Article
Chinese Science Bulletin
SPECIAL ISSUE
Adaptive Evolution and Conservation Ecology of Wild Animals
June 2013 Vol.58 No.18: 2191–2197
doi: 10.1007/s11434-012-5581-5

Characterization of major histocompatibility complex *DRA* and *DRB* genes of the forest musk deer (*Moschus berezovskii*)

LI Ling, ZHU Ying, GE YunFa & WAN QiuHong*

The 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

Received May 7, 2012; accepted August 31, 2012; published online December 24, 2012

The forest musk deer (*Moschus berezovskii*) is one of the most endangered species in China. Over the past decades, extensive hunting and poaching have pushed the forest musk deer to the edge of extinction, and conservation biologists are presently pursuing scientific management plans to rescue this species. The major histocompatibility complex (MHC), a cluster of genes responsible for antigen presentation, is a highly polymorphic genomic region in vertebrates that has become a popular functional marker system for studying adaptive variation. In this study, we developed locus-specific genotyping primers for exon 2 fragments of one *DRA* gene and one *DRB* locus of the forest musk deer using a suite of comprehensive methods that included universal primer amplification, genome walking, single-strand conformation polymorphism (SSCP), heteroduplex (HD) profiling, and sequence analysis. Each forest musk deer showed no more than two sequences per locus, confirming the specificity of our primers. Genotyping with these primers allowed us to identify two *DRA* alleles and six *DRB* alleles in a captive breeding population of the Sichuan Musk Deer Breeding Institution. For the *DRA* locus, we found a slightly higher observed heterozygosity (N_O=0.154) than expected (N_E=0.143). In contrast, the *DRB* locus showed a significant heterozygote deficiency (N_O=0.508; N_E=0.761; P<0.05), which was probably due to inbreeding in the captive population. An obvious excess of nonsynonymous substitutions over synonymous was observed at the antigen-binding positions of the *DRA* and *DRB* loci, showing the presence of positive selection in the forest musk deer *DR* genes. Finally, generation of phylogenetic trees for the *DRA* and *DRB* sequences of the forest musk deer and other ruminants revealed that the *DRA* and *DRB* loci identified in this study had homologous relationships with the known ruminant *DRA* and *DRB* genes. Based on this analysis, and to facilitate future studies, we named these novel loci *Mobe-DRA* and *Mobe-DRB3*.

forest musk deer, adaptive evolution, positive selection, phylogenetic analysis

Citation: Li L, Zhu Y, Ge Y F, et al. Characterization of major histocompatibility complex *DRA* and *DRB* genes of the forest musk deer (*Moschus berezovskii*). Chin Sci Bull, 2013, 58: 2191–2197, doi: 10.1007/s11434-012-5581-5

The major histocompatibility complex (MHC) is a highly polymorphic genomic region that is divided into three tightly linked regions, termed class I, class II, and class III in vertebrates [1]. The class II genes encode glycoproteins that present exogenous degraded peptides of 8–25 residues in length to $CD4^+$ T lymphocytes to induce immune responses [2–4]. A functional MHC Class II molecule is a heterodimer composed of two non-covalently linked chains (one α chain and one β chain), and is primarily expressed on the surface of macrophages, B cells and T cells [5,6]. Both the α and β chains of the MHC class II genes contain three parts: two extracellular domains (α1, α2 and β1, β2), a transmembrane domain and a cytoplasmic tail [7]. The α1 and β1 domains together form the peptide-binding region (PBR), where molecular pockets influence the binding of foreign peptides and subsequent immune responses. Nonsynonymous substitutions significantly exceed synonymous substitutions in the PBR, suggesting that positive selection is acting on the sites and accelerating the diversification of the MHC locus [8]. In the MHC *DR* subregion, the most highly polymorphic region is confined to exon 2 of the β1 domain (*DRB*) [9], whereas exon 2 of the α1 domain (*DRA*)

*Corresponding author (email: qiuhongwan@zju.edu.cn)

is monomorphic in most cases, although it may be dimorphic in some mammals [10].

Musk deer (*Moschus*) are typical solitary ruminants that are mostly distributed in and near the mountains and forests of China. Among the five species of musk deer, the forest musk deer (*Moschus berezovskii*) has the largest population and the widest distribution [11,12]. The secretion from the male, known as musk, is a raw material for perfumes and medicines. As the global demand for musk has increased, the number of *Moschus* in China has decreased sharply, and the expansion of snare use and the over-exploitation/habitat destruction by humans has placed these species on the edge of extinction. Today, all *Moschus* species are included in the Appendices of Convention on International Trade in Endangered Species (CITES) and the first grade of the Chinese State Key Protected Wildlife List [13]. Many captive populations of the forest musk deer have been established in China in the past decades, but the effective population size is still small, in part due to a high susceptibility to infectious disease [14,15]. Thus, we clearly need a functional marker system (such as the MHC genes) that can be used to determine the adaptive evolutionary potential of the forest musk deer.

