Assessment of genetic diversity in the forest musk deer (*Moschus berezovskii*) using microsatellite and AFLP markers

ZHAO ShaSha¹ ², CHEN Xuan¹ & WAN QiuHong¹*

¹ Key Laboratory of Conservation Biology for Endangered Wildlife of the Ministry of Education and State Conservation Center for Gene Resources of Endangered Wildlife, College of Life Sciences, Zhejiang University, Hangzhou 310058, China; ² Sichuan Provincial Laboratory for Natural Resources Protection and Sustainable Utilization, Chengdu 610015, China

Received March 1, 2011; accepted April 27, 2011

Microsatellite genotyping and amplified fragment length polymorphism (AFLP) techniques are often utilized in studies of conservation genetics of endangered animals. To select a more effective marker system for conserving the endangered forest musk deer, we used microsatellite and AFLP markers to estimate levels of genetic diversity of two populations, the pure mother Jinfengshan (JFS) group and the offspring Baisha (BS) group with introduction of new blood. It was expected that JFS would possess significantly higher genetic variability than BS if it is assumed that no new deer were introduced into BS. The results revealed that the 15 microsatellite markers produced 141 alleles and the 22 AFLP primer combinations yielded 908 polymorphic bands, showing that AFLPs had higher resolving power than microsatellites. Nonetheless, compared with other deer, both marker systems indicated that the forest musk deer contains a relatively high level of genetic variation. The 15 microsatellite loci indicated a similar level of genetic diversity was present in the JFS and BS populations, seemingly suggesting that the new blood obviously elevated the level of genetic diversity in BS. However, 22 suites of AFLP markers yielded a significantly higher level of genetic variation in JFS than that in BS, in contrast to the diversity pattern of microsatellites. In view of the difference in resolving power of these two marker systems, we conclude that gene introgression was limited in the BS population, as revealed by the AFLP markers. Furthermore, from the perspective of whole genome variability, we recommend that the AFLP marker system is more suitable for conservation genetics studies on the forest musk deer.

*Moschus berezovskii*, genetic diversity, microsatellite, AFLP

Citation: Zhao S S, Chen X, Wan Q H. Assessment of genetic diversity in the forest musk deer (*Moschus berezovskii*) using microsatellite and AFLP markers. Chinese Sci Bull, 2011, 56: 2565–2569, doi: 10.1007/s11434-011-4572-2

Musk deer (*Moschus* spp.) are solitary ruminants, distributed throughout the forests and mountainous parts of Asia, from just north of the Arctic Circle southward to the northern edge of Mongolia and Korea. Further south, avoiding the Gobi Desert, musk deer occur in China, Burma, northern India, northern Vietnam and the Himalayan region [1]. Presently, four or more species of musk deer are recognized [2], and all have been included in the Appendices of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) since 1979. Generally, the musk deer population has been declining because of over-hunting for musk and loss of habitat. In China, the Sichuan Province started the earliest farming of forest musk deer (*Moschus berezovskii*) in 1958 and now possesses the largest captive population held in the Sichuan Musk Deer Breeding Institution (containing about 1100 individuals; [3]).

The Sichuan Musk Deer Breeding Institution has Jinfengshan (JFS) and Baisha (BS) captive farms and the Maerkang wild base. The JFS farm was established in 1964 while the BS farm was started from one-third of the JFS population in 1982, when the JFS group had developed into a relatively large population (about 150 deer) (Chen S personal communication). To accelerate the captive breeding of the BS group, new blood of forest musk deer from the Maerkang base were introduced to the BS farm several
times over the past two decades. Therefore, in spite of the founding of BS with JFS individuals 18 years later, introduction of new blood would possibly homogenize intra-BS genetic variation and elevate the level of genetic diversity compared to the original BS founders. To identify genetic effects of introduced deer on genetic diversity of the BS population, we estimated the level of genetic variation of BS without introduction of new blood and predicted the difference in genetic diversity between the mother JFS and pure offspring BS populations.

