BMC
Evolutionary Biology

RESEARCH ARTICLE

Balancing selection and genetic drift at major histocompatibility complex class II genes in isolated populations of golden snub-nosed monkey (*Rhinopithecus roxellana*)

Mao-Fang Luo[1,2,3], Hui-Juan Pan[4], Zhi-Jin Liu[1] and Ming Li[1*]

Abstract

Background: Small, isolated populations often experience loss of genetic variation due to random genetic drift. Unlike neutral or nearly neutral markers (such as mitochondrial genes or microsatellites), major histocompatibility complex (MHC) genes in these populations may retain high levels of polymorphism due to balancing selection. The relative roles of balancing selection and genetic drift in either small isolated or bottlenecked populations remain controversial. In this study, we examined the mechanisms maintaining polymorphisms of MHC genes in small isolated populations of the endangered golden snub-nosed monkey (*Rhinopithecus roxellana*) by comparing genetic variation found in MHC and microsatellite loci. There are few studies of this kind conducted on highly endangered primate species.

Results: Two MHC genes were sequenced and sixteen microsatellite loci were genotyped from samples representing three isolated populations. We isolated nine *DQA1* alleles and sixteen *DQB1* alleles and validated expression of the alleles. Lowest genetic variation for both MHC and microsatellites was found in the Shennongjia (SNJ) population. Historical balancing selection was revealed at both the *DQA1* and *DQB1* loci, as revealed by excess non-synonymous substitutions at antigen binding sites (ABS) and maximum-likelihood-based random-site models. Patterns of microsatellite variation revealed population structure. F_{ST} outlier analysis showed that population differentiation at the two MHC loci was similar to the microsatellite loci.

Conclusions: MHC genes and microsatellite loci showed the same allelic richness pattern with the lowest genetic variation occurring in SNJ, suggesting that genetic drift played a prominent role in these isolated populations. As MHC genes are subject to selective pressures, the maintenance of genetic variation is of particular interest in small, long-isolated populations. The results of this study may contribute to captive breeding and translocation programs for endangered species.

Keywords: Balancing selection, Conservation genetics, Gene drift, MHC, *Rhinopithecus roxellana*

* Correspondence: lim@ioz.ac.cn
[1]Key laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1-5 Beixhenxi Road, Chaoyang, Beijing 100101, China
Full list of author information is available at the end of the article

Background

Understanding how levels of genetic variation influence the survival of threatened species is of fundamental interest to evolutionary and conservation biologists because many natural populations are threatened by intense reduction and fragmentation of habitat, leading to isolation, declining populations, and decreasing genetic diversity [1,2]. Loss of genetic diversity may increase the risk of extinction due to decreased reproductive fitness, decreased adaptive flexibility, and increased disease susceptibility [3]. Clarifying the mechanism that determines genetic variation in small, isolated populations is therefore essential for their conservation [4]. An important assumption in conservation genetics is that small, isolated populations are more sensitive to genetic drift and inbreeding [5,6]. Genetic drift is the random fluctuation of allele frequencies over time; thus, adaptive alleles may be lost and deleterious alleles could be fixed in the population. The small population size and fixation of deleterious alleles leads to inbreeding depression and reduction of individual fitness, which decreases viability and compromises a population's evolutionary adaptive potential [6].

However, some functionally important genes that are maintained by balancing selection, such as major histocompatibility complex (MHC) genes, may have a different evolutionary pattern compared with neutral markers. The multi-gene MHC family is found in vertebrates, codes for cell surface glycoproteins, and is important in animal conservation due to its role in resisting pathogens [7]. Compared to nearly neutral markers such as microsatellite loci or mitochondrial DNA, which are informative for phylogenetic and phylogeographic reconstructions [8], MHC variability is believed to determine the capability of individuals to resist continuously evolving pathogens and parasites. Consequently, MHC variability is a reflection of the processes that are related to adaptive evolution within and between populations [9]. Thus, most variation at MHC loci reflects the effects of balancing selection [10], which is the main mechanism for retaining high MHC genetic diversity. Balancing selection includes frequency-dependent selection, overdominance and diversifying selection and promotes long evolutionary persistence of individual alleles and strongly differentiated allelic lineages in mammals [2]. Besides balancing selection, intragenic recombination has also been suggested as one evolutionary mechanism for generation of MHC sequence diversity [11,12]. Still, several significant questions remain. Current results conflict or are unclear with regard to the relative roles of balancing selection and genetic drift in maintaining MHC polymorphism in small, isolated, or severely bottlenecked populations [13,14]. Studies of guppies, Mexican wolves and Namibian leopards found that even when genetic

diversity at neutral markers was poor, polymorphisms in MHC were still maintained by balancing selection [13-17]. In contrast, other research has found low levels of detectable polymorphisms for MHC genes in populations with lower diversity in neutral markers, including studies on fallow deer (*Cervus dama*) [18], northern elephant seals (*Mirounga angustirostris*) [19], great crested newt (*Triturus cristatus*) [20], and the black-footed rock-wallaby (*Petrogale lateralis lateralis*) [21]. These results suggest that compared to genetic drift balancing selection is relatively weak in small populations, leading to reduced variation at some MHC loci [22]. To elucidate the role of balancing selection and genetic drift in populations, F_{ST} outlier analysis [23] is widely used because demographic processes affect neutral loci and lead to population differentiation (measured by F_{ST}) [24]. The F_{ST} values are computed for all genes to distinguish the genes under selection from those non-selective genes [25].

We studied MHC and neutral genetic variation in the golden snub-nosed monkey (*Rhinopithecus roxellana*), an endangered primate endemic to China where it 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 22,000 individuals remain (about 15,000 individuals in SG, 5500 individuals in QL and 1000 individuals in SNJ) [26], and mitochondrial DNA and microsatellite studies indicate that the genetically distinct SNJ population has very low genetic variation [27,28]. The current effective population size (Ne) in SNJ is approximately 80 individuals [29]. The small size and isolation that typically characterize local populations make the golden snub-nosed monkey particularly susceptible to loss of genetic variation through inbreeding and genetic drift [30,31]. In this study, we 1) investigated population structure and polymorphism levels of 16 microsatellite loci and 2 MHC II genes, 2) tested for both selection at the MHC loci and patterns of between population differentiations, and 3) evaluated the role of balancing selection and genetic drift in populations.

Results

Genetic variation at microsatellites

Genotyping of the 16 microsatellite loci revealed that the highest level of expected heterozygosity (H_E) and observed heterozygosity (H_O) was in the SG population, while the lowest level was in the SNJ population (Table 1, Additional file 1: Table S1). The Hardy–Weinberg equilibrium test showed no deviation from equilibrium after Bonferroni correction. No linkage disequilibrium occurred for pairs of loci in any population, and

Figure 1 Distribution of isolated snub-nosed monkey populations.
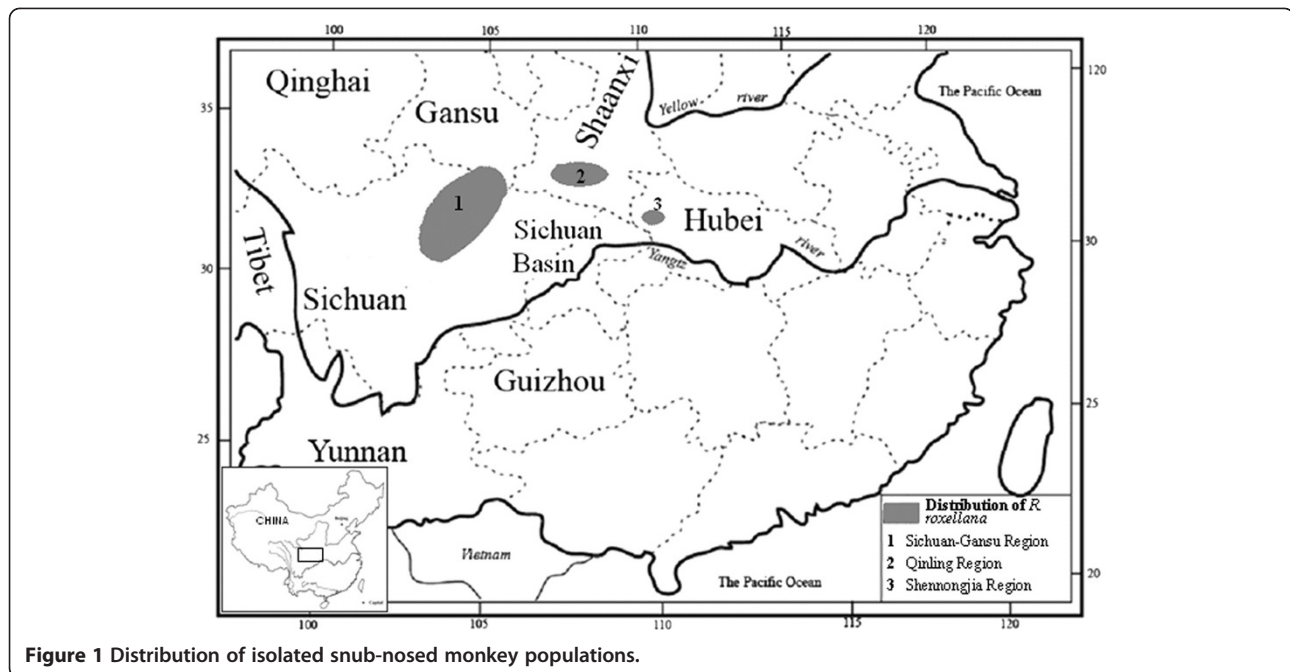
Micro-Checker revealed no evidence of stuttering, null alleles, and allele dropout.

