Open Access
Asian Australas. J. Anim. Sci.
Vol. 29, No. 5 : 631-639 May 2016
http://dx.doi.org/10.5713/ajas.15.0473
www.ajas.info
pISSN 1011-2367 eISSN 1976-5517

Complete Mitochondrial Genome Sequences of Chinese Indigenous Sheep with Different Tail Types and an Analysis of Phylogenetic Evolution in Domestic Sheep

Hongying Fan[1,2,a], Fuping Zhao[2,a], Caiye Zhu[2], Fadi Li[3,*], Jidong Liu[4], Li Zhang[2], Caihong Wei[2], and Lixin Du[1,2,*]

[1] College of Animal Science and Technology, Gansu Agricultural University, Lanzhou, Gansu 730070, China

ABSTRACT: China has a long history of sheep (*Ovis aries* [*O. aries*]) breeding and an abundance of sheep genetic resources. Knowledge of the complete *O. aries* mitogenome should facilitate the study of the evolutionary history of the species. Therefore, the complete mitogenome of *O. aries* was sequenced and annotated. In order to characterize the mitogenomes of 3 Chinese sheep breeds (Altay sheep [AL], Shandong large-tailed sheep [SD], and small-tailed Hulun Buir sheep [sHL]), 19 sets of primers were employed to amplify contiguous, overlapping segments of the complete mitochondrial DNA (mtDNA) sequence of each breed. The sizes of the complete mitochondrial genomes of the sHL, AL, and SD breeds were 16,617 bp, 16,613 bp, and 16,613 bp, respectively. The mitochondrial genomes were deposited in the GenBank database with accession numbers KP702285 (AL sheep), KP981378 (SD sheep), and KP981380 (sHL sheep) respectively. The organization of the 3 analyzed sheep mitochondrial genomes was similar, with each consisting of 22 tRNA genes, 2 rRNA genes (12S rRNA and 16S rRNA), 13 protein-coding genes, and 1 control region (D-loop). The NADH dehydrogenase subunit 6 (ND6) and 8 tRNA genes were encoded on the light strand, whereas the rest of the mitochondrial genes were encoded on the heavy strand. The nucleotide skewness of the coding strands of the 3 analyzed mitogenomes was biased toward A and T. We constructed a phylogenetic tree using the complete mitogenomes of each type of sheep to allow us to understand the genetic relationships between Chinese breeds of *O. aries* and those developed and utilized in other countries. Our findings provide important information regarding the *O. aries* mitogenome and the evolutionary history of *O. aries* inside and outside China. In addition, our results provide a foundation for further exploration of the taxonomic status of *O. aries*. (**Key Words:** Mitochondrial DNA, Mitogenome, Sheep, Phylogenetic, Mitochondrial Control Region)

INTRODUCTION

The origins of domestic sheep (*Ovis aries* [*O. aries*]) breeds remain a controversial topic. Domestic sheep can interbreed with mouflon (*Ovis orientalis orientalis*), urial (*Ovis orientalis vignei*), and argali (*Ovis ammon*), complicating study of the origins of sheep breeds (Guo et al., 2005). China has a centuries-old tradition of sheep domestication, production, and breeding (Ma et al., 2006). In China, sheep are an important source of meat, wool, skin, and dairy products that fulfill agricultural, economic, cultural, and religious roles (Chen et al., 2006). At least 42 sheep breeds are indigenous to China (Resources, 2011); however, the origins and phylogenic relationships of these breeds are uncertain.

Evaluation of the structure and function of mitochondrial DNA (mtDNA) provides information useful in the study of molecular evolution, classification, population genetic analysis, and relationship identification.

* Corresponding Author: Fadi Li. Tel: +86-931-763-1225; Fax: +86-931-763-2468, E-mail: lifd@lzu.edu.cn / Lixin Du. Tel: +86-10-6281-8815, Fax: +86-10-6281-8815, E-mail: lxdu@263.net
[2] Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China.
[3] College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou, Gansu 730020, China.
[4] Institute of Grass Ecology Research, Hulunbuir College, Hulunbuir 021008, Inner Mongolia Autonomous Region, China.
[a] These authors contributed equally to this work.
Submitted May 30, 2015; Revised Jul. 23, 2015; Accepted Aug. 12, 2015

Analyses of mtDNA have been used to illuminate the maternal origins of populations (Bruford et al., 2003). Animal mtDNA is a small extrachromosomal genome that is typically 15 to 20 kb in size. With few exceptions, all animal mitochondrial genomes contain the same 37 genes: 2 ribosomal RNAs (rRNAs), 13 protein-coding genes and 22 transfer RNAs (tRNAs) (Boore, 1999). Previous studies constructed phylogenetic trees using the control region sequence (which includes the D-loop) and mitochondrial cytochrome b (*CYTB*) gene (Pedrosa et al., 2005; Gu et al., 2007). The earliest studies of mtDNA variation in domestic *O. aries* breeds exposed 3 distinct lineages (Wood and Phua, 1996; Hiendleder et al., 1998a, b; Guo et al., 2005).

