Article

Chinese Science Bulletin

SPECIAL ISSUE
Adaptive Evolution and Conservation Ecology of Wild Animals
Progress of Projects Supported by NSFC

June 2013 Vol.58 No.18: 2119–2127
doi: 10.1007/s11434-013-5713-6

MHC II DRB variation and trans-species polymorphism in the golden snub-nosed monkey (*Rhinopithecus roxellana*)

LUO MaoFang[1,2,3] & PAN HuiJuan[1*]

[1] *College of Nature Conservation, Beijing Forestry University, Beijing 100083, China;*
[2] *Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China;*
[3] *Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China*

Received June 10, 2012; accepted August 30, 2012; published online March 13, 2013

Genetic variation is generally believed to be important in studying endangered species' adaptive potential. Early studies assessed genetic diversity using nearly neutral markers, such as microsatellite loci and mitochondrial DNA (mtDNA), which are very informative for phylogenetic and phylogeographic reconstructions. However, the variation at these loci cannot provide direct information on selective processes involving the interaction of individuals with their environment, or on the capability to resist continuously evolving pathogens and parasites. The importance of genetic diversity at informative adaptive markers, such as major histocompatibility complex (MHC) genes, is increasingly being realized, especially in endangered, isolated species. Small population size and isolation make the golden snub-nosed monkey (*Rhinopithecus roxellana*) particularly susceptible to genetic variation losses through inbreeding and restricted gene flow. In this study, we compared the genetic variation and population structure of microsatellites, mtDNA, and the most relevant adaptive region of the MHC II-DRB genes in the golden snub-nosed monkey. We examined three Chinese *R. roxellana* populations and found the same variation patterns in all gene regions, with the population from Shennongjia population, Hubei Province, showing the lowest polymorphism among three populations. Genetic drift that outweighed balancing selection and the founder effect in these populations may explain the similar genetic variation pattern found in these neutral and adaptive genes.

golden snub-nosed monkey, major histocompatibility complex, microsatellite, mitochondrial D-loop

Citation: Luo M F, Pan H J. MHC II DRB variation and trans-species polymorphism in the golden snub-nosed monkey (*Rhinopithecus roxellana*). Chin Sci Bull, 2013, 58: 2119–2127, doi: 10.1007/s11434-013-5713-6

Genetic variation plays an important role in buffering populations against widespread pandemics [1]. Understanding how different levels of genetic variation influence the survival of threatened species is of primal interest to evolutionary and conservation biologists. Dramatic reductions in available habitat, together with increasing habitat fragmentation and isolation threatened many species, leading to population sizes descending and reduction of genetic diversity [2–5]. Loss of genetic diversity is likely to happen in small populations, and this may augment extinction risk due to decreased reproductive fitness and adaptive flexibility, and increased disease susceptibility [6].

Genetic diversity is often measured by neutral markers, such as mitochondrial DNA and microsatellites [5], which are instructive for phylogenetic and phylogeographic reconstructions, such as examining dispersal routes of individuals and classifying individuals by relatedness and paternity analyses [7,8]. However, neutral variation provides little direct information on adaptive evolution within and between populations [9,10]. Meanwhile, the significance of genetic variation at informative adaptive genes is increasingly being realized [11,12], as the aforementioned issues are relevant to evolutionary ecology and conservation [13,14]. Major histocompatibility complex (MHC) genes are famous adaptive significance examples, and are of special relevance to conservation owing to their role in patho-

*Corresponding author (email: phjjanine@yahoo.com.cn)

gen resistance [15,16]. The MHC genes are highly polymorphic, with an important role in the regulation of the immune system and in the recognition and discrimination of self from non-self antigens. MHC variability is believed to measure the ability of individuals to react to continuously evolving pathogens and parasites and thus, reflects evolutionarily-relevant and adaptive processes within and between populations [5,17].

Parasite-driven balancing selection, which encompasses heterozygote advantage hypothesis, frequency-dependent selection hypothesis and fluctuating selection hypothesis, is widely accepted as the main mechanism in maintaining the unusually high level of polymorphism in the MHC genes [18–20]. Other than balancing selection, reproductive mechanisms might be alternative or complementary mechanisms for MHC variation maintenance, such as disassortative mating and maternal-fetal interactions (reviewed in [5]). Two major groups of MHC genes have been widely studied. The MHC class I genes, which are expressed on all nucleated somatic cell surfaces, are crucial in the immune defense against intracellular pathogens by binding peptides mainly derived from viral proteins and cancer-infected cells. MHC class II genes are preponderantly involved in monitoring the extracellular environment and primarily expressed on antigen presenting cells of the immune system, such as B cells and macrophages [21]. In mammals, most researches have focused on DRB exon 2 and other class II genes because they code for parts of the antigen binding sites (ABS) that are of significant function [5,22,23]. The class II genes are closely linked, and variants at these genes are generally in strong linkage equilibrium [24]. Thus, the observed MHC II loci pattern represents a proper index of MHC variation [25].

We studied DRB and neutral genetic variation in the golden snub-nosed monkey (*Rhinopithecus roxellana*), an endangered primate endemic to China, which inhabits three isolated areas: Sichuan and Gansu provinces (SG); the Qinling Mountains, Shaanxi Province (QL); and the Shennongjia Forestry District, Hubei Province (SNJ) (Figure 1). Current census data suggest that fewer than 20000 individuals remain, and previous studies have found that SNJ population has very low genetic variation and is genetically- distinct [26,27]. The small population size and isolated status that typically characterize the SNJ population make it particularly susceptible to genetic variation losses through inbreeding and restricted gene flow. In this study, we described genetic variation at the most studied adaptive region of the DRB genes in the golden snub-nosed monkey, and studied DRB evolution by testing for signatures of balancing selection, recombination, and trans-species polymorphism (TSP). We also compared the genetic variation and population structure at neutral genes and DRB. This study furthers our understanding of the evolutionary significance and conservation implications of MHC in free-ranging monkeys.