MHC variation, expression analyses, and recombination analysis

For each locus, we analysed 894 clones from 64 individuals that represented three monkey populations. We obtained 13 different *DQA1* fragments with lengths that ranged from 436 bp to 444 bp including exon 2 and partial intron 2 and intron 3 [GenBank: JQ217094-JQ217106]. After performing BLAST with the *HLA-DQA1* and *Macaca mulatta DQA1* exon 2, we removed the intron and obtained 9 *DQA1* exon 2 sequences with an equal length of 249 bp. These sequences were labelled as *Rhro-DQA1*01–09* [GenBank: JQ217107-JQ217115] according to the nomenclature of Klein *et al.* [32] and were used for analysis. At the *DQB1* locus, we identified 16 unique *DQB1* exon 2 sequences that we called *Rhro-DQB1*01–16* [GenBank: JQ217116- JQ217131].

Table 1 Population genetic parameters for populations estimated from microsatellite data

Population	N	No. of Loci	A_R	$H_E \pm SD$	$H_O \pm SD$	F_{IS}
SG	25	16	4.635	0.736±0.022	0.714±0.022	0.017
QL	22	16	4.854	0.713±0.036	0.653±0.024	0.071
SNJ	17	16	3.473	0.611±0.038	0.591±0.030	0.030
All	64	16	4.321	0.755±0.022	0.660±0.014	0.041

A_R, allelic richness; H_E, expected heterozygosity; H_O, observed heterozygosity; F_{IS}, inbreeding coefficient.

Fragments of MHC loci were successfully amplified from the cDNA of 2 blood samples that were also used in DNA amplifications. At the *DQA1* locus, 189 bp fragments of *DQA1* exon 2 were obtained from 2 blood samples, and these fragments were part of the alleles obtained from DNA in the same individual [EMBL: HE616682- HE616683]. At the *DQB1* locus longer fragments were obtained, including the entire exon 2, partial exon 1, and exon 3; however, these fragments were only obtained from 1 blood sample [GenBank: JQ217132-JQ217133]. Shorter fragments of *DQB1* exon 2 were obtained from 2 blood samples [EMBL: HE616684-HE616686]. No sequences contained stop codons. The sequences from RNA were identical with that from DNA, which supports the hypothesis that these MHC PCR products were transcribed and expressed.

No more than two alleles were present per individual, suggesting that for each MHC locus one gene copy was sequenced with these primer sets. No stop codon or insertion/deletion was detected in *DQA1* and *DQB1*. The measures of MHC diversity in the 3 populations are summarized in Table 2. All three populations shared four of the nine *DQA1* alleles and two of the among 16 *DQB1* alleles (Additional file 2: Table S2). The MHC sequences were highly divergent: 61 of 249 bp of *DQA1* and 70 of 256 bp of *DQB1* were variable. The Hardy–Weinberg equilibrium tests showed that the QL population deviated from equilibrium (p = 0.006). Linkage disequilibrium occurred in the SG population (p = 0.01). Population recombination analysis using LDhat revealed that the *DQB1* locus had a higher

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

MHC	Population	N	Variable sites	Parsimony- informative sites	Total alleles	Private alleles	H	A$_R$	Over mean distance	F_{IS}
DQA1	SG	25	59	29	6	1	0.798	5.628	0.118±0.016	0.675
	QL	22	48	30	8	3	0.851	7.661	0.090±0.013	0.511
	SNJ	17	43	7	4	0	0.671	4.000	0.090±0.015	0.123
	All	64	61	33	9	-	0.972	7.503	0.102±0.014	0.495
DQB1	SG	25	65	32	7	3	0.768	6.126	0.115±0.015	0.325
	QL	22	57	36	11	6	0.861	9.962	0.085±0.012	0.452
	SNJ	17	47	20	5	2	0.743	5.000	0.092±0.014	0.251
	All	64	70	56	16	-	0.891	9.587	0.097±0.012	0.360

Private alleles, alleles only found in this population; h, genetic variation; A$_R$, allele richness; F_{IS}, inbreeding coefficient.
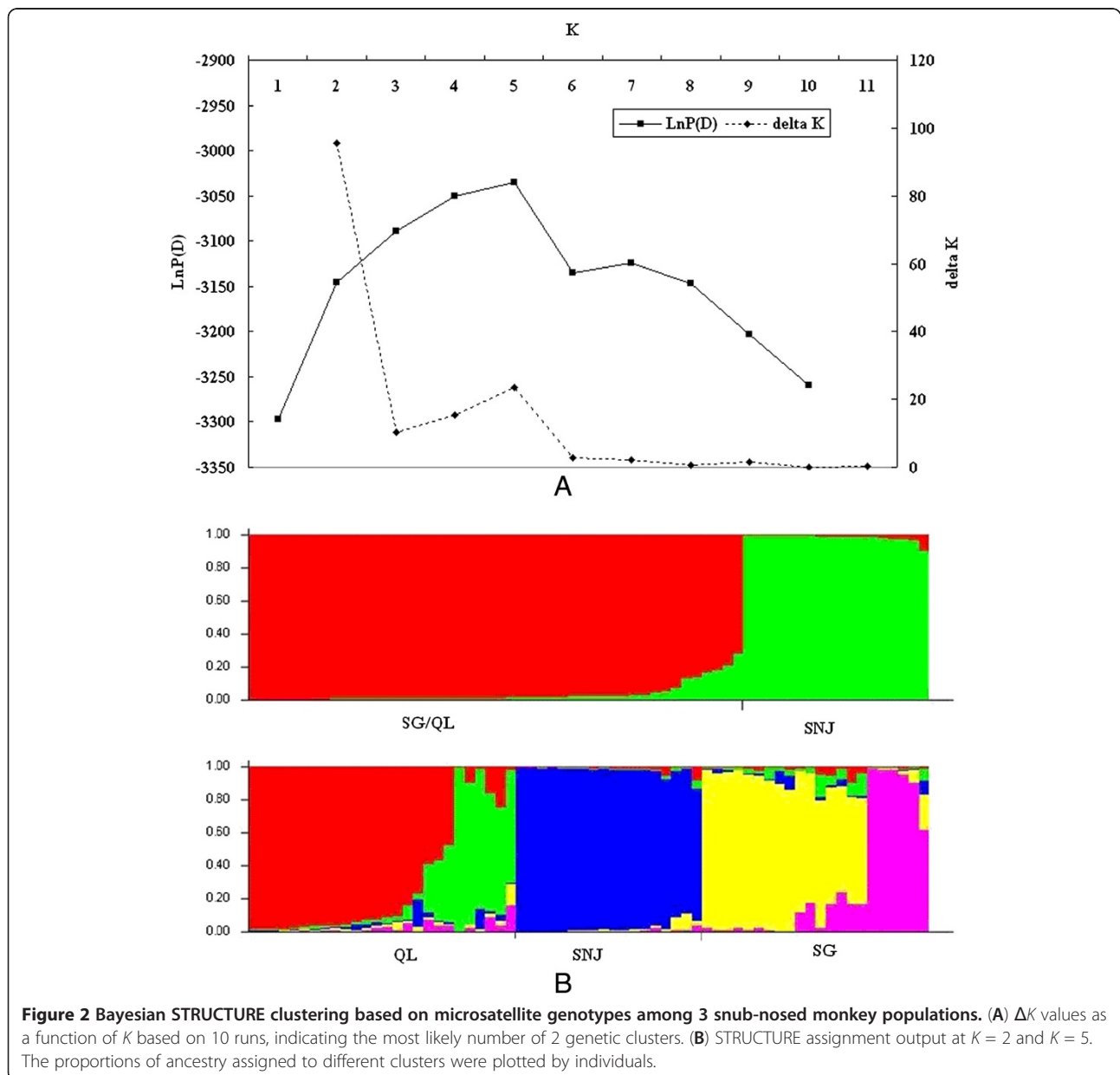

Figure 2 Bayesian STRUCTURE clustering based on microsatellite genotypes among 3 snub-nosed monkey populations. (A) Δ*K* values as a function of *K* based on 10 runs, indicating the most likely number of 2 genetic clusters. (**B**) STRUCTURE assignment output at *K* = 2 and *K* = 5. The proportions of ancestry assigned to different clusters were plotted by individuals.

