Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 37 (2005) 804–814
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Mitochondrial diversity and phylogeographic structure of Chinese domestic goats

Shan-Yuan Chen [a,b], Yan-Hua Su [a,b], Shi-Fang Wu [b], Tao Sha [a], Ya-Ping Zhang [a,b,*]

[a] *Laboratory for Conservation and Utilization of Bio-resource, Yunnan University, Kunming 650091, China*
[b] *Laboratory of Cellular and Molecular Evolution, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*

Received 8 February 2005; revised 18 May 2005; accepted 7 June 2005
Available online 26 July 2005

Abstract

China has numerous native domestic goat breeds, but so far there has been no extensive study on genetic diversity, population demographic history, and origin of Chinese goats. Here, we examined the genetic diversity and phylogeographic structure of Chinese domestic goats by determining a 481-bp fragment of the first hypervariable region of mitochondrial DNA (mtDNA) control region from 368 individuals representing 18 indigenous breeds. Phylogenetic analyses revealed that there were four mtDNA lineages (A–D) identified in Chinese goats, in which lineage A was predominant, lineage B was moderate, and lineages C and D were at low frequency. These results further support the multiple maternal origins of domestic goats. The pattern of genetic variation in goat mtDNA sequences indicated that the two larger lineages A and B had undergone population expansion events. In a combined analysis of previously reported sequences and our sequences belonging to lineage B, we detected two subclades, in which one was unique to eastern Asia and another was shared between eastern and southern Asia. A larger genetic variation in eastern Asia than southern Asia and the pattern of phylogeographic variation in lineage B suggest that at least one subclade of lineage B originated from eastern Asia. There was no significant geographical structuring in Chinese goat populations, which suggested that there existed strong gene flow among goat populations caused by extensive transportation of goats in history.

Keywords: Domestic goats; mtDNA; Origin; Phylogeography

1. Introduction

Domestic goats (*Capra hircus*) provide a full range of useful products to human society (e.g., meat, milk, and fiber), and this making the goat one of the most useful animals that humans have ever domesticated (Porter, 1996). Goats have performed agricultural, economic, cultural, and even religious roles from very early times in human civilization (Joshi et al., 2004). Evidence from archaeological studies showed that goats were probably first domesticated in the Fertile Crescent region of the Near East around 10,000 years ago (Porter, 1996; Pringle, 1998; Zeder and Hesse, 2000), whereas some studies suggested that a second independent domestication in Pakistan gave rise to the Cashmere-like goat breeds (Meaow, 1993; Porter, 1996). Though that goat has played an important role since Neolithic times around the globe (Legge, 1996; Porter, 1996; Pringle, 1998), the origins of domestic goats are still not well understood.

Mitochondrial DNA (mtDNA) has represented the most informative genomic element for untangling the origins of domestic animals (MacHugh and Bradley, 2001). Up to now, mitochondrial sequences have been widely used to study the origin of cattle (Bradley et al., 1996; Loftus et al., 1994; Troy et al., 2001), pig

* Corresponding author. Fax: +86 871 5032804.
E-mail address: zhangyp1@263.net.cn (Y.-P. Zhang).

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.06.014

(Giuffra et al., 2000), sheep (Hiendleder et al., 1998, 2002), horse (Vilà et al., 2001), dog (Savolainen et al., 2002; Vilà et al., 1997), donkey (Beja-Pereira et al., 2004), and goat (Joshi et al., 2004; Luikart et al., 2001; Mannen et al., 2001; Sultana et al., 2003). Previous studies on domestic goats identified at least four major mtDNA lineages (Joshi et al., 2004; Luikart et al., 2001; Sultana et al., 2003). Lineage A is the most diverse and widely distributed across all continents. Lineage B is confined to eastern and southern Asia, including Mongolia, Laos, Malaysia, Pakistan, and India. Lineage C is present in low frequencies in Mongolia, Switzerland, Slovenia, Pakistan, and India. Finally, lineage D is rare and only observed in Pakistani and Indian local goats.

The time since divergence among these four lineages (more than 200,000 years ago) far predated the time of domestication around 10,000 years ago (Joshi et al., 2004; Luikart et al., 2001; Sultana et al., 2003). These results combined with archaeological findings suggested that domestic goats have multiple maternal origins (Porter, 1996; Pringle, 1998; Zeder and Hesse, 2000). However, an additional sampling, especially in regions likely to hold the minor mtDNA lineages, will improve our understanding where the goats originated (MacHugh and Bradley, 2001).

China is one among those regions from where goats were never extensively genetically assessed. This country has numerous local types of all-purpose native goat breeds that range from the high Qinghai-Tibet Plateau with very dry, cold, and harsh environment to the low lands of the Southern China with warm, humid, and mild environment. Due to the large environmental gradient in China, it is natural that there are some phenotypic differences among different native breeds. It was hypothesized that the ancient southwestern mountain area in China might be a domestication center of goat (Sieh, 1985). It is expected that molecular study on Chinese native goats will shed light on understanding the origin and process of goat domestication and will contribute to the resolution of globe goat phylogeny. The previous globe survey conducted by Luikart et al. (2001) involved only two individuals of two breeds from China. Here, we determined the first hypervariable segment (HVSI) of the mtDNA control region sequences of 368 individuals from 18 Chinese native goat breeds and used these data to investigate genetic diversity, origin, and phylogeography of Chinese goat populations.

2. Materials and methods

2.1. Sample collection and DNA extraction

Blood samples from 368 goats (*Capra hircus*) representing 18 pure indigenous Chinese breeds were collected from small remote villages belonging to 10 provinces or regions of China (Fig. 1) and stored at −70 °C. Latter in our laboratory, total genomic DNA

Fig. 1. Geographical and lineage distribution of the Chinese goat breeds sampled and the distribution frequency of lineage B in all Asian countries. The area of the circle is proportional to sample size.

