Journal of Genetics and Genomics 44 (2017) 163–170

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Original research

A cryptic mitochondrial DNA link between North European and West African dogs

Adeniyi C. Adeola [a, b, c], Sheila C. Ommeh [d], Jiao-Jiao Song [a, e], S. Charles Olaogun [f], Oscar J. Sanke [g], Ting-Ting Yin [a, b, c], Guo-Dong Wang [a, b, c], Shi-Fang Wu [a, b], Zhong-Yin Zhou [a, h], Jacqueline K. Lichoti [i], Bernard R. Agwanda [j], Philip M. Dawuda [k], Robert W. Murphy [a, l], Min-Sheng Peng [a, b, c, *], Ya-Ping Zhang [a, b, c, m, *]

[a] State Key Laboratory of Genetic Resources and Evolution, Yunnan Laboratory of Molecular Biology of Domestic Animals, Germplasm Bank of Wild Species, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[b] Sino-Africa Joint Research Center, Chinese Academy of Sciences, Kunming 650223, China
[c] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming 650204, China
[d] Animal Biotechnology Group, Institute of Biotechnology Research, Jomo Kenyatta University of Agriculture and Technology, Nairobi 00200, Kenya
[e] Institute of Health Sciences, Anhui University, Hefei 230601, China
[f] Department of Veterinary Medicine, University of Ibadan, Ibadan 200284, Nigeria
[g] Taraba State Ministry of Agriculture and Natural Resources, Jalingo 660221, Nigeria
[h] Department of Molecular and Cell Biology, School of Life Sciences, University of Science and Technology of China, Hefei 230026, China
[i] State Department of Livestock, Ministry of Agriculture Livestock and Fisheries, Nairobi 00100, Kenya
[j] Department of Zoology, National Museums of Kenya, Nairobi 00100, Kenya
[k] Department of Veterinary Surgery and Theriogenology, College of Veterinary Medicine, University of Agriculture Makurdi, Makurdi 972261, Nigeria
[l] Centre for Biodiversity and Conservation Biology, Royal Ontario Museum, Toronto M5S 2C6, Canada
[m] State Key Laboratory for Conservation and Utilization of Bio-Resources, Yunnan University, Kunming 650091, China

ARTICLE INFO

Article history:
Received 20 May 2016
Received in revised form
17 August 2016
Accepted 21 October 2016
Available online 24 November 2016

Keywords:
Dog
Mitochondrial DNA
Africa
Haplogroup D
Late-glacial dispersal
Hybridization

ABSTRACT

Domestic dogs have an ancient origin and a long history in Africa. Nevertheless, the timing and sources of their introduction into Africa remain enigmatic. Herein, we analyse variation in mitochondrial DNA (mtDNA) D-loop sequences from 345 Nigerian and 37 Kenyan village dogs plus 1530 published sequences of dogs from other parts of Africa, Europe and West Asia. All Kenyan dogs can be assigned to one of three haplogroups (matrilines; clades): A, B, and C, while Nigerian dogs can be assigned to one of four haplogroups A, B, C, and D. None of the African dogs exhibits a matrilineal contribution from the African wolf (*Canis lupus lupaster*). The genetic signal of a recent demographic expansion is detected in Nigerian dogs from West Africa. The analyses of mitochondrial genomes reveal a maternal genetic link between modern West African and North European dogs indicated by sub-haplogroup D1 (but not the entire haplogroup D) coalescing around 12,000 years ago. Incorporating molecular anthropological evidence, we propose that sub-haplogroup D1 in West African dogs could be traced back to the late-glacial dispersals, potentially associated with human hunter-gatherer migration from southwestern Europe.

1. Introduction

Domestic dogs (*Canis lupus familiaris*) have lived with humans for thousands of years in Africa (Gifford-Gonzalez and Hanotte, 2011). They have been kept for hunting, guarding, herding and as a source of meat (MacDonald and MacDonald, 2000). Archaeological records of dog remains dated from 6300 BCE to 1600 AD occur across Africa (Mitchell, 2015). The rock art of the eastern and western Sahara Desert suggests that dogs could have been brought across the Sahara in prehistoric times (Blench, 2000). Various African indigenous breeds, such as the Basenji (American Kennel Club, 2006), as well as numerous village dogs (Boyko et al., 2009), have evolved during this long history.

---

* Corresponding authors.
E-mail addresses: pengminsheng@mail.kiz.ac.cn (M.-S. Peng), zhangyp@mail.kiz.ac.cn (Y.-P. Zhang).

http://dx.doi.org/10.1016/j.jgg.2016.10.008

However, at some early sites in Africa, it has been difficult to differentiate the remains of dogs from those of side-striped jackals (*Canis adustus*) and African wild dogs (*Lycaon pictus*) (MacDonald and MacDonald, 2000). The scarcity of archaeological evidence and relevant research results in an ambiguous history as to when and how dogs were introduced into Africa (Gifford-Gonzalez and Hanotte, 2011).

Genetic analyses can reconstruct the history of dogs (Larson and Bradley, 2014; Wang et al., 2014). Analysis of mitochondrial DNA (mtDNA) and autosomal (SNPs and microsatellites/STRs) data revealed that African village dogs have a complex population structure and high genetic diversity (Boyko et al., 2009). In addition, African village dogs appear to be a mosaic of indigenous dogs that descended from early migrants to Africa, and non-native, breed-admixed individuals (Boyko et al., 2009). Other studies based on autosomal (Parker et al., 2004; Larson and Bradley, 2014; Fan et al., 2016) and Y chromosomal (Brown et al., 2011) markers showed Basenji to be a characteristic, ancient African breed, that separated long ago from other modern breeds (Parker et al., 2004; Larson and Bradley, 2014; Fan et al., 2016), and that potentially harbors primitive genetic characteristics (Parker et al., 2004; Vonholdt et al., 2010; Pickrell and Pritchard, 2012). In addition, genetic admixture is evident for the African indigenous dogs (including Basenji and village dogs) and gray wolf (Pickrell and Pritchard, 2012; Freedman et al., 2014; Fan et al., 2016; Wang et al., 2016). Consequently, modern African dogs may retain ancient genetic components of early domestic dogs and/or gray wolves.

To trace the source(s) of these ancient genetic components in African dogs, we focused on the village dogs from West Africa. Village dogs are less likely to have experienced the intense artificial selection resulting from the closed breeding practices, e.g., characteristic of pet dogs, and thus village dogs may retain signals of indigenous ancestry (Boyko et al., 2009; Shannon et al., 2015). It is expected that certain ancient genetic components detected in Basenji may also be maintained in West African village dogs, especially those that are Basenji-like (Fig. S1). Nevertheless, West African village dogs have been neglected in previous studies. Herein, we included an extensive sampling of Nigerian village dogs in our analyses of mtDNA sequences. Our results not only help further understand the mtDNA variation diversity in African dogs, but also provide novel insights into their introduction into Africa.

