scientific reports

OPEN

Diversity and evolution of the MHC class II *DRB* gene in the *Capra sibirica* experienced a demographic fluctuation in China

Pei-Pei Dong[1,3], Rui-Rui Wang[1,3] & Shamshidin Abduriyim[1,2✉]

The major histocompatibility complex (MHC) genes are the most polymorphic genes in vertebrates, and their proteins play a critical role in adaptive immunity for defense against a variety of pathogens. MHC diversity was lost in many species after experiencing a decline in size. To understand the variation and evolution of MHC genes in the Siberian ibex, *Capra sibirica*, which has undergone a population decline, we analyzed the variation of the second exon of MHC class II *DRB* genes in samples collected from five geographic localities in Xinjiang, China, that belong to three diverged mitochondrial clades. Consequently, we identified a total of 26 putative functional alleles (PFAs) with 260 bp in length from 43 individuals, and found one (for 27 individuals) to three (for 5 individuals) PFAs per individual, indicating the presence of one or two *DRB* loci per haploid genome. The *Casi-DRB1*16* was the most frequently occurring PFA, *Casi-DRB1*22* was found in only seven individuals, 14 PFAs occurred once, 7 PFAs twice, implying high frequency of rare PFAs. Interestingly, more than half (15) of the PFAs were specific to clade I, only two and three PFAs were specific to clades II and III, respectively. So, we assume that the polygamy and sexual segregation nature of this species likely contributed to the allelic diversity of *DRB* genes. Genetic diversity indices showed that PFAs of clade II were lower in nucleotide, amino acid, and supertype diversity compared to those of the other two clades. The pattern of allele sharing and F_{ST} values between the three clades was to some extent in agreement with the pattern observed in mitochondrial DNA divergence. In addition, recombination analyses revealed no evidence for significant signatures of recombination events. Alleles shared by clades III and the other two clades diverged 6 million years ago, and systematic neighbor grids showed Trans-species polymorphism. Together with the PAML and MEME analyses, the results indicated that the *DRB* gene in *C. sibirica* evolved under balancing and positive selection. However, by comparison, it can be clearly seen that different populations were under different selective pressures. Our results are valuable in understanding the diversity and evolution of the *DRB* gene in a mountain living *C. sibirica* and in making decisions on future long-term protection strategies.

The way a species interacts with other species or the environment is related to its genetic diversity. The ability of a species to adapt to human and natural environmental disturbances also depends on variation. The degree of genetic diversity within a species influences the ability of that species to adapt to environmental change[1]. The measurement of these variation, however, cannot be limited to the use of traditional neutral markers. The polymorphisms of Major Histocompatibility Complex (MHC) genes are considered to affect the functional plasticity of immune responses to diverse pathogenic stressors, making them excellent candidates to research adaptive evolutionary processes in natural populations[2,3]. This trait highlights the sensitivity of the immune system to environmental stresses and its importance in elucidating the mechanisms of adaptive genetic variation required for the long-term survival of a species or population[4,5].

The most polymorphic region of the vertebrate genome that evolved under positive and balancing selection is the MHC[5–7], a multigene family of the vertebrate adaptive immune system that contains highly polymorphic motifs that are strongly associated with immune response and disease resistance[8,9]. MHC genes belong to two

[1]College of Life Science, Shihezi University, Shihezi 832003, Xinjiang, China. [2]Xinjiang Production and Construction Corps Key Laboratory of Oasis Town and Mountain-Basin System Ecology, Shihezi University, Shihezi 832003, Xinjiang, China. [3]These authors contributed equally: Pei-Pei Dong and Rui-Rui Wang. ✉email: shamshidin@shzu.edu.cn

main subfamilies, class I and class II, and encode proteins that are necessary for pathogen recognition and presentation to T cells. Functional class II proteins are heterodimers consisting of α and β chains, and DR subclasses are encoded by the *DRA* and *DRB* genes, respectively[10]. The amino acid residues that bind directly to the antigen are called antigen binding sites (ABS). ABSs were located at the α1 domain of α1 chain and the β1 domain of β chain, in which MHC polymorphisms in vertebrates mainly occur[11,12].

Alpine ungulates play a significant role in maintenance the structure of vegetation and the cycling of nutrients in high-mountain ecosystems, as well as a significant source of food for predators[13,14]. However, due to its slow growth, poor rate of reproduction, vulnerability to exploitation by humans, loss of habitat, susceptibility to infectious diseases, and other reasons[15–17], most of them are extremely vulnerable to extinction. Of these, a typical alpine hoofed species of the subfamily Caprinae (family Bovidae), is the Siberian ibex, *Capra sibirica*. This species widely habituated in the alpine regions of Central Asia, from northern India through Pakistan and Afghanistan to Russia (Siberia), and eastward to northwest China and western Mongolia[18]. According to studies, the Siberian ibex, to some extent, suffered threats from various pathogens (lethal bacteria and viruses), endoparasites (helminths) and ectoparasites (mites)[19–23]. Moreover, it shares more than 76 percent of its food with domestic animals in Chinese territory[24], indicating not only fierce food competition but also a greatly increased risk of becoming infected. Despite the importance of MHC genes for immunological fitness, an assessment of the diversity and occurrence of these genes is still lacking in the Siberian ibex, the globally 'Near Threatened' mammal in Central Asia, and locally urgently needs an effective conservation and management programs[18].

Because of anthropogenic impacts, Siberian ibex populations dropped globally and their range shrank drastically in the 1970s[25,26]. In China, particularly, it had been listed as an endangered species and given Class I protection priority in 1998[27]. Since then, Chinese scholars have started to pay attention and carry out field studies on population size and density estimation in several restricted regions in different years. For example, 593 individuals were estimated in Tashkurgan county in 2009[28], 5604 individuals in Bay county in 2010[29]; the population density was estimated to be 2.69 individuals/km² in Tomur National Natural Reserve in 2005[30], 1.27 individuals/km² in Kudi village in Kagilik county in 2010[31] and 0–0.25 individuals/km² in Tashkurgan county in 2012[32]. Later, to the best of our knowledge, there was no single report of population size and density publicly available. However, it was suggested that the population has fortunately started to recover owing to effective conservation and management (through the creation of protected areas, etc.) by the Chinese government in recent years, and thus its protection priority was decreased to Class II in 2021[33]. Generally, reduced genetic diversity is associated with demographic perturbation. Natural populations of many species that underwent a reduction in size exhibited very limited MHC diversity[34–36]. However, both theoretical and empirical studies also showed that a longer timescale of selection maintained higher MHC diversity in a population experienced demographic fluctuations[37,38]. It is thus significant to study the MHC diversity of a highly genetically diverged Siberian ibex populations[39] during a more than half-century period of recovery.

Therefore, our objectives in this study had three facets. To begin with, we aim to comparatively evaluate the MHC diversity in different Siberian ibex populations in Xinjiang, China, and discuss our results with other species that have experienced bottlenecks. In addition, we also try to ascertain if the MHC *DRB1* divergence in different populations was in accordance with the results of mitochondrial genes divergence we reported previously[39]. Finally, to check if the MHC *DRB1* genes in the Siberian ibex that went through population fluctuation resemble the common characteristics of MHC in other vertebrates, such as positive selection, recombination, and transspecies polymorphism, and to clarify the genetic relationships of MHC *DRB1* alleles of the relic species Siberian ibex and its congenerics, including domestic goats. Our results were of importance in understanding the adaptive ability of this species and planning scientific conservation strategies to ensure long-term population development.

