	0.2	-3.2		-1	1.3	-1.5		-1.6	1.2	-2.2
WLR_MN		-0.1	-1.3	-1.4		-1	0.2	-3.6		-0.5	1.3	-2.3		-0.2	1.5	-1.6
WLR_LN		-1.3	0.1	-2		-2.5	0.5	-3.1		-2.9	1.7	-2		-1	1.3	-0.8
WLR_BA		-1.1	-1.2	-1.4		-1.5	-0.1	-2.2		-1.7	0.7	-1.4		-0.8	0.8	-1
WLR_BA_o		0.5	0.4	-0.7		0.7	-0.1	-0.9		0.6	0.6	-0.2		0.6	0.3	-0.1
HMMH_MN		-1.3	-1.5	-1.4		-0.5	-2.1	-3.7		-0.6	-1.4	-2.8		0.3	-0.6	-1.4
AR19K		0.2	-3.4	-3.5		1.1	-3.1	-4.2		1.9	-2.6	-3.2		1.2	-2.7	-2.9
AR_EN		0.6	-1.9	-1.2		-1.6	-2.2	-3.5		-1.7	-3.1	-2.5		-1	-1.7	-2
Boisman_MN		2.4	-0.5	0.1		4.3	0.6	-0.4		6.1	1.7	1.6		3	0.9	0.2
DevilsCave_N		-0.1	-0.4	-0.3		0.6	0.7	-1.1		1.1	1.5	0.2		0.5	1.4	-0.3
AR14K		-0.6	-2.3	-1.7		-1.1	-3.5	-2.8		-0.1	-3	-1.6		-0.9	-2	-1.7
AR13-10K		0.1	-1.5	-0.6		0.8	-0.2	-1.6		1.2	0.4	-0.1		0.4	-0.4	-0.1
ARpost9K		0.4	-1.2	-0.3		-0.4	-1.6	-1.7		0.6	-0.4	0.6		-0.5	0.4	0.5
AR9.2K_o		0.4	-2.7	-0.7		-0.2	-2.5	0		0.4	-2	0.9		-0.8	-2	0.8
Yankovsky_IA		-0.5	-0.4	-1		0.4	1	1.1		1	1.8	2.7		1	1.8	1.9
AR_Xianbei_IA		-0.1	-1.2	-2		-1.3	-2	-2.5		-1.1	-1.7	-1.5		-1.1	-1.3	-1.4
Heshui_Mohe		0.5	-0.6	0		-0.3	-1.8	-3		0.3	-0.8	-1.7		0	-0.8	-2.2
YR_MN_WGM	GS	-2.1	-0.9	-0.7	TL	-3	-2.4	-2.7	XZ	-2.8	-2.5	-3				
YR_MN_XW		-0.6	-0.3	-1.2		-1	-1.3	-3.2		-1.1	-1.7	-3.2				
YR_LN		-1	0.3	-1.3		-2.9	-1.1	-3		-2.9	-1.5	-3.1				
YR_LBIA		-0.5	0.5	-0.4		-2	-1.6	-2.2		-2	-1.5	-2.4				
WLR_MN		-1.2	-0.4	-2.2		-1.2	-0.6	-2.9		-1.5	-0.9	-3				
WLR_LN		-2.1	-0.3	-1.4		-3.5	-0.7	-2		-3.3	-0.7	-2.6				
WLR_BA		-0.2	-0.7	-1.1		-2.3	-0.8	-1.9		-2	-1.2	-1.7				
WLR_BA_o		0.7	-0.1	0.7		-0.2	-1.3	-0.8		-0.3	-1.1	-0.6				
HMMH_MN		-0.6	-2.3	-2.4		-1	-3.6	-3.4		-1.2	-3.4	-3.2				
AR19K		0.5	-2.9	-3.6		1.1	-4.1	-3.8		0.8	-4.4	-3.5				
AR_EN		-0.9	-2.7	-2.3		-2.3	-3.4	-3.2		-2.1	-3.6	-3.3				
Boisman_MN		3.4	0.2	0.9		4.7	-1.2	1.3		4.1	-1.2	0.6				
DevilsCave_N		0.9	0.4	0		0.4	-1	-0.7		0.2	-1.1	-0.8				
AR14K		-0.2	-2.1	-0.9		-0.9	-4.5	-1.9		-1	-5.1	-2.4				
AR13-10K		0	-0.7	0.8		0.2	-1.1	-1.1		-0.1	-1.9	-1.1				
ARpost9K		0.1	-1.1	0.2		-0.8	-2.5	-0.7		-0.8	-3.4	-1.2				
AR9.2K_o		-0.5	-2.7	0.7		0.2	-3	0.9		-0.8	-2.8	0.4				
Yankovsky_IA		0.6	1.4	2.1		0.6	0.4	1.9		0.3	0.2	1.9				
AR_Xianbei_IA		-1.1	-1.4	-0.9		-1.6	-3.1	-1.9		-1.8	-3.5	-2.3				
Heshui_Mohe		0.5	-0.9	-0.3		-0.3	-1.8	-2.3		-0.6	-2.1	-2.7				

Supplementary Figures:

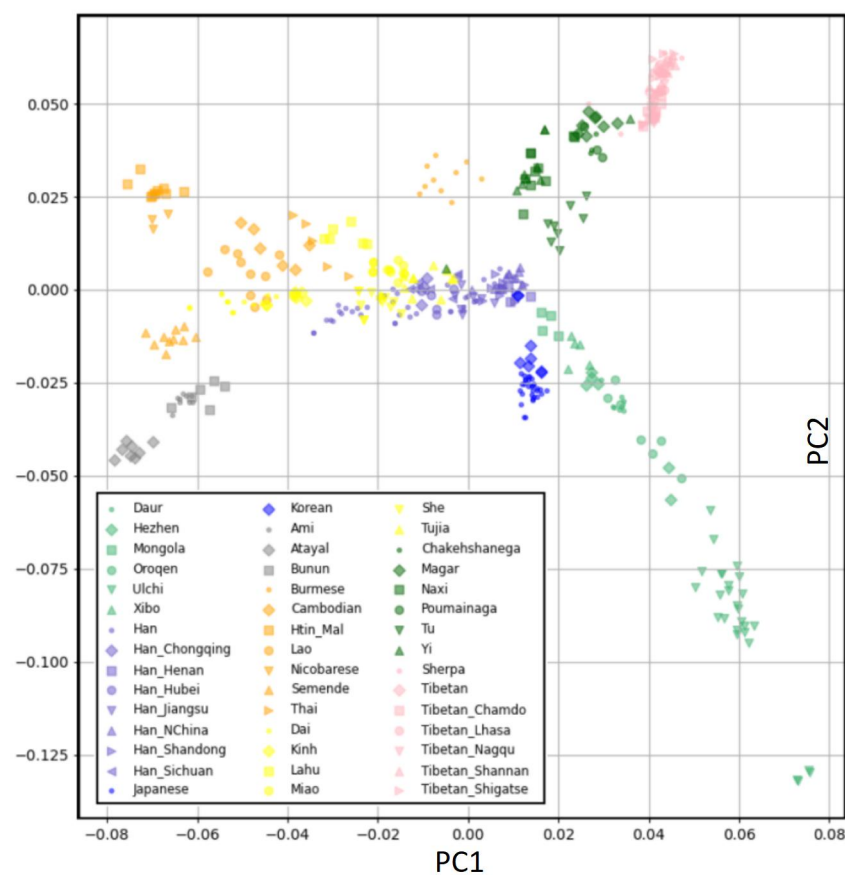


Fig S1. Principal component analysis (PCA) of PC1 and PC2 for present-day East Asian populations. Present-day East and Southeast Asians from northern (bright green, blue), Tibetan (dark green and pink), and southern regions (gray, orange, yellow) form three distinct clines moving away from the origin, where more centrally located populations (dark blue) such as the Han are located.

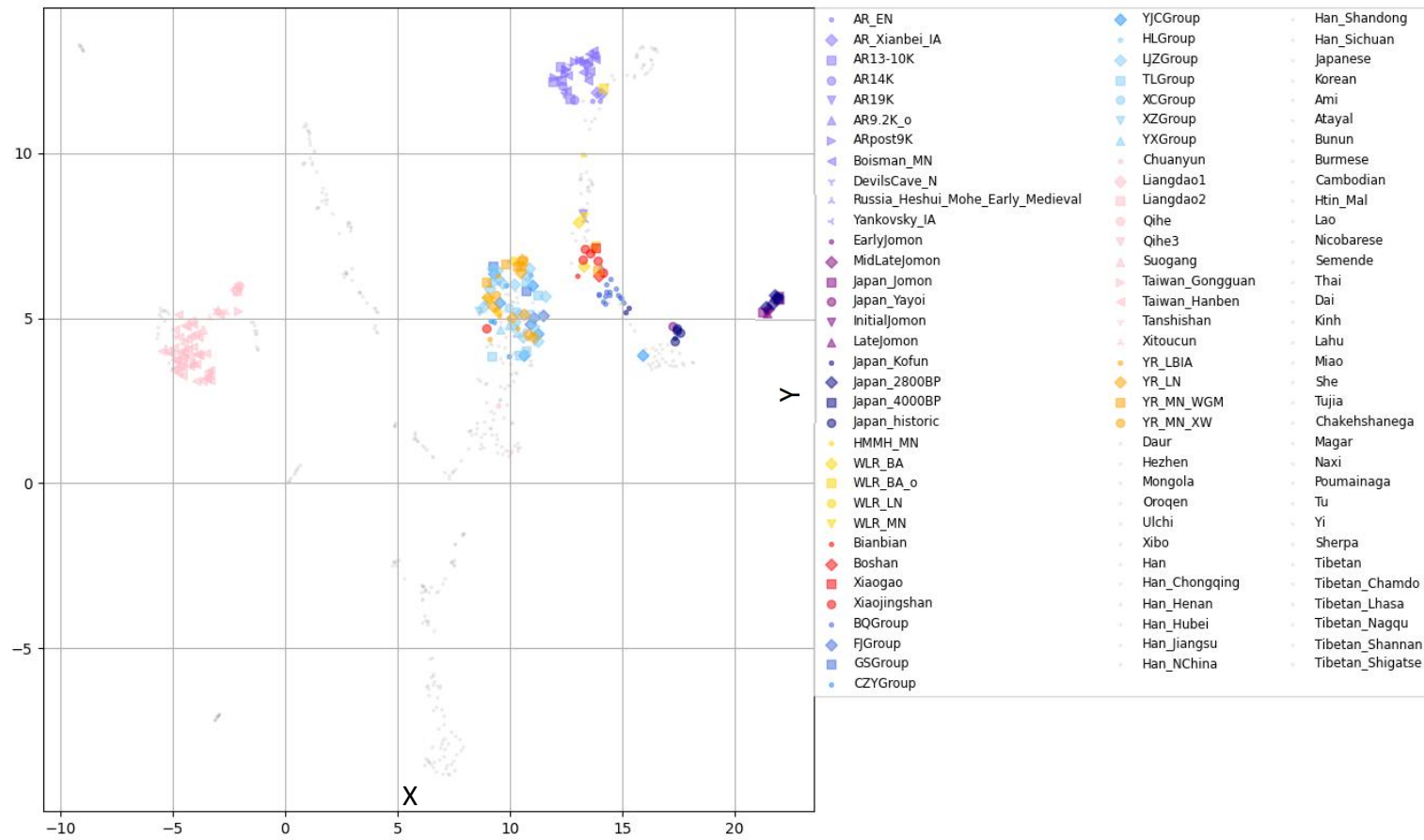


Fig S2A. Dimensionality reduction of PC1-PC10 data using the Umap method, the dataset used was consistent with that in the PCA analysis.