In this study, we used a suite of comprehensive methods, including universal primer amplification, genome walking, single-strand conformation polymorphism (SSCP), heteroduplex (HD) profiling and sequence analysis, to develop locus-specific genotyping techniques for the *DRA* and *DRB* genes of the forest musk deer, in order to supplement the existing adaptive markers for conservation biology studies of this valuable species.

1 Materials and methods

A total of 65 samples were collected from the Sichuan Musk Deer Breeding Institution during the period of 2005–2006. Of them, 11 blood samples were collected during routine examinations and 54 tissues were obtained from individuals that had died of disease. Genomic DNA was isolated using a standard phenol/chloroform-isoamyl alcohol extraction protocol [16].

We designed universal primers for exon 2 of the *DRA* and *DRB* loci based on multiple alignments of homologous sequences, and isolated the corresponding sequences from the forest musk deer. The utilized primer pairs were as follows: DRAup and DRAdn for the *DRA* genes; DRBup1 and DRBdn1 for primary PCR of the *DRB*; and DRBup2 and DRBdn2 for nested PCR of *DRB* (Table 1). The 10 μL PCR reaction mixtures each contained 1 μL of template DNA (30–50 ng/μL), 0.5 U ExTaq DNA polymerase (TaKaRa), 1 μL of 10×ExTaq buffer (TaKaRa), 1 μL of 2.5 mmol/L dNTPs, 0.2 μL of each primer (diluted to 10 μmol/L), and 6.5 μL of double-distilled water. The PCR conditions were as follows: 94°C for 5 min, then 32 cycles of 95°C for 30 s, the optimal annealing temperature of each primer pair (Table 1) for 30 s, 72°C for 45 s, and a final extension at 72°C for 8 min. The PCR products were analyzed by SSCP and HD [1]. Following comparison of the SSCP and HD banding patterns, the PCR products of the representative banding patterns were cloned into the pMD-18T vector (TaKaRa). When mixed PCR products of different representative clones were capable of recovering the banding patterns as the original PCR products, representative clones were selected from at least three separate PCRs and subjected to sequencing.

In order to design the locus-specific primers, the flanking regions of *DRA* exon 2 were obtained using the DNA walking *SpeedUp*™ Premix Kit (Seegene) according to the manufacturer's instructions. Primers SP1, SP2 and SP3 were designed for the 5′ flanking region of exon 2, and primers SPa, SPb and SPc were designed for the 3′ flanking region

Table 1 Primer pairs used for amplification of the *Mobe-DRA* and *DRB* genes and the optimal annealing temperature for each primer pair (T_a)

Locus	Primer	Primer sequence 5′→3′	T_a (°C)	Product length (bp)
DRA	DRAup	AATCATGTGATCATCCAAGCTG	52	244
	DRAdn	CATTGGTGTTTGGGGTGTTGTTG		
	DRAF	CTAGGAGGACTGGAGCAGAG	63	350
	DRAR	GCATCTCACATCCTGGGTTC		
DRB	DRBup1	CCTCTCTCTGCAGCACATTTCCT	60	282
	DRBdn1	CTTGAATTCGCGCTCACCTCGCCGCTG		
	DRBup2	TCTGCAGCACATTTCCTGGAGGGTCT	65	
	DRBdn2	CTCACCTGGGGCGAGCACAGTGAA		
	DRBLF	GTGCGCTTCGACAGCGACTG	68	2162
	DRBLR	TCTTTACAGGATACACAGTCAC		
	DRBF	TAGTTGGGGTGCCGGTGGAG	58	452
	DRBR	TTTAAATTCGCGCTCACCTCGCCGCT		