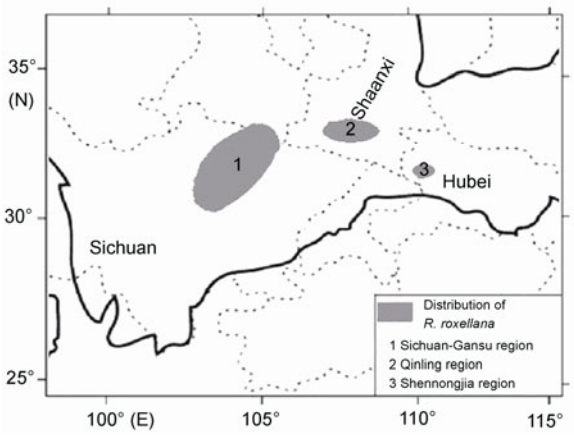

Figure 1 Distribution of isolated *R. roxellana* populations.

1 Materials and methods

1.1 Sample collection and DNA extraction

We collected 64 *R. roxellana* samples (SG population = 25, QL population = 22, and SNJ population = 17) (Figure 1), in compliance with the relevant institutions and laws of China. Muscle and skin samples were also gathered from carcasses provided by local museums and nature reserves. Skin samples were stored dry. Muscle samples were stored in 95% ethanol. Blood samples were collected while trapping individuals for physical examination and were stored at −80°C. Benches and plastic ware were cleaned with 10% bleach and sterile water and then exposed to UV light for 30 min prior to handling to prevent contamination during DNA extraction. The surface of muscle, skin, and hair samples were also exposed to UV light for 30 min. We used eight extraction controls, and none produced positive amplification during subsequent polymerase chain reaction (PCR) analysis.

1.2 MHC amplification, cloning and sequencing

Primers 5′-TTCTCAGGAGGCCGCCCGTGTGA-3′ and 5′-ACCTCGCCGCTGCACTGTGAAGCTC-3′ were designed to amplify 270 bp of the MHC II DRB genes. PCR was performed in 50 μL reaction mix comprising 2.5 mmol/L $MgCl_2$, 10 mmol/L Tris-HCl (pH 8.4), 50 mmol/L KCl, 0.2 mmol/L each dNTP, 0.4 μmol/L each primer, 1.0 unit Hotstart-Taq DNA polymerase (Takara Bio, Otsu, Japan), and 10–100 ng DNA template. The amplification began with incubation at 94°C for 5 min, followed by 35 cycles of 30 s at 94°C, 30 s at 54°C and 30 s at 72°C, offending with anextension at 72°C for 10 min. Wizard PCR Preps DNA Purification Kit (Promega, Madison, WI, USA) was used to purify PCR products according to the manufacturer's protocol. Purified PCR products were cloned into a pMD-18T vector (Takara Bio). Positive transformants containing an

insert of the predicted length were identified by PCR screening and agarose electrophoresis. Eight to 15 clones were sequenced on an ABI 377 or ABI-PRISMTM 3100 Genetic Analyzer (Applied Biosystems Inc.) with the Prism BigDyeTM Terminator Ready Reaction Kit (Applied Biosystems Inc.).

1.3 MHC diversity analyses

All MHC sequences were aligned using ClustalX version 1.83 [28] and translated into the corresponding amino acid sequences using the program Mega 4 [29]. The histocompatibility nature of the sequences was verified through a homology analysis using BlastN at NCBI (http://www.ncbi.nlm.nih.gov). As some sequences obtained could have been artifacts of polymerase error during amplification [30], we considered a new sequence variant to be an allele when it was identified in either two separate PCRs from the same individual, or from PCRs from at least two different individuals [31]. We used Mega 4 to detect the number of variable and parsimony-informative sites (i.e. sites with at least two different nucleotides or amino acids) to compute the mean number of nucleotide differences. This program was also used to derive the overall mean genetic distances of nucleotide sequences based on Kimura's two-parameter evolutionary distances, as well as Poisson-corrected amino acid distances. Standard errors of the estimates were obtained through 1000 bootstrap replicates.

The DRB sequences obtained represented alleles from at least two loci. In accordance with Miller et al. [16], we referred to all sequences as alleles even though they may have come from different loci. Not all alleles could be assigned to each locus in an individual on the basis of their sequence, making it impossible to calculate heterozygosity, allele frequencies, and F_{ST} at DRB genes. Mean number of alleles per individual, total number of alleles per population, and average percent difference were used as measures of within-population genetic variation [16]. The locations of the putative ABS (i.e. residues whose side chains form the antigen-binding groove) and non-ABS were inferred from the human MHC II molecule structure by Reche and Reinherz [32].

1.4 Phylogenetic analysis and recombination analysis

Phylogenetic relationship among MHC alleles was constructed under maximum likelihood parameters using PhyML version 3.0 [33]. Prior to phylogenetic analysis of DRB sequences, the best-fitting models of sequence evolution were chosen on the basis of the Akaike information criterion using Modeltest version 3.7 [34]. The analysis revealed that the model K81uf+I+G was most appropriate for the DRB data, with a gamma shape parameter α = 0.4484. In addition, intra-specific phylogenic structures were inferred using Neighbor-Net method in SplitsTree4 V 4.12.6 [35]. To estimate the rate of population recombination ρ (ρ=4$N_e r$), the composite-likelihood method [36] in LDhat [37] was used. The ρ was calculated by crossing over effective population size (N_e) and rate per generation (r), and was estimated without prior information [37]. Even for sequences evolving under balancing selection in the presence of recombination events, LDhat still works efficiently [38]. Additionally, GARD tests incorporated in the HyPhy package on the Datamonkey website was also used to detect recombination signals [39].

