ORIGINAL PAPER

High levels of diversity characterize mandrill (*Mandrillus sphinx*) Mhc-DRB sequences

Kristin M. Abbott · E. Jean Wickings · Leslie A. Knapp

Abstract The major histocompatibility complex (MHC) is highly polymorphic in most primate species studied thus far. The rhesus macaque (*Macaca mulatta*) has been studied extensively and the *Mhc-DRB* region demonstrates variability similar to humans. The extent of MHC diversity is relatively unknown for other Old World monkeys (OWM), especially among genera other than *Macaca*. A molecular survey of the *Mhc-DRB* region in mandrills (*Mandrillus sphinx*) revealed extensive variability, suggesting that other OWMs may also possess high levels of *Mhc-DRB* polymorphism. In the present study, 33 *Mhc-DRB* loci were identified from only 13 animals. Eleven were wild-born and presumed to be unrelated and two were captive-born twins. Two to seven different sequences were identified for each individual, suggesting that some mandrills may have as many as four *Mhc-DRB* loci on a single haplotype. From these sequences, representatives of at least six *Mhc-DRB* loci or lineages were identified. As observed in other primates, some new lineages may have arisen through the process of gene conversion. These findings indicate that mandrills have *Mhc-DRB* diversity not unlike rhesus macaques and humans.

Keywords Major histocompatibility complex ·
Mandrillus sphinx · DGGE · Sequencing · Phylogenetics

K. M. Abbott (✉) · L. A. Knapp
PrIME, Department of Biological Anthropology,
University of Cambridge,
Downing Street,
Cambridge CB2 3DZ, UK
email: K.M.Abbott.01@cantab.net

E. J. Wickings
Unité de Génétique des Ecosystèmes Tropicaux, CIRMF,
BP 769 Franceville,
Gabon, Africa

Introduction

The major histocompatibility complex (MHC) is well-known for the large number of genes and diversity of alleles found in most vertebrates (Klein and Figueroa 1986; Klein et al. 1993). These genes are divided into three classes, of which classes I and II genes exhibit the most polymorphism (Trowsdale 1996). Class I genes encode molecules that recognise intercellular antigens and class II genes encode molecules that recognise extracellular antigens, such as those of invading parasites and other pathogens. Evidence for the MHC's role in immune function is based, in part, on the patterns of variation concentrated in the peptide-binding region (PBR) (Hughes and Nei 1989). Variation within the PBR suggests that there has been evolutionary pressure for organisms to combat a wide range of immunological challenges.

The human MHC is better known as the human leukocyte antigens (HLA) and the *HLA-DRB* genes consist of four functional loci that encode molecules used in immune response (-*DRB1*, *3*, *4*, and *5*). Three hundred and eighty-nine sequences assigned to the *HLA-DRB1* locus can be divided into *lineages* based on nucleotide and amino acid sequence motifs (i.e. *HLA-DRB1*01*, -*DRB1*02*, etc.). Five pseudogene loci have also been identified (-*DRB2*, *6*, *7*, *8* and *9*) and a total of 470 *HLA-DRB* sequences have been described (Robinson et al. 2003; Marsh et al. 2005).

The MHC appears to be polymorphic in all primate species studied thus far. Molecular studies have shown that rhesus macaque (*Macaca mulatta*) *Mamu-DRB* loci are also highly variable. To date, 134 *Mamu-DRB* sequences have been identified using molecular genetic techniques (Doxiadis et al. 2001; Robinson et al. 2003). *Mamu-DRB* nucleotide sequences have been given locus and lineage designations based on their similarity to human sequences.

Like the human ortholog, the *Mamu-DRB1* locus also consists of several lineages: *Mamu-DRB1*1/10*, *1*03*, *1*04* and *1*07*. Homologs of other *HLA-DRB* loci have also been identified in rhesus macaques. These include the *-DRB3*, *4*, *5* and *6* loci. Interestingly, not all loci and lineages are shared between the two species. For example, the *HLA-DRB1*02* and *-DRB1*09* lineages and the *HLA-DRB2*, *7*, *8* and *9* loci are not present in rhesus macaques. Likewise, rhesus macaques have numerous loci that are not orthologous to humans and, hence, are identified by workshop designations (e.g., *DRB*W201*).

Humans and rhesus macaques also differ in the degree of variation observed at their respective *Mhc-DRB* loci and among lineages. While both species exhibit variation within the *DRB1*03* lineage, variation is also high in the rhesus macaque *-DRB6* and *-DRB*W* loci. As in humans, *Mamu-DRB6* loci are pseudogenes and research demonstrates that some *-DRB1*, *-DRB3* and *-DRB*W* genes are also not transcribed (De Groot et al. 2004). The presence of a large number of non-functional sequences may provide the raw material for creating new functional genes through gene conversion and recombination (Beck and Trowsdale 2000).

While humans and rhesus macaques have been the primary focus of primate MHC research to date, little is known about the MHC of other Old World monkeys (OWM). Only DNA sequences in a few macaque species (*M. arctoides*: Kriener et al. 2000; *M. fascicularis*: Kriener et al. 2000; Leuchte et al. 2004; *M. mulatta*: Robinson et al. 2003; *M. nemestrina*: Zhu et al. 1991; *M. silenus*: Blankenburg et al. 2003), and only one sequence from a baboon (*Papio hamadryas*: Riess et al. 1990) and three sequences from one drill (*Mandrillus leucophaeus*: Gyllensten et al. 1991), have been published in the ImMuno-GeneTics Information System MHC-NHP database. In fact, only three macaque species have been examined thoroughly. Therefore, little is known about the variability of loci, lineages, or even haplotypes between OWM species. In this paper, we describe *Mhc-DRB* diversity in another OWM, the mandrill (*Mandrillus sphinx*). We are currently studying the mandrill to better understand the functions and

complexities of class II MHC genes in primates and to evaluate whether the patterns of *Mhc-DRB* diversity described for rhesus macaques are unique among OWMs.

Materials and methods

Source of DNA

Peripheral blood samples were collected from 11 wild-born mandrills housed at the Centre International de Recherches Medicales, Franceville (CIRMF) in Gabon. Samples, obtained during routine health examinations, were treated with 1×SSC (0.15 M NaCl, 0.015 M sodium citrate). Genomic DNA was extracted in Unité de Génétique des Ecosystèmes Tropicaux (UGENET) as described by Wickings and Dixson (1992). Samples were stored at −60°C until they were transported to the Department of Biological Anthropology, Cambridge University, (DEFRA-CITES: AHZ/2035A/2002/4-250403/01-03 and AHZ/2035C/2004/6-262315/01-03). Tissue samples were also collected from two captive stillborn mandrills at the Chester Zoo (United Kingdom). Samples were stored at −20°C at the Department of Biological Anthropology, Cambridge University. Genomic DNA was extracted using QIAamp DNA Mini Kit (Qiagen, UK).

Polymerase chain reaction (PCR)

Amplification of MHC sequences was performed as described by Knapp et al. (1997), with slight modification to the cycling conditions. Thirty to 40 ng of genomic DNA, in 25 μl of 1×PCR buffer [pH 8.5, MgCl$_2$ 1.5 mM (Invitrogen, Paisley, UK)], 2.5 mM of deoxyribonucleotide triphosphates (dNTPs) (Invitrogen), 25 pmol of each of the oligonucleotide primers 5′MDRB and 3′MDRB+GC (Table 1) and 1 unit of *Taq* polymerase (Invitrogen) was used for amplification by PCR. Optimal cycling conditions consisted of 30 cycles of 90 s denaturation at 94°C, a 90-s annealing at 54°C and a 90-s extension at 74°C. A single

Table 1 Nucleotide primer sequence list

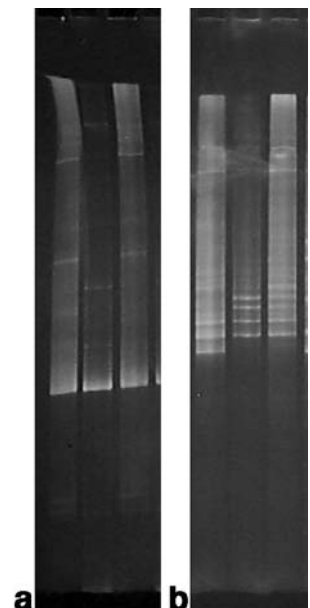| Name	Sequence (5′–3′)	PCR protocol
5′ MDRB	GCC TCG AGT GTC CCC CCA GCA CGT TTC	DGGE/cloning
3′ MDRB	GCA AGC TTT CAC CTC GCC GCT G	Cloning
3′ MDRB+GC	CGC CCG CCG CGC CCC GCG CCC GGC CGCCGG CCG CCC CCG CCC GCA AGC TTT CACCTC GCC GCT G	DGGE
5′ MDRBrM13	CAG GAA ACA GCT ATG ACC TGT CCC CCAGC ACG	Direct sequencing
3′ DRB1BM13	TGT AAA ACG ACG GCC AGT CGC TGC ACTGTG AAG CTC TC	Direct sequencing
M13 forward (−20)	GTAAAACGACGGCCAGT	n/a
M13 reverse (−20)	GCGGATAACAATTTCACACAGG	n/a
M13 reverse-pUC(−40)	CAGGAAACAGCTATGAC	n/a

Table 2 *Masp-DRB* typing results for two unrelated mandrills. The sequence found using only one method is indicated in bold print

Reference Animal	*Masp-DRB* sequences	Cloning/ sequencing	PCR/DGGE/ sequencing
10	1*0403	+	+
	3*0401	+	+
	6*0101	−	+
	6*0401	+	+
	*W2501	+	+
	*W2503	+	+
	*W5101	+	+
18	1*0405	+	+
	5*0301	+	+
	*W301	+	+
	*W302	+	+
	*W5201	+	+

Fig. 2 Using DNA for three identical subjects, **a** shows a 3.75-h DGGE using a gradient of 41–50% and **b** shows a 4.5-h (+15 min longer due to low amperage) DGGE using a gradient of 40–60%

cycle of 10 min at 74°C was carried out, followed by a holding temperature of 4°C. After PCR, 5 μl of the product was electrophoresed on a 1.5% agarose gel to confirm successful amplification of DNA samples.