Table 1
The representation of map numbers, breed names, geographic regions, and sample sizes, number of haplotypes, haplotype diversity, and nucleotide diversity for each breed used in this study

Map No.	Breed	Geographic distribution	Number of individuals	Number of haplotypes	Haplotype diversity (±SE)	Nucleotide diversity (±SE)
1	Heitou	Southwestern region (SW), Yunnan	28	16	0.9418 ± 0.0250	0.0279 ± 0.0144
2	Wujiao	Southwestern region (SW), Yunnan	26	14	0.9323 ± 0.0280	0.0351 ± 0.0180
3	Gui Dairy	Southwestern region (SW), Yunnan	18	15	0.9804 ± 0.0243	0.0315 ± 0.0165
4	Longling Yellow	Southwestern region (SW), Yunnan	24	10	0.9130 ± 0.0285	0.0352 ± 0.0181
5	Yunling Black	Southwestern region (SW), Yunnan	53	25	0.9398 ± 0.0179	0.0329 ± 0.0165
6	Chengdu Brown	Southwestern region (SW), Sichuan	8	5	0.8571 ± 0.1083	0.0355 ± 0.0202
7	Tibetan	Southwestern region (SW), Sichuan	7	5	0.8571 ± 0.1371	0.0490 ± 0.0282
8	Chuandong White	Southwestern region (SW), Sichuan	14	6	0.7473 ± 0.1114	0.0228 ± 0.0124
9	Guizhou White	Southwestern region (SW), Guizhou	14	7	0.9011 ± 0.0465	0.0346 ± 0.0184
10	Guizhou Black	Southwestern region (SW), Guizhou	9	6	0.9167 ± 0.0725	0.0293 ± 0.0165
11	Duan	Southern region (S), Guangxi	14	11	0.9670 ± 0.0366	0.0292 ± 0.0156
12	Longlin	Southern region (S), Guangxi	10	7	0.9111 ± 0.0773	0.0391 ± 0.0214
13	Leizhou	Southern region (S), Guangdong	15	8	0.8762 ± 0.0595	0.0295 ± 0.0157
14	Shaanan White	Northern region (N), Shaanxi	16	7	0.8417 ± 0.0595	0.0305 ± 0.0162
15	Yichang White	Central region (C), Hubei	13	7	0.8974 ± 0.0537	0.0329 ± 0.0177
16	Matou	Central region (C), Hunan	12	3	0.7121 ± 0.0691	0.0321 ± 0.0174
17	Xinjiang	Northwestern region (NW), Xinjiang	40	24	0.9744 ± 0.0099	0.0243 ± 0.0125
18	Liaoning Cashmere	Northeastern region (NE), Liaoning	47	21	0.9204 ± 0.0230	0.0159 ± 0.0084

was extracted from blood by standard phenol–chloroform extraction method. An endeavor was made to collect samples from unrelated individuals based on the information provided by the owners and local farmers. The detailed information of breeds, geographic regions, and sample sizes are given in Table 1.

2.2. PCR amplification and sequencing

The first hypervariable segment (HVSI) of the mtDNA control region sequences was amplified and sequenced. The primers CAP-F (5′-CGTGTATGCA AGTACATAC-3′) and CAP-R (5′-CTGATTAGTCA TTAGTCCATC-3′) (Luikart et al., 2001) were used to amplify a 579-bp DNA fragment. The complete mitochondrial cytochrome *b* gene sequences (1140 bp) of seven goats from the four *C. hircus* lineages (two individuals with different control region sequences from each lineages A–C and one individual from lineage D) were amplified by the primers L14724V (5′-ATGATATGAAAAACCATCG TTG-3′) and H15915V (5′-TCTCCTTCTCTGGTT TACAAGAC-3′) (Luikart et al., 2001). PCR amplifications were conducted in a 50 μl volume containing 5 μl of 10× reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 μM each primer, 1.5 U *Taq* DNA polymerase (TaKaRa Biosystems), and approximately 30 ng genomic DNA. The PCR mixture underwent 3 min at 95 °C, 35 cycles 50 s at 94 °C, 1 min at 51 °C (HVSI) or 53 °C (cytochrome *b*), and 1 min at 72 °C, and 5 min at 72 °C. PCR products were purified by using Watson PCR Purification Kit (Watson BioTechnologies, Shanghai).

A 481-bp HVSI fragment of the PCR product was sequenced using two internal primers CAP-FI (5′-TCCATATAACGCGGACATAC-3′) and CAP-RI (5′-ATGGCCCTGAAGAAAGAAC-3′) (Luikart et al., 2001) with the ABI PRISM BigDye Terminator v3.0 Ready Reaction Cycle Sequencing Kit in 5 μl volumes (Applied Biosystems). The seven sequences of cytochrome *b* were also sequenced using the primers L14724V, H15915V, and the two internal primers CY TB-FI (5′-CGCCTTCCACTTTATCCTCCC-3′) and CYTB-RI (5′-GTCTGATGGAATTCCTGTGGG-3′) with the same Sequencing Kit in 5 μl volumes. The cycling profile for the sequencing reaction consisted of 25 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 4 min. Isopropanol-purified sequencing products were analyzed on the ABI PRISM 3700 DNA Analyzer (Applied Biosystems). DNA sequences were edited using DNASTAR5.0 (DNASTAR).

2.3. Data analysis

All 368 sequences of mtDNA HVSI region, seven sequences from seven wild goat species (GenBank Accession Nos. AJ317864, AJ317870–AJ317872, AJ317874, AJ317875, and AB110590), and three sequences belonging to lineage D from Pakistan goats (GenBank Accession Nos. AB110587–AB110589) were aligned using the Clustal W program (Thompson et al., 1994).

To investigate the origin of previously defined lineage B of domestic goats (Luikart et al., 2001), all reported sequences belonging to the lineage B of domestic goats were also aligned and included in our analyses (Table 2).

