*Immunogenetics 35:* 332–340, 1992

*Immuno-genetics*



# Exonic polymorphism vs intronic simple repeat hypervariability in *MHC-DRB* genes

**Hubert Ammer[1], Franz-Werner Schwaiger[1], Clauda Kammerbauer[1], Maria Gomolka[1], Annette Arriens[2], Sandor Lazary[2], and Jörg T. Epplen[1]**

[1] Max-Planck-Institute for Psychiatry, W-8033 Martinsried, Federal Republic of Germany
[2] Institute of Animal Breeding, School of Veterinary Medicine, University of Bern, CH-3012 Bern, Switzerland

Received August 8, 1991

**Abstract.** Gene products encoded by the major histocompatibility complex often exhibit a high degree of polymorphism. In humans the *HLA-DR* polymorphism is due to more than 50 alleles with varying exon 2 sequences. Each group of *DRB* alleles contains a certain form of the basic simple repeat motif $(gt)_n(ga)_m$ in intron 2. Identical alleles can be differentiated on the basis of the hypervariable repeat. In this study focused on cattle (*Bos taurus*) we identified different *Bota-DRB* alleles in a limited survey by amplification via polymerase chain reaction and sequencing. In addition *DRB* exon 2 sequences were also obtained from eight additional hoofed animal species (seven horned artiodactyls and one pig) revealing artiodactyl-specific polymorphic and nonpolymorphic substitutions. In the genus *Bos* the intronic simple repeat variability was compared with exonic *DRB* polymorphism. As in humans all *Bota-DRB* exons were always associated with specifically organized basic simple repeat structures. Yet the extent of simple repeat variability was lower in cattle compared to humans. Selective breeding in the process of domestication might be responsible for the diminished intronic hypervariability. Nevertheless, the hypermutable simple repeat sequences have been preserved in the same position and with the same principal structure for at least $70 \times 10^6$ years of evolution. Unexpectedly, the rate of intronic simple repeat and exonic changes appear quite similar.

## Introduction

The major histocompatibility complex (MHC) of vertebrates encodes molecules which play a crucial role in antigen presentation to T lymphocytes (Zinkernagel and Doherty 1979). MHC class I and class II molecules are transmembrane glycoproteins expressed on the cell surface (Klein 1986). Class II genes are among the most polymorphic genes in humans (Bodmer et al. 1990). The reasons for this high degree of polymorphism are not fully understood, but may relate to differences in the ability of the different molecules to bind different peptides (Schwartz 1986). The explanation for this excessive polymorphism remains of paramount importance for the understanding of the interacting structures in the trimolecular complex of MHC molecule, antigen, and T-cell receptor (Tcr). Almost all of the variable residues in the heterodimeric class II *DR* molecules are concentrated in the $\beta$ chain (Marsh and Bodmer 1990). These residues have been proposed to be in close contact with either the peptide or the Tcr (Peccoud et al. 1990). Recently the three-dimensional structures of two MHC class I molecules have been resolved (Bjorkman et al. 1987; Garrett et al. 1989) and a model of the class II counterpart has been proposed (Brown et al. 1988). In this model the majority of the variable amino acids reside in a groove, which has been proposed to represent the binding site for peptides.

So far only the human and mouse MHC has been studied in sufficient detail with respect to the extent of polymorphism of the different loci (Marsh and Bodmer 1990; She et al. 1991). Recently, however, a number of bovine MHC (*Bota* complex, *Bota-DRB3**) sequences have been described (Andersson et al. 1991).

We investigated *DRB* exon 2 sequences in a number of artiodactyl species. In both, humans and cattle, intron 2 of the *DRB* genes contains simple repeat sequences with the principal motif $(gt)_n(ga)_m$ (Rieß et al. 1990; Groenen et al. 1990). Normally such elements evolve rapidly (Nürnberg et al. 1989; Nanda et al. 1990; Stallings et al. 1991; Epplen et al. 1991a), but the simple repeat in the

*Address correspondence and offprint requests to:* J. T. Epplen, Molekulare Humangenetik der Ruhr-Universität, Universitätsstr. 150, W-4630 Bochum, FRG.

*DRB* intron 2 seems to have been preserved for at least 70 million years. The origin as well as the functions, if any, of simple repetitive elements are largely unknown (Epplen et al. 1991b). In an earlier publication we compared exonic polymorphism in *HLA-DRB* genes with the intronic hypervariability of a mixed simple repetitive $(gt)_n(ga)_m$ element (Rieß et al. 1990). Many additional simple repeats from humans analyzed in the meantime substantiate the two main conclusions drawn previously. First, there is a group-specific organization of the simple repetitive elements, so that for example, *DRw52* exon 2 sequences can readily be identified by the basic structure of their adjacent repeat. Second, the number of dinucleotide units varies in the simple repeats associated with the same exonic sequence from unrelated individuals. Here, we report on the existence of an analogous situation in cattle.

## Materials and methods

*Amplification of bovine DNA by polymerase chain reaction (PCR).* DNA was isolated from peripheral blood leukocytes (PBL) from a panel of unrelated Simmental cattle (Arriens et al. 1991), Ndama, Zebu, Red Pied, Holstein Friesian, Galloway, Dahome, and other artiodactyl species according to Miller and co-workers (1988). All DNAs were characterized by oligonucleotide fingerprinting (W. Schwaiger, unpublished data). The oligonucleotide primers for PCR were synthesized on an Applied Biosystems synthesizer 381A (Applied Biosystems, Foster City, CA). Primer 1 covers part of the exon 2/intron 2 boundary as determined by Groenen and co-workers (1990); (5′-TCTCTGCAGCACATTTCCTGG-3′). Primer 2 (5′-CAGTCGTACACACTCGACCCA-3′) attaches in the intron 2 sequence. In addition, we designed an internal sequencing primer complementary to the bovine 3′ exonic sequence: 5′-CTCGCCGCTGCACAGTGAAAC-3′. Primers were directly used after detritylation. PCR was carried out using *Taq* polymerase according to the manufacturer's instructions (Perkin Elmer Cetus, Norwalk, CT) usually for 25 cycles at an annealing temperature of 55 °C. Prior to amplification samples were treated with ultraviolet (UV) light according to Sarkar and Sommer (1990) to avoid contaminations.