Cloning and sequencing

Initially, cloning and sequencing, as well as PCR, denaturing gradient gel electrophoresis (DGGE) and direct sequencing, were used to identify *Mhc-DRB* sequences in two wild-born mandrills. Clones were sequenced (Macrogen,
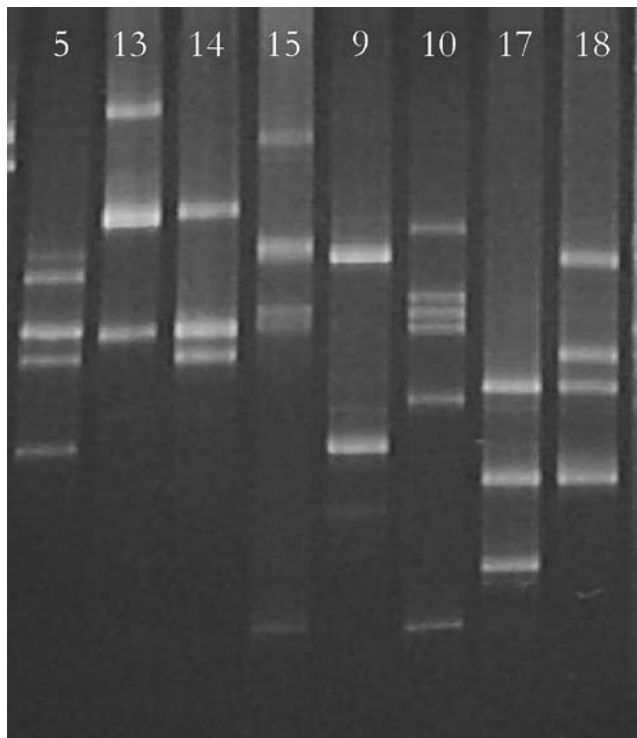
Fig. 1 Parallel DGGE parameters: 40–60% gradient, 4.5 h (+15 min longer due to low amperage), 300 V and 60°C

Korea) in both directions using M13 forward (−20) and M13 reverse (−20) primers (Table 1). A total of 322 and 262 clones were screened from multiple cloning experiments for mandrills 10 and 18, respectively. As DGGE and sequencing was quicker and less costly than cloning and sequencing, and proved to reliably identify all sequences, it was the chosen primary method for the present study.

Denaturing gradient gel electrophoresis

The technique of *Mhc-DRB* allele separation using DGGE established by Knapp et al. (1997) was used with slight modifications of parallel conditions specifically required for identification of mandrill *Mhc-DRB*. A perpendicular DGGE was used to determine optimal conditions for parallel DGGE separation of mandrill *DRB* sequences. The GC-clamped PCR product was electrophoresed in a 12.6% acrylamide (37.5:1 acrylamide to bisacrylamide) gel with an increasing gradient from 0 to 80% denaturant (100% denaturant=7 M urea and 40% formamide) for 3.5 h at 300 constant volts. Perpendicular gels were electrophoresed in a BIORAD D-GENE (BioRad, Hemel Hempstead, UK) apparatus in 1×TAE buffer at 60°C. After electrophoresis, the gel was silver-stained (Bassam et al. 1991) and photographed.

According to perpendicular gel results, 7–20 μl GC-clamped PCR products were electrophoresed on a 41–50% parallel denaturing gradient gel for 3.75 h at 300 constant volts to separate mandrill -*DRB* sequences. As described for the perpendicular DGGE, the 12.6% acrylamide gel was electrophoresed in a BIORAD D-GENE apparatus in 1×TAE buffer at 60°C. After electrophoresis, gels were stained using Sybr-Gold (BioGene, Cambridge, UK) for 10 min and visualised using UV illumination. Gel plugs of

Table 3 *Masp-DRB* sequences for 13 mandrills

Reference animal (wild-born and **captive-born**) and complement of identified *Masp-DRB* sequences

5[a]	6[a]	9	10	12	13	14	15	16[a]	17[a]	18	**A1**[a]	**B1**[a]
1*0305	1*0301	5*0304	1*0403	1*0403	1*0304	1*0301	1*0302	1*0302	*W302	1*0405	1*0302	1*0302
1*0401	1*0404	*W2504	3*0401	3*0401	1*0402	3*0402	5*0302	5*0302	*W2502	5*0301	1*0303	3*0402
3*0402	*W2601		6*0101	6*0401	6*0402	6*0402	6*0102	5*0304	*W5201	*W301	5*0302	5*0302
6*0403			6*0401	*W3701	*W2601	*W401	*W2501	*W2504		*W302	5*0304	6*0404
*W401			*W2501			*W2601	*W2503			*W5201	6*0404	*W401
			*W2503				*W5101					
			*W5101									

[a] PCR products electrophoresed on 40–60% parallel denaturing gradient gel. All other PCR products were electrophoresed on 41–50% parallel denaturing gradient gel

individual bands were excised using wide-bore pipette tips and eluted in 50 µl distilled water at room temperature overnight, as described by Knapp et al. (1997).

DRB sequencing

To generate PCR product for sequencing from DGGE, 15 µl of the eluted DNA was re-amplified in a 25-µl reaction using 10×PCR buffer (Invitrogen), 2.0 mM MgCl$_2$ (Invitrogen), 2.5 mM of dNTPs, 12.5 pmol of each of the hemi-nested primers 5′MDRBr- and 3′DRB1B-M13 (Table 1) and 0.5 unit of *Taq* polymerase. Cycling conditions consisted of 30 cycles of 20 s denaturation at 94°C, 30 s annealing at 60°C and 36 s extension at 72°C, a single cycle of 10 min at 72°C and a final holding temperature of 4°C.

In preparation for sequencing, 5 µl of the PCR product was electrophoresed on a 1.5% agarose gel to confirm successful amplification of DNA samples. Samples with visible primer-dimer on agarose gels were treated with 3 µl ExoSapIt (Amersham Biosciences, UK) and incubated at 37°C activation for 15 min, 80°C deactivation for 15 min and a final temperature of 4°C. All sequencing was performed by Macrogen (Korea). DNA was sequenced in both directions using M13 forward (−20) and M13 reverse-pUC (−40) primers (Table 1).

Masp-DRB nomenclature

Official locus and sequence names were requested and designated by the ImMunoGeneTics Information System MHC-NHP group (Robinson et al. 2003). New sequences have been deposited in GenBank (Accession numbers: DQ103715-47).

Sequence analysis and phylogenetic trees

Mandrill *DRB* sequences were compared by alignment to previously published OWM sequences, including *Mamu-*,

Table 4 Summary of *Masp-DRB* sequences and typing techniques used to identify each gene

Sequence	Technique		Accession number
	DGGE/direct sequencing	Cloning/ sequencing	
DRB1*0301	+		DQ103731
DRB1*0302	+		DQ103732
DRB1*0303	+		DQ103733
DRB1*0304	+		DQ103734
DRB1*0305	+		DQ103735
DRB1*0401	+		DQ103736
DRB1*0402	+		DQ103737
DRB1*0403	+	+	DQ103721
DRB1*0404	+		DQ103738
DRB1*0405	+	+	DQ103722
DRB3*0401	+	+	DQ103723
DRB3*0402	+		DQ103739
DRB5*0301	+	+	DQ103724
DRB5*0302	+		DQ103740
DRB5*0303	+		DQ103741
DRB5*0304	+		DQ103742
DRB6*0101	+		DQ103743
DRB6*0102	+		DQ103744
DRB6*0401	+	+	DQ103725
DRB6*0402	+		DQ103745
DRB6*0403	+		DQ103746
DRB6*0404	+		DQ103747
DRB*W301	+	+	DQ103717
DRB*W302	+	+	DQ103718
DRB*W401	+		DQ103730
DRB*W2501	+	+	DQ103715
DRB*W2502	+		DQ103726
DRB*W2503	+	+	DQ103716
DRB*W2504	+		DQ103727
DRB*W2601	+		DQ103728
DRB*W3701	+		DQ103729
DRB*W5101	+	+	DQ103719
DRB*W5201	+	+	DQ103720

Mafa- (*Macaca fascicularis*), *Maar-* (*Macaca arctoides*), *Male-* (*Mandrillus leucophaeus*) and *Paha-DRB* (*Papio hamadryas*). Phylogenetic trees were constructed using the Jukes and Kantor neighbour-joining method and the computer program MEGA version 2.1 (Kumar et al. 1993). The analyses were based on 252 nucleotides. Bootstrap analyses using 500 replications were performed to determine the repeatability of sequence-clustering patterns. A second analysis was conducted for mandrill and other OWM *Mhc-DRB6* sequences.

