Immunogenetics (2005) 57: 108–115
DOI 10.1007/s00251-005-0784-4

ORIGINAL PAPER

Helmut Schaschl · Franz Suchentrunk · Sabine Hammer · Simon J. Goodman

Recombination and the origin of sequence diversity in the *DRB* MHC class II locus in chamois (*Rupicapra* spp.)

Received: 17 November 2004 / Revised: 18 January 2005 / Published online: 9 March 2005

Abstract We examined the evolutionary processes contributing to genetic diversity at the major histocompatibility complex (MHC) class II *DRB* locus in chamois (*Rupicapra* spp., subfamily Caprinae). We characterised the pattern of intragenic recombination (or homologous gene conversion) and quantified the amount of recombination in the geneaogical history of the two chamois species, Pyrenean chamois (*Rupicapra pyrenaica*) and Alpine chamois (*Rupicapra rupicapra*). We found evidence for intragenic recombination, and the estimated amount of population recombination suggests that recombination has been a significant process in generating *DRB* allelic diversity in the genealogical history of the genus *Rupicapra*. Moreover, positive selection appears to act on the same peptide-binding residues in both analysed chamois species, but not in identical intensity. Recombination coupled with positive selection drives the rapid evolution at the peptide-binding sites in the MHC class II *DRB* gene. Many chamois MHC class II *DRB* alleles are thus much younger than previously assumed.

Keywords MHC class II diversity · Chamois · Recombination · Allelic polymorphism

H. Schaschl · S. J. Goodman
Institute of Zoology, Zoological Society of London,
Regent's Park,
London, NW1 4RY, UK

F. Suchentrunk
Research Institute of Wildlife Ecology,
University of Veterinary Medicine Vienna,
Savoyenstrasse 1,
1160 Vienna, Austria

S. Hammer
Laboratories of Genome Dynamics,
Medical University of Vienna,
Waehringerstrasse 10,
1090 Vienna, Austria

S. J. Goodman
Department of Biology, University of Leeds,
L. C. Miall Building, Clarendon Way,
Leeds, LS2 9JT, UK

H. Schaschl (✉)
Department of Evolutionary Ecology,
Max-Planck-Institute of Limnology,
August-Thienemann-Str. 2,
24306 Ploen, Germany
e-mail: Schaschl@mpil-ploen.mpg.de
Tel.: +49-4522763-258
Fax: +49-4522763-310

Introduction

The genus *Rupicapra* (chamois) is a group within the subfamily Caprinae of the family Bovidae and is thought to have appeared about 20 MYA. It encompasses the two extant chamois species *Rupicapra rupicapra* (Alpine chamois) and *Rupicapra pyrenaica* (Pyrenean chamois) (Lovari 1987; Masini and Lovari 1988). The phylogenetic divergence time between the two chamois species was estimated to range between about 150 kyBP (cf. Masini and Lovari 1988) and 280 kyBP (Hammer et al. 1995). *R. rupicapra* consists of seven subspecies (*rupicapra*, *tatrica*, *carpatica*, *balcanica*, *cartusiana*, *caucasica* and *asiatica*), distributed among the Alps, the Tatra massif (Slovak Republic), the Carpathian Mountains (Romania) and various mountain massifs in the Balkans, Asia Minor and the Caucasus. *R. pyrenaica* consist of three subspecies (*pyrenaica*, *parva* and *ornata*) which are disjunctly distributed in south-west Europe (Pyrenees and Cantabrian Mountains) and the southern Apennine Mountain Range in Italy (Shackleton 1997). Normally, chamois live in small social groups throughout the year, with only older bucks remaining mostly solitary. The population dynamics of *R. rupicapra* populations in large parts of the Eastern Alps as well as populations in the Iberian peninsular (*R. pyrenaica*) and of Ibexes (*Capra ibex, C. pyrenaica*) is influenced by sarcoptic mange epidemics, an infection of *Sarcoptes rupicaprae*. Commonly, high mortality rates (80%) are noted in local populations in the course of regional epidemics (León-Vizcaíno et al. 1999; Rossi et al. 1995). Given this background chamois provide an interesting system for the study of the dynamics of host–parasite interactions and