*Oligonucleotide hybridizations.* The amplification products were separated by electrophoresis on 2% agarose gels and after drying the gels hybridized in situ with the $^{32}$P-labeled internal oligonucleotide 5′-GTCCTTCTGGCTGTTCCAGTA-3′ in 5 × saline-sodium phosphate-EDTA (SSPE), 5 × Denhardt's solution, 0.1% sodium dodecyl sulfate (SDS), 10 µg/µl *Escherichia coli* DNA at 59 °C for 1–2 h (For the bovine-related species the hybridization temperature was occasionally lowered by 5–10 °C in order to allow for 1–2 mismatches). The gels were washed in 6 × standard sodium citrate (SSC) at room temperature for about 3 h, including a 1 min wash at 59 °C in 6 × SSC. Electroblotting and blot hybridization were carried out according to Rieß and co-workers (1990).

*Molecular cloning and sequence analysis.* All routine molecular biological techniques essentially followed the established protocols (Sambrook et al. 1989). Amplification products in a relative mass range of 450 base pairs (bp) were isolated using diethylaminoethanol (DEAE) paper. The PCR products were blunted using T4 DNA polymerase and phosphorylated using T4 polynucleotide kinase (New England Biolabs, Beverly, MA). The fragments were ligated overnight at 16 °C into the *Sma* I-digested and dephosphorylated pUC19 vector. Following transformation, recombinant clones were identified after transfer on Hybond N (Amersham, Amersham, UK) filters by hybridization with the $^{32}$P-labeled internal oligonucleotide. Both strands of positive clones were sequenced using the Sequenase kit (US Biochemicals, Cleveland, OH) and $\alpha^{35}$S-dATP. In order to screen for any differences in the clones obtained from each individual, all plasmids were *Pst* I-digested and "A-tracked". For computational analyses of evolutionary relationships the maximum parsimony method was used (PHYLIP 3.4; Felsenstein 1988). Alteration from an occupied site to a deletion or vice versa was counted as one change. Since the currently available alignment programs do not meet the requirements necessitated by the completely different evolutionary mechanisms of simple repeat slippage (vs the conventional insertions/deletions), a novel strategy for the alignment of intron 2 simple repeats had to be devised. Since simple repeats expand and contract very frequently, their alignment was achieved by introducing gaps between the perfect simple repeat stretches without penalties. Alignments were optimized by sequentially calculating scores according to the number of the perfectly matching simple repeat nucleotides (F.-W. Schwaiger and H. Zucker, unpublished data). This evaluation was accomplished by using a universal data handling program ("Autesp") developed by H. Zucker (Max-Planck-Institute for Psychiatry, Martinsried).

## Results

Our initial attempts to amplify *Bota-DRB* exon 2 and the adjacent introns failed due to erroneous sequence data (Groenen et al. 1990). After resequencing the original clones (kindly supplied by M. A. M. Groenen) and correcting the essential intronic primer attachment site for amplification, we could also amend the sequence of the simple repetitive element. The correct simple repeat $[(gt)_{20}(ga)_9\ (ca(ga)_2)_4\ taga\ (caga)_2]$ showed less similarity to the corresponding human sequence than reported previously (Groenen et al. 1990). After having established the optimal amplification conditions for *Bota-DRB* sequences, the PCR products from the genomes of different cattle specimen were subcloned. Clones which hybridized to the internal probe were characterized by sequence analysis. We obtained amplification products from all bovine DNA samples tested. In certain cases where many amplification products were analyzed, up to four different *Bota-DRB* sequences from one individual were identified. So far we can not assign them to the respective *DRB1*, *−2*, or *−3* loci. Comparisons of several *Bota-DRB* sequences with the corresponding human exons revealed sequence similarities at the nucleotide level of up to 83%. Amino acid replacements were limited to particular residues. Comparison of the *Bota-DRB* exonic sequences with a human consensus sequence (Fig. 1) revealed two bovine-specific residues at positions 34 (G) and 44 (W). Remarkably, only four of the 22 exon 2 sequences are identical to those in the study of Andersson and coworkers (1991): *Bota-DRB01*=DRB3*1A; *Bota-DRB13*=DRB3*5; *Bota-DRB17*=DRB3*3; *Bota-DRB20*=DRB3*13A.