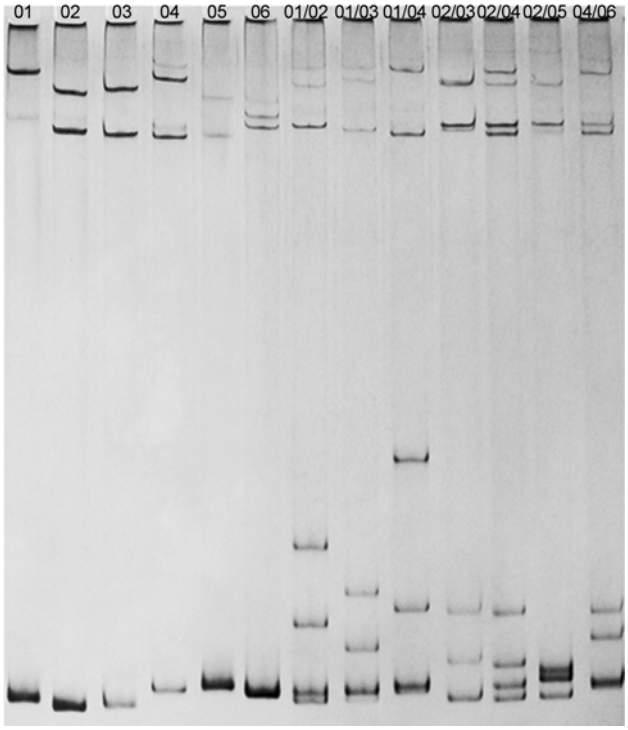

Figure 1 Genotyping profiles of representative homozygotes and heterozygotes of the *DRB* gene. The PCR products were electrophoresed on 12% polyacrylamide gels. Numbers 1–6 represent the *DRB* alleles *Mobe-DRB**01–*Mobe-DRB**06.

(Figure 2(a)). Modified thermal asymmetric interlaced PCR (TAIL-PCR) [17–19] together with long-range PCR were used to obtain the flanking regions of *DRB* exon 2. Primers sp1 to sp5 were designed for the 5′ flanking region (Figure 2(b)), and primer pair DRBLF and DRBLR (Table 1) were designed for the 3′ flanking region. Based on multiple alignments of sequences involving both introns and exons, we designed the intronic locus-specific primers DRAF and DRAR for the *DRA* gene, and DRBF and DRBR for the *DRB* gene (Table 1), and subjected the forest musk deer samples to genotyping based on SSCP and HD analysis.

The *DRA* homozygotes and heterozygotes showed very similar SSCP-HD profiles, so the sequences amplified by DRAF and DRAR were directly sequenced to avoid errors. For *DRB*, in contrast, the PCR products amplified by DRBF and DRBR were first subjected to SSCP and HD analysis, and then representative homozygous samples (i.e., those carrying the simplest HD patterns) were cloned into the pMD-18T vector (TaKaRa). If the original PCR-SSCP and HD patterns of heterozygotes could be reconstituted using the isolated homozygous sequences, we were confident that we had identified the alleles represented by the homozygotes. If not (i.e. if we had not isolated all of the corresponding homozygous sequences), we directly cloned the original PCR products from heterozygotes and determined their alleles (Figure 1) [1]. According to the proposed nomenclature convention [20], we designated the exon 2 alleles as *Mobe-DRA* and *Mobe-DRB,* for *Moschus berezovskii* (*Moschus*), along with identifying numbers.

The expected number of heterozygotes and allele frequencies were calculated using Genepop 4.0 [21]. The allelic sequences were aligned and translated into the corresponding amino acid sequences using MEGA 5.0 [22]. Pairwise nucleotide distances among the *DRA* and *DRB* alleles were estimated using the Kimura-2-parameter method, and the average rate of synonymous (d_S) and nonsynonymous (d_N) substitutions in all of exon 2, the PBR and the non-PBR were calculated using the Nei-Gojobori method [23] with the Jukes-Cantor correction for multiple substitutions. Standard errors were obtained through 1000 bootstrap replicates. Two approaches were used to examine positive selection pressure on *DRB* exon 2. First, a *Z*-test at the 5% level was used to examine the presence of positive selection on each locus by comparing the relative abundance of synonymous and nonsynonymous substitutions in MEGA 5.0. Second, the CODEML program in the PAML 4 package [24,25] was used to test positive selection by comparing the likelihoods of four codon-based models (M1a and M2a, M7 and M8) of sequence evolution. Positively selected codons were identified through the empirical Bayes method [25,26]. Before we reconstructed the phylogenetic trees, we found the best-fitting model for the *DRA* and *DRB* alleles using the modeltest software in MEGA 5.0 [22]. The maximum-likelihood (ML) phylogenetic trees were constructed using the PhyML 3.0 software [27,28], and bootstrap analysis was performed with 1000 replicates to determine the reliability of the branching.

2 Results

2.1 Analysis of the *Mobe-DRA* locus

We obtained flanking sequences spanning 430 bp adjacent to the 5′ end and 115 bp adjacent to the 3′ end of the 246 bp *DRA* exon 2 sequence, and used the full 791 bp *DRA* sequence to design locus-specific primers. Sequencing of the 65 forest musk deer samples identified two *DRA* alleles that differed by only one nucleotide. No deletion, insertion, or stop codon was found in either allele, suggesting that both should represent functional molecules. The allele frequencies of the *Mobe-DRA**01 and *Mobe-DRA**02 alleles were 0.077 and 0.923, respectively. The observed heterozygosity (N_O=0.154) was slightly higher than expected (N_E=0.143). In the 243 bp DRA exon 2, only one site (0.4%) was variable; deduced amino acid position 37 of *Mobe-DRA**01 was aspartic acid, whereas *Mobe-DRA**02 harbored glutamic acid at this position. Comparison of the deduced amino acid sequences of forest musk deer *DRA* exon 2 with those from deer, cattle, sheep and human showed that within the α1 domain, the isoleucine at position 36 and the histidine at position 79 were conserved between the forest musk deer alleles but differed in the other species (Figure 2(a)). Thus,