To predict the level of genetic diversity of BS without new introductions, we employed the equation \(H_t = \frac{1}{2N}H_0\), where \(N\) is the number of generations; and \(H_0\) is the heterozygosity \(t\) generations previously [4]. Since we previously detected microsatellite diversity of the mother JFS population in 2006 [5], we named the expected heterozygosity of JFS in 2006 as \(H_{JFS0}\). Then, we calculated the original JFS heterozygosity \(H_{JFS0}\) in 1984 as \(H_{JFS0} = \frac{1}{2(1/2 \times 100)}\) based on the time span of 1984–2006 and generation time (2 years) for the forest musk deer [6]. Next, based on the allelic data of microsatellites for JFS in 2006, we performed 1000 permutations, as implemented in Genetix software [7], and extracted one-third allelic data to simulate sampling of BS from JFS. The results showed that one-third sampling could result in 3.32% loss of genetic variation. Thus, genetic diversity of the original BS population in 1984 (i.e., \(H_{BS0}\)) would be 96.68% of \(H_{JFS0}\) (about 1.02 \(H_{JFS0}\)) and that of pure BS in 2006 (i.e., \(H_{BS}\)) would be \(\frac{1}{2-1/2 \times 100}\) \(H_{BS0}\) (about 90.84% \(H_{BS0}\)). Finally, we employed Student’s \(t\)-test to compare the difference between the \(H_{JFS0}\) and \(H_{BS}\) values for 15 microsatellite loci and obtained an extremely significant result \((P < 0.005)\). As a result, if there had been no new deer introduced into BS, the current mother JFS would show significantly higher genetic diversity compared to its pure offspring BS.

In this study, one objective was to compare the level of genetic diversity in the mother JFS to that in the mixed offspring BS as measured by microsatellite and amplified fragment length polymorphism (AFLP) marker systems, in order to identify genetic effects of the introduced deer on the BS population. If the current BS showed a similar level of genetic variation compared to JFS \((P > 0.05)\), it was concluded that the introduced deer significantly elevated the level of genetic diversity of JFS-derived BS individuals. A second objective of this study was to screen a better marker system for future studies of forest musk deer by comparing the microsatellite and AFLP results for the JFS and BS populations.

1 Materials and methods

Blood samples and liver tissues of 41 forest musk deer were collected randomly from animals of the Sichuan Musk Deer Breeding Institution. Seventeen of the samples were from the JFS farm and 24 were from the BS farm. Genomic DNA was isolated using a standard protocol [8].

Fifteen pairs of microsatellite primers were developed previously and microsatellite genotyping was conducted on a LI-COR 4200 automated DNA sequencer according to the protocol of Zhao et al. [5]. The number of alleles \((N_a)\) and polymorphic information content (PIC) were computed for each microsatellite in Cervus 2.0 [9] and the level of microsatellite polymorphism was calculated as observed and expected \((H_e\) and \(H_o\)) heterozygosities using GENEPOP 3.4 [10]. Departures from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were also tested in GENEPOP [10]. When allelic frequencies deviated significantly from the HWE expectations, the program MICRO-CHECKER [11] was used to deduce the most probable cause of the HWE departures. The effective number of alleles \((N_e)\) and Shannon’s information index \((Shannon I)\) were estimated in POPGENE 1.32 [12]. Statistically significant differences were determined using \(t\)-tests between the means of the JFS and BS populations.

In this study, AFLP polymorphism was based on combination of the MseI primer \(5’-GATGAGTCTCCTGAGTAGTANNN-3’\) (abbreviated as M-NNN, where NNN are three selective bases)) and fluorescence-labeled EcoRI primer \(5’-GACTGCCTAACATTGNNN-3’\) (abbreviated as E-NNN, where NNN are three selective bases)). The AFLP gel was electrophorated on a LI-COR 4200 automated DNA sequencer and the AFLP bands were scored by Saga™ AFLP® software (LI-COR). We adopted POPGENE 1.32 [12] to calculate AFLP diversity parameters, including percentage of polymorphic loci (PPL), the effective number of alleles \((N_e)\), Nei’s gene diversity \((H_e)\), and Shannon I. In view of multiple loci generated by each primer combination, we performed \(t\)-tests for each primer combination to determine statistically significant differences between the means of \(N_e\), \(H_e\) and Shannon I from the JFS and BS populations.