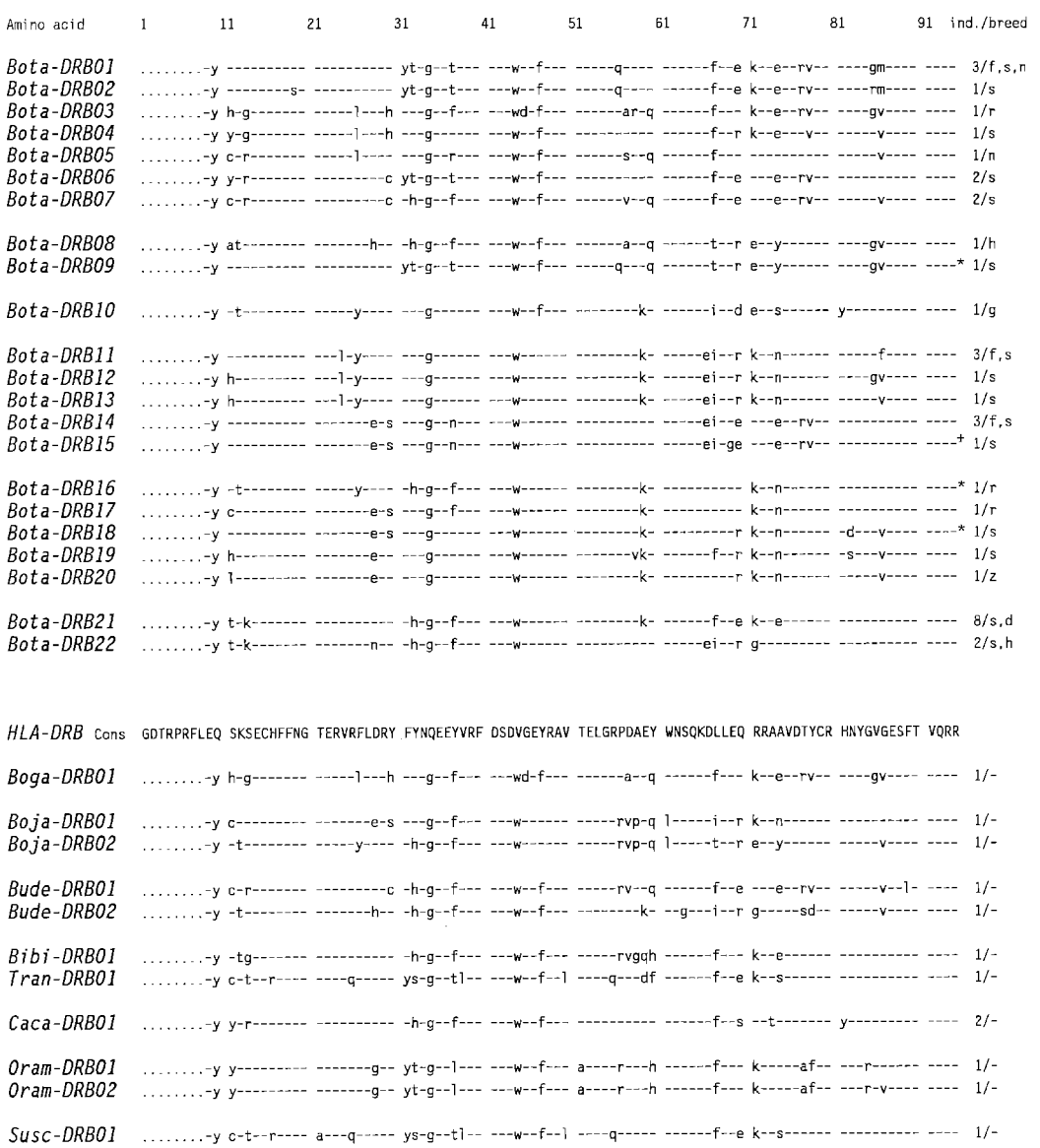
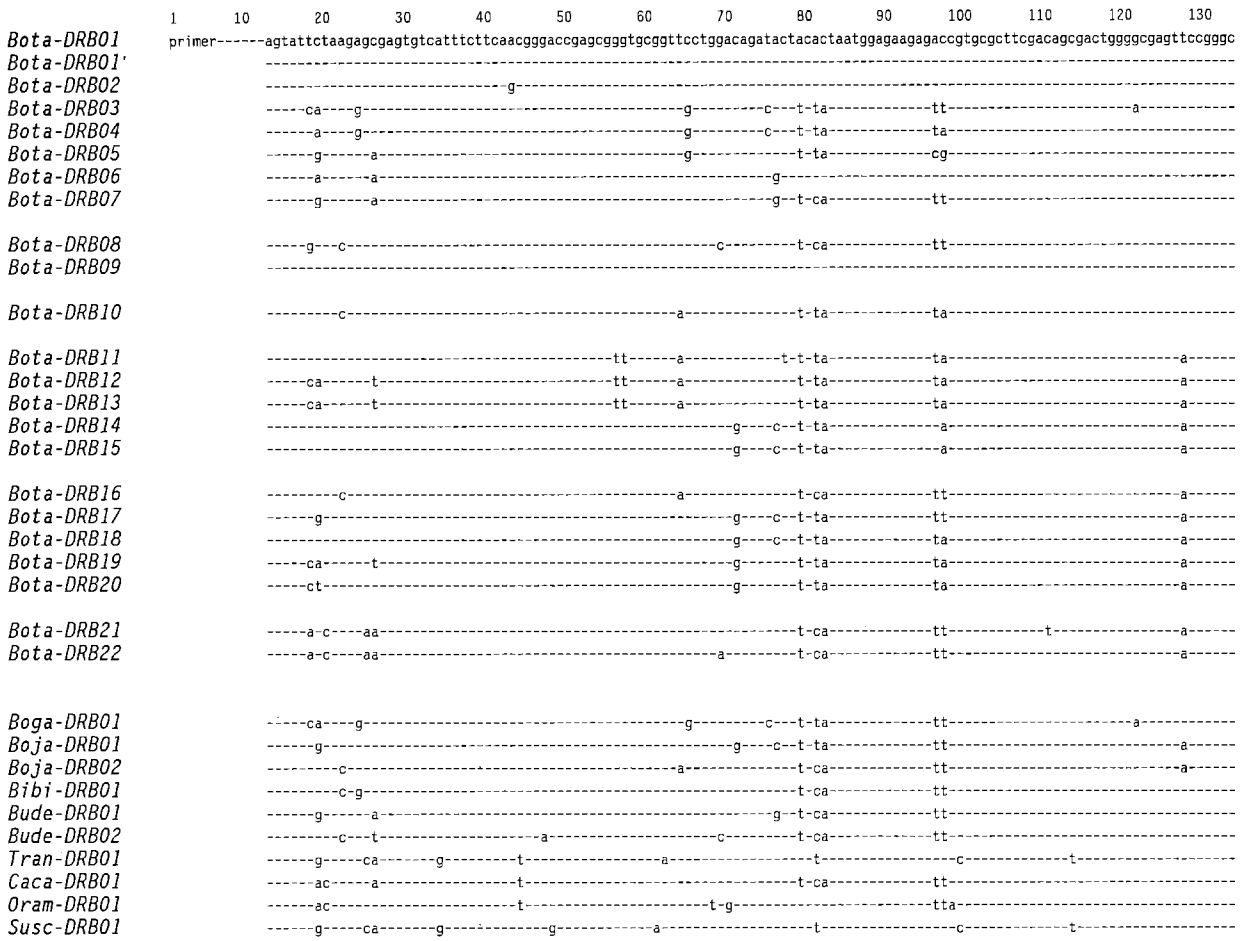



**Fig. 1.** Translated exon 2 sequences of *Bota-DRB* genes compared with human consensus *HLA-DRB*. Usually several clones were investigated from the number of individuals and the breeds given in the right hand columns. The sequences of *Bota-DRB09*, *16*, and *18* were identified only once. *DRB14* represents a possible PCR artifact since the simple repeats are completely identical with *DRB13*. d=Dahome; f=German Simmental; g=Galloway; r=Red Pied; s=Swiss Simmental; h=Holstein Friesian; n=Ndama; z=Zebu.