2. Results

2.1. Genetic diversity of mtDNA D-loop sequences

A total of 1912 D-loop sequences were analyzed, including 382 *de novo* (GenBank: KJ139039–KJ139383 and KU159539–KU159575) and 1530 downloaded from the GenBank (Boyko et al., 2009; Ardalan et al., 2011, 2015). Estimated indexes of genetic diversity in domestic dogs from Africa and neighboring Europe and West Asia (Table 1) revealed that African dogs had high haplotype diversity (0.950) but low nucleotide diversity (0.086) and this discrepancy is mainly caused by the excess of low-frequency mutations. Resampling to adjust for different sample sizes showed that African dogs had high haplotype diversity (0.957; Table S1) but low nucleotide diversity (0.116; Table S1). Within Africa, there was high haplotype diversity (0.948) in West African dogs, yet nucleotide diversity was the lowest (0.075). Both Tajima's *D* and Fu's *Fs* were negative in West African dogs ($P < 0.05$). As for other African dogs, Fu's *Fs* and Tajima's *D* did not deviate significantly from zero for populations from other regions in Africa. Therefore, West African dogs exhibit a unique pattern of genetic diversity.

2.2. Analyses of mtDNA D-loop sequences

NJ, ML and MP trees (Figs. S2–S4) were obtained for 324 haplotypes (Table S2). All haplotypes of Nigerian dogs were assigned to previously defined haplogroups A, B, C or D (Savolainen et al., 2002; Pang et al., 2009) and samples from Lamu, Kenya were assigned to A, B and C (Fig. 1). Although the low bootstrap values failed to provide additional statistical support for particular haplogroups (Figs. S2–S4), the assignment results were supported by haplogrouping with MitoToolPy (Table S3). The haplogroup distribution profile is shown in Fig. 1 and Table 1. Most African dogs (99.0%, 1104/1115) were allocated to haplogroups A, B or C, which exhibited universal Old World distributions (Savolainen et al., 2002; Pang et al., 2009; Ardalan et al., 2011; Oskarsson et al., 2012). Only village dogs from Calabar and Taraba in Nigeria, West Africa have haplotypes from haplogroup D (0.8%, 9/1115) (Tables 1 and S4). The haplogrouping of West African dogs seemed distinct from other African dog populations (Fig. 1 and Table 1), suggesting certain different population structures in African dogs (Table S5). African wolves (Gaubert et al., 2012) based on the matrilineal genetics

Table 1
Genetic diversity of dog population.

Population	Size[a]	nA (%)[b]	nB (%)[b]	nC (%)[b]	nD (%)[b]	nHT[c]	HTdiv (SD)[d]	πdiv (SD)[e]	*D*[f]	*Fs*[g]
Calabar[h]	57	54 (94.7)	1 (1.8)	–	2 (3.5)	25	0.937 (0.018)	0.107 (0.061)	1.662*	14.046*
Ibadan[h]	58	51 (87.9)	7 (12.1)	–	–	27	0.946 (0.015)	0.200 (0.109)	0.235	11.236*
Minna[h]	43	39 (90.7)	3 (7.0)	1 (2.3)	–	24	0.946 (0.020)	0.140 (0.078)	1.353*	12.622*
Ondo[h]	81	80 (98.8)	1 (1.2)	–	–	25	0.931 (0.013)	0.093 (0.054)	1.692*	12.316*
Taraba[h]	77	66 (85.7)	3 (3.9)	1 (1.3)	7 (9.1)	32	0.934 (0.017)	0.137 (0.074)	1.052	13.641*
Uyo[h]	29	29 (100)	–	–	–	11	0.889 (0.027)	0.185 (0.112)	0.929	2.75
Nigeria[h]	345	319 (92.5)	15 (4.3)	2 (0.6)	9(2.6)	75	0.947 (0.006)	0.076 (0.042)	1.479*	25.18*
Lamu/Kenya[h]	37	27 (73.0)	8 (21.6)	2 (5.4)	–	11	0.851 (0.033)	0.282 (0.152)	0.605	1.447
West Africa[i]	364	337 (92.6)	16 (4.4)	2 (0.5)	9 (2.5)	79	0.948 (0.006)	0.075 (0.041)	1.467*	25.188*
North Africa[j]	87	62 (71.3)	12 (13.8)	11 (12.6)	2 (2.3)	30	0.951 (0.009)	0.206 (0.108)	0.118	6.293
South Africa[j]	283	213 (75.3)	55 (19.4)	15 (5.3)	–	36	0.921 (0.007)	0.177 (0.092)	0.280	4.162
East Africa[i]	237	176 (74.3)	48 (20.3)	13 (5.5)	–	30	0.910 (0.009)	0.211 (0.110)	0.777	2.152
Madagascar[j]	144	84 (58.3)	35 (24.3)	25 (17.4)	–	26	0.897 (0.014)	0.251 (0.130)	1.143	0.812
Africa[i]	1115	872 (78.2)	166 (14.9)	66 (5.9)	11 (1.0)	128	0.950 (0.003)	0.086 (0.045)	0.938	24.076*
Europe[j]	452	298 (65.9)	94 (20.8)	34 (7.5)	26 (5.8)	52	0.926 (0.006)	0.166 (0.086)	0.308	9.348
West Asia[j]	345	200 (57.9)	113 (32.8)	22 (6.4)	10 (2.9)	51	0.894 (0.010)	0.161 (0.083)	0.087	10.588

[a]Total number of samples in each region. [b]Number of samples belonging to each of the haplogroups A, B, C and D (proportions of the total samples of the region). [c]Number of haplotypes. [d]Haplotype diversity (standard deviation). [e]Nucleotide diversity (standard deviation). [f]Tajima's *D* test of selective neutrality. [g]Fu's *Fs* test of selective neutrality. [h]*De novo* D-loop sequences in this study. [i]This is the combination of *de novo* and retrieved data. [j]Data retrieved from published studies (Boyko et al., 2009; Ardalan et al., 2011, 2015). Numbers with asterisks are statistically significant at 5% level.

clustered separately from the dogs and coyote outside as the outgroup (Figs. S5–S7).

The networks for haplogroups A, B, C and D (Fig. 2) display the relationships of dogs within Africa as well as with European and West Asian dogs. Of the 171 haplotypes identified among 1912 mtDNA D-loop sequences, 32 haplotypes from haplogroups A, B, and C were shared by African and European/West Asian dogs (Fig. 2 and Table S6). Twenty-one of these were shared by at least two populations within Africa. The shared lineages generally consisted of central “hubs” in the networks, suggesting that they represent the founders of African dogs. Most private haplotypes were differed by a single nucleotide difference from their most similar relative. In haplogroup A, some sequences of the West African village dogs showed a star-like pattern (Fig. S8), that is, a signature of rapid population expansion. This was further indicated by significantly negative values of both Tajima's D -1.425 ($P < 0.05$) and Fu's Fs -12.842 ($P < 0.05$) (Table S4).