Materials and methods
Samples
A total of 43 samples, including 33 feces, 5 muscle, 4 skin, and 1 liver sample, were analysed. Of these, 10 samples collected from Urumqi, 16 from Arturk, 1 from Sawan, 13 from Ulugqat, and 3 from Kagilik (Fig. 1). All samples in this study were came from samples of our previous study[39]. Tissue samples either taken from individuals died of natural causes or dead individuals that were poached. Individual identity of fecal samples was established according to Abduriyim et al.[40,41]. All fecal and tissue samples were preserved in 96% ethanol, and skin samples were directly frozen in plastic bags at −80 °C until use.

Experimental procedures
The total genomic DNA of fecal samples was extracted using an Omega stool DNA extraction kit (Omega Bio-tec, Georgia, USA), and that of muscle and skin samples was extracted using a Tiangen tissue/blood DNA extraction kit (Tiangen Bio-tec, Beijing, China), following the manufacturer's instructions. After electrophoresis detection, DNA concentration and purity were measured by the Thermo Nanodrop 1000 and stored at 4 °C for later use.

Part of MHC class II *DRB1* exon 2 (260 bp, excluding the primer sequences) was PCR amplified using primer pairs of CapDRB1.1F and CapDRB1.2R[42], because this segment is the most polymorphic region and includes all ABSs necessary for pathogen recognition[12,43]. A PCR reaction volume contained 40–150 ng of DNA, 5 pmol of each primer, 12.5 μL of Tiangen's 2 TaqPCR Master Mix, and then adjusted to a final volume of 25 μL with RNase-Free double distilled water. The PCR thermal cycling conditions were as follows: pre-denaturation at 94 °C for 4 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 64 °C for 30 s, extension at 72 °C for 45 s, and final extension at 72 °C for 5 min. The PCR products were verified by 2% agarose gel electrophoresis and green fluorescence dye imaging under ultraviolet irradiation, and those with the expected band size were sent for Sanger sequencing bidirectionally using both the forward and reverse primers (Qingke Biology, Xi-an, China).
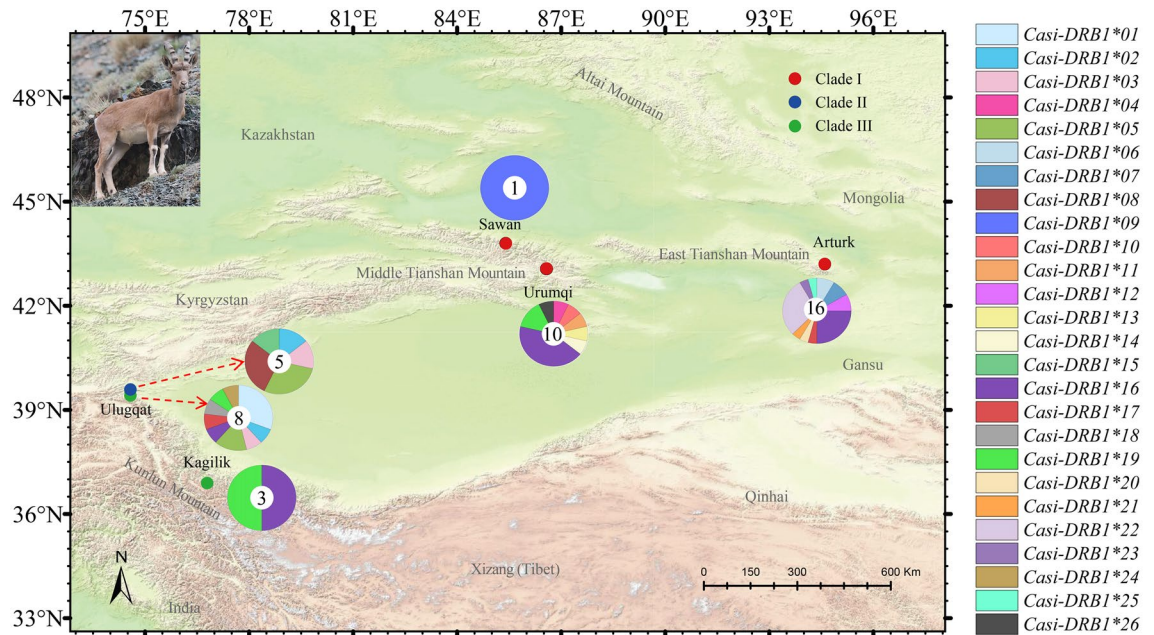
Figure 1. Sampling locations for *Capra sibirica* in Xinjiang, China, in this study. Each small circle on the map indicates sampling locality and different colors signify different clades determined based on the mtDNA analyses[39]. The pie chart shows alleles frequencies (number of alleles) for geographic populations/clades, with each allele in a different color (key at the right). This figure was produced using ArcGIS Pro 2.5.0 (https://www.esri.com/).

PCR products assumed to contain more than one sequences were proceeded to cloning and sequencing for allele isolation. PCR products were recovered using a gel recovery kit (Tiangen, Beijing, China), connected to the PMD™19-T plasmid vector (Takara, Tokyo, Japan), then transformed into *Escherichia coli* (DH5α) receptive cells. For selection of cells with positive plasmids, the bacteria were grown on LB solid medium containing ampicillin, IPTG and X-gal at 37 °C overnight. Bacteria containing plasmids with the target PCR product were screened by blue/white selection and direct-colony PCR amplification using M13 forward and reverse primers with the same PCR condition as described earlier. At least 8 clones per sample were bidirectionally sequenced for each individual.

MHC genotyping

All nucleotide sequences obtained were aligned using MEGA v.6.0[44]. The unique and same sequences were screened using DnaSP v.5.10.01[45]. The final sequences were identified as potentially genuine *DRB1* exon 2 sequences if they matched in the forward and reverse directions, and were detected at least twice in one individual (two independent PCR reactions for one individual) or once each from at least two individuals[46]. Single, unique sequences were omitted, as they may have been PCR chimeras or due to other PCR errors[5,43]. We verified candidate sequences with BLAST searches[47] at the National Center for Biotechnology Information (NCBI) GenBank database. Final verified sequences were named by consulting the conventions of Klein et al.[48], and Ballingall and Todd[49].

Data analyses

The nucleotide, amino acid, super type diversity, and pairwise population fixation indices (F_{ST}) of *DRB1* exon 2 for different populations were calculated by DnaSP v.5.10.01[45], and the neutral selection was analyzed. MEGA v.6.0 were used to estimate the ratio ω (d_N/d_S) of non-synonymous (d_N) to synonymous (d_S) substitution rates[50]; this ratio provides a measure of selective pressure at the level of individual sites[51]. Values of $\omega > 1$ indicate positive selection, while $\omega = 1$ and $\omega < 1$ indicate neutral evolution and purifying selection, respectively. Values of d_N, d_S and ω were calculated separately for presumed ABSs deduced according to Reche and Reinherz[52], non-ABSs, and all sites. HyPhy[53] implemented in MEGA was used to detect signs of positive selection. In order to examine positive selection across all sites based on maximum likelihood methods, CodeML in PAML 4.9[54] was employed as well. The likelihood ratio tests (LRTs) were used to compare the four models: M1a, almost neutral; M2a, positive selection; M7, beta; and M8, beta and ω, and decide which model best fit our data[51,55,56]. Using LRTs, two nested models (M1a vs. M2a; M7 vs. M8) were compared. Using Bayes Empirical Bayes inference[57], positively selected locations were found. In addition, using Datamonkey v.2.0[58], a web-based server for the HyPhy Package, a mixed-effects model of evolution, MEME[59] analysis was carried out to find codons that had been subject to positive selection.