In this study, we evaluated 3 *O. aries* breeds indigenous to China, the Altay sheep, Shandong large-tailed sheep, and small-tailed Hulun Buir sheep, which were of the fat-rump tailed, fat-tailed, and small-tailed varieties, respectively. We sequenced the complete mitogenomes of each of the 3 types of sheep and confirmed the mitochondrial genome structure. The primary features of the sheep mitogenome, including gene structure, gene arrangement, initiation codon, termination codon, and anticodon, were described and compared among the 3 types of sheep. The A/T-content of each mitogenome was calculated by DNA frequency analysis. The mtDNA D-loop region and *CYTB* gene have been widely used to investigate the origins of sheep (Xin et al., 2006; Wang et al., 2007). However, the mtDNA D-loop region is best-suited to studies of the phylogeny of closely related breeds. Therefore, to facilitate the study of the evolutionary relationships of distantly related species, we analyzed the complete mitogenome. In addition, similar to other analyses of complete sheep mitogenomes, we performed a molecular phylogenetic analysis using the complete mitogenome and mtDNA control region via unweighted pair group method with arithmetic means (UPGMA) method (Tamura and Nei, 1993). This study will facilitate further investigations of the phylogenetic relationships of *O. aries* and provides important annotation information for the sheep mitogenome.

MATERIALS AND METHODS

Sample collection and DNA extraction

Blood samples were collected from Altay sheep (AL) from Fuhai County in the Xinjiang Uygur Autonomous Region, Shandong large-tailed sheep (SD) from Liaocheng City in Shandong Province, and small-tailed Hulun Buir sheep (sHL) from the Autonomous County of Evenki in the Inner Mongolia Autonomous Region. The samples were collected only from purebred, healthy, young female sheep.

Blood samples (approximately 10 cc) were drawn from the jugular veins of 40 unrelated domestic sheep into K3 ethylenediaminetetraacetic acid vacuum tubes by licensed veterinarian experts. Total genomic DNA from each individual was extracted using the QIAGEN DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) and stored at –20°C.

Polymerase chain reaction amplification and sequencing

Based on alignments and comparisons of complete mitochondrial sequences of other sheep (Hiendleder et al., 1998a), 19 sets of primers were employed to amplify contiguous, overlapping segments of the mtDNA from each type of sheep (Table 1). Polymerase chain reaction (PCR)

Table 1. PCR primers used for the analysis of sheep mitochondrial genomes

No.	Forward (5′–3′)	Reverse (5′–3′)	Product length (bp)	Tm (°C)
1	AACTTAAAGCAAGGCACT	TTTACTTGAGGAGGGTGA	918	51
2	AAATGACAATTCCCAACC	GATGTAGGGAGAAAATAGTTAGATC	916	54
3	ATGGTTGAGGCCGGAGCA	TCAGAGTATCGTCGTGGT	997	56
4	ATGAGCCAAAATCCACTT	ATTTCTGAGCATTGACCG	1,060	51
5	CAATTCCAGGTCGTTTAA	CTTATTAGTGCAAGGGTG	945	51
6	TGAAACCATCAGCCTATT	TACGAAGTGTCAGTATCAGG	984	54
7	TCCTAATTGTCTGCTTCT	ATAGTCAGGTTAGGGGTA	971	51
8	TACCCCTAACCTGACTAT	CAAGTGCTATGTGGCTAA	958	52
9	AGCTCAATTTGCCTTCGC	GAGGGTTTGGATGGTTAG	985	54
10	TCATGCGCTTTCATCACT	TGTCTTGCTCGTCGTTTA	946	52
11	CCACCTAGCATTCCTTCA	TTGCTGCTTTATACAGTTATGG	940	54
12	TCCACCACAATCACAAGC	GGTAATAACTGTTGCTCCT	1,127	54
13	ATTTGCGACAATAGCCACAG	GTTGATAATGATTCAGGGA	932	55
14	GAACAACCAACCTCCCTA	TGTGACCCAGGTGCCTAT	1,255	55
15	TAATAGCCAAAGGAGGAT	CTTAAACTTGTGCGAGGA	950	51
16	ACTGGACTATTCTATTACTAA	GAACTTTTCGTTCAACTA	1,124	49
17	CCAGTTAAGGTGGCAGAG	GTAATTGGGAGCACGAAG	1,163	55
18	CCCACAAACATAAGAAAT	TTTGAAGGCTCTTGGTCT	1,198	50
19	CTATGGCCGTCTGAGGCCTG	TAAGCAAGGCGTTGTGAG	484	51

PCR, polymerase chain reaction; Tm, temperature.

amplification was performed using an ABI 9700 Thermocycler. Each PCR reaction was carried out in a 25-μL reaction volume consisting of 12.5 μL premixed enzyme, 2.0 μL DNA template, 1.0 μL of each primer (10 ppm), and 8.5 μL sterile deionized water. The PCR conditions were as follows: an initial denaturation at 94°C for 3 min, followed by 35 cycles of 94°C for 30 s (denaturation), 60°C for 30 s (annealing), and 72°C for 1 min (extension), with a final extension at 72°C for 10 min. Each sample of amplified DNA (5 μL) was inspected by electrophoresis on a 1% agarose gel to confirm the length of the amplified fragment.

The second data set used in this study was 16 Chinese sheep and 13 foreign sheep from National Center for Biotechnology Information (NCBI) (Table 2). These data set were used to construct the phylogenetic tree about Chinese and foreign sheep.