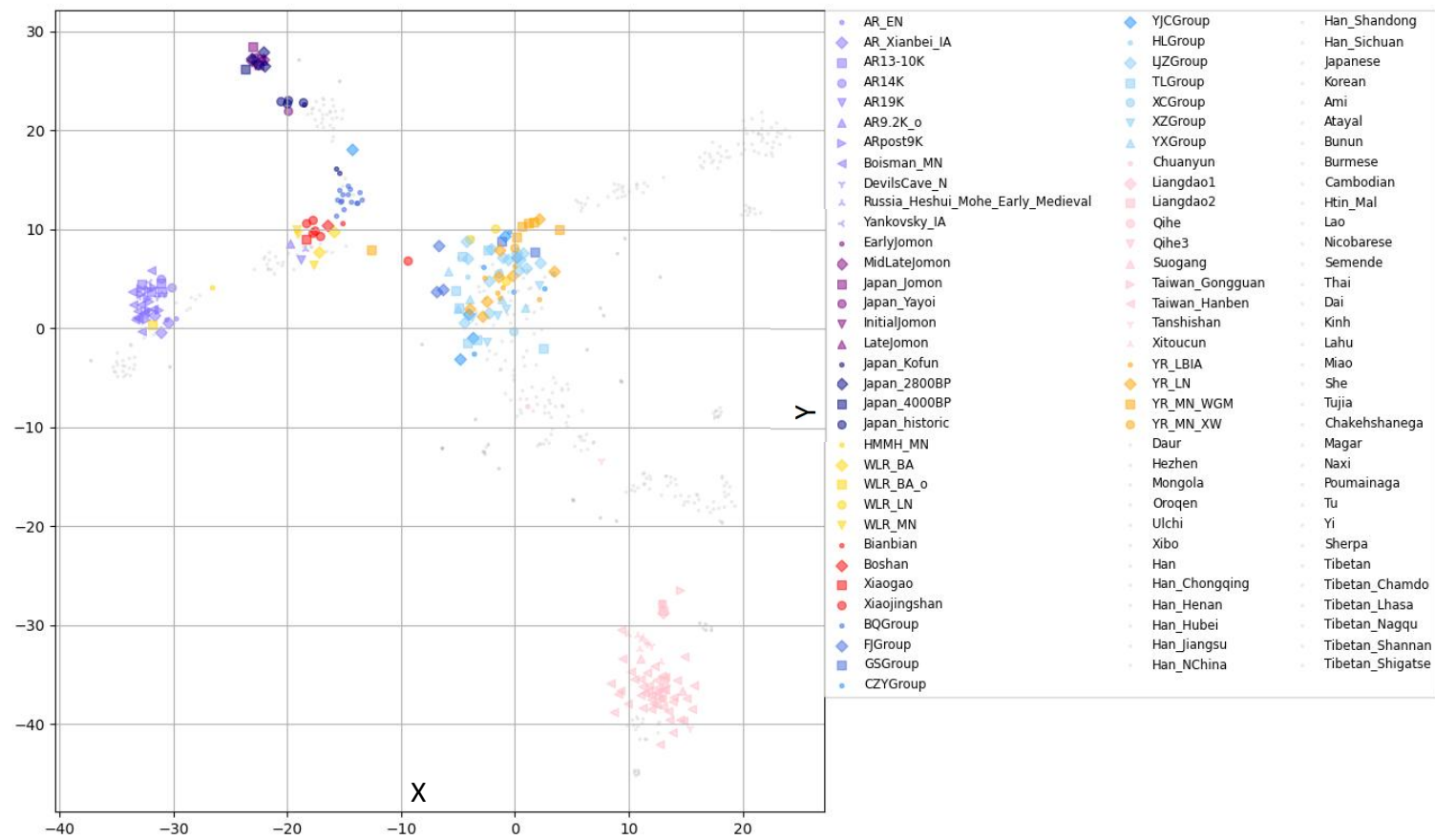


Fig S2B. Dimensionality reduction of PC1-PC10 data using the t-sne method, the dataset used was consistent with that in the PCA analysis.

(a)	LJZ4602	LJZ4607	LJZ4605	LJZ4593	LJZ4595	LJZ4596	LJZ4594	LJZ4608	LJZ4599	LJZ4597	LJZ4598	LJZ4592	LJZ4606	LJZ4601	LJZ4600
LJZ4602	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LJZ4607	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LJZ4605	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LJZ4593	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
LJZ4595	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
LJZ4596	0	0	0	0	0	0	0	3	0	0	0	1	1	0	0
LJZ4594	0	0	0	0	0	0	1	3	0	0	0	0	0	0	8
LJZ4608	0	0	0	0	0	0	1	0	0	0	1	0	0	3	10
LJZ4599	0	0	0	0	3	3	0	0	0	0	1	0	0	1	0
LJZ4597	0	0	0	1	0	0	0	0	0	0	1	0	1	2	0
LJZ4598	0	0	0	1	0	0	1	0	0	0	0	0	1	2	0
LJZ4592	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0
LJZ4606	0	0	0	1	1	0	0	0	0	0	1	0	1	0	8
LJZ4601	0	0	0	1	0	1	0	3	0	1	1	1	1	0	0
LJZ4600	0	2	1	0	0	0	8	10	1	2	2	0	8	0	0

(b)	GS1691	GS1693
GS1691	0	0
GS1693	0	0

(c)	FJ4777	FJ4778	FJ4779
FJ4777	0	0	0
FJ4778	0	0	0
FJ4779	0	0	0

(e)	CZY4657	CZY4656	CZY4654	CZY4652	CZY4653
CZY4657	0	0	0	0	0
CZY4656	0	0	0	0	1
CZY4654	0	0	0	0	1
CZY4652	0	0	0	0	2
CZY4653	0	1	1	2	0

(f)	YJC4339	YJC4341	YJC4340	YJC4338	YJC4659	YJC4337
YJC4339	0	0	0	0	0	0
YJC4341	0	0	0	0	0	0
YJC4340	0	0	0	0	0	1
YJC4338	0	0	0	0	1	0
YJC4659	0	0	0	1	0	1
YJC4337	0	0	1	0	1	0

(g)	YX5196	YX5194	YX4790	YX5195	YX4793
YX5196	0	0	0	0	1
YX5194	0	0	0	0	3
YX4790	0	0	0	1	0
YX5195	0	0	1	0	2
YX4793	1	3	0	2	0

(h)	XZ3471	XZ3468	XZ5197	XZ3469	XZ3467
XZ3471	0	0	0	1	0
XZ3468	0	0	1	0	0
XZ5197	0	1	0	0	2
XZ3469	1	0	0	0	1
XZ3467	0	0	2	1	0

(i)	HL4786	HL4785	HL4788
HL4786	0	0	3
HL4785	0	0	7
HL4788	3	7	0

(j)	XC1711	XC1715
XC1711	0	0
XC1715	0	0

(k)	LJZ4602	LJZ4607	LJZ4605	LJZ4593	LJZ4595	LJZ4596	LJZ4594	LJZ4608	LJZ4599	LJZ4597	LJZ4598	LJZ4592	LJZ4606	LJZ4601	LJZ4600
LJZ4602	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LJZ4607	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LJZ4605	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LJZ4593	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
LJZ4595	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
LJZ4596	0	0	0	0	0	0	0	3	0	0	0	0	1	1	0
LJZ4594	0	0	0	0	0	0	1	3	0	0	0	0	0	0	8
LJZ4608	0	0	0	0	0	1	0	0	0	1	0	0	3	10	0
LJZ4599	0	0	0	0	3	3	0	0	0	0	0	1	0	0	1
LJZ4597	0	0	0	1	0	0	0	0	0	0	1	0	1	2	0
LJZ4598	0	0	0	1	0	0	1	0	0	0	0	0	1	2	0
LJZ4592	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0
LJZ4606	0	0	0	1	1	0	0	0	0	0	1	0	1	0	8
LJZ4601	0	0	0	1	0	1	0	3	0	1	1	1	1	0	0
LJZ4600	0	2	1	0	0	0	8	10	1	2	2	0	8	0	0

