Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 376 (2006) 216–223
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Origin, genetic diversity, and population structure of Chinese domestic sheep

Shan-Yuan Chen [a,b,1], Zi-Yuan Duan [b,1], Tao Sha [a], Jinggong Xiangyu [b], Shi-Fang Wu [b], 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 650223, China*

Received 1 December 2005; received in revised form 6 March 2006; accepted 13 March 2006
Available online 5 April 2006
Received by Takashi Gojobori

Abstract

To characterize the origin, genetic diversity, and phylogeographic structure of Chinese domestic sheep, we here analyzed a 531-bp fragment of mtDNA control region of 449 Chinese autochthonous sheep from 19 breeds/populations from 13 geographic regions, together with previously reported 44 sequences from Chinese indigenous sheep. Phylogenetic analysis showed that all three previously defined lineages A, B, and C were found in all sampled Chinese sheep populations, except for the absence of lineage C in four populations. Network profiles revealed that the lineages B and C displayed a star-like phylogeny with the founder haplotype in the centre, and that two star-like subclades with two founder haplotypes were identified in lineage A. The pattern of genetic variation in lineage A, together with the divergence time between the two central founder haplotypes suggested that two independent domestication events have occurred in sheep lineage A. Considerable mitochondrial diversity was observed in Chinese sheep. Weak structuring was observed either among Chinese indigenous sheep populations or between Asian and European sheep and this can be attributable to long-term strong gene flow induced by historical human movements. The high levels of intra-population diversity in Chinese sheep and the weak phylogeographic structuring indicated three geographically independent domestication events have occurred and the domestication place was not only confined to the Near East, but also occurred in other regions.

Keywords: Domestication; *Ovis aries*; mtDNA; Control region; Phylogeography

1. Introduction

Domestic sheep (*Ovis aries*) have played important roles in diverse human societies as a source of food, hide, and wool, and are one of the major components of agro-pastoral societies since the Neolithic. It is therefore of considerable interest to appreciate what, where, and when of sheep domestication. Archaeological evidence suggested that sheep were probably first domesticated in the Fertile Crescent region of the Near East approximately 8000–9000 years ago (Ryder, 1984). Three major groups of Eurasian wild sheep, mouflon (*O. musimon* or *O. orientalis*), urial (*O. vignei*), and argali (*O. ammon*), have been proposed as ancestors of domestic sheep (Ryder, 1984) or, at least, are believed to have contributed to specific breeds (Reed, 1960; Zeuner, 1963).

Mitochondrial DNA (mtDNA) data from wild and domestic sheep revealed that there were no contributions from urial and argali species to domestic sheep and thus favoring the mouflon species as the only progenitor of domestic sheep (Hiendleder et al., 1998a, 2002). Two maternal lineages A and B were identified in modern domestic sheep breeds sampled from different geographic regions of the world (Hiendleder et al., 1998a, 2002; Meadows et al., 2005; Wood and Phua, 1996). More recently, a novel maternal lineage C, besides lineages A and B, was observed in sheep from Turkey and China (Guo et al., 2005; Pedrosa et al., 2005). The time

Abbreviations: mtDNA, mitochondrial DNA; bp, base pairs; nps, nucleotide positions; CR, control region; Cyt *b*, Cytochrome *b*; AMOVA, analysis of molecular variance; MDS, multidimensional scaling; TMRCA, the most recent common ancestor.

* Corresponding author. Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. Tel./fax: +86 871 5032804.
E-mail addresses: zhangyp1@263.net.cn, zhangyp@mail.kiz.ac.cn (Y.-P. Zhang).
[1] These authors contributed equally to this work.

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.03.009

since divergence among these three lineages in Turkish sheep estimated from the mtDNA cytochrome *b* gene (around 160,000 to 750,000 years ago) (Pedrosa et al., 2005) greatly predated the time of sheep domestication (around 8000–9000 years ago) and suggested that at least three independent sheep domestication events occurred. These results, combined with archaeological data and wide distribution of wild sheep populations, further highlight the importance of the Near East region in the sheep domestication process (Pedrosa et al., 2005). However, the question of how many founders in each lineage, the geographic distribution of these lineages, and the dispersal pattern from the domestication centre to other regions remains poorly understood.

China has 15 autochthonous sheep breeds and numerous local sheep populations, which are distributed from the high Qinghai–Tibet Plateau to the low land of East China (Tu et al., 1989). Based on the morphology and the distribution of wild sheep populations of *O. ammon*, it was hypothesized that Chinese domestic sheep derived from wild sheep populations of *O. orientalis* and *O. ammon* (Tu et al., 1989). However, molecular phylogenetic analyses have excluded *O. ammon* as the maternal ancestor of domestic sheep (Hiendleder et al., 1998a, 2002; Wu et al., 2003). Previous genetic studies identified three mtDNA types in Chinese domestic sheep by restriction fragment length polymorphism and hinting multiple maternal sources (Li et al., 2001; Tu and Zhang, 1998). This point is further confirmed by the recent survey, in which three mtDNA lineages A, B, and C were found in six Chinese indigenous breeds by single-strand conformational polymorphism and mtDNA control region sequence analysis (Guo et al., 2005). Considering archaeological remains showing an early presence of domestic sheep in ancient China (Kuo et al., 1999; Tu et al., 1989), whether there is any specific founder or subclade in the three lineages A, B, and C in Chinese domestic sheep needs to be further investigated. Due to limited geographic sampling in previous studies, little is known about the pattern of genetic diversity, the geographic distribution of these three lineages, and population structure in Chinese domestic sheep.