their effects on the evolution of immune genes in natural populations. Genetic polymorphism in immune-response genes may vary temporally and spatially in host populations, which makes it difficult for pathogens to adapt to all genotypes in a population ('moving targets'). Genes of the major histocompatibility complex (MHC) are usually highly polymorphic and a striking example of selectively maintained polymorphic genes in vertebrates. The MHC genes encode cell-surface glycoproteins playing a key role in the vertebrate adaptive immune system. The main function of the MHC class II molecules is presenting peptide fragments derived from pathogens on the cell surface to T-helper cells. Subsequently, T-helper cells stimulate B cells to generate and secret antibodies. However, MHC molecules also play an important role in the shaping of the T-cell receptor repertoire during T-cell maturation in the thymus (Janeway et al. 2001). MHC class II molecules are α/β heterodimers and are mainly expressed on specialized antigen-presenting cells such as macrophages, B cells or dendritic cells. The α_1 and β_1 domains form the peptide-binding region (PBR) in which peptides are bound and recognized by T-helper cell receptors. The β_1 domain is encoded by exon 2 of β-genes, and its genetic diversity can be extremely high at the population level, both in terms of the number of alleles present and in the extent of sequence diversity among alleles (Apanius et al. 1997). In humans, for example, 362 MHC class II *DRB1* alleles are currently assigned (Robinson et al. 2003). In the PBR the number of non-synonymous substitutions usually exceeds the number of synonymous substitutions, an indication that positive selection is acting at PBR sites within genes, driving the diversification of MHC loci. It is thought primarily that some form of balancing selection (heterozygote advantage or negative frequency dependent selection) maintains the high polymorphism at the MHC (Hughes and Yeager 1998). Balancing selection may act to maintain ancient alleles and amino acid motifs in a trans-specific manner in mammals, that is, allelic lineages are passed from species to species and persist in the populations over long periods of evolutionary time (Klein et al. 1993). MHC loci are becoming increasingly well characterised for a growing number of ungulate species. However, the relative importance of different molecular mechanisms behind the generation of allelic diversity in the MHC remains contentious. Intragenic recombination (or homologous gene conversion) has been suggested as an important evolutionary mechanism for the generation of MHC sequence diversity (Andersson and Mikko 1995; Bergstrom et al. 1998; Gyllensten et al. 1991; Richman et al. 2003a,b). On the other hand, it has also been suggested that steady accumulation of point mutations and occasional convergence evolution cannot be ruled out as an important mode to create new MHC alleles (Klein et al. 1993; Klein and O'hUigin 1995; O'hUigin 1995; Takahata and Satta 1998). The direct measurement of recombination rate by sperm typing or pedigree analyses at a high resolution is technically difficult and time consuming. Moreover for wild mammals it is often impracticable, only really being feasible for laboratory or domestic species. However, the application of recently developed statistical methods to DNA sequences from natural populations now permits indirect, but quantitative, estimation of recombination rate from population–genetic data (Hudson 2001; McVean et al. 2002), allowing quantitative assessments of the relative contributions of different molecular mechanisms in generating allelic diversity at the MHC.

Mutation provides a constant influx of new MHC variants into the population, but a high recombination rate would speed up this influx. Therefore, high recombination rate in the MHC enables the host to keep up with the usually faster evolving parasites and might have beneficial fitness effects for the host. Therefore, we aimed to provide insights into the scale of the population recombination rate in chamois as compared to point mutations. After the divergence of the genus *Rupicapra* into Pyrenean chamois and Alpine chamois, both species might have experienced different intensity of adaptive evolution at the *DRB* locus. A further main objective was therefore to test for the intensity of diversifying selection in the PBR in both chamois species and to identify those sites that are probably under strong positive selection.

Materials and methods

Samples, DNA isolation and sequence data from GenBank

This study is based on the combined analysis of new data on 18 Pyrenean chamois (*R. pyrenaica*) from two Spanish populations in the Pyrenees (*Rupicapra p. pyrenaica*) and the Cantabrian Mountains (*R. p. ornata*) as well as one Italian population from the Abruzzo Mountains (*R. p. parva*) (for sample overview see Table 1) and the following sequences of MHC class II second exon alleles obtained from GenBank: 19 alleles from Alpine chamois (*Rupicapra r. rupicapra*, accession nos. AF324840–AF324861, Schaschl et al. 2004) and nine alleles from Pyrenean chamois (*R. pyrenaica*, accessions nos. AY212149–AY212157, Alvarez-Busto and Jugo, unpublished). Genomic DNA from the Pyrenean chamois samples was extracted with a DNA

Table 1 List of *Rupicapra pyrenaica* (Pyrenean chamois) subspecies used in the present study; observed Pyrenean chamois *DRB* alleles (*Rupy-DRB*) are given

Species	Sub-species	Geographic location	No. of samples	Observed *DRB* alleles
R. pyrenaica	*R. p. pyrenaica*	BenásValley, Spain	5	*Rupy-DRB02* *Rupy-DRB04* *Rupy-DRB10* *Rupy-DRB11*
	R. p. parva	Asturias, Spain	4	*Rupy-DRB01* *Rupy-DRB02* *Rupy-DRB04*
	R. p. ornata	Abruzzo, Italy	9	*Rupy-DRB012* *Rupy-DRB013*

extraction kit (Deasy Tissue Kit, Qiagen) according to the manufacturer's protocol.

PCR amplification, cloning and sequencing

PCR amplification of the *DRB* exon 2 was achieved by following the protocol of Schaschl et al. (2004). PCR products were cloned into the pCR 2.1 TOPO plasmid (Invitrogen). Between five and eight clones per individual were sequenced. Sequences were determined on both DNA strands using BigDey Terminator Cycle Sequencing Kit v3.1 (Applied Biosystem) and an ABI 3100 DNA Sequencer. Samples with poor cloning efficiency were discarded, and sequences were only considered when they were found in two or more clones. Preliminary sequence processing and analysis was performed with BioEdit (Hall 1999).