*Comparisons of HLA-DRB and Bota-DRB sequences with regard to highly polymorphic residues.* Comparison of the *Bota-DRB* sequences indicates that several amino acid positions varied considerably among the sequences of the individuals investigated: **11**, 12, **13**, 26, **28**, **30**, 31, 32, **37**, 47, **57**, 58, 60, **67**, **70**, **71**, **74**, 77, 78, 81, 85, and **86** (positions with four variant residues are in bold face, those with even more variants are in addition underlined). These positions correspond to the polymorphic amino acid positions in humans (Marsh and Bodmer 1990) and those from cattle published recently (Andersson et al. 1991). In addition, other polymorphic residues in *Bota-DRB* exon 2 are apparently invariant in *HLA-DRB*, i. e., the positions 24, 56, and 66. None of the *Bota-DRB* alleles showed a significantly higher similarity to any of the known *HLA-DRB* sequences. This is in contrast to the results deduced from two cDNA sequences from the miniature swine (Gustafsson et al. 1990). Quite analogous to this study, in artiodactyl *DRB* exon 2 the silent nucleotide substitutions are about five times more abundant in the inter-species vs the intra-species comparison (Fig. 2).

*Intronic hypervariability versus exonic polymorphism.* The analysis of intron 2 showed hypervariability of the




*Bota-DRB*-associated $(gt)_n(ga)_m$ elements (Fig. 3) similar to that observed in humans (Rieß et al. 1990): Identical *HLA-DRB* exon 2 sequences from unrelated individuals were associated with length variation in the repetitive elements of intron 2. However, this association is considerably less pronounced in cattle than in humans (e. g., *Bota-DRB21*, where eight individuals from two breeds were analyzed). Interestingly, in different cattle breeds the same exonic sequence also occurs with the same simple repeat stretch. In both species, humans and cattle, the simple repetitive elements are inherited according to Mendelian principles as revealed by electroblot hybridization (Rieß et al. 1990, and data not shown). The same cattle exon 2 sequence was associated with only slight variation of the $(gt)_n$ and $(ga)_m$ repeats. With only one exception the simple repetitive element always started strictly 34 bases after the exon/intron boundary, invariably with the conserved sequence ccagtgtggagca.

In cattle the simple repetitive elements are generally less perfectly repeated compared to humans, as indicated for example by "degenerate" simple repetitive parts like gcgtgc and $([ca(ga)_2]_4$; Fig. 3). In analogy to humans, however, a group-specific organization of the bovine simple repeat stretches was also evident. Approximately 2% difference in the single copy spacer region between exon 2 and the intronic simple repeat was observed (22 sequences per 34 bp). Surprisingly, the grouping of the alleles according to the basic structure of intronic simple repeats (Fig. 3) resulted in a nearly identical classification as that obtained by using the exonic and deduced amino acid sequences. Even more important was the fact that after suitable alignment according to the criteria described in Materials and methods, the PHYLIP 3.4 programs yielded very similar (unrooted) trees for the translated exon 2 and simple repeat sequences (Fig. 4).

*DRB exon 2 sequences of other artiodactyls.* The same PCR procedure as developed for *Bota-DRB* sequences was also successfully applied to amplify *DRB* exonic and intronic repeats from the gaur (*Bos gaurus*), the banteng (*Bos javanicus*), the bison (*Bison bison*), nyala (*Tragelaphus angasi*), and from the species that supposedly retained most morphological characteristics of the probable evolutionary forerunner of *Bovinae*, the anoa

