

Molecular study of *Mhc-DRB* in wild chacma baboons reveals high variability and evidence for trans-species inheritance

Elise Huchard · Guy Cowlshaw · Michel Raymond ·
Mylene Weill · Leslie A. Knapp

Received: 31 January 2006 / Accepted: 24 August 2006 / Published online: 5 October 2006
© Springer-Verlag 2006

Abstract The MHC class II genes of many primate species were investigated extensively in recent years. However, while *Mhc-DRB* genes were studied in Old World monkeys such as rhesus macaques, the *Mhc-DRB* of baboons was only studied in a limited way. Because of their close anatomical and physiological relationship to humans, baboons are often used as models for reproduction and transplantation research. Baboons are also studied as a model species in behavioural ecology. Thus, identification of MHC genes would provide a foundation for studies of *Mhc*, biology and behaviour. Here, we describe the use of PCR, cloning, denaturing gradient gel electrophoresis (DGGE) and sequencing to identify *Mhc-DRB* sequences in wild chacma baboons (*Papio ursinus*). We amplified the highly variable second exon of baboon *Mhc-DRB* sequen-

ces using generic DRB primers. To validate and optimize the DGGE protocol, four DNA samples were initially studied using cloning and sequencing. Clones were screened using a novel RFLP approach to increase the number of clones identified for each individual. Results from cloning and sequencing were used to optimise DGGE conditions for *Mhc-DRB* genotyping of the remaining study subjects. Using these techniques, we identified 16 *Paur-DRB* sequences from 30 chacma baboons. On the basis of phylogenetic tree analyses, representatives of the *Mhc-DRB1* and *Mhc-DRB5* loci, and 13 different DRB lineages were identified. Evidence for trans-species inheritance of some *Mhc-DRB* sequences comes from high identity between the new *Paur-DRB* sequences and sequences from *Papio cynocephalus*, *Macaca mulatta* and possibly *Galago moholi*.

Nucleotide sequence data reported are available in the GenBank/EMBL/DDBJ databases under the accession numbers DQ339722–DQ339737.

E. Huchard · M. Raymond · M. Weill
Equipe Génétique et Environnement, Institut des Sciences de l'Evolution (CNRS-UMR 5554), Université Montpellier II (CC 65),
Place Eugène Bataillon,
34095 Montpellier, France

G. Cowlshaw
Institute of Zoology, Zoological Society of London,
Regent's Park,
London NW1 4RY, UK

L. A. Knapp (✉)
Department of Biological Anthropology,
University of Cambridge,
Downing Street,
Cambridge CB2 3DZ, UK
e-mail: lak26@cam.ac.uk

Keywords Old world monkey · Chacma baboon · *Mhc-DRB* · Denaturing gradient gel electrophoresis · Cloning · Sequencing · RFLP

Introduction

Class II major histocompatibility complex (MHC) molecules are heterodimeric cell surface glycoproteins expressed on immunocompetent cells. They play a critical role in the vertebrate immune system by binding “self” and “non-self” peptides and presenting them to T cells, which triggers immune reactions in response to foreign antigens. MHC genes are the most polymorphic loci of all nuclear-encoding genes in vertebrates. High levels of variation are found not only in the number of alleles, but also in the extent of sequence variation between alleles (Hughes and Yeager

1998). In general, the antigen binding sites (ABS) of the MHC molecules display more non-synonymous than synonymous substitutions, thereby changing the amino acid sequence of the ABS. Thus, individuals with different MHC alleles should respond to a variety of different pathogens (Brown et al. 1988, 1993).

In the past two decades, the MHC class II genes of many primate species were investigated extensively (Bodmer et al. 1994; Kenter et al. 1992; Lekutis and Letvin 1995; Marsh and Bodmer 1995; Sliereendregt et al. 1993, 1994). While *Mhc-DRB* genes were studied in the Old World monkeys such as rhesus macaques for many years, the class II MHC of baboons was only studied in a limited way (Alberts 1999; Gaur et al. 1997). Baboons are frequently used as models in medical research, especially in reproduction and transplantation because of their close anatomical and physiological relationship to humans. Moreover, baboons are extensively studied as a model species in evolutionary ecology. Therefore, knowledge of the MHC genes in this species will benefit both biomedical and evolutionary science, the latter of which was closely interested in the link between MHC and behaviour during the last decade (Bernatchez and Landry 2003).

In primates, the class II DP, DQ, and DR sub-regions encode one α and one or more β chains. This heterodimeric structure allows the formation of a pocket, the ABS in which the peptide being presented is held. The pocket consists of a groove formed by two α helices sitting on top of a β -pleated sheet. The antigen binding properties of the MHC molecule are determined by the amino acids lining the antigen pocket (Bjorkman et al. 1987). In humans and in macaques, the DQ region genes are polymorphic. The *Macaca mulatta* (*Mamu*) DPA1 gene is monomorphic, while 15 *Mamu-DPB1* sequences were described (Otting et al. 2000). The organization of the primate *Mhc-DR* region is more complex with a higher degree of polymorphism, although the *Mhc-DRA* gene is invariant in the primate species examined (Klein et al. 1993; Otting et al. 2000). Within the DR sub-region, nine different DRB loci were described in humans (Bodmer et al. 1994). Four of these loci, *DRB1*, *DRB3*, *DRB4* and *DRB5* are functional and capable of presenting peptides to T cells. While all of these functional DRB loci are characterized by allelic variation in the ABS, encoded by the second exon, *HLA-DRB1* exhibits the most extensive degree of polymorphism. The allelic variation is primarily maintained by positive selection pressure because different sequences will result in different peptides binding abilities (Hughes and Yeager 1998). The remaining loci, *DRB2*, *DRB6*, *DRB7*, *DRB8* and *DRB9* are pseudogenes, containing one or more structural abnormalities that render them non-functional. Nucleotide sequence analysis of diverse primate DRB alleles suggests that particular DRB lineages were inherited in a trans-species

manner (Klein 1987) with distantly related primates possessing nearly identical DRB alleles. Functional and orthologous *DRB1*, *DRB3*, *DRB4* and *DRB5* loci were described in several non-human primate species (reviewed in Sliereendregt et al. 1994). Furthermore, the presence of certain lineages, such as *DRB1*01*, *DRB1*03*, *DRB1*07*, *DRB1*10*, *DRB3*, and *DRB5* in humans, apes and Old World monkeys (Sliereendregt et al. 1994), suggests the inheritance of DRB alleles from a common primate ancestor.