1.5 Detecting balancing selection

Two methods were used to detect historical selection. First, sliding window calculation of Tajima's D was calculated in DNAsp [40] using a sliding window size of 6 bp, and a step size of 2 bp. Second, the Codeml subroutine in the PAML 4 program suite was used [41]. This procedure, which is believed to be more sensitive than other methods for detecting selection [42], uses maximum likelihood estimation to examine heterogeneity in ω (ω= d_N/d_S) [43] among codons within a sequence (ω= d_N/d_S >1 indicating positive selection). ω were estimated following the protocol of Yang et al. [44]. Six different models (M0, M1a, M2a, M3, M7, and M8) that integrated different selection intensities among sites (and deduced from the data) were tested in this study [44,45]. Likelihood-ratio tests comparing nested models (M0 vs. M3, M1a vs. M2a, and M7 vs. M8), in which the alternative models (M2a, M3, and M8) suggest the presence of sites with ω>1.

2 Results

A total of 679 clones derived from 64 individuals, representing three populations of the golden snub-nosed monkey, were examined. We identified 37 unique DRB alleles (Table 1) and these alleles were named *Rhro*-DRB*01–37 (GenBank accession numbers JQ863322–JQ863358) according to the nomenclature of Klein et al. [46]. Two to four alleles were found in each sample, indicating multi-locus amplification. Three among the 37 alleles identified in this study were commonly shared among the three populations. However, differences in frequency for these shared alleles were observed between populations. The nucleotide alignment of DRB sequences revealed a total of 80 (29.63%) variable sites. No indels causing shifts of the reading frame and/or stop codons were detected. The putative amino acid translation of this fragment corresponded to 89 amino acids. Of these amino acid sites, 35 (39.33%) were parsimony-informative sites; 41 (46.07%) were variable, and of those variable sites, 13 were located in putative important antigen-binding positions. There was no indication that either locus was a pseudogene. Measures of MHC diversity in three populations are summarized in Table 2. Though the

Table 1 Alignment of the deduced amino acid sequences of MHC *Rhro*-DRB exon 2 sequences and alleles' distribution in populations[a)]

Alleles	1 1 1 1 1 1 1 1 1 1 12222222222233333333334444444444555555555566666666667777777777888888888812345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789 * * * * * * ** * * * * * * * * * **	SG	QL	SNJ	Number of individuals
DRB*01	RILRVPTARFLEQFKSECHFFNGTERVRYLQRYFYNQEEYVRFDSDVGEFRAVTELGRPVAENFNSQKDFLEQRRAQVDNYCRHNYGVV	10	0	8	18
DRB*02	.....V.F.....F.E.RVH.....A.....E.....D..YW.G...L..D...S.....	8	9	2	19
DRB*03	.....V.Y.....L.E.H.....FL.....Y.....W..RR.Y.....A..T.....	6	6	7	19
DRB*04	Q.....A.....F.D..IH.R.....D..YW..R..I..RA..A..TV.....R.D	4	10	0	14
DRB*05	.....YST.....F.D.....Y.....D...W.....I..D...A..TF.....R.G	1	12	0	13
DRB*06	.....V.F.....F.E.RVH.....A.....E.....D..YW.G...L..D...A.....G	7	5	0	12
DRB*07	.....V.H.....F.D..IS....N.....Y.....D...W.....V...E..A..T.....R.G	4	1	8	13
DRB*08	.....W.P.R.....F.D.....Y.E.....RS..YW.....I..RA..A..T.....	5	3	0	8
DRB*09	G.....Q.G.A.....Q..E..I.....Y.....D...W.....L..D...S..T.....	4	3	0	7
DRB*10	G.....Q.G.A.....Q..E..I.....LL.....W.....I..D...R..-I.....	5	0	0	5
DRB*11	G.....Q.G.A.....Q..E.HI.....FL.....W.....L..D...S..T.....	1	0	7	8
DRB*12	.....YST.....F.D.....D..YW.....L..RK..E..TF.....R.G	5	0	0	5
DRB*13	G.....Q.G.A.....Q..E.HI.....F.....W.....L..D...S..T.....	1	0	4	5
DRB*14	G.V.....A.C.....F.E.RVH.R...A.....Y.....D.KYW.....L..R...A..TF.....	0	4	0	4
DRB*15	.....W.Y.P.....L.E.....N.....Y.....D...W..RR.L.....E..TV.....RI.	5	0	0	5
DRB*16	.....V.H.....F.D.....Y.E.....RS..YW.....I..RA..A..T.....	3	1	0	4
DRB*17	.....W.P.R.....F.D.....F.....Y.....L..R.....AL.A.....A	3	1	0	4
DRB*18	Q.....YTT.....F.D.....RS..YW.....D...S..T.....	0	1	0	1
DRB*19	G.....Q.G.A.Y.....Q..E.HI.....FA.....W.....L..D...S..T.....	0	0	1	1
DRB*20	.....YST.....F.D.....Y.E.....D..YW..R..I..DK..A..TV.....G	0	1	0	1
DRB*21	Q.....A.....F.D.....RS..YW.....L..RKQ.A..T.....	0	0	1	1
DRB*22	Q.....A.....F.D..IH.R.....D..YW..R..I..RA..A..TV.....R.D	0	2	0	2
DRB*23	..VP.....K.D.H.....H.DI.....DL.....Y.....D..YW.....E..K-.....	1	1	0	2
DRB*24	.....A.....C...	1	0	0	1
DRB*25	.....Q.G.A.....Q..E.HI.....F.....W.....L..D...S..T.....	0	0	2	2
DRB*26	.....YST.....F.D.....E.....RS..YG.....I..RA..A..T.....	0	2	0	2
DRB*27	.....V.H.....F.D..IS....N.....Y.....D...W.....V...E..A..T.....R.G	0	0	1	1
DRB*28	G.....V.H.....F.D..IS....N.....Y.....D...W.....V...E..A..T.....R.G	0	0	1	1
DRB*29	Q.....A.....	1	0	0	1
DRB*30	.....V.Y.....L.E.H.....FL.....Y.....W..RR.Y.....A..T.....R.G	0	0	1	1
DRB*31	Q.....YTT.....F.D.....RS..YW.G.....D...S..T.....	1	1	0	2
DRB*32	.....S.W.Y.P.....L.E.....N.....Y.....D...W..RR.L.....E..TV.....RI.	1	0	0	1
DRB*33	.....A.Y.....L.E.H.....FL.....RS..YW.....A..T.....	1	1	0	2
DRB*34	Q.....A.....L.E.H.....FL.....Y.....W..RR.Y.....A..T.....	1	1	0	2
DRB*35	.....V.H.....F.D..IS....N.....Y.....D...W.....L..D...S..T.....	0	0	2	2
DRB*36	.....P..W.P.R.....F.D.....Y.E.....RS..YW.....I..RA..A..T.....	1	0	0	1
DRB*37	..VP.....K.D.H.....H.DI.....DL.....Y.....D..YW.....I..D...A..T.....	0	1	0	1