To address these issues, we here examined the mtDNA sequence variation in 19 Chinese autochthonous sheep breeds/populations from 13 geographic regions in China, combined with the 44 sequences of Chinese native sheep reported in a previous study (Guo et al., 2005). To know more about the degree of diversity found in Chinese sheep and haplotype sharing between sheep from China and other geographic regions, we compared our data with the available sheep sequences in GenBank database. Meadows et al. (2005) found weak structuring between Asian and European sheep owing to high levels of gene flow. Given the few sequences from Asian sheep during that time, we also further investigated the phylogeographic structure between Asian and European sheep by adding the Chinese sheep sequences.

2. Materials and methods

2.1. Sampling and DNA extraction

Blood samples from 405 domestic sheep (*O. aries*) representing 17 populations from 13 Chinese autochthonous breeds were collected from small remote villages belonging to 11 geographic regions in China (Fig. 1 and Table 1), and stored at −70 °C. An effort was made to collect samples from unrelated individuals based on the information from the owners and local farmers.

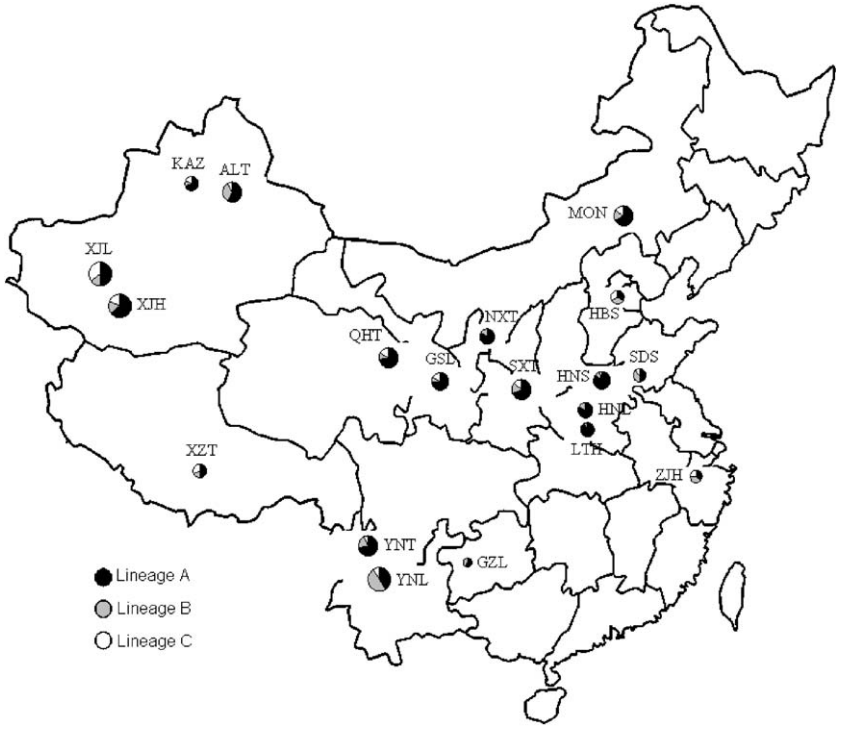

Fig. 1. Geographic distribution of samples and of the three mtDNA lineages A, B, and C in Chinese domestic sheep breeds/populations. Breed/population codes are as reported in Table 1. Circle area is proportional to the sample size.

Table 1
The sample information and some diversity indices for each breed/population used in the study

Breed/population	Code	Location	Sample size[a]	Number of haplotypes	Lineage observed	Haplotype diversity (SE)	Nucleotide diversity (SE)
Tan	NXT	Ningxia	12	4	A, B	0.4545±0.1701	0.0069±0.0042
Small Tailed Han	SDS	Shandong	8	8	A, B, C	1.0000±0.0625	0.0200±0.0116
Small Tailed Han	HNS	Henan	21	11	A, B, C	0.8762±0.0581	0.0106±0.0059
Small Tailed Han	HBS	Hebei	9 (gb)	8	A, B, C	0.9722±0.0640	0.0223±0.0127
Mongolian	MON	Inner Mongolia	28+9 (gb)	28	A, B, C	0.9640±0.0222	0.0154±0.0081
Hu	ZJH	Zhejiang	8 (gb)	8	A, B, C	1.0000±0.0625	0.0239±0.0138
Lanzhou Large Tailed	GSL	Gansu	22	15	A, B, C	0.9437±0.0330	0.0170±0.0091
Lanzhou Large Tailed	HNL	Henan	16	7	A, B	0.7917±0.0886	0.0106±0.0060
Large Tailed Han	LTH	Henan	11	10	A, B	0.9818±0.0463	0.0104±0.0061
Tong	SXT	Shaanxi	24+5 (gb)	21	A, B, C	0.9729±0.0173	0.0161±0.0085
Tibetan	QHT	Qinghai	29	16	A, B, C	0.8941±0.0467	0.0142±0.0076
Tibetan	XZT	Xizang	9+7 (gb)	13	A, B, C	0.9750±0.0295	0.0173±0.0094
Tibetan	YNT	Yunnan	28	8	A, B, C	0.5794±0.1041	0.0101±0.0056
Yunnan Local	YNL	Yunnan	51	20	A, B, C	0.8941±0.0274	0.0145±0.0076
Guizhou Local	GZL	Guizhou	5	4	A, B	0.9000±0.1610	0.0139±0.0092
Kazakh Fat-rumped	KAZ	Xinjiang	10+6 (gb)	13	A, B, C	0.9750±0.0295	0.0148±0.0082
Altay Fat-rumped	ALT	Xinjiang	35	19	A, B, C	0.9092±0.0385	0.0141±0.0075
Hetian	XJH	Xinjiang	48	22	A, B, C	0.9025±0.0339	0.0132±0.0070
Xinjiang Local	XJL	Xinjiang	48	29	A, B, C	0.9619±0.0147	0.0151±0.0079