Genome organization and structure analysis

All of the sequences were inspected and assembled using DNASTAR (DNASTAR, Inc., Madison, WI, USA) and DNAMAN 7.0 (Lynnon LLC., San Ramon, CA, USA) software, after which the target sequences were assessed by basic local alignment search tool (BLAST) searches of the NCBI database. Thirteen typical protein-coding genes and 2 rRNA genes were identified by BLAST searches of the NCBI database (Hu and Gao, 2014). Twenty-two tRNA genes and anticodons were annotated by tRNAscan-SE Search Server v.1.21 online (http://lowelab.ucsc.edu/tRNAscan-SE/) (Lowe and Eddy, 1997). Initiation and termination codons were identified by comparison of the obtained genomes with fully sequenced *O. aries* mitogenomes. The complete mitochondrial genome was annotated using the program Sequin (National Center for Biotechnology Information, Bethesda, MD, USA). A nucleotide structure chart was generated for the entire mitogenome using OGDRAW (http://ogdraw.mpimp-golm.mpg.de/). The A+T-content of the entire mitogenome was calculated using the EditSeq program included in the Lasergene software package (DNASTAR, Inc., USA), after which C, G, A, and T frequencies were calculated. Skewness was evaluated using the following formulas: GC-skew = (G–C)/(G+C) and AT-skew = (A–T)/(A+T).

Table 2. List of complete sheep mitogenomes available in the NCBI GenBank database

No.	Species	Genbank No.	Length (bp)
1	Tashkurgan sheep	KF938337.1	16,618
2	Yecheng sheep	KF938338.1	16,618
3	Qira Black sheep	KF938326.1	16,617
4	Sunite sheep	KF938317.1	16,543
5	Bashbay sheep	KF938330.1	16,618
6	Bayinbuluke sheep	KF938331.1	16,618
7	Tan sheep	KF938336.1	16,618
8	Small tailed Han sheep	KF977847.1	16,617
9	Duolang sheep	KF938332.1	16,618
10	Kirghiz sheep	KF938334.1	16,617
11	Qinghai Tibetan sheep	KF938325.1	16,617
12	Ujimqin sheep	KF938319.1	16,620
13	Hetian sheep	KF938322.1	16,620
14	Baerchuke sheep	KF938321.1	16,620
15	Turfan Black sheep	KF938324.1	16,620
16	Lanzhou large-tailed sheep	KF938335.1	16,617
17	Sahelian sheep	KF977846.1	16,617
18	Finnsheep	KF938355.1	16,617
19	Kulunda sheep	KF938358.1	16,617
20	Andi sheep	KF938340.1	16,617
21	Oxford down sheep	KF938359.1	16,617
22	Swiniarka sheep	KF938349.1	16,617
23	Pramenka sheep	KF938347.1	16,617
24	Mountain carpathian sheep	KF938357.1	16,617
25	Djallonke sheep	KF977845.1	16,617
26	Karakul sheep	KF938348.1	16,618
27	Minxian Black Fur sheep	KF938318.1	16,620
28	Ammon hodgsoni sheep (mouflon)	JX101654.1	16,688
29	Ammon isolate sheep (argali)	HM236188.1	16,613

NCBI, National Center for Biotechnology Information.

Phylogenetic analysis

The complete mitogenomes of 16 Chinese sheep breeds and 13 foreign sheep breeds were obtained from the National Center for Biotechnology Information (Table 2) and used to construct a phylogenetic tree. Clustering analysis was performed using MEGA 6.0 software (Tamura et al., 2013).

RESULTS

Genome organization and structure

To characterize the mitochondrial genomes of sHL, AL, and SD sheep, 19 sets of primers were employed to amplify contiguous, overlapping segments of complete mtDNA from each breed. The complete mitochondrial genomes of AL, SD, and sHL sheep have been submitted to the NCBI GenBank database (accession numbers: KP981380 [sHL], KP702285 [AL], and KP981378 [SD]). The mitochondrial genomes of the sHL, AL, and SD breeds were circular molecules that were 16,617 bp, 16,613 bp, and 16,613 bp in size, respectively.

The mitochondrial genomes of the sHL, AL, and SD breeds were similar in size to the mitogenomes of other sheep breeds (Hu and Gao, 2014). In addition, the gene arrangements and transcriptional directions of the mitochondrial genomes of the sHL, AL, and SD breeds were similar to those of typical *O. aries* mitogenomes (Hiendleder et al., 1998a, b; Hu and Gao, 2014). The mitochondrial genes and D-loop region arrangements of the

mitochondrial genomes of the sHL, AL, and SD breeds were identical. The nucleotide structure of the sHL mitogenome is shown Figure 1.

The nucleotide skewness of the coding strands of the mitogenomes of the sHL, AL, and SD breeds (AT-skew = 0.10, GC-skew = –0.33) was biased toward A and T (Table 3). The A-T rich pattern of the mitogenomes of the sHL, AL, and SD breeds was typical of vertebrate mitogenomes (Mayfield and McKenna, 1978; Hiendleder et al., 1998b). The nucleotide composition of the coding strand was biased toward T and A in C, similar to most *O. aries* mitogenomes (Jiang et al., 2013). In addition, some variation in the nucleotide skewness of the whole mitogenome and control region was identified among the sHL, AL, and SD breeds (Table 3).

Four gene overlaps were found, all of which were on the H-strand. In addition, 17 intergenic spacers were detected, with 11 on the heavy strand (H-strand) and 6 on the light strand (L-strand). The overlaps were 1 to 40 bp in size, whereas the intergenic spacers were 1 to 32 bp in size (Table 4). The longest overlap (40 bp) was located between ATP synthase protein 8 (ATP8) and ATP synthase protein 6 (ATP6). The longest intergenic spacer (32 bp) occurred between tRNA-Asn and tRNA-Cys.

