Ancient DNA evidence from China reveals the expansion of Pacific dogs

Ming Zhang[1,2,3], Guoping Sun[4], Lele Ren[5], Haibing Yuan[6], Guanghui Dong[7], Lizhao Zhang[1], Feng Liu[1], Peng Cao[1], Albert Min-Shan Ko[1], Melinda A. Yang[8], Songmei Hu[9*], Guo-Dong Wang[10,11*], Qiaomei Fu[1,2,3*]

[1]Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing, China
[2] CAS Center for Excellence in Life and Paleoenvironment, Beijing, China
[3]University of Chinese Academy of Sciences, Beijing, China
[4]Zhejiang Provincial Institute of Relics and Archaeology, Hangzhou, China
[5]School of History and Culture, Lanzhou University, Lanzhou, China
[6]National Demonstration Center for Experimental Archaeology Education, Department of Archaeology, Sichuan University, Chengdu, China
[7]Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China
[8]Department of Biology, University of Richmond, Richmond, Virginia, USA
[9]Shaanxi Academy of Archaeology, Xi'an, China
[10]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China
[11]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China
***Corresponding authors:** E-mail: fuqiaomei@ivpp.ac.cn; wanggd@mail.kiz.ac.cn; 940779538@qq.com.

Data from this article are provided as Supplementary Material online. Mitochondrial sequences are available on GeneBank under accession numbers MN699608-MN699634.

Abstract

The ancestral homeland of Australian dingoes and Pacific dogs is proposed to be in South China. However, the location and timing of their dispersal and relationship to dog domestication is unclear. Here, we sequenced 7,000 to 2,000-year-old complete mitochondrial DNA (mtDNA) genomes of 27 ancient canids (one gray wolf and 26 domestic dogs) from the Yellow River and Yangtze River basins (YYRB). These are the first complete ancient mtDNA of Chinese dogs from the cradle of early Chinese

civilization. We found that most ancient dogs (18/26) belong to the haplogroup A1b lineage that is found in high frequency in present-day Australian dingoes and pre-colonial Pacific Island dogs, but low frequency in present-day China. Particularly, a 7,000-year-old dog from the Tianluoshan site in Zhejiang province possesses a haplotype basal to the entire haplogroup A1b lineage. We propose that A1b lineage dogs were once widely distributed in the YYRB area. Following their dispersal to South China, and then into Southeast Asia, New Guinea and remote Oceania, they were largely replaced by dogs belonging to other lineages in the last 2,000 years in present-day China, especially North China.

Keywords: Ancient DNA, dogs, mitogenome, China, replacement, Pacific islands

Introduction

Dogs have played a central role in human history, both because they were the first domesticated animal and because they accompanied humans as they dispersed across the world (Clutton-Brock 1995). Scientists have proposed several places of origin for dogs, including Europe (Thalmann et al. 2013), the Middle East (Vonholdt et al. 2010), and southern East Asia (Wang et al. 2016). Ancient mitochondrial DNA suggests that dogs may have been independently domesticated in Western Eurasia but replaced by Eastern lineages 14,000–6,400 years ago (Frantz et al. 2016). However, ancient nuclear genomes show no evidence of a major population replacement (Botigue et al. 2017). The “Asia South of the Yangtze” (ASY) model describes a domestication center south of the Yangtze River as early as 14,000 years ago (Pang et al. 2009). However, archaeological and morphological evidence of domesticated dogs in China only dates up to 8,000 years ago (Liu and Chen 2012), largely because the region is hot and humid, making preservation of ancient remains difficult.

Ancient mitogenomes provide a powerful way of understanding the evolution of dogs (Thalmann et al. 2013). So far, ancient DNA (aDNA) studies have sequenced partial

regions of mitochondrial DNA (mtDNA) (Peng et al. 2015), but they were unable to provide the level of differentiation needed to address questions at finer geographical and temporal scales attuned to archaeological research questions (Paijmans et al. 2013). Sampling complete mitochondrial genomes from China will provide more powerful evidence for this area, which has been virtually unsampled in previous ancient DNA studies of dogs.

Results and Discussion

In this study, we used aDNA capture techniques (Fu et al. 2013) and successfully obtained 27 high-quality complete mitogenomes of 7,000 to 2,000-year-old specimens identified as either dog or wolf. They come from an important time period in the Yellow and Yangtze River basins (YYRB, fig. 1A and supplementary table S1). We obtained 9.0- to 288.6-fold (mean 104.5-fold) coverage for the complete mtDNA sequences used in this study (supplementary table S1). The 7,000 year-old ancient sample from the Tianluoshan archaeological site in Zhejiang province, China, represents the oldest dog sample found south of the Yangtze River. We modified a DNA extraction protocol (Dabney et al. 2013) to extract aDNA from the specimens, which were then used to make double-stranded libraries. After aDNA capture (Fu et al. 2013) and reconditioning of shotgun libraries, we sequenced the libraries on one lane of the Illumina Miseq. Raw sequence reads were processed using a pipeline for mapping reads to the dog mitochondrial reference (NC_002008) (Kim et al. 1998). For comparison, we collected 1,029 complete or nearly complete mitochondrial sequences from the National Center for Biotechnology Information (NCBI). This dataset includes four coyotes, 61 wolves or wolf-like species, 964 dogs or dog-like species, 15 complete mitogenomes from a previously published whole-genome dataset (ten dingoes, two New Guinea Singing dogs, and three dogs from Indonesia) (Zhang et al. 2018), and 72 ancient dogs from Siberia and North America (Leathlobhair et al. 2018) (supplementary table S2).