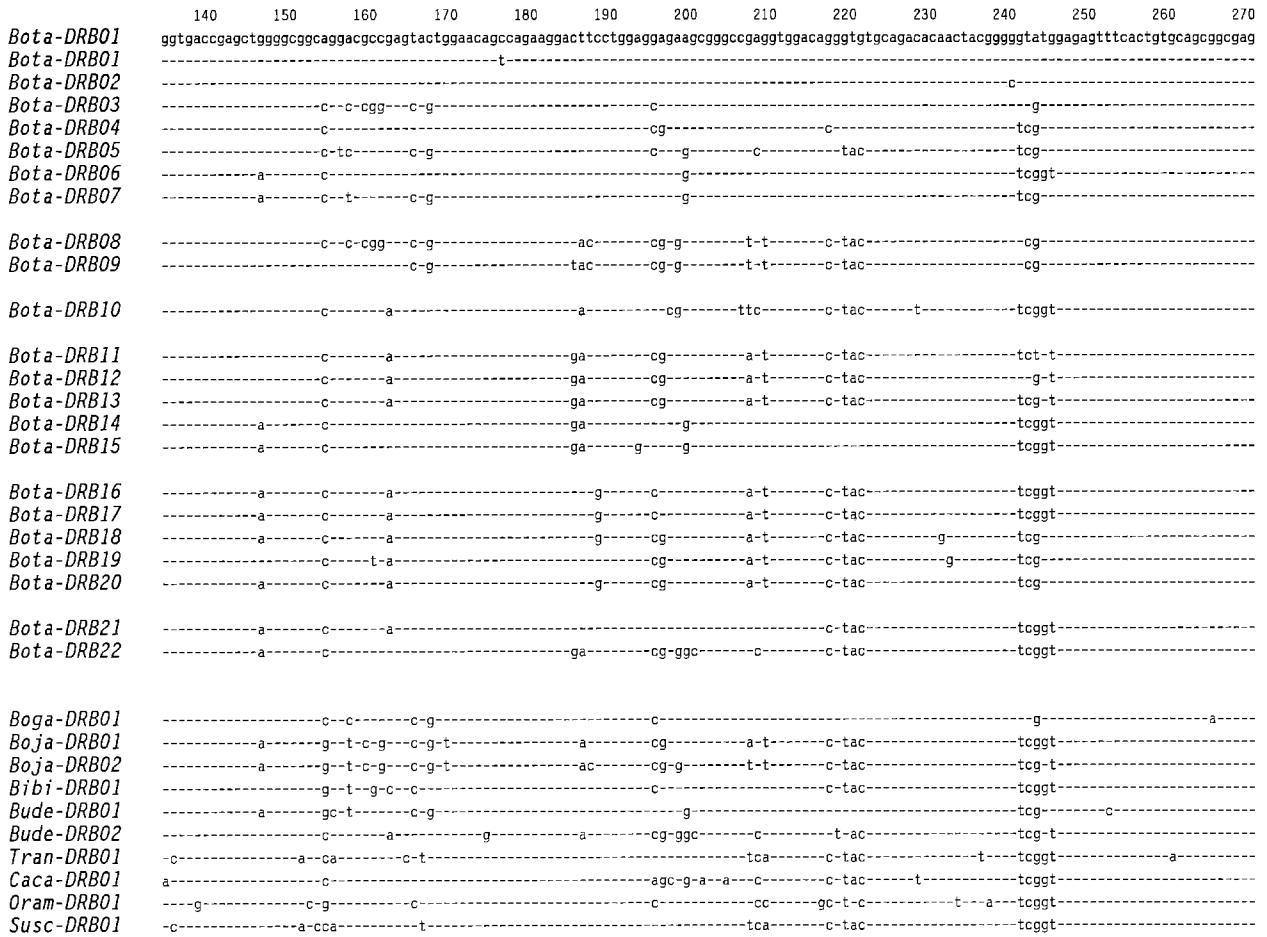



**Fig. 2.** Nucleotide sequences of *Bota-DRB* exons 2. Both strands were sequenced in all cases.

(*Bubalus depressicornis*). In all cases we obtained characteristic amplification products containing exon and intron sequences in the range of 450 bp. In addition, using two exonic primers we obtained amplification products from goat (*Capra capra*), mountain goat (*Oreamnus americanus*), and a wild boar (*Sus scrofa*). As in cattle, the amino acids G and W appear in these artiodactyl species at positions 34 and 44. Unexpectedly, the derived *Susc-DRB* sequence differs considerably from that described for the miniature swine (Gustafsson et al. 1990).

The other species of the genus Bovidae share the artiodactyl-specific residues at positions 34 and 44 and the Y at position 10. Since it was not possible to obtain unrelated individuals from the gaur, banteng, anoa, bison, or nyala no conclusions could be drawn on the *DRB* variability of these species. All exon 2 sequences were found highly similar to those of cattle, although they contained several residues or combinations not found in *Bota-DRB* alleles. Again, specific compositions of the adjacent simple repeat elements were observed, each being slightly more similar to cattle than to humans.

## Discussion

In this report the variation in *DRB* genes has been analyzed by sequencing PCR-amplified DNA from different artiodactyl species, cattle breeds, and unrelated Simmental individuals. A comparison of the translated protein sequences with the human consensus sequence showed that high variability was concentrated to the same positions in cattle and humans. At a given position several of the highly variable *Bota-DRB* amino acid residues occur also in *HLA-DRB* proteins. Other substitutions are species-specific, suggesting the existence of different *DRB* gene pools with which the species initiated their evolution or reflecting the different pathogens relevant for the species. All artiodactyl $\beta$ chains have a Y at position 10 as do human w52 group chains. None of our bovine alleles showed a significant higher similarity to any of the human alleles. This finding stands in stark contrast to a recent report on a limited number of porcine (*Susc-DRB*) cDNA sequences (Gustafsson et al. 1990): Higher homologies in the transspecies comparison (humans vs minipig) has