The sHL, AL, and SD mitogenomes each contained 13 protein-coding genes, mitochondrially encoded NADH dehydrogenase 1–6 (ND1–6), mitochondrially encoded NADH dehydrogenase 4L (ND4L), mitochondrially encoded cytochrome c oxidase I (CO1), mitochondrially encoded cytochrome c oxidase II (CO2), mitochondrially encoded cytochrome c oxidase III (CO3) and ATP6, ATP8, and CYTB. ND1–6 were encoded on the L-strand, whereas the other 12 genes were encoded on the H-strand (Table 4). The start codon of ND2, ND3, and ND5 was ATA, whereas the start of the other 10 genes was ATG. ND1, ND2, ATP6, COX3, ND3, and ND4 were inferred to end with an incomplete stop codon (T--), whereas the stop codon of the other genes was TAA. With the exceptions of COX2 and ATP8, the different A/T content of protein-coding genes is different from the other breeds. ATP8 had the highest A/T content (69.15%, 69.15%, and 69.15% in the sHL, AL, and SD breeds, respectively), whereas COX3 had the lowest A/T content (54.72%, 54.21%, and 54.21% in the sHL, AL, and SD breeds, respectively).

The 22 tRNA encoded in the complete mitochondrial genome of the sHL, AL, and SD breeds formed a typical clover-leaf secondary structure, with the exception of tRNAser (AGN), in which the DHU arm formed a simple loop. The 22 tRNA genes ranged in size from 66 to 75 bp. Fourteen of the tRNA genes were encoded on the H-strand, whereas 8 tRNA genes were encoded on the L-strand (Table 4). With the exceptions of tRNA-Asn, tRNA-Asp, tRNA-

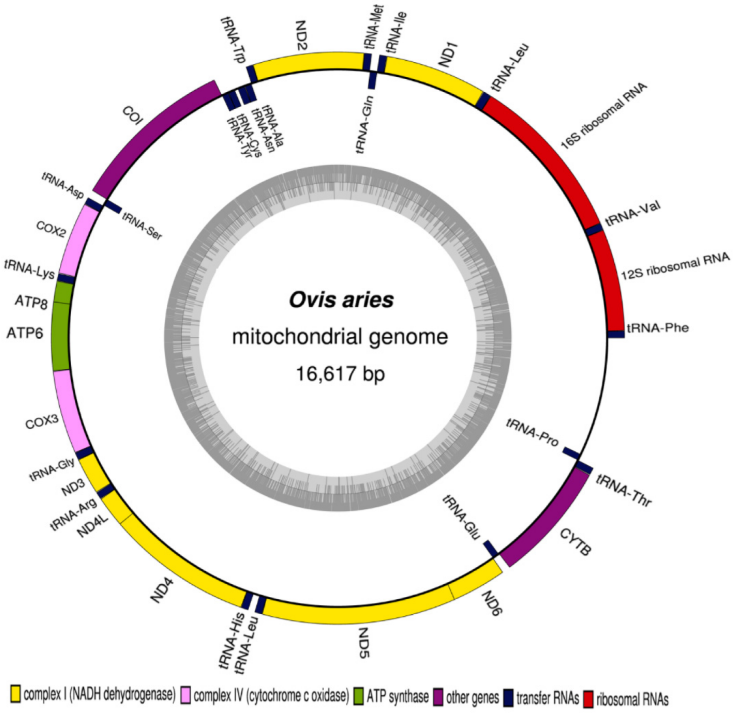

Figure 1. Mitochondrial gene arrangements of *Ovis aries*. Gene organization of the small-tailed Hulun Buir sheep for instance. Different color areas represent 13 protein-coding, 22 tRNA genes, 12S and16S rRNA genes and D-loop region. All protein-coding genes are encoded on the H-strand, with the exception of ND6 which is encoded on the L-strand. The two ribosomal RNA genes are encoded on the H-strand. Transfer RNA genes are designated by single-letter amino acid codes. Transfer RNA genes encoded on the H-strand and L-strand are shown outside and inside the circular gene map, respectively.