Gene recombination analysis of *DRB1* exon 2 sequences was performed in RDP4[60]. Specific methods were first used in RDP[61], GENECONV[62], MaxChi[63] and Bootscan[64], which use default Settings to detect recombination events using Bonferroni correction for multiple comparisons. Recombination events detected by at least three

of these methods were then rechecked using all RDP methods available[61]. In addition, we also use the GARD[65], provided by the Datamonkey webserver[66], to detect the signatures of recombination breakpoints. In order to avoid the impact of possible gene replication, conversion, and recombination on phylogenetic analysis, we chose Splitstree4 v.4.14.5 to construct a neighbor network of *DRB1* sequences[5,67].

Results
Diversity of *DRB1* alleles

Our analytical sequences were 260 bp in length, encode 86 amino acids including 20 ABSs, accounting for 91% of the *DRB1* β1 domain (Fig. 2). We identified 26 presumably functional alleles (PFA) in a total of 43 individuals belong to three mtDNA clades[39]. None of these sequences were pseudogenes. The number of PFA found in a
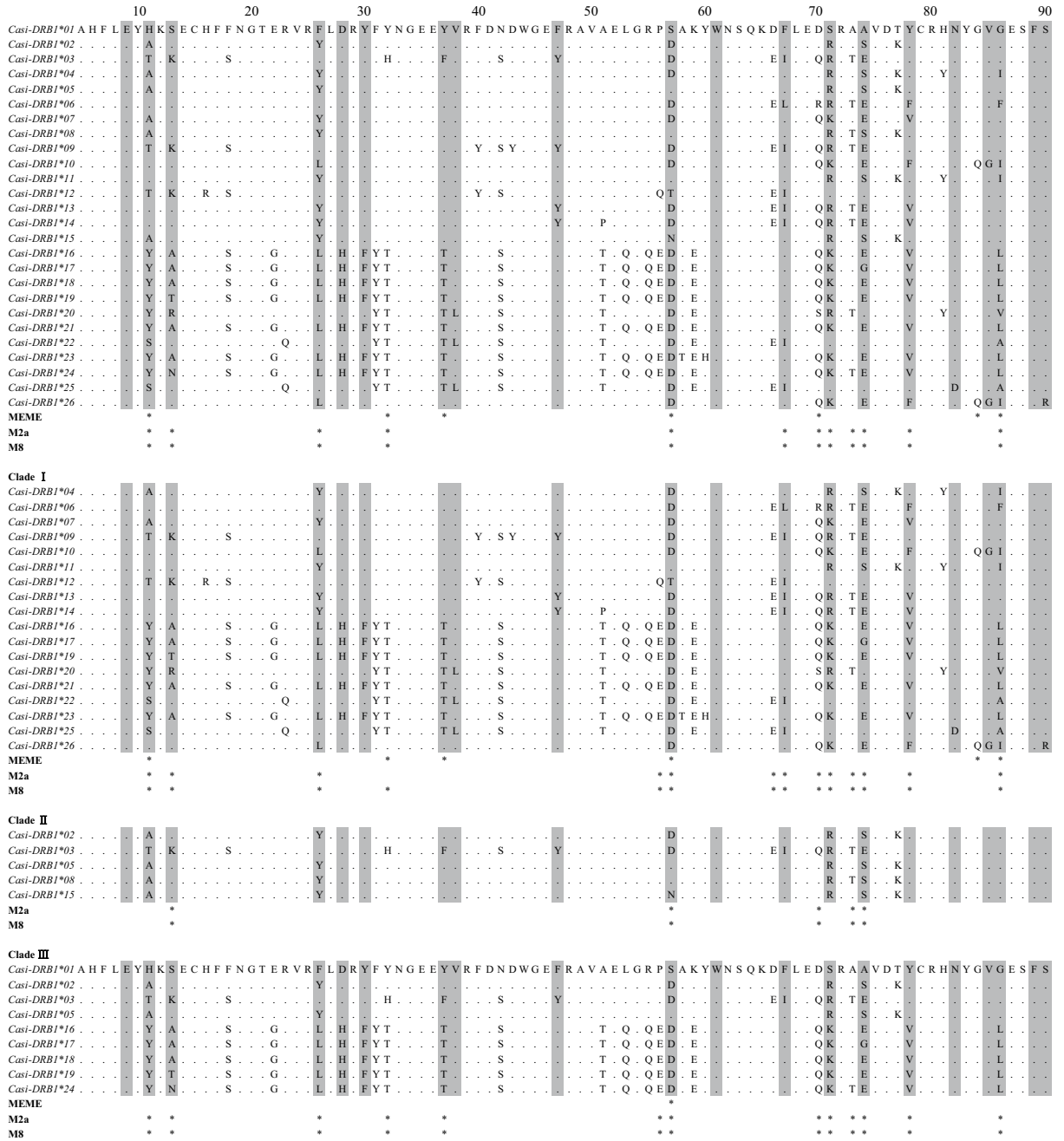

Figure 2. Alignment of deduced amino acid sequences encoded by exon 2 of MHC class II *DRB1* alleles in *Capra sibirica*. Numbers above the amino acid sequences indicate positions in the β1-domain of the DR protein β-chain. Dots indicate amino acids identical to those in *Casi-DRB1*01*. Putative ABSs as determined by Reche & Reinherz[52] are shaded. * signs at the bottom of the table indicate sites inferred to be under positive selection by MEME analysis and Bayes Empirical Bayes inference (BEB) using PAML. For the M2a and M8 models in BEB, only significant results are indicated by + (*P* > 95%). Clade II did not find positive selection sites by MEME analysis and is therefore not shown.

single individual ranged from 1 to 3, indicating existence of 1 or 2 loci per haploid genome in the Siberian ibex. 26 individuals possess only one PFA, implying that these individuals were homozygous. 12 individuals with two PFAs were highly likely heterozygous individuals. Only five individuals had three PFLs (Table A1). Summarizing, most of the individuals in Siberian ibex Xinjiang populations had one *DRB1* locus and could be homozygous.

The most common PFA was *Casi-DRB1*16*, occurred in 17 individuals that belong to Clade I (12 individuals) and Clade III (5 individuals). The PFAs *Casi-DRB1*22* and *19* came after, but we found only in seven and five individuals, respectively. *Casi-DRB1*16* and *17* were shared by Clades I and III, *Casi-DRB1*02*, *03*, and *05* by Clades II and III. Besides, *Casi-DRB1*08* and *15* were specific to Clade II, while *Casi-DRB1*01*, *18*, and *24* were only found in Clade III, and the remaining ones were exclusively occupied by Clade I (Fig. 1, Table A1).

Diversity indices showed a high level of genetic diversity at the nucleotide, amino acid and supertype sequence level for all clades. By comparison, the nucleotide diversity of individuals from Clade I and Clade III was similar, and both were higher than that from CladeII. In terms of amino acid and super type, Clade II had an overall lower level of diversity than Clade I and Clade III as well (Table 1). Tajima's *D* values were positive, except for Clade II, though none of these values were significant (Table 1). Altogether, this indicates that the MHC class II *DRB1* locus in different clades was likely at the different the stage of bottleneck or selection pressures.