A neighbor-joining (NJ) tree was constructed using PAUP* 4.0b10 (Swofford, 2003) under the HKY+I+Γ model optimized for our data set (146 haplotypes of Chinese goats, seven sequences of wild goat species, and three sequences belonging to lineage D from Pakistan goats) by Modeltest 3.06 (Posada and Crandall, 1998)

Table 2
Number and proportion of individuals, number of haplotypes, and unique haplotypes for lineage B in different Asian countries

Country	Number of individuals (proportion, %)	Number of haplotypes (unique types)	Reference
China	92 (25)	25 (23)	This study
Mongolia	1 (6.7)	1 (0)	Luikart et al. (2001)
Laos	8 (66.7)	5 (4)	Mannen et al. (2001)
Malaysia	12 (75)	4 (3)	Luikart et al. (2001)
Pakistan	15 (19.2)	5 (4)	Luikart et al. (2001) and Sultana et al. (2003)
India	6 (1.6)	5 (4)	Luikart et al. (2001) and Joshi et al. (2004)
Total	134	40	

based on the likelihood ratio tests. Settings for HKY+I+Γ model were as follows: base frequencies (A 0.3347, C 0.2283, G 0.1359, and T 0.3010); transition/transversion ratio 36.4695; proportion of invariable sites 0.4028; and the shape parameter of the gamma distribution 0.4632. Support for nodes was assessed using non-parametric bootstrapping (Felsenstein, 1985) with 1000 replicates. Bayesian phylogenetic analysis was performed using MrBayes 3.04 (Huelsenbeck and Ronquist, 2001). Two MCMC runs starting from different random trees were completed using the HKY+I+Γ model, each with five million generations and four chains. All sample points prior to reaching convergence were discarded as burn-in samples. The remaining samples were used to generate a majority rule consensus tree, where percentage of samples recovering any particular clade represented the clade's posterior probability (Huelsenbeck and Ronquist, 2001).

Median-joining network (Bandelt et al., 1999) was drawn using the program Network 4.0 (www.fluxus-engineering.com) to investigate the possible relationships among haplotypes of Chinese domestic goats. Haplotype diversity and its standard error (SE), nucleotide diversity, mismatch analysis (Schneider and Excoffier, 1999), and analysis of molecular variance (AMOVA) (Excoffier et al., 1992) were computed using the software ARLEQUIN version 2.000 (http://anthropologei.unige.ch/arlequin) (Schneider et al., 2000). A Mantel test was also carried out in ARLEQUIN version 2.000 to estimate the correlation between geographical distances (geographical distances between the center of distribution of sampling breeds) and genetic distances between breeds.

According to Luikart et al. (2001) and Beja-Pereira et al. (2004), a molecular-clock likelihood-ratio test (for heterogeneity in substitution rates) was conducted using the seven Chinese goats and two domestic sheep (*Ovis aries*) (GenBank Accession Nos. AF034730 and X56284) cytochrome *b* sequences analyzed for the 380 nucleotides at third codon positions. The test was not significant ($P > 0.05$), allowing us to use the amount of divergence between sheep and goat sequences to estimate the approximate time to the most recent common ancestor (TMRCA) of Chinese goats. All sequences are deposited in GenBank (Accession Nos. DQ089106–DQ089480).

3. Results

3.1. MtDNA variation in Chinese domestic goats

There were no insertions/deletions in 368 sequences of HVSI of the control regions. The HVSI sequences were highly polymorphic. Our 368 sequences gave 146 different haplotypes with 119 variable sites defined (only one mutation site was a transversion). The largest haplotype group consisted of 20 individuals, three haplotype groups included more than 10 individuals, and two haplotypes included 10 individuals. The number of haplotypes detected in each breed varies from 3 to 25, and haplotype diversity values from 0.7121 ± 0.0691 in the Matou to 0.9804 ± 0.0243 in the Gui Dairy (Table 1). The Tibetan breed displays the highest nucleotide diversity value (0.0490 ± 0.0282), while in Liaoing Cashmere breed holds lowest one (0.0159 ± 0.0084) (Table 1).

When compared mtDNA variation in all Chinese domestic goats with those of Indian and Pakistani domestic goats, the level of diversity within China (nucleotide diversity, $\pi = 0.0352 \pm 0.0174$) is higher than that of India ($\pi = 0.0212 \pm 0.0108$) and slightly lower than that of Pakistan ($\pi = 0.0365 \pm 0.0184$) based on reanalysis of data of previous studies (Joshi et al., 2004; Sultana et al., 2003). Interestingly, the level of diversity within China is also higher than that of goat samples across the Old World ($\pi = 0.0278 \pm 0.0139$) (Luikart et al., 2001).

3.2. Phylogenetic tree construction

Comparing these haplotypes with previously well-defined lineages A, B, C (Luikart et al., 2001), and D (Joshi et al., 2004; Sultana et al., 2003), the NJ tree constructed with 146 haplotypes of all Chinese goats and seven wild goat species clearly showed that Chinese domestic goats were divided into four distinct mtDNA lineages A–D (Fig. 2). The topology of Bayesian tree of the same data set was very similar to that of NJ tree except for minor incongruence among internal branches. The lineages A, B, and C included 117, 25, and 3 haplotypes comprising 269, 92, and 6 individuals, respectively, whereas the lineage D included a single haplotype, only presented in one individual. The median-joining network (Fig. 3), at a fine-scale, showed the relationships among 146 haplotypes and revealed substantial diversity in Chinese goat breeds. As expected, the network also clearly identified four