Statistical analyses

Masp-DRB sequences were analysed with MEGA version 2.1 software (Kumar et al. 1993). Variable amino acid positions, presumed to represent the PBR (Brown et al. 1993), were assigned as for rhesus macaques (Slierendregt et al. 1992). Rates of non-synonymous (d_N) and synonymous (d_S) substitutions within the PBR and non-PBR were calculated according to Nei and Gojobori (1986), with the Jukes and Cantor (1969) correction. Significance levels were determined using a Z test ($p<0.05$).

Results

Cloning and sequencing

Multiple exon 2 *Masp-DRB* sequences were amplified using 5′MDRB and 3′MDRB primers. Cloning and sequencing identified 11 different sequences, 252 nucleotides in length, in two individuals (Table 2). No sequences were identical to those previously published for other primates.

Separation of Masp-DRB sequences using DGGE

The optimal denaturing gradient for mandrills (41–50%) was determined using perpendicular DGGE for one mandrill used in the cloning and sequencing experiments. As with cloning and sequencing, multiple *Masp-DRB* sequences were amplified using the 5′MDRB and 3′MDRB (this time with a GC-clamp) primers. A direct comparison of results from cloning and sequencing and DGGE and direct sequencing revealed the same 11 sequences, with one additional sequence obtained using DGGE (Table 2). Authenticity of this sequence was confirmed by multiple DGGE experiments.

```
                   10          20          30          40          50          60          70          80          90
                    ↓           ↓           ↓           ↓           ↓           ↓           ↓           ↓           ↓
HLA-DRB1*010101    LWQLKFECHF  FNGTERVRLL  ERCIYNQEES  VRFDSDVGEY  RAVTELGRPD  AEYWNSQKDL  LEQRRAAVDT  YCRHNYGVGE  SFTVQRR
Masp-DRB1*0301     -EYSTS----  --------F-  D-YF-----Y  ---------F  ---S------  -----GR--I  --DE--S---  --------A*  *******
Masp-DRB1*0302     -EYSTS----  --------F-  D-YFH----L  ---------F  ---S-----E  --S---R---  ----------  ------R--*  *******
Masp-DRB1*0303     -EYSTS----  --------F-  D-YF-----Y  ---------F  ---S----LI  ---G--R--Y  --DE------  ------V*  *******
Masp-DRB1*0304     -EYSTS----  --------Y-  D-YF-----N  ------L--F  ----------  --N------I  --DE--S---  F-------V*  *******
Masp-DRB1*0305     -EYSTA----  --------F-  D-YFH----L  ----------  ---S-----E  --S---R---  ---K--S---  F-----R--*  *******
Masp-DRB1*0401     -E-V-H--R-  --------Y-  --LF-----F  LH----L--F  Q--------V  --NL--R---  ---K--E---  F-----R-V*  *******
Masp-DRB1*0402     -E-V-Y--R-  --------Y-  --LF-----F  ------L--F  ---------V  --N--R----  --DD--S---  F--------*  *******
Masp-DRB1*0403     -E-V-H----  --------F-  D-YF-----Y  ---------H  ----------  ------R--I  --DQ------  F--Y----V-  ----***
Masp-DRB1*0404     -E-V-H----  --------F-  D-YFH----L  ----------  ----------  --S-------  --RE------  ------V-  ----***
Masp-DRB1*0405     -E-A-H----  --------Y-  I-Y-H----L  ----------  ---S------  ----------  --R-------  ------R-V*  *******
Masp-DRB3*0401     -E-F-S----  -D------Y-  Q-HF-----F  ----------  ----------  ---L-----Y  ------QA-N  --------F-  ----***
Masp-DRB3*0402     -E-G-S----  --------Y-  Q-HF-----F  ----------  ----------  --------Y  ------Q--N  ---Y----V*  *******
Masp-DRB5*0301     -K-E------  ----------  H-Y------D  A---------  ----------  ----------  ----------  -------IV-  ----***
Masp-DRB5*0302     -K-E------  ----------  H-Y------D  A---------  ----------  ----------  --R---E---  --------V*  *******
Masp-DRB5*0303     -K-D-Y----  --------F-  H-Y------D  A---------  W---------  ---------I  ------E---  V-------V*  *******
Masp-DRB5*0304     -K-E------  ----------  D-Y------D  A---------  ----------  ----------  ----------  ---V--***  *******
Masp-DRB6*0101     -E-G-S--QI  ---R----Y-  N-N-HKR--N  L--------F  Q--M-V---V  --N-----GI  --EK-DK---  ---Y--R-F*  *******
Masp-DRB6*0102     -E-G----QI  ---R-W--Y-  N-N-HKR--N  L--------F  Q--M-V---V  --N-----GI  --EK-DK---  ---Y--R-F*  *******
Masp-DRB6*0401     -E-A-C---I  --------Y-  N-N-HKR--N  L--H--L--F  Q-----E--V  --N-----SI  --EK-DK---  ---Y--R-F-  ---E***
Masp-DRB6*0402     -E-A-CV--I  -------QY-  N-Y-HKR--N  L--H---L-F  Q-----EW-V  --N-----GI  --EK-DK-TP  TADTITEF**  *******
Masp-DRB6*0403     -E-A-C----  --------QY  N-Y-HKR-GN  L--H--L--F  Q-----EW-V  --N-----GI  --EKWDK--A  ---N----F*  *******
Masp-DRB6*0404     -E-A-C---I  -----P-QH-  N-YV-KR--N  L--H--L--F  !-EM---W-V  --N-----GI  --EK-DK--S  ---Y----F*  *******
Masp-DRB*W301      -E-A-R----  --------F-  D-YFH----Y  A--------F  --------RS  ---F-----F  ---A-T---N  ---Y------  ----***
Masp-DRB*W302      -E-A-R----  --------F-  D-YFH----Y  A--------F  --------RS  ---F-----F  ---A-T---N  ---Y------  ----***
Masp-DRB*W401      ---F-P----  --------F-  D-YF-----F  A--------F  ---S-----E  ------R---  --DE------  F-------L*  *******
Masp-DRB*W2501     -E-A-A----  --------F-  D-YVH----Y  A---------  ----------  ------R--V  --RA-T---N  ---Y------  ----***
Masp-DRB*W2502     -E-R-A----  --------Q--  D-FVH----Y  -----H---F  ---S-----V  -D-G--R--Y  ---E------  F------GV*  *******
Masp-DRB*W2503     -E-A-S----  --------F-  D-Y-H----L  ---------F  ----------  --S---R---  --RA------  V--Y--RG--  ----***
Masp-DRB*W2504     -E-A-S----  ----------  D-Y-H----F  ----------  ----------  --S---R--V  --RE------  -------GV*  *******
Masp-DRB*W2601     -E-A-C----  --------F-  --HF-----Y  ---------F  ----------  ---------Y  --D--V--.  I-------A*  *******
Masp-DRB*W3701     -E-A-S----  --------F-  D-Y-H----F  ---------F  ----------  --S--G---I  ---K--S---  --------F*  *******
Masp-DRB*W5101     -E-V-S----  --------F-  --YFH----F  ----------  --------RS  ------R---  --K-GQ--N  ---Y----F-  ----***
Masp-DRB*W5201     -E-A-S----  -D---Q--F-  Q-YF-----F  ---------F  ----------  ------R--I  --RE--E---  F-------V-  ---E***
```

Fig. 3 Predicted amino acid sequences of mandrill DRB (exon 2). *Numbering* of amino acids is shown above the consensus of *Masp-DRB* sequences; *bold* typeface indicates inferred PBR sites. Identity to consensus is shown by *dashes* (–). Deletions are indicated by *dots* (.) and a stop codon is indicated by an *exclamation point* (!). An *asterisk* (*) denotes an un-sequenced region