Figure 3 Differences of pairwise F_{ST} values among populations in MHC genes and microsatellite loci.
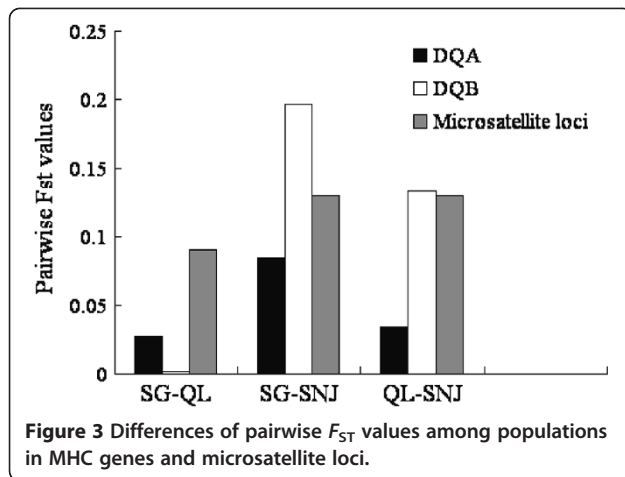
recombination rate (ρ = 7) when all alleles were included between the MHC loci (all DQA alleles: ρ = 2).

Population structure and phylogenetic analysis

Microsatellite-based analyses detected significant genetic structures among the populations (Figure 2). Two genetic clusters were found (SG-QL, SNJ) with a clear maximum ΔK (ΔK = 95.500 at K = 2). The pairwise F_{ST} was the largest between SNJ and the other 2 populations among all of the F_{ST} values (Figure 3), which showed that the differences between SG and QL were minimal compared to those between SG-QL and SNJ. However, there was also a small ΔK peak at K = 5 (ΔK = 23.59), and results showed that in addition to the divergence between SG and QL, there was also divergence within the SG and QL populations. These results agree with earlier studies, where the differences within the SG and QL populations resulted from their different origins and habitat fragmentation [33,34]. The clear population structure observed in microsatellites was not found at either the *DQA1* locus or the *DQB1* locus (Additional file 3: Figure S1). Similar to previous studies [7], the relationship among MHC alleles did not agree with the geographical locations of the 3 populations, as observed in the microsatellite results (Figure 4). *DQA1/DQB1* alleles were not grouped by species but were mixed with other species (trans-species polymorphism (TSP). No clear clade could be identified from the phylogenetic
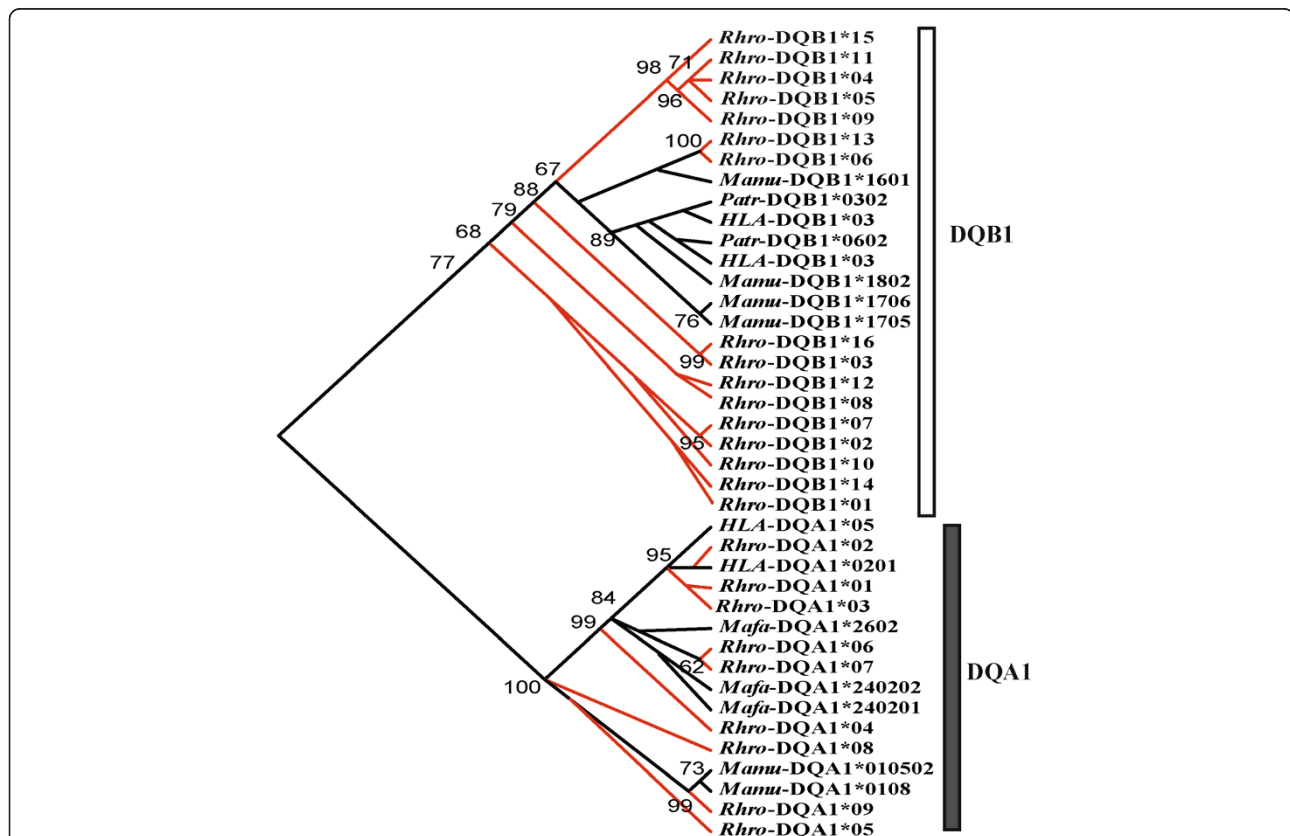

Figure 4 Phylogenetic relationships among snub-nosed monkey MHC alleles. The phylogenetic tree was reconstructed using the ML inference method. Bootstrap values above 60 are given for each clade. The alleles obtained in this study are shown as red bars. Other allelic sequences, which were downloaded from GenBank are included in the analyses: *Homo sapiens* (HSU97555, AY375917, GQ422610, AY334565); *Macaca fascicularis* (AM086058-AM086060); *Macaca mulatta* (M81297, M81292, AJ308047, AJ308046); *Pan troglodytes* (M81260, M81262).

tree at these two loci. Populations had more private alleles at *DQB1* than *DQA1* (Table 2). Four among 9 *DQA1* exon 2 alleles were commonly shared among the 3 populations, while just 2 (*Rhro-DQB1*01, 03*) among 16 *DQB1* exon 2 alleles were shared among all of the populations (Additional file 2: Table S2). However, frequency differences of these shared alleles were found between all population pairs (except *Rhro-DQA1*01, 02*). Comparisons of pairwise F_{ST} values of MHC and microsatellite sites are shown in Figure 3. The F_{ST} values showed lower divergence among populations compared with microsatellite loci (except SG-SNJ and QL-SNJ at the *DQB1* locus). Half of the F_{ST} values at the MHC loci were higher than 0.05, indicating that differences existed among populations, especially between SG and SNJ. The population difference was more distinct at the *DQB1* locus: the differentiations between the SG and SNJ populations and between the QL and SNJ populations were even larger than the differentiation at the microsatellite.