Sequence analysis and nomenclature

All nucleotide and amino acid sequences were aligned with the program ClustalX (Thompson et al. 1997). Phylogenetic and molecular evolutionary analyses were conducted using MEGA, version 2.1 (Kumar et al. 2001). The relative frequencies of non-synonymous (d_N) and synonymous (d_S) substitution in the exon 2 were calculated according to Nei and Gojobori (1986) and applying Jukes and Cantor's correction for multiple substitutions (Jukes and Cantor 1969). The significance of the difference between these rates was tested with a *Z*-test of selection at the 5% level, whereby the *P*-values are the probability of rejecting the null hypothesis of neutrality (d_N=d_S; Nei and Kumar 2000). In accordance with the proposed nomenclature for MHC in nonhuman species (Klein et al. 1990), we designated the exon 2 alleles *Rupy-DRB* for Pyrenean chamois (*R.pyrenaica*) with serial numbers attached.

Recombination/gene conversion analysis

Intragenic recombination involves the exchange of sets of DNA segments from the same gene, generating DNA blocks of compatible sites that are incompatible with contiguous DNA blocks.

The programme Geneconv, version 1.81 (Sawyer 1989; Sawyer 1999; available at http://www.math.wustl.edu/~sawyer/mbprogs/), was employed to find the most likely candidate alleles for intragenic recombination/gene conversion events in the genealogical history the genus *Rupcapra*. This method uses pairwise comparison of sequences in the alignment to find blocks of sequence pairs that are more similar than would be expected by chance. Genconv finds and ranks the highest-scoring fragments globally for the entire alignment. Global permutation test *P*-values of <0.05 (derived from BLAST-like global scores using 10,000 replicates) were considered as evidence of intragenic recombination. These global permutation test *P*-values have an intrinsic multiple-comparison correction for all sequence pairs in the alignment. The underlying method in Geneconv has a high statistical power for detecting recombination when recombination is present or likely to be present, while the risk of obtaining false positive results is low (Posada 2002).

The population recombination rate (ρ=$4N_e r$) in chamois was estimated using the programme LDhat (see for details McVean et al. 2002). This programme implements Hudson's (2001) composite–likelihood estimate approach to estimate the population recombination rate conditioned on the estimate of mutation rate per site (θ=$4N_e\mu$) from an approximate finite-sites version of the Watterson estimate. This method has been extended by McVean et al. (2002) to take into account high rates of recurrent mutations in sequences. The estimate of $4N_e r$ is taken as the value that has the highest composite likelihood estimate (McVean et al. 2002). We used the implemented likelihood permutation test to test the null hypothesis of no recombination (ρ=0). Extensive computer simulation, carried out by Richman et al. (2003b), revealed that the LDhat estimates recombination rates fairly accurately with regard to sequences evolving under symmetric balancing selection. Furthermore, we calculated the ratio ρ/θ as an estimate of the relative amount of recombination compared to mutation, which is robust against several violations of the underlying coalescent model (Fearnhead and Donnelly 2001).

Test for positive selection using maximum-likelihood analysis

We used the programme CODEML of the PAML, version 3.14, package (Yang 1997) to test for the presence of codon sites affected by positive selection and to identify those sites. Positive selection is indicated by ω=d_N/d_S>1. The models considered in this study were M7 (beta), and M8 (beta and ω) (Yang et al. 2000). Under the model M7 (beta) the ω ratio various according to the beta distribution and does not allow for positive selected sites (0<ω<1) and thus serves as the null model by comparing with model M8 (beta and ω). Model M8 adds an additional site class to the beta model to account for sites under positive selection (ω>1). The models M7 and M8 can be compared in pairs using the likelihood-ratio test (LRT) (Nielsen and Yang 1998). The LRT statistics calculates twice the log-likelihood difference compared with a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters between the two compared models. The best tree for both species by maximum-likelihood search was in accordance with the one-ratio model (M0) used to provide phylogenetic information. A Bayesian approach implemented in CODEML was used to identify residues under positive selection in the MHC class II *DRB* sequences.