The population history in YYRB

Haplogroup A accounts for 75% of the matrilineal haplogroups of domestic dogs worldwide (supplementary fig. S1 and supplementary table S2). Asia has the highest diversity of matrilineal sub-haplogroups within haplogroup A (fig. 1B). The proportion of haplogroup A in our ancient dogs (96%) is similar to that found in Southeastern Asian (Thailand) dogs (100%), Australian dingoes (100%) and pre-colonial Pacific island dogs (100%; supplementary table S2 and supplementary fig. S1). Most of our 7,000 to 2,000-year-old dogs belong to sub-haplogroup A1b (69.2%, supplementary table S1). The highest frequency of A1b is found in present-day Australian dingoes (100% belong to haplogroup A1b4) and pre-colonial Pacific Island dogs (95%, 40 out of 42 belong to haplogroup A1b2). Moreover, our ancient Zhejiang sample is the central haplotype in a star-like expansion and basal to the entire A1b haplogroup network of 117 dogs from Asia and Oceania (fig. 1C and supplementary fig. S2). Thus, the ancient Zhejiang sample possibly represents a population that is related to the ancestor of the Alb haplogroup. We found that the genetic patterns significantly correlate with their geography using an AMOVA. Among all investigated combinations, the highest variance among populations (Fct=52.54%, *P*-value < 0.001) is explained by the following groups: Ancient China (A_China), Modern Mainland Asia (MMA), Indonesia (IN), New Guinea (NG), Remote Oceania (RO), Northwestern dingo (DG_NW), and Southeastern dingo (DG_SE) (supplementary table S3). The results also show the close relationship between ancient Chinese dogs and pre-colonial Pacific island dogs (see details in the Supporting Information Results and Discussion).

Between 7,000 and 2,000 years ago, the YYRB dogs show at least four A sub-haplogroups (A1b, A5, A6, and one Unassigned A, fig. 1D and supplementary fig. S3). Only one individual from Jinchankou, Qinghai (~4,000 BP) belongs to haplogroup C, which found primarily in European dogs, suggesting a connection to them (fig. 1D and supplementary fig. S1) (Frantz et al. 2016; Ollivier et al. 2018). None of our ancient dogs from China belongs to the A1a haplogroup, which

constitutes 62% of A haplogroup lineages found in present-day domesticated dogs worldwide. However, present-day dogs found north of the Yangtze River mostly belong to the A1a haplogroup and rarely belong to the A1b haplogroup. To explain this pattern, we propose a 'replacement' hypothesis of the A1b haplogroup by other haplogroups within the last 2,000 years.

After estimating Bayesian phylogenetic trees, we found that major sub-haplogroups of A1b possibly emerged around 9,500 (95% HPD, 10,852-8,173 years ago) to 8,500 (95% HPD, 9,725-7,463 years ago) years ago (fig. 1D and supplementary fig. S3). From a Bayesian skyline plot, we observe a rapid population size expansion of dogs around 7,500 years ago (fig. 1E), which is consistent with the emergence of the major branches of the A1b haplogroup. From 7,000 to 2,000 years ago, a large fraction of the dogs in the YYRB belonged to the A1b haplogroup (69.2%, supplementary table S1). Furthermore, two ancient samples (~1,750BP) found in western Chukotka, Russia (Leathlobhair et al. 2018) are placed at the base of the A1b4 haplogroup (fig. 1D). They have a direct connection to the Tianluoshan sample in the median-joining network analysis (fig. 1C), which suggests that they represent one lineage of dogs that arrived in the polar region before 1,750 years ago. Ameen et al. (2019) found that the A1b haplogroup is also found in the region spanning Alaska to Greenland, which suggests that this dog lineage might have dispersed farther and reached Greenland before five centuries ago. A major replacement likely occurred in the last 2,000 years, which almost replaced the native dogs in the YYRB, especially in the Northern region. For instance, only one individual belongs to the A1b haplogroup in present-day Shaanxi dogs from the Northern region (n = 11). There are no ancient mtDNA after 2,000 years ago in the YYRB, making the time of replacement uncertain.

The emergence of the major branches of the A1b haplogroup coincides with a shift in food production in China around 10,000–8,000 years ago and the rapid spread of agricultural communities across the region (Lu, et al. 2009; Yang, et al. 2012). The population expansion of dogs around 7,500 years ago (Figure 1E) might be caused by

a growth in food supply. As the food supply improved, not only did the human population size increase (Zheng, et al. 2011), but also the population size of animals that lived alongside humans. Millet (C_4 plant) domestication originated in Northern China (Lu, et al. 2009; Yang, et al. 2012), and millet is the only C_4 crop widely cultivated in Shaanxi, Gansu, and Qinghai – provinces in northern China. It was a major food source for ancient humans in these areas before 5,000 years ago. Our samples from Shaanxi, Gansu and Qinghai have strong C_4 signals (supplementary table S1), indicating that these dogs fed on millet byproducts, human leftovers, and/or garbage. The food source of the Zhejiang sample was mostly C_3, more consistent with societies that practiced rice agriculture.