Historical selection

Evidence for historical balancing selection was detected. First, in the ABS, the number of nonsynonymous (d_N) mutations was greater than the number of synonymous (d_S) mutations using the Z-test (*DQA1*: d_N/d_S = 1.15, p = 0.405; *DQB1*: d_N/d_S = 4.41, p = 0.01) (Table 3). Although p-value was not significant for *DQA1* locus in Mega, ABS sites under significant selection were detected in PAML 4 (Additional file 4: Table S3). After comparisons of various codon evolution models using CODEML in PAML 4, models integrating positive selection (M2a, M8, and M3) matched MHC better than the other models based on the Akaike information criterion (AIC) values (Tables 4 and 5). Under model M2a, 9 *DQA1* sites and 5 *DQB1* sites were exposed to significant selection. Under model M8, 11 *DQA1* sites and 9 *DQB1* sites were identified. Moreover, most of these sites were ABS sites and others located near ABS sites (Additional file 4: Table S3, Additional file 5: Table S4). The existing TSP at these loci provided additional

Table 3 Average non-synonymous substitutions per non-synonymous site (d_N) and synonymous substitutions per synonymous sites (d_S)

MHC	Sites	N	d_N	d_S	d_N/d_S	Z	P
DQA1	ABS	20	0.152±0.048	0.132±0.080	1.15	0.247	0.405
	Non-ABS	62	0.081±0.020	0.113±0.031	0.72	−0.882	0.380
	All	82	0.100±0.020	0.114±0.026	0.88	−0.426	0.671
DQB1	ABS	23	0.247±0.066	0.056±0.033	4.41	2.477	0.01
	Non-ABS	62	0.063±0.014	0.092±0.022	0.68	−1.104	0.272
	All	85	0.103±0.018	0.083±0.017	1.24	0.709	0.480

evidence in support of historical balancing selection (Figure 4).

F_{ST} outlier analysis

The F_{ST} outlier analysis did not reveal a high or low level of differentiation among populations at the *DQA1* and *DQB1* loci compared with 16 microsatellite loci (Figure 5A). One microsatellite locus (D14S306) was located outside the 95% confidence interval. For each population, the MHC loci were also located within the 95% confidence interval of the microsatellites. Some satellites existed outside the 95% confidence interval: in the SG population, D14S306 was lower than the neutral level, and D1S1665 was at the edge of the neutral level (Figure 5B); in the QL population, D6S1056 was under the candidate positive selection level (Figure 5C); and in the SNJ population, D1S1665 was above the neutral level while D6S474 was below (Figure 5D).

Discussion

Habitat isolation generates barriers to gene flow among populations that often result in loss of genetic diversity through genetic drift and inbreeding [2]. For vertebrates, genetic variation is of special importance in MHC genes due to the significant role they play in immune functions [35]. MHC diversity is presumed to improve parasite resistance, reproductive success, and population viability [36], and has been studied in species such as western gorilla (*Gorilla gorilla*) [37], brown bear (*Ursus arctos*) [38], Ethiopian wolf (*Canis simensis*) [39], European bison (*Bison bonasus*) [40], Bengal tiger (*Panthera tigris tigris*) [41], and Namibian leopard (*Panthera pardus pardus*) [17]. In this study, a relatively high level of MHC variation was found in the golden snub-nosed monkey, with 9 *DQA1* and 16 *DQB1* alleles found in 64 individuals. However, we found lower MHC variation in the SNJ population, which also showed lower genetic diversity in microsatellites and mitochondrial genes [27,34]. Small sample size cannot account for the reduced variability, because sample collected across the entire distribution of the SNJ population showed similar patterns of genetic diversity [29].

For each MHC locus, regardless of how many clones were sequenced from an individual, no more than 2 alleles were observed in an individual, a strong indicator that we amplified single loci in all cases. We assumed that all alleles were from a single functional gene. The assumption that our sequences were from functional genes was supported by three findings: 1) sites inferred to have been exposed to significant selection, most of which were ABS sites, indicating historical selection at functional genes; 2) no reading frame or stop codon disruptions found in any alleles; and 3) the

Table 4 Results of maximum-likelihood models for exon 2 of the *DQA1* gene

Model code	P	Log-likelihood	Parameter estimates	Positively selected sites
M0(one ratio)	1	−774.550	$\omega=0.643$	None
M1a (nearly neutral)	1	−760.293	$p_0= 0.723$ $(p_1= 0.277)$	Not allowed
M2a (positive selection)	3	−752.443	$p_0= 0.911$, $p_1= 0.033$ $(p_2= 0.056)$ $\omega_2= 5.148$	*13T*,**16F,**35G,42R,*45E*,46L,48K,**50G,54P,**56G,59R,**61L,63T,64S,70I,**71M,74R
M3 (discrete)	4	−752.206	$p_0= 0.736$, $p1= 0.244$ $(p_2= 0.018)$ $\omega_1= 1.298$, $\omega_2= 8.174$	Not allowed
M7(beta)	2	−761.268	$p= 0.008$, $q= 0.028$	Not allowed
M8(beta and omega)	4	−752.248	$p_0= 0.974$ $(p_1= 0.025)$ $p= 0.008$, $q= 0.023$, $\omega= 7.292$	10L,**13T,16F,29Q,35G,42R,43W,45E,**46L,48K,**50G,**51G,54P,56G,59R,**61L,***63T,64S*,70I,71M,74R,79A

P is the number of parameters in the ω distribution, 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. Positively selected sites were identified in models M2a and M8 by the Bayes empirical Bayes procedure (Yang *et al.* 2005). Sites inferred under selection at the 99% level are listed in bold, and those inferred at the 95% level are shown in italics.

sequences identified from cDNA were parts of or included sequences obtained from DNA.

Historical balancing selection

Golden snub-nosed monkey MHC genes reflect historical balancing selection in that an excess of non-synonymous substitutions was mainly concentrated in the ABS (Table 3, Additional file 4: Table S3, Additional file 5: Table S4). According to neutrality theory [42], the synonymous nucleotide substitution rate is larger than the non-synonymous substitution rate because a change in amino acid sequence has a greater possibility of being deleterious. The elevated rate of non-synonymous substitutions at the ABS provided clear evidence of positive selection [43,44] shaping genetic variation [45]. The p-value was not significant at *DQA1* ABS may due to a weaker recombination within *DQA1* as a lower recombination rate has been shown before. Without higher recombination and stronger selection, some *DQA1*
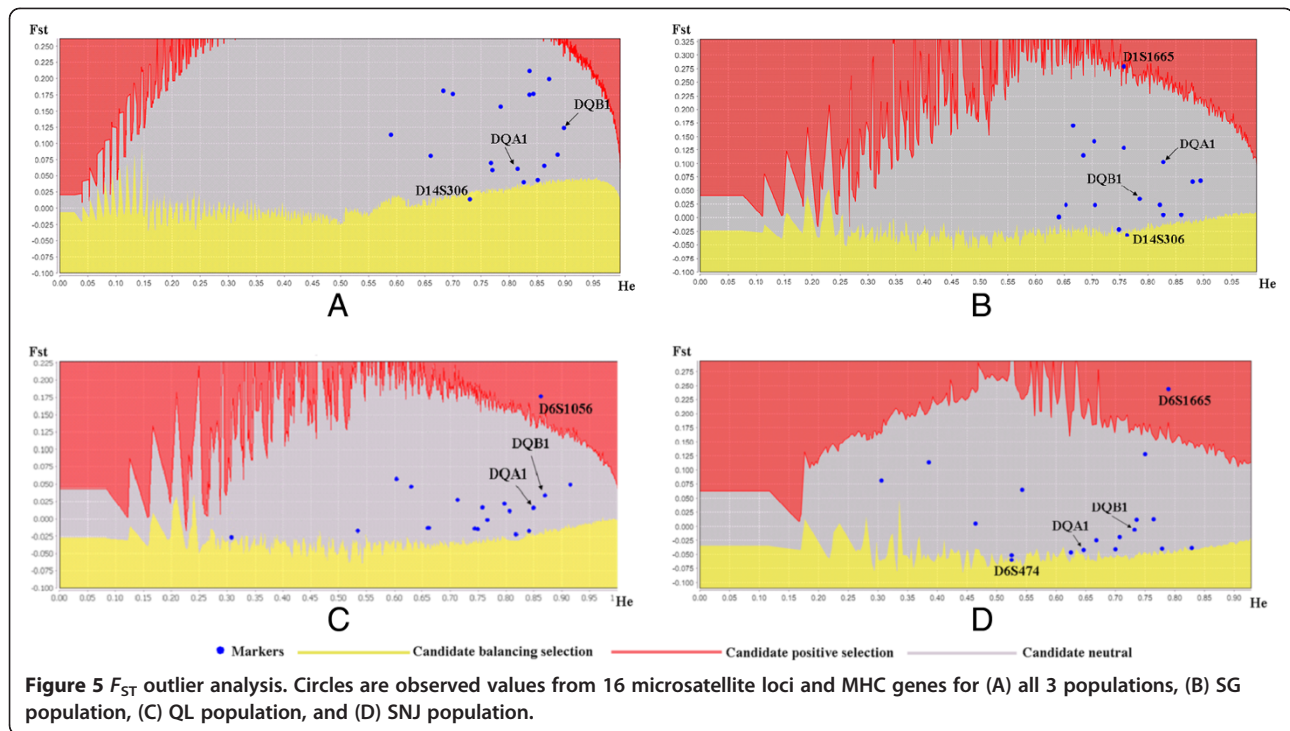
alleles might be lost when population size decreased. The sharing of MHC alleles among populations also indicates that MHC alleles may have been conserved by selection [19]. Second, random site models analysis proved the existence of historical selection based on the maximum likelihood method, which revealed that, for MHC genes, the models including selection (M2a, M3, and M8) match MHC alleles better than models without selection (Table 4,and 5). Under the M2a and M8 models, some sites of the 2 MHC loci were under significant selection pressure. Furthermore, trans-species evolution of the MHC alleles revealed historical balancing selection. Under balancing selection, some MHC alleles or allelic lineages are reported in other species, which indicates that they are ancestral alleles [46].