Results

Allelic polymorphism at the MHC class II *DRB* locus

We obtained seven MHC class II *DRB* exon 2 sequences (*Rupy-DRB*) from *R. pyrenaica* (Pyrenean chamois) samples (Fig. 1). A GenBank search revealed that three of our observed sequences were identical to three of the nine previously published Pyrenean chamois *DRB* exon 2 sequences (*Rupy-DRB01*, *Rupy-DRB02*, and *Rupy-DRB04*; see "Materials and methods" for details and accession numbers). The four novel *DRB* exon 2 sequences in our study were named in accordance with the published sequences as *Rupy-DRB10* to *Rupy-DRB13* (see Table 1), and were submitted to GenBank (GenBank accession numbers AY898752–AY898755). We included in our subsequent analyses the published *R. pyrenaica* sequences as well as published *DRB* exon 2 sequences from *R. rupicapra* (Alpine chamois). Nucleotide sequence variation among all pairwise comparisons of *Rupy-DRB* sequences, corrected for multiple substitutions, ranged from 0.9% to 8.6%, with a mean of 4.8±1.0% (±standard deviation). Mean nucleotide sequence variation among all pairwise comparisons of Alpine chamois *DRB* alleles is 4.4±0.9% and thus similar to Pyrenean chamois variation. The nucleotide divergence between Pyrenean chamois and Alpine chamois is 4.7 ±0.9%. In both chamois species, d_N occurred significantly more frequently than d_S at the PBR sites (*Rupy-DRB*: d_N=0.188±0.051, d_S=0.026±0.017, d_N/d_S=7.2, P=0.001; *Ruru-DRB*: d_N=0.142±0.044, d_S=0.010±0.007, d_N/d_S=14.2, P=0.0003). Among all the sequences in the current study, two Pyrenean chamois sequences were identical with two Alpine chamois sequences (*Rupy-DRB02* and *Ruru-DRB01* (also identical at the nucleotide level), and *Rupy-DRB04* and *Ruru-DRB13* (different in one d_S at the codon position 84). In both chamois species the d_S was far below the d_S found in most other studied ungulates (Gutierrez-Espeleta et al. 2001; Jugo and Vicario 2000; Mikko et al. 1999; Van Den Bussche et al. 1999). This finding is probably a consequence of a young phylogenetic age of the alleles (that is, chamois *DRB* alleles have evolved more recently than that of other ungulates so far studied). It potentially reflects a complex demographic history of chamois populations (see "Discussion").

Level of intragenic recombination

The Geneconv analysis shows that intragenic recombination (or homologous gene conversion) events at the *DRB* locus have occurred in both species. In fact, intragenic recombination events were not only detected within segmental variants of each species but also between alleles of the two different species. Table 2 shows the *Rupy-DRB* and *Ruru-DRB* alleles that may have been involved in intragenic recombination events as well as the relative position of the DNA segments from 83 bp to 100 bp in length, which were involved in recombination. In total the analysis revealed that five *Rupy-DRB* and nine *Ruru-DRB* alleles were found to be involved in recombination events. Apparently, some sequence blocks (for example, DNA block 98–196 and DNA block 147–234, see Table 2) were repeatedly involved in recombination events and may have served as recombination hot spots. Further evidence for recombination in the chamois *DRB* sequences comes from the r^2 correlation test for recombination (Awadalla et al. 1999) implemented in the LDhat programme. This test detected a significant (P<0.05) decay of linkage disequilibrium with pairwise distance, suggesting recombination in the analysed sequences. The population recombination rate (ρ) was estimated as ρ=78 in Pyrenean chamois and ρ=37 in Alpine chamois (Table 3). In both cases these values are very high, being an order of magnitude greater than the corresponding population mutation estimates (θ),