The expansion of Australian dingoes and Pacific dogs

Many studies have concluded that Australian dingoes, New Guinea Singing Dogs (NGSDs), and ancient Polynesian dogs descended from East Asian dogs (Savolainen et al. 2004; Oskarsson et al. 2012; Freedman et al. 2014; Greig et al. 2015; Greig et al. 2018a; Greig et al. 2018b). Piper (2017) claimed that "at least two lineages of dogs were apparently introduced to the island of Southeast Asia (ISEA), the dingo and village dog, whose origins and routes of trans-location across the region remain unknown". (Piper 2017, p. 264). Assuming a substitution rate estimated from the complete mtDNA of dogs included in this study (8.1941×10^{-8} substitutions per site per year, supplementary table S4 and Supporting Information Results and Discussion), the most recent common ancestor (TMRCA) of the A1b haplogroup is estimated to date back to ~10,000 years ago (fig. 1D and supplementary fig. S3). The TMRCA of all sub-haplogroups belonging to A1b dates to around 8,000-6,500 years ago (fig. 1D and supplementary fig. S3). The sub-haplogroup A1b2 includes Pacific dogs who separated from ancient YYRB dogs around 7,200 years ago, and Southeast Asian (SEA) dogs from around 6,000 years ago. Ancient dogs in the YYRB are related to an ~5,500-year-old population related to the ancestors of the SEA dogs, an ~3,300 year-old population related to the ancestors of dogs from Cook Island and other Pacific Islands, and an ~2,700-year-old population related to the ancestors of dogs

from New Zealand and Hawaii (marked in orange in figs. 1D and 2). The divergence time of each branch is earlier than the archaeological records (the presence of domesticated dog in SEA from ~4,000 BP) (Larson et al. 2012; Jones et al. 2019).One of the earliest archaeological sites with evidence of dogs in the Pacific Islands dates back to ~3,000 years ago (Gonzalez et al. 2016), which is consistent with our results.

The sub-haplogroup A1b2 has a wide geographic distribution that corresponds to human populations with Austronesian-related ancestry. The timing of human dispersal seems to have begun ~6,000 years ago in Taiwan or the nearby mainland (Blust 1995; White 2000; Gray et al. 2009; Ko et al. 2014), and human presence can be found in remote Polynesia associated with the Lapita cultural complex (appeared at ~3,300 years ago) (Summerhayes et al. 2010; Skoglund et al. 2016; Lipson et al. 2018; Posth et al. 2018). This suggests that the origins of proto-Austronesians, who likely gave rise to both present-day Austronesian speakers and the Lapita population, may have been in mainland China. These proto-Austronesians then migrated to Taiwan, the Philippines, western Melanesia, and eventually Remote Oceania. This is consistent with a dispersal route of sub-haplogroup A1b2 dogs with more recent dispersals in the Pacific (figs. 1D and 2, supplementary fig. S3). Previous studies also reveal that the dispersal of dogs across the Pacific is inseparably linked to the relationships between dogs and people (Savolainen et al. 2004; Oskarsson et al. 2012; Freedman et al. 2014; Greig et al. 2015; Greig et al. 2018a; Greig et al. 2018b).

Combining all the evidence, we propose that the pre-colonial Pacific dogs possibly originated from the YYRB, dispersed through mainland SEA, and then through Indonesia to the Pacific Islands with people related to the Lapita culture complex. The TMRCA of the majority of Australian dingoes (sub-haplogroup A1b4) dates to 6,844 years ago (95% HPD, 8,048-5,609 years ago). We found a strong Bayesian posterior value (fig. 1D) supporting the separation of Australian dingoes into two groups, also found in network (fig. 1C) and AMOVA analyses (supplementary table S3). One is a southeastern group that clusters with three NGSDs and one Hunan dog, while the

other is a northwestern group that clusters with one ancient dog from Taiwan. Since we are lacking ancient samples that cluster with Australian dingoes, the dispersal route of both dingo groups is still uncertain. The oldest direct date for dingoes in Australia is ~3250 BP from the southern end of the continent, with a suggested colonization time period of ~500 to 1500 years to disperse across the continent (Balme et al. 2018). This suggests that the Australian dingoes migrated from ISEA to Australia, before the Lapita population migration into Remote Oceania.

Conclusions

China possesses one of the earliest agricultural centers in the world, with the domestication of millet (C_4 plant) originating in Northern China before 8,000 years ago (Barton et al. 2009; Bettinger et al. 2010). We find a population expansion for dogs around 7,500 years ago (fig. 1E) that coincides with this shift in food production and a human population expansion in China (Zheng et al. 2011). This may indicate that the A1b haplogroup dogs rapidly dispersed across the entire YYRB with the spread of agricultural communities, and they expanded farther with agricultural populations. The A1b haplogroup is found in ancient dogs from the YYRB, but they appear to have been largely replaced by A1a haplogroup dogs in the last 2,000 years.