2.3. Phylogenetic analyses of haplogroup D

The highly resolved phylogeny based on mitochondrial genomes depicted novel sub-haplogroup D1b, which contained dog samples 1643, 1646, 1662, and 1676 from Nigeria in West African (Fig. 3). Its sister group, D1a, contained Scandinavian dogs. Nigerian dog sample 1695 clustered with the Spanish GalgoEspañol (EU789655) and then with the Turkish Kangal (EU789654) forming D2. Basal lineages D3–D5 represented modern or ancient specimens of European gray wolves. Haplogroup D6 (KF661092) consisted of a 33.5 thousand years ago (kya) incipient dog from the Altai Mountains (Ovodov et al., 2011). The Bayesian inference tree (Fig. S9) and the phylogeny (Fig. 3) were congruent. The coalescence time for dog sub-haplogroup D1'2 was estimated at 12.2 kya (95% highest posterior density [HPD]: 15.3–9.4 kya) (Fig. S9 and Table 2). The divergence time for African (sub-haplogroup D1b) and Scandinavian dogs (D1a) was approximately 11.6 kya (95% HPD: 14.7–8.8 kya) (Table 2).

3. Discussion

Our analyses provide novel insights into the demographic history of African dogs. As compared with other African dogs, village dogs from Nigeria, West Africa, have high haplotype diversity due to the abundance of rare haplotypes (Fig. 2 and Table S6). The excess of low-frequency mutations defines large numbers of haplotypes but accounts for the lowest nucleotide diversity. This pattern is often attributed to a rapid demographic expansion from a small founder population (Avise, 2000). The signature of expansion is also consistent with the significantly negative values of Tajima's D and Fu's Fs tests ($P < 0.05$; Table 1) as well as the star-like pattern in the network of haplogroup A (Fig. 2 and Table S4). Thus, the observed pattern in the Nigerian dog population of West Africa is likely due to a recent demographic expansion.

All mtDNA lineages of African dogs belonging to haplogroups A–D (Fig. 1) support the fact that they originated from the grey wolf. The clustering of African wolves outside of the dogs is an indication that they made no contribution to the modern maternal gene pool of African dogs. At least from the matrilineal genetic perspective, no potential introgression was detected. Although recent findings reported gene flow from grey wolf to African village dogs (Wang et al., 2016), this discrepancy may be due to the limited

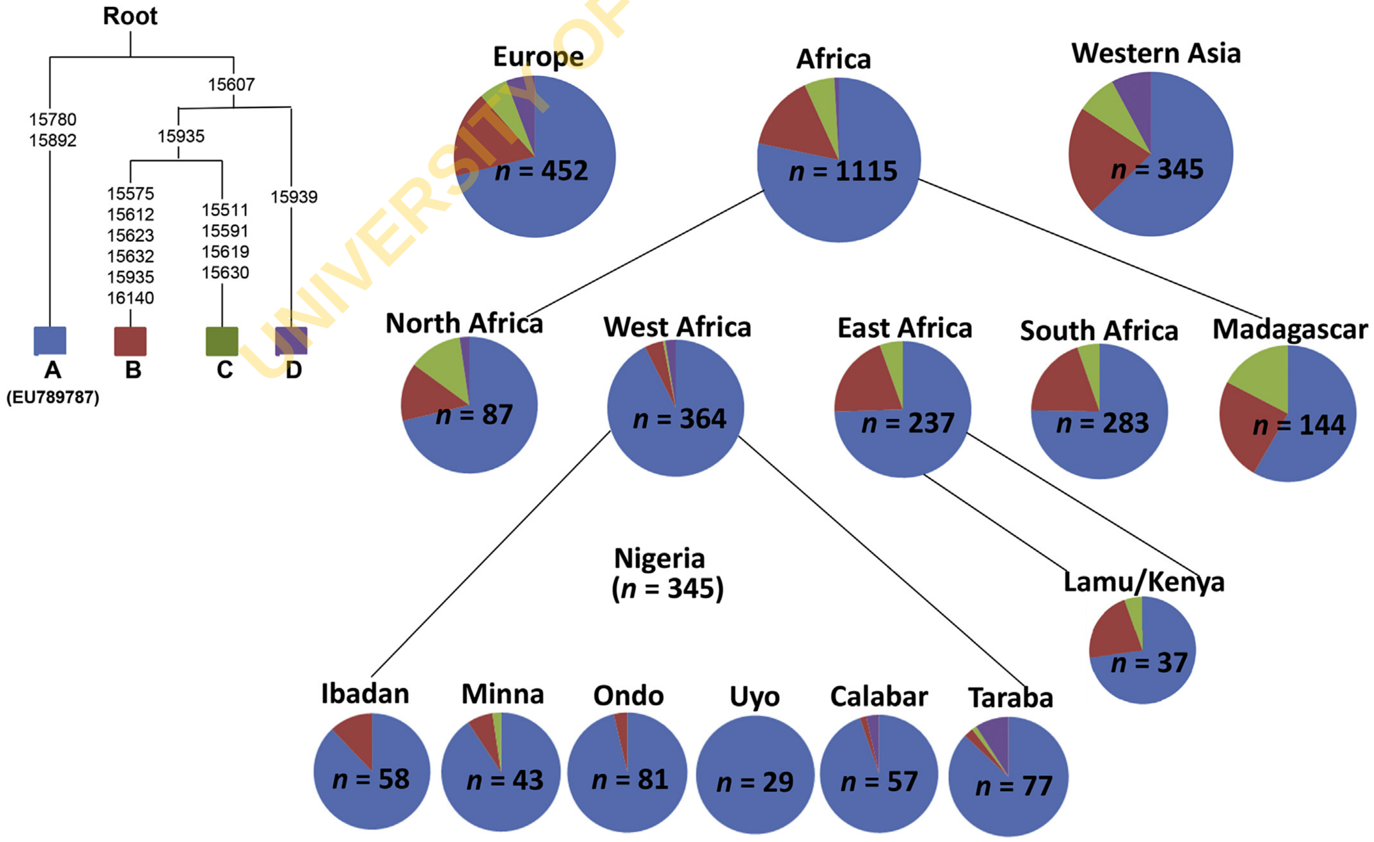

Fig. 1. Profiles of mtDNA haplogroups of dogs across Africa, Europe and West Asia. The diagnostic mutational motifs of 581-bp D-loop sequence for each of the haplogroups are shown on the branches of haplogroup tree. Pie diagrams show the proportion of individuals having haplogroups A (blue), B (red), C (green) and D (purple).