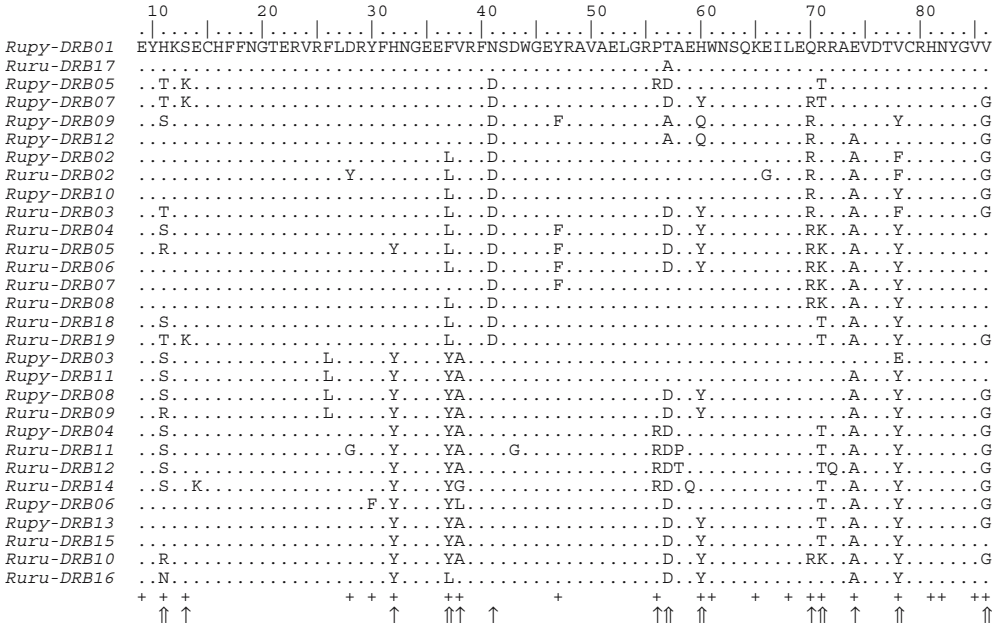

Fig. 1 Alignment of the putative amino acid sequences for MHC class II *DRB* exon 2 from *R. pyrenaica* (*Rupy-DRB*) (Pyrenean chamois) and *R. rupicapra* (*Ruru-DRB*) (Alpine chamois). Sequences were arranged to display similar *DRB* alleles together. *Dots* indicate identity in the amino acid sequence of the *Rupy-DRB01*, and a *cross* indicates codons involved in peptide binding regions (PBRs) in human (Brown et al. 1993). *Thin arrows* indicate PBRs under strong positive selection in *R. pyrenaica*, and *thick arrows* indicate PBRs under strong positive selection in *R. pyrenaica* and in *R. rupicapra* (>0.95 posterior probabilities)

Table 2 Statistical test for recombination for chamois MHC class II *DRB* exon 2 alleles as assessed by the programme Geneconv, version 1.81

MHC class II *DRB* alleles[a]	KA (*P*-value)[b]	Aligned offsets			Num Poly[c]	Num Dif[c]	Tot Difs[c]	MisM Pen[c]
		Begin	End	Length				
Rupy-DRB03, *Ruru-DRB19*	0.035	98	187	90	13	0	14	None
Rupy-DRB04, *Rupy-DRB05*	0.040	98	196	99	15	0	12	None
Rupy-DRB04, *Ruru-DRB19*	0.045	147	234	88	16	0	11	None
Rupy-DRB05, *Ruru-DRB07*	0.035	16	115	100	13	0	14	None
Rupy-DRB05, *Ruru-DRB13*	0.024	98	196	99	15	0	13	None
Rupy-DRB06, *Ruru-DRB19*	0.027	147	234	88	16	0	12	None
Rupy-DRB07, *Ruru-DRB07*	0.022	16	115	100	13	0	15	None
Rupy-DRB07, *Ruru-DRB09*	0.038	98	184	87	11	0	16	None
Rupy-DRB07, *Ruru-DRB10*	0.035	98	187	90	13	0	14	None
Ruru-DRB11, *Ruru-DRB18*	0.036	149	232	84	14	0	13	None
Ruru-DRB11, *Ruru-DRB19*	0.008	149	234	86	15	0	15	None
Ruru-DRB14, *Ruru-DRB19*	0.036	152	234	83	14	0	13	None

[a]*Rupy-DRB* designates the *R. pyrenaica* (Pyrenean chamois) alleles and *Rupy-DRB* the *R. rupicapra* (Alpine chamois) alleles
[b]*KA P*-values (Karlin-Altschul *P*-values (Karlin and Altschul 1990) are Bonferroni-corrected
[c]*Num Poly* Number of polymorphic sites in the fragment, *Num Dif* number of mismatches within the fragment, *Tot Difs* number of mismatches between two sequences, *MisM Pen* penalty per mismatch for these two sequences

Table 3 Statistic and *P*-values for population mutation (Watterson's $\theta=4N\mu$) and population recombination rate ($\rho=4Nr$, McVean et al. 2002)

Species	No. of alleles	θ	ρ	ρ/θ	*P*
R. pyrenaica	13	8.57	78	9.1	<0.001
R. rupicapra	19	3.72	37	9.9	<0.001

and indicate a large contribution from recombination in the history of these sequences. The likelihood permutation test showed that the ρ estimates were significantly different from those expected under the null hypothesis of no recombination ($P<0.001$).

Detecting positive selection at sites using maximum-likelihood analysis

The LRT statistic comparing the two models indicates that M8 fitted the data significantly ($P<0.01$) better than M7. The estimates from M8 suggested that about 23% of the sites were under strong positive selection in the Pyrenean chamois sequences ($\omega=20.49$) and 13% in the Alpine chamois sequences ($\omega=15.23$) (Table 4). Bayes identification of sites under positive selection is listed in Table 4. Those sites that were identified as positively selected sites are mostly in accordance with the human PBR sites (*HLA-DRB1* gene) (Brown et al. 1993; for PBR sites see Fig. 1).