[a] The suffix of sample size (gb) indicates that these sequences were extracted from GenBank database (Guo et al., 2005).

Details of breeds, geographic regions, and sample sizes are given in Table 1. Total genomic DNA was extracted from blood by standard phenol–chloroform extraction method.

2.2. PCR amplification and sequencing

To avoid the tandem repeats in the sheep mtDNA control region (CR), a 531-bp fragment of control region between nucleotide positions (nps) 16007 to 16537 of the sheep reference sequence AF010406 (Hiendleder et al., 1998b) was amplified and sequenced. The primers CR400F (5′-ACTGCTTGACCGTACA TAGTAC-3′) and CR1099R (5′-AGTATTGAGGACGGGGTA A-3′) were used to amplify the 531-bp CR sequence fragment. The 1140 bp cytochrome *b* (Cyt *b*) was sequenced for six sheep arbitrarily chosen (two individuals with different CR sequences from each lineage) by using the primers L14724V (5′-ATGA TATGAAAAACCAT CGTTG-3′) and H15915V (5′-TCTCCTT CTCTGGTTTACAAGAC-3′) (Luikart et al., 2001). PCR amplifications were conducted in a 50 μl volume containing 5 μl 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 (CR sequence fragment) or 53 °C (Cyt *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 Inc., Shanghai) and were directly sequenced by using BigDye™ Terminator v3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems) on ABI PRISM 3700 DNA Analyzer according to the manufacturer's manual. The 531-bp CR sequence fragment was sequenced using the primers CR400F and CR1099R. The Cyt *b* gene sequences were sequenced by the primers L14724V and H15915V and additional two internal primers CYTB-FI (5′-CGCCTTCCACTT TATCCTCCC-3′) and CYTB-RI (5′-GTCTGATGGAATTCCT GTGGG-3′) (Chen et al., 2005). DNA sequences were edited using DNAstar 5.0 package (DNASTAR Inc.).

2.3. Data analysis

All 405 sequences of 531-bp CR sequence fragment and 44 sequences previously reported from Chinese indigenous breeds (GenBank accession nos. AY829376, AY829383–AY829405, and AY829411–AY829430) (see Table 1) were aligned using the Clustal W program (Thompson et al., 1994). In our analysis, we also included mtDNA CR sequences from different geographic regions available from GenBank with the accession numbers: New Zealand [Z35228–Z35268, and Z35293] (Wood and Phua, 1996); Mexico [AY582800–AY582820]; Germany [AF039577–AF039578] (Hiendleder et al., 1998a); Turkey [AY091495–AY091497], Tajikistan [AY091498], and Kazakhstan [AY091499–AY091500] (Hiendleder et al., 2002); India [DQ073049–DQ073050, DQ073053, and DQ087255]; Turkey [DQ097431–DQ097468] (Pedrosa et al., 2005); Austria [AY879343, AY879347–AY879373, AY879409–AY879432, AY879442–AY879443, and AY879451–AY879457], imported to Austria with European origins [AY879397, AY879400, AY879407–AY879408, AY879441, AY879449–AY879450], Indonesia [AY879374–AY879387 and AY879444–AY879447], Aland islands [AY879344–AY879346, AY879396, and AY879440], Finland [AY879399, AY879401–AY879403, and AY879405], Russia [AY879388–AY879395, AY879398, AY879404, and AY879448], Spain [AY879434–AY879439], Tibet [AY879433, and AY879458–AY879462], Central Asia [AY879406], and Mongolia [AY879463] (Meadows et al., 2005).

All median joining (MJ) networks (Bandelt et al., 1999) were drawn using the program Network 4.1.0.9 (www.fluxus-engineering.com) to investigate the possible relationships among haplotypes. Haplotype diversity, nucleotide diversity, mismatch analysis, F_{ST} distances, Fu's (1997) Fs values, and AMOVA (analysis of molecular variance) (Excoffier et al., 1992) were computed using ARLEQUIN 2.0 software (http://anthropologei.unige.ch/arlequin) (Schneider et al., 2000). Population pairwise F_{ST} genetic distances were displayed in two-dimensional space by

means of a multidimensional scaling (MDS) analysis, using the SPSS 10.0 software package.

According to Luikart et al. (2001), Beja-Pereira et al. (2004), and Chen et al. (2005), a molecular-clock likelihood-ratio test (for heterogeneity in substitution rates) was performed by using the six Chinese sheep and two domestic goats (*Capra hircus*) (GenBank accession nos., X56289 and AF533441) Cyt *b* sequences analyzed for the 380 nucleotides at third codon positions (i.e., synonymous positions unlikely to be under selection). The test was not significant ($P>0.05$), permitting us to use the amount of divergence between domestic sheep and goat sequences to estimate the approximate time to the most recent common ancestor (TMRCA) of Chinese sheep. All sequences in this study are deposited in GenBank (accession nos. DQ308611–DQ309021).