While a 41–50% parallel denaturing gradient could be used to separate most mandrill sequences (Fig. 1), this gradient was not optimal for some sequences in other study individuals and second gradient (40–60%) was required for four individuals (Fig. 2a,b; Table 3).

```
                      10                            20                            30
HLA-DRB1*010101 TTG TGG CAG CTT AAG TTT GAA TGT CAT TTC TTC AAT GGG ACG GAG CGG GTG CGG TTG CTG GAA AGA TGC ATC
Masp-DRB1*0301  --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T--
Masp-DRB1*0302  --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T--
Masp-DRB1*0303  --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T--
Masp-DRB1*0304  --- GA- T-C TC- -CA -C- --G --- --C --- --- --- --- --- --- --- --- -AC --- --C --- -A- T--
Masp-DRB1*0305  --- GA- T-C TC- -CA GC- --G --- --C --- --- --- --- --- --- --- --- --C --- --C --- -A- T--
Masp-DRB1*0401  --- GA- --- G-- --A CAC --G --- -G- --- -G- --- --C --- --- --- --- -AC --- --G --- CT- T--
Masp-DRB1*0402  --- GA- --- G-- --A -A- --G --- -G- --- -G- --- --C --- --- --- --- -AC --- --G --- CT- T--
Masp-DRB1*0403  --- GA- --- G-- --A CA- --G --- --- --- --- --- --C --- --- --- --- --C --- --C --- -A- T--
Masp-DRB1*0404  --- GA- --- G-- --A CA- --G --- --- --- --- --- --C --- --- --- --- --C --- --T --- -A- T--
Masp-DRB1*0405  --- GA- --- GC- --A CA- --G --- --C --- --- --- --- --- --- --- --- -AC --- ATC --- -A- ---
Masp-DRB3*0401  --- GA- --- T-- --- -C- --G --- --C --- --- G-C --- --- --- --- --- -AC --- C-G --- CA- T--
Masp-DRB3*0402  --- GA- --- GG- --- -C- --G --- --- --- --- --- --- --- --- --- --- -AC --- C-G --- CA- T--
Masp-DRB5*0301  --- AA- --- GAA --- --- --G --- --C --- --- --C --- --- --- --- --- --- --- C-C --- -A- ---
Masp-DRB5*0302  --- AA- --- GAA --- --- --G --- --C --- --- --C --- --- --- --- --- --- --- C-C --- -A- ---
Masp-DRB5*0303  --- AA- --- GA- --- -A- --G --- --C --- --- --C --- --- --- --- --- --C --- C-C --- -A- ---
Masp-DRB5*0304  --- AA- --- GAA --- --- --G --- --C --- --- --C --- --- --- --- --- --- --- C-C --- -A- ---
Masp-DRB6*0101  --- GAA --- GG- --- -C- --G --- --A A-- --- --C --- -G- --- --- --- -AC --- A-C --- AA- ---
Masp-DRB6*0102  --- GAA --- GG- --- --- --G --- --A A-- --- --C --- -G- --- T-- --- -AC --- A-C --- AA- ---
Masp-DRB6*0401  --- GA- --- GC- --A -G- --G --- A-- --- --- --T --- --- --- --- --- -AC --- A-C --- AA- ---
Masp-DRB6*0402  --- GA- --- GC- --A -G- -TG --- A-- --- --- --T --- --- --- -A- -AC --- A-C --- -A- ---
Masp-DRB6*0403  --- GA- --- GC- --A -GC --G --- A-- --- --- --T --- --- --- -A- -AC --- A-C --- -A- ---
Masp-DRB6*0404  --- GA- --- GC- --- -G- --G --- A-- --- --- --T --- -C- --- -A- CAC --- A-C --- -A- G--
Masp-DRB*W301   --- GA- --- GC- --A CG- --G --- --- --- --- --C --- --- --- --- --- --C --C --C --- -A- T--
Masp-DRB*W302   --- GA- --- GC- --A CG- --G --- --- --- --- --C --- --- --- --- --- --C --C --C --- -A- T--
Masp-DRB*W401   --- --- --- T-- --- CC- --G --- --- --- --- --C --- --- --- --- --- --C --- --C --- -A- T--
Masp-DRB*W2501  --- GA- --- GC- --- GC- --G --- --C --- --- --C --- --- --- --- --- --C --- --G --- C-- G--
Masp-DRB*W2502  --- GA- --- -G- --- GC- --G --- --C --- --- --- --- --- --- --- -A- C-C --- --C --- -T- G--
Masp-DRB*W2503  --- GA- --- GC- --- -C- --G --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- ---
Masp-DRB*W2504  --- GA- --- GC- --- -C- --G --- --- --- --- --C --- --- --- --- --- --C --- --C --- -A- ---
Masp-DRB*W2601  --- GAA --- GC- --- -G- --G --- --- --- --- --C --- --- --- --- --- --C --- --G --- CA- T--
Masp-DRB*W3701  --- GA- --- GC- --- -C- --G --- --- --- --- --C --- --- --- --- --- --C --- --T --- -A- ---
Masp-DRB*W5101  --- GA- --- G-- --- -C- --G --- --- --- --- --C --- --- --- --- --- --C --- --G --- -A- T--
Masp-DRB*W5201  --- GA- --- GC- --- -C- --G --- --C --- --- G-C --- --- --- -A- A-- --C --- C-G --- -A- T--


                      40                            50
HLA-DRB1*010101 TAT AAC CAA GAG GAG TCC GTG CGC TTC GAC AGC GAC GTG GGG GAG TAC CGG GCG GTG ACG GAG CTG GGG CGG
Masp-DRB1*0301  --- --- --G --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- T-A --- --- --- ---
Masp-DRB1*0302  C-- --- --G --- --- -TA --- --- --- --- --- --- --- --- --- -T- --- --- --- T-- --- --- --- ---
Masp-DRB1*0303  --- --- --G --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- T-- --- --- --- ---
Masp-DRB1*0304  --- --- --- --- --- AA- --- --- --- --- --- --- C-- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB1*0305  C-- --- --G --- --- -TA --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- ---
Masp-DRB1*0401  --- --- --G --- --- -T- C-- -A- --- --- --T C-- --- --- --- -T- -A- --- --- --- --- --- --- ---
Masp-DRB1*0402  --- --- --G --- --- -T- --- --- --- --- --- C-- --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB1*0403  --- --- --G --- --- -A- --- --- --- --- --- --- --- --- --- C-- --- --- --- --- --- --- --- ---
Masp-DRB1*0404  C-- --- --G --- --- -TA --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB1*0405  C-- --- --G --- --- CT- --- --- --- --- --- --- --- --- --- --- --- --- --- T-A --- --- --- ---
Masp-DRB3*0401  --- --- --G --- --- -T- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB3*0402  --- --- --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- T-- --- --- ---
Masp-DRB5*0301  --- --- --- --- --- GA- -C- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB5*0302  --- --- --- --- --- GA- -C- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB5*0303  --- --- --- --- --- GAT -C- --- --- --- --- --- --- --- --- T-- --- --- --- --- --- --- --- ---
Masp-DRB5*0304  --- --- --- --- --- GA- -C- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB6*0101  C-- --A -GG --- --- AA- C-- --- --- --- --- --- --- --- --- -T- -A- --- --- -T- --A G-- --- ---
Masp-DRB6*0102  C-- --A -GG --- --- AA- C-- --- --- --- --- --- --- --- --- -T- -A- --- --- -T- --A G-- --- ---
Masp-DRB6*0401  C-- --A -GG --- --- AA- C-- --- --- C-- --- --- C-- --- --- -T- -A- --- --- --A --- -A- --- ---
Masp-DRB6*0402  C-- --A -GG --- --- AA- C-- --- --- C-- --- --- --C --- --- -T- -A- --- --- --A --A --- -A- T--
Masp-DRB6*0403  C-- --A -GG --- -G- AA- C-- --- --- C-- --- --- C-- --- --- -T- -A- --- --- --A --- -A- T-- ---
Masp-DRB6*0404  --- --A -GG --- --- AAT C-- --- --- C-- --- --- C-- --- --- -T- TA- -A- --- -T- --A --A --- T--
Masp-DRB*W301   C-- --- --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --T --- --- --- --- --- ---
Masp-DRB*W302   C-- --- --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --A --- --- --- --- --- ---
Masp-DRB*W401   --- --- --G --- --- -T- --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- ---
Masp-DRB*W2501  C-- --- --- --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB*W2502  C-- --- --- --- --- -A- --- --- --- --- --- C-- --- --- --- -T- --A --- --- T-- --- --- --- ---
Masp-DRB*W2503  C-- --- --- --- --- -TA --- --- --- --- --- --C --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB*W2504  C-- --- --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB*W2601  --- --- --G --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB*W3701  C-- --- --G --- --- -T- --- --- --- --- --A --- --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB*W5101  C-- --- --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- ---
Masp-DRB*W5201  --- --- --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- ---
```

Fig. 4 Nucleotide sequence of mandrill DRB (exon 2). *Codon numbering* is shown above the consensus of *Masp-DRB* sequences. Identity to consensus is shown by *dashes* (–). A deletion is indicated by *dots* (.) and *asterisks* (*) denote unsequenced nucleotides