Table 3. Base compositions of sheep mitochondrial genomes

Region	A/T content (%)	Base composition (%)				AT-skewness	GC-skewness
	sHL/AL/SD	A	G	T	C		
Protein-coding gene							
ND1	59.37/60.00/60.00	31.73/32.25/32.25	11.31/10.89/10.89	27.64/27.75/27.75	29.32/29.11/29.11		
ND2	64.01/63.63/63.63	37.14/36.37/37.04	8.25/8.35/8.35	26.87/26.58/26.58	27.74/28.02/28.02		
COI	59.48/59.35/59.35	29.13/28.93/28.93	16.12/16.44/16.44	30.36/30.42/30.42	24.40/24.21/24.21		
COX2	62.28/62.28/62.28	35.23/35.09/35.09	13.16/13.30/13.30	27.05/27.19/27.19	24.56/24.42/24.42		
ATP8	69.15/69.15/69.15	41.79/41.79/41.79	5.97/5.97/5.79	27.36/27.36/27.36	24.88/24.88/24.88		
ATP6	60.09/59.65/59.65	31.52/31.52/31.52	11.78/11.63/11.63	28.57/28.13/28.13	28.13/28.72/28.72		
COX3	54.72/54.21/54.21	26.15/25.64/25.64	15.31/15.56/15.56	28.57/28.57/28.57	29.97/30.23/30.23		
ND3	57.80/58.38/58.38	30.64/30.64/30.64	11.85/11.85/11.85	27.17/27.75/27.75	30.35/29.77/29.77		
ND4L	60.27/60.61/60.61	30.30/30.30/30.30	12.46/12.46/12.46	29.97/30.30/30.30	27.27/26.94/26.94		
ND4	60.96/60.81/60.81	31.93/32.00/32.00	10.74/10.67/10.67	29.03/28.81/28.81	28.30/28.52/28.52		
ND5	61.18/61.07/61.07	33.77/33.66/33.66	10.21/10.16/10.16	27.40/27.40/27.40	28.61/28.78/28.78		
ND6	64.96/64.77/64.77	42.99/42.99/42.99	7.01/6.82/6.82	21.97/21.78/21.78	28.03/28.41/28.41		
Cyt b	58.68/58.60/58.60	31.49/31.40/31.40	12.89/12.98/12.98	27.19/27.19/27.19	28.42/28.42/28.42		
tRNA gene							
tRNA-Phe	61.76/61.76/61.76	36.76/36.76/36.76	19.12/19.12/19.12	25.00/25.00/25.00	19.12/19.12/19.12		
tRNA-Val	70.15/70.15/70.15	37.31/37.31/37.31	13.43/13.43/12.43	32.84/32.84/32.84	16.42/16.42/16.42		
tRNA-Leu	60.00/60.00/60.00	32.00/32.00/32.00	17.33/17.33/17.33	28.00/28.00/28.00	22.67/22.67/22067		
tRNA-Ile	73.91/73.91/73.91	39.13/39.13/39.13	15.94/15.94/15.94	34.78/34.78/34.78	10.14/10.14/1014		
tRNA-Gln	61.11/61.11/61.11	36.11/36.11/36.11	9.72/9.72/9.72	25.00/25.00/25.00	29.17/29.17/29.17		
tRNA-Met	55.07/55.07/55.07	27.54/27.54/27.54	18.84/18.84/18.84	27.54/27.54/27.54	26.09/26.09/26.09		
tRNA-Trp	64.18/64.18/64.18	37.31/37.31/37.31	16.42/16.42/16.42	26.87/26.87/26.87	19.40/19.40/19.40		
tRNA-Ala	66.67/66.67/66.67	39.13/39.13/39.13	10.14/10.14/10.14	27.54/27.54/27.54	23.19/23.19/23.19		
tRNA-Asn	54.79/56.16/56.16	31.51/31.51/31.51	15.07/15.07/15.07	23.29/24.66/24.66	30.14/28.77/28.77		
tRNA-Cys	58.82/58.82/58.82	27.94/27.94/27.94	19.12/19.12/19.12	30.88/30.88/30.88	22.06/22.06/22.06		
tRNA-Tyr	66.18/66.18/66.18	29.41/29.41/29.41	14.71/14.71/14.71	36.76/36.76/36.76	19.12/19.12/19.12		
tRNA-Ser	57.97/57.97/57.97	33.33/33.33/33.33	15.94/15.94/15.94	24.64/24.64/24.64	26.09/26.09/26.09		
tRNA-Asp	67.65/67.65/66.67	39.71/39.71/39.13	14.71/14.71/14.49	27.94/27.94/27.54	17.65/17.65/18.84		
tRNA-Lys	66.18/63.24/63.24	33.82/30.88/30.88	14.71/17.65/17.65	32.35/32.35/32.35	19.12/19.12/19.12		
tRNA-Gly	66.67/66.67/66.67	33.33/33.33/33.33	14.49/14.49/14.49	33.33/33.33/33.33	18.84/18.84/18.84		
tRNA-Arg	74.63/74.63/75.00	40.30/40.30/39.41	11.94/11.94/11.76	34.33/34.33/35.29	13.43/13.43/13.24		
tRNA-His	73.91/73.91/73.91	39.13/39.13/39.13	13.04/13.04/13.04	34.78/34.78/34.78	13.04/13.04/13.04		
tRNA-Ser	62.30/63.93/63.93	32.79/34.43/34.43	16.39/14.75/14.75	29.51/29.51/29.51	21.31/21.31/21.31		
tRNA-Leu	67.61/67.61/67.61	39.44/39.44/39.44	16.90/16.90/16.90	28.17/28.17/28.17	15.49/15.49/15.49		
tRNA-Glu	65.22/65.22/65.22	39.13/39.13/39.13	11.59/11.59/11.59	26.09/26.09/26.09	23.19/23.19/23.19		
tRNA-Thr	60.00/60.00/60.00	35.71/35.71/35.71	17.14/17.14/17.14	24.29/24.29/24.29	22.86/22.86/22.86		
tRNA-Pro	57.58/57.58/57.58	34.85/34.85/34.85	13.64/13.64/13.64	22.73/22.73/22.73	28.79/28.79/28.79		
12SrRNA	59.08/59.08/59.08	36.85/36.85/36.85	17.75/17.75/17.75	22.23/22.23/22.23	23.17/23.17/23.17		
16SrRNA	62.19/62.19/62.19	37.43/37.43/37.43	17.06/17.06/17.06	24.76/24.76/24.76	20.75/20.75/20.75		
D-loop	62.49/62.63/62.63	32.85/33.05/33.05	14.56/14.41/14.41	29.64/29.58/29.58	22.95/22.97/22.97		
Entire mitogenome	61.04/60.09/61.08	33.65/33.61/33.66	13.14/13.14/13.09	25.82/25.87/25.83	27.39/27.38/27.42	0.10/0.10/0.10	–0.33/–0.33/–0.33

sHL, small-tailed Hulun Buir sheep; AL, Altay sheep; SD, Shandong large-tailed sheep.

Lys, and tRNA-Arg, the tRNAs have equal A/T content in the sHL, AL, and SD breeds. tRNA-Arg had the highest A/T content (74.63%, 74.63%, and 75.00% in the sHL, AL, and SD breeds, respectively), whereas tRNA-Asn had the lowest A/T content (54.79%, 56.16%, and 56.16% in the sHL, AL, and SD breeds, respectively). In the sHL, AL, and SD breeds, the A/T content of the 16S rRNA was 59.08%, whereas that of the 12S rRNA was 62.19%. These findings provide an important foundation for further exploration of the taxonomic status of *O. aries*.