Discussion

In this study we found extensive sharing of amino acid motifs between the *DRB* alleles of the two extant chamois species. Further, two Alpine chamois *DRB* alleles are identical to two Pyrenean chamois alleles. If these two alleles do not result from convergent evolution, they could present shared ancestral alleles. Alternatively, relatively recent extensive hybridisation and introgression between the two chamois species (e.g. events of reticulate evolution) could explain these findings. However, the two spe-

Table 4 Log-likelihood values and parameter estimates for the MHC class II *DRB* alleles of chamois

Model	lnL[a]		Estimates of parameters[a]		Positively selected sites[b]	
	Pyrenean chamois	Alpine chamois	Pyrenean chamois	Alpine chamois	Pyrenean chamois	Alpine chamois
M7 (beta)	−601.90	−684.75	p=0.0050 q=0.0076	p=0.0163 q=0.0317	Not allowed	Not allowed
M8 (beta and ω)	−580.27	−659.08	p_0=0.771 (p_1=0.229) p=0.117 q=2.0322 ω=20.49	p_0=0.869 (p_1=0.131) p=2.490 q=0.296 ω=15.23	11H** 13S** 32* 37F* 38V** 41** 56P** 57T** 60H** 70Q** 71R** 74E** 78V** 86V**	11H** 37L** 57T** 60H** 71R** 78F** 86G**

[a]*LnL* is the log-likelihood value, ω is the selection parameter and p_n is the proportion of sites that fall into ω_n site class
[b]Sites inferred to be under positive selection are given at the 95% (*) and 99% (**) confidence interval level

cies are geographically isolated by considerable distance and unsuitable habitat. There is no evidence of recent gene flow from neutral DNA marker studies (Hammer et al. 1995; Pérez et al. 2002), but some 20,000 years ago, during the late glacial maximum, chamois roamed over wide areas in central Europe (Sägesser and Krapp 1986). The late Pleistocene distribution of both species could have resulted in temporary contact or overlap of ranges and might have thus enabled introgression. Preliminary mtDNA and nuclear gene sequence data (Hammer et al., unpublished data) suggest such episodes of reticulate evolution in chamois. One of the two pairs of shared alleles among the Pyrenean chamois and Alpine chamois alleles is *Rupy-DRB02* and *Ruru-DRB01*. The allele *Ruru-DRB01* is the most common Alpine chamois *DRB* allele, with an overall frequency of 0.297 (Schaschl et al. 2004), while the *Rupy-DRB02* is one of the most common Pyrenean chamois allele with a frequency of 0.222. Among the Pyrenean chamois subspecies, *R. p. pyrenaica* and *R. p. parva*, five *DRB* alleles were identified in this study plus six alleles previously. In contrast in the Pyrenean chamois subspecies *R. p. ornata* from the Italian population, only two *DRB* (*Rupy-DRB12* and *Rupy-DRB13*) alleles have been detected, which were not found in the Pyrenean chamois samples from Spain. Hence the *DRB* alleles from the Apennine subspecies may represent novel alleles not present in the ancestral population.

The nucleotide variation in Pyrenean chamois is slightly higher than in Alpine chamois, a pattern that is repeated for other types of markers. In allozyme surveys Pyrenean chamois have marginally higher values of allozyme diversity than Alpine chamois from the Eastern Alps (Nascetti et al. 1985; Schaschl et al. 2003). A striking feature of the pattern of sequence variation in both chamois species is the low level of silent variation. This suggests a disproportionate loss of silent variation in species that may have been caused by one ore more bottlenecks, sometime during their population history. Similar observations and conclusions have been recorded for European and North American moose (*Alces alces*) populations by Mikko and Andersson (1995), for Madagascan lemur species (Go et al. 2002), and also for deer mouse (*Peromyscus*) species (Richman et al. 2003a). Chamois likely have experienced a complex demographic history, probably governed by multiple processes acting over ancient and contemporary time scales, generating changes in population size. This includes population expansions and contractions associated with Pleistocene glacial cycles, contemporary habitat fragmentation and reduction in population size due to human hunting pressure, and disease epizootics.

Population bottlenecks that reduce levels of allelic diversity at the MHC do not appear to be uncommon (Hedrick et al. 2000; Mikko and Andersson 1995; Mikko et al. 1999; Schmulder et al. 2003). Following such population bottlenecks, intragenic recombination provides a mechanism that could regenerate allelic diversity rapidly (Andersson and Mikko 1995).

In has been shown for the human genome that recombination rate occurs in a higher frequency in non-coding region than within genes (McVean et al. 2004). McVean et al. (2004) found also that 80% of recombination occurs in less than 10% of the human MHC sequence. Consequently, the human MHC is thought to be a recombination hotspot. Thus, density and intensity of recombination rate might be optimised for different sections in the genome. This study revealed incompatible sequence blocks in the sequences and a significant ($P<0.05$) decay in linkage disequilibrium with distance in several chamois *DRB* alleles, suggesting recombination in the MHC class II *DRB* gene. In both chamois species the estimated population recombination rate (ρ) differs significantly ($P<0.001$) from that expected under the null hypothesis of no recombination. In addition, the estimated recombination rate exceeds the estimated mutation rate (θ) by an order of magnitude. This indicates that the accumulation of new recombinant alleles greatly exceeds that of alleles derived from new point mutations and that intragenic recombination may have an adaptive significance in the evolution of the MHC class II *DRB* gene. Apparently, within exon 2, two DNA segments (positions 98–196 and positions 147–234) may have acted as hot spots for recombination. These DNA blocks were found in several alleles in either chamois species. Putative intragenic recombination events have taken place also between the two sets of species–specific alleles, suggesting frequent segmental sequence exchanges among the *DRB* alleles in the common history of the genus *Rupicapra*. The very high population recombination rate (ρ) in either chamois species indicates also that in their short intraspecific evolutionary time a rapid accumulation of novel alleles generated by recombination has taken place and the origin of some of these alleles might be relative recent. This fact also has consequences for phylogenetic inferences made from MHC sequence data, because the occurrence of recombination means that different parts of the alleles have different phylogenies. Schierup and Hein (2000) showed that with recombination the length of terminal branches, and the total branch lengths are larger and the time to the most recent common ancestor is smaller than for a phylogenetic tree reconstructed with no recombination. Therefore, conclusions based on MHC class II gene phylogenies should be considered with caution (e.g. trans-species polymorphism).