DNA sequence analyses have expanded the number of DRB alleles identified for both human and non-human primates. In humans, at least 503 *HLA-DRB* sequences were described using cloning and sequencing or other molecular techniques, such as sequence-specific oligotyping or sequence-specific priming (reviewed in Marsh et al. 2005). While cloning and sequencing represents the “gold standard” for defining *Mhc-DRB* polymorphisms in populations, it is impractical for genotyping large numbers of individuals, as required by some medical projects or evolutionary studies. More rapid typing techniques such as allele-specific polymerase chain reaction remain impossible without prior knowledge of DNA sequence data when working on a new species. Here we described the use of cloning and sequencing combined with PCR-RFLP, as well as denaturing gradient gel electrophoresis (DGGE) and direct sequencing, to identify new *Mhc-DRB* sequences in wild chacma baboon *Papio ursinus* individuals. To date, no *Mhc-DRB* sequences were described in this taxon. These data are used to examine variability from an intra-specific and inter-specific perspective and will represent the first step in describing *Mhc-DRB* haplotypes in baboons, which requires an extended family-based study.

Materials and methods

DNA samples

For *Mhc-DRB* genotyping, genomic DNA was extracted from tissue samples from 30 wild chacma baboons (*P. ursinus*) living in Tsaobis Leopard Park, on the edge of the Namib Desert in Namibia, Southern Africa (for details of the site and population, see, e.g., Cowlshaw 1999). In the present study, samples were obtained from one solitary male and individuals belonging to four study troops, including three sets of mother–father–offspring, during trapping operations in 2000 and 2001 (unpublished data). All DNA samples were extracted using a DNeasy Tissue Kit (Qiagen, Crawley, UK) following the manufacturer’s instructions. Note that the taxonomy of *Papio* baboons is controversial, and the five members of the genus (*anubis*, *cynocephalus*, *hamadryas*, *papio*, and *ursinus*) can be treated as either individual species

or as subspecies of a single superspecies (named either *cynocephalus* or *hamadryas*). Here, we consider each taxon as an individual species following Groves (2001).

Cloning, RFLP analysis and sequencing

The complete set of *Mhc-DRB* sequences were characterised for four unrelated individuals. These sequences and their corresponding migration on a DGGE gel were next used as standards to identify the *Mhc-DRB* sequences from other individuals. A PCR was performed on individual genomic DNA with the universal MDRB primers 5'MDRB (5'-GCCTCGAGT GTCCCCCAGCACGTTTC-3') and 3'MDRB (5'-GCAAGCTTTCACCTCGCCGCTG-5') amplifying multiple DRB sequences for each individual. 30 to 40 ng of genomic DNA in 50 µl of F Buffer (Bioline, Paris, France), 1.5 mM Mg, 2.5 mM of each of the four deoxyribonucleotide triphosphates (dGTP, dATP, dTTP, and dCTP) (Invitrogen, Paisley, UK), 25 pmol of each of the primers and 1 unit of *Taq* polymerase (Goldstar, Eurogentec) were used for amplification. Cycling conditions consisted of 35 rounds of 90 s denaturation at 94°C, 90 s annealing at 54°C, and 90 s extension at 74°C. 5 µl of the 351 bp PCR product containing all the DRB genes of an individual were purified on a 1.5% agarose gel before cloning using a Topo TA Cloning kit (Invitrogen) following the manufacturer's instructions.

To increase the number of clones analysed to acquire a complete repertoire and to reduce the number of sequencing reactions performed, a RFLP screen approach was developed. For each individual at least 30 bacterial colonies were directly amplified with the MDRB primers. Cycling conditions consisted of 30 rounds of 20 s denaturation at 94°C, 30 s annealing at 65°C, and 36 s extension at 72°C, followed by a 10-min additional extension at 72°C. 10 µl of each PCR product was then digested independently with three restriction enzymes known to cut frequently (*MboII*, *HaeIII* and *Alu I*) and electrophoresed on 1.5% agarose gel to establish its restriction profile. For each restriction profile, three clones were purified with a Qiagen plasmid preparation kit and sequencing was directly performed on an ABI prism 310 sequencer using the Big Dye Terminator kit and the primers M13R-pUC (5'-CAGGAAACAGCTA TGAC-3') and M13F (5'-GTAAAACGACGGCCAGT-3'). When less than three clones corresponding to a specific profile were available, more colonies were RFLP screened.

Initially, a range of restriction enzymes were tested on the basis of previously published primate *Mhc-DRB* sequences. Once baboon DRB sequences were acquired, the choice of restriction enzymes was based on the analysis of variability comparing each DNA sequence, to find enzymes with restriction sites corresponding to variable sites using the program DNA Strider (Marck 1988). *Alu I*

was then replaced by *Rsa I*, which was more efficient in differentiating particular sequences.

Polymerase chain reaction for DGGE analysis

30 to 40 ng of genomic DNA in 50 µl of F Buffer (Bioline), 1.5 mM Mg, 2.5 mM of each of the four deoxyribonucleotide triphosphates (dGTP, dATP, dTTP, and dCTP (Invitrogen), 25 pmol of each of the primers 5'MDRB and 3'MDRB+GC (5'-CGCCCGCCGCGCCCCGCGCCCCGCCCCGCCCCGCCCGCCCCGCCCCGCAAGCTTTCACCTCGCCGCTG-3') and 1 unit of polymerase (Applied Biosystems, Warrington, UK) was used for amplification by PCR. Cycling conditions consisted of 35 rounds of 90 s denaturation at 94°C, 90 s annealing at 54°C, and 90 s extension at 74°C. After PCR, 5 µl of the 351 base-pair product was electrophoresed on a 1.5% agarose gel to confirm successful amplification of DNA samples.

Optimal conditions for the separation of DRB sequences using DGGE were established using perpendicular denaturing gradient as described by Knapp et al. (1997b). Using this information, parallel DGGE was used to separate DRB sequences. 30 µl of the GC-clamped PCR products were mixed with 15 µl of loading buffer and electrophoresed on a 12.6% acrylamide (37.5:1 acrylamide:bisacrylamide) gel with an increasing gradient from 35% to 60% denaturant (100% denaturant=7 M urea and 40% formamide) parallel denaturing gradient according to the method described by Myers et al. (1987). DGGEs were electrophoresed on a BIORAD D-GENE (Biorad, Hemel, Hempstead, UK) apparatus in 1×TAE buffer at 300 V constant voltage and 60°C constant temperature, during 5 h. Individual bands on the DGGE gels were visualized using SYBR Green stain (BioWhittaker, Wokingham, UK) and UV illumination. Individual gel plugs were removed using wide-bore pipette tips and eluted in 50 µl distilled water at room temperature overnight.