a) Identical amino acids are shown by points, deletions by asterisks and * represent ABS site.

heterozygosity at DRB loci was not able to be calculated, mean number of DRB sequences per individual could be a good estimator of heterozygosity, since homozygous individuals at two loci will have two alleles while heterozygous individuals have four [16]. Similar to neutral variation, MHC allelic richness was lower in the SNJ population (Tables 2 and 3).

The alleles phylogenetic relationships are shown in Figures 2 and 3. The alleles relationship was not consistent with the population geographical distribution. Also, phylogenetic analyses showed alleles did not cluster into different species, but were intermixed with each other, which is known as TSP (Figure 2). The pseudogenes that download from GenBank clustered together, and none of the alleles identified in this study clustered with those pseudogenes. Population recombination analysis in LDhat revealed that the DRB locus had a high recombination rate (ρ = 16). And the GARD test showed significant evidence for a recombination breakpoint within DRB allele (P<0.01).

Clear signals of historical selection for amino acid replacements in the codons involved in antigen binding were detected. First, Tajima's D analysis across the exon 2 sequence showed three regions with a significantly positive (P<0.05) D value (Figure 4), indicating balancing selection has been acting on these regions. The regions (between 37–52, 93–98 and 215–224) encompass putative ABS sites. Second, through comparisons of codon evolution models in PAML 4, similar results were obtained (Table 4). On the basis of the LRT tests, models integrated positive selection (M2a, M8, and M3) fitted our data significantly better than other models that did not (Table 5).

3 Discussion and conclusions

3.1 Genetic variation of DRB in golden snub-nosed monkey

In the past two decades, the MHC class II genes of some primate species were investigated extensively, especially those in the rhesus macaque (*Macaca mulatta*), which is widely used in biomedical research to study organ transplantation or human pathogens because of its similarity to humans [47]. More recently, extensive research has been conducted in non-model primates to assess the adaptive genetic variation in populations. For example at DRB loci, 0.24 alleles per individual were found in the grey mouse lemur (*Microcebus murinus*) [48], 2.06 alleles per individual were identified in Cynomolgus macaques (*Macaca fascicularis*) [49], and 0.34 alleles per individual were isolated in fat-tailed dwarf lemurs (*Cheirogaleus medius*) [18]. We found that the MHC genetic diversity in the golden snub-nosed monkey is relatively high (0.58 alleles per individual) compared with these previous studies. Two to four alleles were found in individuals, indicating at least two DRB loci were sequenced in this study. Multiple amplifications are common in DRB research because gene duplication is a major reason for high polymorphism in MHC genes, and it is difficult to sequence a single locus from DRB without genome data owing to the similarity among loci [19]. Some duplicated genes become pseudogenes or have altered

Table 2 Summary of DRB variation in *R. roxellana* populations

	SG	QL	SNJ
N	25	22	17
Number of alleles	24	20	13
Number of specific alleles	7	6	7
Mean number of alleles per individual	0.92	0.91	0.76
Gene diversity	0.95±0.01	0.91±0.02	0.89±0.02
Mean number of pairwise differences	26.97±11.93	23.53±10.48	19.55±8.81
Nucleotide diversity (average over loci)	0.10±0.05	0.09±0.04	0.07±0.04

Table 3 Summary of neutral variation in *R. roxellana* populations[a)]

Pop	Control region						Microsatellites (unpublished data)					
	n	*N*	*h*±SD	π±SD	Tajima's D (*P*-value)	F_s (*P*-value)	*N*	No. of loci	A_R	H_E±SD	H_O±SD	F_{IS}
SG	41	15	0.90±0.02	0.02±0.01	1.45 (0.95)	7.67(0.97)	25	16	4.635	0.736±0.022	0.714±0.022	0.017
QL	23	10	0.84±0.05	0.01±0.007	1.02 (0.87)	3.29 (0.90)	22	16	4.854	0.713±0.036	0.653±0.024	0.071
SNJ	22	5	0.56±0.10	0.002±0.001	–0.08 (0.51)	1.04 (0.76)	17	16	3.473	0.611±0.038	0.591±0.030	0.030
All	86	30	0.94±0.01	0.020±0.010	0.42 (0.72)	2.64 (0.86)	64	16	4.321	0.755±0.022	0.660±0.014	0.041

a) *n*, Number of individuals; *N*, number of haplotypes; *h*, gene diversity; π, nucleotide diversity; A_R, allelic richness; H_E, expected heterozygosity; H_O, observed heterozygosity.