We calculated the genetic differentiation by F_{ST} values, both including and excluding shared alleles between the three Clades. The F_{ST} value between clades I and III was negative (Table 2), indicating that within-population genetic differentiation was higher than between-population genetic differentiation. After excluding the shared alleles, the value was positive but very low, implying that the shared alleles were more divergent than the unique alleles in these clades. When including shared alleles, the F_{ST} values between clades I and II, II and III were 0.267 and 0.343, respectively (Table 2), indicating a very high level of differentiation.

Recombination and selection on *DRB1*

It was hardly evident that significant recombination signatures exist in our analyses of the *DRB1* exon 2 sequences of the Siberian ibex. Hence, we used all sequences in the downstream analyses. To evaluate selection pressure, we calculated the ω ratio of non-synonymous to synonymous substitution rates for positions in the presumed ABSs, non-ABS codons and all codons for three Clades. The ω ratio value for ABS codons in *C. sibirica DRB1* was greater than one. Our result thus indicates that variation at the ABS codons were generated and maintained by positive selection. Comparatively, ω value for Clade II was nearly twice of that for clades I and II, indicating that the selection intensity on these clades was different (Table 3). This was in line with the results of PAML and MEME analyses that provide evidence for positive selections at the single codon level (Table 4). Likelihood ratio tests (LRT) showed that the M2a and M8 models provided significantly better fits to our data than models without selection (Table 4). The M2a and M8 models identified 13 and 14 positively selected sites in Clade I, respectively, 13 from each model in Clade III, while only five from each model in Clade II (Fig. 2, Table 4), with most of the sites occurring in presumed ABS codons. Finally, the MEME analysis showed six codons under positive selection in Clade I, only one in Clade III, and none in Clade II (Fig. 2).

Phylogeny of *DRB1* alleles

In the phylogenetic neighbor grid, the *DRB1* exon 2 sequences of *C. sibirica* did not form separate groups according to the geographical population results, but merged into sequences of other species of genus *Capra*, forming groups of A, B, C, D, E, F, and G (Fig. 3). Trans-species polymorphism (TSP) was clearly evident, with some exon

Clade	Sample Size	Number	Diversity indices			
			π	aa	Supertype	Tajima's *D*
I	27	18/16/14	0.092	0.190	0.427	0.184
II	5	5/5/4	0.035	0.078	0.228	−0.793
III	11	9/8/8	0.084	0.177	0.402	0.705
Total	43	26/23/20	0.089	0.179	0.406	0.277

Table 1. Genetic diversity at nucleotide, amino acid and supertype levels and neutrality test on the MHC *DRB1* gene for three clades of *Capra sibirica* in Xinjiang, China. The Number column shows, in order, the number of alleles, the number of amino acids, and the number of supertypes.

Clade	I	II	III
I		0.364	0.030
II	0.267		0.560
III	−0.013	0.343	

Table 2. Genetic differentiation (F_{ST}) of the MHC *DRB1* gene among the three clades of *Capra sibirica* in Xinjiang, China. The F_{ST} Values were calculated including (below the diagonal) and excluding (above the diagonal) shared alleles between clades.

Substitution type		Number of codons	Clade I	Clade II	Clade III
d_N	ABS	20	1.203 ± 0.209	0.244 ± 0.149	0.690 ± 0.186
	Non-ABS	66	0.288 ± 0.140	0.051 ± 0.084	0.132 ± 0.096
	Overall	86	0.514 ± 0.127	0.101 ± 0.081	0.257 ± 0.100
d_S	ABS	20	0.658 ± 0.293	0.087 ± 0.288	0.392 ± 0.298
	Non-ABS	66	0.159 ± 0.170	0.015 ± 0.122	0.068 ± 0.208
	Overall	86	0.293 ± 0.158	0.037 ± 0.144	0.167 ± 0.180
ω	ABS	20	1.828	2.805	1.760
	Non-ABS	66	1.811	3.400	1.941
	Overall	86	1.754	2.730	1.539

Table 3. Rates ($\pm$ standard error) of non-synonymous (d_N) and synonymous (d_S) substitutions and their ratio (ω) for the presumed antigen binding sites (ABS), non-ABSs, and sites overall in the β1-domain of the MHC class II *DRB1* genes for *Capra sibirica* in Xinjiang, China.

Clade	Models	lnL	Parameter estimates	PSS	LRT	d.f.	P value
I	M1a	-1010.99	$P_0 = 0.860$, $P_1 = 0.140$, $\omega_0 = 0.041$, $\omega_1 = 1.000$		M1a vs M2a	2	< 0.01
	M2a	-972.69	$P_0 = 0.572$, $P_1 = 0.401$, $P_2 = 0.026$, $\omega_0 = 0.099$, $\omega_1 = 1.000$, $\omega_2 = 13.181$	11, 13, 26, 56, 57, 66, 67, 70, 71, 73, 74, 78, 86			
	M7	-1014.03	$P = 0.025$, $q = 0.155$		M7 vs M8	2	< 0.01
	M8	-972.72	$P_0 = 0.974$, $P = 0.107$, $q = 0.116$, $P_1 = 0.026$, $\omega = 13.366$	11, 13, 26, 32, 56, 57, 66, 67, 70, 71, 73, 74, 78, 86			
II	M1a	-445.45	$P_0 = 0.524$, $P_1 = 0.476$, $\omega_0 = 0.000$, $\omega_1 = 1.000$		M1a vs M2a	2	0.0103
	M2a	-440.88	$P_0 = 0.858$, $P_1 = 0.000$, $P_2 = 0.142$, $\omega_0 = 0.000$, $\omega_1 = 1.000$, $\omega_2 = 11.496$	13, 57, 70, 73, 74			
	M7	-446.65	$P = 1.970$, $q = 0.005$		M7 vs M8	2	< 0.01
	M8	-440.88	$P_0 = 0.859$, $P = 0.005$, $q = 2.990$, $P_1 = 0.142$, $\omega = 11.496$	13, 57, 70, 73, 74			
III	M1a	-640.12	$P_0 = 0.737$, $P_1 = 0.263$, $\omega_0 = 0.000$, $\omega_1 = 1.000$		M1a vs M2a	2	< 0.01
	M2a	-622.57	$P_0 = 0.963$, $P_1 = 0.000$, $P_2 = 0.037$, $\omega_0 = 0.546$, $\omega_1 = 1.000$, $\omega_2 = 18.992$	11, 13, 26, 32, 37, 56, 57, 70, 71, 73, 74, 78, 86			
	M7	-640.26	$P = 0.005$, $q = 0.012$		M7 vs M8	2	< 0.01
	M8	-622.47	$P_0 = 0.969$, $P = 0.008$, $q = 0.005$, $P_1 = 0.031$, $\omega = 20.626$	11, 13, 26, 32, 37, 56, 57, 70, 71, 73, 74, 78, 86			
All	M1a	-1132.10	$P_0 = 0.898$, $P_1 = 0.102$, $\omega_0 = 0.039$, $\omega_1 = 1.000$				
	M2a	-1081.75	$P_0 = 0.978$, $P_1 = 0.000$, $P_2 = 0.022$, $\omega_0 = 0.468$, $\omega_1 = 1.000$, $\omega_2 = 14.736$	11, 13, 26, 32, 57, 67, 70, 71, 73, 74, 78, 86	M1a vs M2a	2	< 0.01
	M7	-1134.53	$P = 0.016$, $q = 0.103$				
	M8	-1079.73	$P_0 = 0.979$, $P = 0.021$, $q = 0.026$, $P_1 = 0.021$, $\omega = 14.775$	11, 13, 26, 32, 57, 67, 70, 71, 73, 74, 78, 86	M7 vs M8	2	< 0.01

Table 4. The results of codon based positive selection analyses using maximum likelihood models in CodeML for MHC *DRB1* exon 2 sequences from *Capra sibirica*. Positively selected sites (PSS), log-likelihood (lnL), the likelihood ratio test (LRT) and probability (P) values were presented. ω equals d_N to d_S ratio; P_n is the proportion of amino acids in the ω_n site class; P and q are parameters of the beta distribution.