Sequence analysis and phylogenetic tree

MEGA (Kumar et al. 2001) was used to calculate the relative rate of non-synonymous and synonymous substitutions according to Nei and Gojobori (1986), applying the correction of Jukes and Cantor (1969) for multiple hits. *Paur-DRB* sequences were named by the IMGT-NHP (Immunogenetics-Non-human Primate) Nomenclature Committee and were submitted to the GenBank/EMBL/DBJ databases (accession numbers DQ339722–DQ339737). Accession numbers for sequences in the phylogenetic tree are provided in Fig. 5. *Paur-DRB* sequences were aligned with 80 non-human primate alleles published in the GenBank and the *HLA-DRB1*010101* as a reference for alignment. Sequences were compared in a

phylogenetic tree constructed by the neighbor-joining method based upon distances simulated by the Jukes and Kantor method using the computer software MEGA (Kumar et al. 1993). The analyses were based upon 219 nucleotides. Bootstrap analyses using 500 replications were performed to determine the repeatability of the DRB sequence alignment.

Results

Amplification of baboon DRB sequences

We amplified the highly variable second exon of multiple class II DRB sequences in 30 chacma baboons, using modified versions of the PCR primer pair, which amplifies all known *Mamu-DRB* sequences. Using the generic DRB primers 5'MDRB and 3'MDRB+GC, we identified 16 different DRB sequences in our study subjects.

Identification of DRB sequences using cloning and RFLP analysis

Cloning, in concert with RFLP analysis and sequencing, was first performed using four unrelated individuals to identify their complete repertoire of *Mhc-DRB* sequences. An example of the profiles for *Paur-DRB*W2701* and *Paur-DRB*W5501* is illustrated in Fig. 1. When less than three clones corresponding to a specific profile were available in this first set, more colonies were RFLP screened. The frequency of each profile in a set of bacterial colonies is highly variable. Each reference animal presents either four or five restriction profiles. Each restriction profile was sequenced and corresponded to a different *Mhc-DRB* sequence, which means that at least three loci were amplified. Using three different restriction enzymes, *HaeIII*, *MboII* and *Rsa I*, 14 different profiles were isolated.

Validation of DRB alleles using parallel DGGE

For each reference animal, each sequence obtained by cloning was first electrophoresed in a separate lane on a parallel denaturing gradient gel to determine if the sequences had unique positions on a DGGE. Then, a mix of the four or five sequences from each reference animal obtained by cloning was electrophoresed with the PCR product mixture amplified directly from the genomic DNA, on a parallel denaturing gradient. Figure 2 demonstrates a 100% correspondence of the bands characterized by cloning and sequencing with the bands obtained by direct amplification of the genomic DNA. An additional band in the IM01 PCR product represents a heteroduplex band, formed between complementary strands of different allele sequences during PCR amplification of DNA from heterozygous samples. Overall, the frequency of heteroduplex bands was less than 10%.

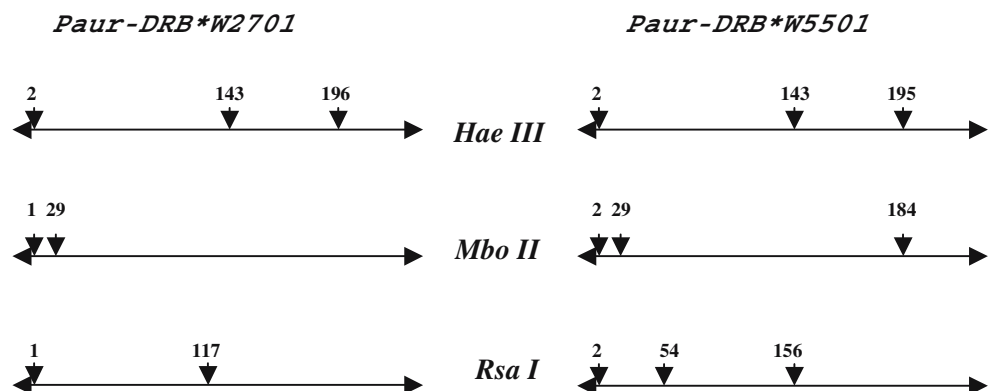
Separation of DRB sequences using parallel DGGE

The optimum gradient was identified using a perpendicular DGGE. PCR products for all 26 remaining animals were electrophoresed on a 35–60% parallel denaturing gradient gel. The reference animals were used as standards for identification of *Mhc-DRB* sequences of other individuals. After DGGE, individual bands corresponding to new bands were extracted, re-amplified, and directly sequenced. Two additional sequences were identified using DGGE.

Analysis of *Mhc-DRB* variability

Thus far, 16 different *Mhc-DRB* sequences were identified in our population, based on 92 (36.5%) variable nucleotide positions in a 252 bp sequence (Fig. 3). These sequences show wide-ranging levels of divergence with an average of 29.2 (11.6%) nucleotides differences (minimum: 3 substitutions, maximum: 41 substitutions) between sequences.

Fig. 1 Restriction enzymes were used to define the sequence profiles when cloning and sequencing. The profiles for *Paur-DRB*W4801* and *Paur-DRB*W2701* are shown in this example



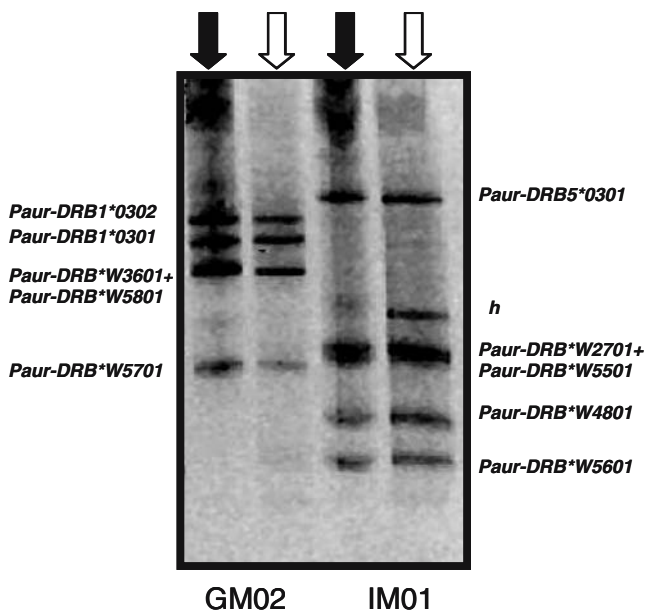


Fig. 2 Validation of the DRB typing technique. A SYBR Green-stained 35–59% parallel denaturing gradient gel showing DRB typing results from two chacma baboons (GM02 and IM01). *Black arrows* show a mix of all the *Mhc-DRB* sequences obtained by cloning for each individual. *White arrows* show the separation of the DRB sequences when the PCR product (5'MDRB/3'MDRB+GC) is electrophoresed on the denaturing gradient gel for each individual. A small *h* is positioned to the right of heteroduplex bands. Some bands can be difficult to separate by DGGE, as *Paur-DRB*W3601* and *Paur-DRB*W5801* or *Paur-DRB*W2701* and *Paur-DRB*W5501* on this gel

Each sequence has a unique amino acid sequence (Fig. 4) and the absence of stop codons suggests that all sequences could encode functional proteins. 40 (47.6%) variable sites are identified on a 84 amino acids sequence. Analysis of amino acid differences between sequences reveals between 3 (3.6%) and 26 (30.9%) with an average of 18.23 (21.7%) amino acids differences. Of the 16 sites predicted to be involved in antigen recognition (Brown et al. 1993; Slierendregt et al. 1992), all were variable, whereas 24 of 68 (35.3%) non-ABS sites are polymorphic. These variable non-ABS sites are mostly located next to an ABS site. The rate of non-synonymous (d_N) and synonymous (d_S) substitutions were estimated for both ABS and non-ABS amino acid positions (Table 1).