3. Results

3.1. mtDNA variation in Chinese domestic sheep

All the 449 Chinese sheep sequences represented 19 populations from 13 geographic regions (Fig. 1 and Table 1). A total of 170 haplotypes of 449 sequences of 531-bp CR sequence fragment was defined by 91 variable sites (see Supplementary Fig. S1), in which there were 87 substitutions (10 transversions) and 4 insertions/deletions. The number of haplotypes identified in each population ranges from four to twenty-nine, and haplotype diversity values from 0.4545±0.1701 in Tan breed to 1.0000±0.0625 in Hu breed and Small Tailed Han breed from Shandong (Table 1). The Hu breed shows the highest nucleotide diversity value (0.0239±0.0138), whereas the Tan breed displays the lowest one (0.0069±0.0042) (Table 1).

3.2. Phylogeny of haplotypes

In comparison to previously well-defined sheep mtDNA lineages A, B (Hiendleder et al., 1998a), and C (Guo et al., 2005; Pedrosa et al., 2005), all these 170 haplotypes of Chinese domestic sheep can be clearly grouped into these three lineages A, B, and C with 91 (a1–a91), 48 (b1–b48), and 31 (c1–c31) haplotypes, respectively (see Supplementary Fig. S1). To discern the phylogenetic relationships among haplotypes within lineages, we constructed the median joining (MJ) networks for each lineage (Figs. 2 and 3). Within the network profile of each lineage, there is one major haplotype in the centre (a1 for lineage A, b10 for lineage B, and c6 for lineage C) as shown in Figs. 2 and 3. From the haplotype a1 in lineage A, there were eleven and ten mutation steps to b10 of lineage B and c6 of lineage C, respectively. Interestingly, one larger subclade with central haplotype a1 (here defined as subclade a1) and one smaller subclade with two major central haplotypes a5 and a22 (considering a5 connected more one-mutation-step haplotypes and a22 also derived from a5, here defined it as subclade a5) can be identified in lineage A, and there are eight mutation steps between a1 and a5 (Fig. 2).

3.3. Population demography and structure

The star-like phylogeny of Chinese sheep lineages A, B, and C in the MJ network (Figs. 2 and 3) is consistent with a population expansion (Slatkin and Hudson, 1991). The sign of a population expansion was also demonstrated by the bell-shaped curve of the mismatch pairwise distributions (Rogers and Harpending, 1992) within each lineage (data not shown). As expected, lineages A, B, and C yielded significantly negative values

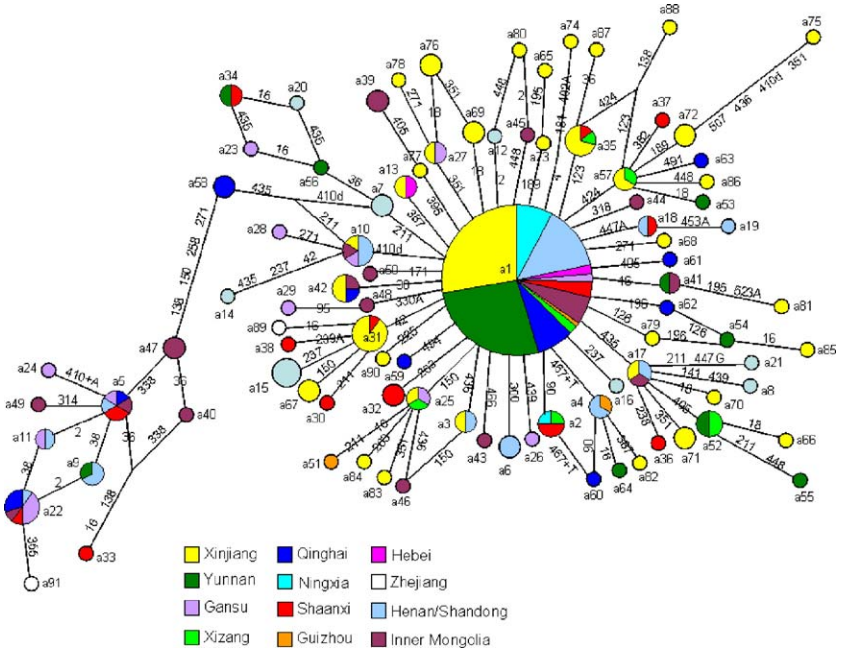

Fig. 2. The median joining network of 91 haplotypes in the control region 531-bp fragment of 281 Chinese sheep sequences. Suffixes A, C, G, and T indicate transversion, "d" indicates deletion, and "+" indicates insertion. The nucleotide positions 1 to 531 correspond to the nucleotide positions 16007 to 16537 of *Ovis aries* [AF010406] (Hiendleder et al., 1998b).