2 sequences from particular *C. sibirica* being more closely related to sequences from other *Capra* species than to those from the same species. The E and G groups are composed of four types of *Capra*: *C. sibirica*, *C. aegagrus*, *C. hircus*, *C. pyrenaica DRB* sequences; The remaining A, B, C, D, and F groups contain only *DRB* sequences of *C. sibirica*, *C. aegagrus*, and *C. hircus*.

Discussion

In this study, we examined the sequence diversity of the MHC Class II *DRB1* exon 2 of *C. sibirica* from the eastern Tianshan, the middle Tianshan, and the Kunlun Mountains, with a maximum distance of about 2000 km and an average of 4500–5250 m-height mountain peaks[68], which seriously hindered the genetic exchange between populations[39]. In addition, anthropogenic factors led these populations to a drastic decline in size.

Some may wonder allelic dropout could happen when fecal samples used for genetic studies. Nonetheless, it was reported that the detection of MHC alleles using fecal DNA was generally consistent with the results of blood DNA[69]. Considering the worst case, to prevent the possible allelic dropout, we first used the fresh fecal samples and preserved them in 96% ethanol at −80 °C to avoid DNA degradation; we then bidirectionally sequenced at least eight (at most 24) independent clones for an individual; and all sequences from all individuals together were used for searching identical nucleotide sequences which were true alleles, as different individuals probably carry same alleles[43,69]. Moreover, the allele numbers obtained from tissue samples were consistent with that from fecal samples (Table A1). These indicate the obtained alleles in our present study were reliable.
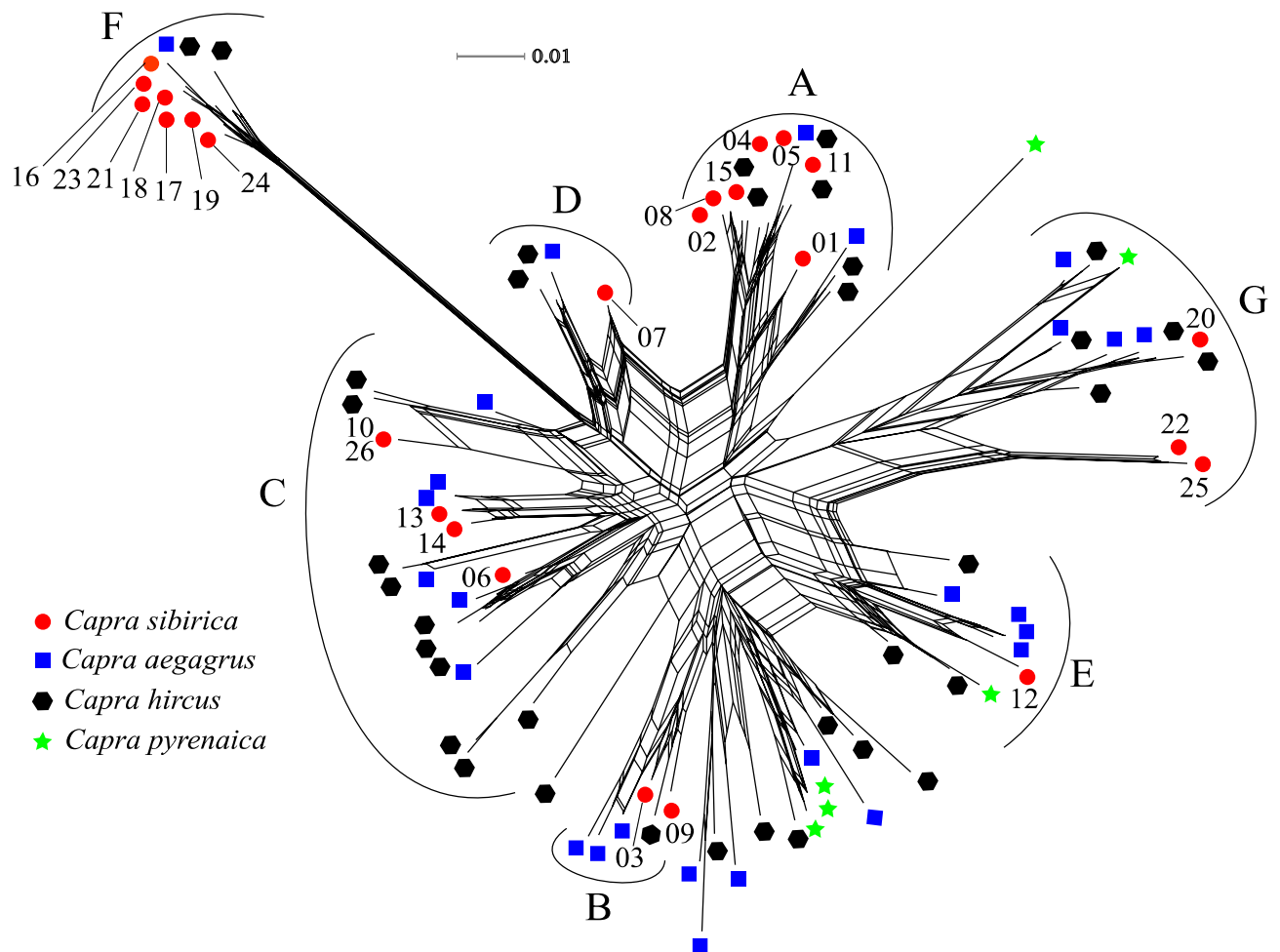
Figure 3. Phylogenetic neighbor network of MHC II *DRB1* exon 2 (233 bp) sequences from *Capra* species, including *C. sibirica* in this study and the remaining three species sequences downloaded from GenBank. The numbers represent allele names with prefix *Casi-DRB1**.

MHC *DRB1* diversity and divergence

Indirect indicators of the immunological fitness of populations, MHC genes are adaptive genetic markers useful in wild animal populations of concern for protection[3,70]. Many species, which went through severe bottlenecks, show very low levels of genetic diversity at the MHC, for example, mountain goats, *Oreamnos americanus*[35] and Galà pagos penguin, *Spheniscus mendiculus*[36]. Conversely, despite a rinderpest epidemic-induced bottleneck, high allelic diversity for the *DRB3* gene was reported for the African buffalo, *Syncerus caffer*[71]. Our study on MHC class II *DRB1* exon 2 allowed, for the first time, a comparison of genetic variation among *C. sibirica* populations that genetically highly diverged and underwent population reduction in size in Xinjiang, China[18,39]. We found higher allelic diversity of MHC class II *DRB1* loci in *C. sibirica* compared to other congenerics. Although the 26 PFAs we detected in 43 *C. sibirica* individuals (Fig. 1, Table A1) seem to be lower than the 22 PFAs among 25 samples reported for its domestic counterpart from six different breeds[72], only seven PFAs were found among 132 individuals of *Capra pyrenaica* with two subspecies, *C. p. hispanica* and *C. p. victoria*[73]. This high number of alleles is mainly attributed the pathogens and parasites they suffer from[22,23], and to some degree to this species' sexual segregation and preference for different habitats and diets for both genders[74,75]. Though the reduction in size in *C. sibirica* may have an impact on the heterozygosity of the MHC *DRB* locus, since more than half of the studied individuals (26 out of 43 samples) possess a single PFA (Table A1).