Fig. 4 (continued)

```
                              60                                      70
HLA-DRB1*010101 CCT GAT GCC GAG TAC TGG AAC AGC CAG AAG GAC CTC CTG GAG CAG AGG CGG GCC GCG GTG GAC ACC TAC TGC
Masp-DRB1*0301  --- --C --- --- --- --- --- G-- -G- --- --- A-- --- --A G-C GA- --- --- T-- --A --- --- --- ---
Masp-DRB1*0302  --- --A --- --- -C- --- --- -G- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB1*0303  -T- ATC --G --- --- G-- --- --- -G- --- --- TA- --- --A G-C GA- --- --- --- --- --- --- --- ---
Masp-DRB1*0304  --- --C --- --- A-- --- --- --- --- --- --- A-- --- --A G-C GA- --- --- T-- --- --- --- --- -T-
Masp-DRB1*0305  --- --A --- --- -C- --- --- -G- --- --- --- --- --- --- --A --- --- --- T-- --A --- --- --- -T-
Masp-DRB1*0401  --- -TC --- --- A-- -T- --- --T -G- --- --- --- --- --- --- -A- A-- --- -A- --- --- --- --- -T-
Masp-DRB1*0402  --- -TC --- --- A-- --- --- --T -G- --- --- --- --- --A G-C --- --- --- T-- --- --- --- --- -T-
Masp-DRB1*0403  --- --C --- --- --- --- --- -G- --- --- A-- --- --A G-C CA- --- --- --- --- --- --- --- -T-
Masp-DRB1*0404  --- --C --- --- -C- --- --- --- --- --- --- --- --- --- -G- GA- --- --- --- --- --- --- --- ---
Masp-DRB1*0405  --- --C --- --- --- --- --- --- --- --- --- --- --- --- -G- --- --- --- --- --- --- --- --- ---
Masp-DRB3*0401  --- -TC --- --- -T- --- --- --- --- --- TA- --- --- --- --- --- CA- -C- --- -A- --- --- --- ---
Masp-DRB3*0402  --- --C --- --- --- --- --- --- --- --- TA- --- --- --- --- --- CA- --- --- -A- --- --- --- ---
Masp-DRB5*0301  --- --C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB5*0302  --- --C --- --- --- --- --- --- --- --- --- --- --- --- -G- --- --- --- -A- --- --- --- --- ---
Masp-DRB5*0303  --- --C --- --- --- --- --- --- --- --- A-- --- --- --- --- --- --- -A- --- --- --- GTG ---
Masp-DRB5*0304  --- --C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Masp-DRB6*0101  --- -TC --G --- A-- --- --- --- --- --- -G- A-- --- --- G-- -A- --- -A- AA- --- --- --- --- ---
Masp-DRB6*0102  --- -TC --G --- A-- --- --- --- --- --- -G- A-- --- --- G-- -A- --- -A- AA- --- --- --- --- ---
Masp-DRB6*0401  --- -TC --G --- A-- --- --- --- --- --- AG- A-- --- --- G-- -A- --- -A- AA- --- --- --- --- ---
Masp-DRB6*0402  --- -TC --G --- A-- --- --- --- --- --- -G- A-- --- --- G-- -A- --- -A- AA- --. --- T-- --- ---
Masp-DRB6*0403  --- -TC --G --- A-- --- --- --- --- --- -G- A-- --- --- G-- -A- T-- -A- AA- --- --- G-- --- ---
Masp-DRB6*0404  --- -TC --G --- A-- --- --- --- --- --- -G- A-- --- --- G-- -A- --- -A- AA- --- --- T-- --- ---
Masp-DRB*W301   -G- AGC --G --- --- -TC --- --- --- --- --- T-T --- --- --- GC- --- A-- --- --- --- -AT --- ---
Masp-DRB*W302   -G- AGC --G --- --- -TC --- --- --- --- --- T-T --- --- --- GC- --- A-- --- --- --- -AT --T ---
Masp-DRB*W401   --- --A --- --- --- --- --- --- -G- --- --- --- --- --A G-C GA- --- --- --- --- --- --- -T- ---
Masp-DRB*W2501  --- --C --- --- --- --- --T -G- --- --- G-- --- --- -G- GC- --- A-- --- --- --- -A- --- --- ---
Masp-DRB*W2502  --- -TC --- --C --- G-- --- --- -G- --- --- TA- --- --- --- GA- --- --- --- --- --- -T- --- ---
Masp-DRB*W2503  --- --C --- --- -C- --- --T -G- --- --- --- --- --- -G- GC- --- --- --- --- --- GTG ---
Masp-DRB*W2504  --- -TC --- --- -C- --- --T -G- --- --- G-- --- --- -G- GA- --- --- --- --- --- --- --- ---
Masp-DRB*W2601  --- --C --- --- --- --- --- --- --- --- TA- --- --A G-C --- --- --- -T- --- --- -.. .T- ---
Masp-DRB*W3701  --- --C --- --- -C- --- G-- --- --- --- A-- --- --- --- -A- --- --- T-- --- --- --- --- ---
Masp-DRB*W5101  -G- AGC --- --- --- --- --T -G- --- --- --- --- --- --- -A- --- -G- CA- --T --- -A- --- ---
Masp-DRB*W5201  --- --C --- --- --- --- --- -G- --- --- A-- --- --- -G- GA- --- --- -A- --- --- --- -T- ---


                80                                      90
HLA-DRB1*010101 AGA CAC AAC TAC GGG GTT GGT GAG AGC TTC ACA GTG CAG CGG CGA G
Masp-DRB1*0301  --- --- --- --- --- --- -CG *** *** *** *** *** *** *** *** *
Masp-DRB1*0302  --- --- --- --- C-- --- --- *** *** *** *** *** *** *** *** *
Masp-DRB1*0303  --- --- --- --- --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB1*0304  --- --- --- --- --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB1*0305  --- --- --- --- C-- --- --- *** *** *** *** *** *** *** *** *
Masp-DRB1*0401  --- --- --- --- C-- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB1*0402  --- --- --- --- --- --- --- *** *** *** *** *** *** *** *** *
Masp-DRB1*0403  --- T-- --- --- --- --- -TG --- --- --- --- --- *** *** *** *
Masp-DRB1*0404  --- --- --- --- --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB1*0405  --- --- --- --- C-- --- -TG --- --- --- --- --- *** *** *** *
Masp-DRB3*0401  --- --- --- --- --- --- TT- --- --- --- --- --- *** *** *** *
Masp-DRB3*0402  --- T-- --- --- --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB5*0301  --- --- --- --T --- A-- -TG --- --- --- --- --- *** *** *** *
Masp-DRB5*0302  --- --- --- --T --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB5*0303  --- --- --- --- --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB5*0304  --- --- --- --T --- --- -TG *** *** *** *** *** *** *** *** *
Masp-DRB6*0101  --- T-- --T --- C-- --- TT- *** *** *** *** *** *** *** *** *
Masp-DRB6*0102  --- T-- --T --- C-- --- TT- *** *** *** *** *** *** *** *** *
Masp-DRB6*0401  --- T-- --T --- A-- --- TT- --- --- --- --- -A- *** *** *** *
Masp-DRB6*0402  --- T-- --- --- --A --- TT- *** *** *** *** *** *** *** *** *
Masp-DRB6*0403  --- A-- --T --- --- --- TT- *** *** *** *** *** *** *** *** *
Masp-DRB6*0404  --- T-- --T --- --- --- TT- *** *** *** *** *** *** *** *** *
Masp-DRB*W301   --- T-- --- --- --- --- --- --- --- --- --- --- *** *** *** *
Masp-DRB*W302   --- T-- --- --- --- --- --- --- --- --- --- --- *** *** *** *
Masp-DRB*W401   --- --- --- --- --- --- CT- *** *** *** *** *** *** *** *** *
Masp-DRB*W2501  --- T-- --- --- --- --- --- --- --- --- --- --- *** *** *** *
Masp-DRB*W2502  --- --- --- --- --- -G- -TG *** *** *** *** *** *** *** *** *
Masp-DRB*W2503  --- T-- --- --- C-- -G- --- --- --- --- --- --- *** *** *** *
Masp-DRB*W2504  --- --- --- --- --- -G- -TG *** *** *** *** *** *** *** *** *
Masp-DRB*W2601  --- --- --- --- --- --- -CG *** *** *** *** *** *** *** *** *
Masp-DRB*W3701  --- --- --- --- --- --- TT- *** *** *** *** *** *** *** *** *
Masp-DRB*W5101  --- T-- --- --T --- --- TT- --- --- --- --- --- *** *** *** *
Masp-DRB*W5201  --- --- --- --- --- --- -TG --- --- --- --- -A- *** *** *** *
```

Using DGGE and direct sequencing, 22 additional sequences were identified in the remaining 11 individuals.