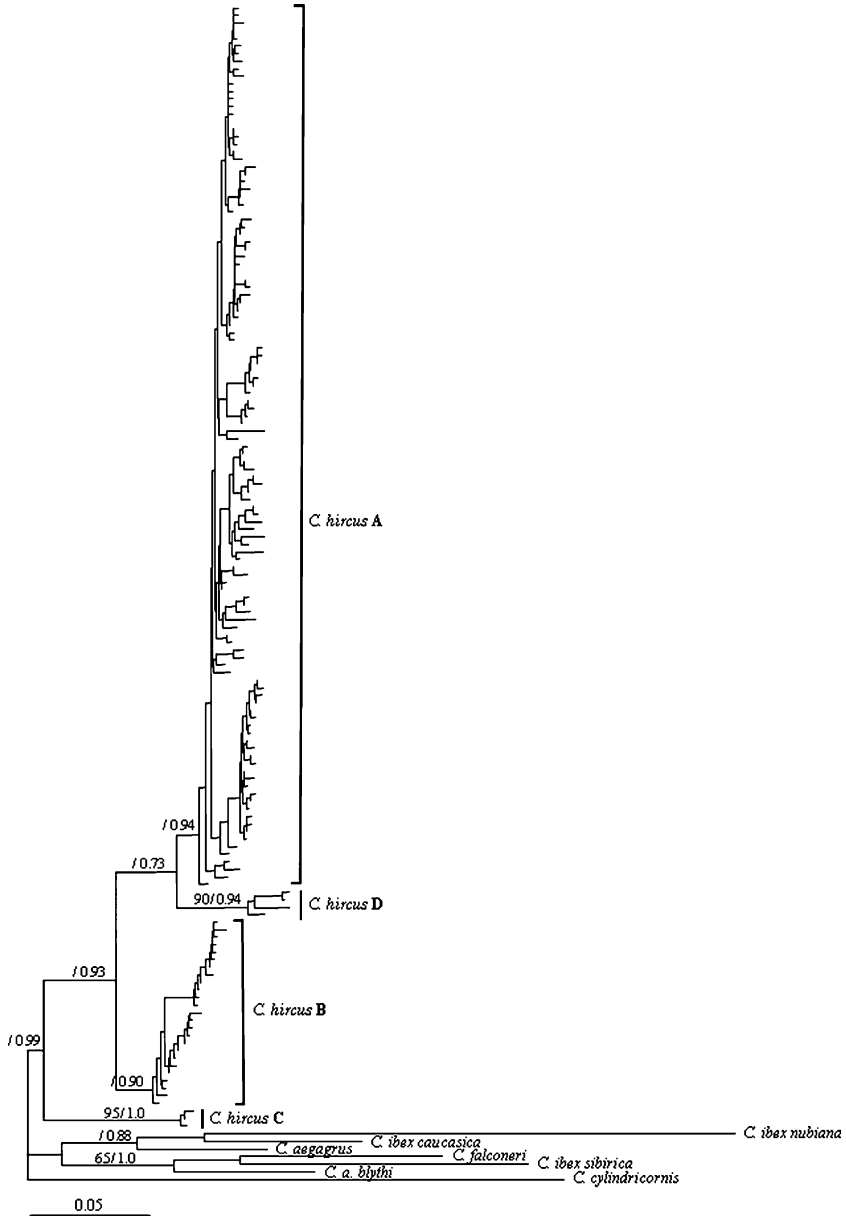

Fig. 2. Neighbor-joining tree of 146 haplotypes of Chinese domestic goat sequences, seven wild goat sequences, and three sequences belonging to lineage D from Pakistan goats. Tree was constructed under HKY+I+Γ model (with $\alpha = 0.4632$ and $I = 0.4028$). Numbers at the major clades denote the bootstrap percentages of 1000 replicates (only those $\geqslant 50\%$ are shown) and posterior probabilities.

lineages A–D and showed that 17, 31, and 18 mutation steps separate lineage A from lineages B, C, and D, respectively. Lineage A was found in all breeds. Lineage B was presented in all breeds except in the two Northern Xinjiang and Liaoning Cashmere breeds (Fig. 3). On the other hand, lineage C haplotypes were found only in Xinjiang and Tibetan breeds. Finally, the single haplotype of lineage D was found in Laoning Cashmere breed.

In a separate analysis, we combined all previously published sequences belonging to the *C. hircus* mtDNA lineage B with 92 sequences of Chinese domestic goats in our data. The analysis revealed 40 haplotypes in all 134 sequences, and the phylogenetic relationships among these haplotypes are shown in a minimum-spanning network (Fig. 4). The network comprised two subclades with star-like phylogeny, in which two high-frequency haplotypes B1 (16 individuals, 11.9% of the total of lineage B) and B3 (33 individuals, 24.6% of the total of lineage B) located in the center. The haplotype B3 shared among individuals from China (20 individuals), Laos (3 individuals), Malaysia (4 individuals), Pakistan (4 individuals), and India (2 individuals), whereas the haplotype B1 included individuals only from China. The single one individual from Mongolia combined with five individuals from China as haplotype B11, with one mutation step to haplotype B1.

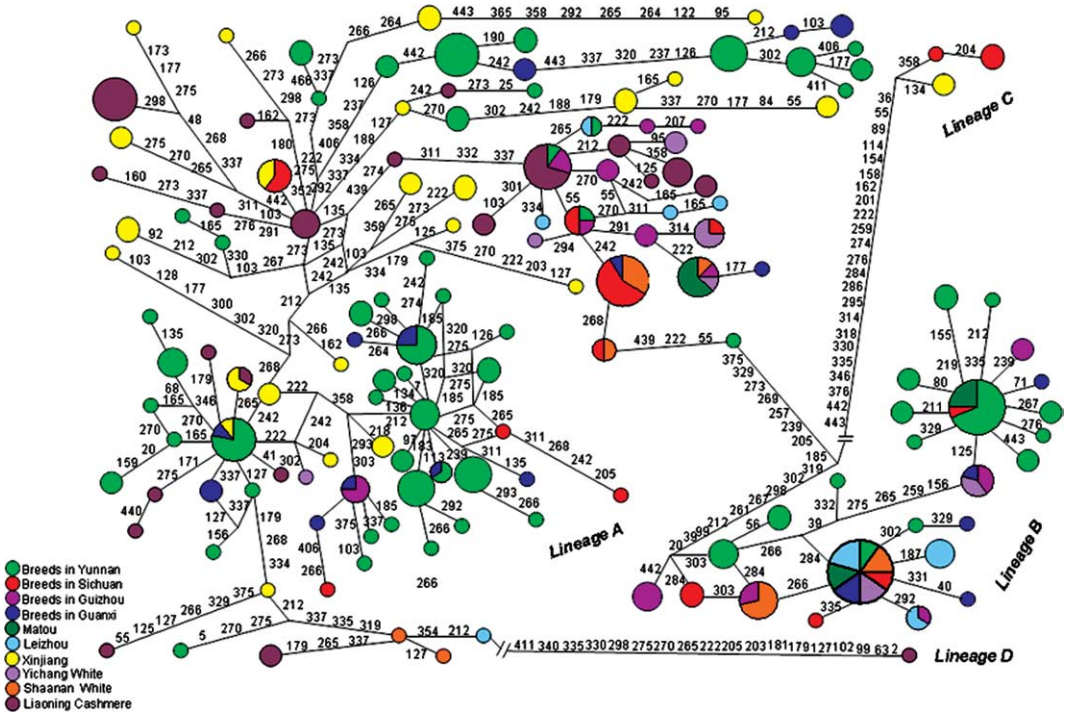

Fig. 3. Median-joining network of 146 mtDNA haplotypes of Chinese goats. The area of circle is proportional to haplotype frequency.