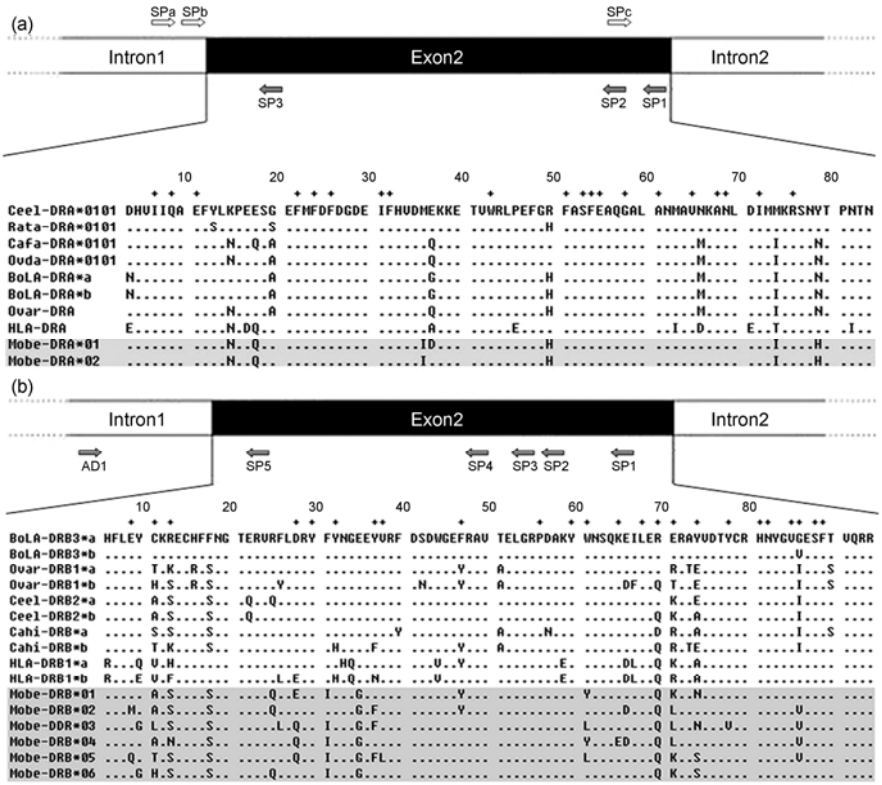

Figure 2 Genome-walking primers and alignment of multiple amino acid sequences for the *DRA* (a) and *DRB* (b) loci. The sequences of the forest musk deer are shaded. Dots indicate identical amino acids and crosses represent putative peptide-binding sites. (a) The utilized *DRA* sequences from other animals were red deer *Ceel-DRA**0101 (FM986347.1), reindeer *Rata-DRA**0101 (FM986348.1), markhor *Cafa-DRA**0101 (FM986346.1), Dall sheep *Ovda-DRA**0101 (FM986341.1), cattle *BoLA-DRA**a (DQ821713.1) and *BoLA-DRA**b (FQ482091.6), sheep *Ovar-DRA* (Z11600.1) and human *HLA-DRA* (NM_019111.4). The genome-walking primers for *DRA* were as follows: SP1, 5′-TGTTTGGGGTGTTGTTGGAGC-3′; SP2, 5′-GATGTCCAGGTTGGC-TTTGTT-3′; SP3, 5′-CATAAACTCGCCTGATTGCT-3′; SPa, 5′-GGAGTGACGGAGAGTAGGGA-3′; SPb, 5′-GGAACCAGAAGAGGAAGATGG-3′; SPc, 5′-CAGGTTGGCTTTGTTCACAGC-3′. (b) The utilized *DRB* sequences from other animals were cattle *BoLA-DRB3**a (U78548.1) and *BoLA-DRB3**b (FQ482091.6), sheep *Ovar-DRB1**a (AM884913.1) and *Ovar-DRB1**b (AM884914.1), red deer *Ceel-DRB2**a (EU573258.1) and *Ceel-DRB2**b (EU573257.1), goat *Cahi-DRB**a (AB008353.1) and *Cahi-DRB**b (AB008358.1), human *HLA-DRB1**a (HM067845.1) and *HLA-DRB1**b (XM_003120517). The genome-walking primers for *DRB* were as follows: sp1, 5′-CTCAMCGACACCGTAGTTGTGTCTGC-3′; SP2, 5′-TGCTCCAGGATC-TCCTTCTGrCTGTT-3′; SP3, 5′-TTGGCGTCCGGCCGCCCCAGCTCsGT-3′; SP4, 5′-GwACTCGCCCCAGTCGCTGTCGAAGC-3′; SP5, 5′-CCGTTGGAG-AAATGACACTCACTCTT-3′; AD1, NTCGASTWTSGWGTT [17].

these amino acids may be species-specific to the forest musk deer. The peptide-binding sites were conservative in the two *DRA* alleles, and the variable site represented a nonsynonymous substitution in the non-PBR.

