遗 传 学 报 *Acta Genetica Sinica*, December 2006, 33 (12): 1081–1086 ISSN 0379-4172

Analysis of the Genetic Diversity and the Phylogenetic Evolution of Chinese Sheep Based on Cyt *b* Gene Sequences

WANG Xin[1,2], MA Yue-Hui[2,①], CHEN Hong [1,3,①]

1. *College of Animal Science and Technology*, *Northwest Sci-Tech University of Agriculture and Forestry*, *Yangling* 712100, *China*;

2. *Institute of Animal Science*, *Chinese Academy of Agricultural Sciences*, *Beijing* 100094, *China*;

3. *Institute of Biotechnology*, *Xuzhou Normal University*, *Xuzhou* 221116, *China*

Abstract: The complete sequences of Cyt *b* gene from 20 individuals belonging to eight Chinese indigenous sheep breeds and one foreign breed were studied. The results showed that the hapolotype diversity of Chinese sheep breeds was 97.1%. The mean nucleotide composition of all the sequences was 27.1% T, 28.5% C, 31.4% A, and 13.0% G. The nucleotide diversity was 0.602%. A total of 43 mutation sites were detected, including 40 transitions and 3 transversions. Fu's test of selective neutrality showed that the sheep populations had no population demographic expansion ($0.10 > P > 0.05$). The different clustering methods, namely neighbor-joining, minimum evolution, and unweighted pair group method with arithmetic means, all showed a similar result, which indicated that Chinese local sheep had three maternal resources.

Key words: sheep; Cyt *b*; phylogenetic evolution

Mitochondrial DNA (mtDNA) is the genetic material that exists outside the nucleus in eukaryotic cells. It has a simple molecular structure. It does not undergo recombination with nuclear DNA and has no identical sequence with nuclear DNA. It has multiple copies, has a rapid evolutionary rate, and follows maternal inheritance. Cytochrome *b* gene (Cyt *b*) is one of the genes that is coded by mtDNA, and its gene product plays an important role in electron transfer in the respiration chain. Cyt *b* gene has a moderate evolutionary rate and a clear evolutionary pattern that makes it suitable for the studies on the phylogenetic evolution at the intra- and interspecific levels[1-3].

China has a centuries-old history of breeding sheep. It has abundant sheep breed resources, with more than 40 local sheep breeds, which serve as important genetic resources for the sustainable development of animal husbandry and for the preservation of biological diversities[4]. For example, Ganjia and Oula sheep in Gansu Province belonging to local breeds of Tibetan sheep can adapt to the plateau environment and endure coarse feeding. The Bashibai sheep, which is found in Xijiang Province, has delicate and tasty meat that is easier to digest and has a higher calcium and iron content. In addition, there are other local sheep breeds, such as Henan big-tail sheep and Heiqiupi sheep that possess unique features. Previous studies on these breeds were carried out only at the morphological level. There are still no reports on some of the breeds at the molecular level. The purpose of this study was to investigate the genetic diversity and phylogenetic evolution of Chinese sheep based on the analysis of the complete sequence of the Cyt *b* gene. This will be helpful for the conservation, utilization, and exploitation of the genetic resources of the indigenous Chinese sheep.

Received: 2005-11-23; Accepted: 2006-05-20

This work was supported by the National Natural Science Foundation of China (No. 30371026), Doctor Startup Fund (No. 01140503) and Topnotch Person with Personal Ability Support Project of Northwest A&F University (No. 01140101).

① Corresponding author. E-mail: yuehuima@263.net ; chenhong1212@263.net

1 Materials and Methods

1.1 Sample collection and DNA extraction

According to the simple random sampling method, ear samples from 382 individuals of nine different sheep breeds were collected at livestock farms throughout China. The breeds were chosen from Xinjiang Province (Bashibai sheep, *n*=48; Tashikuergan sheep, *n*=47), Gansu Province (Lanzhou big tail sheep, *n*=42; Ganjia sheep, *n*=47; Oula sheep, *n*=48; and Heiqiupi sheep, *n*=48), Shaanxi Province (Hanzhong sheep, *n*=50), Henan Province (Henan big tail sheep, *n*=52), and Beijing City (skudde, *n*=30). DNA was extracted from these specimens using phenol–chloroform method[5].

