Structure of mitochondrial DNA control region and genetic diversity of *Moschus berezovskii* populations in Shaanxi Province

H. Feng¹², C.L. Feng², Y. Huang¹ and J. Tang²

¹The College of Life Sciences, Shaanxi Normal University, Xi’an, China
²Shaanxi Institute of Zoology, Xi’an, China

Corresponding author: Y. Huang
E-mail: fenghui84@126.com

Received September 2, 2015
Accepted December 8, 2015
Published April 7, 2016
DOI http://dx.doi.org/10.4238/gmr.15027578

ABSTRACT. In China, *Moschus berezovskii* (forest musk deer), a first-class national protected animal, was earlier widely distributed. However, wild populations of the forest musk deer have declined because of human activity and habitat loss. In order to gather useful information for its conservation and management, we investigated the genetic diversity and population structure of this species by analyzing a 632-bp fragment of mitochondrial DNA (mtDNA) control region in three wild populations in Shaanxi Province, China. The average A+T content (64.1%) of mtDNA was higher than that of G+C (35.9%). A total of 178 variable sites (about 28.03% of the total nucleotides in the sequence) were detected in 71 individuals. The nucleotide diversity (θ) in the 71 individuals was 0.04688, and the average nucleotide differences (K) were 21.238. The 71 individuals belonged to 33 haplotypes according to the determined sequences. The average genetic distance (P) among the haplotypes of the species was 0.169. The phylogenetic tree constructed using the neighbor-joining method revealed that these individuals were clustered into three groups, but the individual distribution in those groups was disordered. These data indicated the variation and rich genetic diversity in the three populations of
M. berezovskii. Compared to the wild population in Longxian, those in Liuba and Fengxian had a close kinship. The present results indicated no signs of inbreeding or a decline in genetic diversity in the wild M. berezovskii population.

Key words: Moschus berezovskii; Mitochondrial DNA; Genetic diversity; Control region

INTRODUCTION

Moschus berezovskii (Class: Mammalia; Order: Artiodactyla; Family: Moschidae; Sheng, 1992; Wang et al., 2006), commonly called forest musk deer, musk deer, river deer, donkeys, or roe, are famous for secretions called musk that the males of the species secrete from their musk glands. Musk has been an indispensable ingredient in many traditional Chinese medicines. Data indicate that over 70% musk and musk-associated products in the world come from China (Sheng, 1996; Zou et al., 2005). They are first-class nationally protected animals (Le and Chen, 1998) with widest distribution in China. However, populations of forest musk deer have declined dramatically as a result of poaching for musk pods and habitat destruction that has resulted from deforestation and other human disturbances (Yang et al., 2003; Meng et al., 2006; Sheng and Liu, 2007). Wild forest musk deer are only scattered in the Fengxian, Longxian, and Liuba counties of Shaanxi Province (Jiang, 1997). Their populations dropped from 80,000 to about 4000 (Liu et al., 2006) and they are critically endangered (Wu and Wang, 2006). Therefore, we evaluated the genetic diversity of forest musk deer, a significant investigation for conservation of the forest musk deer resources.

Mitochondrial DNA (mtDNA) is an extra-nuclear genetic material with simple and stable structure that is widely used in genetic diversity research because of its maternal genetic inheritance, more intuitive conservation population catastrophe characteristics, and faster evolution without genetic recombination between generations (Zhou et al., 2001). mtDNA control region is a non-coding region of mitochondrial genome, also called a control area, rich in A and T bases, which belong to the area of high genetic variability and quick evolutionary speed (Sun et al., 2007). Therefore, mtDNA control can be used to analyze the evolution of different populations within the same species. In the present study, three wild forest musk deer populations (from Fengxian, Longxian, and Liuba) from the Qinling mountain of Shaanxi Province were investigated and the genetic diversity of partial sequence of the mtDNA control region among 71 forest musk deer was analyzed using polymerase chain reaction (PCR) to obtain scientific data for studies of the genetic diversity of wild forest musk deer population in Shaanxi and for their conservation.