(l)	TL4770	TL4775	TL4774	TL5193	TL4776	TL4772	TL5192
TL4770	0	0	0	0	0	0	0
TL4775	0	0	1	0	0	0	2
TL4774	0	1	0	0	0	0	1
TL5193	0	0	0	0	1	3	1
TL4776	0	0	0	1	0	1	1
TL4772	0	0	0	3	1	0	4
TL5192	0	2	1	1	1	4	0

Fig S3. Grouping analysis of individuals within each archaeological site. Yellow highlighting shows where more than five Pairwise D analysis show $|Z|>3$, indicating that the corresponding individual does not show similar genetic patterns with other ancient individuals from the site. These individuals were labeled as outliers, denoted ‘_o’ in Table S1.

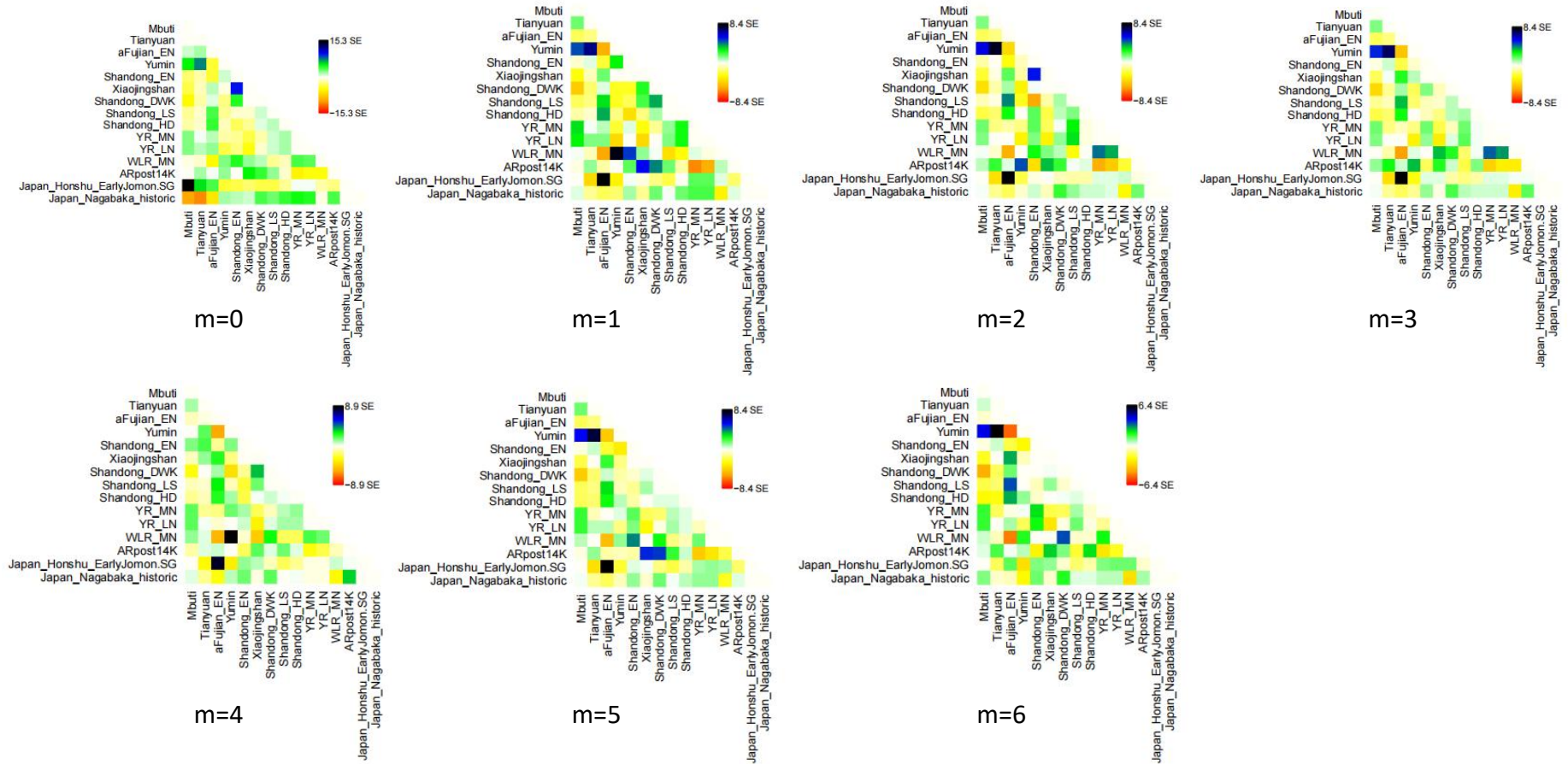


Fig S4. Residuals heatmap ($m=0$ to $m=6$) generated using Treemix.