The genotypes of the 30 animals are described in Table 2. Only two pairs of animals, FF10 and HM05 and GM05 and GM06 share the same genotype. Data on relatedness for three pairs of parents (FF01 with FM07, IM01 with IF02 and IM01 with IF08) with infants (FM04, IM04 and IF11, respectively) are presented in Table 2 and suggest that *Mhc-DRB* sequences are vertically transmitted by haplotypes made up of no more than three loci (for example, the *Paur-DRB*W5601*–*Paur-DRB5*0301* haplotype is transmitted from IM01 to IM04).

Analysis of the phylogenetic tree

According to IMGT nomenclature, at least two loci and 13 lineages were identified. These sequences were assigned to *DRB1* and *DRB5* loci, to lineages that were described for other Old World primates and to new lineages that were not identified in other primates (Fig. 5). We did not identify any baboon DRB sequence assigned to the pseudogene locus, *DRB6*. BLAST analyses of sequences revealed a high percentage of identity with some other primate sequences. For example, *Paur-DRB*W101* was 100% identical with all 236 nucleotides reported for a *Papio cynocephalus* sequence, *Paca-DRB4*04c*, and *Paur-DRB1*0301* was 99% identical to *Paca-DRB1*13a*. A high degree of identity was also identified when comparisons were made between chacma baboons and rhesus macaques with some sequences having as much as 98% identity across the entire exon 2 sequence (e.g. *Mamu-DRB*W402* and *Paur-DRB*W401* or *Mamu-DRB*W101* with *Paur-DRB*W101*). It is surprising to note that *Paur-DRB*W3601* was found to be 99% identical with 249 bases of a *Galago moholi Mhc-DRB* sequence (Figuroa et al. 1994).

Three deleted nucleotides corresponding to a codon were found in two chacma baboons DRB sequences: *Paur-DRB*W3601* and *Paur-DRB*W5801*, for the positions 209, 210 and 211. This deletion is also found in some other primate species such as *G. moholi*, *Macaca nemestrina*, *Macaca fascicularis* and perhaps *Papio cynocephalus anubis*.

Discussion

Previously, identification of MHC alleles in non-human primates has relied upon ambiguous serological methods, equivocal isoelectric focusing, or labour-intensive cloning and sequencing techniques. Using the PCR and DGGE techniques (Knapp et al. 1997a), however, we rapidly and unambiguously identified 16 class II DRB in 30 wild chacma baboons. Our technique was first validated by identifying DRB sequences in four animals using cloning, RFLP and sequencing. For these four animals, DGGE correctly identified all 14 DRB sequences identified by cloning and sequencing. In addition, we identified two new sequences by typing the 26 remaining animals using PCR, DGGE technique and direct sequencing.

The combined use of RFLP, cloning and sequencing for a few individuals, followed by PCR, DGGE and direct sequencing for a larger population, represents a significant improvement over traditional class II typing methods. First, it allows the researcher to study a new species for whom no DRB sequences are known, by first identifying a limited set of DRB sequences using cloning for just a few individuals. Then,

HLADRB1*10101 TTGTGGCAGCTTAAGTTTGAATGTCATTTCTTCAATGGGACGGAGCGGGTGC GGTTGCTGGAAAGATGCATCTATAAC
Paur-DRB*W4801 ...GA...GG...C...G...C...C...C...C...C...A...T...C...
Paur-DRB*W2701 ...GA...GC...C...G...C...C...AC...C...A...C...
Paur-DRB*W5501 ...T...GG...G...C...GA...C...TC...A...
Paur-DRB*W5601 C...CA...T...C...G...C...C...AC...C...CA...T...
Paur-DRB5*0301 ...AA...GA...A...G...C...C...C...A...
Paur-DRB1*0301 ...GA.T.CTC...CA...G...C...C...T...C...C...A...T...
Paur-DRB1*0302 ...GA.T.CTC...CA...G...C...C...C...C...A...T...
Paur-DRB*W3601 ...CGA.T.CTG...C...CA...G...C...C...AC...A...TC...CTTT...
Paur-DRB*W5701 ...GA...GC...C...G...C...C...C...C...A...T...
Paur-DRB*W5801 ...CA...G...GA.T.TCC...G...A...AC...G...A...T...
Paur-DRB*W401 ...GA...CG...C...G...C...C...C...G...CA...T...
Paur-DRB*W301 ...GA...GC...ACG...G...C...C...C...C...A...T...C...
Paur-DRB*W5901 ...GAT...G...A...G...C...C...GC...CA...T...
Paur-DRB*W5602 C...CA...T...C...G...C...C...g...AC...C...CA...T...
Paur-DRB1*0303 ...GAGTACTC.AC...C...G...c...c...c...AC...C...A...T...
Paur-DRB*W101 C.TCAGTACTG.A...G...C...C...T...T...AC...ATC...A...T...CA...
79 156

HLADRB1*10101 CAAGAGGAGTCCCGTGCCTTCGACAGCGACGTGGGGGAGTACCGGGCGGGTACCGGAGCTGGGGCGGCCTGATGCCGAG
Paur-DRB*W4801 ...G...T...T...C...C...C...C...C...C...C...C...
Paur-DRB*W2701 ...G...T...C...C...C...C...C...C...C...C...
Paur-DRB*W5501 ...G...T...C...C...C...C...C...C...C...C...
Paur-DRB*W5601 ...G...T...T...C...C...C...C...C...C...C...C...
Paur-DRB5*0301 ...GA.C.T...C...C...C...C...C...C...C...C...
Paur-DRB1*0301 ...G...A...T...T...T...G.AGC...G...
Paur-DRB1*0302 ...G...TA...T...T...A...A...
Paur-DRB*W3601 ...G...AT...A...T.A...A...A...
Paur-DRB*W5701 ...G...A...T...T...T...C...C...
Paur-DRB*W5801 ...G...T.C...C...T...T...T...A...
Paur-DRB*W401 ...G...AA.C...A...A...A...A...
Paur-DRB*W301 ...G...A.C...T...G.AGC...G...
Paur-DRB*W5901 ...G...TA.C...T...T...A...A...
Paur-DRB*W5602 ...G...T...T...T...TC...C...
Paur-DRB1*0303 ...AA...A...A...C...C...
Paur-DRB*W101 ...A...A...A...TC...
157 234