Finally, positive selection was determined for the exon 2 in both chamois species. In the Pyrenean chamois sequences about 23% of the sites were identified to be under strong positive selection ($w=20.42$), whereas in the Alpine chamois sequences, only 13% of sites were found to be under strong positive selection ($w=15.23$). We used only the M7 model, which acts as the null model versus the alternative model M8 (Yang et al. 2000). It has been shown that these two models are much more robust against the occurrence of recombination in the sequences than the other implemented models in CODEML (Anisimova et al. 2003). Furthermore, the Bayes' prediction of sites under positive selection appears to be robust to recombination effects (Anisimova et al. 2003). The sites that were identified as positively selected sites were mostly in accordance with the putative human PBR sites (Brown et al.

1993). In the Pyrenean chamois sequences 14 sites were identified as positively selected sites out of 22 putative PBR sites. From those 14 sites only two (amino acid position 41 and 57) are not considered as putative PBR sites. In Alpine chamois only seven PBR sites were found to be under strong diversifying selection (Fig. 1; Table 4). However, the positively selected PBR sites shared between both species are identical.

In summary, the study revealed that the contribution of intragenic recombination for generating sequence polymorphism in the chamois *DRB* gene is about ten times high than point mutations. Thus, intragenic recombination coupled with strong positive selection are the main forces in generating sequence diversity in the MHC class II *DRB* gene.

Acknowledgements We would like to thank TBH Reusch from the Max-Planck Institute of Limnology for helpful comments on a previous version of this manuscript and A Hasselmeyer for assistance in sequencing.

References

Andersson L, Mikko S (1995) Generation of MHC class II diversity by intra- and intergenic recombination. Immunol Rev 143:5–12

Anisimova M, Nielsen R, Yang Z (2003) Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics 164:1229–1236

Apanius V, Penn D, Slev PR, Ruff LR, Potts WK (1997) The nature of selection of the major histocompatibility complex. Crit Rev Immunol 17:179–224

Awadalla P, Eyre-Walker AA, Maynard Smith J (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. Science 286:2524–2525

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

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 64:33–39

Fearnhead P, Donnelly P (2001) Estimating recombination rates from population genetic data. Genetics 159:1299–1318

Go Y, Satta Y, Kawamoto Y, Rakotoarisoa G, Randrianjafy A, Koyama N, Hirai H (2002) *Mhc-DRB* genes evolution in lemurs. Immunogenetics 54:403–417

Gutierrez-Espeleta G, Hedrik PW, Kalinowski ST, Garrigan D, Boyce WM (2001) Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation? Heredity 86:439–450

Gyllensten UB, Sundvall M, Erlich HA (1991) Allelic diversity is generated by intraexon sequence exchange at the DRB1 locus of primates. Proc Natl Acad Sci USA 88:3686–3690

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor an analysis program for Windows 96/98/NT. Nucleic Acids Symp Ser 41:95–98

Hammer S, Nadlinger K, Hartl GB (1995) Mitochondrial DNA differentiation in chamois (genus *Rupicapra*): implication for taxonomy, conservation, and management. Acta Theriol [Suppl 3]:145–155

Hedrick PW, Parker KM, Gutierrez Espeleta GA, Rattink A, Lievers K (2000) Major histocompatibility complex variation in the Arabian oryx. Evolution 54:2145–2151

Hudson RR (2001) Two-locus sampling distributions and their application. Genetics 159:1805–1817

Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet 32:415–435

Janeway CA, Travers P, Walport M Capra JD (2001) Immunobiology: the immune system in health and disease, 5th edn. Garland, Edinburgh

Jugo BM, Vicario A (2000) Single-strand conformational polymorphism and sequence polymorphism of *Mhc-DRB* in Latxa and Karrantzar sheep: implication for Caprinae phylogeny. Immunogenetics 51:887–897

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic, pp 21–32

Karlin S, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequences by using general scoring schemes. Proc Natl Acad Sci USA 87:2264–2268

Klein J, O'hUigin C (1995) Class IIB Mhc motifs in an evolutionary perspective. Immunol Rev 143:89–111

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

Klein JN, Takahata N, Ayala FJ (1993) Mhc diversity and human origins. Sci Am 269:46–51