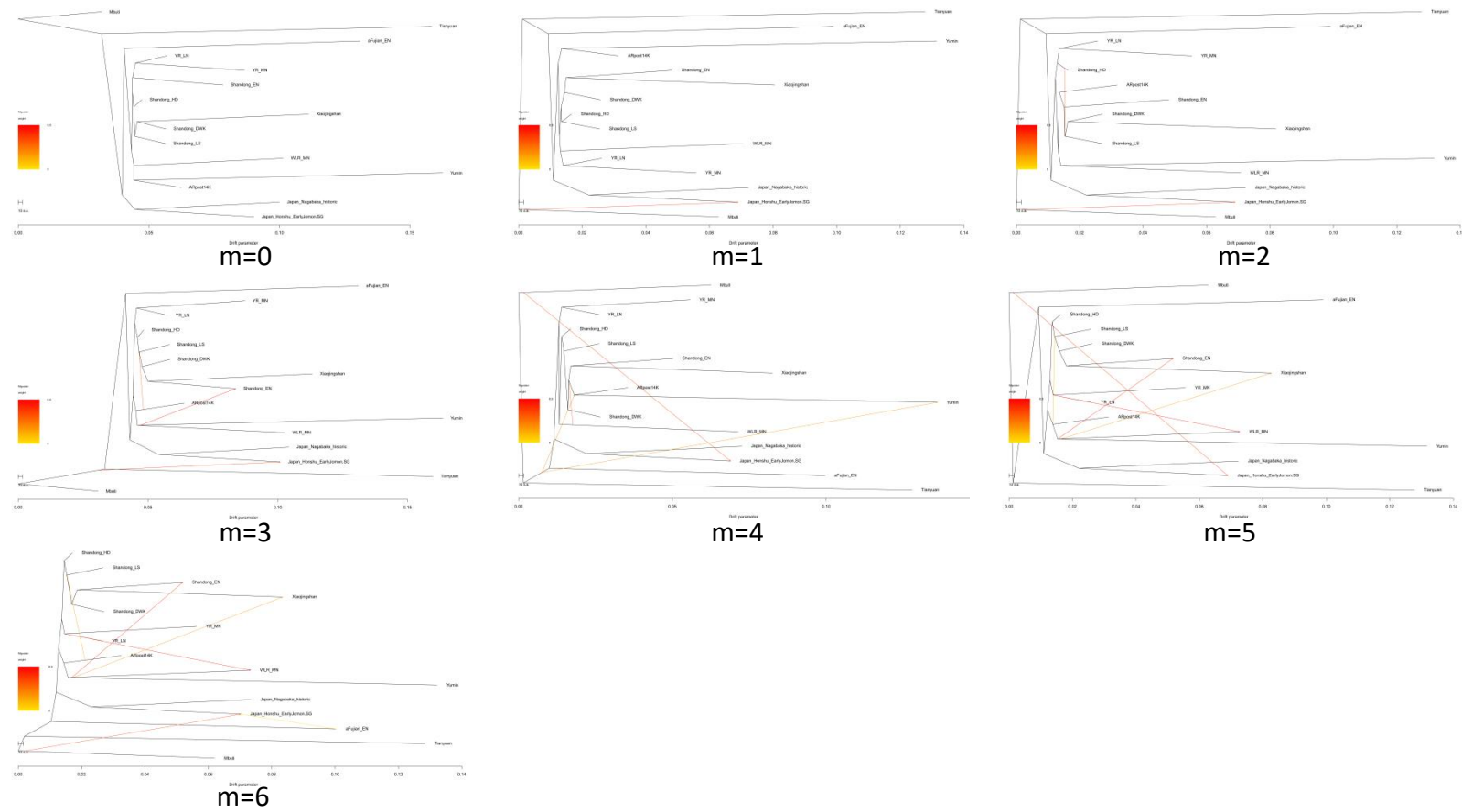


Fig S5. Maximum likelihood phylogenies ($m=0$ to $m=6$) generated using Treemix.

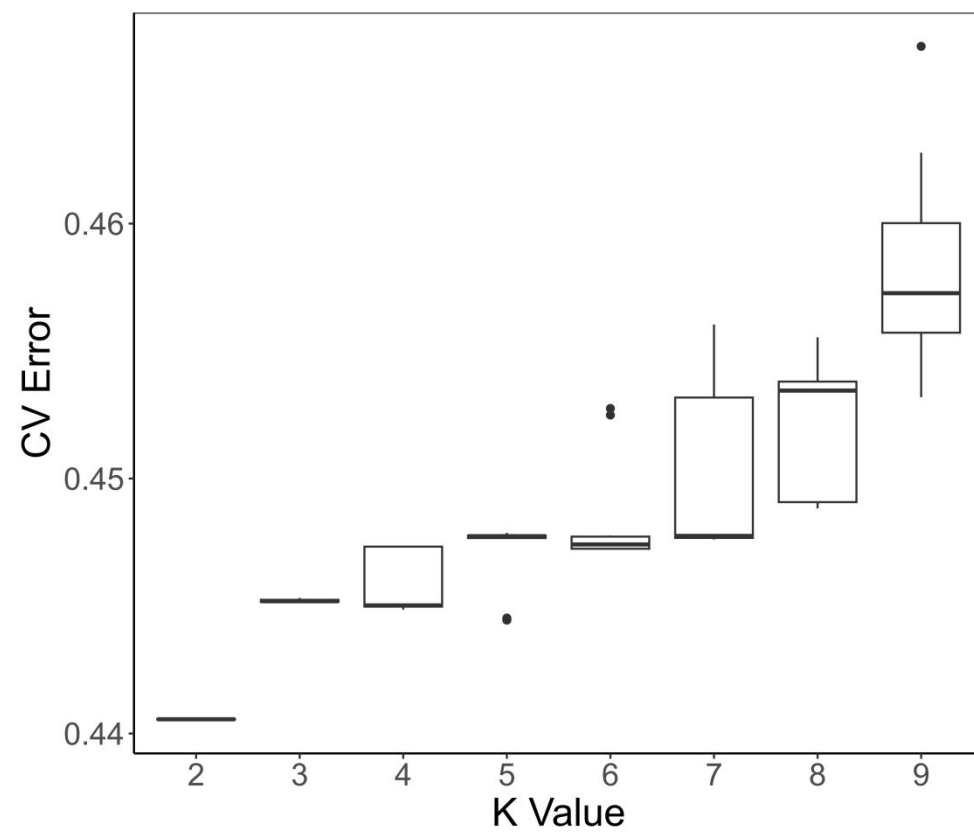


Fig S6. CV values from the ADMIXTURE analysis to evaluate the chosen K. Vertical lines indicate 0-100% of the range of values, upper and lower limits of the rectangle indicate 25-75% of the range of values, horizontal lines within the rectangle indicate the median of the values, and dots indicate outliers.