MATERIAL AND METHODS

Sample collection

Samples were sourced from three wild breeding population groups from Fengxian, Liuba, and Longxian. Thirty-nine individual hair samples or muscle tissue from dead forest musk deer were collected from Fengxian, 17 samples of wild musk deer muscle were from Liuba, and 15 samples were from Longxian. Muscle samples were fixed in ethanol and stored at -20°C in a laboratory refrigerator. Chelex-100 (for extracting DNA from hair) was purchased from Bio-Rad Corporation (Beijing, China) and the PCR primers were synthesized by Shanghai Sangon Biological Engineering Technology Services Limited (Shanghai, China).
mtDNA and genetic diversity in *Moschus berezovskii*

**mtDNA extraction**

Extraction of mtDNA was done as described by Zhou et al. (2003). All the glassware and plastic materials were sterilized at high temperature and pressure. To prevent partial degradation of obsolete hair collected from ground or contamination through external sources during the experiment, blank extraction control experiment was carried out, i.e., parallel control experiments without tested samples were conducted.

Total DNA from the muscle was extracted according to the methods reported by Sambrook and Russell (2001) and Cai et al. (2006) for the extraction of high molecular weight DNA from mammalian cells; the procedure involved treatment with protease K and phenol separation.

**PCR and sequencing of the amplified product**

The PCR primers used were as follows: L: 5’-CAACTAACCTCCTAAGACTTCAAG-3’ and H: 5’-CCAAATGTATGACAGCACAGTTATG-3’. The total volume of PCR mixture was 50 μL, which included 5.0 μL 10X PCR buffer, 1.0 μL dNTPs, 2 μL DNA template, 2 μL primers, 0.4 μL Taq enzyme, and 37.6 μL ultrapure sterilized water. The PCR conditions were as follows: initial denaturation at 95°C for 4 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 45 s, extension at 68°C for 1 min and a final incubation at 68°C for 10 min. The amplified products were stored at 4°C until they were examined by agarose gel electrophoresis on a 1% gel and were subsequently sent to Shanghai Invitrogen Biotechnology Co., Ltd. (Shanghai, China) for bidirectional sequencing.

**Data analysis**

The amplified sequences were compared with a reference musk deer sequence (GenBank accession No. AY835375.1) using the Clustal X1.81 software (Yue et al., 2004). The results were compared by the Seeview software (Galtier et al., 1996). The statistics of population nucleotide diversity (Pi), the number of haplotypes of sequence (h), haplotype diversity (Hd), the average number of nucleotide differences (K), and population distribution were conducted by DNASP 4.10 (Rozas et al., 2003). The base composition, diversity, and genetic distance between the individuals were calculated by the Mega 5.05 software (Tamura et al., 2011). The genetic distances in Kimura 2-parameter model between and within the populations were calculated by the MEGA 5.0 software. The base composition and the substitution number of nucleotide positions were analyzed. Neighbor-joining phylogenetic tree was built with 1000 times Bootstrap replicate analysis. The kinship network diagram to analyze the evolutionary relationship among various haplotypes was established by the Network software (Bandelt et al., 1999).

**RESULTS**

**PCR results and sequence analysis**

In the present study, fragments of mtDNA control region in 71 musk deer were amplified using PCR. As shown in Figure 1, length of the amplified fragment was 635-637 bp. Subsequently, these fragments were sequenced by bi-directional DNA sequencing method and aligned with the reference sequence. The results showed that the average content of the four nucleotides (A, T,
G, and C) were 33.5, 30.6, 14.4, and 21.5%, respectively. The content of A+T (64.1%) was higher than that of G+C (35.9%). This indicated that the partial sequence of mtDNA control region of forest musk deer was rich in nucleotides A and T, revealing the bias in base composition. Thirty-three haplotypes were identified from 71 sequences based on nucleotide variation (Figure 2); the sequences of all the haplotypes were deposited in GenBank (accession Nos. KT582203-KT5822345). There were 178 diversity sites, among which 42 were single-variable sites and 136 were parsimony-informative sites, accounting for 28.03% of the total analyzed sites.

Genetic diversity of mtDNA control region sequences and molecular phylogenetic tree

The genetic diversity index of the mtDNA control region sequences from the 71 individual forest musk deer, calculated by the DNAsp software, is presented in Table 1. The haplotype diversity ($H_d$), nucleotide diversity ($P_i$), and the average number of nucleotide differences ($K$) were 0.931, 0.04688, and 21.238, respectively, suggesting a high genetic diversity in wild forest musk deer in Shaanxi.
According to the control region sequence (Figure 2), genetic distance between 33 haplotypes calculated with the MEGA software was in the 0.006-0.796 range and the average genetic haplotype distance was 0.169, with a standard error of 0.016. The genetic distance between Hap 9 and Hap 12, Hap 15 and Hap 17, Hap 18 and Hap 19, Hap 24 and Hap 30 was lowest, while that between Hap 13 and Hap 25 was highest.