Phylogenetic analysis

We constructed a phylogenetic tree using the complete mtDNA (including the control region) of 32 breeds of sheep via the UPGMA method (Tamura et al., 1993). The complete dataset for the 32 sheep breeds included 2 parts: one part was from our sequencing data, whereas the other part was downloaded from the NCBI. The bootstrap consensus tree inferred from 1,000 replicates was taken to represent the taxonomy of the analyzed sheep breeds (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% of bootstrap replicates were collapsed. The evolutionary distances were computed using the maximum composite likelihood method (Tamura et al., 2004) and provided as the number of base substitutions per site. The analysis involved 32 nucleotide sequences. First, second, third, and noncoding codon positions were included. All positions containing gaps and missing data were

Table 4. Summary of sheep mitochondrial genomes

	Length (bp)	Position		Codons		Anticodon	Intergenic nucleotide	Strand
	sHL/AL/SD	Start	End	Start	Stop			
tRNA-Phe	68/68/68	1/1/1	68/68/68			GAA	0	H
rRNA	958/958/958	69/69/69	1,026/1,026/1,026				0	H
tRNA-Val	67/67/67	1,027/1,027/1,027	1,093/1,093/1,093/1,093			TAC	0	H
rRNA	1,574/1,571/1,571	1,094/1,094/1,094	2,667/2,664/2,664				0	H
tRNA-Leu	75/75/75	2668/2665	2,742/2,739/2,739			TAA	2	H
ND1	955/955/955	2745/2742/2742	3,699/3,696/3,696	ATG	T--		1	H
tRNA-Ile	69/69/69	3701/3698/3698	3,769/3,766/3,766			GAT	–1	H
tRNA-Gln	72/72/72	3767/3764/3764	3,838/3,835/3,835			TTG	2	L
tRNA-Met	69/69/69	3841/3838/3838	3,909/3,906/3,906			CAT	0	H
ND2	1,042/1,042/1,042	3910/3907/3907	4,951/4,948/4,948	ATA	T--		0	H
tRNA-Trp	67/67/67	4952/4949/4949	5,018/5,015/5,015			TCA	1	H
tRNA-Ala	69/69/69	5020/5017/5017	5,088/5,085/5,085			TGC	1	L
tRNA-Asn	73/73/73	5090/5087/5087	5,162/5,159/5,159			GTT	32	L
tRNA-Cys	68/68/68	5195/5192/5192	5,262/5,259/5,259			GCA	0	L
tRNA-Tyr	68/68/68	5263/5260/5260	5,330/5,327/537			GTA	1	L
COI	1,545/1,545/1,545	5332/5329/5329	6,876/6,873/6,873	ATG	TAA		–1	H
tRNA-Ser	69/69/69	6874/6871/6871	6,942/6,939/6,939			TGA	7	L
tRNA-Asp	68/68/68	6950/6947/6947	7,017/7,014/7,014			GTC	1	H
COX2	684/684/684	7019/7016/7016	7,702/7,699/7,699	ATG	TAA		3	H
tRNA-Lys	68/68/68	7706/7703/7703	7,773/7,770/7,770			TTT	1	H
ATP8	201/201/201	7775/7772/7772	7,975/7,972/7,972	ATG	TAA		–40	H
ATP6	679/679/679	7936/7933/7933	8,614/8,611/8,611	ATG	T--		1	H
COX3	784/784/784	8616/8613/8613	9,399/9,396/9,396	ATG	T--		0	H
tRNA-Gly	69/69/69	9400/9397/9397	9,468/9,465/9,465			TCC	0	H
ND3	346/346/346	9469/9466/9466	9,814/9,811/9,811	ATA	T--		2	H
tRNA-Arg	67/67/67	9817/9814/9814	9,883/9,880/9,880			TCG	1	H
ND4L	297/297/297	9885/9882/9882	10,181/10,178/10,178	ATG	TAA		7	H
ND4	1,378/1,378/1,378	10,175/10,172/10,172	11,552/11,549/11,549	ATG	T--		0	H
tRNA-His	69/69/69	11,553/11,550/11,550	11,621/11,618/11,618			GTG	0	H
tRNA-Ser	61/61/61	11,622/11,619/11,619	11,682/11,679/11,679			GCT	0	H
tRNA-Leu	71/71/71	11,683/11,680/11,680	11,753/11,750/11,750			TAG	0	H
ND5	1821/1821	11,754/11,751/11,751	13,574/13,571/13,571	ATA	TAA		–17	H
ND6	528/528/528	13,558/13,555/13,555	14,085/14,082/14,082	ATG	TAA		0	L
tRNA-Glu	69/69/69	14,086/14,083/14,083	14,154/14,151/14,151			TTC	4	L
Cyt b	1139/1139	14,159/14,156/14,156	15,298/15,295/15,295	ATG	TAA		3	H
tRNA-Thr	70/70/70	15,302/15,299/15,299	15,371/15,368/15,368			TGT	0	H
tRNA-Pro	65/65/65	15,371/15,368/15,368	15,436/15,433/15,433			TGG	0	L
D-loop	1,180/1,180/1,180	15,437/15,434/13,434	16,617/16,613/16,613					

sHL, small-tailed Hulun Buir sheep; AL, Altay sheep; SD, Shandong large-tailed sheep.

eliminated. There were a total of 16,530 positions in the final dataset. Clustering analyses were conducted using MEGA6.0 (Tamura et al., 2013).