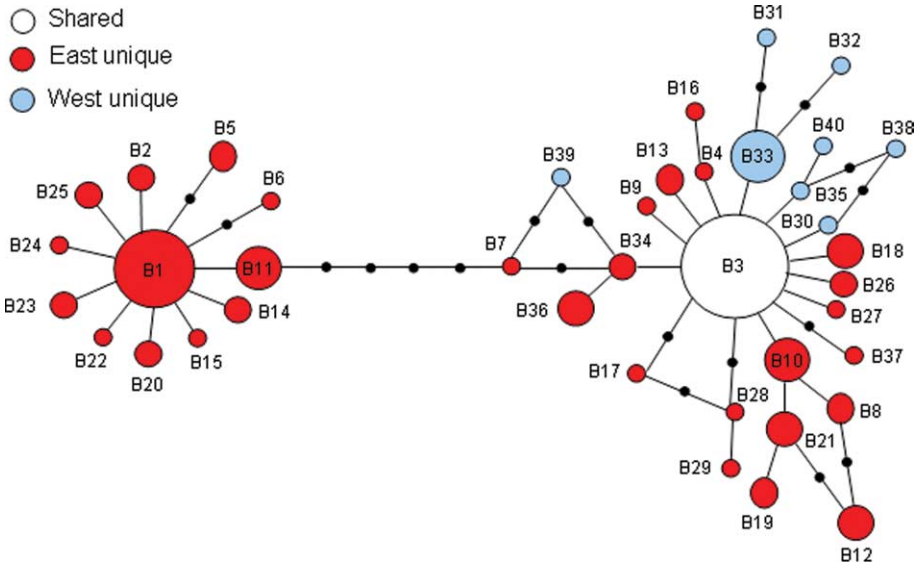

Fig. 4. A minimum spanning network showing genetic relationships among all mtDNA goat haplotypes of lineage B. Black dots are hypothetical intermediates. Circles denote haplotypes found in both East and West (white), unique to East (red), and unique to West (blue). The sizes of circles are proportional to haplotype frequency. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

3.3. Population structure

The median-joining network (Fig. 3) of 146 haplotypes of Chinese goats showed that breeds from different geographic regions intermingled. Haplotypes of each breed from different geographic regions did not cluster together and separate from those from other regions. Some haplotypes were shared by individuals of different breeds from different geographical regions. In addition, most breeds from one geographic region (e.g., Yunnan) distributed throughout the whole network. These results indicated that there was no correspondence between the geographic regions of origin and relationships among breeds.

To further investigate geographical structuring, we performed several AMOVA analyses at different hierarchical levels. A global AMOVA estimated that 91.25%

($P < 0.001$) of the variation was within breeds and only 8.75% ($P < 0.001$) was among breeds. When all goat breeds were grouped into four regions (Southwestern versus South versus Central versus Northern + Northwestern + Northeastern), 0.07% ($P = 0.1779$) of variation was among groups and 8.70% ($P < 0.001$) was among breeds within groups. Furthermore, a Mantel test revealed that there was no correlation between the genetic distances (F_{ST} values) and the geographical distances ($r = -0.29$, $P = 0.9348$, 10,000 permutations). All these results consistently demonstrated that there was no significant geographical structuring among Chinese goat breeds.

3.4. Population expansions

Historical demography of populations appeared to have had profound effect on patterns of genetic variation of mtDNA and genetic diversity (Donnelly et al., 1996; Harpending et al., 1998; Reich and Goldstein, 1998; Shriver et al., 1997). Previous studies demonstrated that there happened population expansion events in domestic goats (Joshi et al., 2004; Luikart et al., 2001). Generally, two different approaches were used to examine the traces of population expansion (Excoffier and Schneider, 1999). The first one approach is Fu's *Fs* statistic (Fu, 1997), which is based on the probability of having a number of alleles greater or equal to the observed number in a sample drawn from a stationary population. It provides a sensitive test for population expansion. We carried out 5000 simulations to obtain the null distribution of *Fs* statistic and its *P* value. Significantly large negative *Fs* values were interpreted as an evidence for population expansion (Fu, 1997). The other approach to detect expansion is the distribution of the number of pairwise differences between sequences in a sample (mismatch distribution) (Rogers and Harpending, 1992). This model is based on an infinite-site model and assumes that a stepwise expansion occurred some time in the past from a stationary population to a large stationary population. For that, we conducted a mismatch analysis with 5000 simulations. The overall validity of the estimated demographic model was tested by comparing the distribution of a test statistic sum of squared differences (SSDs) between the observed and the estimated mismatch distribution by a bootstrap approach (Excoffier and Schneider, 1999; Schneider and Excoffier, 1999). Significant SSD values can be taken as an evidence for departure from the estimated demographic model, which can be either an expanding or a stationary population (Excoffier and Schneider, 1999).

Because of the small sample size in most breeds in our sampling (<20 individuals), the detection of population expansion was not performed at individual breed level. The complete data set of all Chinese goat breeds had a significantly large negative *Fs* value (−23.57, $P < 0.01$) and this result was congruent with a demographic model showing a large and sudden expansion as inferred from the mismatch distribution (Fig. 5). The mismatch distribution of complete data set showed that there were two major peaks at around 12 and 28 and one smaller peak at around three differences. These results suggested at least three expansion events in the population demographic history of Chinese goat population.