Genes vary in terms of the level of selection, and both *DQA1* and *DQB1* revealed different patterns of selection. Each population had unique *DQB1* alleles, while not every population had unique *DQA1* alleles (Table 2).

Table 5 Results of maximum-likelihood models for exon 2 of the *DQB1* gene

Model code	P	Log-likelihood	Parameter estimates	Positively selected sites
M0(one ratio)	1	−1067.274	$\omega=0.496$	None
M1a(nearly neutral)	1	−1007.026	$p_0= 0.869$ $(p_1= 0.131)$	Not allowed
M2a(positive selection)	3	−989.138	$p_0=0.818$, $p_1=0.176$ $(p_2= 0.006)$ $\omega_2= 8.276$	**1V,**6M,**18L,**20T,48S,*49S*,67L,*72R*,**77L,**79L,81T
M3(discrete)	4	−994.628	$p_1= 0.059$ $(p_2= 0.018)$ $\omega_1= 0.076$, $\omega_2= 3.390$	Not allowed
M7(beta)	2	−1007.600	$p= 0.006$, $q= 0.036$	Not allowed
M8(beta and omega)	4	−989.309	$p_0= 0.994$ $(p_1= 0.006)$ $p= 0.005$, $q= 0.020$, $\omega= 8.643$	**1V,**6M,**18L,***20T*,22Y,29Y,30A,38E,*48S*,**49S,**62G,63T,67L,*72R*,**77L,***79L*, *81T*

P is the number of parameters in the ω distribution, 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. Positively selected sites were identified in models M2a and M8 by the Bayes empirical Bayes procedure (Yang *et al.* 2005). Sites inferred under selection at the 99% level are listed in bold, and those inferred at the 95% level are shown in italics.

Figure 5 F_{ST} outlier analysis. Circles are observed values from 16 microsatellite loci and MHC genes for (A) all 3 populations, (B) SG population, (C) QL population, and (D) SNJ population.

Population divergence, measured as pairwise F_{ST}, was larger in *DQB1* than in *DQA1* except for that between SG and QL (Figure 3). Similar results were reported in water vole, where balancing selection pressure was different at MHC genes in continuous populations [47]. MHC genes are assumed to be closely linked [48]. In our study, however, linkage disequilibrium between *DQA1* and *DQB1* was only observed in the SG population. This weak linkage disequilibrium and the different selection pressures on these loci could be a result of recombination, which is common at the MHC genes [49,50]. The higher recombination rate that was found in *DQB1* genes may explain their larger allelic richness compared to *DQA1* genes. Recombinants maintained by selection may counteract the linkage of closely linked genes [51] and play an important adaptive role in *DQB1* evolution. In the present study, historical selection was found, but this does not conclusively indicate that balancing selection is acting on current populations. First, an excess of non-synonymous mutations requires a long time to accumulate [47]. Once present, this pattern would take a long time to vanish after the disappearance of selection [10]. Hence, we investigated whether selection continues to play a major role at present.

Patterns of selection and drift
Although selection historically maintained MHC diversity, recent population isolation and fragmentation has increased the role of genetic drift in shaping patterns of MHC variation in snub-nosed monkeys. First, compared with neutral forces, balancing selection is supposed to diminish population differentiation as measured by conventional pairwise F_{ST} [22,52]. Thus, the population structure of genes under balancing selection should not be pronounced [53]. However, in the present study, half of the pairwise F_{ST} values were greater than 0.05, and two F_{ST} values at DQB were even greater than those at microsatellites (Figure 3). Second, F_{ST} outlier analysis showed that the structure in the MHC loci was within the neutrality level for all populations and for each population. Considering all populations, one microsatellite (D14S306) showed a F_{ST} value that was lower than the neutral level, indicating its linkage with other genes under selection [22]. Lastly, a positive correlation was found between allelic richness in MHC and microsatellites. The QL population had the highest allelic richness in microsatellites and MHC, while SNJ had the lowest. This positive correlation indicates that genetic drift plays a significant role in maintaining MHC diversity for snub-nosed monkeys [25]. Maintenance of MHC variation through balancing selection may be hampered in small, isolated populations because of their lower effective recombination rate [54]. In all, our results indicate that even though selection acts on MHC, it is overwhelmed by genetic drift in small, isolated populations.

The positive correlation in the allelic richness of MHC and microsatellites, together with other evidence, indicates that genetic drift has a great influence on the maintenance of MHC variations in small, isolated populations of snub-nosed monkeys [7,25,55]. No evidence showed that MHC polymorphism had increased in populations that contained low neutral variation [27,42]. Under neutral evolution theory [56], alleles are expected to be neutral when s < 1/2 Ne (s = selection coefficient, Ne = effective population size). Therefore, the smaller Ne becomes, the greater the likelihood of genetic drift [57]. The SNJ population is subject to more genetic drift than the other two populations as found in a previous study [29]. Other animals whose patterns of MHC polymorphism have been contributed to drift over selection include the great crested newt (*Triturus cristatus*) [19], black-footed rock-wallaby (*Petrogale lateralis lateralis*) [20], tuatara (*Sphenodon spp.*) [7], and the Egyptian vulture (*Neophron percnopterus*) [58]. These results suggest that selection on MHC is not strong enough to counteract drift that results from population fragmentation, isolation and bottleneck.

Conclusion

Variation at MHC loci is widely accepted as being maintained by balancing selection (reviewed in 2), even with a low level of neutral variability in some species [29,59]. This shows the importance of balancing selection for maintaining variation in the field and exposes the problem of using neutral genes as substitutes for variation in fitness-related genes [32]. However, in small, isolated populations or bottlenecked populations, balancing selection is overwhelmed by drift [21]. In this study, we found the same genetic variation pattern both at neutral and MHC markers, suggesting that genetic drift was stronger than selection, thus leading to a reduction in MHC diversity in the most isolated populations. Such findings may contribute to the conservation of endangered species such as snub-nosed monkeys both in captive breeding and translocation programs. Though the relationships between MHC with mate choice and pregnancy outcomes still remain controversial [60], it could be helpful to examine MHC variation in captive individuals that are involved in mating programs. In the field, translocation of individuals from demographically and genetically healthy populations to populations that suffer from reduced genetic diversity can improve the chances of genetic recovery [61,62]. Recently, researchers have found that inbred populations could thrive with the import of migrants as part of a conservation management program [63-65]. Further, the introduction of individuals from western populations of golden snub-nosed monkeys could restore genetic diversity to the relatively homogenous SNJ groups, reducing the likelihood of inbreeding depression.

Methods

Sample collection and DNA extraction

We collected 64 *R. roxellana* samples (muscle, skin, and blood) from the current population (SG population = 25, SNJ population = 17, and QL population = 22) (Figure 1 and Table 1). All collections complied with the relevant animal welfare institutions and laws of China. Also, the Institute of Zoology (Chinese Academy of Sciences) provided ethical approval for this study. Muscle and skin samples were collected from carcasses that were provided by local museums and nature reserves. Skin samples were stored dry, muscle samples were stored in 95% ethanol, and blood samples were collected while trapping individuals for physical examination and were stored in a refrigerator at –80°C. During DNA extraction, benches and plasticware were washed with 10% bleach and sterile water and were exposed to UV light for 30 min prior to treatment to prevent contamination during DNA extraction. During the subsequent polymerase chain reaction (PCR), 8 extraction controls were used without any positive amplification.

Microsatellite analyses

Samples were genotyped at 16 microsatellite loci [33]. Thirty-five cycles of PCR amplification were executed at the same time for up to 3 loci with reliable genotypes, with combinations selected by fluorescent dye (HEX, ISMRA, or FAM), Tm, and fragment size using the QIAGEN Mutliplex PCR kit following the manufacturer's protocols at optimized annealing temperatures (55°C). An ABI 377 prism automated sequencer was used to resolve products, which were then analysed using GeneScan v3.1.2 and Genotyper 2.5 (Applied Biosystems).