Due to the use of hemi-nested primers for gel plug amplification, the sequences reported are 237 nucleotides

Fig. 5 Phylogenetic tree of 33 *Masp-DRB* sequences and other selected OWM sequences. All non-mandrill sequences are located in the IMGT MHC-NHP database (Robinson et al. 2003). The tree configuration was derived from nucleotide sequences using the neighbor-joining method of the MEGA program. *Numbers* on the branches refer to confidence probability values
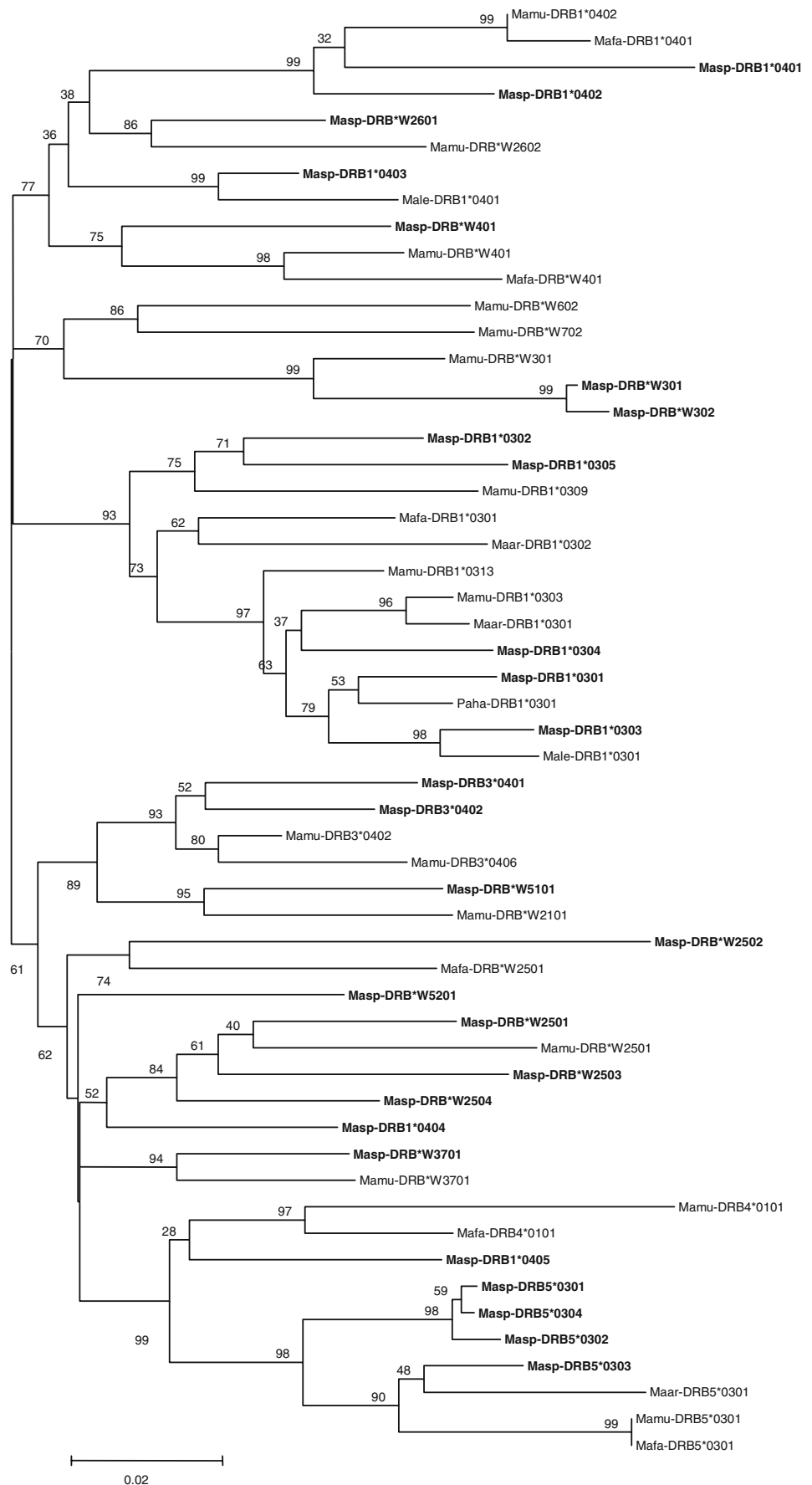
in length. Although sequences derived from cloning had no ambiguities and were longer in length, the DGGE method identified the complete sequence repertoire for both subjects tested. As some bands migrated to identical places on the DGGE, assumptions of sequence identity based on migration patterns were not possible for mandrills.

Identification of mandrill DRB alleles

A total of 33 different *Masp-DRB* sequences were identified using cloning and sequencing and/or DGGE and direct sequencing (Table 4). Each individual had a unique DGGE banding pattern (Fig. 1). Predicted amino acid sequences and nucleotide sequences are presented in Figs. 3 and 4, respectively. Up to seven different sequences were identified for a single individual (Table 3).

Phylogenetic analysis of the *Masp-DRB* sequences was performed to identify the locus and lineage of each new sequence. Trees included the new sequences and published OWM *Mhc-DRB* sequences (Fig. 5). Mandrill representatives of -DRB1, -DRB3, -DRB5, -DRB6 and -DRB*W loci and lineages were recognised. Five mandrill sequences clustered in the -DRB1*03 lineage and four sequences clustered both in the -DRB5 and -DRB*W25 loci. Two mandrill sequences clustered with other -DRB3 and -DRB*W301 sequences. Other single sequences clustered with other -DRB*W sequences. Sequences assigned to the -DRB1*04 lineage clustered with an assortment of other OWM sequences (i.e. -DRB1*04, -DRB4*01, -DRB*W101). Two sequences were assigned novel designations.

Six *Masp-DRB6* sequences were also identified and a separate phylogenetic tree was created to compare these with other pseudogene sequences (Fig. 6). One *Masp-DRB6* sequence clustered with the rhesus macaque *Mamu-DRB6*0101*, three with *Mamu-DRB6*0110*, and two clustered with the cynomolgus macaque *Mafa-DRB6*0101*.

Patterns of nucleotide substitution

Most nucleotide substitutions were confined to residues representing the PBR. Based on work by Brown et al. (1993), 17 positions were identified as most variable (i.e. four or more amino acid changes) and assumed to make direct contact with peptides. The variable residues were at positions 11, 13, 23, 26, 28, 30 and 37 of the β sheet and 57, 61, 67, 70, 71, 73, 74, 77, 78 and 86 of the α helix (Fig. 7). These positions are directly comparable to those described for rhesus macaques by Slierendregt et al. (1992).

Estimated rates of non-synonymous and synonymous substitutions within the PBR and non-PBR were also examined. The PBR of non-DRB6 sequences (Table 5) contained significantly more non-synonymous changes than synonymous changes ($p < 0.0001$). However, the region outside the PBR contained fewer non-synonymous than synonymous substitutions, and this difference was not significant ($p = 0.109$). Pseudogenes were examined separately, and no significant differences between frequency of non-synonymous and synonymous substitution were found (Table 6).

Discussion

Using DGGE and direct sequencing, as well as cloning and sequencing, a total of 33 *Masp-DRB* sequences were identified in 13 mandrills. A simple comparison of the first 30

Fig. 6 Phylogenetic tree of the *Masp-DRB6* sequences and other selected OWM-*DRB6* sequences. All non-mandrill sequences are located in the IMGT MHC-NHP database (Robinson et al. 2003)
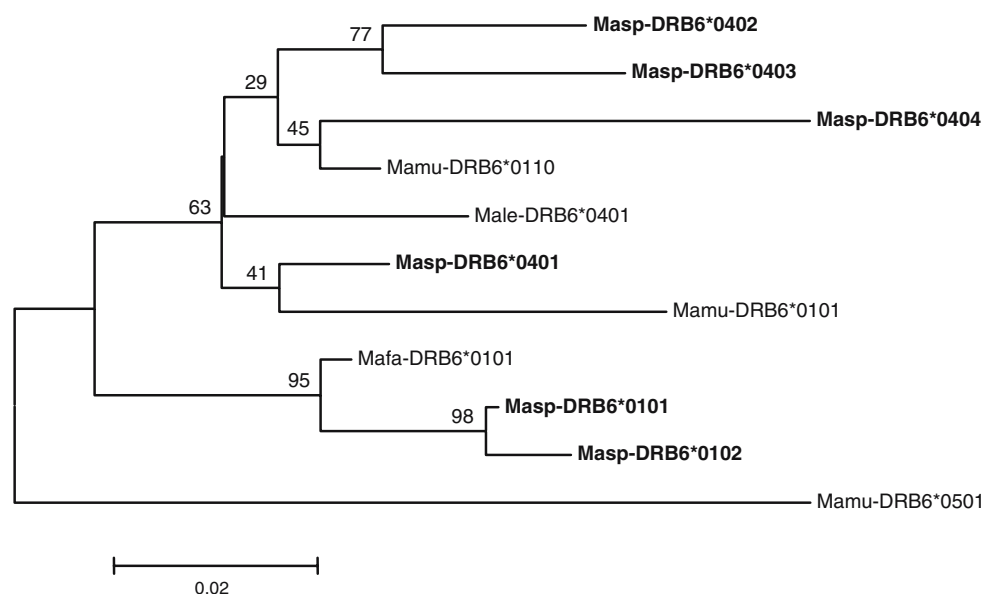
Fig. 7 Amino acid variability plot for 33 *Masp-DRB* alleles. The *horizontal axis* shows the amino acid position, and the *vertical axis* shows the number of different amino acids at a given position. The most variable sites in the rhesus macaque MHC (Slierendregt et al. 1992) are indicated by *asterisks* (*)
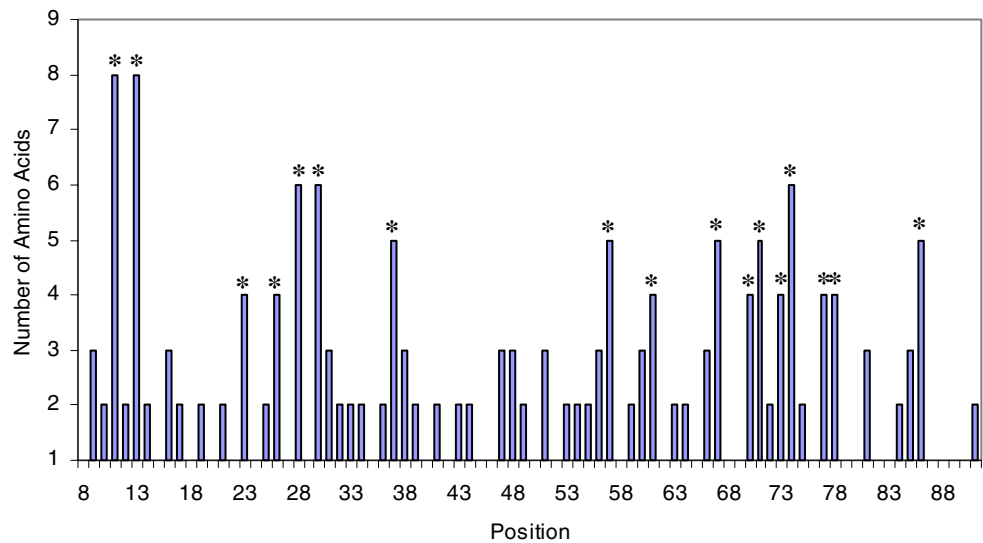
nucleotides of exon 2 demonstrates more than 91% shared sequence similarity between mandrills and other OWMs. Two sequences were given novel DRB*W designations.