2 Results

A total of 141 alleles were revealed for the 15 microsatellite loci of the forest musk deer. Genetic polymorphism indices \((N_a, N_e, H_e, H_0\) and Shannon I) of the microsatellite marker system showed some differences between the JFS and BS populations (Table 1). For the JFS population, the \(N_a\) and \(N_e\) per locus ranged from 5 to 13 and from 2.276 to 8.892, respectively, whereas the \(H_0\) and \(H_e\) values ranged from 0.294 to 0.941 and from 0.592 to 0.909, respectively. Shannon I ranged from 1.082 to 2.328. For the BS population, the ranges of \(N_a\), \(N_e\), \(H_0\), \(H_e\) and Shannon I were 6–14, 3.200–9.216, 0.417–0.958, 0.645–0.910 and 1.375–2.306, respectively. Nevertheless, the \(t\)-test failed to determine statistically significant differences between the JFS and BS populations (all \(P > 0.05\); Table 1), suggesting that (1) the
current mother JFS and the mixed offspring BS had similar levels of genetic variation and (2) the introduced deer have significantly elevated the level of genetic diversity in the pure offspring BS. No linkage disequilibrium was detected among loci. The PIC values of different microsatellites across the two populations ranged from 0.664 to 0.882 (Table 1), indicating that these microsatellite markers were highly polymorphic. Four loci (Mb10, Mb32, Mb39, and Mb40) in JFS and one locus (Mb40) in BS showed significant deviation from HWE ($P < 0.001$) (Table 1), suggestive of a deficiency of heterozygotes. Analysis of homozygote size classes carried out in MICROCHECKER revealed that the deficiency of heterozygotes at these loci resulted from the presence of null alleles. There was no LD among loci ($P > 0.05$).

The use of 22 AFLP primer combinations resulted in 908 loci at the species level, ranging from 21 to 67 with an average of 41.27 loci per primer combination (Table 2). Of these 908 bands, the JFS and BS populations had four and 13 population-specific monomorphic bands, respectively, which resulted in four and seven primer combinations showing non-100% PPL for 10 primer combinations ($P > 0.05$; Table 2), consistent with the expected result, suggesting that the introduced Maerkang deer failed to significantly increase intra-BS genetic variation at the whole-genome level.

3 Discussion

The microsatellite and AFLP markers have distinct respective genetic diversity measures. The level of microsatellite-based genetic diversity is mainly assessed by $N_a$, heterozygosity (usually $H_j$), while AFLP-based variation is primarily estimated by PPL, $H_j$ and Shannon I. Compared with the level of microsatellite genetic diversity in other endangered deer species, such as red deer ($N_a = 5.78–7.28$,

### Table 1 Comparison of genetic diversity based on 15 polymorphic microsatellite loci between the Jinfengshan (JFS) and Baisha (BS) populations