1.2 Amplification, purification, cloning, and DNA sequencing

Primers for Cyt *b* gene were designed using Primer 5.0 software and synthesized by Shanghai Sangon Bio-tech Co., Ltd. The length of the amplified fragment was about 1 610 bp, including the complete Cyt *b* gene (1 140 bp) and more than 230 bp sequence in both flanks. The sequences of the primers were forward: 5′-ACACCCAACCCCACCAC-3′, reverse: 5′-GTGGGTGGTTGTGCTTTTCT-3′. The volume of the PCR amplification reaction system was 60 μL, consisting of genomic DNA 50 ng, dNTPs 200 μmol/L, primers 10 pmol, $MgCl_2$ 250 μmol/L, and *Taq* DNA polymerase 1 U. The reaction conditions included an initial denaturation at 95℃ for 5 min, followed by 30 cycles, each consisting of 30 s denaturation at 95℃, primer annealing at 56℃ for 45 s, extension at 72℃ for 60 s, and then a final extension at 72℃ for 8 min. The PCR products were electrophoresed using 2.0% (wt/vol) agarose gel, which was stained with ethidium bromide solution.

The amplified products were purified with a DNA purification kit according to the manufacturer's instructions (TW-Biotech Co., Ltd.). The purified fragments were cloned into pGEM-T easy vector and subsequently transformed into *E. coli* Top10. After 16–20 hours, single colonies were inoculated to obtain recombinant plasmid. The recombinant DNA was extracted and sequenced using an ABI model 3730 automated sequencer. The results that were indicative of unique polymorphisms were confirmed by resequencing of the independent clones.

1.3 Data analysis

The sequences were aligned using the BioEdit 7.0 software. Identical sequences were considered as belonging to the same haplotype. The DnaSP 4.1 program was used to analyze the polymorphism of the haplotypes and to estimate the degree of variation and the substitution frequency of the nucleotides. The molecular phylogenetic trees were constructed to analyze evolution of sheep using the MEGA 2.0 program with a Kimura 2-parameter model[6]. The bootstrap value was 1 000.

2 Results

2.1 Nucleotide analysis of Cyt *b* gene

When compared with the control sequence (Accession No. NC001941), a total of 43 polymorphic sites were obtained from the 21 sequences（from top to bottom in Fig. 1, the GenBank accession numbers range from DQ903208 to DQ903227 except NC001941), including 29 single variable sites and 14 parsimony informative sites (Fig. 1). Transversions occurred only at three positions: 603 (C/G), 693 (T/G), and 1078 (C/A) and in all the other positions, transitions occurred (G/A, 16 and T/C, 24). The transitions of T/C and A/G at positions 309 and 495 could be regarded as characteristics of haplotype B, and transitions of T/C at positions 207, 393, 396, and 696; C/T at 328, 735, and 990; A/G at 476 and 813; and G/A at 939 could be regarded as characteristic of haplotype C.

The variance of nucleotide diversity was 3.77%, the transition rate was 3.51% and the transversion rate was 0.26%. There were 24 synonymous and 19 nonsynonymous substitutions among the 43 variations, and the nucleotide diversity of synonymous and

```
[Domain=Data;
[                                                          1 111]
[          111223333 3444445566 6666777888 8889999990 001]
[         8009030239 9667896902 2499349155 5673339997 890]
[         1123729813 6586358733 9336516314 8843690178 054]
 BSB3     TGCTTATCGT TTTACAGCCC GATTCTTATC TGCGAGCGTC TAA
 GJ12     .......... .......... ......C... .......... ..G
 HN36     C......... .......... .....CC... .......... ...
 HZ1      .......... .......... ..G...C... .......... ...
 LZ5      ..T..G.... ......A... ......C... .......... ...
 LZ6      .A........ ....T..... ......C... .......... ...
 NC001941 ......C... .....G.... ......C... .......... ...
 HZ2      ......C... .....G.... ......C.CT .A..G....A .G.
 GJ14     ......C... .....G.... ......C... .......... ...
 HN31     ......C... .....G.... .G....C... .......A.. .G.
 BSB22    ......C... .....G.... ......C... .......... .G.
 LZ10     ...C..C... .....G.... ......C... .......... .G.
 OL33     ......C... .....G.... ......C... .......... ...
 QP6      ......C.A. ..C..G...T ......C... .........C. C..
 QP11     ......C... .....G.... ......C... C......... ...
 SFK25    ......C... .C...G.... ......C... .......... ...
 SFK30    ......C... .....G.... ......C... .......... ...
 TSK18    ......C... .....G.... ......C... .........C. ...
 BSB7     ....C..T.C C..G...... ...CT.CG.. ...A.AT... ...
 HN32     ....C..T.C C..G...TG. A..CT.CG.. .....AT... ...
 QP9      ....C..T.C C..G...... ...CT.CG.. ..T..AT... ...
```