At least some dogs carrying the A1b haplogroup successfully spread to Australia and the Pacific Islands before population replacement occurred in China. Our analysis reveals that Australian dingoes and pre-colonial Pacific dogs ancestries can be traced to the YYRB around 10,000 years ago, which suggests that Australian dingoes possibly migrated across the ISEA to Australia prior to Lapita populations (Ko et al. 2014; Skoglund et al. 2016; Lipson et al. 2018; Posth et al. 2018). However, the pre-colonial Pacific dogs dispersed from the YYRB through mainland SEA and Indonesia to the islands in the Pacific Ocean, possibly alongside people ancestral or closely related to Austronesian-speakers (Ko et al. 2014; Skoglund et al. 2016; Lipson et al. 2018; Posth et al. 2018). In the future, more ancient samples and nuclear genome data from East Asia and adjacent regions, particularly from the southern part

of East Asia, would help to clarify the population history and origin of dogs.

Materials and Methods

Sample collections for DNA analyses

A total of 28 ancient samples, spanning the Middle Neolithic to the Iron Age (ca. ~7,000 to 2,000 BP), was obtained from seven archaeological sites in Qinghai (Jinchankou, n=3; Yangqu, n=1), Gansu (Shannashuza, n=4), Shaanxi (Quanhucun, n=8; Yangguanzhai, n=3; Xinjie, n=8), and Zhejiang (Tianluoshan, n=1) provinces of China (fig. 1A and supplementary table S1).

Radiocarbon dating and stable isotope biochemistry

Thirteen specimens representing all time periods were selected for accelerator mass spectrometry (AMS) radiocarbon dating at the BETA laboratory, Miami, USA (supplementary table S1). We ensured that at least the oldest and youngest samples from the same archaeological site were selected for dating. We also tested the δN and δC isotope values to assess the type of food sources consumed by the sampled dogs (supplementary table S1).

DNA extraction and amplification

Bone powder was obtained from each specimen using a Dremel tool and single-use drill bits. All samples were processed in a dedicated ancient DNA laboratory at the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP), Chinese Academy of Sciences in Beijing, China. We used a DNA extraction protocol modified from Dabney et al. (2013). All samples were prepared as double-stranded libraries.

Capture and sequencing of mtDNA

Mitochondrial DNA (mtDNA) were captured by hybridizing the libraries with oligonucleotide probes synthesized by Agilent Technologies (California, USA), as described in SI 3.2 of (Fu et al. 2013). Reconditioned shotgun libraries were sequenced on one lane on the Illumina Miseq at IVPP, Beijing, China.

Authenticity criteria for ancient DNA

Negative controls were used during DNA extractions (every tenth sample) and in all PCR reactions. All reagents used were molecular biology grade and working areas and equipment were decontaminated using bleach, and/or UV irradiation. Only samples that were consistent for repeated extractions and amplifications were included in the analyses. Raw sequence reads were processed using a pipeline for mapping reads to the dog mitochondrial reference (NC_002008) (Kim et al. 1998). We obtained a mean 104.5-fold (range: 9.0 to 288.6) coverage for the 28 complete mtDNA sequences. The damage patterns are a prominent feature of ancient DNA, and all samples showed high C>T frequencies at the 5' end of fragments except libraries treated with half UDG (uracil-DNA-glycosylase). We used a likelihood-based method that has been used previously for estimating present-day contamination in ancient mitochondrial DNA (Fu et al. 2013). The likelihood-based method simultaneously estimates contamination and error by comparing sequenced mtDNA with that found in 644 present-day dogs, whose mtDNA haplotypes are treated as the contaminating population. The contamination rates estimated for all samples are lower than 5%, except for the contamination rate estimated for L3665 (7.3% on average). To exclude the possibility of modern contamination in library L3665, we only used deaminated fragments for further analyses, and still obtained 66.84-fold coverage (supplementary table S1). One sequence contains 77.98% unknown sites, so we discarded this sequence for subsequent analysis (supplementary table S1). The complete mtDNA sequences in this study have been deposited in GenBank (accession numbers: MN699608-MN699634).

Comparison with present-day dogs and population-level analyses

We downloaded 1029 complete or nearly complete mitochondrial sequences from NCBI (https://www.ncbi.nlm.nih.gov/), including four coyotes, 61 wolves or wolf-like species, 964 dogs or dog-like species, and 15 complete mitochondrial genomes from a whole genome dataset consisting of ten dingoes, two New Guinea Singing dogs, and

three dogs from Indonesia (Zhang et al. 2018). We also retrieved mtDNA data for 72 ancient dogs (Leathlobhair et al. 2018). Origin countries of most breed dogs are provided in supplementary table S2. The arrival of European dogs in the Pacific resulted in widespread interbreeding between European and Pacific dogs. As a result, Pacific dogs are no longer identifiable as a distinct breed in the present-day dog population. Thus, we did not use present-day Pacific dog data and only used mtDNA for Pacific dogs sampled from archaeological sites in the Pacific.