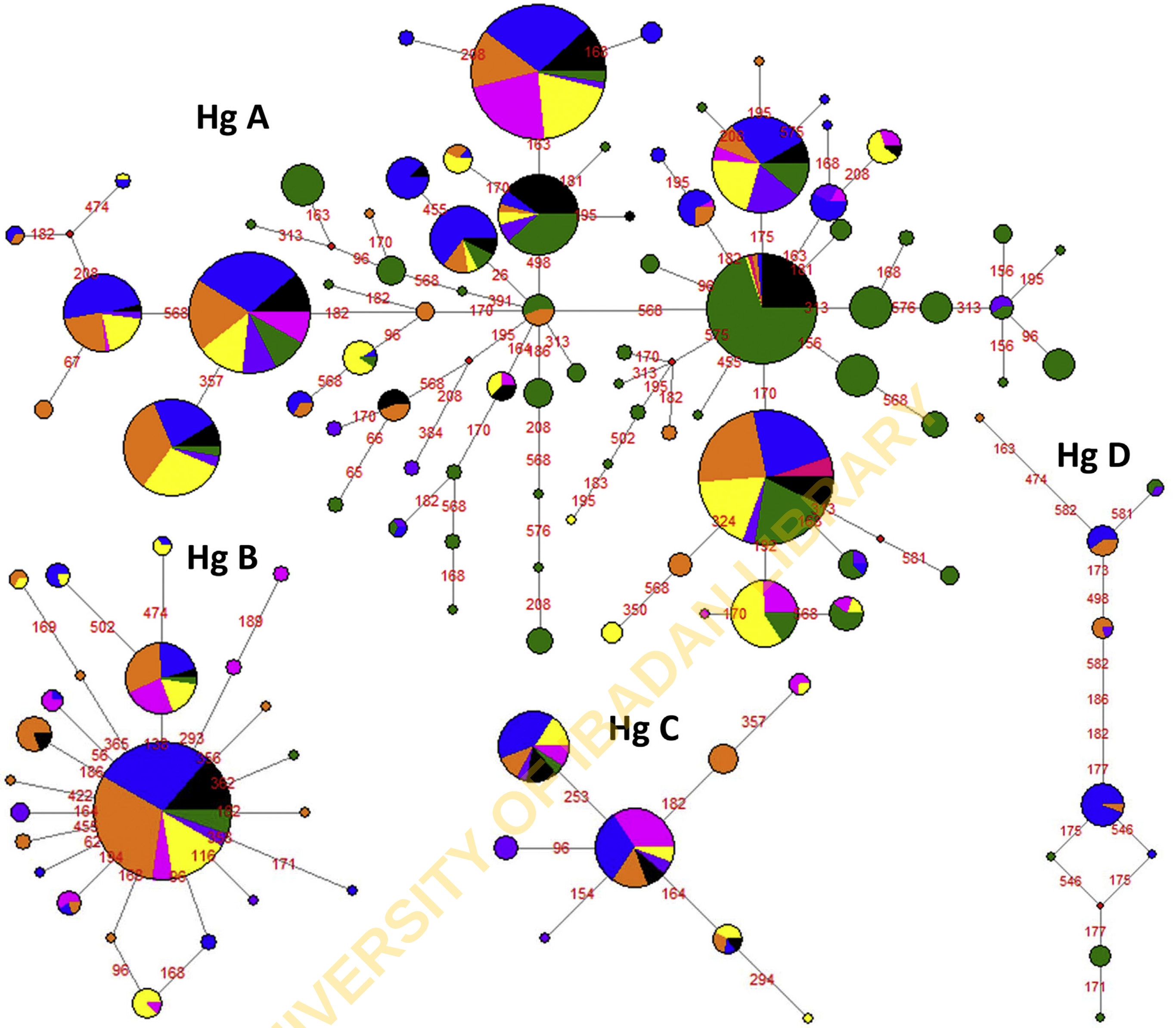

Fig. 2. Median-joining networks for haplogroups (Hg) A, B, C and D of dogs. Sizes of the circles are proportional to haplotype frequencies. The mutations are noted on each of the branches. Colors indicate the sampling locations: West Africa (green), North Africa (purple), East Africa (black), South Africa (yellow), Europe (blue), and West Asia (orange).

information from mtDNA as compared with autosomal markers that detected the gene flow. So far, no evidence has shown the occurrence of haplogroup D in East Asia and Southeast Asia, suggesting an origin outside eastern Asia (Pang et al., 2009; Ardalan et al., 2011; Klütsch et al., 2011). The presence of haplogroup D in West African dogs reflects a different evolutionary history. The high-resolution phylogeny based on mitochondrial genomes reveals that out-groups of sub-haplogroup D1′2 are distributed in modern (i.e., D4) or ancient specimens (i.e., D3 and D5) of European gray wolves (Fig. 3). This suggests that sub-haplogroup D1′2 of modern dogs originated in Europe (Fig. 3). The estimated timescale (15.3–9.4 kya, 95% HPD; Table 2) was after the Last Glacial Maximum (Clark et al., 2009), before the Neolithic diffusion into Europe from West Asia (Soares et al., 2010). The sequence of sub-haplogroup D2 in West Africa clustering with those from the Mediterranean and West Asia (Fig. 3) can be explained as a recent gene flow from those regions (Ardalan et al., 2011).

Most importantly, sub-haplogroup D1 contains dogs from Scandinavia and Nigeria (Fig. 3), which today live >5000 km apart. This cryptic mtDNA link between North European and West African dogs may be explained by a similar mtDNA link observed in human populations. The analysis of human mtDNA haplogroup U5b1b of European origin directly links North Europeans and Africans; the Saami of Scandinavia and the Yakut of Siberia formed a minor branch, as did the Berber of North Africa and the Fulbe of West Africa (Achilli et al., 2005). Both branches clustered with Spanish and Italian representatives into haplogroup U5b1b, with an estimated coalescence time of 8.6 ± 2.4 kya (Achilli et al., 2005). Taking into account standard error, the estimated timescale overlapped with the estimated time of coalescence of dog sub-haplogroup D1 with 95% HPD (14.7–8.8 kya; Table 2). Meanwhile, human haplogroup U5b1b frequencies in the Saami (47.6%) (Tambets et al., 2004) are comparable to sub-haplogroup D1 frequencies in their associated dogs (Swedish Lapphund: 100%, Finnish Lapphund:

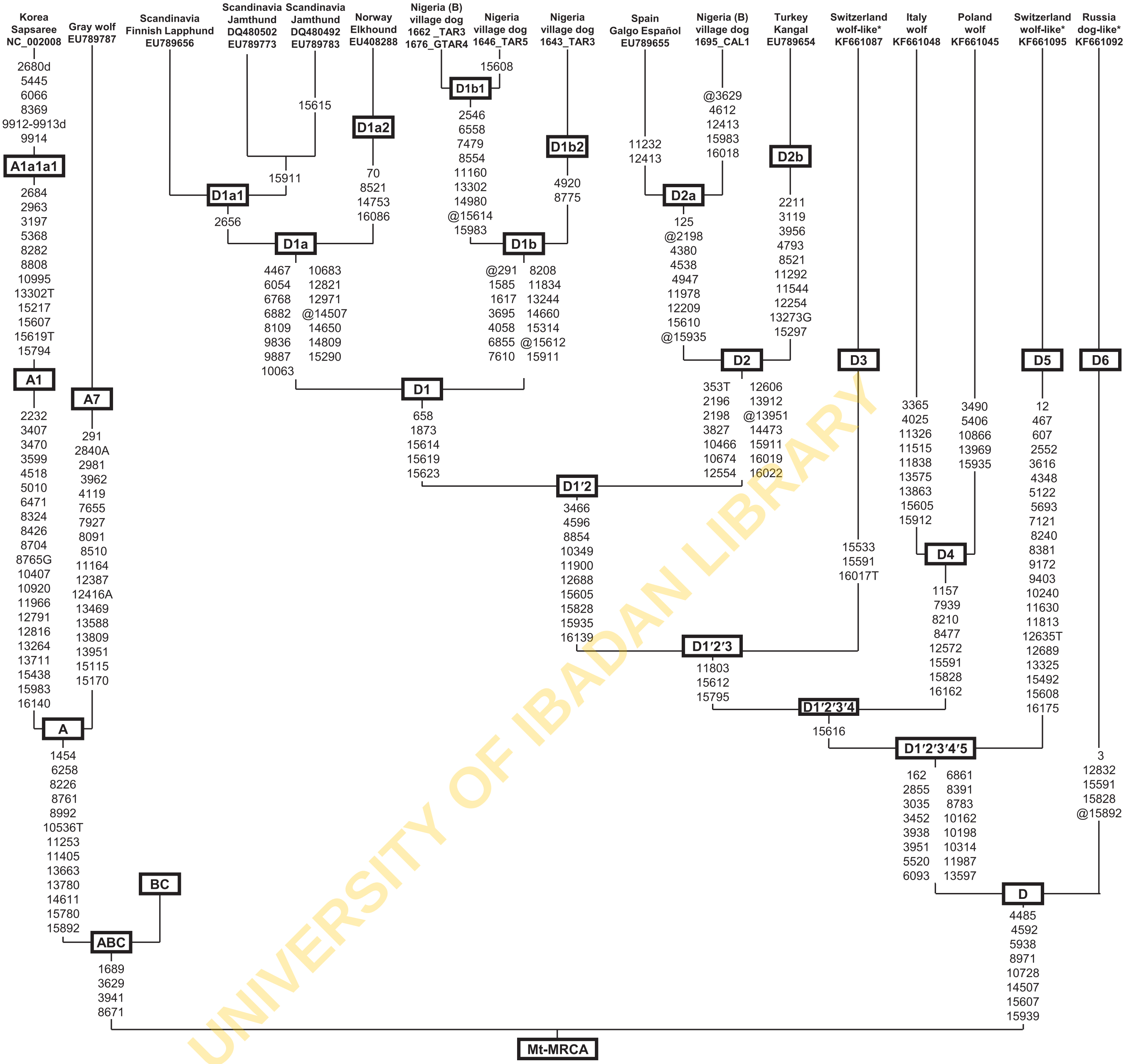

Fig. 3. The phylogeny (matrilineal genealogy) of haplogroup D in dogs. Nucleotide positions were scored relative to reference sequence EU789787 (Pang et al., 2009). Mutations are shown on the branches and transversions are denoted by suffixes. Deletions are indicated by "d" following the position. Mutations toward the bases of EU789787 are indicated with the prefix @. The undetermined mutations 15918d and 16193–19196d are not considered in the analyses. Basenji-like dogs are indicated by "(B)". Asterisks note samples of ancient DNA.

64.7%, and Lapponian Herder: 75%) (Klütsch et al., 2011). Further, U5b1b occurs in low frequencies in West Africa, such as 2.7% in Senegambia (Rosa et al., 2004) and Cameroon (Coia et al., 2005), similar to frequencies of D1 in indigenous dogs (about 2.0% in Nigeria). Thus, the synthesis of our discoveries and previous analyses reveals similar genographic patterns for human U5b1b and dog D1.

The mirror images of humans and dogs shed considerable light on the evolutionary history of dogs. Haplogroup U5b1b was associated with the late-glacial dispersals of human hunter-gatherers from the Franco-Cantabrian refuge area of southwestern Europe (Achilli et al., 2005). Some populations migrated northwards to Scandinavia and others dispersed southwards into Africa by crossing the Strait of Gibraltar (Achilli et al., 2005). The late-glacial "back-to-Africa" migrations were also supported by the analyses of genome-wide SNPs (Henn et al., 2012).

It is expected that sub-haplogroup D1 in dogs likely accompanied human populations in the late-glacial dispersals**.** As the distributions of haplogroup D in modern dogs have been proposed to result from dog-wolf hybridization (Pang et al., 2009; Ardalan et al., 2011; Klütsch et al., 2011), one possibility is that haplogroup D1′2 wolves might have scavenged (Wang et al., 2013) and afterwards co-migrated with humans. Dog-wolf hybridization might have occurred during this process. Indeed, the gene flow from wolves to

Table 2
Estimated timescale for mtDNA haplogroup D based on 12 protein-coding genes.

Haplogroup	Number of sequences	Median age (kya)[a]	95% HPD[b] (kya)
D1	10	11.6	14.7–8.8
D1a	6	1.5	2.8–0.5
D1b	4	3.9	6.4–1.9
D2	3	7.3	10.1–4.5
D1′2	13	12.2	15.3–9.4
D1′2′3	14	16.6	19.8–14.5
D4	2	6.4	10.6–2.8
D1′2′3′4′5	17	21.2	24.4–18.5

[a]kya, thousand years before present. [b]95% HPD boundaries for each of the haplogroup.

the African village dogs was uncovered in the most recent genomic analysis of re-sequencing data (Wang et al., 2016). Meanwhile, another possibility that should be considered is the co-migration of dogs and humans. According to the same genomic data, dogs from Africa and Europe shared a most recent common ancestor, and their origin was traced back to the migration from East Asia starting around 15,000 years ago. Dogs arrived in Europe about 10,000 years ago (Wang et al., 2016). And archaeological evidence exists for late-glacial (15,000–11,500 cal BP (calibrated years before present)) domestic dogs in the Franco-Cantabrian region (Pionnier-Capitan et al., 2011). Further analyses of the ancient DNA will be helpful in deciding between these possibilities.

Our results also provide some insights into deciphering ancient DNA. The previous analysis of mtDNA of 24 ancient dog samples from four Neolithic (about 5300–4500 BP) and four Medieval (about 1000–500 BP) sites in Scandinavia did not detect haplogroup D (Malmström et al., 2008). The coalescence time was estimated as about 2800–500 BP (95% HPD) (Table 2) for sub-haplogroup D1a in modern Scandinavian dogs, suggesting that the dogs of sub-haplogroup D1a might have arrived in Scandinavia after the Neolithic period. When just comparing the distribution of haplogroup D between 253 modern Scandinavian dogs (63/253, number of lineages having a subclade d1 haplotype compared with the total number of samples (Klütsch et al., 2011)) and nine Medieval samples (0/9, none of the nine medieval samples was assigned to haplogroup D (Malmström et al., 2008)), the difference was not significant ($P = 0.086$, Pearson's *Chi*-squared test). As a result, it is not surprising that dogs with subhaplogroup D1a were not found in 24 Scandinavian ancient samples.

4. Materials and methods

4.1. Animal sampling

Peripheral blood samples were collected from 345 village dogs with the consent of the owners from six different States in Nigeria (Fig. 1). Additional blood samples of 37 village dogs were sampled from Lamu County in Kenya after receiving authorization for research from the State Department of Livestock, Ministry of Agriculture, Livestock and Fisheries under permit number RES/POL/VOL.XXVII/162. All experimental procedures were approved by the Animal Care and Use Committee of Kunming Institute of Zoology (approval ID: SYDW-2013021). The methods were carried out in accordance with the approved guidelines.