Phylogenetic analysis of the complete mitogenome showed that the Chinese local sheep breeds were clustered into a single group, whereas the foreign breeds were closely clustered into another group (Figure 2). The mouflon (*O. orientalis orientalis*) (ammon_isolate sheep) and argali (*O. ammon*) (ammon_hodgsoni sheep) were placed into a separate branch as outgroups. The phylogenetic tree constructed using the D-loop of the Chinese and foreign sheep breeds showed a disordered distribution (Figure 3). These results indicate that evolutionary analyses utilizing full mitogenomes can provide information regarding the relationships between distantly related species that cannot be generated by analysis of the D-loop region only.

DISCUSSION

The complete mitochondrial genomic DNA sequences

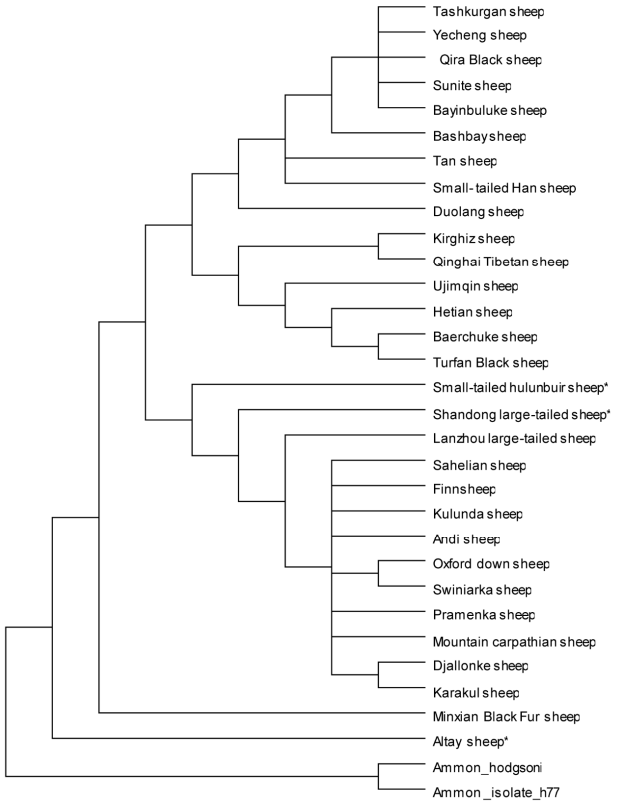

Figure 2. Evolutionary relationships of sheep based on the mitogenomes. Phylogenetic trees of *Ovis aries* constructed by the complete mitochondrial genome of 30 sheep. Ammon hodgsoni and Ammon isolate h77 were selected out groups. The evolutionary distances were computed using the Maximum Composite Likelihood method. Tree topologies was inferred using the UPGMA method. The asterisk indicates the sequence generated in this research.

Figure 3. Evolutionary relationships of sheep based on the D-loop region. Results of phylogenetic analyses via UPGMA method using analysis indicated evolutionary relationships among 30 individuals based on mtDNA D-loop region. Ammon hodgsoni and Ammon isolate h77 were selected out groups. The evolutionary distances were computed using the Maximum Composite Likelihood method. The asterisk indicates the sequence generated in this research. UPGMA, unweighted pair group method with arithmetic means; mtDNA, mitochondrial DNA.

of 3 breeds of indigenous Chinese sheep, AL, SD, and sHL, were 16,613 to 16,617 bp in length and contained 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes, and 1 putative control region. The organization and structure of the mitogenomes of the AL, SD, and sHL sheep were similar to those of other domestic sheep, including the Texel and Merinolandschaf breeds (Hiendleder et al., 1998a, b; Hu and Gao, 2014). The minor differences in the sizes of the genomes of domestic sheep were due primarily to changes in the control region.

The AL, SD, and sHL mitogenomes had similar A/T content (approximately 60% to 61%), which was also similar to that of the Texel, Merino, and Merinolandschaf breeds (Hiendleder et al., 1998a, b; Lancioni et al., 2013; Hu and Gao, 2014). The gene overlaps and intergenic spacers among 1 bp to 7 bp at the large, other than tRNA-Ala and tRNA-Asn, ATP8, and ATP6, also ND5 and ND6 were 32, –40, –17, respectively.

The relative organization of the 37 genes and control regions comprising the mitogenomes of the sHL, AL, and SD breeds was identical with that of other domestic sheep breeds (Hiendleder et al., 1998a, b; Hu and Gao, 2014). The 13 protein-coding genes in the mitogenomes of the sHL, AL, and SD breeds began with initiation codons ATG or ATA, but only 3 genes (ND2, ND3, and ND5) used the latter. About half of the 13 protein-coding genes had incomplete termination codons, whereas the others had the typical TAA stop codon. Incomplete termination codon (T) is in the process of transcription by TAA after the formation affixed Poly (A) terminal translated (Clayton, 1991). Incomplete stop codons are commonly found in many vertebrate mitochondrial genes (Wang et al., 2014).

The phylogenetic tree based on the mitogenome clusters the sHL and SD sheep into the same clade, suggesting they have closer relationship than AL. This agrees with the fact that both of SD and sHL are fat–tailed sheep while AL are fat-rump tailed sheep. In addition to details regarding the organization of the mitochondrial genome of indigenous Chinese sheep, this study provides information regarding the evolutionary relationship between the mitogenomes of indigenous Chinese sheep and foreign sheep. The NJ, ML,

and UPGMA trees constructed according to the maximum composite likelihood method showed the same results; therefore, we showed only the UPGMA tree. Our phylogenetic tree, constructed using the entire mitogenome, clearly shows the evolutionary separation between indigenous Chinese sheep and foreign sheep, whereas the phylogenetic tree constructed using only the control region did not show this relationship.