MHC amplification, cloning, sequencing, and expression analyses

The PCR was carried out in a 50 μL solution including 10 mM Tris–HCl (pH 8.4), 50 mM KCl, 2.5 mM MgCl$_2$, 0.4 μM each primer, 0.2 mM each dNTP, 1.0 unit Hotstart-Taq DNA polymerase (Takara), and 10–100 ng DNA template. The amplification profile consisted of 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 56°C (*DQA1*) or 58°C (*DQB1*), and 30 s at 72°C, with a final extension of 10 min at 72°C. The following primers were used for amplifying exon 2: 5'*DQA1*-AAGCCCA TAATATT TGAAAGTCAGT and 3'*DQA1*-TATGTGATTTTAGAGATGGGAGATG, or 5'*DQB1*-TGTAAAC GACGGCCAGTTCCCCGCA-GAGGATTTCGTG and 3'*DQB1*-TGCTCTAGAGGG-CGACGACGACGCCTCACCTC [66]. The *DQA1* primers were designed based on BLAST sequences from relative

species. According to the manufacturer's protocols, a Wizard PCR Preps DNA Purification Kit (Promega) was used to purify the PCR products. Purified PCR products were cloned using the pMD-18T vector (Takara) following manufacturer's instructions. Ten to twenty clones containing inserts from each individual 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.).

Expression analyses were conducted to validate the expression of the obtained MHC II-DQA and DQB alleles. The RNA was isolated from 2.5 mL whole blood using Trizol (Invitrogen) following the manufacturer's instructions. To ensure that the genomic DNA was removed from the isolated RNA, a second DNA digestion was performed using the DNase I RNase-free Set (Promega). The cDNA synthesis was obtained using 200 U of M-MuLV Reverse Transcriptase (Promega) in a 25 μL reaction tube containing 1μL Oligo(dT)12–18 primer (0.5 μg/μL; Invitrogen), 2 μg total RNA as a template, and 25 U ribonuclease inhibitor, RNase-free water, 5 μL dNTP mix (10 mM), and 5 μL 5× reaction buffer. The reactions were incubated at 42°C for 60 min. The cDNA was acquired from 2 samples and was amplified by PCR using the primer sets (Additional file 6: Table S5). The primer sets were exon-spanning to detect the amplification of genomic DNA contaminants based on the BLAST results of human and *Macaca mulatta* exons. All amplified cDNA products were analysed by cloning and sequencing as described above.

Data analysis
Microsatellite loci diversity analyses
Genetic diversity was measured as observed (H_O) and expected heterozygosities (H_E) [67]. The above analyses were computed using Arlequin v3 [68]. Inbreeding coefficients (F_{IS}) and allelic richness (A_R) analyses were performed in FSTAT version 2.9.3 [69]. The A_R calculations were set to the smallest population size correction. Pairwise F_{ST} among populations and Hardy–Weinberg equilibrium were analysed across all loci for each population in an exact probability test in GENEPOP 4.0 [70]. Using the Bonferroni correction, significance values were adjusted for multiple comparisons. Linkage disequilibrium (LD) was calculated across all loci with the Weir correlation coefficient [71] in GENEPOP 4.0 [70]. To determine which LD values were significant, a permutation test was used. The presence of stuttering, null alleles, and allele dropout were examined using Micro-Checker [72].

MHC diversity analyses
Sequences were aligned and translated into the amino acid sequences in MEGA 4 [73]. The gene identity was verified through homology with publicized MHC alleles of other species using BLASTN (http://blast.ncbi.nlm.nih.gov/Blast.cgi) from NCBI. Because some obtained sequences are false alleles potentially corresponded to PCR amplification artifacts [47], a new sequence was assumed to be an allele when it was identified from 3 separate amplications from the same sample or from at least 3 different samples [74]. We used MEGA 4 to detect the variable and parsimony-informative sites (sites with at least 2 different nucleotides or amino acids), to compute the mean number of nucleotide differences, and to derive the overall mean genetic distances of nucleotide sequences. Standard estimate errors were detected using 1,000 bootstrap replicates. Hardy–Weinberg equilibrium tests, pairwise F_{ST} values, and LD were calculated in GENEPOP. The A_R and F_{IS} of *DQA1* and *DQB1* were compared among populations using the comparison among groups of samples option in FSTAT version 2.9.3 [69]. The A_R estimates were adjusted to the smallest sample size.

Phylogenetic analysis and recombination analysis
To detect the genetic structures of microsatellite loci, Bayesian clustering was used in STRUCTURE 2.1 [75]. Because of the subtle population structure, correlated allele frequencies between populations and the admixture model were chosen [76]. The clusters (K) tested range was set from 1 to 10, and 10 independent runs were executed for each analysis. The Markov Chain Monte Carlo (MCMC) iterations lengths and burn-ins were set at 5,000,000 and 200,000, respectively. For most situations, ΔK had a mode at true K; thus, true K was selected using the ΔK statistic, which was calculated according to the rate of the log probability change of the data between successive K values [77].

The relationship of the MHC allele phylogenesis was constructed using a maximum likelihood method in PhyML 3.0 [78]. Before phylogenetic analysis, the most appropriate evolution models of sequence were estimated based on Akaike Information Criterion (AIC) in MODELTEST 3.7 [79]. The K80 + G model was suggested to be the optimum model for MHC sequences with the gamma shape parameter α = 0.6532. The reliability of the obtained tree topology structure was carried out with 1,000 bootstrap replications. In addition, intraspecific phylogenetic structures were inferred using median-joining networks in NETWORK 4.600 [80]. To estimate the rate of population recombination ρ (ρ = $4N_e r$), the composite-likelihood method [81] in LDhat [82] was used. The ρ is calculated using the crossing over for an effective population size (N_e) and rate per generation (r), and is estimated without prior information [82]. Even for the sequences that evolved

under balancing selection in the existence of recombination events, the LDhat still works efficiently [83].

Detecting historical selection

Two methods were used to detect historical selection. First, we calculated the rates of non-synonymous and synonymous substitutions at all amino acid sites, ABS sites, and non-ABS in Mega 4 by the Nei-Gojobori method with the Jukes-Cantor correction [84] and 1,000 bootstrap replicates to obtain standard errors. The putative ABS and non-ABS locations were derived according to the structure of human MHC II [85]. Compare with the method that used HLA-DR1 [86] structure to derive ABS sites, Reche and Reinherz's research provided a better way to avoid the emergence of uncertain gaps while blasting DQ1 genes with DR1 genes, and was also adopted by other studies [87,88]. Historical selection evidence was obtained with Codeml in the Paml 4 package [89]. This procedure examined heterogeneity in ω ($\omega = d_N/d_S$) [90] among codons based on the maximum likelihood method, with positive selection indicated by $\omega = d_N/d_S > 1$. Six models (M0, M2a, M3, M7, M8) allowing for different selection intensity among sites were tested [91,92]. The sites revealed to be under selection were compared with ABS sites.

Detecting recent selection

The F_{ST} outlier analysis [22] was used to examine whether MHC genes had significantly different population differentiation measures compared with microsatellite loci. Loci outside the confidence intervals of neutral F_{ST} values were likely under selection. The F_{ST} was estimated for *DQA1*, *DQB1*, and 16 microsatellite loci using Lositan [22,93]. Simulations involving 100 demes were sampled, and H_E and F_{ST} were printed for each of the 100,000 simulations. The infinite allele model was used in simulations that were performed using Fdist2. The aforementioned F_{ST} outlier analysis was also carried out on each population respectively. For each population, simulation parameters were similar using the same parameter set.

Additional files

Additional file 1: Table S1. Microsatellite data of 64 samples.

Additional file 2: Table S2. MHC haplotype distributions in populations.

Additional file 3: Figure S1. Median-joining networks for MHC alleles of the snub-nosed monkeys. A) Network for *DQA1* alleles; B) Network for *DQB1* alleles. The circles represent alleles (SG population-red, QL population-yellow, SNJ population-blue), with the area proportional to the frequency of the alleles in the 3 populations.

Additional file 4: Table S3. Alignment of the deduced amino acid sequences of *Rhro*-DQA1 exon 2 sequences. Identical amino acids are

shown by points, * represent ABS site, and sites revealed to be under significant selection in PAML are shown by dash.

Additional file 5: Table S4. Alignment of the deduced amino acid sequences of *Rhro*-DQB1 exon 2 sequences. Identical amino acids are shown by points, * represent ABS site, and sites revealed to be under significant selection in PAML are shown by dash.

Additional file 6: Table S5. cDNA primer sets for *DQA* and *DQB* (primer sets with * were those that amplified successfully in cDNA). Their location within the gene is shown in Additional file 7: Figure S2.

Additional file 7: Figure S2. Schematic representation of the position of cDNA primer sets used in the study of MHC II variation.