| Alleles | Simple repeat motifs | | | | | |
|---|---|---|---|---|---|---|
| *Bota-DRB01* | $(gt)_{20}$ | $(ga)_9$ | | | $[ca(ga)_2]_4taga$ | $(caga)_2$ |
| *Bota-DRB01'* | $(gt)_{20-21}$ | $(ga)_9$ | | | $[ca(ga)_2]_4taga$ | $(caga)_2$ |
| *Bota-DRB02* | $(gt)_{18}$ | $(ga)_9$ | | | $[ca(ga)_2]_4taga$ | $(caga)_2$ |
| *Bota-DRB03* | $(gt)_{17}$ | $(ga)_5$ | | | $[ca(ga)_2]_4taga$ | $(caga)_2$ |
| *Bota-DRB04* | $(gt)_{15}$ | $(ga)_6$ | | $gtgg(ga)_6$ | $[ca(ga)_2]$ | $(caga)_3$ |
| *Bota-DRB05* | $(gt)_{15}$ | $(ga)_6$ | | $gtgg(ga)_6$ | $[ca(ga)_2]$ | $(caga)_3$ |
| *Bota-DRB06* | $(gt)_{11}$ | $(ga)_8$ | | | $[ca(ga)_2]_3$ | $(caga)_3$ |
| *Bota-DRB07* | $(gt)_{12}$ | $(ga)_{10}$ | | | $[ca(ga)_2]_3$ | $(caga)_3$ |
| *Bota-DRB08* | $(gt)_2gcgtttgtgt$ | $ga(gt)_{10}$ | $(ga)_{13}$ | $gtgg(ga)_6$ | | $aa(ga)_3(caga)_2$ |
| *Bota-DRB09* | $(gt)_2gcgtttgtgt$ | $ga(gt)_{10}$ | $(ga)_{12}$ | $gtgg(ga)_6$ | | $aa(ga)_3(caga)_2$ |
| *Bota-DRB10* | $(gt)_2gc$ | $(gt)_{11}gagt$ | $(ga)_9$ | $gt\ gtag(ga)_6$ | $[ca(ga)_2]$ | $(caga)_3$ |
| *Bota-DRB11* | $(gt)_2gc$ | $(gt)_{11}$ | $(ga)_{11}$ | $gt(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2$ |
| *Bota-DRB12* | $(gt)_2gcgtgc$ | $(gt)_{12-13}$ | $(ga)_{12}$ | $gt(ga)_3gtgg(ga)_4$ | $[caggga\ ]aa(ga)_3$ | $(caga)_2$ |
| *Bota-DRB13* | $(gt)_2gcgtgc$ | $(gt)_{13}$ | $(ga)_{10}$ | $gt(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2^+$ |
| *Bota-DRB14* | $(gt)_2gc$ | $(gt)_{14}$ | $(ga)_{6-10}$ | $gt(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2$ |
| *Bota-DRB15* | $(gt)_2gc$ | $(gt)_{14}$ | $(ga)_{10}$ | $gt(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2$ |
| *Bota-DRB16* | $(gt)_2gc$ | $(gt)_{10}$ | $(ga)_{11}$ | $at(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2$ |
| *Bota-DRB17* | $(gt)_2gc$ | $(gt)_{10}$ | $(ga)_{11}$ | $at(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2$ |
| *Bota-DRB18* | $(gt)_2gc$ | $(gt)_7$ | $(ga)_{14}$ | $gg(ga)_{15}$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_3$ |
| *Bota-DRB19* | $(gt)_2gc$ | $(gt)_7$ | $(ga)_{14}$ | $gg(ga)_{14}$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_3$ |
| *Bota-DRB20* | $(gt)_2gc$ | $(gt)_7$ | $(ga)_{14}$ | $gg(ga)_{14}$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_3$ |
| *Bota-DRB21* | $(gt)_2gc\ ttgtgctt$ | $(gt)_{8-9}$ | $gaaa(ga)_{13}$ | $ggga\ gg(ga)_5$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_3$ |
| *Bota-DRB22* | $(gt)_2gc\ tt$ | $(gt)_9$ | $gaaa(ga)_{20-22}$ | $ggga\ gg(ga)_5$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_3$ |
| *Boga-DRB01* | $(gt)_{15}$ | $(ga)_5$ | | | $[ca(ga)_2]_4taga$ | $(caga)_2$ |
| *Boja-DRB01* | $(gt)_{13}$ | $(ga)_{11}$ | | $gt(ga)_3gtgg(ga)_4$ | $[ca(ga)_2]aa(ga)_3$ | $(caga)_3$ |
| *Boja-DRB02* | $(gt)_{12}(ga)_6gg(ga)_3gg(ga)_6gt(ga)_3gtgg(ga)_4$ | | | | $[ca(ga)_2]aa(ga)_3$ | $(caga)_2$ |
| *Bude-DRB01* | $(gt)_{11}$ | $(ga)_{10}$ | | | $[ca(ga)_2]_3$ | $(caga)_3$ |
| *Bude-DRB02* | $(gt)_{13}\ ttgt$ | $(ga)_{24}$ | | | $aa(ga)_2\ aa(ga)_3$ | $(caga)_3$ |
| *Bibi-DRB01* | $(gt)_5$ | $(ga)_9$ | | $ctgg(ga)_6$ | $[ca(ga)_2]$ | $(caga)_3$ |
| *Tran-DRB01* | $(gt)_{18}$ | $aa(ga)_3$ | | $gtgg(ga)_4$ | | $(caga)_1$ |

**Fig. 3.** Intron 2 simple repeats of the *Bota-DRB* genes. Adjacent to the *DRB13* exon a second different simple repeat was obtained probably due to a PCR artifact.

been taken as evidence in favor of the trans-species theory of MHC gene evolution. Nevertheless we found strong support for the latter theory in the *Bota-DRB03* and *Boga-DRB01* alleles as well as in *Bota-DRB07* and *Bude-DRB01* (Fig. 1 and 2).

The high degree of polymorphism observed in class II genes across species barriers appears to be the result of positive selection for variability in the antigen binding region (Hughes and Nei 1988, 1989). Indeed, most of the polymorphic residues we found in *Bota-DRB* sequences code for amino acids of the antigen binding region in the model of Brown and co-workers (1988). These residues are probably either relevant for binding of the antigenic peptide or for the interaction with the Tcr (Peccoud et al. 1990). The presence of identical amino acid replacements among the alleles of different species might be the result of convergent evolution.

It has been hypothesized that some of the expressed polymorphisms in present-day mammals developed very early during mammalian evolution (Klein and Takahata 1990). We identified several novel amino acid replacements which have not been reported in either artiodactyls (Andersson et al. 1991; Gustafsson et al. 1991) or in canine *DRB* sequences (Sarmiento et al. 1990). It is therefore not yet possible to determine whether selection pressure provided by antigens derived from common pathogens might be responsible for the appearance of similar functional amino acid residues in MHC class II molecules. As our approach covers potentially both genes and pseudogenes, we are currently investigating mRNA expression by northern blot hybridizations, reverse transcription (RT), and PCR. The PCR approach revealed that the genes of the reported genomic *Bota-DRB* exons are transcribed in the B-cells of one cow (H. Ammer and