Figure 3 shows the neighbor-joining molecular phylogenetic tree of forest musk deer populations based on the different haplotypes of the Qinling Mountains constructed with the software. The figure and Table 1 indicate that all the haplotypes could be divided into three branches, but the distribution of the haplotypes in each population was disordered. Branch 1 contained 22 haplotypes, among which 14 (Hap 1, 15, 17, 18, 19, 20, 21, 24, 25, 27, 28, 29, 30, and 31) belonged to Fengxian, nine (Hap 2, 3, 18, 22, 23, 24, 29, 32, and 33) belonged to Liuba, and six (Hap 15, 16, 17, 18, 26, and 31) belonged to Longxian. Moreover, Hap 18 was shared by three populations, Hap 15, 17, 18, and 31 were shared by Fengxian and Longxian populations, and Hap 24 was shared by Fengxian and Liuba populations. Branch 2 contained nine haplotypes (Hap 5, 6, and 9 from Fengxian; Hap 4, 9, and 10 from Liuba; Hap 4, 7, 8, 11, and 12 from Longxian), among which Hap 4 was shared by the Liuba and Longxian populations and Hap 9 was shared by the Fengxian and Liuba populations. Branch 3 contained two haplotypes (Hap 13 and 14); Hap 13 was shared by the Longxian and Fengxian populations.

Table 1. Demographic parameters estimated from three Moschus berezovskii populations.

<table>
<thead>
<tr>
<th>Population</th>
<th>Number of samples</th>
<th>Haplotype variability (H)</th>
<th>Number of variability locus</th>
<th>Number of haplotypes (h)</th>
<th>Average number of nucleotide differences (K)</th>
<th>Nucleotide diversity (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fengxian</td>
<td>39</td>
<td>0.910</td>
<td>125</td>
<td>20</td>
<td>16.831</td>
<td>0.03659</td>
</tr>
<tr>
<td>Liuba</td>
<td>17</td>
<td>0.925</td>
<td>93</td>
<td>11</td>
<td>26.765</td>
<td>0.05765</td>
</tr>
<tr>
<td>Longxian</td>
<td>15</td>
<td>0.981</td>
<td>102</td>
<td>12</td>
<td>24.390</td>
<td>0.05291</td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>0.931</td>
<td>178</td>
<td>33</td>
<td>21.238</td>
<td>0.04688</td>
</tr>
</tbody>
</table>

Figure 3. Neighbor-joining phylogenetic tree of mtDNA control region haplotypes of Moschus berezovskii in Qinling Mountains. The haplotype codes refer to those in Figure 2.
Figure 4 shows the network structure of haplotypes established by the median-joining method. As shown in Figure 4, distribution of the 33 haplotypes revealed stellate situation. It did not divide the 33 haplotypes into monophyletic group corresponding to different geographical regions or geographical populations. Hap 6 and Hap 21 were located in the center of the stellate graph, connecting the other haplotypes by 1 to 33 step mutations, in which Hap 4, 5, 7, 8, 9, 10, 11, 12, and 13 were mutated from Hap 6 and the other haplotypes were mutated from Hap 21.

**Figure 4.** Genetic relationship network among haplotypes of mtDNA control region sequence for the three *Moschus berezovskii* populations. Missing haplotypes in the network are represented by dots.

### Genetic structure of the three populations

Genetic diversity indices of mtDNA control region sequences of forest musk deer from three different regions in the Qinling Mountains, calculated by the DNAsp software, are shown in Table 1. The wild musk deer in the Fengxian district had haplotype diversity (Hd) of 0.910, haplotype number (h) of 20, nucleotide diversity (P) of 0.03659, and 12 polymorphic sites. The forest musk deer in the Liuba district had a Hd of 0.926, h of 11, P of 0.05768, and 93 polymorphic sites. The Longxian population had an Hd of 0.971, h of 12, P of 0.05291, and 102 polymorphic sites.
mtDNA and genetic diversity in Moschus berezovskii

Genetic distances between and within the populations, calculated with Kimura 2-parameter model, are shown in Table 2. The genetic distance between the Fengxian and Liuba population was smaller than that between the Longxian and Liuba population. The genetic distance reflects the phylogenetic relationship between populations. It, therefore, could be concluded that Liuba population was distantly related to Longxian population. Table 3 shows the genetic differentiation index ($F_{ST}$) obtained by pairwise comparison of the three populations. The $F_{ST}$ (0.07994) between the Liuba and Longxian population was the highest while that between the Fengxian and Liuba (0.01727) was the smallest. The result of genetic distance among populations was consistent with that of genetic differentiation index. The $F_{ST}$ and gene flow ($N_m$) among the wild populations, calculated by the Arlequin 3.51 software, were 0.0239 and 10.21, respectively. The $N_m$ in the three populations was higher than 4, indicating that the three populations had a closely related gene flow and the impact of genetic drift by geographical isolation on the population structure was less (Millar and Libby, 1991).