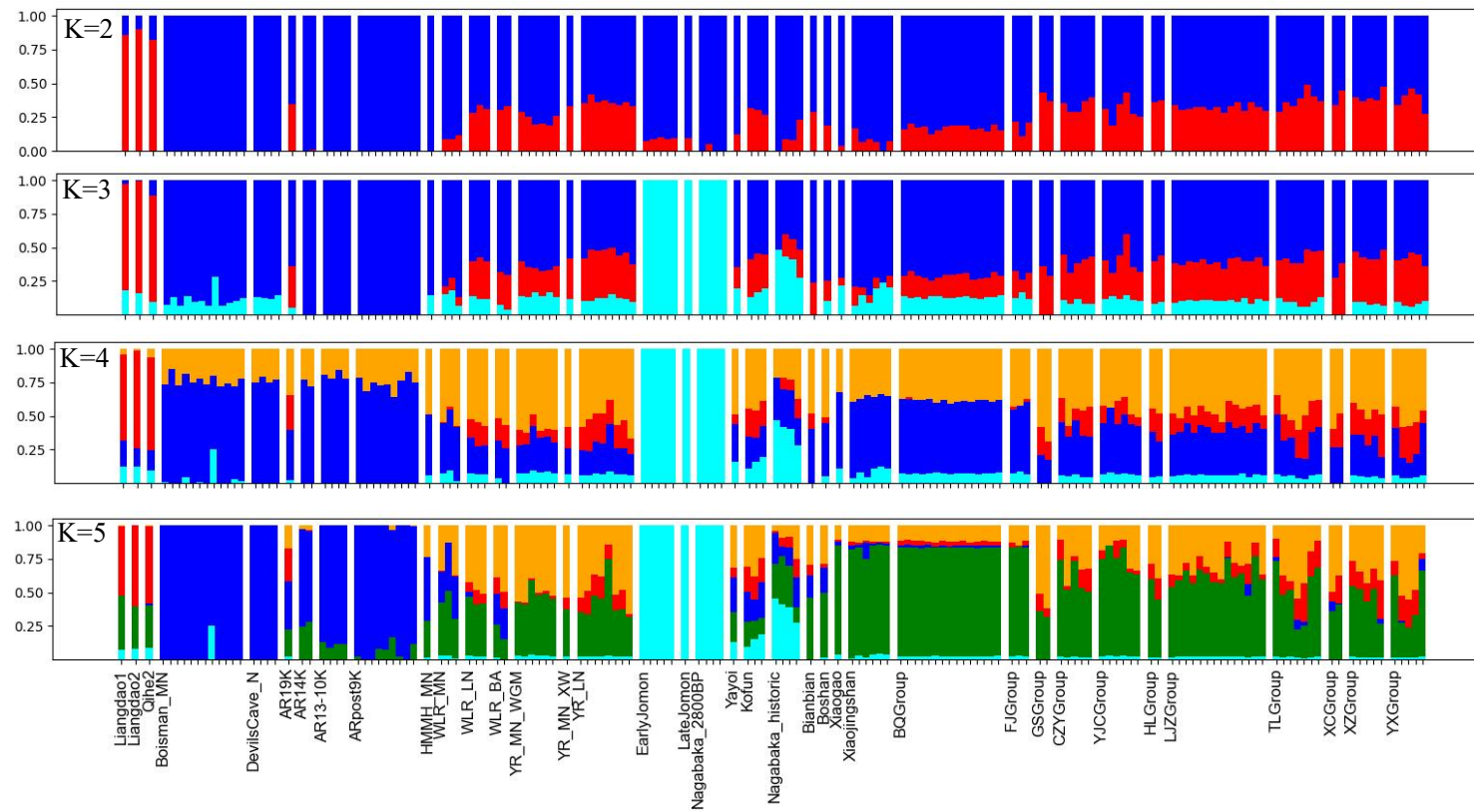


Fig S7. ADMIXTURE results for K = 2 to K = 5.

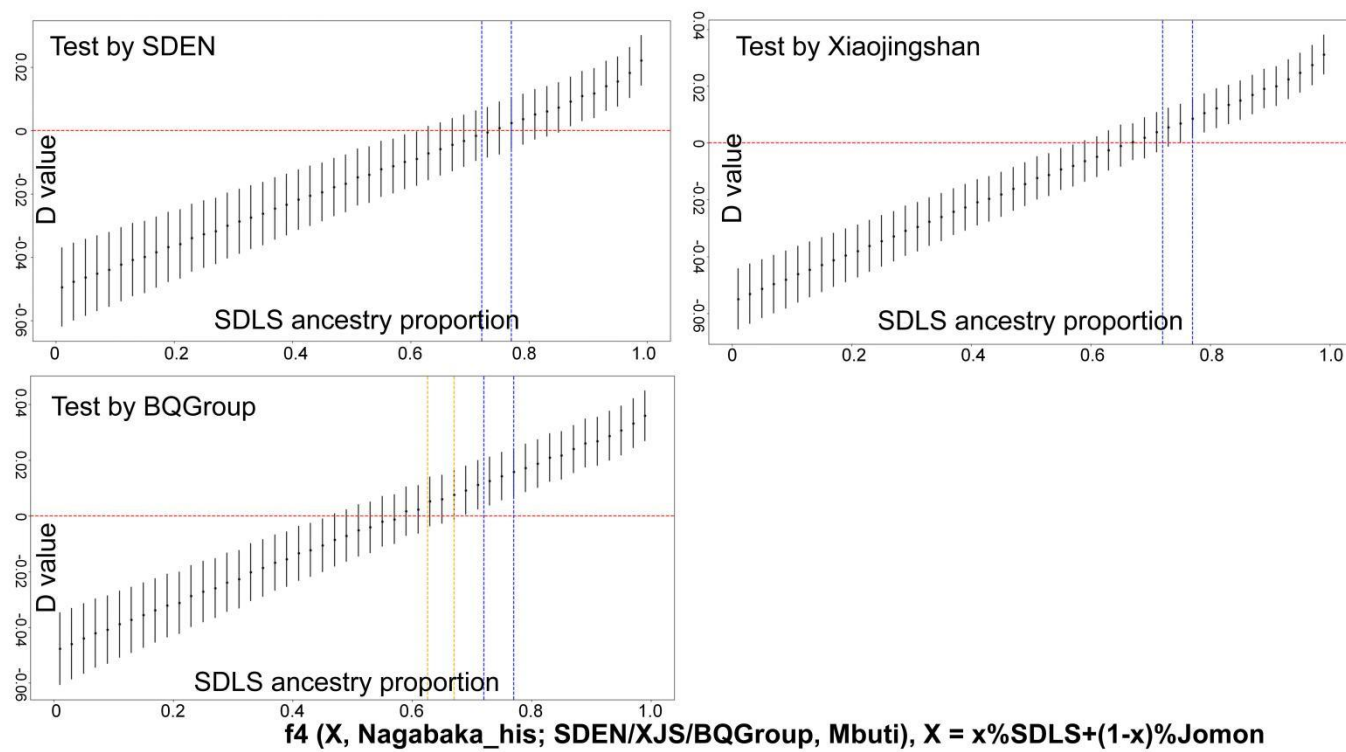


Fig S8. $f_4(X, \text{Nagabaka_his}; \text{SDEN/XJS/BQGroup}, \text{Mbuti})$, while X was a simulated population $((1-x)\% \text{ Jomon} + x\% \text{ Longshan ShanDong populations (YJCGrou, CZYGroup)})$. The measure of centres (circular pattern) are the average values of f_4 , and the error bars were the 3x standard error.