4.2. Analysis of mtDNA D-loop sequences

Genomic DNA was extracted from whole blood by the standard phenol/chloroform method. Protocols for PCR amplification and sequencing of mtDNA D-loop are as previously reported (for details see Supplementary data; Savolainen et al., 2002). Both light and heavy chains were sequenced. Electropherograms for the sequences were visualized, edited and aligned by SEQMAN PRO of LASERGENE 7.1.0 (DNAStar, USA) against the reference sequence EU789787 (Pang et al., 2009). The variants in the D-loop sequences were scored relative to the reference sequence EU789787 (Pang et al., 2009) and then were classified into specific haplogroups (Wu et al., 2007; Miao et al., 2013; Shi et al., 2014; Song et al., 2016) by using MitoToolPy (http://www.mitotool.org/mp.html) (Peng et al., 2015).

For comparative D-loop mtDNA diversity study, 1530 DNA sequences (from African ($n = 733$), West Asian ($n = 345$), and European ($n = 452$)) from previous studies (Boyko et al., 2009; Ardalan et al., 2011, 2015) were retrieved as haplotype data from GenBank (GenBank: HQ261489, HQ452418–HQ452423, HQ452432–HQ452433, HQ452466–HQ452477, GQ375164–GQ375213, KR069086 and KP295480–KP295494). Downloaded haplotypes were assigned to their original groups (Boyko et al., 2009; Ardalan et al., 2011, 2015). All 1912 sequences (382 *de novo* and 1530 from GenBank range 581–764 bp) were aligned and trimmed to 581 bp for analysis. Nine D-loop sequences of 269 bp from the African wolf (i.e., Egyptian jackal; *C. l. lupaster*: (GenBank: JQ088675-JQ088676, JQ088678–JQ088684)), as well as one from gray wolf in Egypt (GenBank: JQ088677), were also downloaded from GenBank (Gaubert et al., 2012). Coyote (*C. latrans*) (GenBank: AF098153, AF098154, AF008158) (Vilà et al., 1997) was used as out-group.

The 1912 D-loop sequences (382 *de novo* (GenBank: KJ139039–KJ139383 and KU159539–KU159575) and 1530 from the GenBank) were initially aligned with CLUSTALX 2.1 (Larkin et al., 2007) and then checked by eye. Comparisons of sequences and identification of haplotypes were performed with DnaSP 5.10.1 (Librado and Rozas, 2009). Indexes for nucleotide and haplotype diversity (Nei, 1987), Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997) tests and population F_{ST} were calculated using Arlequin 3.5.1.3 (Excoffier and Lischer, 2010). The model of substitution and related parameters were determined by the Bayesian information criterion (Schwarz, 1978) in JMODELTEST 2.1.4 (Darriba et al., 2012). NJ, ML and MP trees were constructed from the 312 haplotype data haplotypes (Table S2) to visualize overall similarity using MEGA6 (Tamura et al., 2013) with TrN + G model of substitution which was the best model estimated by JMODELTEST 2.1.4. A median-joining network (Bandelt et al., 1999) for lineages A, B, C and D was constructed with NETWORK 4.6.11 (http://www.fluxus-engineering.com). To discern the potential genetic contribution from indigenous African wolves, NJ, ML and MP trees were constructed based on haplotypes defined by the 269-bp mtDNA D-loop segment.

4.3. Analysis of mitochondrial genomes

According to the NJ, ML and MP trees, five samples representing four D-loop haplotypes assigned to haplogroup D were selected for sequencing the entire mitochondrial genome. The PCR primers, protocols for PCR amplification and sequencing of mtDNA genome have been described previously (for details see Supplementary data; Pang et al., 2009). Both light and heavy chains were sequenced. We followed caveats for quality control in mtDNA genome study (Shi et al., 2014). The five mitochondrial genome sequences generated in this study have been deposited in GenBank (GenBank: KJ139384–KJ139388). Variants in the *de novo* mtDNA genomes were scored relative to the reference sequence EU789787 (Pang et al., 2009). Incorporating the published sequences (Björnerfeldt et al., 2006; Pang et al., 2009; Webb and Allard, 2009; Thalmann et al., 2013), the phylogeny of haplogroup D was constructed based on the parsimony-like method as described elsewhere (Wu et al., 2007; Achilli et al., 2008; Miao et al., 2013). The hierarchical nomenclature for each sub-haplogroup conformed to the dog and grey wolf mtDNA tree of DomeTree (http://www.dometree.org/trees/dog.htm) (Peng et al., 2015).

4.4. Estimation of divergence times

We performed a Bayesian analysis of mitochondrial genome sequences of haplogroup D at a global context and a total of 160 sequences were used (Thalmann et al., 2013). We only considered the 12 protein-coding genes in the light chain with the length of 10731 bp, whereas *ND6* gene in the heavy chain was excluded due to its unusual patterns of substitutions and nucleotide composition (Saccone et al., 1999). The 12 protein-coding genes were able to provide enough information for time estimation (Endicott and Ho, 2008). Previously reported approximate ages of 18 ancient specimens were used as tip-priors (Thalmann et al., 2013). The SRD06 model was selected using the Akaike information criterion (Akaike, 1974) in JMODELTEST 2.1.4. We adopted the strict molecular clock and Bayesian skygrid model (Gill et al., 2013) implemented in BEAST 1.8.0 (Drummond et al., 2012) as suggested by Thalmann et al. (2013). The Markov Chain Monte Carlo (MCMC) (Drummond et al., 2002) was run for 30 million iterations with sampling every 3000 steps and discarding the initial 100 samplings as burn-in. We evaluated the convergence of MCMC with TRACER 1.5 (http://tree.bio.ed.ac.uk/software/tracer/). The effective sample size for the parameter of interest was over 200. The Maximum Clade Credibility (MCC) tree was summarized using TREEANNOTATOR in the BEAST 1.8.0 package and visualized in FIGTREE 1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Data accessibility

Sequences are deposited in GenBank under accessions no. KJ139039–KJ139383, KU159539–KU159575 for D-loop and KJ139384–KJ139388 for mitochondrial genome sequences.

Acknowledgments

We appreciate all volunteers who assisted in sampling. We thank Ms. Ni-Ni Shi for technical assistance. This work was supported by Sino-Africa Joint Research Center, Chinese Academy of Sciences (Nos. SAJC201611 and SAJC201306) and the National Basic Research program (Nos. 2013CB835200 and 2013CB835202). The Youth Innovation Promotion Association, Chinese Academy of Sciences, provided support to M.S.P. and G.D.W. This work was supported by the Animal Branch of the Germplasm Bank of Wild Species, Chinese Academy of Sciences (the Large Research Infrastructure Funding). In addition, this work was also supported, in part, by a Visiting Professorship for Senior International Scientists from the Chinese Academy of Sciences, and manuscript preparation by a Natural Sciences and Engineering Research Council of Canada Discovery Grant A3148 to R.W.M.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jgg.2016.10.008.