Through analysis of the peptide-binding region, significantly higher rates of non-synonymous than synonymous substitutions within the *Masp-DRB* were found. Hughes and Nei (1989) argue that substitutions favouring change in amino acid configuration within the PBR indicate evidence for positive selection of Mhc diversity. These findings have important implications for our understanding of *Mhc-DRB* evolution in primates, but new research by De Groot et al. (2004) suggests that not all sequences may be expressed. De Groot and co-workers demonstrated that despite the assumption of their functionality, certain rhesus macaque sequences (i.e. -DRB1*0309, -DRB3*0403 and -W307) were not transcribed, rendering them non-functional.

Klein's (1987) trans-species inheritance hypothesis argues that many primate *Mhc-DRB* sequences share sequence similarities due to common ancestry. Thus, a comprehensive understanding of the primate MHC can only be gained by studying related species in depth. Based upon extensive work on chimpanzees (*Pan troglodytes*) and

rhesus macaques, it is increasingly obvious that there are major differences in frequency and variability of MHC genes of primates. According to Bontrop et al. (1999), the -DRB3, -DRB4, -DRB5 and -DRB6 loci, as well as the -DRB1*03, -DRB1*04, -DRB1*07 and -DRB1*10 lineages, are shared by humans, common chimpanzees and rhesus macaques. Some OWM and ape *Mhc-DRB* regions contain loci with no human equivalent (e.g. -DRB*W loci). Thus, some loci, or even lineages, may have been lost, while others will arise through duplication after divergence from a common ancestor. Many of these loci might be unique to some species. In the present study, two sequences are unique to mandrills (-DRB*W5101 and -DRB*W5201).

Knapp et al. (1997) identified 34 different sequences from 12 unrelated rhesus macaques. Similarly, 33 mandrill DRB sequences were identified from only 11 unrelated and two related individuals. The captive-born individuals shared five sequences with four of the wild-born individuals. Shared sequences between the two groups, as well as among the unrelated wild-born mandrills, indicate that some sequences may have high frequencies in mandrills. Analyses of the *Masp-DRB* sequences revealed distinct

Table 5 Estimated rates of non-synonymous (d_N) and synonymous (d_S) substitutions for protein binding (PBR) and non-protein binding (Non-PBR) regions for *DRB* exon 2 of mandrill sequences (excluding -DRB6)

Positions	# of codons	d_N	d_S	d_N/d_S	p
PBR	17	0.52±0.07	0.14±0.06	3.71	$p<0.0001$
Non-PBR	67	0.05±0.01	0.11±0.03	0.45	ns
Overall	84	0.13±0.02	0.11±0.02	1.18	ns

Significance values were calculated using the Z test, $p<0.05$
ns Not significant

Table 6 Estimated rates of non-synonymous (d_N) and synonymous (d_S) substitutions for PBR and non-PBR regions for *DRB* exon 2 of mandrill -DRB6 sequences only

Positions	No. of codons	d_N	d_S	d_N/d_S	p
PBR	17	0.10±0.04	0.07±0.07	1.43	ns
Non-PBR	67	0.07±0.02	0.09±0.04	0.78	ns
Overall	84	0.07±0.01	0.09±0.03	0.78	ns

Significance values were calculated using the Z test, $p<0.05$
ns Not significant

patterns of amino acid sequences (i.e. motifs) shared between humans, apes or monkeys. While shared motifs can be identified in many sequences, only those found at the amino end of the beta-1 domain are used to differentiate between loci and lineages (Gyllensten et al. 1991; O'hUigin 1995). For example, the amino acid motif of EYST occurs at positions 9–12 on -DRB1*03 alleles in humans, apes, Old and New World monkeys (Gyllensten et al. 1991; Klein and O'hUigin 1994; Bontrop et al. 1999). The presence of this motif in many different primates suggests descent either from a single ancestral allele or through convergent evolution in the case of New World monkeys (Antunes et al. 1998). In the mandrill, five sequences share the -DRB1*03 EYST motif, encoded by the same nucleotide sequence (GAG TAC TCT ACA), with the human and rhesus macaque. As no more than two -DRB1*03 sequences are found in any single individual, it is possible that these five sequences are allelic. However, this concept should be interpreted with caution, as reports have shown that macaques could have two -DRB1*03 sequences on a single haplotype (Doxiadis et al. 2003; Penedo et al. 2005). Further study of related individuals should provide important information about locus identity.

All other mandrill sequences belong to other -DRB1 lineages and -DRB3, 5 and 6 and -DRB*W loci. Although the -DRB1*01/10 and -DRB1*07 lineages are found in both humans and rhesus macaques, mandrills seem to lack these sequences. A study on cynomolgus macaques also reported the presence of the -DRB1*07 lineage (Leuchte et al. 2004). The mandrills in the present study also lack -DRB4 sequences. Despite the large number of sequences identified in the present study, representatives from missing lineages may be uncovered through further study of additional individuals.

While official sequence names are given on the basis of nucleotides found in the first part of exon 2, phylogenetic trees highlight difficulties in the nomenclature process. Three examples are apparent in Fig. 5. First, Masp-DRB*W5101 clusters with Mamu-DRB*W2101. Despite

their similarities (Fig. 8), distinct nucleotides found in Mamu-DRB*W21 alleles are not present in the Masp-DRB*W5101 sequence. Second, the Masp-DRB1*0405 sequence did not cluster with other Mhc-DRB1*04 sequences, but instead grouped with a Mamu-DRB4 sequence. Examination of the first 20 amino acids of Masp-DRB1*0405 demonstrates the similarities with other -DRB1*04 sequences, but other regions of the sequence share motif similarities with Mamu-DRB4*0101. These motifs may have arisen by positive selection or by other processes such as gene conversion or recombination. Lastly, the Masp-DRB1*0404 clusters with Mhc-DRB*W25 sequences and most closely with the Masp-DRB*W2504 sequence. In Fig. 8, these two sequences are compared with two similar Mamu sequences and their similarity to Mhc-DRB1*04 sequences are apparent. Interestingly, phylogenetic trees constructed with only non-PBR residues do not improve resolution of the Masp-DRB1*04 allelic lineage (data not shown). As mentioned previously, sequences are given workshop designations based on their similarity to HLA sequences. The observations made in the present study suggest that an MHC nomenclature system based on the HLA (a relatively recent primate species) may complicate naming of MHC sequences in the non-human species, and lineage identity does not imply similarity across an entire sequence.