Kumar S, Tamura K, Jokobsen I, Nei M (2001) MEGA 2.1: Molecular evolutionary genetics analysis software. Version 2.1. Arizona State University, Tempe

León-Vizcaíno L, Ruíz de Ybáñz MR, Cubero MJ, Ortíz JM, Espinosa J, Pérez L, Simón MA, Alonso F (1999) Sarcoptic mange in Spanish Ibex from Spain. J Wildlife Dis 35:647–659

Lovari S (1987) Evolutionary aspects of the biology of chamois, *Rupicapra* spp. (Bovidae, Caprinae). In: Soma H (ed) The biology and management of *Capricornis* and related mountain antelopes. Croom–Helm, London, pp 51–61

Masini F, Lovari S (1988) Systematics, phylogenetic relationship and dispersal of the chamois (*Rupicapra* spp.) Quaternary Res 30:339–349

McVean GAT, Awadalla P, Fearnhead P (2002) A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics 160:1231–1241

McVean GAT, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P (2004) The fine-scale structure of recombination rate variation in the human genome. Science 304:581–584

Mikko S, Andersson L (1995) Low major histocompatibility complex class II diversity in European and North American moose. Proc Natl Acad Sci USA 92:4259–4263

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

Nascetti G, Lovari S, Lanfranchi P, Berducou C, Mattiucci S, Rossi L, Bullini L (1985) Revision of *Rupicapra* genus III: electrophoretic studies demonstrating species distinction of chamois populations of the Alps from those of the Apennines and Pyrenees. In: Lovari S (ed) Biology and management of mountain ungulates. Croom–Helm, London, pp 57–62

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, Oxford

Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929–936

O'hUigin C (1995) Quantifying the degree of convergence in primate *Mhc-DRB* genes. Immunol Rev 143:123–140

Pérez TM, Albornoz J, Domínguez A (2002) Phylogeography of chamois (*Rupicapra* spp.) inferred from microsatellites. Mol Phylogenet Evol 25:524–534

Posada D (2002) Evaluation of methods for detecting recombination from DNA sequences: empirical data. Mol Biol Evol 19:708–717

Richman AD, Herrera LG, Nash D (2003a) Evolution of MHC class II Eβ diversity within the genus *Peromyscus*. Genetics 164: 197–289

Richman AD, Herrera LG, Nash D, Schierup HM (2003b) Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in *Peromyscus maniculatus*. Genet Res Camb 82:89–99

Robinson J, Waller MJ, Parham P, de Groot N, Bontrop R, Kennedy LJ, Stoehr P, Marsh SGE (2003) IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res 31:311–314

Rossi L, Meneguz PG, De Martin P, Rodolfi M (1995) The epizootiology of sarcoptic mange in Chamois *Rupicapra rupicapra*, from the Italian Eastern Alps. Parassitologia 37:233–240

Sägesser H, Krapp F (1986) *Rupicapra rupicapra* (Linnaeus, 1758), Gämse In: Niethammer J, Krapp F (eds) Handbuch der Säugetiere Europas. Aula Verlag, Wiesbaden, pp 316–348

Sawyer SA (1989) Statistical tests for detecting gene conversion. Mol Biol Evol 6:526–538

Sawyer SA (1999) GENECONV: a computer package for statistical detection of gene conversion. Available at http://www.math.wustl.edu/~sawyer/mbprogs/.

Schaschl H, Kaulfus D, Hammer S, Suchentrunk F (2003) Spatial patterns of mitochondrial and nuclear gene pools in Chamois (*Rupicapra r. rupicapra*) from the Eastern Alps. Heredity 91: 125–135

Schaschl H, Goodman SJ, Suchentrunk F (2004) Sequence analysis of the MHC class II DRB alleles in Alpine chamois (*Rupicapra r. rupicapra*). Dev Comp Immunol 28:265–277

Schierup MH, Hein J (2000) Consequences of recombination on traditional phylogenetic analysis. Genetics 156:879–891

Schmulder MJM, Snoek LB, Booy G, Vosman B (2003) Complete loss of MHC genetic diversity in the common hamster (*Cricetus cricetus*) population in The Netherlands: consequences for conservation strategies. Conserv Genetics 4:441–451

Shackleton DM (1997) Wild sheep and goats and their relatives: status survey and conservation action plan for Caprinae. IUCN/SSC Caprinae specialist group, Gland

Takahata N, Satta Y (1998) Selection, convergence, and intragenic recombination in HLA diversity. Genetica 157:157–169

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Van Den Bussche RA, Hoofer SR, Lochmilr RL (1999) Characterisation of *Mhc-DRB* allelic diversity in white-tailed deer (*Odocoileus virginianus*) provides insight into *Mhc-DRB* allelic evolution within Cervidae. Immunogenetics 49:429–437

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Cabios 13:555–556

Yang Z, Nielsen R, Goldman N, Pedersen A-MK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155:431–449