CONCLUSION

The present study provides the complete mitochondrial genome DNA sequences of the small-tailed Hulun Buir sheep, Altay sheep, and Shandong large-tailed sheep, which are indigenous to China. The mitogenomes of the small-tailed Hulun Buir sheep, Altay sheep, and Shandong large-tailed sheep were 16,617 bp, 16,613 bp, and 16,613 bp in length, respectively, and included a typical set of 37 genes: 13 protein-coding genes, 2 rRNA genes, and 22 tRNA genes, as well as a control region. The gene annotations and genome organizations of the mitogenomes of small-tailed Hulun Buir sheep, Altay sheep, and Shandong large-tailed sheep have been submitted to the GenBank database. Phylogenetic relationships determined via analysis of complete mitochondrial genomes may be more precise than those determined via analysis of only the mitochondrial control region.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Agriculture Science and Technology Innovation Program (no. ASTIP-IAS-TS-6) and the National Modern Agricultural Industry Technology Fund for Scientists in Sheep Industry System (no. CARS-39-04B).

REFERENCES

Boore, J. L. 1999. Animal mitochondrial genomes. Nucl. Acids Res. 27:1767-1780.

Bruford, M. W., D. G. Bradley, and G. Luikart. 2003. DNA markers reveal the complexity of livestock domestication. Nat. Rev. Genet. 4:900-910.

Chen, S. Y., Z. Y. Duan, T. Sha, J. Xiangyu, S. F. Wu, and Y. P. Zhang. 2006. Origin, genetic diversity, and population structure of Chinese domestic sheep. Gene 376:216-223.

Clayton, D. A 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. 7:453-478.

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

Gu, Z., X. Zhao, N. Li, and C. Wu. 2007. Complete sequence of the yak (*Bos grunniens*) mitochondrial genome and its evolutionary relationship with other ruminants. Mol. Phylogenet. Evol. 42:248-255.

Guo, J., L. X. Du, Y. H. Ma, W. J. Guan, H. B. Li, Q. J. Zhao, X. Li, and S. Q. Rao 2005. A novel maternal lineage revealed in sheep (*Ovis aries*). Anim. Genet. 36:331-336.

Hiendleder, S., H. Lewalski, R. Wassmuth, and A. Janke 1998a. 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., K. Mainz, Y. Plante, and H. Lewalski. 1998b. 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. Hered. 89:113-120.

Hu, X. D. and L. Z. Gao. 2014. The complete mitochondrial genome of domestic sheep, *Ovis aries*. Mitochondrial DNA 27:1425-1427.

Jiang, L., G. Wang, S. Tan, S. Gong, M. Yang, Q. Peng, R. Peng, and F. Zou 2013. The complete mitochondrial genome sequence analysis of Tibetan argali (*Ovis ammon hodgsoni*): Implications of Tibetan argali and Gansu argali as the same subspecies. Gene 521:24-31.

Lancioni, H., P. Di Lorenzo, S. Ceccobelli, U. A. Perego, A. Miglio, V. Landi, M. T. Antognoni, F. M. Sarti, E. Lasagna, and A. Achilli. 2013. Phylogenetic relationships of three Italian merino-derived sheep breeds evaluated through a complete mitogenome analysis. PLoS One 8:e73712.

Lowe, T. M. and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl. Acids Res. 25:955-964.

Ma, Y.-H., S.-Q. Rao, S.-J. Lu, G.-Y. Hou, W.-J. Guan, H.-B. Li, X. Li, Q.-J. Zhao, and J. Guo. 2006. Phylogeography and Origin of Sheep Breeds in Northern China. Conserv. Genet. 7:117-127.

Mayfield, J. E. and J. F. McKenna. 1978. A-T rich sequences in vertebrate DNA. Chromosoma 67:157-163.

Pedrosa, S., M. Uzun, J. J. Arranz, B. Gutierrez-Gil, F. San Primitivo, and Y. Bayon. 2005. Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. Proc. Biol. Sci. 272:2211-2217.

Resources (China National Commission of Animal Genetic) 2011. Animal Genetic Resources in China: Sheep and Goats records: China Agriculture Press, Beijing, China.

Tamura, K. and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10:512-526.

Tamura, K., M. Nei, and S. Kumar. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Natl. Acad. Sci. USA. 101:11030-11035.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30:2725-2729.

Wang, X., Y. H. Ma, H. Chen, and W. J. Guan. 2007. Genetic and phylogenetic studies of Chinese native sheep breeds (*Ovis*

aries) based on mtDNA D-loop sequences. Small Rumin. Res. 72:232-236.

Wang, X., Y. Q. Zhang, D. C. He, X. M. Yang, B. Li, D. C. Wang, J. Guang, F. Xu, J. Y. Li, X. Gao, H. J. Gao, L. P. Zhang, and X. Z. Zhang. 2014. The complete mitochondrial genome of Bos taurus coreanae (Korean native cattle). Mitochondrial DNA 27:120-121.

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

Xin, W., M. A. Yue-Hui, and C. Hong. 2006. Analysis of the genetic diversity and the phylogenetic evolution of Chinese sheep based on Cyt b gene sequences. Acta Genet. Sin. 33: 1081-1086.