HLADRB1*10101 TACTGGAACAGCCAGAAGGACCTCCTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACCTACGGGGTT
Paur-DRB*W4801 ...C...T.G...A...A...G.CA...T...A...T...
Paur-DRB*W2701 ...T.G...A...AG.CCA...T...T...C...
Paur-DRB*W5501 A...G...A...A...C...C...
Paur-DRB*W5601 ...T...TA...GCA...A...T...C...
Paur-DRB5*0301 ...G...G...A...TGTG...A...G...
Paur-DRB1*0301 ...G...TA...AG.CGA...A...T...T...
Paur-DRB1*0302 ...C...G...G...G.GC...A...A...A...T...C...
Paur-DRB*W3601 ...C...T...C...A.GT...A...--T...
Paur-DRB*W5701 ...T...AG.C.T...A...--T...GG...C...
Paur-DRB*W5801 ...C...G...G...AG.C...CG...--T...GG...C...
Paur-DRB*W401 ...C...A...G...A...CG...AT...T...C...
Paur-DRB*W301 ...TC...T.T...GC...A...AT...T...
Paur-DRB*W5901 A...T...CT.G...A...G.GC...GTG...C...
Paur-DRB*W5602 ...G...T...GCA...A...T...C...
Paur-DRB1*0303 ...A...A...G.CA...A...A...
Paur-DRB*W101 A...A...A...A...
235 252

HLADRB1*10101 GGTGAGAGCTTACAGTG
Paur-DRB*W4801 .TG...
Paur-DRB*W2701 TT...
Paur-DRB*W5501 ...
Paur-DRB*W5601 ...
Paur-DRB5*0301 .TG...
Paur-DRB1*0301 .TG...
Paur-DRB1*0302 ...
Paur-DRB*W3601 .TG...
Paur-DRB*W5701 ...
Paur-DRB*W5801 .TG...
Paur-DRB*W401 ...
Paur-DRB*W301 ...
Paur-DRB*W5901 ...
Paur-DRB*W5602 ...
Paur-DRB1*0303 ...
Paur-DRB*W101 ...T...

Fig. 3 Nucleotide sequences of chacma baboons (*P. ursinus*) *Mhc-DRB* sequences. Nucleotide number is shown above the *HLA-DRB1*10101* allele. Identity to this allele is shown by *dots*, while differences are given by *letter substitutions*, and gaps by *dashes*. Antigen binding sites are shown in *grey*

RFLP represents a cost-effective approach to screening a large set of bacterial colonies cloned from one individual because different restriction profiles, corresponding to all the DRB sequences of the individual, can be identified. This screening allows the researcher to select only three copies of each profile for sequencing. The DRB sequences were then used to validate and optimize the conditions of the PCR and DGGE technique, by electrophoresing the cloned sequences of one individual on the same gel as the PCR product amplified directly from individual's genomic DNA. As a final step, the sequences can be used as a standard for identification of *Mhc-*

DRB sequences of other individuals by systematically electrophoresing the new individuals and standards on denaturing gels. Only novel bands could be directly sequenced after the DGGE, allowing for rapid and economical genotyping of large numbers of sample. Thus, this method would be suitable for protocols requiring the study of many animals, such as population genetics studies or biomedical research.

In the present study, we identified 16 different sequences. In some cases, however, we observed two or three bands that were difficult to separate in our original 4-h denaturing gradient gels (Fig. 2). Three potential solutions were identified. First, improved separation of some bands was observed without compromising the resolution of other bands when denaturing gels were electrophoresed for a longer period of time. Alternatively, two co-migrating bands could be resolved using alternative sequencing

	1			70
HLADRB1*10	LWQLKFECHF	FNGTERVRL	ERCIYNQEE	SRFDSVGEY RAVTELGRP
Paur-DRB*W	.E.G.S....F.	D.YFH...FS...R...K.GQ..N
Paur-DRB*W	.E.A.S....Y.	D.Y.H...FH.....R..I...DQ.....
Paur-DRB*W	...F.G....F.	V.Y....FH.....N.....R...E...
Paur-DRB*W	.Q.F.S....Y.	Q.HF...FF.....V...L....Y...Q..N
Paur-DRB5*	.K.D.Y....H.	Y....D	A.....V.....E...
Paur-DRB1*	.EYSTS....	...M...F.	D.YF...YF...S...RS.....R..Y..DE.....
Paur-DRB1*	.EYSTS....F.	D.YF...LS...E..S...R..V..RA.T...N
Paur-DRB*W	FEYCTH....Y.	V.LF..R..Y	...N...F Q.....E..S.....KV..E..-
Paur-DRB*W	.E.A.S....F.	D.YF...YF.....A.....F..DS.....
Paur-DRB*W	.QRDYP....QY.	..YF...F	L...H...F...S...E..S...R..V..D...R..-
Paur-DRB*W	.EHV.S....F.	..HF...N	L.....E...S...E..S...R..V..EK..R...N
Paur-DRB*W	.E.A.R....F.	D.YFH...Y	A.....F.....RS...F...F..A.T...N
Paur-DRB*W	.D.V.Y....G.	H.F...L	A.....F...S...E..NL.TR..I..RA.....
Paur-DRB*W	.Q.F.S....Y.	Q.HF...FF.....V...G...F...Q..N
Paur-DRB1*	.EYSTS....Y.	D.YF...NK.GQ..N
Paur-DRB*W	.QYC.....L.LY.	I.YFH...Y	..Y.....V..N.....E..N
	71	84		
HLADRB1*10	YCRHNYGVGE	SFTV		
Paur-DRB*W	...Y...V.		
Paur-DRB*W	F..Y..R.F.		
Paur-DRB*WR...		
Paur-DRB*W	...Y..R...		
Paur-DRB5*	V.....RGV.		
Paur-DRB1*	...Y...V.		
Paur-DRB1*	...Y..R...		
Paur-DRB*W	-.....V.		
Paur-DRB*W		
Paur-DRB*W	-..R..R.V.		
Paur-DRB*WR...		
Paur-DRB*W	...Y.....		
Paur-DRB*W	V.....R...		
Paur-DRB*W	...Y..R...		
Paur-DRB1*		
Paur-DRB*W		