Individuals of *C. sibirica* clade II had low levels of diversity at the allelic, nucleotide, amino acid, and supertype levels relative to those of Clades I and III (Table A1, Table 1), indicating that the impact of population declines and/or environmental pathogenic pressures on the different geographic populations was different[76]. We also cannot exclude the possibility that this difference is due to the low number of samples analyzed; thus, dense sampling is needed for further related studies.

Although we did not find a single allele shared by all three *C. sibirica* clades, but found alleles common to two clades. For instance, the alleles *Casi-DRB1*16* (the most frequent one identified in individuals from east and middle Tianshan mountains, and Kunlun Mountains), *Casi-DRB1*17* and *Casi-DRB1*19* were shared by clades III and I, while alleles *Casi-DRB1*02* and *Casi-DRB1*05* were shared by clades III and II (Fig. 1 and Table A1). The radiation of these clades dates back around 6.75 million years ago[39], indicating preservation of these alleles

in *C. sibirica* for such a long evolutionary time. An allele was conserved in the genus *Meles* for nearly 2 million years[43]; two alleles were even shared by multiple species from different genera in mustelids[77], diverged more than 11 million years ago[78]; some MHC allele surprisingly preserved among different family[5]. This is probably because *C. sibirica* populations in Xinjiang, China subjected to same pathogenic burdens for a long evolutionary time, as polymorphism in MHC gene was pathogen driven[79].

Meantime, we found more population- or clade-specific alleles (Fig. 1, Table A1), implying high differentiation at this locus. The long radiation time (3.3–6.7 million years) of these populations or clades probably illuminates this phenomenon[39]. MHC genes also showed genetic differentiation between populations in some mammal species[76,80,81]. The F_{ST} values between Clade I and Clade II and Clade II and Clade III were greater than 0.25 (Table 2), indicating large genetic differences. What is puzzling is that the F_{ST} value between the Clade I and Clade III was slightly negative. This is because that the differences within populations were greater than the differences between populations[82]. The negative value of F_{ST} generally interpreted as 0[83], which means clade I and III were not differentiated. However, after excluding the shared alleles between these clades in F_{ST} estimation, it turns to be low differentiation (Table 2). Overall, this in part supports the pattern of mitochondrial DNA[39].

The number of PFAs identified for our studied individuals shows 1 or 2 loci of the *DRB* gene (Table A1), with a low frequency of 2 loci (5 out of 43 individuals), though. Many species in the genus *Capra*, including Alpine ibex (*Capra ibex*), Spanish ibex (*Capra pyrenaica*), and Himalayan tahr (*Hemitragus jemlahicus*), have only one locus of *DRB*[84]. It can be seen that the ancestor of *Capra* species is supposed to possess a single locus at the MHC *DRB* gene. Despite the small portion of individuals with two loci, they split into clades I and III, respectively. We assume that the one locus likely emerged from the other locus through gene duplication[6]. Even if no evidence supports the occurrence of recombination events due likely to the shortness of our analytical sequences, intergenic recombination or gene conversion may explain this phenomenon as well[85], and they might happen twice independently in these two clades. A population genome study on the MHC class II region will help us demonstrate this notion.

Evolution of the *DRB1* gene

Generally, MHC gene polymorphism were generated and retained by gene recombination[84,86], gene duplication[6], balancing selection[79,87], and/or positive selection[43,79,87]. In our study, we did not find any significant signature of recombination events, convincing us that gene recombination was not the reason for generation of MHC diversity. Nonetheless, we found more rare alleles than shared or high frequency alleles (Table A1). This is suggestion of balanced polymorphisms include negative-frequency-dependent selection, where rare alleles are favored. Besides, we also found a notable excess of nonsynonymous over synonymous substitutions at ABSs, in different clades (Table 3). In our phylogenetic relationship analysis, *DRB1* sequences of *C. sibirica* were grouped with the sequences of its counterparts (Fig. 3), suggesting that some alleles are phylogenetically more closely related to the alleles of other species than to those of its own, a typical trans-species polymorphism[88], which is reported for MHC genes of many species[5,43,79,87,89]. All of these were the evidence supporting the presence of long-term balancing selection in the *C. sibirica*, considering that the *Capra* species were diverged approximately 6 million years ago[39]. Moreover, the PAML CodeML and MEME analyses identified up to 12 positively selected sites, most of which coincide with the ABSs (Fig. 2 and Table 4), suggested that the sequence variation of *DRB1* genes was driven by positive selection due to pathogenic burdens[19–23]. In sum, our results together indicate that selection was the main force shaping and maintaining *DRB1* gene polymorphism in *C. sibirica*.

It is worthy to mention that we as well as observed an exceeded nonsynonymous relative to synonymous substitutions at the none-ABSs in all clades of *C. sibirica* (Table 3), which is in line with the positive selection analyses that showed several positively selected sites out of ABSs (Fig. 2 and Table 4). This is consistent with the results of Abduriyim et al.[43] in a species of Canidae. Considering that all MHC studies deduce the ABS locations based on human MHC structure[52], the actual location of ABSs in the MHC Class II DR β-chain of *C. sibirica*, radiation from humans took place as far back as 95 million years[90], may be different. This leaves an open question if ABSs of MHC molecule in all mammals were overlapped.

Conclusions

Despite the level of genetic diversity in clade II is lower than that in other clades, and thus requires close attention in future conservation plans, the overall diversity (i.e., allelic, nucleotide, amino acid and supertype diversity) of MHC class II *DRB1* genes in *C. sibirica* Xinjiang populations after a bottleneck have not rapidly been lost. The differential preference for habitat and food of two sexes might contribute to generation and retain of MHC diversity. The genetic differentiation of clades/populations was to some extent in support of the results by Wang et al.[39] on mtDNA. The diversity of MHC *DRB1* genes in *C. sibirica* was shaped and maintained by selection, both positive and balancing selection.

Data availability

The sequences we obtained have been deposited in the NCBI databases under accession numbers OR257668–OR257693.