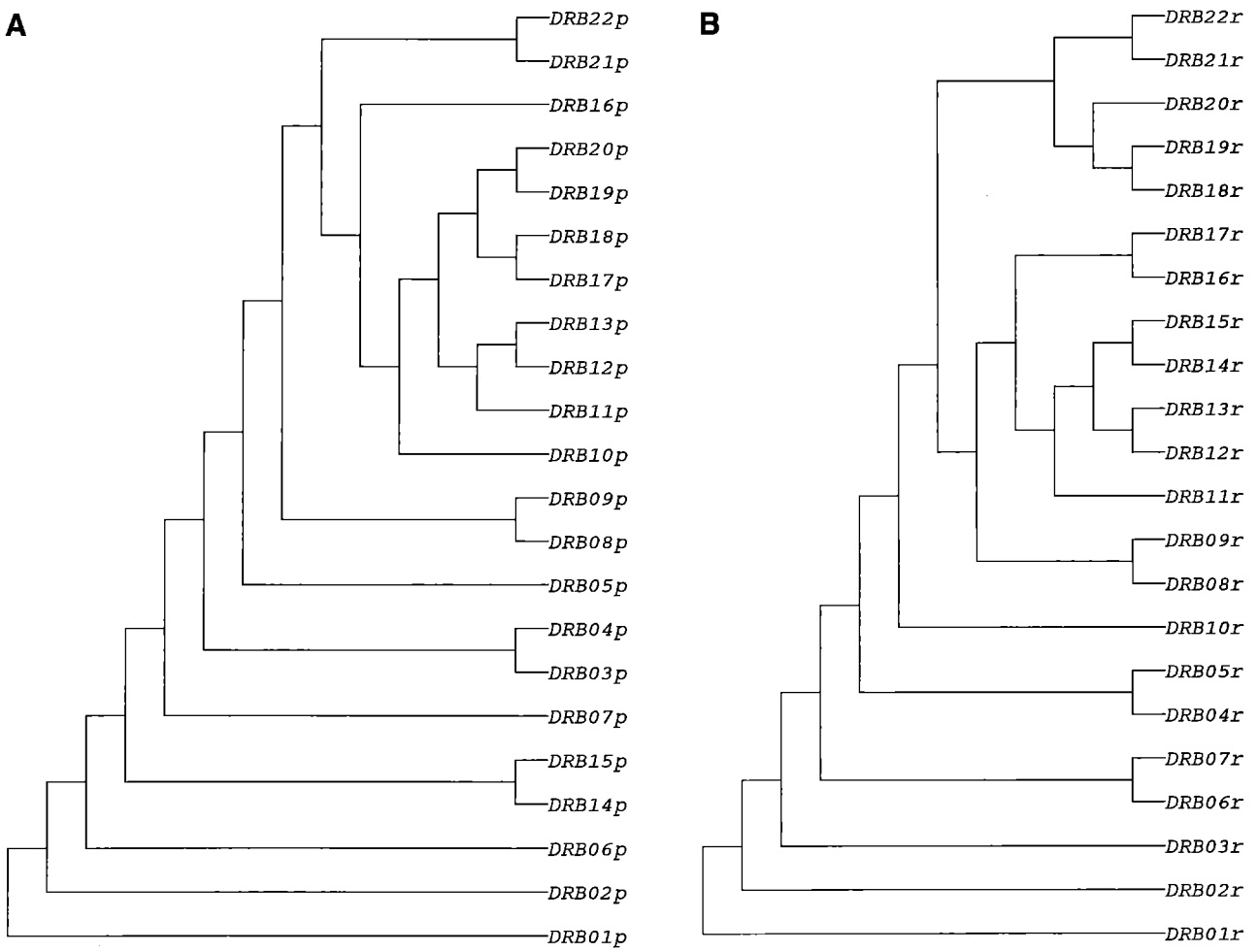



**Fig. 4A, B.** "Unrooted trees" of the *Bota-DRB* exon 2-encoded protein (**A**; *DRBxxp*) and the adjacent simple repetitive sequences (**B**; *DRBxxr*). Sequences were aligned (for description of the simple sequence alignment see Materials and methods) and analyzed by using the maximum parsimony method (PHYLIP 3.4, Felsenstein 1988). In **A** 1 of 2, in **B** 1 of 6 equivalent trees are depicted. Distances are not drawn to any scale of evolutionary time spans.

C. Epplen, unpublished data). Having established the PCR conditions for MHC genes in farm animals it will be possible to test the hypothesis that disease susceptibility is associated with defined alleles in these species (Giphart et al. 1990).

We analyzed also simple repetitive elements in the *Bota-DRB* intron 2. These elements share the hypervariable character with the corresponding sequences at the same position in the *HLA* complex. Intra-experimental variability due to slippage of the *Taq* polymerase was observed (by electroblotting of the amplified exons and internal oligonucleotide hybridization). Sequence analyses revealed the presence of amplification products differing in the number of their simple repeat units – a well known phenomenon with simple repeat motifs of 2–4 bases (Rieß et al. 1990). Only the longest stretch is reported here since it reflects the original template as revealed from PCR experiments of cloned DNA.

As in humans, identical or very similar *Bota-DRB* alleles from different individuals possess identical (or very similar) simple repetitive elements. This finding suggests that whenever a certain type of simple repetitive element is associated with a certain *Bota-DRB* allele, then both elements do not appear to evolve independently. The degree of "hypervariability" in the bovine simple repeat stretches adjacent to the same exon was considerably less in comparison to humans (also in animals from different breeds). This could be the result of a high inbreeding coefficient, although the grandparents of most of our Simmental cattle were unrelated according to the breeder's records.

It has been assumed that little purifying selection is operating on the vast bulk of simple repetitive elements [for discussion see Stallings and co-workers (1991); Epplen and co-workers (1991b)]. This contention is also substantiated in part by the extensive repeat differences

in the genus Bos. The common denominator of the simple repeat structure in cattle is an abrupt transition of alternating purine (pu) and pyrimidine residues to pure poly(pu). Multiple expansions and contractions of simple di-, tetra-, and hexanucleotide motifs occur, apparently without any effect on their bearers. The abrupt transition is also present in other species like human, baboon (Rieß et al. 1990), and goat (F.-W. Schwaiger, unpublished data), but it has been modified in the mouse *H-2Eb* intron to $(gt)_{31}(pu)_n$ . . . $(caga)_9$ (Braunstein and Germain 1986).

Within one species, however, coding and noncoding DNA appear to change in their overall structure with comparable speed as evidenced by the concordant phylogenetic trees from exon deduced protein and simple repeat sequences (Fig. 4). Hence the preservation of this simple repetitive element in mammals prompts questions on the functional meaning of this elusive structure. Initial protein binding assays resulted in band shifts for the *MHC-DRB* intron 2 simple repetitive elements indicating specific interaction with yet unidentified nuclear proteins (W. Mäueler, unpublished results).

*Acknowledgments.* We thank Dr. M. A. M. Groenen for providing *Bota-DRB* phages and subclones, Dr. J. Buitkamp and Dr. H. Wiesner (Tierpark Hellabrunn, Munich), and the Rotterdam Zoo for blood samples, Dr. H. Zucker for help with alignments, Dr. C. Epplen for cell culture and Dr. C. Linington, Dr. E. Albert, and Dr. J. Klein for comments on the manuscript. This work was supported by DFG (Ep 7/6-1) and Fresenius AG, FRG.