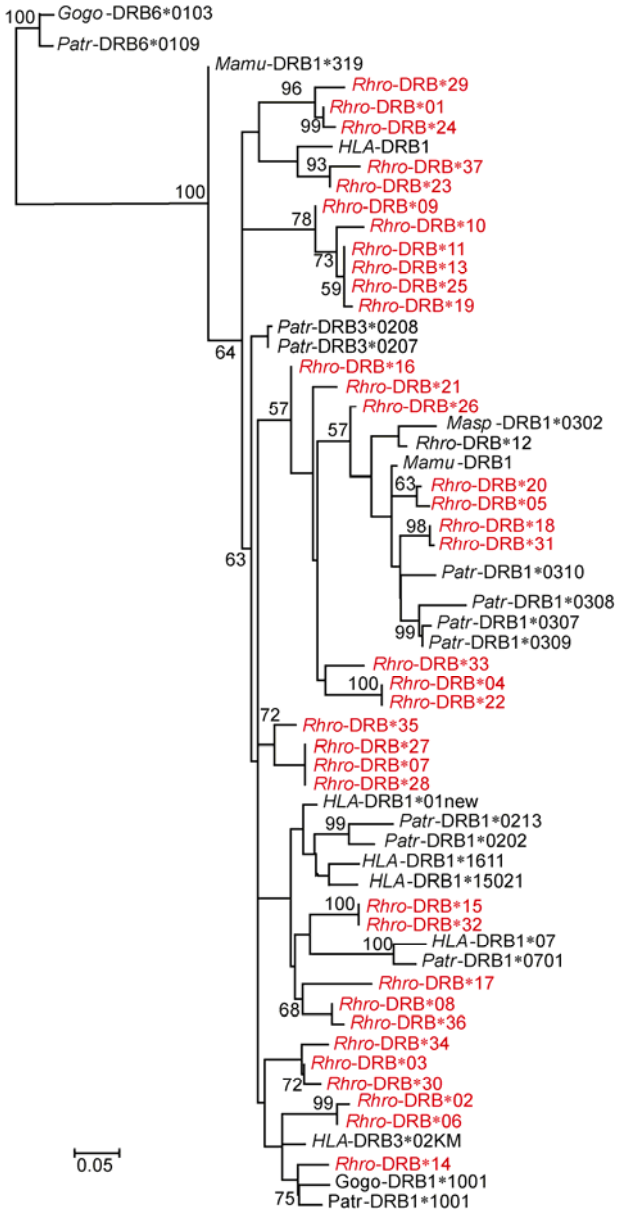

Figure 2 Maximum likelihood of *Rhro*-DRB alleles and a representative set of other primates alleles. Bootstrap values above 50% are shown. Red lines indicate *Rhro*-DRB allelehaplotypes from this study. Other allelic sequences were downloaded from GenBank, which are: *Homo sapiens* (FJ442950, DQ837166, AY271987, AJ311892, AJ293861, AJ238155), *Pan troglodytes* (M96121-M96123, M94944, M96077, M96089, M96084-M96087, M94950), *Gorilla gorilla* (AF031271, AF031275), *Macaca mulatta* (AF17531, AF175315), *Mandrillus sphinx* (DQ103732).

functions because of mutations, insertions or deletions [50]. We found two alleles (H10 and H23) that had lost three bases at 239–241 bp and 242–244 bp, resulting in a loss of an amino acid at site 80 and site 81, respectively (Table 1). They may come from another locus; as the deletions were in-frame and these two genes did not cluster with the pseudogenes during the phylogenetic construction (Figure 2), it appears they are still functional.

Within populations, the SNJ population had the lowest polymorphism at the MHC loci, as was observed for the microsatellite and mitochondrial gene analysis (Tables 2 and 3). Both the number of alleles and the mean number of pair-wise differences in the SNJ population were lower than other populations and the result was consistent with previous study [26,51]. Though DRB genes usually display high diversity, in cases of small and isolated or bottlenecked populations, a lower variation is expected [52]. Similar results were found in other species such as northern elephant seals (*Mirounga angustirostris*) [53], great crested newt (*Triturus cristatus*) [54], and the black-footed rock-wallaby (*Petrogale lateralis lateralis*) [55]. Genetic drift is the reason for the reduced MHC variation in these populations because compared with balancing selection, which usually has a great influence on MHC genes, genetic drift becomes relatively stronger in small, isolated populations, leading to reduced variation at the MHC loci [56]. Beyond genetic drift, a foundation event may have also contributed to the low genetic diversity present in the SNJ population. Among the three populations, the SG population is predicted to be the ancestral population, with the SNJ population arising from the QL population [57]. Microsatellites (unpublished data) and mitochondrial data [57] suggested little gene flow between SNJ and the other two populations. This may indicate that the genetic variation in the SNJ population was reduced compared with the other populations either at its foundation or soon after [52].

3.2 Balancing selection and recombination

MHC genes have been recognized of important role in evolutionary genetics because they are believed to be a good example of the effects of balancing selection [19,58]. Evidence of balancing selection acting on DRB genes was revealed in this study. First, three regions which encompassing putative ABS sites with a significantly positive (P<0.05) D value were identified in Tajima's D test across the exon 2 sequences (Figure 4), suggesting balancing selection exists on these regions [16,59]. Additionally, the sharing of MHC alleles among populations also indicates that MHC alleles may have been conserved by selection [54]. Second, random sites model analysis in PAML showed the existence of base selection in the maximum likelihood method. Analysis suggested that the models including selection (M2a, M3 and M8) matched DRB alleles better than those without selection (Tables 4 and 5). Under the M2a and M8 models, some DRB sites were exposed to significant selection. Further evidence for balancing selection was provided by trans-species evolution of the DRB alleles. TSP is the occurrence of alleles that are more similar in related species than alleles within each species, except cases in which the sameness happened by convergent evolution [60], and is commonly observed in MHC genes. Under balancing selection, TSP is generated by the passage of alleles from ancestral to descendant species. Under selection, some MHC alleles or