Author details

[1]Key laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1-5 Beixhenxi Road, Chaoyang, Beijing 100101, China. [2]Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China. [3]Graduate School of the Chinese Academy of Sciences, Beijing 100049, China. [4]College of Nature Conservation, Beijing Forestry University, Beijing 100083, China.

References

1. Peacock MM, Smith AT: **The effects of habitat fragmentation on dispersal patterns, mating behavior and genetic variation in a pica (*Ochotona princeps*) metapopulation.** *Oecol* 1997, **112**:524–533.
2. Sommer S: **The importance of immune gene variability (MHC) in evolutionary ecology and conservation.** *Front Zool* 2005, **2**:16.
3. Allendorf FW, Luikart G: *Units of conservation. In: Conservation and the genetics of populations.* Oxford: Blackwell Publishing; 2007:380–420.
4. Frankham R: **Genetics and extinction.** *Biol Conserv* 2005, **126**:131–140.
5. Frankham R, Ballou JD, Briscoe DA: *Introduction to Conservation Genetics (2nd edn).* Cambridge: Cambridge University Press; 2010:1–618. ISBN ISBN 978-0-521-7027-1-3.
6. Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW: **Conservation genetics in transition to conservation genomics.** *Trends Genet* 2010, **26**:177–187.
7. Miller HC, Allendorf F, Daugherty CH: **Genetic diversity and differentiation at MHC genes in island populations of tuatara (*Sphenodon* spp.).** *Mol Ecol* 2010, **19**:3894–3908.
8. Morin PA, Luikart G, Wayne RK, the SNP workshop group: **SNPs in ecology, evolution and conservation.** *Trends Ecol Evo* 2004, **19**:208–216.
9. Spurgin LG, Richardson DS: **How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings.** *Proc R Soc B* 2010, **277**:979–988.
10. Garrigan D, Hedrick PW: **Perspective: detecting adaptive molecular polymorphism: lessons from the MHC.** *Evolution* 2003, **57**:1707–1722.
11. Andersson L, Mikko S: **Generation of MHC class II diversity by intra- and intergenic recombination.** *Immunol Rev* 1995, **143**:5–12.

12. Bergstrom TF, Josefsson A, Erlich HA, Gyllensten U: **Recent origin of HLA-DRB1 alleles and implications for human evolution**. *Nat Genet* 1998, **18**:237–242.

13. Van Oosterhout C, Joyce DA, Cummings SM, Blais J, Barson NJ, Ramnarine IW, Mohammed RS, Persad N, Cable J: **Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*)**. *Evolution* 2006, **60**:2562–2574.

14. Sutton JT, Nakagawa S, Robertson BC, Jamieson IG: **Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes**. *Mol Ecol* 2011, **20**:4408–4420.

15. Miller KM, Kaukinen KH, Beacham TD, Withler RE: **Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon**. *Genetica* 2001, **111**:237–257.

16. Hedrick PW, Lee RN, Buchanan C: **Canine parvovirus enteritis, canine distemper, and major histocompatibility complex genetic variation in Mexican wolves**. *J Wildlife Dis* 2003, **39**:909–913.

17. Castro-Prieto A, Wachter B, Melzheimer J, Thalwitzer S, Sommer S: **Diversity and evolutionary patterns of immune genes in free-ranging Namibian leopards (*Panthera pardus pardus*)**. *Heredity* 2011, **102**:653–665.

18. Mikko S, Roed K, Schmutz S, Andersson L: **Monomorphism and polymorphism at MHC DRB loci in domestic and wild ruminants**. *Immunol Rev* 1999, **167**:169–178.

19. Weber DS, Stewart BS, Schienman J, Lehman N: **Major histocompatibility complex variation at three class II loci in the northern elephant seal**. *Mol Ecol* 2004, **13**:711–718.

20. Babik W, Pabijan M, Arntzen JW, Cogalniceanu D, Durka W, Radwan J: **Long-term survival of a urodele amphibian despite depleted major histocompatibility complex variation**. *Mol Ecol* 2009, **18**:769–781.

21. Mason RAB, Browning TL, Eldridge DB: **Reduced MHC class II diversity in island compared to mainland populations of the black-footed rock-wallaby (*Petrogale lateralis lateralis*)**. *Conserv Genet* 2011, **12**:91–103.

22. Biedrzycka A, Radwan J: **Population fragmentation and major histocompatibility complex variation in the spotted suslik, *Spermophilus suslicus***. *Mol Ecol* 2008, **17**:4801–4811.

23. Beaumont MA, Nichols RA: **Evaluating loci for use in the genetic analysis of population structure**. *Proc R Soc B* 1996, **263**:1619–1626.

24. Radwan J, Biedrzycka A, Babik W: **Does reduced MHC diversity decrease viability of vertebrate populations?** *Biol Conserv* 2010, **143**:537–544.

25. Aguilar A, Garza JC: **A comparison of variability and population structure for major histocompatibiltiy complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum)**. *Mol Ecol* 2006, **15**:923–937.

26. Zhang RZ: In *The primates of China: biogeography and conservation status—past, present and future*. Edited by Department of Wildlife Conservation, State Forestry Administration, the People's Republic of China. Beijing: China Forestry Publishing House; 2002.

27. Li M, Liu ZJ, Gou JX, Ren BP, Pan RL, Su YJ, Funk SM, Wei FW: **Phylogeography and population structure of the golden monkeys (*Rhinopithecus roxellana*): inferred from mitochondrial DNA sequences**. *Am J Primatol* 2007, **69**:1195–1209.

28. Pan D, Hu HX, Meng SJ, Men ZM, Fu YX, Zhang YP: **A high polymorphism level in *Rhinopithecus roxellana***. *Int J Primatol* 2009, **30**:337–351.

29. Chang ZF, Luo MF, Liu ZJ, Yang JY, Xiang ZF, Li M, Vigilant L: **Human influence on the population decline and loss of genetic diversity in a small and isolated population of Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*)**. *Genetica* 2012, **140**:105–114.

30. Boessenkool S, Taylor SS, Tepolt CK, Komdeur J, Jamieson IG: **Large mainland populations of South Island robins retain greater genetic diversity than offshore island refuges**. *Conserv Genet* 2007, **8**:705–714.

31. Hedrick PW, Parker KM, Miller EL, Miller PS: **Major histocompatiblity complex variation in the endangered Przewalski's horse**. *Genetics* 2001, **152**:1701–1710.

32. Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK, Watkins DI: **Nomenclature for the major histocompatibility complexes of different species: A proposal**. *Immunogenetics* 1990, **31**:217–219.

33. Chang ZF, Liu ZJ, Yang JY, Li M, Vigilant L: **Noninvasive genetic assessment of the population trend and sex ratio of the Shennongjia population of Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*)**. *Chinese Sci Bull* 2011, **57**:1135-1141.

34. Luo MF, Liu ZJ, Pan HJ, Zhao L, Li M: **Historical geographic dispersal of the golden snub-nosed monkey (*Rhinopithecus roxellana*) and the influence of climatic oscillations**. *Am J Primatol* 2012, **74**:91–101.

35. Bijlsma R, Bundgaard J, Boerema AC: **Does inbreeding affect the extinction risk of small populations? Predictions from *Drosophila***. *J Evolution Biol* 2000, **13**:502–514.

36. Thoß M, Ilmonen P, Musolf K, Penn DJ: **Major histocompatibility complex heterozygosity enhances reproductive success**. *Mol Ecol* 2011, **20**:1546–1557.

37. Lukas D, Bradley BJ, Nsubuga AM, Doran-Sheehy D, Robbins MM, Vigilant L: **Major histocompatibility complex and microsatellite variation in two populations of wild gorillas**. *Mol Ecol* 2004, **13**:3389–3402.

38. Goda1 N, Mano T, Kosintsev P, Vorobie A, Masuda R: **Allelic diversity of the MHC class II DRB genes in brown bears (*Ursus arctos*) and a comparison of DRB sequences within the family Ursidae**. *Tissue Antigens* 2010, **76**:404–410.

39. Kennedy LJ, Randall DA, Knobel D, Brown JJ, Fooks AR, Argaw K, Shiferaw F, Ollier WER, Sillero-Zubiri C, Macdonald DW, Laurenson MK: **Major histocompatibility complex diversity in the endangered Ethiopian wolf (*Canis simensis*)**. *Tissue Antigens* 2010, **77**:118–125.

40. Radwan J, Demiaszkiewicz AW, Kowalczyk R, Lachowicz J, Kawalko A, Wojcik JM, Pyziel AM, Babik W: **An evaluation of two potential risk factors, MHC diversity and host density, for infection by an invasive nematode *Ashworthius sidemi* in endangered European bison (*Bison bonasus*)**. *Biol Conserv* 2010, **143**:2049–2053.