Fig. 4 Protein sequences of chacma baboons (*P. ursinus*) DRB alleles (exon 2). The protein number is shown above the *HLA-DRB1*10101* allele. Identity to this allele is shown by *dots*, while differences are

given by *letter substitutions*, and gaps by *dashes*. Antigen binding sites are shown in *grey*

Table 1 The estimated rates (\pm SE) of non-synonymous (d_N) and synonymous (d_S) substitutions for antigen (ABS) and non-antigen (non-ABS) binding sites and their ratio for DRB–exon 2 sequences in *P. ursinus*

Positions	<i>n</i>	d_N	d_S	d_N/d_S	<i>P</i>
ABS	16	0.515 \pm 0.091	0.268 \pm 0.090	1.92	<0.001
Non-ABS	68	0.062 \pm 0.016	0.076 \pm 0.024	0.81	<0.001
All	84	0.132 \pm 0.024	0.105 \pm 0.025	1.26	<0.001

n is the number of codons in each category and *P* is the probability that d_N and d_S are different using a *t* test

chemistries, such as *Taq* FS Dye Primer, designed for sequencing mixed (heterozygous) samples. Thirdly, an economical solution for systematic analysis of a large number of samples is excision of the co-migrating bands for RFLP analysis, as described for screening clones.

Occasionally, the presence of heteroduplex bands was noted on our parallel DGGEs (Fig. 2). A number of studies have described heteroduplex formation between complementary strands of different alleles during PCR amplification of DNA from heterozygous samples (Keohavong et al. 1991; Myers and Maniatis 1986; Uhrberg et al. 1994). The heteroduplex can be identified in every case when forward and reverse sequences are compared. In our experience, this process is highly repeatable because two sequences will always produce heteroduplex bands when they migrate together in the same lane on a parallel gel, as for example *Paur-DRB*W4801* and *Paur-DRB*W5501* (see Fig. 2). Thus, identification is straightforward when working on a limited set of sequences.

Sixteen different MHC class II DRB-exon 2 alleles were found in 30 individuals. Each baboon possesses at least two sequences, but no more than six (Table 2), suggesting that up to three DRB loci were amplified. The *Paur-DRB* sequences showed a high level of divergence at the intra-specific level and all had unique amino acid sequences. Polymorphism was higher in the functionally important antigen binding sites. In the ABS, significantly more non-synonymous than synonymous substitutions were found (ratio: 1.92, see Table 1). Previous authors (Hughes and Nei 1989; Hughes and Yeager 1998) have considered this high rate of non-synonymous substitutions to be a clear indication of positive selection and it is characteristic of proteins with antigen-presenting functions (Bergstrom and Gyllenstein 1995). By contrast, purifying selection is observed to act in the non-antigen binding sites (Nei and Gojobori 1986) and the ratio between non-synonymous and synonymous substitutions was significantly smaller than one (0.81, see Table 1). d_N was 8.31 times higher in the antigen-binding sites than in the non-ABS. This indicates selection processes that maintain polymorphism in the functionally important regions of the MHC.

The nomenclature and phylogenetic analysis demonstrates that chacma baboons *Mhc-DRB* sequences are highly variable. Instead of clustering together, our *P. ursinus* sequences are spread throughout the phylogenetic tree (Fig. 5). In contrast with the pattern illustrated in the *Paca-DRB* sequences, this result suggests the possibility of trans-species inheritance of some sequences rather than new ones arising after divergence from a common ancestral sequence. In addition, at least one chacma baboon sequence could be very ancient, as it is nearly identical to a prosimian species, *Galago moholi*. While bootstrap values are generally low in the phylogenetic tree, the presence of loci and lineages that were identified in other non-human primates seems likely: three *DRB1* sequences and one *DRB5* sequence were found. As in most other primate studies, it is difficult to be sure of locus identification and family studies for haplotype identification are required to confirm that alleles belong to the same locus.

It is interesting to note that all sequences appear to be structurally functional, as they do not show stop codons, deletions or insertions, which change the reading frame. This lack of *DRB6* pseudogenes is unusual in primates (but see Schad et al. 2004) and is unlikely to result from our methodology because we used a relatively low annealing temperature to ensure amplification of multiple DRB sequences in every individual and previous use of our primer set has led to successful identification of *DRB6* sequences in other primate species. Nevertheless, one should be cautious in assuming that all sequences were amplified in our study subjects. Pseudogenes are well-known for their contribution to the evolution and creation of new alleles by gene conversion processes. The lack of pseudogenes and the fact that few sequences are very similar to one another suggests that *Mhc-DRB* sequences are well-established in this population of wild chacma baboons. Indeed, pseudogenes could have been lost in the absence of selection acting to create new alleles by duplication and gene conversion. The high degree of similarity between one baboon sequence and one from *Galago moholi*, whose geographical distribution in southern Africa widely overlaps the chacma baboon, further suggests that some *Mhc-DRB* genes may be very stable because they are closely adapted to local selective pressures. As a consequence of the limited *Mhc-DRB* repertoire in this species, the creation of unique genotypes may require other evolutionary processes than duplication and gene conversion, such as recombination or assortative mating between individuals. Supporting this hypothesis, we also observed that only two pairs of individuals out of 30 chacma baboons share the same genotype. Moreover, data on relatedness for three pairs of parents with infants suggest that *Paur-DRB* sequences are vertically transmitted by haplotypes made up of no more than three sequences.

Table 2 *Paur-DRB* sequences identified using cloning, PCR/DGGE, and direct sequencing for 18 individuals

Ind.	Mo.	Fa.	W4801	W2701	W5501	W5601	5*0301	I*0301	I*0302	W3601	W5701	W5801	W401	W301	W5901	W 5602	I*0303	W101	
FF02																			
FF03		X			X	X	X	X					X						
FF05		X	X		X	X	X												
FF09		X		X	X	X	X												
FF10		X		X	X	X	X												
FM07		X		X	X	X	X	X											
FM04	FF01	FM07			X*	X*	X*												
FF01					X	X	X	X					X						
HM05		X		X	X	X	X												
HM15					X	X	X						X		X				
HM17		X	X							X	X								
HF02										X	X								
HF10		X		X															
HF11																			
HF17										X	X			X					X
GM02										X	X			X					
GM04					X	X	X	X						X					
GM05					X	X	X	X											
GM06					X	X	X	X											X
GF04										X	X			X					
GF12																			X
GF13					X	X	X	X											X
GF14					X	X	X	X											X
IM01		X		X	X	X	X	X											
IM04	IF02	IM01			X*	X*	X*	X*											X°
IF02		X		X															X
IF11	IF08	IM01	X*	X*									X°	X°					X°
IF08																			X
IM02										X	X								X
ZM02										X	X								X

Individuals used for cloning are indicated in boldface. For individuals FM04, IM04 and IF11, sequences inherited from the father are indicated by an asterisk and sequences inherited from the mother are indicated by a circle.