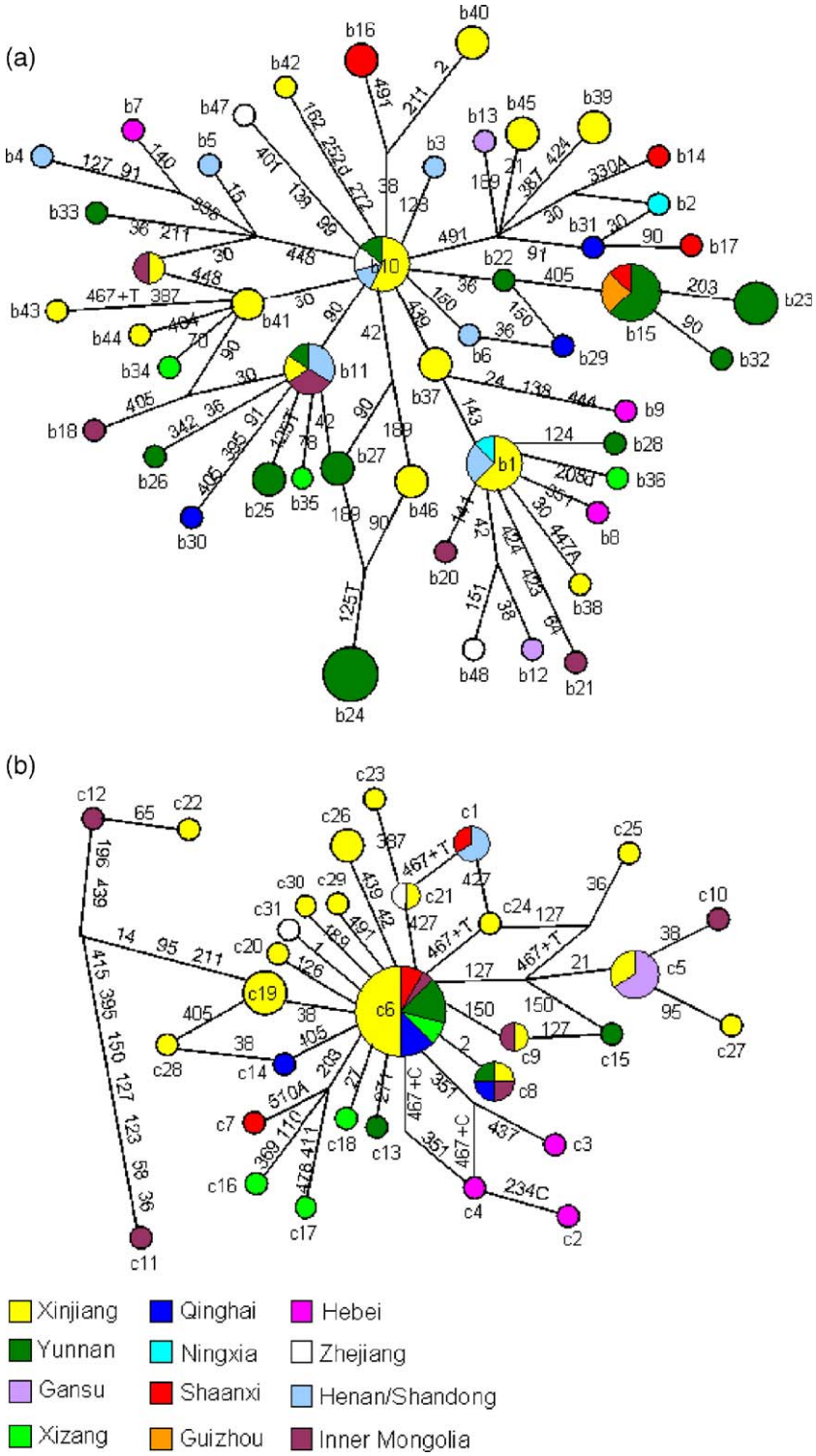

Fig. 3. The median joining networks of 48 and 31 haplotypes in the control region 531-bp fragment of lineages B (a) and C (b) from 98 and 70 Chinese sheep sequences, respectively. Suffixes A, C, G, and T indicate transversion, "d" indicates deletion, and "+" indicates insertion. The nucleotide positions 1 to 531 correspond to the nucleotide positions 16007 to 16537 of *Ovis aries* [AF010406] (Hiendleder et al., 1998b).

for Fu's (1997) Fs test (−26.20, −25.89, and −25.31 for lineages A, B, and C, respectively). These analyses uniformly indicated lineages A, B, and C had undergone population expansion events.

The network profile showed that populations from different geographic regions in China intermixed (Figs. 2 and 3). Some haplotypes (e.g., the largest haplotype a1) were shared by individuals from different populations across wide geographic regions. The MDS plot of population pairwise F_{ST} genetic distances displayed that the relationships among populations were not in harmony with their geographic locations, for instance, the population of Lanzhou Large Tailed breed from Hennan closely

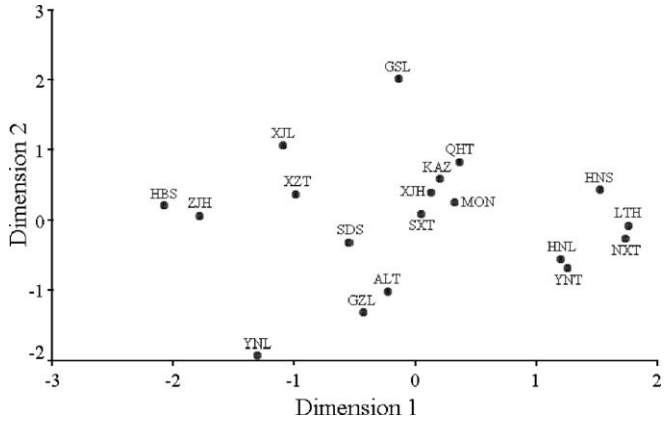

Fig. 4. Two-dimensional scaling plot of breed/population pairwise F_{ST} distances. Breed/population codes are as reported in Table 1.