References

Achilli, A., Olivieri, A., Pellecchia, M., Uboldi, C., Colli, L., Al-Zahery, N., Accetturo, M., Pala, M., Hooshiar Kashani, B., Perego, U.A., Battaglia, V., Fornarino, S., Kalamati, J., Houshmand, M., Negrini, R., Semino, O., Richards, M., Macaulay, V., Ferretti, L., Bandelt, H.J., Ajmone-Marsan, P., Torroni, A., 2008. Mitochondrial genomes of extinct aurochs survive in domestic cattle. Curr. Biol. 18, R157–R158.

Achilli, A., Rengo, C., Battaglia, V., Pala, M., Olivieri, A., Fornarino, S., Magri, C., Scozzari, R., Babudri, N., Santachiara-Benerecetti, A.S., Bandelt, H.J., Semino, O., Torroni, A., 2005. Saami and Berbers—an unexpected mitochondrial DNA link. Am. J. Hum. Genet. 76, 883–886.

Akaike, H., 1974. A new look at the statistical model identification. IEEE T. Autom. Contr. 19, 716–723.

American Kennel Club, 2006. The Complete Dog Book, twentieth ed. Ballantine Books, NY.

Ardalan, A., Kluetsch, C.F., Zhang, A.B., Erdogan, M., Uhlen, M., Houshmand, M., Tepeli, C., Ashtiani, S.R., Savolainen, P., 2011. Comprehensive study of mtDNA among Southwest Asian dogs contradicts independent domestication of wolf, but implies dog-wolf hybridization. Ecol. Evol. 1, 373–385.

Ardalan, A., Oskarsson, M.C., van Asch, B., Rabakonandriania, E., Savolainen, P., 2015. African origin for Madagascan dogs revealed by mtDNA analysis. R. Soc. Open Sci. 2, 140552.

Avise, J.C., 2000. Phylogeography: the History and Formation of Species. Harvard University Press, Cambridge.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Björnerfeldt, S., Webster, M.T., Vilà, C., 2006. Relaxation of selective constraint on dog mitochondrial DNA following domestication. Genome Res. 16, 990–994.

Blench, R.M., 2000. African minor livestock species. In: Blench, R.M., MacDonald, K.C. (Eds.), The Origins and Development of African Livestock: Archaeology, Genetics, Linguistics and Ethnography. Routledge, London, pp. 314–338.

Boyko, A.R., Boyko, R.H., Boyko, C.M., Parker, H.G., Castelhano, M., Corey, L., Degenhardt, J.D., Auton, A., Hedimbi, M., Kityo, R., Ostrander, E.A., Schoenebeck, J., Todhunter, R.J., Jones, P., Bustamante, C.D., 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. Proc. Natl. Acad. Sci. U. S. A. 106, 13903–13908.

Brown, S.K., Pedersen, N.C., Jafarishorijeh, S., Bannasch, D.L., Ahrens, K.D., Wu, J.T., Okon, M., Sacks, B.N., 2011. Phylogenetic distinctiveness of Middle Eastern and Southeast Asian village dog Y chromosomes illuminates dog origins. PLoS One 6, e28496.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W., McCabe, A.M., 2009. The Last glacial maximum. Science 325, 710–714.

Coia, V., Destro-Bisol, G., Verginelli, F., Battaggia, C., Boschi, I., Cruciani, F., Spedini, G., Comas, D., Calafell, F., 2005. Brief communication: mtDNA variation in North Cameroon: lack of Asian lineages and implications for back migration from Asia to sub-Saharan Africa. Am. J. Phys. Anthropol. 128, 678–681.

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

Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161, 1307–1320.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Endicott, P., Ho, S.Y., 2008. A Bayesian evaluation of human mitochondrial substitution rates. Am. J. Hum. Genet. 82, 895–902.

Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564–567.

Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A.S., Schweizer, R.M., Ramirez, O., Pollinger, J., Galaverni, M., Ortega Del-Vecchyo, D., Du, L., Zhang, W., Zhang, Z., Xing, J., Vilà, C., Marques-Bonet, T., Godinho, R., Yue, B., Wayne, R.K., 2016. Worldwide patterns of genomic variation and admixture in gray wolves. Genome Res. 26, 163–173.

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vila, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E.A., Marques-Bonet, T., Wayne, R.K., Novembre, J., 2014. Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. 10, e1004016.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Gaubert, P., Bloch, C., Benyacoub, S., Abdelhamid, A., Pagani, P., Djagoun, C.A., Couloux, A., Dufour, S., 2012. Reviving the African wolf *Canis lupus lupaster* in North and West Africa: a mitochondrial lineage ranging more than 6,000 km wide. PLoS One 7, e42740.

Gifford-Gonzalez, D., Hanotte, O., 2011. Domesticating animals in Africa: implications of genetic and archaeological findings. J. World Prehist. 24, 1–23.

Gill, M.S., Lemey, P., Fari, N.R., Rambaut, A., Shapiro, B., Suchard, M.A., 2013. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30, 713–724.

Henn, B.M., Botigue, L.R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J.K., Fadhlaoui-Zid, K., Zalloua, P.A., Moreno-Estrada, A., Bertranpetit, J., Bustamante, C.D., Comas, D., 2012. Genomic ancestry of North Africans supports back-to-Africa migrations. PLoS Genet. 8, e1002397.

Klütsch, C.F., Seppälä, E.H., Fall, T., Uhlén, M., Hedhammar, A., Lohi, H., Savolainen, P., 2011. Regional occurrence, high frequency but low diversity of mitochondrial

DNA haplogroup d1 suggests a recent dog-wolf hybridization in Scandinavia. Anim. Genet. 42, 100–103.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and clustal x version 2.0. Bioinformatics 23, 2947–2948.

Larson, G., Bradley, D.G., 2014. How much is that in dog years? The advent of canine population genomics. PLoS Genet. 10, e1004093.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

MacDonald, K.C., MacDonald, R.H., 2000. The origins and development of domesticated animals in arid West Africa. In: Blench, R.M., MacDonald, K.C. (Eds.), The Origins and Development of African Livestock: Archaeology, Genetics, Linguistics and Ethnography. Routledge, London, pp. 127–162.

Malmström, H., Vilà, C., Gilbert, M.T., Storå, J., Willerslev, E., Holmlund, G., Gotherstrom, A., 2008. Barking up the wrong tree: modern northern European dogs fail to explain their origin. BMC Evol. Biol. 8, 71.

Miao, Y.W., Peng, M.S., Wu, G.S., Ouyang, Y.N., Yang, Z.Y., Yu, N., Liang, J.P., Pianchou, G., Beja-Pereira, A., Mitra, B., Palanichamy, M.G., Baig, M., Chaudhuri, T.K., Shen, Y.Y., Kong, Q.P., Murphy, R.W., Yao, Y.G., Zhang, Y.P., 2013. Chicken domestication: an updated perspective based on mitochondrial genomes. Heredity 110, 277–282.