References

1. Ellegren, H. & Galtier, N. Determinants of genetic diversity. *Nat. Rev. Genet.* **17**, 422–433 (2016).

2. Kaufman, J. Unfinished business: Evolution of the MHC and the adaptive immune system of jawed vertebrates. *Annu. Rev. Immunol.* **36**, 383–409 (2018).
3. Sommer, S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front. Zool.* **2**, 16 (2005).
4. Meyers, L. A. & Bull, J. J. Fighting change with change: Adaptive variation in an uncertain world. *Trends Ecol. Evol.* **17**, 551–557 (2002).
5. Abduriyim, S. *et al.* Evolution of MHC class I genes in Eurasian badgers, genus *Meles* (Carnivora, Mustelidae). *Heredity* **122**, 205–218 (2019).
6. Abduriyim, S., Zou, D. H. & Zhao, H. B. Origin and evolution of the major histocompatibility complex class I region in eutherian mammals. *Ecol. Evol.* **9**, 7861–7874 (2019).
7. Mungall, A. J. *et al.* The DNA sequence and analysis of human chromosome 6. *Nature* **425**, 805–811 (2003).
8. Wieczorek, M. *et al.* Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. *Front. Immunol.* **8**, 292–305 (2017).
9. Wolin, A. *et al.* Snp variants in major histocompatibility complex are associate with sarcoidosis susceptibility—A joint analysis in four European populations. *Front. Immunol.* **8**, 422 (2017).
10. Alberts, B. *et al. Molecular Biology of the Cell* 5th edn. (Garland Science, 2007).
11. Brown, J. H. *et al.* Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33–39 (1993).
12. Sin, Y. W., Dugdale, H. L., Newman, C., Macdonald, D. W. & Burke, T. MHC class II genes in the European badger (*Meles meles*): Characterization, patterns of variation, and transcription analysis. *Immunogenetics* **64**, 313–327 (2012).
13. Karanth, K. U., Nichols, J. D., Kumar, N. S., Link, W. A. & Hines, J. E. Tigers and their prey: Predicting carnivore densities from prey abundance. *Proc. Natl. Acad. Sci. USA* **101**, 4854–4858 (2004).
14. Suryawanshi, K. R. *et al.* Impact of wild prey availability on livestock predation by snow leopards. *R. Soc. Open Sci.* **4**, 170026 (2017).
15. Schaller, G. *et al.* Status of large mammals in the Taxkorgan Nature Reserve, Xinjiang, China. *Arid Zone Res.* **42**, 53–71 (1987).
16. Wang, X. F. Diagnosis and treatment of viral keratoconjunctivitis in *Capra sibirica*. *Heilongjiang Anim. Sci. Vet. Med.* **307**, 145 (2007).
17. Zhu, X. S., Wang, M. Y., Yang, W. K. & David, B. Ecology and biology of *Capra sibirica*: Current situation of studies. *Chin. J. Ecol.* **34**, 3553–3559 (2015).
18. Reading, R. Michel, S. Suryawanshi, K. & Bhatnagar, Y. V. *Capra sibirica*. The IUCN Red List of Threatened Species 2020, e.T42398A22148720. https://doi.org/10.2305/IUCN.UK.2020-2.RLTS.T42398A22148720.en. Accessed on 31 May 2023.
19. Thermistor, H. Diagnosis of streptococcal disease in *Capra sibirica*. *Chin. J. Anim. Husb. Vet. Med.* **4**, 42 (2008).
20. Chen, W. D., Li, T. & Yakp,. Laboratory diagnosis and prevention of death of wild *Capra sibirica* caused by Streptococcus pneumoniae. *Xinjiang Anim. Husb.* **7**, 55–56 (2012).
21. Bayendrigan, & An, N. Diagnosis and control of anthrax infection in wild *Capra sibirica*. *J. Anim. Sci. Vet. Med.* **34**, 124 (2015).
22. Ahmed, S. Parasites of markhor, urial and Chiltan wild goat in Pakistan. *Ann. Parasitol.* **66**, 3–12 (2020).
23. Sarwar, M. M. On the synonymy of *Marshallagia orientalis* (Bhalerao, 1932) with *Marshallagia marshalli* and a record of its occurrence from the mountainous region of the Punjab. *Parasitology* **42**, 125–125 (1952).
24. Zhu, X. S. *Studies on Feeding Habits and Homoerotic Clusters of Capra sibirica* (University of Chinese Academy of Sciences, 2016).
25. Yang, Q. S. & Feng, Z. J. *China Red Book of Endangered Animals: Mammals* 314–317 (Beijing Science Press, 1998).
26. Reading, R. & Shank, C. *Capra sibirica*. IUCN Red List of Threatened Species 2009.
27. Wang, S. & Xie, Y. *Red List of Chinese Species* (Higher Education Press, 2004).
28. Abdukadir, A., Dai, Z., Shi, J. & Wang, Z. Investigation of population resources of Argali and Ibex on the higher-mountains in Xinjiang Taxkorgan in winter. *J. Arid Land Res. Environ.* **11**, 174–179 (2010).
29. Abdukadir, A. *et al.* Distribution and population resources of argali sheep and ibex in Baicheng of Aksu Tianshan Mountains. *Chin. J. Wildl.* **31**, 270–275 (2010).
30. Xu, F., Ma, M. & Wu, Y. Q. Population density and habitat utilization of ibex (*Capra ibex*) in Tomur National Nature Reserve, Xinjiang, China. *Zool. Res.* **1**, 53–55 (2007).
31. Dai, K. *The Population and Distribution Research of Blue Sheep (Pseudois navaur) and Ibex (Capra sibirica) in West Kunlun Mountain Area* (Beijing Forestry University, 2012).
32. Wang, J., Shi, K. & Riordan, P. Study on population density of ungulates in Taxkorgan, Xinjiang, China. *Chin. J. Wildl.* **3**, 113–117 (2012).
33. National Forestry and Grassland Administration, National Park Administration. http://www.forestry.gov.cn (2021).
34. Radwan, J., Kawako, A., Wójcik, J. M. & Babik, W. Mhc-drb3 variation in a free-living population of the European bison, *Bison bonasus*. *Mol. Ecol.* **16**, 531–540 (2010).
35. Mainguy, J., Worley, K., Côté, S. D. & Coltman, D. W. Low MHC *DRB* class II diversity in the mountain goat: Past bottlenecks and possible role of pathogens and parasites. *Conserv. Genet.* **8**, 885–891 (2007).
36. Bollmer, J. L., Vargas, F. H. & Parker, P. G. Low MHC variation in the endangered Galápagos penguin (*Spheniscus mendiculus*). *Immunogenetics* **59**, 593–602 (2007).
37. Ejsmond, M. J. & Radwan, J. MHC diversity in bottlenecked populations: A simulation model. *Conserv. Genet.* **12**, 129–137 (2011).
38. Oliver, M. K. & Piertney, S. B. Selection maintains MHC diversity through a natural population bottleneck. *Mol. Biol. Evol.* **29**, 1713–1720 (2012).
39. Wang, R. R., Dong, P. P., Hirata, D. & Abduriyim, S. Mitochondrial DNA analyses revealed distinct lineages in an alpine mammal, Siberian ibex (*Capra sibirica*) in Xinjiang, China. *Ecol. Evol.* **13**, e10288 (2023).
40. Abduriyim, S., Nabi, A. & Halik, M. Low genetic diversity in the goitered gazelle *Gazella subgutturosa* (Guldenstadt, 1780) (Artiodactyla: Bovidae) in North-western China as revealed by the mitochondrial cytochrome *b* gene. *Acta Zool. Bulgar.* **70**, 211–218 (2018).
41. Abduriyim, S., Zibibulla, G., Eli, S., Ismayil, Z. & Halik, M. Phylogeny and genetic structure of the goitered gazelle (Artiodactyla, Bovidae) in north-western China indicated by the hypervariable mitochondrial control region. *Syst. Biodivers.* **16**, 527–537 (2018).
42. Amills, M., Francino, O. & Sanchez, A. Nested PCR allows the characterization of Taql and Pstl RFLPs in the second exon of the caprine MHC class II *DRB* gene. *Vet. Immunol. Immunopathol.* **48**, 313–321 (1995).
43. Abduriyim, S. *et al.* Diversity and evolution of MHC class II *DRB* gene in the Eurasian badger genus *Meles* (Mammalia: Mustelidae). *Biol. J. Linn. Soc.* **122**, 258–273 (2017).
44. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
45. Librado, P. & Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
46. Kennedy, L. J. *et al.* Nomenclature for factors of the dog major histocompatibility system (DLA), 1998: First report of the ISAG DLA Nomenclature Committee. *Anim. Genet.* **31**, 52–61 (2000).
47. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
48. Klein, J. *et al.* Nomenclature for the major histocompatibility complexes of different species: A proposal. *Immunogenetics* **31**, 217–219 (1990).