| Locus  | $N_a$  | $N_e$  | PIC    | $H_0$  | $H_e$  | Shannon I |
|--------|-------|-------|--------|-------|-------|-----------|
| Mb06   | 5     | 6     | 2.792  | 3.245 | 0.595 | 0.645     | 0.352 | 0.417 | 0.661 | 0.678 | 1.239 | 1.375 |
| Mb10   | 8     | 8     | 4.000  | 3.571 | 0.714 | 0.678   | 0.294 | 0.417 | 0.729 | 0.645 | 1.596 | 1.508 |
| Mb18   | 10    | 10    | 4.737  | 4.461 | 0.759 | 0.741   | 0.529 | 0.750 | 0.720 | 0.728 | 1.734 | 1.599 |
| Mb30   | 5     | 6     | 2.848  | 3.200 | 0.608 | 0.645   | 0.470 | 0.583 | 0.592 | 0.587 | 1.287 | 1.396 |
| Mb32   | 8     | 8     | 4.512  | 5.848 | 0.741 | 0.807   | 0.353 | 0.625 | 0.797 | 0.847 | 1.614 | 1.878 |
| Mb33   | 6     | 11    | 2.276  | 3.785 | 0.515 | 0.850   | 0.647 | 0.792 | 0.577 | 0.883 | 1.082 | 2.143 |
| Mb34   | 10    | 8     | 5.069  | 3.986 | 0.777 | 0.709   | 0.765 | 0.792 | 0.780 | 0.765 | 1.811 | 1.556 |
| Mb37   | 11    | 11    | 8.500  | 9.216 | 0.871 | 0.882   | 0.941 | 0.958 | 0.909 | 0.910 | 2.257 | 2.306 |
| Mb38   | 9     | 9     | 3.461  | 3.699 | 0.690 | 0.701   | 0.529 | 0.667 | 0.649 | 0.715 | 1.643 | 1.621 |
| Mb39   | 8     | 9     | 3.010  | 3.557 | 0.623 | 0.685   | 0.294 | 0.417 | 0.688 | 0.644 | 1.296 | 1.581 |
| Mb40   | 13    | 14    | 8.892  | 7.680 | 0.878 | 0.856   | 0.764 | 0.708 | 0.914 | 0.888 | 2.328 | 2.167 |
| Mb41   | 12    | 12    | 6.215  | 6.330 | 0.821 | 0.824   | 0.647 | 0.750 | 0.865 | 0.860 | 2.047 | 2.049 |
| Mb43   | 9     | 9     | 4.376  | 5.011 | 0.735 | 0.773   | 0.529 | 0.500 | 0.750 | 0.717 | 1.584 | 1.778 |
| MbA1   | 11    | 11    | 7.918  | 6.491 | 0.860 | 0.829   | 0.706 | 0.833 | 0.900 | 0.829 | 2.117 | 2.045 |
| MbA3   | 6     | 7     | 3.635  | 5.031 | 0.675 | 0.772   | 0.882 | 0.875 | 0.747 | 0.818 | 1.386 | 1.726 |
| Average| 8.73  | 9.27  | 4.809  | 5.247 | 0.724 | 0.760   | 0.580 | 0.672 | 0.752 | 0.768 | 1.668 | 1.782 |

*t*-test $P = 0.543$ | $P = 0.553$ | $P = 0.314$ | $P = 0.198$ | $P = 0.688$ | $P = 0.372$

a) $N_a$ and $N_e$ are the number of alleles and the effective number of alleles, respectively. $H_0$ and $H_e$ are observed and expected heterozygosities, respectively. PIC and Shannon I are the polymorphism information content and Shannon’s information index, respectively. The loci showed significant deviation from HWE ($P < 0.001$). If these loci were excluded for $H_e$ comparison, the $P$-value of the *t*-test changed from 0.198 to 0.570.
the Sichuan Musk Deer Breeding Institution (genetic diversity was exhibited in the forest musk deer from the Sichuan Musk Deer Breeding Institution (\(N_a = 8.73–9.27; H_E = 0.752–0.768; \text{Table 1}\)). The AFLP technique is utilized widely in plants, but no AFLP data from deer is available for comparison. Furthermore, heterozygosity for AFLP markers was limited within the range 0.0–0.50 owing to the dominant nature of the markers, making it unsuitable to compare the level of genetic variation between the two marker systems. Nonetheless, the JFS and BS populations showed 18 and 15 primer combinations showing 100% PPL (Table 2), which is unusual for AFLP observations in other species, thus providing evidence for a high level of genetic variation in the forest musk deer.