To further explore a detailed information on demographic history of Chinese goat population, we also estimated the *Fs* statistic values and mismatch distribution for mtDNA lineages A and B of Chinese goats (because of few sequences belonging to lineages C and D, and these two lineages were not considered). Overall, results

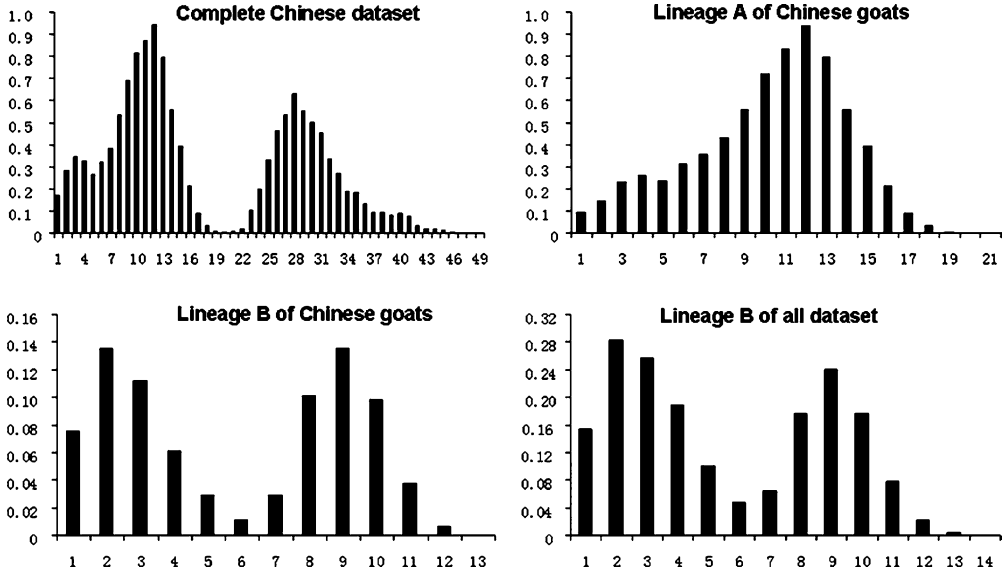

Fig. 5. Mismatch distributions of mtDNA types of Chinese goats and all reported data of lineage B.

drawn from *F*s statistic closely paralleled to those drawn from the mismatch analysis. *F*s values were −24.17 ($P < 0.01$) and −6.13 ($P < 0.05$) for lineages A and B, respectively, which indicated recent population growth, and were similar to mismatch distribution (Fig. 5). In addition, we examined the demographic expansion for combined data set of Chinese goats and previously reported sequences belonging to lineage B, and the result also showed population expansion as detected by significant negative *F*s value (*F*s value = −21.20, $P < 0.01$) and mismatch distribution (Fig. 5).

3.5. Molecular dating

Using five or seven million years for the sheep–goat split derived from fossil records (Carroll, 1988; Savage and Russell, 1983), the age of the TMRCA of the four Chinese domestic goat mtDNA lineages (A–D) was dated back to roughly 332,800 or 465,900 years. This time was higher than that of three mtDNA lineages A–C (roughly 201,380 or 281,932 years) estimated by Luikart et al. (2001), but was lower than that of the same four mtDNA lineages (427,006 or 597,806 years) estimated by Sultana et al. (2003). Furthermore, we also estimated the divergence time between the most recently evolved mtDNA lineages A and D, and compared it with the value obtained by Sultana et al. (2003). The estimated divergence time between lineages A and D was about 226,900 or 317,700 years, which was lower (but non-significant) than the value of 265,038 or 371,052 years from Sultana et al. (2003).

To gain an insight into the domestication of the goat mtDNA lineage B, we estimated the divergence time between the two central haplotypes B1 and B3 (Fig. 4). The estimated divergence time was about 100,800–141,200 years ago.

4. Discussion

4.1. Origin of goat mtDNA lineage B

The *C. hircus* mtDNA lineage B was first found in samples from Mongolia, Pakistan, India, and Malaysia in eastern and southern Asia, but not in Europe, Africa, or the Middle/Near East, where sampling was extensive (Luikart et al., 2001). Thus, Luikart et al. (2001) hypothesized that the lineage B was likely to originate in Asia. Due to small samples in lineage B, their data set did not provide answer to whether the lineage B domesticated in eastern or southern Asia, and the exact number of founders and approximate time of domestication events. Hereafter, new data belonging to lineage B were detected from samples from Laos (Mannen et al., 2001), Pakistan (Sultana et al., 2003), and India (Joshi et al., 2004). Additional 92 sequences belonging to lineage B were also found in this study. The frequencies of lineage B in different countries are shown in Fig. 1 and Table 2, showing higher frequency in eastern Asia than southern Asia. This suggested that lineage B might have arisen in eastern Asia.

To determine whether the origin of lineage B was in eastern or southern Asia, we compared the number of haplotypes and genetic diversity of sequence data from eastern and southern Asia, respectively. When an ancestral population and a derived population (derived from a subset of the genetic types of the ancestral population) are compared, genetic loci from the ancestral population are expected to show higher haplotypic and nucleotide diversity (Savolainen et al., 2002; Troy et al., 2001). The amount of genetic variation can be roughly measured by the mean pairwise differences. Mean pairwise differences between the mtDNA haplotypes from eastern Asia were 4.74 (SD = 2.34) and were from southern Asia 1.82 (SD = 1.09). Furthermore, we found 32 haplotypes in eastern Asia, which were larger than nine haplotypes in southern Asia. After correcting the sample size bias by resampling with replacements, according to the method of Savolainen et al. (2002), there were 13.05 (SD = 2.05) haplotypes among 21 goats in eastern Asia, which were significantly larger than nine haplotypes found among 21 goats in southern Asia ($P < 0.01$; 1000 replicates).

The detailed information on the pattern of genetic variation in eastern and southern haplotypic data was shown in the minimum spanning network of lineage B (Fig. 4). The minimum spanning network composed of two major subclades: one included 12 haplotypes with B1 in the center (subclade B1) and another included the remaining 28 haplotypes with B3 in the center (subclade B3). The eastern haplotpyes were found in both subclade B1 and subclade B3, whereas southern haplotypes only in subclade B3.

To further study the larger genetic variation in eastern Asia and the possibility of an origin of goat lineage B from eastern Asia rather than southern Asia, we compared the eastern (East) and southern (West) Asia, defined here as the regions east and west of the Himalayas Mountain. Only one haplotype (B3) was shared between East and West, whereas 31 were unique to East and 8 to West (Fig. 4). When we considered only the subclade B3, there were 11 unique haplotypes in East having at least two-step distance from the shared haplotype B3, whereas 5 unique haplotypes in West having two or three steps from the haplotype B3. All haplotypes in the subclade B1 were found only in China and Mongolia with 34.5% frequency in sample of lineage B in East. It would be expected that there should be haplotypes from samples of West presented in subclade B1 with low frequency. In subclade B1, however, there was no any haplotype from West. Although the sample size

in lineage B, up to now, was still not enough to draw decisive conclusion, our data indicated that the mtDNA haplotypes of lineage B in West (India and Pakistan) were likely to derive from the introduction of a subset of East Asian haplotypes, from which the haplotypes unique to West were latter formed. Notably, the network revealed two subclades with two founders (haplotypes B1 and B3), in which haplotype B1 was unique to eastern Asia. On the basis of all these results, we can safely say that at least one subclade of lineage B was originated from eastern Asia. Since haplotypes in subclade B1 were mainly distributed in China (one individual from Mongolia shared with individuals from China), especially in southwest China, this result supported the previous hypothesis that the ancient southwestern mountain area of China might be a domestication center of goat (Sieh, 1985).