Mitchell, P., 2015. Did disease constrain the spread of domestic dogs (*Canis familiaris*) into Sub-Saharan Africa? Azania Archaeol. Res. Afr. 50, 92–135.

Nei, M., 1987. Molecular Evolutionary Genetics. Colombia University Press, NY.

Oskarsson, M.C., Klutsch, C.F., 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. Proc. Biol. Sci. 279, 967–974.

Ovodov, N.D., Crockford, S.J., Kuzmin, Y.V., Higham, T.F., Hodgins, G.W., van der Plicht, J., 2011. A 33,000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the Last Glacial Maximum. PLoS One 6, e22821.

Pang, J.F., Kluetsch, C., Zou, X.J., Zhang, A.B., Luo, L.Y., Angleby, H., Ardalan, A., Ekstrom, C., Skollermo, A., Lundeberg, J., Matsumura, S., Leitner, T., Zhang, Y.P., Savolainen, P., 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, 2849–2864.

Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A., Kruglyak, L., 2004. Genetic structure of the purebred domestic dog. Science 304, 1160–1164.

Peng, M.S., Fan, L., Shi, N.N., Ning, T., Yao, Y.G., Murphy, R.W., Wang, W.Z., Zhang, Y.P., 2015. DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals. Mol. Ecol. Resour. 15, 1238–1242.

Pickrell, J.K., Pritchard, J.K., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 8, e1002967.

Pionnier-Capitan, M., Bemilli, C., Bodu, P., Célérier, G., Ferrié, J., Fosse, P., Garciá, M., Vigne, J.D., 2011. New evidence for upper palaeolithic small domestic dogs in south-western Europe. J. Arch. Sci. 38, 2123–2140.

Rosa, A., Brehm, A., Kivisild, T., Metspalu, E., Villems, R., 2004. MtDNA profile of West Africa guineans: towards a better understanding of the Senegambia region. Ann. Hum. Genet. 68, 340–352.

Saccone, C., De Giorgi, C., Gissi, C., Pesole, G., Reyes, A., 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238, 195–209.

Savolainen, P., Zhang, Y.P., Luo, J., Lundeberg, J., Leitner, T., 2002. Genetic evidence for an East Asian origin of domestic dogs. Science 298, 1610–1613.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Shannon, L.M., Boyko, R.H., Castelhano, M., Corey, E., Hayward, J.J., McLean, C., White, M.E., Abi Said, M., Anita, B.A., Bondjengo, N.I., Calero, J., Galov, A., Hedimbi, M., Imam, B., Khalap, R., Lally, D., Masta, A., Oliveira, K.C., Perez, L., Randall, J., Tam, N.M., Trujillo-Cornejo, F.J., Valeriano, C., Sutter, N.B., Todhunter, R.J., Bustamante, C.D., Boyko, A.R., 2015. Genetic structure in village dogs reveals a Central Asian domestication origin. Proc. Natl. Acad. Sci. U.S. A 112, 13639–13644.

Shi, N.N., Fan, L., Yao, Y.G., Peng, M.S., Zhang, Y.P., 2014. Mitochondrial genomes of domestic animals need scrutiny. Mol. Ecol. 23, 5393–5397.

Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V., Bandelt, H.J., Torroni, A., Richards, M.B., 2010. The archaeogenetics of Europe. Curr. Biol. 20, R174–R183.

Song, J.J., Wang, W.Z., Otecko, N.O., Peng, M.S., Zhang, Y.P., 2016. Reconciling the conflicts between mitochondrial DNA haplogroup trees of *Canis lupus*. Forensic Sci. Int. Genet. 23, 83–85.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tambets, K., Rootsi, S., Kivisild, T., Help, H., Serk, P., Loogvali, E.L., Tolk, H.V., Reidla, M., Metspalu, E., Pliss, L., Balanovsky, O., Pshenichnov, A., Balanovska, E., Gubina, M., Zhadanov, S., Osipova, L., Damba, L., Voevoda, M., Kutuev, I., Bermisheva, M., Khusnutdinova, E., Gusar, V., Grechanina, E., Parik, J., Pennarun, E., Richard, C., Chaventre, A., Moisan, J.P., Barac, L., Pericic, M., Rudan, P., Terzic, R., Mikerezi, I., Krumina, A., Baumanis, V., Koziel, S., Rickards, O., De Stefano, G.F., Anagnou, N., Pappa, K.I., Michalodimitrakis, E., Ferak, V., Furedi, S., Komel, R., Beckman, L., Villems, R., 2004. The western and eastern roots of the Saami—the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. Am. J. Hum. Genet. 74, 661–682.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpre, M.B., Sablin, M.V., Lopez-Giraldez, F., Domingo-Roura, X., Napierala, H., Uerpmann, H.P., Loponte, D.M., Acosta, A.A., Giemsch, L., Schmitz, R.W., Worthington, B., Buikstra, J.E., Druzhkova, A., Graphodatsky, A.S., Ovodov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfli, K.P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., Wayne, R.K., 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science 342, 871–874.

Vilà, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J., Wayne, R.K., 1997. Multiple and ancient origins of the domestic dog. Science 276, 1687–1689.

Vonholdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E., Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D.A., Auton, A., Reynolds, A., Bryc, K., Brisbin, A., Knowles, J.C., Mosher, D.S., Spady, T.C., Elkahloun, A., Geffen, E., Pilot, M., Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.L., Zhang, Y.P., Bustamante, C.D., Ostrander, E.A., Novembre, J., Wayne, R.K., 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464, 898–902.

Wang, G.D., Zhai, W., Yang, H.C., Fan, R.X., Cao, X., Zhong, L., Wang, L., Liu, F., Wu, H., Cheng, L.G., Poyarkov, A.D., Poyarkov Jr., N.A., Tang, S.S., Zhao, W.M., Gao, Y., Lv, X.M., Irwin, D.M., Savolainen, P., Wu, C.I., Zhang, Y.P., 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. Nat. Commun. 4, 1860.

Wang, G.D., Xie, H.B., Peng, M.S., Irwin, D., Zhang, Y.P., 2014. Domestication genomics: evidence from animals. Annu. Rev. Anim. Biosci. 2, 65–84.

Wang, G.D., Zhai, W., Yang, H.C., Wang, L., Zhong, L., Liu, Y.H., Fan, R.X., Yin, T.T., Zhu, C.L., Poyarkov, A.D., Irwin, D.M., Hytonen, M.K., Lohi, H., Wu, C.I., Savolainen, P., Zhang, Y.P., 2016. Out of southern East Asia: the natural history of domestic dogs across the world. Cell Res. 26, 21–33.

Webb, K.M., Allard, M.W., 2009. Mitochondrial genome DNA analysis of the domestic dog: identifying informative SNPs outside of the control region. J. Forensic Sci. 54, 275–288.

Wu, G.S., Yao, Y.G., Qu, K.X., Ding, Z.L., Li, H., Palanichamy, M.G., Duan, Z.Y., Li, N., Chen, Y.S., Zhang, Y.P., 2007. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. Genome Biol. 8, R245.