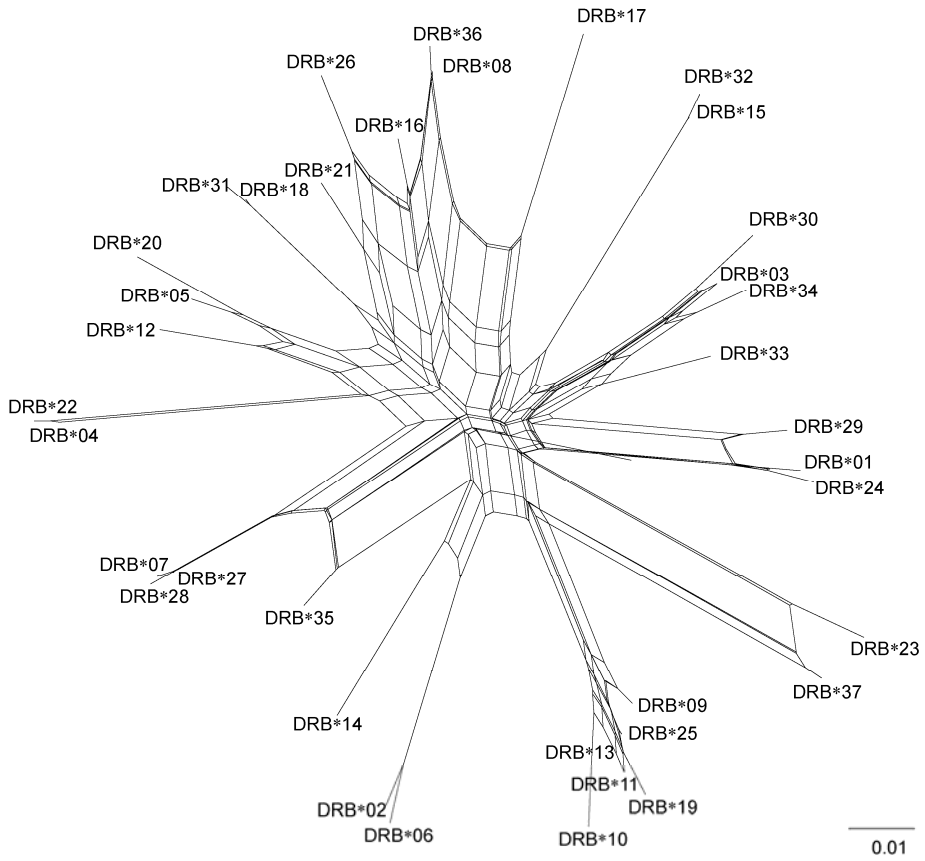

Figure 3 Network of DRB alleles using Neighbor-Net method in SplitsTree.

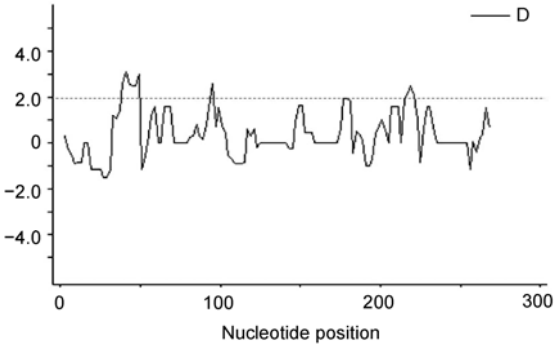

Figure 4 Sliding window calculation of Tajima's D for exon 2 of DRB (window size 6 bp, step 2 bp). The threshold for $P < 0.05$ is shown by the dotted line.

allelic lineages are found in other species, which indicates they existed prior to speciation and were inherited from ancestral species [60]. For example, two DRB exon two sequences are shared by cynomolgus (crab-eating) macaques (*M. fascicularis*) and rhesus macaques (*M. mulatta*) [61]. Mass of sharing of both MHC alleles and allele lineages was also perceived among 28 species of cetacean [62].

While balancing selection maintains high MHC genetic variation, recombination is an important mechanism for generating MHC polymorphism [63]. In human MHC, recombination is thought to have played a major role in generation of novel alleles at various human leucocyte antigen (HLA) loci [64]. In our study, a high recombination rate was found (ρ = 16) in DRB and the recombination event was further confirmed by GARD test. In the LDhat software, the likelihood model excludes the effects of selection, increasing the likelihood that balancing selection influenced the estimated result. However, a previous study simulated a model of symmetric balancing selection with recombination, indicating that LDhat works well even when a great deal of recombination and mutation expected to accumulate among sequences maintained by balancing selection [38].

3.3 Conclusions

Control region DNA and microsatellites are believed to be neutral or nearly neutral genes and are representatives of mitochondrial DNA and nuclear DNA respectively; while DRB genes are a good example of adaptive related genes and the polymorphism are mainly maintained by balancing selection [22,23]. Despite the differences among these genes, the same genetic variation pattern was found in the populations of golden snub-nosed monkey with the lowest polymorphism in SNJ population. The results revealed that the long isolated status of the small populations making genetic drift may play a greater role in different kinds of genes regardless of different evolutionary mechanisms. This

Table 4 Results of maximum-likelihood models for exon 2 of the DRB gene

Model code	*P*	Log-likelihood	Parameter estimates	Positively selected sites
M0 (one ratio)	1	−2429.84	$\omega = 0.748$	None
M1a (nearly neutral)	1	−2455.86	$p_0 = 0.916$ ($p_1 = 0.084$)	Not allowed
M2a (positive selection)	3	−2404.20	$p_0 = 0.858$, $p_1 = 0.120$ ($p_2 = 0.022$) $\omega_2 = 4.904$	**1R**,3L,12E,**14F**,**16S**,33Y,35Y,40Y,41V,*50F*,*59P* **60V**,*63N*,64F,**70F**,**73Q**,**74R**,**77Q**,*81Y*,*89V*
M3 (discrete)	4	−2391.34	$p_0 = 0.930$, $p_1 = 0.069$ ($p_2 = 0.001$) $\omega_1 = 3.150$, $\omega_2 = 28.757$	Not allowed
M7 (beta)	2	−2456.47	$p = 0.010$, $q = 0.066$	Not allowed
M8 (beta and omega)	4	−2404.41	$p_0 = 0.975$ ($p_1 = 0.024$) $p = 0.011$, $q = 0.072$, $\omega = 4.673$	**1R**,3L,*12E*,**14F**,**16S**,29Y,31Q,33Y,*35Y*,*40Y*,*41V*, **50F**,**59P**,**60V**,**63N**, 64F,67Q, **70F**, **73Q**,**74R**, **77Q**, 80N, **81Y**,87G, **89V**