### Table 2. Genetic distances within (on the diagonal) and between (under the diagonal) the three Moschus berezovskii populations.

<table>
<thead>
<tr>
<th>Population</th>
<th>Fengxian</th>
<th>Liuba</th>
<th>Longxian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fengxian</td>
<td>0.039</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Liuba</td>
<td>0.056</td>
<td>0.063</td>
<td>-</td>
</tr>
<tr>
<td>Longxian</td>
<td>0.056</td>
<td>0.058</td>
<td>0.058</td>
</tr>
</tbody>
</table>

### Table 3. Genetic differentiation among the different Moschus berezovskii populations.

<table>
<thead>
<tr>
<th>Pop1</th>
<th>Pop2</th>
<th>$H_s$</th>
<th>$K_Y$</th>
<th>$G_{ST}$</th>
<th>$F_{ST}$</th>
<th>$D_{xy}$</th>
<th>$D_{a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fengxian</td>
<td>Liuba</td>
<td>0.90485</td>
<td>23.37104</td>
<td>0.00361</td>
<td>0.01727</td>
<td>0.05159</td>
<td>0.00412</td>
</tr>
<tr>
<td>Fengxian</td>
<td>Longxian</td>
<td>0.91968</td>
<td>21.08718</td>
<td>0.01897</td>
<td>0.03452</td>
<td>0.04655</td>
<td>0.00161</td>
</tr>
<tr>
<td>Liuba</td>
<td>Longxian</td>
<td>0.94734</td>
<td>25.84706</td>
<td>0.01638</td>
<td>0.07994</td>
<td>0.05706</td>
<td>0.00099</td>
</tr>
</tbody>
</table>

### Demographic history

Neither the selective neutrality test nor the mismatch distribution test supported the hypothesis that the captive musk deer had passed through a bottleneck or population expansion. No statistical significance for Fu’s FS or Tajima’s D was observed for all the individuals or for each population (Table 4). Results from the mismatch distribution revealed that the three populations have not suffered from a bottleneck because the shapes of the distributions were all ragged and multimodal, as shown in Figure 5. The model of recent population expansion was also rejected by significant sum of square deviations and r values for all the tested datasets (Table 4).

### Table 4. Estimation of sudden population expansion of Moschus berezovskii.

<table>
<thead>
<tr>
<th></th>
<th>Fengxian</th>
<th>Liuba</th>
<th>Longxian</th>
<th>All populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fu’s FS (P)</td>
<td>0.575 (0.135)</td>
<td>3.015 (0.0083)</td>
<td>0.487 (0.2529)</td>
<td>-0.394 (0.0096)</td>
</tr>
<tr>
<td>Tajima’s D (P)</td>
<td>-1.803 (P &gt; 0.10)</td>
<td>-0.365 (P &gt; 0.10)</td>
<td>-1.135 (P &gt; 0.10)</td>
<td>-1.898 (P &lt; 0.10)</td>
</tr>
<tr>
<td>SSSD (P)</td>
<td>0.009 (0.560)</td>
<td>0.032 (0.299)</td>
<td>0.017 (0.370)</td>
<td>0.007 (0.18)</td>
</tr>
<tr>
<td>Raggedness (P)</td>
<td>0.008 (0.890)</td>
<td>0.045 (0.140)</td>
<td>0.022 (0.720)</td>
<td>0.004 (0.820)</td>
</tr>
</tbody>
</table>
DISCUSSION

Sequence features of mtDNA control region of forest musk deer in the Qinling Mountains

There were 178 variable sites in the mtDNA control region from the three forest musk deer populations in the Qinling Mountains, as revealed by the sequencing analysis. Of these, 42 sites were single-variable sites and 136 were parsimony-informative sites. They accounted for 28.03% of the total analyzed sites. This result proved that the evolutionary rate of mtDNA control region was higher than those of the other areas.