41. Pokorny I, Sharma R, Goyal SP, Mishra S, Tiedemann R: **MHC class I and MHC class II DRB gene variability in wild and captive Bengal tigers (*Panthera tigris tigris*)**. *Immunogenetics* 2010, **62**:667–679.

42. Kimura M: **Preponderance of synonymous changes as evidence for the neutrality theory of molecular evolution**. *Nature* 1977, **267**:275–276.

43. Hughes AL, Nei M: **Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection**. *Nature* 1988, **335**:167–170.

44. Hughes AL, Nei M: **Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection**. *Proc Natl Acad Sci USA* 1989, **86**:958–962.

45. Hughes AL: **Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level**. *Heredity* 2007, **99**:364–373.

46. 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.

47. Bryja J, Charbonnel N, Berthier K, Galan M, Cosson JF: **Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles**. *Mol Ecol* 2007, **16**:5084–5097.

48. Li HS, Ligons DL, Rose NR: **Genetic complexity of autoimmune myocarditis**. *Autoimm reviews* 2008, **7**:168–173.

49. Schaschl H, Suchentruk F, Hammer S, Goodman SJ: **Recombination and the origin of sequence diversity in the DRB MHC class II locus in chamois (*Rupicapra* spp.)**. *Immunogenetics* 2005, **57**:108–115.

50. Xu SX, Ren WH, Zhou XM, Zhou KY, Yang G: **Sequence polymorphism and geographical variation at a positively selected MHC-DRB gene in the finless porpoise (*Neophocaena phocaenoides*): Implication for recent differentiation of the Yangtze finless porpoise?** *J Mol Ecol* 2010, **71**:6–22.

51. Kauppi L, Jeffreys AJ, Keeney S: **Where the crossovers are: recombination distributions in mammals**. *Nat Rev Genet* 2004, **5**:413–424.

52. Bonneaud C, Chastel O, Federici P, Westerdahl, Sorci G: **Complex Mhc-based mate choice in a wild passerine**. *Proc R Soc B* 2006, **273**:1111–1116.

53. Muirhead CA: **Consequences of population structure on genes under balancing selection**. *Evolution* 2001, **55**:1532–1541.

54. Richman AD, Herrera LG, Nash D: **Evolution of MHC class II E beta diversity within the genus Peromysus**. *Genetics* 2003, **164**:289–297.

55. Miller HC, Lambert DM: **Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae)**. *Mol Ecol* 2004, **13**:3709–3721.

56. Kimura M: *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press; 1983.

57. Maureen BP, Thomas FT: **Genetic variation of the major histocompatibility complex (MHC class II b gene) in the threatened Gila trout, Oncorhynchus gilae gilae**. *Conserv Genet* 2008, **9**:257–270.

58. Agudo R, Alcaide M, Rico C, Lemus JA, Blanco G, Hiraldo F, Donázar JA: **Major histocompatibility complex varian in insular populations of the Egyptian vulture: inferences about the roles of genetic drift and selection.** *Mol Ecol* 2011, **20**:2329–2340.
59. Oliver MK, Lambin X, Cornulier T, Piertney SB: **Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations.** *Mol Ecol* 2009, **18**:80–92.
60. Ujvari B, Belov K: **Major histocompatibility complex (MHC) markers in conservation biology.** *Int J Mol Sci* 2011, **12**:5168–5186.
61. Tallmon DA, Luikart G, Waples RS: **The alluring simplicity and complex reality of genetic rescue.** *Trends Ecol Evol* 2004, **19**:489–496.
62. Edmands S: **Between a rock and a hard place: Evaluating the relative risks of inbreeding and outbreeding for conservation and management.** *Mol Ecol* 2007, **16**:463–475.
63. Vilà C, Sundqvist AK, Flagstad O, Seddon J, Björnerfeldt S, Kojola I, Casulli A, Sand H, Wabakken P, Ellegren H: **Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant.** *Proc R Soc Lond B* 2003, **270**:91–97.
64. Madsen T, Ujvari B, Olsson M: **Novel genes continue to enhance population growth in adders (*Vipera berus*).** *Biol Conserv* 2004, **120**:145–147.
65. Bouzat JL, Johnson JE, Toepfer JE, Simpson SA, Esker TL, Westemeier RL: **Beyond the beneficial effects of translocations in an effective tool for the genetic restoration of isolated populations.** *Conserv Genet* 2009, **10**:191–201.
66. Doxiadis GGM, Otting N, de Groot NG, Groot N, Rouweler AJM, Noort R, Verschoor EJ, Bontjer I, Bontrop RE: **Evolutionary stability of MHC class II haplotypes in diverse rhesus macaque populations.** *Immunogenetics* 2003, **55**:540–551.
67. Nei M: **Estimation of average heterozygosity and genetic distance from a small number of individuals.** *Genetics* 1978, **89**:583–590.
68. Excoffier L, Schneider S: **Arlequin ver. 3.0: An integrated software package for population genetics data analysis.** *Evol Bioinform* 2005, **1**:47–50.
69. Goudet J: *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)*. 2001. Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).
70. Rousset F: **Genepop'007: a complete re-implementation of the genepop software for Windows and Linux.** *Mol Ecol Res* 2008, **8**:103–106.
71. Weir BS: **Inferences about linkage disequilibrium.** *Biometrics* 1979, **35**:235–254.
72. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P: **MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data.** *Mol Ecol Notes* 2004, **4**:535–538.
73. Tamura K, Dudley J, Nei M, Kumar S: **MEGA 4: Molecular evolutionary genetics analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**:1596–1599.
74. Kennedy LJ, Ryvar R, Gaskell RM, Addie D, Willoughby K, Carter S, Thomson W, Ollier W, Radford A: **Sequence analysis of MHC DRB alleles in domestic cats from the United Kingdom.** *Immunogenetics* 2002, **54**:348–352.
75. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155**:945–959.
76. Falush D, Stephens M, Pritchard J: **Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies.** *Genetics* 2003, **164**:1567–1587.
77. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study.** *Mol Ecol* 2005, **14**:2611–2620.
78. Guindon S, Gascuel O: **A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003, **52**:696–704.
79. Posada D, Crandall KA: **Modeltest: Testing the model of DNA substitution.** *Bioinformatics* 1998, **14**:817–818.
80. Bandelt HJ, Forster P, Röhl A: **Median-joining networks for inferring intraspecific phylogenies.** *Mol Biol Evol* 1999, **16**:37–48.
81. Hudson RR: **Two-locus sampling distributions and their application.** *Genetics* 2001, **159**:1805–1817.
82. McVean G, Awadalla P, Fearnhead P: **A coalescennt-based method for detecting and estimating recombination from gene sequences.** *Genetics* 2002, **160**:1231–1241.
83. Richman AD, Herrera LG, Nash D, Schierup MH: **Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in *Peromyscus maniculatus*.** *Genet Res* 2003, **82**:89–99.
84. Nei M, Gojobori T: **Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions.** *Mol Biol Evol* 1986, **3**:418–426.
85. Reche PA, Reinherz EL: **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.
86. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC: **3-Dimensional structure of the human class-II histocompatibility antigen HLA-DR1.** *Nature* 1993, **364**:33–39.
87. Kuduk K, Johanet A, Allaine D, Cohas A, Radwan J: **Contrasting patterns of selection acting on MHC class I and class II DRB genes in the Alpine marmot (*Marmota marmota*).** *J Evolutiona Biol* 2012, doi:10.1111/j.1420-9101.2012.02537.x.
88. Kloch A, Baran K, Buczek M, Konarzewski M, Radwan J: **MHC influences infection with parasites and wintersurvival in the root vole *Microtus oeconomus*.** *Evol Ecol* 2012. doi:10.1007/s10682-012-9611-1.
89. Yang Z: **PAML 4: phylogenetic analysis by maximum likelihood.** *Mol Biol Evol* 2007, **24**:1586–1591.
90. Goldman N, Yang Z: **A codon-based model of nucleotide substitution for protein-coding DNA sequences.** *Mol Biol Evol* 1994, **11**:725–736.
91. Yang Z, Nielsen R, Goldman N, Pedersen AMK: **Codon substitution models for heterogeneous selection pressure at amino acid sites.** *Genetics* 2000, **155**:431–449.
92. Yang Z, Wong WSW, Nielsen R: **Bayes Empirical Bayes inference of amino acid sites under positive selection.** *Mol Biol Evol* 2005, **22**:1107–1111.
93. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G: **LOSITAN: a workbench to detect molecular adaptation based on a *Fst*-outlier method.** *BMC Bioinformatics* 2008, **9**:323.