All sequences were aligned to the dog mtDNA reference (NC_002008) (Kim et al. 1998) using MUSCLE (v3.8.31) (Edgar 2004). Haplogroups were called using mitotoolpy-seq.py, a python script that is available at Dometree (http://www.dometree.org/trees/dog.htm) (Peng et al. 2015). Several clades are found in both Peng et al. (2015) and Ameen et al. (2019), though using different nomenclature. We have followed the Ameen et al. (2019) nomenclature, where we changed A1 (Peng et al. 2015) to A1a (Ameen et al. 2019), A2 (Peng et al. 2015) to A1b (Ameen et al. 2019), and A4 (Peng et al. 2015) to A2a/A2b (Ameen et al. 2019). Otherwise, we used nomenclature similar to that found in Peng et al. (2015). We also divided the A1b haplogroup into A1b1, A1b2, A1b3, and A1b4 to better correspond to geographical groupings (figs. 1C and 1D). We investigated the genetic population structure of the selected dogs by constructing a median-joining network in PopART (v1.7.1) (Leigh and Bryant 2015) and conducted an analysis of molecular variance (AMOVA) in Arlequin 3.5 to calculate the genetic variation between and among each group, using a significance test based on 10,000 permutations (Excoffier and Lischer 2010) (supplementary table S3).

We used calibrated radiocarbon dates for samples in this study and a dataset from Thalmann et al. (2013) (excluding three low-quality sequences) as well as the mtDNA data to infer a new substitution rate for dogs. We used BEAST 1.8.4 (Drummond et al. 2012) for different partitions: nearly complete (16039 bp, a large number of the downloaded sequences were missing the repetitive and difficult-to-align regions),

protein-coding regions (11409 bp), control region (581 bp), tRNAs (1520 bp), and rRNAs (2534 bp) (supplementary table S4). The best substitution model was selected using the Bayesian information criterion (BIC) in jModelTest 2 (Darriba et al. 2012). Bayesian phylogenetic trees were generated by BEAST 1.8.4 (Drummond et al. 2012) using a Bayesian skyline with a piecewise-linear tree model for 50 million iterations, sampling at every 5,000 iterations (fig. 1D and supplementary fig. S3). All effective sample sizes exceeded 300 in Tracer (v1.6.0) (Rambaut et al. 2014). The phylogenetic tree was visualized in FigTree (v1.4.0) with 20% of the trees discarded as burn-in (fig. 1D and supplementary fig. S3).

Supplementary Materials

Supplementary data are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Supplementary files include:

Supplementary file 1 — Supporting Information for results and discussion.

Supplementary file 2 — Supplementary Figures, including Supplementary Figures S1-S3.

Supplementary file 3 — Supplementary Table 1, Ancient Samples Information.

Supplementary file 4 — Supplementary Table 2, Samples from Published Data.

Supplementary file 5 — Supplementary Table 3, AMOVA Analysis.

Supplementary file 6 — Supplementary Table 4, Substitution Rate Analysis.

Supplementary files 1 and 2 are provided in PDF format while Supplementary files 3-6 are provided in XLSX format.

Acknowledgments

We thank Greger Larson, Laurent Frantz, and Minsheng Peng for giving comments. This work was supported by the Strategic Priority Research Program (B) (XDB26000000) of CAS, NSFC (41925009, 91731303, 41672021, 41630102), Chinese Academy of Sciences (XDB13000000, QYZDBSSW-DQC003, XDA19050102), National Key R&D Program of China (2016YFE0203700), Tencent Foundation through the XPLORER PRIZE, and the Howard Hughes Medical Institute (grant no. 55008731). G.-D.W is supported by the National Youth Talent Support Program and the Program for the Innovative Research Team of Yunnan Province. G.D. is supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP), Grant No. 2019QZKK0601. H.Y. is supported by the National Social Science Fund of China (19BKG038).

Author Contributions

Q.F. designed the research project. Q.F., G.-D.W. and S.H. managed the project. S.H., G.P., L.R. and G.D. provided the samples. M.Z. and L.Z. collected and drilled samples. M.Z., F.L., and P.C. performed DNA experiments. Q.F. and M.Z. did the data processing. M.Z. and Q.F. analyzed the data. M.Z., A.M.X.K., G.-D.W., M.Y., and Q.F. wrote the manuscript. All authors discussed, critically revised, and approved the final version of the manuscript.

References

Ameen C, Feuerborn TR, Brown SK, Linderholm A, Hulme-Beaman A, Lebrasseur O, Sinding M-HS, Lounsberry ZT, Lin AT, Appelt M, et al. 2019. Specialized sledge dogs accompanied Inuit dispersal across the North American Arctic. *P Roy Soc B-Biol Sci.* 286(1916): 20191929. DOI: 10.1098/rspb.2019.1929

Balme J, O'Connor S, Fallon S. 2018. New dates on dingo bones from Madura Cave provide oldest firm evidence for arrival of the species in Australia. *Sci Rep.* 8: 9933. DOI: 10.1038/s41598-018-28324-x

Barton L, Newsome SD, Chen FH, Wang H, Guilderson TP, Bettinger RL. 2009. Agricultural origins and the isotopic identity of domestication in northern China. *Proc Natl Acad Sci U S A.* 106(14): 5523-5528. DOI: 10.1073/pnas.0809960106

Bettinger RL, Barton L, Morgan C. 2010. The origins of food production in North China: A different kind of agricultural revolution. *Evol Anthropol.* 19(1): 9-21.