a) *P* is the number of parameters in the ω distribution, ω is the selection parameter and pn is the proportion of sites falling into the ωn site class. For models M7 and M8, *p* and *q* are the shape parameters of the β function. Positively selected sites were identified in models M2a and M8 by the Bayes empirical Bayes procedure [45]. Sites inferred under selection at the 99% level are listed in bold, and those inferred at the 95% level are shown in italics.

Table 5 Summary of test statistics for the likelihood-ratio test of codon evolution at DRB exon 2

Models compared	d*f*	Test statistic	Significance (*P*)
M2a vs. M1a	2	103.32	<0.001
M3 vs. M0	4	77	<0.001
M8 vs. M7	2	104.12	<0.001

research may contribute to making effective management decisions. Though the population size of golden snub-nosed is largest among the snub-nosed monkeys [65], as a flagship endangered species, it is also important to keep genetic health other than population increase [66].

This work was supported by the National Natural Science Foundation of China (31130061 and 30970427), the Project of Public Benefit (201104073), the Beijing Forestry University Young Scientist Fund, and State Forestry Administration of China. We thank two anonymous reviewers for their helpful comments on the former version of the manuscript. Thanks to Ren B P, Chang Z F, Yang B H, Wang B S and Liu Z J for sampling and lab assistance. Special thanks to Dr. Li Ming and Alicia Krzton for English polishing on an earlier draft.

1 Altizer S, Harvell D, Friedle E. Rapid evolutionary dynamics and disease threats to biodiversity. Trends Ecol Evol, 2003, 18: 589–596
2 Wahlberg N, Moilanen A, Hanski I. Predicting the occurrence of endangered species in fragmented landscapes. Science, 1996, 273: 1536–1538
3 Meffe G K, Carroll C G, contributors: Principles of Conservation Biology. 2nd ed. Massachusetts: Sinauer Associates, 1997
4 Peacock M M, Smith A T. The effects of habitat fragmentation on dispersal patterns, mating behavior and genetic variation in a pica (*Ochotona princeps*) metapopulation. Oecol, 1997, 112: 524–533
5 Sommer S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front Zool, 2005, 2: 16
6 Allendorf F W, Luikart G. Conservation and the Genetics of Populations. Malden: Blackwell Publishing, 2007
7 Brumfield R T, Beerli P, Nickerson D A, et al. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol Evol, 2003, 18: 249–256
8 Morin P A, Luikart G, Wayne R K. The SNP workshop group: SNPs in ecology, evolution and conservation. Trends Ecol Evol, 2004, 19: 208–216
9 Meyers L A, Bull J J. Fighting change with change: Adaptive variation in an uncertain world. Trends Ecol Evol, 2002, 17: 551–557
10 Van Tienderen P H, de Haan A A, van der Linden G, et al. Biodiversity assessment using markers for ecologically important traits. Trends Ecol Evol, 2002, 17: 577–582
11 Kohn M H, Murphy W J, Ostrander E A, et al. Genomics and conservation genetics. Trends Ecol Evol, 2006, 21: 629–637
12 Bonin A, Nicole F, Pompanon F, et al. Population adaptive index: A new method to help measure intraspecific genetic diversity and prioritize populations for conservation. Conserv Biol, 2007, 21: 697–708
13 Crandall K A, Bininda-Emonds O R P, Mace G M, et al. Considering evolutionary processes in evolutionary biology. Trends Ecol Evol, 2000, 15: 290–295
14 Stockwell C A, Hendry A P, Kinnison M T. Contemporary evolution meets conservation biology. Trends Ecol Evol, 2003, 18: 94–101
15 Siddle H V, Kreiss A, Eldridge M D B, et al. Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA, 2007, 104: 16221–16226
16 Miller H C, Allendorf F , Daugherty C H. Genetic diversity and differentiation at MHC genes in island populations of tuatara (*Sphenodon* spp.). Mol Ecol, 2010, 19: 3894–3908
17 Spurgin L G, Richardson D S. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc Roy Soc B-Biol Sci, 2010, 277: 979–988
18 Schwensow N, Fietz J, Dausmann K H. Neutral versus adaptive genetic variation in parasite resistance: Importance of major histocompatibility complex supertypes in a free-ranging primate. Heredity, 2007, 99: 265–277
19 Bernatchez L, Landry C. MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? J Evol Biol, 2003, 16: 363–377
20 Spurgin L G, Richardson D S. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc Roy Soc B, 2010, 277: 979–988
21 Dengjel J, Schoor O, Fischer R, et al. Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. Proc Natl Acad Sci USA, 2005, 102: 7922–7927
22 Ohta T. On the pattern of polymorphisms at major histocompatibility complex loci. J Mol Evol, 1998, 46: 633–638
23 Radwan J, Biedrzycka A, Babik W. Does reduced MHC diversity decrease viability of vertebrate populations? Biol Conserv, 2010, 143: 537–544
24 Stenzel A, Lu T, Koch W A, et al. Patterns of linkage disequilibrium in the MHC region on human chromosome 6p. Hum Genet, 2004, 114: 377–385
25 Kelley J, Walter L, Trowsdale J. Comparative genomics of major histocompatibility complexes. Immunogenetics, 2005, 56: 683–695
26 Li M, Liu Z J, Gou J X, et al. Phylogeography and population struc-