Lineage A

Lineage B

Lineage C

Fig. 1 Variable sites of Cyt *b* gene

HZ: Hanzhong sheep; HN: Henan big-tail sheep; QK: Qiaoke sheep; OL: Oula sheep; LZ: Lanzhou big-tail sheep; GJ: Ganjia sheep; QP: Minshan black sheep; BSB: Bashibai sheep; TSK: Tashikuergan sheep; SFK: Skudde. The figures that follow them indicate the individuals among the different breeds. The Arabic numbers at the top of the figure indicate the variable sites.

nonsynonymous substitutions was 1.45% and 0.31%, respectively. Transitions at positions 309 and 495 (haplotype B) and at positions 207, 393, 396, 939, and 328 (haplotype C) were synonymous. In contrast, transition of A/G at position 1095 in haplotype B changed the coding amino acid from an Ile to a Met.

2. 2 Construction of the phylogenetic tree

The NJ (Neighbor-joining), ME (Minimum Evolution), and UPGMA (Unweighted Pair Group Method with Arithmetic Means) trees were constructed according to the Nei-Gojobori method using MEGA 2.0 program (Bootstrap value was 1 000). All the three types of trees showed the same result, and only the NJ tree is shown (Figs. 2 and 3). The 21 individuals belonged to three unattached haplotype groups, and they were named lineages A, B, and C. In the tree, lineage A had 6 individuals, B had 12, and C had 3. The result indicated that Chinese sheep were derived from three maternal resources.

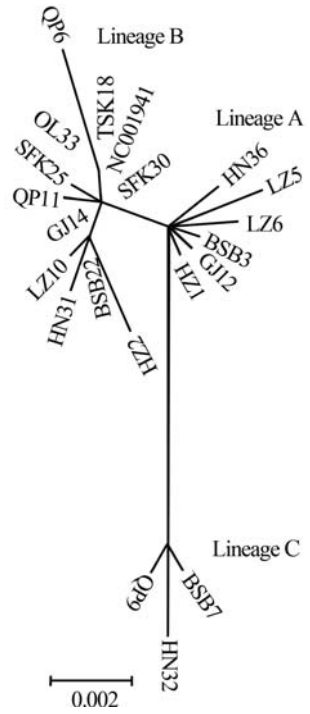

Fig. 2 The NJ phylogenetic tree of 21 individuals at the Cyt *b* gene site

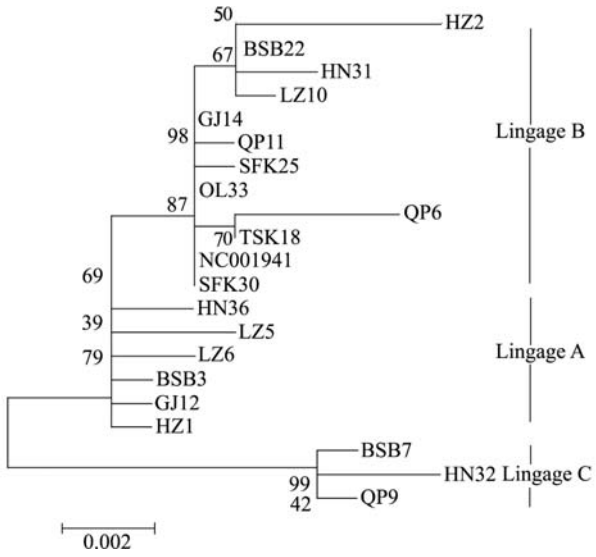

Fig. 3 The NJ rectangle tree of 21 individuals at the Cyt *b* gene site

2. 3 Sequence diversity analysis of Cyt *b* gene

A total of 18 haplotypes were detected in the 21 individuals. BSB7, QP9, OL33, GJ14, and NC001941 had similar haplotypes. The haplotype diversity was 97.1% ± 0.09% and the average number of nucleotide differences (k) was 6.867. The nucleotide diversity (π) was 0.602%. The indices of genetic differentiations among the lineages are shown in Table 1.

In Table 1, Gst is the relative index that measures gene differentiation among the various populations. It refers to the proportion of the mean heterozygosities of between populations to the total heterozygosities of between- and within populations. The smaller the genetic differentiation index is, the nearer is the relationship between the two populations. As the genetic differentiation index between lineages A and C was the lowest, the relationship between them was the nearest, whereas that between lineages B and C was the farthest. Fu's test of selective neutrality showed that the differences were not significant ($D = -1.6835$, $0.05 < P < 0.10$).