It would be expected that there should be a star-like phylogeny, with the founder in the center and newly developed haplotypes radially distributed in a domestication event with a subsequent population expansion (Savolainen et al., 2002). The networks of the two subclades in lineage B were star-like and Fu's *F*s values for subclades B1 and B3 (−7.25 and −22.50, respectively) were significant ($P < 0.01$); these results supported population expansion happened in these two subclades. The divergence time between the central haplotypes B1 and B3 (roughly 100,800 to 141,200 years ago) was considerably older than the time of domestication (around 10,000 years). Since goat domestication was highly unlikely to have been much before 10,000 years according to the comprehensive archaeological record (Luikart et al., 2001; Zeder and Hesse, 2000), the domestication of goat lineage B were involved at least two maternally related goat populations carrying considerable preexisting mtDNA diversity.

4.2. Population structure of Chinese goat breeds

Our phylogeographic analyses revealed no significant geographical structuring of mtDNA variation among Chinese domestic goat breeds. This result did not conflict with the previous study, showing weak phylogeographic structure compared with other domestic animals (Luikart et al., 2001).

Two possible hypotheses can be used to explain our result. First, mtDNA variation seems to be a poor assay for analyzing population genetic structure at the breed level (Bradley et al., 1996). When the latter authors conducted a single African and European division excluding Indian breeds, they found that the variation within-continent and between-breed could account for only less than 4% of variance, despite the presence of very distinct Zebu and Taurine breeds in the African sample. However, this case was not true for Chinese domestic goat breeds, because Joshi et al. (2004) found significant genetic structure among Indian goat breeds using the mitochondrial first hypervariable region sequences.

Second, there was strong gene flow among goat populations in history. Since our samples including only indigenous breeds collected from remote geographic locations as Luikart et al. (2001) done, it was unreasonable that no significant geographical structuring in Chinese goat breeds was due to the very recent transport of goats among different regions in China. Therefore, it would be expected that transport of goats occurred commonly in history. Previous study demonstrated that goat genetic history was likely to be linked to human history (Joshi et al., 2004). Thus, information on human history can provide indirect clues for goat genetic history. According to the historical records, there were continuous southward movements of people (Han people) due to warfare and famine in northern China (Ge et al., 1997). Substantial data from Y chromosome and mtDNA also revealed that historical movements of human populations happened frequently in China (Su et al., 1999, 2000; Wen et al., 2004; Yao et al., 2002). It is justifiable that people would prefer to carry goats during their migratory movements, as goats have the ability to survive in the most adverse circumstances and supply a full range of useful products (Porter, 1996). In addition, people in early times treated goats as currency (Clutton-Brock, 1981), and this would be putatively responsible for gene flow among different goat populations in different regions.

5. Conclusions

In this study, we observed substantial mtDNA diversity in Chinese domestic goats. The level of diversity within China is higher than that of India (Joshi et al., 2004) and even the Old World (Luikart et al., 2001), but is slightly lower than that of Pakistan (Sultana et al., 2003). The time to the TMRCA of Chinese goat mtDNA, although rough estimate, is far older than the previously established domestication time of goats around 10,000 years (Porter, 1996; Pringle, 1998; Zeder and Hesse, 2000). The four mtDNA lineages A–D found in Chinese goat breeds further support the previous view of multiple maternal origins of domestic goats (Joshi et al., 2004; Luikart et al., 2001; Sultana et al., 2003).

Previously published sequences and our sequences belonging to lineage B, when combined, permitted us to find two subclades suggesting two founders and a larger genetic variation in eastern Asia than southern Asia in lineage B when corrected for sample size. Further, these results indicate that at least one subclade in lineage B originated from eastern Asia. Further studies on extensive sampling in Pakistan, southern Indian, and Southeast Asian countries will help to clarify the origin of lineage B.

As expected, no significant geographical structuring was found among Chinese goat breeds, and this was probably due to extensive transport of goats along with the frequent movements of human population in ancient China. This observation, however, was inferred only from the maternal genetic materials, further information from paternal Y chromosome and more extensive sampling needed to confirm this initial observation.

Acknowledgments

The work was supported by the State Key Basic Research and Development Plan (G20000161), National Natural Science Foundation of China (30021004), and Natural Science Foundation of Yunnan Province. We thank Dr. G. Luikart and Dr. A. Beja-Pereira for valuable comments. We also thank two anonymous reviewers for their constructive comments and suggestions. We are grateful to Professor A. Abdukadir (Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences) for his help in sampling. We are also grateful to H.-B. Xie for technical assistance.

References

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

Beja-Pereira, A., England, P.R., Ferrand, N., Jordan, S., Bakhiet, A.O., Abdalla, M.A., Mashkour, M., Jordana, J., Taberlet, P., Luikart, G., 2004. African origins of the domestic donkey. Science 304, 1781.

Bradley, D.G., MacHugh, D.E., Cunningham, P., Loftus, R.T., 1996. Mitochondrial diversity and the origin of the African and European cattle. Proc. Natl. Acad. Sci. USA 93, 5131–5135.

Carroll, R.L., 1988. Vertebrate Palaeontology and Evolution. Freeman, New York.

Clutton-Brock, J., 1981. Domesticated Animals from Early Times. University of Texas Press, Texas.

Donnelly, P., Tavare, S., Balding, D.J., Griffiths, R.C., 1996. Estimating the age of the common ancestor of men from the ZFY intron. Science 272, 1357–1359.