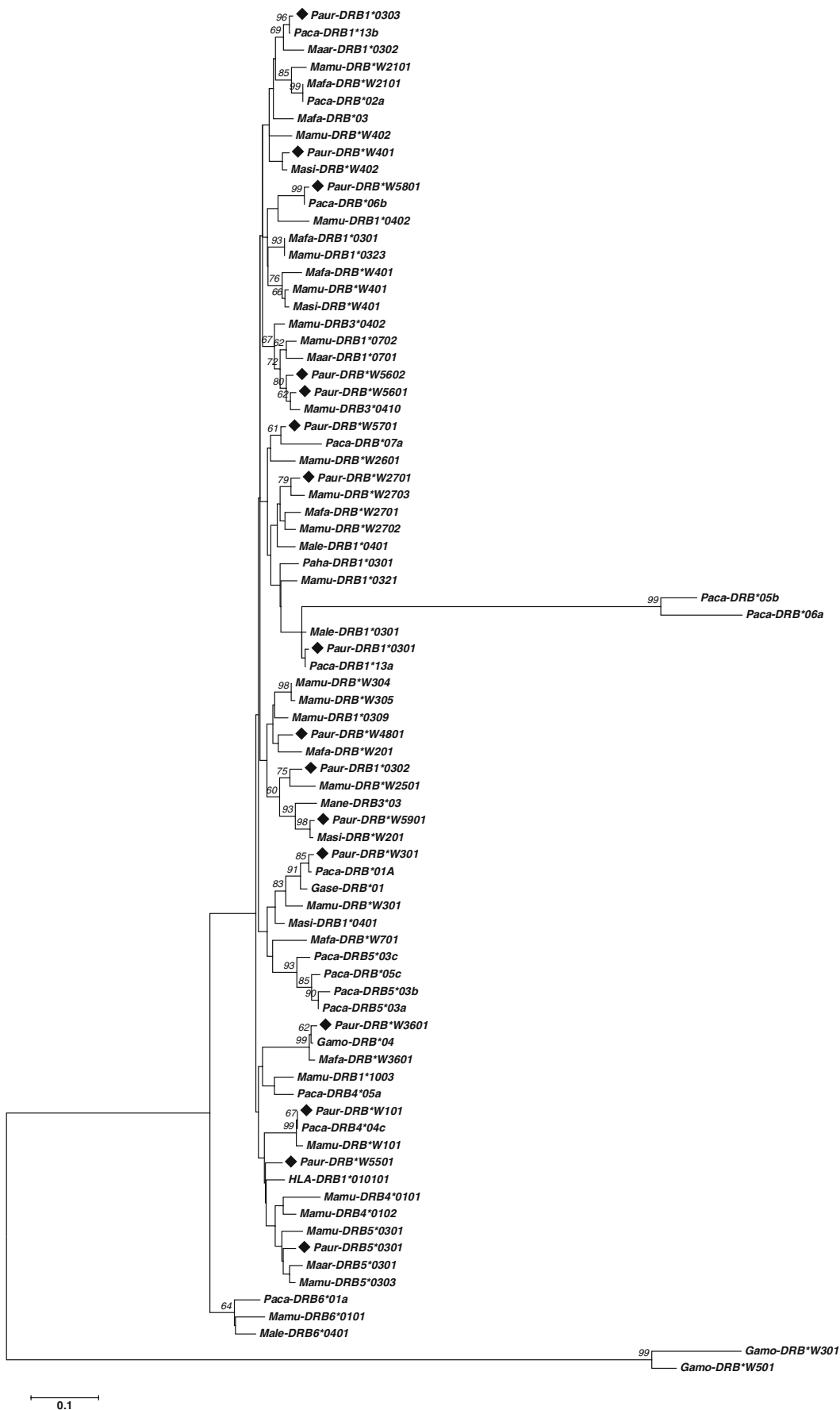


Fig. 5 Phylogenetic tree of *Paur-DRB* sequences and selected primate *Mhc-DRB* sequences published in the GenBank. All new *Paur* sequences are displayed with ♦. The tree configuration was derived from nucleotide sequences using the neighbor-joining method of the MEGA program. Numbers beside the branches refer to bootstrap values. Accession numbers or references, if there is no accession number, of sequences are as follows: *HLA-DRB1*010101*: AF029288; *Mamu-DRB*: AF163272–AF163288, AF031269, AF173354, AJ601352–AJ601367, AY487268–AY487276, U57943–U57947, Z26143–Z26174; *Mafa-DRB*: AF492277, AY340670–AY340699; *Mane-DRB*: M60061; *Maar-DRB*: AF173351–AF173356; *Masi-DRB*: AY172138–AY172145; *Male-DRB*: Gyllensten et al. 1991; *Paca-DRB*: L77186–L77192, L79974–L79985; *Paha-DRB*: M57589; *Gamo-DRB*: AF197238, AF197240, Z27133; *Gase-DRB*: Z27157

This transmission pattern seems very different from those described in other Old World monkeys such as rhesus macaques, where individuals have extremely different number of sequences, and haplotypes are composed of up to six different genes (Slierendregt et al. 1994). The creation of new genotypes could be made possible in chacma baboons by combining these small haplotypes, but this will have to be confirmed using a large sample of related individuals.

The baboon represents an important biomedical tool for human disease, and a model system for evolutionary ecology. The rapid and unambiguous identification of *Mhc-DRB* sequences in baboons will therefore be a valuable contribution to ongoing studies in these fields. In the future, the application of DGGE for DRB genotyping in the different baboon taxa should facilitate the identification and characterization of unique alleles and, perhaps, additional loci in large samples. Thus, our method is especially suitable for population genetics studies. When applied to family-based haplotype studies, DGGE should also prove useful in the investigations of the organization and evolution of baboon DRB haplotypes. The 16 sequences characterized in our study represent new sequences with different antigen-presenting abilities. Our study is the first systematic study conducted on wild chacma baboons. These new sequences should, therefore, make a valuable contribution to the understanding of *Mhc-DRB* evolution in Old World primates. Moreover, given the limited sequence data available for *Mhc-DRB* in other baboon taxa, this study describes a reliable and efficient method that can be used to study baboons and other primates.