We then constructed a phylogenetic tree of the *DRA* locus with *DRA* sequences from human, horse, pig and other ruminants, with the human *DRA* sequence as the outgroup. Our results showed that the ruminant *DRA*s shaped a highly supported clade (97.9%) and the *Mobe-DRA**01 and *Mobe-DRA**02 clustered together with a high bootstrap value (98.1%) (Figure 3(a)). The ruminant *DRA* group branched into three parallel clusters corresponding to *Moschidae*, *Cervidae* and *Bovidae* (Figure 3(a)), suggesting that there is a clear sister phylogenetic relationship among these three ruminant families.

2.2 Analysis of the *Mobe-DRB* locus

The 696 bp sequence adjacent to the 5′ end of *DRB* exon 2 (267 bp) and the 2162 bp adjacent to the 3′ end were obtained, and intronic sequences were used to design locus-specific primers. These primers amplified six *DRB* alleles from the forest musk deer samples. No more than two alleles were detected in any sample, indicating that we had amplified a single locus. The allele frequencies were 0.015 (*Mobe-DRB**06), 0.046 (*Mobe-DRB**05), 0.169 (*Mobe-DRB**04), 0.154 (*Mobe-DRB**03), 0.292 (*Mobe-DRB**02) and 0.323 (*Mobe-DRB**01). The observed level of heterozygosity (N_O=0.508) was significantly lower than expected (N_E=0.761) (P<0.05). Within *DRB* exon 2, 33 nucleotide sites (12.36%) were variable and 18 of 89 amino acids (20.22%) were polymorphic (Figure 2(b)). Comparison of the deduced amino acid sequences of forest musk deer, cattle, sheep, deer and human revealed that positions 15 and 79 were conserved cysteine residues in all species; these may play an important role in the correct folding of the *DR* molecule. Potential N-glycosylation signals at positions 19, 20 and 21 were also conserved in all of the compared species.