ture of the golden monkeys (*Rhinopithecus roxellana*): Inferred from mitochondrial DNA sequences. Am J Primatol, 2007, 69: 1195–1209
27 Pan D, Hu H X, Meng S J, et al. A high polymorphism level in *Rhinopithecus roxellana*. Int J Primatol, 2009, 30: 337–351
28 Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res, 1997, 25: 4876–4882
29 Tamura K, Dudley J, Nei M, et al. MEGA 4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol, 2007, 24: 1596–1599
30 Bryja J, Charbonnel N, Berthier K, et al. Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles. Mol Evol, 2007, 16: 5084–5097
31 Kennedy L J, Ryvar R, Gaskell R M, et al. Sequence analysis of MHC DRB alleles in domestic cats from the United Kingdom. Immunogenetics, 2002, 54: 348–352
32 Reche P A, Reinherz E L. Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. J Mol Biol, 2003, 331: 623–641
33 Guindon S, Gascuel O. A simple, fast and accurate algorithm to estimate large phylogenies by max-imum likelihood. Sys Biol, 2003, 52: 696–704
34 Posada D, Crandall K A. Modeltest: Testing the model of DNA substitution. Bioinformatics, 1998, 14: 817–818
35 Huson D H, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol, 2006, 23: 254–267
36 Hudson R R. Two-locus sampling distributions and their application. Genetics, 2001, 159: 1805–1817
37 McVean G, Awadalla P, Fearnhead P. A coalescennt-based method for detecting and estimating recombination from gene sequences. Genetics, 2002, 160: 1231–1241
38 Richman A D, Herrera L G, Nash D, et al. Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in *Peromyscus maniculatus*. Genet Res Camb, 2003, 82: 89–99
39 Delport W, Poon A F, Frost S D, et al. Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics, 2010, 26: 2455–2457
40 Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 2009, 25: 1451–1452
41 Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol, 2007, 24: 1586–1591
42 Anisimova M, Nielsen R, Yang Z. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics, 2003, 164: 1229–1236
43 Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol, 1994, 11: 725–736
44 Yang Z, Nielsen R, Goldman N, et al. Codon substitution models for heterogeneous selection pressure at amino acid sites. Genetics, 2000, 155: 431–449
45 Yang Z, Wong W S W, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol, 2005, 22: 1107–1118
46 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
47 Lafont B A P, McGraw C M, Stukes S A, et al. The locus encoding an oligomorphic family of MHC-A alleles (Mane-A*06/Mamu-A*05) is present at high frequency in several macaque species. Immunogenetics, 2007, 59: 211–223
48 Schwensow N, Eberle M, Sommer S. Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. Proc Roy Soc B, 2008, 275: 555–564
49 Doxiadis G G M, Rouweler A J M, Groot N G, et al. Extensive sharing of MHC class II alleles between rhesus and cynomolgus macaques. Immunogenetics, 2006, 58: 259–268
50 Nei M, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA, 1997, 94: 7799–7806
51 Pan D, Hu H X, Meng S J, et al. Population analysis of golden monkey using mitochondrial control region: High level of polymorphism and its implications. Chin Sci Bull, 2005, 50: 2489–2494
52 Bollmer J L, Hull J M, Ernest H B, et al. Reduced MHC and neutral-variation in the Galapagos hawk, an island endemic. BMC Evol Biol, 2011, 11: 43
53 Weber D S, Stewart B S, Schienman J, et al. Major histocompatibility complex variation at three class II loci in the northern elephant seal. Mol Evol, 2004, 13: 711–718
54 Babik W, Pabijan M, Arntzen J W, et al. Long-term survival of a urodele amphibian despite depleted major histocompatibility complex variation. Mol Evol, 2009, 18: 769–781
55 Mason R A B, Browning T L, Eldridge D B. Reduced MHC class II diversity in island compared to mainland populations of the black-footed rock-wallaby (*Petrogale lateralis lateralis*). Conserv Gene, 2011, 12: 91–103
56 Biedrzycka A, Radwan J. Population fragmentation and major histocompatibility complex variation in the spotted suslik, *Spermophilus suslicus*. Mol Evol, 2008, 17: 4801–4811
57 Luo M F, Liu Z J, Pan H J, et al. Historical geographic dispersal of the golden snub-nosed monkey (*Rhinopithecus roxellana*) and the influence of climatic oscillations. Am J Primatol, 2012, 74: 91–101
58 Garrigan D, Hedrick P W. Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. Evolution, 2003, 57: 1707–1722
59 Thomas J C, Godfrey P A, Feldgarden M, et al. Candidate targets of balancing selection in the genome of *Staphylococcus aureus*. Mol Biol Evol, 2012, 29: 1175–1186
60 Klein J, Sato A, Nikolaidis N. MHC, TSP, and the origin of species: From immunogenetics to evolutionary genetics. Annu Rev Genet, 2007, 41: 281–304
61 Doxiadis G G M, Rouweler A J M, Groot N G de, et al. Extensive sharing of MHC class II alleles between rhesus and cynomolgus macaques. Immunogenetics, 2006, 58: 259–268
62 Xu S X, Ren W H, Li S Z, et al. Sequence polymorphism and evolution of three cetacean MHC genes. J Mol Evol, 2009, 69: 260–275
63 Kauppi L, Sajantila A, Jeffreys A J. Recombination hotspots rather than population history dominate linkage disequilibrium in the MHC class II region. Hum Mol Genet, 2003, 12: 33–40
64 Carnngton M. Recombination within the human MHC. Immunol Rev, 1999, 167: 245–256
65 Pan H J, Shi F L, Chang Z F, et al. Mitochondrial DNA variation analysis suggests extreme low genetic diversity in Guizhou snub-nosed monkeys (*Rhinopithecus brelichi*). Chin Sci Bull, 2011, 56: 2541–2544
66 Fang S G. Does population increase alone ensure the long-term survival of endangered species? Chin Sci Bull, 2011, 56: 2521–2522