Acknowledgements For the assistance with sample collection in the field, we gratefully thank all those individuals who helped with our trapping operations in 2000 and 2001, especially Crispian Barlow, Andre and Tana Burger, Fredi Devas, Anna Randall, H.O. Reuter, Marcus Rowcliffe, and Hartmut Winterbach. Special thanks also go to Anna Randall and Stephan Funk for conducting the DNA extractions (with additional assistance from Dada Gotelli) in London. We also thank Polla and Chrija Swart for their permission to work at the Tsaobis Leopard Park, the Desert Research Foundation of Namibia for affiliation, and the Ministry of Environment and Tourism for research permission. We confirm that we have adhered to the Guidelines for the Use of Animals in Behavioural Research and Teaching (Animal

Behaviour 2003, 65:249–255) and the legal requirements of the country in which the work was carried out (Namibia). This work was funded by a Natural Environment Research Council (UK) grant awarded to GC, and a Ministère de l'Éducation et de la Recherche (France) Studentship awarded to EH. GC was funded by a NERC Advanced Fellowship during the writing of this paper. This paper is a publication of the ZSL Institute of Zoology's *Tsaobis Baboon Project*. Contribution ISEM 2006, #XXX.

References

- Alberts SC (1999) Thirteen Mhc-DQA1 alleles from two populations of baboons. *Immunogenetics* 49:825–827
- Bergstrom T, Gyllensten U (1995) Evolution of Mhc class II polymorphism: the rise and fall of class II gene function in primates. *Immunol Rev* 143:13–31
- Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol* 16:363–377
- Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC (1987) Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329:506–512
- Bodmer JG, Marsch, SGE, Albert E, Bodmer ED, Dupont WF, Erlich HA, Mach B, Parham WR, Sasazuki P, Schreuder T, Strominger GMT, Svejgaard A, Terasaki PI (1994) Nomenclature for factors of the HLA system. *Tissue Antigens* 44:1–18
- Brown JH, Jardetzky T, Saper MA, Samaraoui PJ, Bjorkman PJ, Wiley DC (1988) A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature* 332:845
- 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* 364:33–39
- Cowlshaw G (1999) Ecological and social determinants of spacing behaviour in desert baboon groups. *Behav Ecol Sociobiol* 45:67–77
- Figuroa F, O'hUigin C, Tichy H, Klein J (1994) The origin of the primate Mhc-DRB genes and allelic lineages as deduced from the study of prosimians. *J Immunol* 152:4455–4465
- Gaur LK, Nepom GT, Snyder KE, Anderson J, Pandarpurkar M, Yadock W, Heise ER (1997) Mhc-DRB allelic sequences incorporate distinct intragenic trans-specific segments. *Tissue Antigens* 49:342–355
- Groves C (2001) Primate taxonomy. Smithsonian series in comparative evolutionary biology. Smithsonian Institution Press, Washington and London
- Gyllensten U, Sundvall M, Ezcurra I, Erlich HA (1991) Genetic diversity at class II DRB loci of the primate MHC. *J Immunol* 146:4368–4376
- Hughes AL, Nei M (1989) Nucleotide substitutions at MHC class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA* 93:5380–5383
- Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. *Annu Rev Genet* 32:415–432
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian protein metabolism*. Academic, New York, pp21–132
- Kenter M, Otting N, Anholts J, Jonker M, Schipper R, Bontrop RE (1992) Mhc-DRB diversity of the chimpanzee (*Pan troglodytes*). *Immunogenetics* 37:1–11
- Keohavong P, Liu VF, Thilly WG (1991) Analysis of point mutations induced by ultraviolet light in human cells. *Mutat Res* 249:147–159
- Klein J (1987) Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Hum Immunol* 19:155

- Klein J, O'hUigin C, Figueroa F, Mayer WE, Klein D (1993) Different modes of Mhc evolution in primates. *Mol Biol Evol* 10:48–59
- Knapp LA, Cadavid LF, Eberle ME, Knechtle SJ, Bontrop RE, Watkins DI (1997a) Identification of new Mamu-DRB alleles using DGGE and direct sequencing. *Immunogenetics* 45:171–179
- Knapp LA, Lehmann E, Hennes L, Eberle ME, Watkins DI (1997b) High-resolution HLA-DRB typing using DGGE and direct sequencing. *Tissue Antigens* 50:170–177
- Kumar S, Tamura K, Nei M (1993) Molecular evolutionary genetics analysis, Pennsylvania State University, University Park
- Kumar S, Koichiro T, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17: 1244–1245
- Lekutis C, Letvin NL (1995) Biochemical and molecular characterization of rhesus monkeys major histocompatibility complex class II DR. *Hum Immunol* 43:72–80
- Marck C (1988) 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res* 11:1829–1836
- Marsh SG (2005) WHO nomenclature committee for factors of the HLA system: nomenclature for factors of the HLA system, update June 2005. *Hum Immunol* 66:1025–1027
- Marsh SGE, Bodmer JG (1995) HLA class II region nucleotide sequences. *Eur J Immunogenet* 22:526a–527b
- Myers RM, Maniatis T (1986) Recent advances in the development of methods for detecting single-base substitutions associated with human genetic diseases. *Cold Spring Harbor Symp Quant Biol* 51 Pt 1:275–284
- Myers RM, Fischer SG, Lerman LS, Maniatis T (1987) Nearly all single base substitutions in DNA fragments joined to GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Res* 13:3131–3145
- Nei M, Gojobori T (1986) Simple methods for estimating the number of synonymous and non-synonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
- Otting N, de Groot NG, Noort MC, Doxiadis GGM, Bontrop RE (2000) Allelic diversity of Mhc-DRB alleles in rhesus macaques. *Tissue Antigens* 56:58–68
- Schad J, Sommer S, Ganzhorn JU (2004) MHC variability of a small lemur in the littoral forest fragments of southeastern Madagascar. *Conservation Genetics* 5:299–309
- Slierendregt BL, van Noort JT, Bakas RM, Otting N, Jonker M, Bontrop RE (1992) Evolutionary stability of transspecies major histocompatibility complex class II DRB lineages in humans and rhesus monkeys. *Hum Immunol* 35:29–39
- Slierendregt BL, Kenter M, Otting N, Anholts J, Jonker M, Bontrop RE (1993) Major histocompatibility complex class II haplotypes in a breeding colony of chimpanzees (*Pan troglodytes*). *Tissue Antigens* 42:55–61
- Slierendregt BL, Otting N, van Besouw N, Jonker M, Bontrop RE (1994) Expansion and contraction of rhesus macaque DRB regions by duplication and deletion. *J Immunol* 152:2298–2307
- Uhrberg M, Enczmann J, Wernet P (1994) Rapid DNA crossmatch analysis of HLA class II genotypic polymorphisms by temperature gradient gel electrophoresis in unrelated bone marrow donor selection. *Eur J Immunogenet* 21:313–324