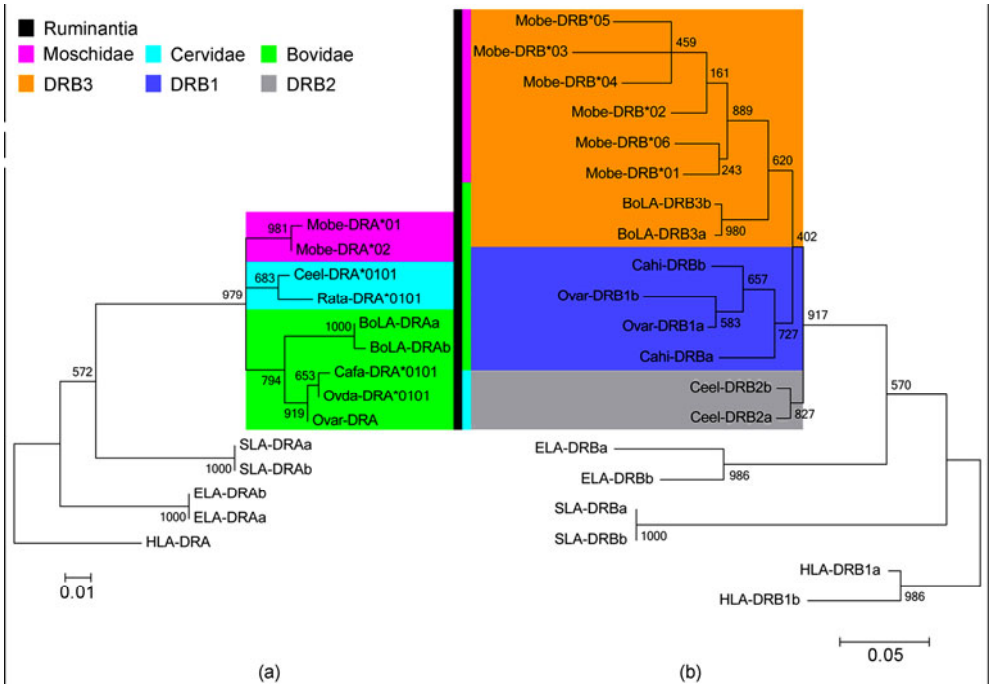

Figure 3 The maximum-likelihood (ML) phylogenetic trees generated for the forest musk deer *DRA* (a) and *DRB* (b) loci. The accession numbers of the other sequences were as follows: sheep *Ovar-DRA* (Z11600.1), Dall sheep *Ovda-DRA**0101 (FM986341.1), markhor *Cafa-DRA**0101 (FM986346.1), cattle *BoLA-DRA*a (DQ821713.1) and *BoLA-DRA*b (FQ482091.6), reindeer *Rata-DRA**0101 (FM986348.1), red deer *Ceel-DRA**0101 (FM986347.1), pig *SLA-DRA*a (NM_001113706.1) and *SLA-DRA*b (EF143987.1), horse *ELA-DRA*a (L47171.1) and *ELA-DRA*b (HQ637392.1), human *HLA-DRA* (NM_019111.4), cattle *BoLA-DRB*3a (U78548.1) and *BoLA-DRB*3b (FQ482091.6), goat *Cahi-DRB*a (AB008353.1) and *Cahi-DRB*b (AB008358.1), sheep *Ovar-DRB*1a (AM884913.1) and *Ovar-DRB*1b (AM884914.1), red deer *Ceel-DRB*2a (EU573258.1) and *Ceel-DRB*2b (EU573257.1), horse *ELA-DRB*a (NM_001142816.1) and *ELA-DRB*b (JN035622.1), pig SLA-DRBa (NM_001113695.1) and SLA-DRBb (NM_001113695.1), human *HLA-DRB*1a (HM067845.1), and *HLA-DRB*1b (XM_003120517). The human MHC *DR* sequences were rooted as the outgroup. Numbers indicate bootstrap values (1000 replicates). The scale bar indicates the number of substitution sites.

Among the non-conserved positions that varied across species, three potential residues were found for positions 9, 28, 61 and 74, and four potential residues were identified for position 11 (Figure 2(b)).

In the PBR, we found that 13 of 24 (54.17%) amino acid positions were variable in the musk deer (Figure 2(b)). The amino acid replacements per site among all pairwise comparisons in the PBR ranged from 0.208 to 0.375 (mean± standard error=0.306±0.066). When considering the entirety of exon 2, the nucleotide variation per site among all pairwise comparisons of the six alleles ranged from 0.042 to 0.087 (0.065±0.011) and the amino acid variation per site ranged from 0.067 to 0.135 (0.109±0.024). The variation on the nucleotide level was lower than that on the amino acid level, and the amino-acid level variation in the PBR was much higher than that in the non-PBR. The rates of non-synonymous substitutions in the PBR were significantly higher than the rates of synonymous substitutions (d_N/d_S= 4.11, P<0.05). The same tendency was also found when we considered the entirety of exon 2 (d_N/d_S=1.35), though the differences between d_N and d_S were not significant (Z= 0.222, P>0.05). However, the value of synonymous substitutions was 2.5 times greater than that of non-synonymous substitutions in the non-PBR, indicating that a strong purifying selection had occurred in the non-PBR.

Assuming that a fraction of the codons had been affected by positive selection, likelihood-ratio tests (LRTs) demonstrated that model M8 fit the data better than M7 (P<0.01), and M2a fit the data better than M1a (P<0.01) (Table 2). The first pair of models identified nine positively selected codons (9, 10, 11, 28, 37, 61, 71, 74 and 78; posterior probability>95%), eight of which (9, 11, 28, 37, 61, 71, 74 and 78) were at putative peptide-binding sites (PBS). The second pair of models identified seven positively selected codons (10, 11, 28, 61, 71, 74 and 78), six of which (11, 28, 61, 71, 74 and 78) were at putative PBS positions. Thus, the PBS sites of the forest musk deer *DRB* appear to have undergone strong positive selection.

Similar to that of *DRA,* the phylogenetic tree generated for the *DRB* sequences produced a large highly supported (91.7%) clade for ruminants (Figure 3). The *DRB* sequences of the forest musk deer formed a single cluster with a high bootstrap value (88.9%) and were 62.0%-supportively grouped with the cow *DRB*3 sequences (Figure 3(b)). The remaining ruminant *DRB* sequences were split into two clusters: the 82.7%-supported *DRB2* of Cervidae and the 72.7%-supported *DRB1* of Bovidae (Figure 3(b)). The *DRB3*-harboring cluster involved *Moschidae* and *Bovidae*, suggesting that

Table 2 Likelihood ratio test (LRT) comparing models M1a–M2a and M7–M8 for evidence of positive selection and model parameter estimates[a)]