related to the population of Tibetan breed from Yunnan (Fig. 4). Furthermore, a global AMOVA (i.e., the 19 sheep populations were treated as a single group) estimated that 94.66% ($P<0.001$) of genetic variance was within populations and 5.34% ($P<0.001$) was among populations. When all sheep populations were divided into four geographic groups ([North China] [Central and East China] [Northwest China] [Southwest China]) (see Table 3), only 2.13% ($P<0.05$) of the genetic variation could be attributed to differences among regional groups. There results consistently demonstrated that there was no significantly geographical structuring among Chinese sheep populations.

3.4. Divergence time estimation

Using a calibration of 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 three Chinese domestic sheep mtDNA lineages was traced back to roughly 481,700 to 674,400 years, which was much longer than the ca 8000–9000 years of sheep domestication history (Ryder, 1984). The estimated divergence times between lineages C and A (419,900 to 587,900 years), and lineages C and B (542,600 to 759,600 years) were almost identical with those estimated by Pedrosa et al. (2005). The divergence time between lineages A and B (190,900 to 267,300 years) in Chinese sheep was higher than the value of 160,000 to 170,000 years in Turkish sheep (Pedrosa et al., 2005). To give a clue to the domestication process of lineage A, we estimated the divergence time between the two central haplotypes a1 and a5 (Fig. 2). The estimated time was approximately 105,800 to 148,100 years ago.

3.5. Comparative analysis with sheep from other regions

We analyzed currently available 684 mtDNA CR sequences (Table 2). Based on the 445-bp sequence fragment between nps 16093 to 16537 of *O. aries* [AF010406], 85 and 97 haplotypes were identified in 344 and 258 sequences belonging to lineages A and B, respectively. 38 haplotypes were found in 82 sheep

Table 2
Diversity measures and frequency of sheep mtDNA lineages A, B, and C in different geographic regions (mtDNA from nps 16093 to 16537)

Region	Haplotype diversity (SE)	Nucleotide diversity (SE)	Lineage A (%)	Lineage B (%)	Lineage C (%)
China (N=455)[a]	0.8970 (0.0124)	0.0116 (0.0062)	286 (62.8)	99 (21.8)	70 (15.4)
Indonesia (N=18)	0.7582 (0.0787)	0.0061 (0.0038)	4 (22.2)	1 4 (77.8)	0
India (N=5)	0.8000 (0.1640)	0.0072 (0.0052)	5 (100)	0	0
Turkey (N=82)[b]	0.9695 (0.0143)	0.0139 (0.0075)	21 (25.6)	4 4 (53.7)	17 (20.7)
Central Asia (N=5)[c]	0.7000 (0.2184)	0.0103 (0.0071)	3 (60)	2 (40)	0
New Zealand (N=42)	0.8850 (0.0360)	0.0112 (0.0062)	17 (40.5)	25 (59.5)	0
Mexico (N=21)	0.5524 (0.1215)	0.0018 (0.0015)	0	21 (100)	0
Austria (N=61)	0.8842 (0.0361)	0.0074 (0.0043)	9 (14.8)	52 (85.2)	0
Europe (N=36)[d]	0.9365 (0.0217)	0.0091 (0.0052)	6 (16.7)	30 (83.3)	0

[a] 455 sequences include 405 sequences in this study, 44 sequences from Guo et al. (2005), and six sequences from Tibet of China reported by Meadows et al. (2005).

[b] Pedrosa et al. (2005) examined 79 individuals, but only 38 sequences were submitted to GenBank database. Three individuals were reported by Hiendleder et al. (2002). Diversity measures were calculated based on the 41 sequences.

[c] Five sequences include three sequences (two from Kazakhstan and one from Tajikistan) reported by Hiendleder et al. (2002) and two sequences (one from northern Mongolia and one from Central Asia) reported by Meadows et al. (2005).

[d] 36 European sequences include 34 sequences (five from Aland, five from Finland, eleven from Russia, six from Spain including four mouflon, and seven from imported animals with European origin) reported by Meadows et al. (2005) and two of Germany reported by Hiendleder et al. (1998a,b).

Table 3
Different hierarchical distribution of sheep mtDNA variation within and among populations under AMOVA analyses (% of total)

Grouping[a]	Within populations[b]	Among populations within groups[b]	Among groups[c]
No grouping in China	94.66	5.34	
Four regions in China	94.22	3.64	2.13
China versus other Asian countries	93.23	5.15	1.62
China versus Europe	67.51	6.22	26.27
Asia versus Europe	68.61	11.79	19.61

[a] Four regions in China indicate North, Northwest, Central and East, and Southwest China. North China includes six provinces (Ningxia, Qinghai, Gansu, Shaanxi, Hebei, and Inner Mongolia). Northwest China includes one province (Xinjiang). Central and East China includes three provinces (Henan, Shandong, and Zhejiang). Southwest China includes three provinces (Guizhou, Yunnan, and Xizang). Asia includes China, Indonesia, India, Turkey, and Central Asia, all the remaining countries or regions in Table 2 represent Europe, including sheep from European countries or regions and sheep with European origins. The groupings in China are based on 531 bp sequences (nps 16007–16537). Other groupings are based on 445 bp sequences (nps 16093–16537).