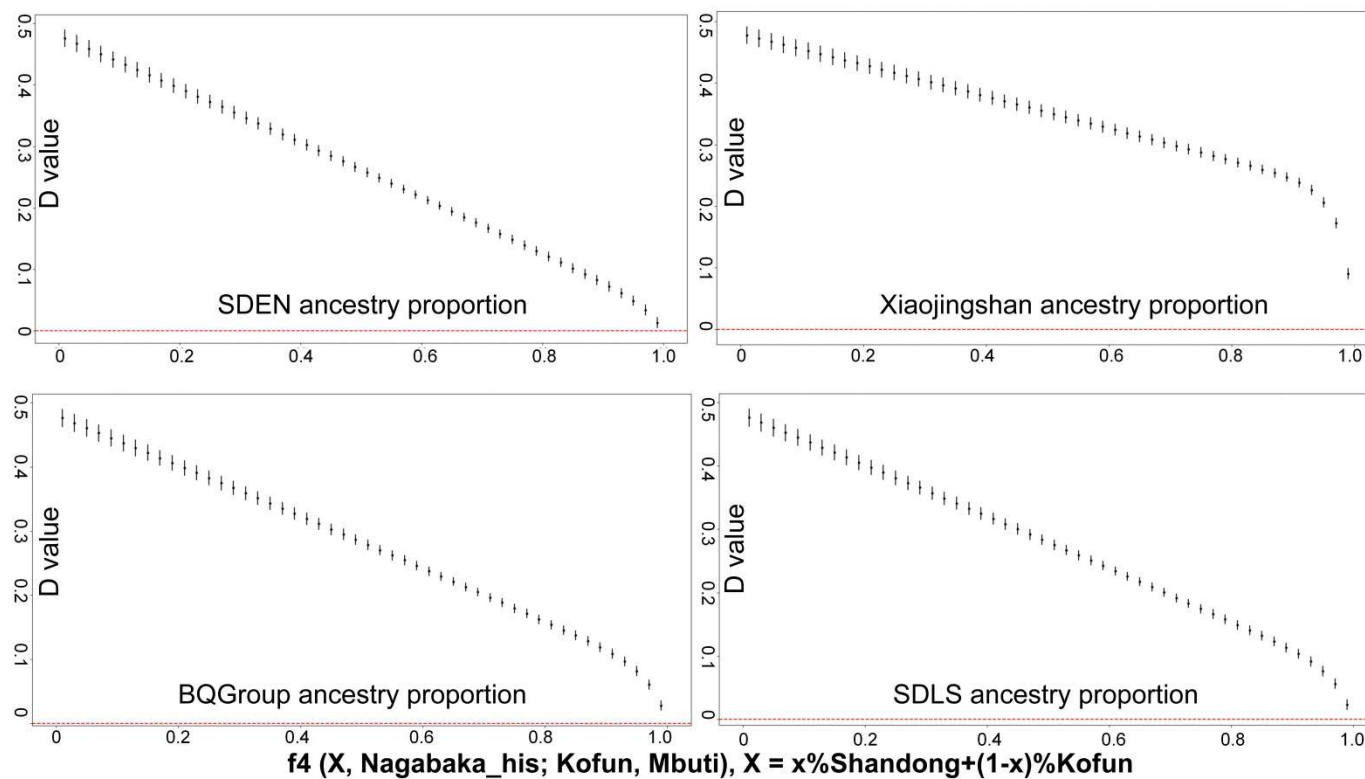


Fig S9. $f_4(X, \text{Nagabaka_his}; \text{Kofun}, \text{Mbuti})$ on simulated data. In this simulation, X is a simulated population ((1-x)% Kofun+x% ShanDong populations, sample size = 30). The measure of centres (circular pattern) are the average values of f_4 , and the error bars were the 3x standard error

References

1. Ramsey CB, Lee S. Recent and Planned Developments of the Program OxCal. *Radiocarbon* **55**, 720-730 (2013).
2. Reimer PJ, Austin WEN, Bard E, Bayliss A, Talamo S. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0-55 kcal BP). *Radiocarbon*, (2020).
3. Wang F, *et al.* Briefing on the 2007 Excavations at Beiqian Site, Jimo City, Shandong, China. *Archaeology*, 3-23+113+197-100 (2011).(in Chinese)
4. Fan Y. A Trial Analysis of the Health Condition of the Dawenkou Culture Period Residents at the Beiqian Site. (2013).(in Chinese)
5. Zhu C, *et al.* Dawenkou Culture Burial Ground, Gangshang Site South, Tengzhou City, Shandong, China. *Archaeology*, 62-64+128+165-181+122 (2023).(in Chinese)
6. Li Z, *et al.* Excavation of the Fujia Site in Guangrao County, Shandong Province. *Archaeology*, 36-44+103+102 (2002).(in Chinese)
7. Zhu C, *et al.* Jinan Zhangqiu District, Chengziyi Site 2013-2015 Excavation Bulletin. *Archaeology*, 3-24+22 (2019).(in Chinese)
8. Cai F, *et al.* The First Trial Excavation of Yinjiacheng in Sishui, Shandong Province. *Archaeology*, 11-17+31+99 (1980).(in Chinese)
9. Li J, *et al.* A preliminary study of the origin, development and social nature of the typical Longshan culture. *Chinese Cultural Relics*, 56-62 (1979).(in Chinese)
10. Yan W. Exploration of the Dongyi Culture. *Chinese Cultural Relics*, 1-12 (1989).(in Chinese)
11. Dong X. Briefing on the Fourth Excavation of Yinjiacheng Site in Sishui, Shandong, China. *Archaeology*, 289-294+376 (1987).(in Chinese)
12. Lan Y. Report on the Excavation of Yinjiacheng Site in Surabaya, 2000. *Archaeology of Haidai*, 353-399+546-556 (2007).(in Chinese)
13. Ren X, *et al.* The Fifth Excavation Briefing of Yinjiacheng Site in Sishui, Shandong Province. *Archaeology*, 399-406+422 (1989).(in Chinese)
14. Yu H, *et al.* Briefing on the Second and Third Excavations of Yinjiacheng Site in Sishui, China. *Archaeology*, 595-601+632 (1985).(in Chinese)