Model	Log-likelihood	Estimate parameters	2ΔL
M1a (nearly neutral)	–651.861	p_0=0. 88020 p_1=0.11980 ω_0=0.00000 ω_1=1.00000	50.23 (M1a vs. M2a; P<0.01)
M2a (positive selection)	–626.748	p_0=0.80593 p_1=0.18947 (p_2=0.00460) ω_0=0.00000 ω_1=1.00000 ω_2=25.83029	
M7 (beta)	–652.067	p=0.005000 q=0.046608	50.62 (M7 vs. M8; P<0.01)
M8 (beta and omega)	–626.757	p_0=0.99541 (p_1=0.00459) p=0.00500 q=0.02027 ω=26.49383	

a) Omega (ω) is the selection parameter and p_n is the proportion of sites falling into the ω_n site class. For models M7 and M8, p and q are the shape parameters of the β function.

there were at least two copies of *DRB* (*DRB*1 and *DRB*3; Figure 3(b)) in the ancestral *Bovidae*, and the forest musk deer *DRB* resembled the cattle *DRB3* gene.

3 Discussion

PCR-based locus isolation uses homologous primers to amplify the corresponding loci from a target species, but often amplifies more than one locus due to the high degree of similarity among different loci. For example, when *BoLA-DRB3* primers were used to amplify *DRB* exon 2 from Alpine chamois (*Rupicapra r. rupicapra*) and Chinese muntjac (*Muntiacus reevesi*), more than two sequences per locus were generated [9,29]. To obtain locus-specific sequences, it is necessary to obtain intronic sequences surrounding the target locus. As a relatively fast and reliable procedure, genome-walking can be used to isolate unknown regions flanking a known DNA sequence, without time-consuming screening of a genomic or BAC library [30,31]. In this study, we successfully used genome walking to obtain intronic sequences for the forest musk deer *DR* loci, and used this information to design locus-specific primers. Ultimately, we identified two *DRA* alleles and six *DRB* alleles from 65 captive forest musk deer.

Within the captive population, we identified *Mobe-DRA**02 homozygotes and *Mobe-DRA**01/02 heterozygotes, but not *Mobe-DRA**01 homozygotes. The difference between *Mobe-DRA**01 and *Mobe-DRA**02 was the presence of a D versus an E (aspartic vs. glutamic acids). The D residue was unique to *Mobe-DRA**01, whereas the E residue was shared by *Mobe-DRA**02, *Ceel-DRA**0101 and *Rata-DRA**0101 (Figure 2(a)). We observed a higher-than-expected observed heterozygosity at the *DRA* locus, suggesting that *Mobe-DRA**01 might be a harmful allele that is controlled in the heterozygous state. However, given the relatively small number of samples, a lack of representative individuals might also have caused the apparent deficiency in *Mobe-DRA**01. Regarding the *DRB* loci, the nonsynonymous substitutions significantly exceeded the synonymous substitutions in the PBR, suggesting that the *Mobe-DRB* locus has experienced strong positive selection. As pathogens are thought to be the main factor driving positive selection and influencing polymorphism of the MHC loci [32], we speculate that the excess of nonsynonymous substitutions favored by positive selection was driven by the need for pathogen resistance.

In the phylogenetic trees for both *DRA* and *DRB*, the forest musk deer grouped with the other ruminants with a high level of support (Figure 3). Among the ruminants, the *DRA* tree produced three sister branches for the *Moschidae*, *Cervidae* and *Bovidae* (Figure 3(a)). However, the ruminant *DRB*s shaped three parallel clusters corresponding to *DRB*3, *DRB*1, *DRB*2, respectively, rather than the three families (Figure 3(b)). The *DRB*3 cluster contained the forest musk deer (i.e. *Moschidae*) and cattle (i.e. *Bovidae*), while *DRB1* and *DRB2* were derived from bovids and cervids, respectively (Figure 3(b)). Hence, the *DRB* genes of the ancestral ruminants appear to have undergone birth-death events, and that of the forest musk deer resembles the surviving *DRB3* locus in cattle. To facilitate future MHC studies, we named the forest musk deer *DRA* and *DRB* genes *Mobe-DRA* and *Mobe-DRB3*, respectively.

Previous microsatellite studies of this species showed that it had a high level of allelic diversity, with 6–14 alleles per locus [11]. The lower allelic diversity of *DRA* and *DRB* might reflect differences in the mechanisms driving the polymorphism of these two different marker systems. In combination with the previous microsatellite genotyping method, however, the reliable genotyping techniques for functional *DR* loci established herein should provide a powerful new tool for protecting this endangered species.