3 Discussion

A total of 18 haplotypes were detected in this study, and the haplotype diversity was found to be 97.1% ± 0.09%. The G+C content was 41.5%. It was 4.02% higher than that of the D-loop region[7]. The result obtained was different from that of Zhang *et al*[8]. This difference was probably caused by the difference in the length of the sequences that were studied. The authors of this article studied the complete sequence of the Cyt *b* gene, whereas Zhang *et al.* only studied 450 bp of the gene. Only 43 mutations were found among the 21 sequences, and the nucleotide diversity was 0.602%. It was far lower than that of the D-loop region[9,10], indicating that the Cyt *b* gene is relatively conserved. Most base substitutions did not change the coding of the amino acid. Other studies on Chinese sheep using RFLP marker of mtDNA showed that the genetic diversity of Chinese sheep breeds was lower[11–13]. In addition, Jia *et al.*[14] studied eight sheep breeds in the Xinjiang Province by investigating their nuclear DNA using microsatellite markers, and the result showed that the mean polymorphism information content (*PIC*) of the eight breeds were lower than that of foreign breeds. Taken together, it may be concluded that the genetic diversity of Chinese sheep breeds is low.

Table 1 Results of genetic differentiation among populations

Lineage	Lineage	Gst	Da	Dxy	Fst test
A	B	0.0277	0.0019	0.0047	0.3949
B	C	0.0505	0.0106	0.0135	0.7895
C	A	0.0137	0.0088	0.0117	0.7500

Da: Number of net nucleotide substance per site between lineages; Dxy: Average number of nucleotide substance per site between lineages; Fst test: Fu's test statistic between lineages.

There has been a controversy regarding the origin of the domestic sheep. On the basis of the study on the mtDNA D-loop region of sheep, Hiendleder *et al*.[15] stated that domestic sheep had two different maternal resources. In this study, different clustering methods, namely NJ, UPGMA, and ME, all showed a similar result. All haplotypes of the nine sheep breeds could be divided into three lineages. The result was consistent with that of Guo *et al*.[10], who studied the mtDNA D-loop region of six Chinese sheep breeds. The result indicated that Chinese sheep may have three maternal resources. The Cyt *b* gene has been used to study other aspects, such as intra- or interspecific relationships and gene flow[16-19]. It is generally recognized that the domestic animals experience a bottle-neck effect after domestication. But in this study, none of the breeds experienced population expansion events irrespective of the size of the population. There was no population demographic expansion in the sheep populations, and all the variations were neutral. During the course of the migration and development of sheep, the sign of population expansion was perhaps maintained chronically in a small population. It also may be caused by the quantitative decrease of the population, so the sign was lost subsequently[20].

In conclusion, the genetic diversity of Chinese local breeds is low and evolutionary analysis suggests that Chinese sheep breeds are derived from three different maternal resources.

References:

[1] Irwin D M, Kocher T D, Wilson A C. Evolution of cytochrome *b* gene of mammals. *J Mol Evol,* 1999, 32(2): 128−144.

[2] Kocher T D, Thomas W K, Meyer A, Edwards S V, Pääbo S, Villablanca F X, Wilson A C. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc Natl Acad Sci USA,* 1989, 86(16): 6196−6200.

[3] Vladimir O A, Poltoraus L A, Zhivotovsky V. Mitochondrial DNA sequence diversity in Russians. In: FEBS Lett, 1999, 197−201.

[4] Du L. Several reasons on the appreciation and preservation of Chinese local livestock and birds. *China Anim Husbandry and Vet Med,* 2003, 30(2): 3−5 (in Chinese).

[5] Sambrook J, Fritsch E F, Maniatis T. Molecular cloning: a laboratory manual (third edition). Cold spring Harbor Laboratory, Cold Spring Harbor, NY, 2001.

[6] Kumar S, Tamura K, Jakobsen I B, Nei M. Molecular evolutionary genetics analysis software. (Version 2.1). Bioinformatics. 2001.

[7] Wang X. Study on mtDNA genetic diversity and phylogeny evolution of 9 indigenous sheep breeds in the middle-western of China. Dissertation for doctor degree of Northwest A&F University. Shaanxi, Yangling, 2005 (in Chinese).