Microsatellites are codominant markers, present as biallelic markers in an individual and multiallelic markers in a population. Thus, a small number of microsatellite loci could yield abundant polymorphism. For example, this suite of 15 microsatellite markers was able to produce about 140 polymorphic bands. In contrast, AFLPs are a dominant marker system, in which the polymorphism in an individual is the absence (0-state) or presence (1-state) of a particular band. At each AFLP locus, the absence of a band represents a 0/0 homozygote but the presence of a band indicates either a 1/1 homozygote or a 0/1 heterozygote. This bistate characteristic in a population results in one AFLP primer combination pair generating tens of polymorphic bands at most, making it essential to adopt several pairs of primer combinations together to identify as much genetic variation as possible. In this study, the first three primer combinations (M-CAT/E-AAG, M-CAT/E-ACC and M-CAA/E-AAC; Table 2) yielded a total number of 140 polymorphic bands, attaining a similar resolving power to the 15 microsatellites. These three sets of AFLP loci gave a consistent result with the suite of microsatellites, i.e. the JFS and BS populations showed similar levels of genetic diversity (\(P > 0.05\)). However, when we utilized other primer combinations, we found that the non-significant results (\(P > 0.05\)) of the M-CAT, M-CAA, M-CAC primer combinations for comparison of genetic variation between JFS and BS were converted to a significant \(P\)-value (\(P < 0.05\)) across all primer combinations if the M-CAG data were included (Table 2). As a consequence, the results were closely related to the number of markers utilized.

Microsatellites are distributed throughout the genome. Fifteen microsatellite loci were just a small fraction of genome-level microsatellites, less representative of genetic variation at the genome level. The AFLP technique, which provides information on whole-genome variability, produced 908 polymorphic bands in the forest musk deer. Compared with the AFLP loci, the 15 microsatellites yielded 140 polymorphic bands, comprising about 15.42% of the AFLP bands generated by 22 primer combinations. As a result, we argue that at the whole-genome level these

**Table 2** Comparison of genetic diversity based on 22 AFLP primer combinations between the Jinfengshan (JFS) and Baisha (BS) populations

| Unlabeled Labeled | No. of loci | PPL (%) | Mean of \(N_E\) | Mean of \(H_E\) | Mean of Shannon I |
|-------------------|-------------|---------|----------------|----------------|------------------|
|                   | JFS BS      | JFS BS  | JFS BS         | JFS BS         | JFS BS           |
| M-CAT             |             |         |                |                |                  |
| E-AAG             | 41          | 100     | 100            | 1.679          | 1.704            | 0.392            | 0.402            | 0.577            | 0.588            |
| E-ACC             | 63          | 100     | 100            | 1.722          | 1.765            | 0.408            | 0.425            | 0.594            | 0.614            |
| E-AAC             | 36          | 100     | 100            | 1.701          | 1.696            | 0.401            | 0.394            | 0.588            | 0.576            |
| M-CAA             |             |         |                |                |                  |
| E-AAG             | 31          | 100     | 96.77          | 1.826*         | 1.694            | 0.446*           | 0.396            | 0.637*           | 0.578            |
| E-ACA             | 21          | 100     | 100            | 1.633          | 1.540            | 0.378            | 0.334            | 0.563            | 0.507            |
| E-ACC             | 25          | 100     | 100            | 1.756          | 1.715            | 0.420            | 0.405            | 0.607            | 0.592            |
| M-CAC             |             |         |                |                |                  |
| E-AAC             | 31          | 100     | 77.42          | 1.492**        | 1.144            | 0.316**          | 0.117            | 0.491**          | 0.214            |
| E-AAG             | 34          | 100     | 97.06          | 1.666*         | 1.525            | 0.386*           | 0.325            | 0.569*           | 0.496            |
| E-ACA             | 41          | 100     | 97.56          | 1.709          | 1.658            | 0.401            | 0.370            | 0.585            | 0.542            |
| E-ACC             | 43          | 100     | 100            | 1.826          | 1.839            | 0.442            | 0.450            | 0.631            | 0.640            |
| E-ACG             | 32          | 100     | 100            | 1.591          | 1.535            | 0.350            | 0.336            | 0.525            | 0.513            |
| E-ACT             | 39          | 100     | 100            | 1.743          | 1.786            | 0.415            | 0.433            | 0.603            | 0.623            |
| E-AGC             | 40          | 100     | 100            | 1.772**        | 1.474            | 0.422**          | 0.307            | 0.609**          | 0.478            |
| E-AGG             | 32          | 100     | 96.88          | 1.793**        | 1.546            | 0.432**          | 0.332            | 0.621**          | 0.502            |
| M-CAG             |             |         |                |                |                  |
| E-AAC             | 65          | 100     | 98.44          | 1.712**        | 1.501            | 0.402**          | 0.308            | 0.587**          | 0.474            |
| E-AAG             | 63          | 100     | 100            | 1.666**        | 1.421            | 0.384**          | 0.271            | 0.566**          | 0.427            |
| E-ACA             | 54          | 98.20   | 100            | 1.630**        | 1.426            | 0.361**          | 0.271            | 0.535**          | 0.429            |
| E-ACC             | 67          | 98.37   | 100            | 1.741**        | 1.504            | 0.409**          | 0.312            | 0.591**          | 0.481            |
| E-ACG             | 44          | 97.77   | 100            | 1.546**        | 1.384            | 0.333**          | 0.259            | 0.505**          | 0.417            |
| E-ACT             | 30          | 100     | 100            | 1.476*         | 1.349            | 0.306*           | 0.243            | 0.477*           | 0.397            |
| E-AGC             | 37          | 97.30   | 100            | 1.623          | 1.540            | 0.361            | 0.331            | 0.537            | 0.502            |
| E-AGG             | 49          | 97.96   | 100            | 1.786**        | 1.594            | 0.425**          | 0.356            | 0.609**          | 0.535            |