Blust R. 1995. The prehistory of the Austronesian-speaking peoples: A view from language. *J World Prehist.* 9: 453-510. DOI: 10.1007/BF02221119

Botigue LR, Song S, Scheu A, Gopalan S, Pendleton AL, Oetjens M, Taravella AM, Seregely T, Zeeb-Lanz A, Arbogast RM, et al. 2017. Ancient European dog genomes reveal continuity since the Early Neolithic. *Nat Commun.* 8: 16082. DOI: 10.1038/Ncomms16082

Clutton-Brock J. 1995. Origins of the dog: domestication and early histroy. In: Serpell J, editor. The domestic dog, its evolution, behavior and interaction with people. Cambridge: Cambridge University Press. p. 7-20.

Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paabo S, Arsuaga JL, et al. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A.* 110(39): 15758-15763. DOI: 10.1073/pnas.1314445110

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 9: 772.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 29(8): 1969-1973. DOI: 10.1093/molbev/mss075

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5): 1792-1797. DOI: 10.1093/nar/gkh340

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 10(3): 564-567. DOI: 10.1111/j.1755-0998.2010.02847.x

Frantz LAF, Mullin VE, Pionnier-Capitan M, Lebrasseur O, Ollivier M, Perri A, Linderholm A, Mattiangeli V, Teasdale MD, Dimopoulos EA, et al. 2016. Genomic and archaeological evidence suggests a dual origin of domestic dogs. *Science* 352(6290): 1228-1231. DOI: 10.1126/science.aaf3161

Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B, et al. 2014. Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet.* 10(1): e1004016. DOI: 10.1371/journal.pgen.1004016

Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, Pääbo S. 2013. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl Acad Sci U S A.* 110(6): 2223-2227. DOI: 10.1073/pnas.1221359110

Gonzalez A, Clark G, O'Connor S, Matisoo-Smith L. 2016. A 3000 year old dog burial in Timor-Leste. *Aust Archaeol.* 76(1): 13-20. DOI: 10.1080/03122417.2013.11681961

Gray RD, Drummond AJ, Greenhill SJ. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323(5913): 479-483. DOI: 10.1126/science.1166858

Greig K, Boocock J, Allen MS, Matisoo-Smith E, Walter R. 2018a. Ancient DNA evidence for the introduction and dispersal of dogs (*Canis familiaris*) in New Zealand. Journal of Pacific Archaeology 9(1): 1-10.

Greig K, Boocock J, Prost S, Ann Horsburgh K, Jacomb SC, Walter R, Matisoo-Smith E. 2015. Complete mitochondrial genomes of New Zealand's first dogs. *PLoS ONE* 10(10): e0138536. DOI: 10.1371/journal.pone.0138536

Greig K, Gosling A, Collins CJ, Boocock J, McDonald K, Addison DJ, Allen MS, David B, Gibbs M,

Higham CFW, et al. 2018b. Complex history of dog (*Canis familiaris*) origins and translocations in the Pacific revealed by ancient mitogenomes. *Sci Rep.* 8(9130): 1-9. DOI: 10.1038/s41598-018-27363-8

Jones RK, Piper PJ, Wood R, Nguyen AT, Oxenham MF. 2019. The Neolithic transition in Vietnam: Assessing evidence for early pig management and domesticated dog. *J Archaeol Sci.* 28: 102042. DOI: 10.1016/j.jasrep.2019.102042

Kim KS, Lee SE, Jeong HW, Ha JH. 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. *Mol Phylogenet Evol.* 10(2): 210-220. DOI: 10.1006/mpev.1998.0513

Ko AMS, Chen CY, Fu Q, Delfin F, Li M, Chiu HL, Stoneking M, Ko YC. 2014. Early austronesians: Into and out of Taiwan. *Am J Hum Genet.* 94: 426-436. DOI: 10.1016/j.ajhg.2014.02.003

Larson G, Karlsson EK, Perri A, Webster MT, Ho SY, Peters J, Stahl PW, Piper PJ, Lingaas F, Fredholm M, et al. 2012. Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci U S A.* 109(23): 8878-8883. DOI: 10.1073/pnas.1203005109

Leathlobhair MN, Perri AR, Irving-Pease EK, Witt KE, Linderholm A, Haile J, Lebrasseur O, Ameen C, Blick J, Boyko AR, et al. 2018. The evolutionary history of dogs in the Americas. *Science* 361(6397): 81-85. DOI: 10.1126/science.aao4776

Leigh JW, Bryant D. 2015. POPART: Full-feature software for haplotype network construction. *Methods Ecol Evol.* 6(9): 1110-1116. DOI: 10.1111/2041-210X.12410

Lipson M, Skoglund P, Spriggs M, Valentin F, Bedford S, Shing R, Buckley H, Phillip I, Ward GK, Mallick S, et al. 2018. Population turnover in remote Oceania shortly after initial settlement. *Curr Biol.* 28(7): 1157-1165 e1157. DOI: 10.1016/j.cub.2018.02.051

Liu L, Chen X. 2012. The archaeology of China: from the late paleolithic to the early bronze age. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139015301

Lu H, Zhang J, Liu KB, 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. *Proc Natl Acad Sci U S A.* 106(18): 7367-7372. DOI: 10.1073/pnas.0900158106

NCBI [https://www.ncbi.nlm.nih.gov/].