[8] Zhang C S, Geng L Y, Wan H W, Li J T, Du L X. Study on sequence variation of mitochondrial cytochrome b gene and phylogenetic relationships of four native sheep breeds. *Acta Vet et Zootech Sinica,* 2005, 36(4): 313−317 (in Chinese with an English abstract).

[9] Hiendleder S. A low rate of replacement substitutions in two major Ovis aries mitochondrial genomes. *Anim Genet*, 1998a, 29(2): 116−122.

[10] Guo J, Du L X, Ma Y H, Guan W J, Li H B, Zhao Q J, Li X, Rao S Q. A novel maternal lineage revealed in sheep (Ovis aries). *Anim Genet*, 2005, 36(4): 331−336.

[11] Lan R, Hong Q H, Gao H Y , Zhang J, Su B, Wang W, Liu A H, Zhang Y P, Shi L M. Mitochondrial DNA polymorphism in sheep from Yunnan. *Hered*, 1998, 20(1): 20−23 (in Chinese with an English abstract).

[12] Jia Y H, Jian C S, Zhu W S, Zhang Y P, Shi X W, Yu Y H, Liao Z L, Li T Q. Mitochondrial DNA polymorphism and genetic relationship of Guizhou sheep. *Southwest China J Agric Sci,* 1998, 12(1): 100−105 (in Chinese with an English abstract).

[13] Li X L, Tian Q Y, Liu Z Z, Sun N Q, Ma G X. Study on the RFLP of mtDNA in several native sheep breeds. *Acta Vet et Zootech Sinica,* 2001, 32(4): 295−298 (in Chinese with an English abstract).

[14] Jia B, Chen J, Zhao R Q, Luo Q J, Yan G Q, Chen J. Microsatellite analysis of genetic and phylogenetic relationship of eight sheep breeds in Xinjiang. *Acta Genetica Sinica,* 2003, 30(9): 847−854 (in Chinese with an English abstract).

[15] Hiendleder S, Mainz K, Plante Y, Lewalski H. 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,* 1998b, 89(2): 113−120.

[16] Cao L R, Wang X M, Fang S G. A molecular phylogeny of bharal and dwarf blue sheep based on mitochondrial

cytochrome b gene sequences. *Acta Zoologica Sinica,* 2003, 49(2): 198−204 (in Chinese with an English abstract) .

[17] Guryev V, Makarevitch I, Blinov A, Martin J. Phylogeny of the Genus Chironomus (Diptera) Inferred from DNA Sequences of Mitochondrial Cytochrome b and Cytochrome oxidase I. *Mol Phylogenet Evol,* 2001, 19(1): 9−21.

[18] Alves E, Ovilo C, Rodriguez M C and Silio L. Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. *Anim Genet,* 2003, 34(5): 319−324.

[19] Alves P C, Ferrand N, Suchentrunk F and Harris D J. Ancient introgression of *Lepus timidus* mtDNA into *L. granatensis* and *L. europaeus* in the Iberian Peninsula. *Mol Phylogenet Evol,* 2003, 27(1): 70−80.

[20] Excoffier L, Schneider S. Why hunter-gatherer populations do not show sign of Pleistocene demographic expansions. *Proc Natl Acad Sci,USA*. 1999, 96(19): 10597−10602.

绵羊 Cyt *b* 基因序列多态性及系统进化研究

王 昕[1,2]，马月辉[2]，陈 宏[1,3]

1. 西北农林科技大学动物科技学院，杨凌 712100；

2. 中国农业科学院畜牧研究所，北京 100094；

3. 徐州师范大学生物技术研究所，徐州 221116

摘 要: 以 8 个中国地方绵羊品种和 1 个外来品种的 20 个个体为研究对象，通过对 Cyt *b* 基因的全序列测定，结果表明：绵羊的单倍型多样性为 97.1%。所有序列的平均碱基组成为 27.1% T，28.5%C，31.4%A 及 13.0%G，G+C 含量为 41.5%，核苷酸多样性为 0.602%。在所有序列中共检测到 43 个变异位点，其中包括 40 处转换和 3 处颠换。Fu's 中性检验表明差异不显著(0.10>*P*>0.05)，说明绵羊群体未发生群体扩张事件。NJ、ME 及 UPGMA 聚类结果均表明，我国绵羊可分为 3 个单倍型组，这提示我国绵羊有 3 个母系起源。

关键词: 绵羊；Cyt *b*；系统进化

作者简介: 王昕（1975−），女，山东莒南人，博士，研究方向：动物遗传资源研究。E-mail: wxwza@126.com