[b] $P<0.001$.

[c] $P<0.05$, except for China versus other Asian countries ($0.05<P<0.1$).

belonging to lineage C from China and Turkey, based on the common 531-bp sequence fragment corresponding to nps 16007 to 16537 of *O. aries* [AF010406]. There were five and seven haplotypes of Chinese sheep shared by other regions in lineage A and B, respectively (see Supplementary Figs. S2 and S3). However, only two haplotypes were shared between sheep from China and Turkey in lineage C (see Supplementary Fig. S4). It is worthy of notice that the level of genetic diversity observed in China (nucleotide diversity value, 0.0116±0.0062) was the second highest one in all countries or regions (Table 2), except for Turkey (0.0139±0.0075).

When all sheep populations were grouped according to continents (Asia and Europe), the fraction of genetic variance between these two continents was 19.61% ($P<0.05$) (Table 3), which was larger than 2.7% estimated by Meadows et al. (2005). When comparing China with other Asian countries, there was only 1.62% ($0.05<P<0.1$) of genetic variation that could be attributed to differences between groups (Table 3). However, the proportion of genetic variation between the grouping China and Europe was 26.27% ($P<0.05$) (Table 3).

4. Discussion

Three star-like major mtDNA lineages A, B, and C were observed in modern domestic sheep, based on all the available CR sequence data. The time since divergence among these three lineages (roughly 160,000–760,000 years) estimated by us and Pedrosa et al. (2005) far predated the time of sheep domestication history (around 8000–9000 years) (Ryder, 1984). This result suggests that at least three independent domestication events have occurred. Together all three mtDNA lineages found in Turkish sheep, with abundant archaeological data throughout the Near East, Pedrosa et al. (2005) highlighted the Near East region in the sheep domestication process. However, all three mtDNA lineages are not only found in sheep from the Near East, but also detected in all sampled Chinese sheep populations in this study, except for the absence of lineage C in four populations (Table 1). The high levels of intra-population diversity in Chinese domestic sheep and the weak population substructure observed in modern domestic sheep (Table 3) suggest that three geographically independent domestication events have occurred and the presence of all lineages in Chinese and Turkish sheep is probably due to human mediated animal movements and introgression. Molecular studies in other domestic species such as Taurine cattle (Bradley et al., 1996; Troy et al., 2001) and goats (Luikart et al., 2001; Chen et al., 2005) have also revealed that animal domestication was not confined to the Near East. In addition, one should recognize that the Near Eastern origin of major domestic livestock supported by archaeological data can be primarily attributed to an artifact of the history of archaeological exploration (Loftus et al., 1999). Considering the dominance of lineage B and archaeological evidence in the Near East, it is suggested that sheep lineage B was probably first domesticated in the Near East region. So far, no sufficient sequence data can be used to resolve the where of sheep lineages A and C domestication, because domestic and wild sheep in wide geographic regions are still poorly sampled. However, our results strongly support the hypothesis that three geographically independent domestication events have occurred in domestic sheep (including other regions other than the Near East).

Interestingly, the pattern of genetic variation in lineage A of Chinese domestic sheep revealed two subclades (subclades a1 and a5) with star-like phylogeny (Fig. 2), in which the subclade a5 was only present in China, especially North China, but not in other regions. Individuals belonging to haplotypes of subclade a5 in Chinese sheep are present at a total frequency of 6.2% (28/449) (see Supplementary Fig. S1). It would be expected that haplotypes belonging to subclade a5 can be observed in other regions with lower frequency. However, there is no any haplotype belonging to subclade a5 found in the 229 sheep sequences from other regions. Assuming the domestication of lineage A from the wild ancestral population including two divergent maternally related sheep, it is unreasonable that the haplotypes of subclade a5 disappeared in sheep from other regions due to random genetic drift, but were only maintained in Chinese sheep. It is worthwhile to note that the mutation steps between the two central haplotypes a1 and a5 were eight (Fig. 2), which were only less than three or two mutation steps from haplotype a1 to central haplotype b10 of lineage B and to central haplotype c6 of lineage C (Fig. 3), respectively. In addition, the estimated divergence time based on the Cyt *b* gene between the two founder haplotypes a1 and a5 was about 105,800 to 148,100 years ago, which was much longer than the time of sheep domestication around 8000 to 9000 years ago (Ryder, 1984). These results suggested that two independent domestication events have occurred in sheep lineage A. Given the lack of samples of domestic sheep and *O. orientalis* from Indus Valley and Central Asia, the quite confused taxonomy of *Ovis* genus, and no records on the distribution of *O. orientalis* in North China, data presented here does not permit us to infer geographic locations of lineage A domestication.