## References

Andersson, L., Sigurdardottir, S., Borsch, C., and Gustafsson, K.: Evolution of MHC polymorphism: extensive sharing of polymorphic sequence motifs between human and bovine *DRB* alleles. *Immunogenetics 33:* 235–243, 1991

Arriens, M. A., Hesford, F., Ruff, G., and Lazary, S.: Production of alloantibodies against bovine B-lymphocyte antigens. *Animal Genet,* in press, 1991

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C.: Structure of the human class I histocompatibility antigen HLA-A2. *Nature 327:* 512–518, 1987

Bodmer, J. G., Marsh, S. G. E., and Albert, E.: Nomenclature for factors of the *HLA* system. *Immunol Today 11:* 3–10, 1990

Braunstein, N. S. and Germain, R. N.: The mouse *Eβ2* gene: a class II MHC β gene with limited intraspecies polymorphism and an unusual pattern of transcription. *EMBO J 5:* 2469–2476, 1986

Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J., and Wiley, D. C.: A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature 332:* 845–850, 1988

Epplen, J. T., Ammer, H., Epplen, C., Kammerbauer, C., Mitreiter, R., Roewer, L., Schwaiger, W., Steimle, V., Zischler, H., Albert, E., Andreas, A., Beyermann, B., Meyer, W., Buitkamp, J., Nanda, I., Schmid, M., Nürnberg, P., Pena, S. D. J., Pöche, H., Sprecher, W., Schartl, M., Weising, K., and Yassouridis, A.: Oligonucleotide fingerprinting using simple repetitive repeat motifs: a convenient, ubiquitously applicable method to detect hypervariability for multiple purposes. *In* T. Burke, G. Dolf, A. J. Jeffreys, and R. Wolff, (eds.): *DNA fingerprinting: approaches and applications,* pp. 50–69, Birkhäuser, Basel, 1991a

Epplen, J. T., Ammer, H., Kammerbauer, C., Schwaiger, W., Schmid, M., and Nanda, I.: On the meaning of hypervariable simple repetitive simple repeat loci in eukaryote genomes: an initial attempt for a basic theoretical assessment. *Adv Mol Genet 4:* 301–310, 1991b

Felsenstein, J.: Phylogeny from molecular sequences: interference and reliability. *Annu Rev Genet 22:* 521–565, 1988

Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L., and Wiley, D. C.: Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature 342:* 692–695, 1989

Giphart, M. J., Groenen, M. A. M., and van der Poel, J. J.: Homologies between the major histocompatibility complex of man and cattle: consequences for disease resistance and susceptibility. *Vet Quart 12:* 202–209, 1990

Groenen, M. A. M., van der Poel, J. J., Dijkhof, R. J. M., and Giphart, R. J.: The nucleotide sequence of bovine MHC class II *DQB* and *DRB* genes. *Immunogenetics 31:* 37–44, 1990

Gustafsson, K., Germana, S., Hirsch, F., Pratt, K., LeGuern, C., and Sachs, D. H.: Structure of miniature swine class II *DRB* genes: conservation of hypervariable amino acid residues between distantly related mammalian species. *Proc Natl Acad Sci USA 87:* 9798–9802, 1990

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

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

Klein, J.: *Natural History of the Major Histocompatibility Complex.* Wiley, New York, 1986

Klein, J. and Takahata, N.: The major histocompatibility complex and the quest for origins. *Immunol Rev 113:* 5–25, 1990

Marsh, S. G. E. and Bodmer, J. G.: *HLA-DRB* nucleotide sequences, 1990. *Immunogenetics 31:* 141–144, 1990

Miller, S. A., Dykes, D. D., and Polesky, H. F.: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res 16:* 1215, 1988

Nanda, I., Feichtinger, W., Schmid, M., Schröder, J. H., Zischler, H., and Epplen, J. T.: Simple repetitive sequences are associated with differentiation of the sex chromosomes in the guppy fish. *J Mol Evol 30:* 456–462, 1990

Nürnberg, P., Roewer, L., Neitzel, H., Sperling, K., Pöpperl, A., Hundrieser, J., Pöche, H., Epplen, C., Zischler, H., and Epplen, J. T.: DNA fingerprinting with the oligonucleotide probe $(CAC)_5/(GTG)_5$: somatic stability and germline mutations. *Hum Genet 84:* 75–78, 1989

Peccoud, J., Dellabona, P., Allen, P., Benoist, C., and Mathis, D.: Delineation of antigen contact residues on an MHC class II molecule. *EMBO J 9:* 4215–4223, 1990

Rieß, O., Kammerbauer, C., Roewer, L., Steimle, V., Andreas, A., Albert, E., Nagai, T., and Epplen, J. T.: Hypervariability of intronic simple $(gt)_n(ga)_m$ repeats in *HLA-DRB* genes. *Immunogenetics 32:* 110–116, 1990

Sambrook, J., Fritsch, E. F., and Maniatis, T.: *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989

Sarkar, G. and Sommer, S. S.: More light on PCR contamination. *Nature 347:* 340–341, 1990

Sarmiento, V. M., Sarmiento, J. I., and Storb, R.: Allelic variation in the *DR* subregion of the canine major histocompatibility complex. *Immunogenetics 32:* 13–19, 1990

Schwartz, R. H.: Immune response (Ir) genes of the murine major histocompatibility complex. *Adv Immunol 38:* 31–201, 1986

She, J. X., Boehme, S. A., Wang, T. W., Bonhomme, F., and Wakeland, E. K.: Amplification of major histocompatibility complex class II gene diversity by intraexonic recombination. *Proc Natl Acad Sci USA 88:* 453–457, 1991

Stallings, R. L., Ford, A. F., Nelson, D., Torney, D. C., Hildebrand, C. E., and Moyzis, R. K.: Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. *Genomics 10:* 807–815, 1991

Zinkernagel, R. M. and Doherty, P. C.: MHC-restricted T cells: studies on the biological role of the polymorphic major transplantation antigens determining T-cell restriction-specificity, function, and responsiveness. *Adv Immunol 27:* 221–292, 1979