15. Wei C. Introduction to the Tonglin Tianwang Site. *GUAN ZI JOURNAL*, 82+97 (1987).(in Chinese)
16. Wang X. Ontogeny of the Tonglin (Tianwang) Culture Site. *GUAN ZI JOURNAL*, 117-120 (2010).(in Chinese)
17. Wei C, *et al.* Briefing on the Excavation of Linzi Liangol Cemetery. *Archaeology of Haidai*, 274-282+414 (1989).(in Chinese)
18. Wang Y, *et al.* Briefing on the Third and Fourth Excavations at Houli Site, Linzi, Shandong, China. *Archaeology*, 97-112 (1994).(in Chinese)
19. Wang Y, *et al.* Briefing on the First and Second Excavations at Houli Site, Linzi, Shandong, China. *Archaeology*, 987-996 (1992).(in Chinese)
20. Xiao W. Introduction to "Changyi Xinzhi - Report on Burial Excavations from 2010 to 2013". *Archaeology*, 56 (2021).(in Chinese)
21. Guo J, *et al.* Briefing on the Excavation of Shang Dynasty Burials M121 and M122 at Liujiazhuang Site, Jinan, China. *Journal of National Museum of China*, 81-119 (2016).(in Chinese)
22. Guo J, *et al.* Report on the Excavation of Shang Dynasty Burials at Liujiazhuang Site, Jinan, China. *Archaeology of Haidai*, 243-334+585-592+646-656 (2018).(in Chinese)
23. Han H. Xichen Ruins, Qufu, Shandong. *Popular Archaeology*, 12-13 (2020).(in Chinese)
24. Pinhasi R, *et al.* Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. *PLoS One* **10**, e0129102 (2015).
25. Pinhasi R, Fernandes DM, Sirak K, Cheronet O. Isolating the human cochlea to generate bone powder for ancient DNA analysis. *Nat Protoc* **14**, 1194-1205 (2019).
26. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012).
27. Meyer M, *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222-226 (2012).
28. Gansauge MT, Meyer M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat Protoc* **8**, 737-748 (2013).

29. Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos Trans R Soc Lond B Biol Sci* **370**, 20130624 (2015).
30. Yang MA, *et al.* Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* **369**, 282-288 (2020).
31. Fu Q, *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216-219 (2015).
32. Haak W, *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015).
33. Yang MA, *et al.* 40,000-Year-Old Individual from Asia Provides Insight into Early Population Structure in Eurasia. *Curr Biol* **27**, 3202-3208.e3209 (2017).
34. Renaud G, Stenzel U, Kelso J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res* **42**, e141 (2014).
35. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
36. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131 (2012).
37. Fu Q, *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl Acad Sci U S A* **110**, 2223-2227 (2013).
38. Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).
39. Briggs AW, *et al.* Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A* **104**, 14616-14621 (2007).
40. Skoglund P, *et al.* Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc Natl Acad Sci U S A* **111**, 2229-2234 (2014).
41. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
42. Patterson N, *et al.* Ancient admixture in human history. *Genetics* **192**, 1065-1093 (2012).

43. Lu D, *et al.* Ancestral Origins and Genetic History of Tibetan Highlanders. *Am J Hum Genet* **99**, 580-594 (2016).
44. Cooke NP, *et al.* Ancient genomics reveals tripartite origins of Japanese populations. *Sci Adv* **7**, eabh2419 (2021).
45. Mao X, *et al.* The deep population history of northern East Asia from the Late Pleistocene to the Holocene. *Cell* **184**, 3256-3266.e3213 (2021).
46. Ning C, *et al.* Ancient genomes from northern China suggest links between subsistence changes and human migration. *Nat Commun* **11**, 2700 (2020).
47. Robbeets M, *et al.* Triangulation supports agricultural spread of the Transeurasian languages. *Nature* **599**, 616-621 (2021).
48. Wang T, *et al.* Human population history at the crossroads of East and Southeast Asia since 11,000 years ago. *Cell* **184**, 3829-3841.e3821 (2021).
49. Wang CC, *et al.* Genomic insights into the formation of human populations in East Asia. *Nature* **591**, 413-419 (2021).
50. McInnes L, Healy J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *The Journal of Open Source Software* **3**, 861 (2018).
51. Maaten LJPVD, Hinton GE. Visualizing High-Dimensional Data using t-SNE. *Journal of Machine Learning Research* **9**, 2579-2605 (2008).
52. Durand EY, Patterson N, Reich D, Slatkin M. Testing for ancient admixture between closely related populations. *Mol Biol Evol* **28**, 2239-2252 (2011).
53. Raghavan M, *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87-91 (2014).
54. Monroy Kuhn JM, Jakobsson M, Günther T. Estimating genetic kin relationships in prehistoric populations. *PLoS One* **13**, e0195491 (2018).
55. Allentoft ME, *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-172 (2015).
56. de Barros Damgaard P, *et al.* The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, (2018).
57. Fu Q, *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449 (2014).

58. Mondal M, *et al.* Genomic analysis of Andamanese provides insights into ancient human migration into Asia and adaptation. *Nat Genet* **48**, 1066-1070 (2016).
59. Patterson N, *et al.* Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* **601**, 588-594 (2022).
60. Sikora M, *et al.* The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182-188 (2019).
61. Skoglund P, *et al.* Genetic evidence for two founding populations of the Americas. *Nature* **525**, 104-108 (2015).
62. Yu H, *et al.* Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia. *Cell* **181**, 1232-1245.e1220 (2020).
63. Pickrell JK, Pritchard JK. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* **8**, e1002967 (2012).
64. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-1664 (2009).
65. Purcell S, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575 (2007).
66. Yüncü E, *et al.* False discovery rates of qpAdm-based screens for genetic admixture. *bioRxiv*, (2023).
67. Chintalapati M, Patterson N, Moorjani P. The spatiotemporal patterns of major human admixture events during the European Holocene. *Elife* **11**, (2022).
68. Moorjani P, Sankararaman S, Fu Q, Przeworski M, Patterson N, Reich D. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc Natl Acad Sci U S A* **113**, 5652-5657 (2016).