Ollivier M, Tresset A, Frantz LAF, Brehard S, Balasescu A, Mashkour M, Boroneant A, Pionnier-Capitan M, Lebrasseur O, Arbogast RM, et al. 2018. Dogs accompanied humans during the Neolithic expansion into Europe. *Bio Letters.* 14(10). DOI: 10.1098/rsbl.2018.0286

Oskarsson MCR, Klütsch CFC, Boonyaprakob U, Wilton A, Tanabe Y, Savolainen P. 2012. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. *P Roy Soc B-Biol Sci.* 279(1730): 967-974. DOI: 10.1098/rspb.2011.1395

Paijmans JLA, Gilbert MTP, Hofreiter M. 2013. Mitogenomic analyses from ancient DNA. *Mol Phylogenet Evol.* 69(2): 404-416. DOI: 10.1016/j.ympev.2012.06.002

Pang JF, Kluetsch C, Zou XJ, Zhang AB, Luo LY, Angleby H, Ardalan A, Ekström C, Sköllermo A, Lundeberg J, et al. 2009. mtDNA data indicate a single origin for dogs south of yangtze river, less than 16,300 years ago, from numerous wolves. *Mol Biol Evol.* 26(12): 2849-2864. DOI: 10.1093/molbev/msp195

Peng MS, Fan L, Shi NN, Ning T, Yao YG, Murphy RW, Wang WZ, Zhang YP. 2015. DomeTree: A canonical toolkit for mitochondrial DNA analyses in domesticated animals. *Mol Ecol Resour.*15(5): 1238-1242. DOI: 10.1111/1755-0998.12386

Piper PJ. 2017. The origins and arrival of the earliest domestic animals in mainland and island South East Asia: a developing story of complexity In: Piper PJ, Matsumura H, Bulbeck D, editors. New Perspectives in Southeast Asian and Pacific Prehistory. Canberra: ANU Press. p. 251–273.

Posth C, Nagele K, Colleran H, Valentin F, Bedford S, Kami KW, Shing R, Buckley H, Kinaston R, Walworth M, et al. 2018. Language continuity despite population replacement in Remote Oceania. *Nat Ecol Evol.* 2: 731-740. DOI: 10.1038/s41559-018-0498-2

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1. 6. Computer program and documentation distributed by the author, website http://beast.bio.ed.ac.uk/Tracer.

Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E, Lundeberg J. 2004. A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc Natl Acad Sci U S A.* 101(33): 12387-12390. DOI: 10.1073/pnas.0401814101

Skoglund P, Posth C, Sirak K, Spriggs M, Valentin F, Bedford S, Clark GR, Reepmeyer C, Petchey F, Fernandes D, et al. 2016. Genomic insights into the peopling of the Southwest Pacific. *Nature* 538: 510-513. DOI: 10.1038/nature19844

Summerhayes G, Matisoo-Smith E, Mandui H, Allen J, Specht JR, Hogg N, McPherson S. 2010. Tamuarawai (EQS): An Early Lapita Site on Emirau, New Ireland, PNG. Journal of Pacific Archaeology 1(1): 62-75.

Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL, Germonpré MB, Sablin MV, López-Giráldez F, Domingo-Roura X, et al. 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* 342(6160): 871-874. DOI: 10.1126/science.1243650

Wang GD, Zhai W, Yang HC, Wang L, Zhong L, Liu YH, Fan RX, Yin TT, Zhu CL, Poyarkov AD, et al. 2016. Out of southern East Asia: The natural history of domestic dogs across the world. *Cell Res.* 26: 21-33. DOI: 10.1038/cr.2015.147

White JP. 2000. On the road of the winds: An archaeological history of the Pacific islands before European contact (review). *Asian Perspectives* 39(1): 189-190. DOI: 10.1353/asi.2000.0016Yang 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. *Proc Natl Acad Sci U S A.* 109(10): 3726-3730. DOI: 10.1073/pnas.1115430109

Zhang SJ, Wang GD, Zhang LL, Yin TT, Liu YH, Otecko NO, Wang M, Mao B, Savolainen P, Zhang YP. 2018. Genomic analysis of dingoes identifies genomic regions under reversible selection during domestication and feralization. bioRxiv.

Zheng HX, Yan S, Qin ZD, Wang Y, Tan JZ, Li H, Jin L. 2011. Major population expansion of East Asians began before Neolithic time: evidence of mtDNA genomes. *PLoS ONE* 6(10): e25835. DOI: 10.1371/journal.pone.0025835

Figure legends

Figure 1. **(A)** Locations of different archaeological sites. The red dots depict all sites from which we retrieved ancient sequences, where the number of specimens with mtDNA results are listed after the location. The inset shows the location of the samples within China. **(B)** A map of China, Southeast Asia, and the Pacific islands shows the location of the specimens and associated sub-haplogroups belonging to haplogroup A. The pre-colonial Pacific island samples are from New Zealand, New Guinea, and remote Oceania, etc. Colors represent the sub-haplogroup, orange is A1a; red is A1b, green is A2a/A2b; blue is A3; brown is A5; purple is A6. **(C)** The median-joining network of haplogroup A1b for the mitochondrial genomes of dogs. Haplotypes are represented by circles whose sizes are proportional to the number of individuals. The network was constructed from 117 dog samples. The black points represent hypothetical sequences that have not been sampled thus far. The ancient dog from Tianluoshan, Zhejiang (TLS_6957) is located in the center, indicating that it has a basal A1b haplogroup. **(D)** The simplified tree was based on all the ancient samples in this study. The A1b sub-tree was enlarged from supplementary fig. S3. The ancient samples of this study were marked in italic, bold and red, other ancient samples marked in bold, and the number after each ancient sample was the average date before present. The whole tree can be seen in supplementary fig. S3 and abbreviations in supplementary table S1. **(E)** A Bayesian skyline plot, showing two expansions, one around 7,500 years ago and another around 1,500 years ago. This analysis uses the same dataset from supplementary fig. S3.