49. Ballingall, K. T. & Todd, H. An official nomenclature for the major histocompatibility complex allele sequences from the domestic goat (*Capra hircus*). *HLA* **93**, 36–38 (2019).
50. Nei, M. & Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–426 (1986).
51. Yang, Z. H. & Nielsen, R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* **19**, 908–917 (2002).
52. 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.* **331**, 623–641 (2003).
53. Pond, S. L. K., Frost, S. D. W. & Muse, S. V. HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
54. Yang, Z. H. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
55. Nielsen, R. & Yang, Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**, 929–936 (1998).
56. Yang, Z. H. & Bielawski, J. P. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**, 496–503 (2000).
57. Yang, Z. H., Wong, W. S. W. & Nielsen, R. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* **22**, 1107–1118 (2005).
58. Weaver, S. *et al.* Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Mol. Biol. Evol.* **35**, 773–777 (2018).
59. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764 (2012).
60. Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **1**, 1–5 (2015).
61. Martin, D. P. *et al.* RDP3: A flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463 (2010).
62. Padidam, M., Sawyer, S. & Fauquet, C. M. Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218–225 (1999).
63. Smith, J. M. Analyzing the mosaic structure of genes. *J. Mol. Evol.* **34**, 126–129 (1992).
64. Martin, D. P., Posada, D., Crandall, K. A. & Williamson, C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res. Hum. Retrov.* **21**, 98–102 (2005).
65. Pond, S. L. K., Posada, D., Gravenor, M. B., Woelk, C. H. & Frost, S. D. W. GARD: A genetic algorithm for recombination detection. *Bioinformatics* **22**, 3096–3098 (2006).
66. Delport, W., Poon, A. F. Y., Frost, S. D. W. & Pond, S. L. K. Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26**, 2455–2457 (2010).
67. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2006).
68. Zhang, X. J. *et al.* Spatiotemporal distribution of cloud liquid water volume over three main mountains in Xinjiang. *Arid Zone Res.* **35**, 846–854 (2018).
69. Yang, B. *et al.* Major histocompatibility complex and mate choice in the polygynous primate: The Sichuan snub-nosed monkey (*Rhinopithecus roxellana*). *Integr. Zool.* **9**, 598–612 (2014).
70. Ujvari, B. & Belov, K. Major histocompatibility complex (MHC) markers in conservation biology. *Int. J. Mol. Sci.* **12**, 5168–5186 (2011).
71. Wenink, P. W., Groen, A. F., Roelke-Parker, M. E. & Prins, H. H. T. African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks. *Mol. Ecol.* **7**, 1315–1322 (1998).
72. Schwaiger, F. W. *et al.* The paradox of MHC-*DRB* exon/intron evolution: α-helix and β-sheet encoding regions diverge while hypervariable intronic simple repeats coevolve with β-sheet codons. *J. Mol. Evol.* **37**, 260–272 (1993).
73. Angelone, S. *et al.* Hidden MHC genetic diversity in the Iberian ibex (*Capra pyrenaica*). *BMC Genet.* **19**, 28 (2018).
74. Han, L. *et al.* Diet differences between males and females in sexually dimorphic ungulates: A case study on Siberian ibex. *Eur. J. Wildl. Res.* **66**, 1–10 (2020).
75. Han, L., Wang, Z., Blank, D., Wang, M. Y. & Yang, W. K. Different environmental requirements of female and male Siberian ibex, *Capra sibirica*. *Sci. Rep.* **11**, 6064 (2021).
76. Awadi, A. *et al.* Positive selection and climatic effects on MHC class II gene diversity in hares (*Lepus capensis*) from a steep ecological gradient. *Sci. Rep.* **8**, 11514 (2018).
77. Sugiyama, Y. *et al.* Diversity of the MHC class II *DRB* gene in the wolverine (Carnivora: Mustelidae: *Gulo gulo*) in Finland. *PLoS One* **17**, e0267609 (2022).
78. Koepfli, K. P. *et al.* Multigene phylogeny of the Mustelidae: Resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biol.* **6**, 10 (2008).
79. Nishita, Y. *et al.* Genetic variation of the MHC class II *DRB* genes in the Japanese weasel, *Mustela itatsi*, endemic to Japan, compared with the Siberian weasel, *Mustela sibirica*. *Tissue Antigens* **86**, 431–442 (2015).
80. Amaike, Y., Nishita, Y., Uraguchi, K. & Masuda, R. Genetic diversity of MHC class II *DRB1* exon 2 in the red fox (*Vulpes vulpes*) on Hokkaido, Japan. *Zool. Sci.* **35**, 402–410 (2018).
81. Mason, R. A. B., Browning, T. L. & Eldridge, M. D. B. Reduced MHC class II diversity in island compared to mainland populations of the black-footed rock-wallaby (*Petrogale lateralis lateralis*). *Conserv. Genet.* **12**, 91–103 (2011).
82. Smaragdov, M. G., Kudinov, A. A. & Uimari, P. Assessing the genetic differentiation of Holstein cattle herds in the Leningrad region using Fst statistics. *Agric. Food Sci.* **27**, 96–101 (2018).
83. Meirmans, P. G. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **60**, 2399–2402 (2006).
84. Schaschl, H., Wandeler, P., Suchentrunk, F., Obexer-Ruff, G. & Goodman, S. J. Selection and recombination drive the evolution of MHC class II DRB diversity in ungulates. *Heredity* **97**, 427–437 (2006).
85. Reusch, T. B. & Langefors, Å. Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined stickleback *Gasterosteus aculeatus*. *J. Mol. Evol.* **61**, 531–541 (2005).
86. Bartocillo, A. M. F., Nishita, Y., Abramov, A. V. & Masuda, R. Molecular evolution of MHC class II *DRB* exon 2 in Japanese and Russian raccoon dogs, *Nyctereutes procyonoides* (Carnivora: Canidae). *Biol. J. Linn. Soc.* **129**, 61–73 (2020).
87. Nishita, Y. *et al.* Diversity of MHC class II *DRB* alleles in the Eurasian population of the least weasel, *Mustela nivalis* (Mustelidae: Mammalia). *Biol. J. Linn. Soc.* **121**, 28–37 (2017).
88. Klein, J., Sato, A., Nagl, S. & Oh'Uigín, C. Molecular trans-species polymorphism. *Annu. Rev. Immunol.* **29**, 1–21 (1998).
89. Bartocillo, A. M. F., Nishita, Y., Abramov, A. V. & Masuda, R. Evolution of MHC class I genes in Japanese and Russian raccoon dogs, *Nyctereutes procyonoides* (Carnivora: Canidae). *Mamm. Res.* **66**, 371–383 (2021).
90. Eizirik, E., Murphy, W. J. & O'Brien, S. J. Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* **92**, 212–219 (2001).

Acknowledgements

Author contributions

S.A. conceived and designed the study; D.P., W.R., and S.A. did the fieldwork and samples collection; D.P. and W.R. conducted lab experiments, D.P. and W.R. performed data analysis; D.P. and S.A. prepared the first draft, S.A. edited the manuscript. All authors approved the final version of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-46717-5.

Correspondence and requests for materials should be addressed to S.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.