The analyses of population genetic structure showed a striking degree of homogeneity among different Chinese sheep populations. This is revealed not only by the MJ network (Figs. 2 and 3) and the MDS plot (Fig. 4), but also by the AMOVA results (Table 3). In other words, it indicated that there was no significant geographical structuring of mtDNA variation among Chinese sheep populations. This result was in agreement with phylogeographic patterning observed in Chinese indigenous goats (Chen et al., 2005) and Portuguese autochthonous goats (Pereira et al., 2005). Because our sampling only included the autochthonous breeds from remote geographic locations as what Luikart et al. (2001) and Chen et al. (2005) did, it is unlikely that weak structuring among Chinese domestic sheep was due to the very recent transport of sheep among different regions in China. Notably, the intercontinental subdivision (Asia and Europe) accounted for ≈20% of the total mtDNA variation in sheep (Table 3), which was slightly higher than that observed in domestic goats ≈14% (Luikart et al., 2001), but was much lower than that found in cattle ≈84% (Bradley et al., 1996). The fundamental factor that could invoke to explain weak structuring is the extensive levels of gene flow induced by human migratory movements and commercial trade in history (Luikart et al., 2001).

In summary, we observed considerable mitochondrial diversity in Chinese domestic sheep. The high levels of intra-population diversity (most Chinese sheep populations contain all

three lineages) and the observation of weak population substructure suggested that three geographically independent domestication events (including other regions other than the Near East) have occurred. Noticeably, two subclades with two central founder haplotypes were identified in Chinese sheep sequences belonging to lineage A. The pattern of genetic variation in lineage A, combined with the divergence time between the two central haplotypes, indicated that two independent domestication events have occurred in domestic sheep lineage A. Our findings have significant implications that the origins of domestic sheep were more complicated and fascinating than previously thought and that the domestication place was not only confined to the Near East, but also occurred in other regions. Further extensive sampling across worldwide geographic regions, especially, Indus Valley and Central Asia, will provide new insights into the origins of lineages A and C. Weak structuring of domestic sheep was observed either among Chinese indigenous sheep populations or between continental grouping, and this can be attributable to strong gene flow induced by historical human movements. It suggests that domestic sheep, like goats, might have played important roles in historical human migration, trade and commerce.

Acknowledgements

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

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2006.03.009.

References

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

Beja-Pereira, A., et al., 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. U. S. A. 93, 5131–5135.

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

Chen, S.Y., Su, Y.H., Wu, S.F., Sha, T., Zhang, Y.P., 2005. Mitochondrial diversity and phylogeographic structure of Chinese domestic goats. Mol. Phylogenet. Evol. 37, 804–814.

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.

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

Guo, J., et al., 2005. A novel maternal lineage revealed in sheep (*Ovis aries*). Anim. Genet. 36, 331–336.

Hiendleder, S., Mainz, K., Plante, Y., Lewalski, H., 1998a. 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., Lewalski, H., Wassmuth, R., Janke, A., 1998b. The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. J. Mol. Evol. 47, 441–448.

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.

Kuo, F., Needham, J., Chhêng, C.T., 1999. The History of Zoology in China. Science Press, Beijing. In Chinese.

Li, X., Tian, Q., Liu, Z., 2001. Study on the RFLP pf mtDNA in several native sheep breeds. Acta Vet. Zootech. Sin. 32, 295–298 (in Chinese).

Loftus, R.T., MacHugh, D.E., Cunningham, P., Bradley, D.G., 1999. Animal domestication. Science 283, 327.

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. U. S. A. 98, 5927–5932.

Meadows, J.R.S., et al., 2005. Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe. J. Heredity 96, 494–501.

Pedrosa, S., Uzun, M., Arranz, J.J., Gutierrez-Gil, B., Primitivo, F.S., Bayon, Y., 2005. Evidence of three maternal lineages in near eastern sheep supporting multiple domestication events. Proc. R. Soc. Lond., B 272, 2211–2217.

Pereira, F., Pereira, L., Van Asch, B., Bradley, D.G., Amorim, A., 2005. The mtDNA catalogue of all Portuguese autochthonous goat (*Capra hircus*) breeds: high diversity of female lineages at the western fringe of European distribution. Mol. Ecol. 14, 2313–2318.

Reed, C., 1960. A review of the archaeological evidence on animal domestication in the prehistoric Near East. In: Braidwood, R.J., Howe, B. (Eds.), Prehistoric Investigations in Iraqi Kurdistan, Studies in Ancient Oriental Civilization. Univ. of Chicago Press, Chicago, pp. 119–145.

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

Ryder, M.L., 1984. Sheep. In: Mason, S.L. (Ed.), Evolution of Domesticated Animals. Longman, London, pp. 63–85.

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

Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin: a Software for Population Genetics Data Analysis. Ver 2.000. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129, 555–562.

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 penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Troy, C.S., et al., 2001. Genetic evidence for Near-Eastern origins of European cattle. Nature 410, 1088–1091.

Tu, Z., Zhang, Y., 1998. Mitochondrial DNA polymorphism in Zang sheep based on restriction endonuclease cleavage patterns. Acta Vet. Zootech. Sin. 29, 132–135 (in Chinese).

Tu, Y., et al., 1989. Sheep and Goat Breeds in China. Shanghai Scientific & Technical Publishers, Shanghai. In Chinese.

Wood, N.J., Phua, S.H., 1996. Variation in the control region sequence of the sheep mitochondrial genome. Anim. Genet. 27, 25–33.

Wu, C.H., Zhang, Y.P., Bunch, T.D., Wang, S., Wang, W., 2003. Mitochondrial control region sequence variation within the argali wild sheep (*Ovis ammon*): evolution and conservation relevance. Mammalia 67, 109–118.

Zeuner, F.E., 1963. A History of Domesticated Animals. Hutchinson, London.