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

1 Wan Q H, Zhang P, Ni X W, et al. A novel Hurrah protocol reveals high numbers of monomorphic MHC class II loci and two asymmetric multi-locus haplotypes in the Père David's Deer. PLoS One, 2011, 6: e14518

2 Swarbrick P A, Schwaiger F W, Epplen J T, et al. Cloning and sequencing of expressed *DRB* genes of the red deer (*Cervus elaphus*) MHC. Immunogenetics, 1995, 42: 1–9
3 Klein J. Natural History of the Major Histocompatibility Complex. New York: John Wiley and Sons, 1986
4 Gruszczyńska J, Brokowska K, Charon K M, et al. Single strand conformation polymorphism in exon 2 of the *Ovar-DRB1* gene in two Polish breeds of sheep. Animal Science Papers and Reports, 2005, 23: 207–312
5 Wei H F, Wang H, Hou S, et al. *DRB* genotyping in cynomolgus monkeys from China using polymerase chain reaction-sequence-specific primers. Hum Immunol, 2007, 68: 135–144
6 Takeshima S N, Sarai Y, Saitou N, et al. MHC class II DR classification based on antigen-binding groove natural selection. Biochem Biophys Res Commun, 2009, 385: 137–142
7 Van der Poel J J, Groenen M A M, Dijkhof R J M, et al. The nucleotide sequence of the bovine MHC class II alpha genes: *DRA*, *DQA*, and *DYA*. Immunogenetics, 1990, 31: 29–36
8 Hughes A L, Yeager M. Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet, 1998, 32: 415–435
9 Schaschl H, Goodman S J, Suchentrunk F. Sequence analysis of the MHC class II *DRB* alleles in Alpine chamois (*Rupicapra r. rupicapr*). Dev Comp Immunol, 2004, 28: 265–277
10 Sena L, Schneider M P, Brenig B, et al. Polymorphisms in MHC-*DRA* and -*DRB* alleles of water buffalo (*Bubalus bubalis*) reveal different features from cattle DR alleles. Anim Genet, 2003, 34: 1–10
11 Zhao S S, Chen X, Fang S G, et al. Development and characterization of 15 novel microsatellite markers from forest musk deer (*Moschus berezovskii*). Conserv Genet, 2008, 9: 723–725
12 Yang Q S, Meng X X, Xia L, et al. Conservation status and causes of decline of musk deer (*Moschus* spp.) in China. Biol Cons, 2003, 109: 333–342
13 Zhou Y J, Meng X X, Feng J C, et al. Review of the distribution, status and conservation of musk deer in China. Folia Zool, 2004, 53: 129–140
14 Lv X H, Qiao J Y, Wu X M, et al. A review of mainly affected on musk deer diseases: Purulent, respiratory system and parasitic diseases (in Chinese). J Econo Animal, 2009, 13: 104–107
15 Luo Y, Cheng J G, Wang J M, et al. Histopathological observations of forest musk deer died in pneumonia and suppurative diseases (in Chinese). Prog Veter Med, 2009, 30: 122–123
16 Sambrook J, Fritsch E, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989
17 Zhou J, Huang H, Meng K, et al. Cloning of a new xylanase gene from *Streptomyces* sp. TN119 using a modified thermal asymmetric interlaced-PCR specific for GC-rich genes and biochemical characterization. App Biochem Biotechnol, 2010, 160: 1277–1292
18 Liu Y G, Whittier R F. Thermal asymmetric interlaced PCR: Automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. Genomics, 1995, 25: 674–681
19 Hwang I T, Kim Y J, Kim S H, et al. Annealing control primer system for improving specificity of PCR amplification. Biotechniques, 2003, 35: 1180–1191
20 Klein J, Bontrop R E, Dawkins R L, et al. Nomenclature for the major histocompatibility complexes of different species: A proposal. Immunogenetics, 1990, 31: 217–219
21 Raymond M, Rousset F. GENEPOP version 1.2: Population genetics software for exact tests and ecumenicism. J Hered, 1995, 86: 248–249
22 Tamura K, Peterson D, Peterson N, et al. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol, 2011, 28: 2731–2739
23 Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol, 1986, 3: 418–426
24 Anisimova M, Nielsen R, Yang Z H. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics, 2003, 164: 1229–1236
25 Yang Z H. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol, 2007, 24: 1586–1591
26 Zhang J Z, Nielsen R, Yang Z H. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol, 2005, 22: 2472–2479
27 Guindon S, Dufayard J F, Lefort V, et al. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol, 2010, 59: 307–321
28 Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol, 2003, 52: 696–704
29 Yu J, Liang W, Zhu F, et al. Maintenance of polymorphism of Chinese muntjac (*Muntiacus reevesi*) MHC-*DRB* gene (in Chinese). Acta Theriol Sin, 2010, 30: 52–57
30 Rishi A, Nelson N D, Goyal A. Genome walking of large fragments: An improved method. J Biotech, 2004, 111: 9–15
31 Reddy P S, Mahanty S, Kaul T, et al. A high-throughput genome-walking method and its use for cloning unknown flanking sequences. Anal Biochem, 2008, 381: 248–253
32 Mikko S, Anderson L. Extensive MHC class II *DRB3* diversity in African and European cattle. Immunogenetics, 1995, 42: 408–413