Excoffier, L., Schneider, S., 1999. Why hunter–gatherer populations do not show signs of Pleistocene demographic expansions. Proc. Natl. Acad. Sci. USA 96, 10597–10602.

Excoffier, L., Smouse, P.E., Quattiro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

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

Ge, J.X., Wu, S.D., Chao, S.J., 1997. Zhongguo yimin shi (The migration history of China). Fujian People Press, Fuzhou.

Giuffra, E., Kijas, J.M.H., Amarger, V., Carlborg, Ö, Jeon, J.T., Anderson, L., 2000. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics 154, 1785–1791.

Harpending, H.C., Batzer, M.A., Gurvens, M., Jorde, L.B., Rogers, A.R., Sherry, S.T., 1998. Genetic traces of ancient demography. Proc. Natl. Acad. Sci. USA 95, 1961–1967.

Hiendleder, S., Mainz, K., Plante, Y., Lewalski, H., 1998. Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources. No evidence for contributions from Urial and Argali sheep. J. Heredity 89, 113–120.

Hiendleder, S., Kaupe, B., Wassmuth, R., Janke, A., 2002. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. Proc. R. Soc. Lond. B 269, 893–904.

Huelsenbeck, J.P., Ronquist, F.R., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Joshi, M.B., Rout, P.K., Mandal, A.K., Tyler-Smith, C., Singh, L., Thangaraj, K., 2004. Phylogeography and origin of Indian domestic goats. Mol. Biol. Evol. 21, 454–462.

Legge, A.J., 1996. The beginning of caprine domestication in Southwest Asia. In: Harris, D.R. (Ed.), The Origins and Spread of Agriculture and Pastoralism in Eurasia. University College London Press, London, pp. 238–262.

Loftus, R.T., MacHugh, D.E., Bradley, D.G., Sharp, P.M., Cunningham, E.P., 1994. Evidence for two independent domestications in cattle. Proc. Natl. Acad. Sci. USA 91, 2757–2761.

Luikart, G., Giellly, L., Excoffier, L., Vigne, J.D., Bouuvet, J., Taberlet, P., 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. Proc. Natl. Acad. Sci. USA 98, 5927–5932.

MacHugh, D.E., Bradley, D.G., 2001. Livestock genetic origins: goats buck the trend. Proc. Natl. Acad. Sci. USA 98, 5382–5384.

Mannen, H., Nagata, Y., Tsuji, S., 2001. Mitochondrial DNA reveal that domestic goat (*Capra hircus*) are genetically affected by two subspecies of bezoar (*Capra aegagurus*). Biochem. Genet. 39, 145–154.

Meaow, R.H., 1993. Animal domestication in the Middle East: a revised view from the eastern margin. In: Possehl, G.L. (Ed.), Harappan Civilization. Oxford and IBH, New Delhi, pp. 295–320.

Porter, V., 1996. Goats of the world. Farming Press, Ipswich, UK.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Pringle, H., 1998. Neolithic agriculture: reading the signs of ancient animal domestication. Science 282, 1448.

Reich, D.E., Goldstein, D.B., 1998. Genetic evidence for a Paleolithic human population expansion in Africa. Proc. Natl. Acad. Sci. USA 96, 8119–8123.

Rogers, A.R., Harpending, H., 1992. Population growth curves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9, 552–569.

Savage, D.E., Russell, D.E., 1983. Mammalian Palaeofaunas of the World. Addison-Wesley, Reading, MA.

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.

Schneider, S., Excoffier, L., 1999. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152, 1079–1089.

Schneider, S., Roessli, D., Excoffier, L., Arlequin: A Software for Population Genetics Data Analysis. Ver 2.000. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Shriver, M.D., Jin, L., Ferrell, R.E., Deka, R., 1997. Microsatellite data support an early population expansion in Africa. Genome Res. 7, 586–591.

Sieh, C.H., 1985. History of Raising Ruminant Livestocks in China. Agricultural Publishing House, Beijing.

Su, B., Xiao, C., Deka, R., Seielstad, M.T., Kangwanpong, D., Xiao, J., Lu, D., Underhill, P., Cavalli-Sforza, L., Chakraborty, R., Jin, L., 2000. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. Hum. Genet. 107, 582–590.

Su, B., Xiao, J., Underhill, P., Deka, R., Zhang, W., Akey, J., Huang, W., Shen, D., Lu, D., Luo, J., Chu, J., Tan, J., Shen, P., Davis, R.,

Cavalli-Sforza, L., Chakraborty, R., Xiong, M., Du, R., Oefner, P., Chen, Z., Jin, L., 1999. Y chromosome evidence for a northward migration of modern humans in East Asia during the last ice age. Am. J. Hum. Genet. 65, 1718–1724.

Sultana, S., Mannen, H., Tsuji, S., 2003. Mitochondrial DNA diversity of Pakistani goats. Anim. Genet. 34, 417–421.

Swofford, D.L., 2003. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4.0b10. Sinauer Associates, Sunderland, MA.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalities and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Troy, C.S., MacHugh, D.E., Bailey, J.F., Magee, D.A., Loftus, R.T., Cunningham, P., Chamberlain, A.T., Sykes, B.C., Bradley, D.G., 2001. Genetic evidence for Near-Eastern origins of European cattle. Nature 410, 1088–1091.

Vilà, C., Savoleinen, 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 dog. Science 276, 1687–1689.

Vilà, C., Leonard, J.A., Götherström, A., Marklund, S., Sandberg, K., Liden, K., Wayne, R.K., Ellegren, H., 2001. Widespread origins of domestic horse lineages. Science 291, 474–477.

Wen, B., Li, H., Lu, D., Song, X., Zhang, F., He, Y., Li, F., Gao, Y., Mao, X., Zhang, L., Qian, J., Tan, J., Jin, J., Huang, W., Deka, R., Su, B., Chakraborty, R., Jin, L., 2004. Genetic evidence supports demic diffusion of Han culture. Nature 431, 302–305.

Yao, Y.G., Kong, Q.P., Bandelt, H.J., Kivisild, T., Zhang, Y.P., 2002. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am. J. Hum. Genet. 70, 635–651.

Zeder, M.A., Hesse, B., 2000. The initial domestication of goats (*Capra hircus*) in the Zagros Mountain 10,000 years ago. Science 287, 2254–2257.