The results also demonstrated that the content of the base G in forest musk deer was lower than those of the other three bases, indicating an obvious bias for the bases. This was because the variation in the third base of codons in the protein coding genes suffered from less natural selection pressure. It proved that the third codon point could clearly reveal the in homogeneity of nucleotide composition in the mitochondrial genome (Meyer, 1993; Feng et al., 2012).

Genetic diversity of forest musk deer population in the Qinling Mountains

Genetic diversity is an important indicator for estimating the long-term survival probability of populations. It is also one of the contents, necessary for protection plans in biology. The loss of genetic diversity can cause the decrease of individual fitness. In general, when the genetic diversity is studied by mtDNA sequencing technology, several important indicators, like $P$, $P'$,
mtDNA and genetic diversity in Moschus berezovskii

$H_D$ (Neigel and Avise, 1993; Li et al., 2010), are used to measure the degree of mtDNA genetic variability for a population. In majority of the mammals, P values more than 0.01 are considered as large variations (Peng and Dai, 2009). The experimental results showed that the average genetic distance of haplotypes among the three populations was 0.043, greater than 0.01, indicating a large genetic variation in forest musk deer population, in the Qinling Mountains. $H_D$ denotes the haplotype diversity, depending on the frequency of random drawing of two different haplotypes. $P_I$ denotes nucleotide diversity, depending on the average nucleotide difference of each site between the sequences. The greater the $H_D$ and $P_I$ values, the higher is the degree of polymorphism and richer is the genetic diversity (Zhang et al., 2010). According to the definition given by Neigel and Avise (1993), the smaller $P_I$ value of two randomly selected mtDNA sequences in a given population indicate lower genetic diversity. Considering the various proportions of mtDNA haplotypes in the populations, the $P_I$ value is more reliable in reflecting the genetic polymorphism of mtDNA in the population than the simple average genetic distance. Lan and Shi (1993) figured that when $P_I$ value was in the range of 0.0015 to 0.0047, the genetic diversity was lower. In the present study, $P_I$ for musk deer mtDNA control region in the three populations was 0.04688. Moreover, it was maximum for the Liuba population and minimum for the Fengxian population, but was still higher than 0.0047. Compared to the nucleotide diversity in the partial sequences of the control region of other Cervidae animals, $P_I$ was larger (0.0265) than that for the original musk deer ($M. moschiferus$; Ma, 2012), whereas it was 0.0106 for sika deer ($Cervus nippon$; Wu et al., 2004), 0.00562 for black muntjac ($Muntiacus muntjak$; Wu and Fang, 2005), and 0.0014-0.0024 for Hainan deer ($Cervus eldi$; Balakrishnan et al., 2003). It can be concluded that there is abundant genetic diversity of forest musk deer in the Qinling Mountains. This is mainly because of the improvement of ecological environment and the construction of nature reserve, and local residents starting farming musk deer instead of hunting. In the phylogenetic tree and the median-joining network, each population had a set of unique haplotypes but all the 33 haplotypes were promiscuously distributed. This was mainly because the gene flow between the three populations became slower. Furthermore, the barrier of gene flow among the populations might also have contributed to each population possessing a set of unique haplotypes. Severe inbreeding could occur in the Fengxian population because it had a haplotype that was shared by more than 10 individuals. The effect of human activities like herb-collection, highways, and tourism could be the probable barriers.

Table 4 demonstrates that the $F_{ST}$ value of 0.01727 between the Fengxian and Longxian population was the lowest, while that of 0.07994 between the Liuba and Longxian population was the highest. It suggests that the genetic differentiation between the Fengxian and Liuba population was small, and the two populations had a high level of gene communication. It might have resulted from the closer distance between the two districts. However, the gene communication among the Longxian and Liuba populations was lower, due to the districts being far apart and the effects of human activities such as tourism, road construction, presence of herbs, and the wild hunt. Otherwise, according to the data published by Shaklee et al. (1982), the classification criteria of fish was raised from three levels of the genus, species, and populations, and the P values of genetic distance were respectively 0.90, 0.30, and 0.05. The average genetic distance among the haplotypes was 0.169, so the genetic differentiation had reached the level of population differentiation.

Conflicts of interest

The authors declare no conflict of interest.
ACKNOWLEDGMENTS

Research supported by the Major Scientific and Technological Project (#2010ZDKG-54), the Research and Development Planned Project about Science and Technology (#2012KW-21), the Youth Talent Training Project of the Academy of Sciences (#2012K-027), and the Science and Technology Special Project of the Academy of Sciences in Shaanxi Province (#2009K-02).

REFERENCES


mtDNA and genetic diversity in *Moschus berezovskii*