Figure 2. Possible origin place (red area) of Australian dingoes, New Guinea Singing dogs (NGSDs), and pre-colonial Pacific dogs. The areas marked with different colors indicate where several dogs were sampled. Orange arrows indicate the proposed route for the introduction of dogs (sub-haplogroup A1b2) to the Pacific islands. In the box is a simplified phylogenetic tree for haplogroup A1b (based on fig. 1D), where the red branches include ancient samples from the Yangtze and Yellow River basins (AYYRB), the red stars at the nodes indicate a posterior greater than 0.90. Other abbreviations include: TLS_6957, Tianluoshan; SC, Southern China; SEA, Southeastern Asia; CK, Cook Islands; W_Chukotka, West Chukotka; Dingo_SE, southeastern group of dingoes; Dingo_NW, northwestern group of dingoes.

Figure 1

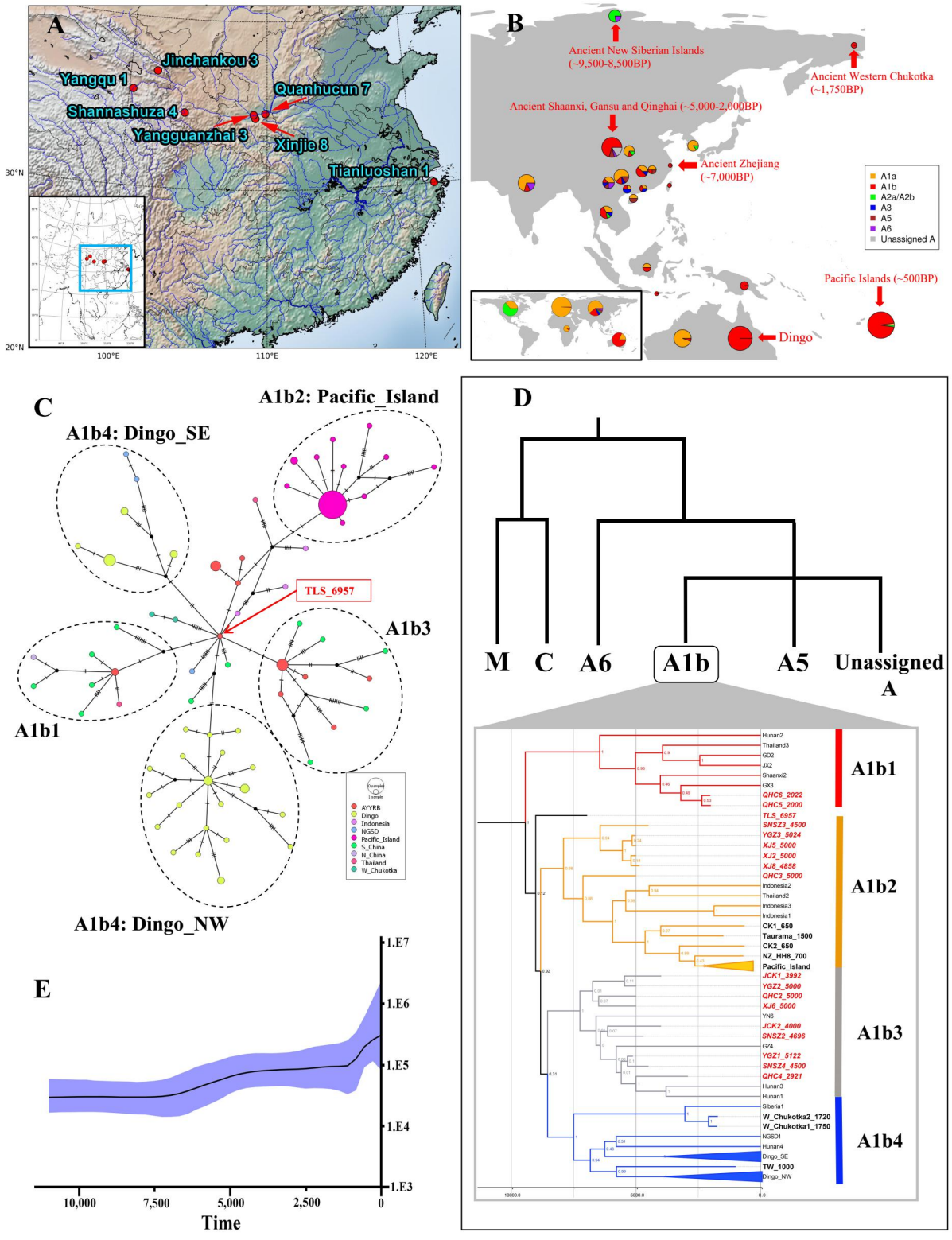

Figure 2