|             | Average     |         |                |                |                  |
|-------------|-------------|---------|----------------|----------------|------------------|
| Unlabeled   | 41.27       | 99.65   | 98.35          | 1.581          | 1.535            | 0.390*           | 0.361*           | 0.570*           | 0.506            |

\(a\) Parameters comprise the number of loci, the percentage of polymorphic loci (PPL), the effective number of alleles (\(N_a\)), Nei’s gene diversity (\(H_E\)) and Shannon’s Information index (Shannon I). In view of multiple loci generated by each primer combination, here we only listed the means of \(N_a\), \(H_E\) and Shannon I for each primer combination. * (\(P < 0.05\)) and ** (\(P < 0.005\)) indicate that the means of the diversity parameters for JFS are significantly higher than those of BS.
microsatellites gave a biased result, as revealed by partial AFLP markers. The 22 sets of primer combinations detected the whole-genome variability and demonstrated significantly higher level of genetic variation in JFS than in BS ($P < 0.05$; Table 2), indicating that genes from introduced Maerkang deer did not widely introgress into the JFS-derived BS individuals.

In the present study, the suite of microsatellite markers and 10 sets of loci generated by 10 pairs of AFLP primer combinations successfully identified genetic effects of the introduced Maerkang deer on increasing genetic variation of the JFS-derived BS individuals. However, the other 12 pairs of AFLP primers achieved expected results in the case of new blood not being introduced (Table 2). Why did these 12 primer combination pairs yield a discordant result with the other 10 pairs? The diversity parameters examined here, except $N_A$ and PPL, were calculated with inclusion of allelic frequency. Thus, these parameters account for both abundance and evenness of the alleles present [17], being useful for understanding allelic structure. For the 12 special AFLP primer combinations, the results indicated that their loci generated significantly more 0 alleles (or 0/0 homozygotes) in BS than in JFS ($P < 0.001$), resulting in a significantly higher level of genetic variation in JFS than in BS. In contrast, the remaining 10 AFLP primer combinations and microsatellites presented similar allelic frequencies between JFS and BS, leading to a similar level of genetic variation between the populations. Hence, the contrasting patterns of genetic diversity between JFS and BS revealed by the different loci can be attributed to the respective allelic structure of the loci used in these two populations.

In conclusion, gene introgression from the introduced deer into the BS population was successfully identified by 15 microsatellite loci and 10 AFLP primer combinations. Nevertheless, from the perspective of whole-genome variability, the AFLP marker system should be more suitable for conservation genetics studies of the forest musk deer.

We thank the Sichuan Forest Musk Deer Breeding Institution for providing all the samples used in this study. This work was supported by a grant from the National Natural Science Foundation of China (30870362) and a special grant from the State Forestry Administration of China and by the Fundamental Research Funds for the Central Universities of China.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.