The mandrill sequences with deletions and premature stop codons are also interesting. Two mandrill sequences contain nucleotide deletions. Masp-DRB6*0402 has a one nucleotide deletion at position 75 and Masp-DRB*W2601 has a three nucleotide deletions at positions 77–78. Another sequence, Masp-DRB6*0404, does not contain a deletion but a premature stop codon at position 48. Although deletions and premature stop codons are characteristic of pseudogenes, the nucleotide sequence of Masp-DRB*W2601 does not suggest it should be grouped with other pseudogenes. According to Paz-Artal et al. (1996), three-nucleotide deletions in -DRB6 alleles typically occur at position 71 in rhesus macaques (Mamu-DRB6*0112) and

```
HLA-DRB1*010101  LWQLKFECHF FNGTERVRLL ERCIYNQEES VRFDSDVGEY RAVTELGRPD AEYWNSQKDL LEQRRAAVDT YCRHNYGVGE SFTV
Masp-DRB*W5101   E.V.S.... .......F. ..YFH ...F ......... ........RS ......R... ..K.GQ..N ...Y..F. ....
Mamu-DRB*W2101   E.V.S.... .......F. ..YF. ...Y ......M... ...S...... ....GR.GY ..K.GQ..N ...Y..R... .--

HLA-DRB1*010101  LWQLKFECHF FNGTERVRLL ERCIYNQEES VRFDSDVGEY RAVTELGRPD AEYWNSQKDL LEQRRAAVDT YCRHNYGVGE SFTV
Mamu-DRB1*0402   .E.V.H..R. .......Y. ..HF. ...F LH....L..F ......... ...L...... .D...G... .......A. .--
Masp-DRB1*0405   .E.A.H.... .......Y. I.Y.H ...L ......... ...S...... ......... .R....... ......R.V. ....
Mamu-DRB4*0101   .E.A.C.... .......WN. I.Y.. ...Y A.YN..L... Q......... ......... .R...E... ....Y..... .--

HLA-DRB1*010101  LWQLKFECHF FNGTERVRLL ERCIYNQEES VRFDSDVGEY RAVTELGRPD AEYWNSQKDL LEQRRAAVDT YCRHNYGVGE SFTV
Mamu-DRB1*0404   .E.V.H... .......F. D.YH ...F ......... ......... ....F .D.Q... V .......V ...-
Masp-DRB1*0404   .E.V.H... .......F. D.YFH ...L ......... ......... ...S...... RE...... ......V. ----
Masp-DRB*W2504   .E.A.S... .......D.Y.H ...F ......... ......... ......V .S...R..V RE...... ...GV- ----
Mamu-DRB*W2501   .E.A.A... .......F. ..HF....L ......... ......... ......T .S...R..V RA.T.A. ......RG. ..—
```

Fig. 8 Selected Masp-DRB amino acid sequences and other OW locus/lineage representatives and phylogenetic clustered sequences. Bold typeface indicates variable residues for each species. Boxes enclose areas of similarities between sequences

at positions 77–78 in orangutans (*Pongo pygmaeus*: *Popy-DRB6*0201* and **0202*). Two-nucleotide deletions were reported for three chimpanzee -*DRB6* alleles (*Patr-DRB6*0106*, **0107* and **0108*) at position 77. Unlike the deletions reported by Paz-Artal et al. (1996), *Masp-DRB6*0402* has a one-nucleotide deletion at position 75. Although the nucleotide sequences of the two sequences are not similar, *Masp-DRB*W2601* possesses a three-nucleotide deletion at positions 77–78 like the orangutan pseudogenes. These findings suggest the possibility of trans-species inheritance of pseudogenes from the common ancestor of orangutans and mandrills.

According to Hughes and Nei (1989), non-synonymous substitutions, or changes that alter encoded amino acids, within the PBR should be more frequent than synonymous (non-altering) changes in functioning Mhc-DRB alleles. Concomitantly, a greater rate of non-synonymous substitution would be indicative of positive selection. Our analyses of nucleotide substitution in the mandrill sequences, based on interlocus comparisons, support Hughes and Nei's model (Table 5). As expected, the *Mhc-DRB* PBR contained significantly more non-synonymous substitutions than synonymous substitutions ($p<0.0001$). Therefore, as in most species, positive selection may be a major force in maintaining sequence diversity in mandrill DRB genes. Future studies on DRB inheritance in mandrills will offer important further information about allele assignment that can be used for intralocus substitution analyses. In addition, studies of these loci using mRNA should clarify which sequences remain functional in the mandrill -*DRB* repertoire.

Considerably less is known about the MHC of OWM other than rhesus macaques. Representatives of the -*DRB1*, -*DRB4*, -*DRB5*, -*DRB6* and -*DRB*W* loci have been described for some, but not all, OWM. Taken together, *Mhc-DRB* sequences of OWMs suggest trans-species inheritance of only some alleles. The present study of mandrill *Mhc-DRB* suggests that, like the rhesus macaque, other OWMs may have high levels of allelic variation with duplication, gene conversion and recombination, providing opportunities for new alleles to arise. Further study will shed light on structural similarities with other primates and clarify haplotype diversity in the mandrill.

References

Antunes SG, de Groot NG, Brok H, Doxiadis G, Menezes AA, Otting N, Bontrop RE (1998) The common marmoset: a New World primate species with limited Mhc class II variability. Proc Natl Acad Sci USA 95:11745–11750

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196:80–83

Beck S, Trowsdale J (2000) The human major histocompatibility complex: lessons from the DNA sequence. Annu Rev Genomics Hum Genet 1:117–137

Blankenburg A, Kaup FJ, Sauermann U (2003) Mhc class II DRB sequences of lion-tailed macaques (*Macaca silenus*). Tissue Antigens 62(3):267–269

Bontrop RE, Otting N, de Groot NG, Doxiadis GG (1999) Major histocompatibility complex class II polymorphisms in primates. Immunol Rev 167:339–350

Brown JH, Jardetzky T, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional structure of the class II histocompatibility antigen HLA-DR1. Nature 364:33–39

de Groot N, Doxiadis GG, de Groot NG, Otting N, Heijmans C, Rouweler AJM, Bontrop RE (2004) Genetic makeup of the *DR* region in rhesus macaques: gene content, transcripts, and pseudogenes. J Immunol 172:6152–6157

Doxiadis GGM, Otting N, De Groot NG, Bontrop RE (2001) Differential evolutionary MHC class II strategies in humans and rhesus macaques: relevance for biomedical studies. Immunol Rev 183:76–85

Doxiadis GGM, Otting N, De Groot NG, De Groot N, Rouweler AJ, Verschoor EJ, Bontjer I, Bontrop RE (2003) Evolutionary stability of the MHC class II haplotypes in diverse rhesus macaque populations. Immunogenetics 55(8):540–551

Gyllensten UB, 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 substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci USA 86:958–962

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic, New York, pp 21–132

Klein J (1987) Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. Hum Immunol 19:155–162

Klein J, Figueroa F (1986) Evolution of the major histocompatibility complex. CRC Crit Rev Immunol 6:295–386

Klein J, O'hUigin C (1994) MHC polymorphism and parasites. Philos Trans R Soc Lond B 346:351–358

Klein J, Satta Y, O'hUigin C, Takahata N (1993) The molecular descent of the major histocompatibility complex. Immunol Rev 11:213–244

Knapp LA, Cadavid LF, Eberle ME, Knechtle SJ, Bontrop RE, Watkins DI (1997) Identification of new Mamu-DRB alleles using DGGE and direct sequencing. Immunogenetics 45:171–179

Kriener K, O'hUigin C, Tichy H, Klein J (2000) Convergent evolution of major histocompatibility complex molecules in humans and New World monkeys. Immunogenetics 51:169–178

Kumar S, Tamura K, Nei M (1993) Molecular evolutionary genetics analysis. Pennsylvania State University, University Park

Leuchte N, Berry N, Köhler B, Almond N, LeGrand R, Thorstensson R, Titti F, Sauermann U (2004) MhcDRB-sequences from cynomologus macaques (*Macaca fascicularis*) of different origin. Tissue Antigens 63:529–537

Marsh SG, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Geraghty DE, Hansen JA, Hurley CK, Mach B, Mayr WR, Parham P, Petersdorf EW, Sasazuki T, Schreuder GMTh, Strominger JL, Svejgaard A, Terasaki PI, Trowsdale J (2005) Nomenclature for factors of the HLA system, 2004. Tissue Antigens 65:301–369

Nei M, Gojobori T (1986) Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Mol Bio Evol 3:418–426

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

Paz-Artal E, Corell A, Varela P, Martinez-Laso J, Gomez-Casado E, Fernandez-Soria VM, Moreno MA, Arnaiz-Villena A (1996) Primate DRB6 gene expression and evolution: a study in *Macaca mulatta* and *Cercopithecus aethiops*. Tissue Antigens 47:222–227

Penedo MCT, Bontrop RE, Heijmans CMC, Otting N, Noort R, Rouweler AJM, de Groot NG, Ward T, Doxiadis GGM (2005) Microsatellite typing of the rhesus macaque MHC region. Immunogenetics 57(3–4):198–209

Riess O, Kammerbauer C, Roewer L, Steimle V, Andreas A, Albert E, Nagai T, Epplen JT (1990) Hypervariability of intronic simple (gt) n(ga)m repeats in HLA-DRB genes. Immunogenetics 32:110–116

Robinson J, Waller MJ, Parham P, de Groot N, Bodmer JG, 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

Slierendregt BL, van Noort JT, Bakas RM, Otting N, Jonker M, Bontrop RE (1992) Evolutionary stability of trans-species major histocompatibility complex class II DRB lineages in humans and rhesus monkeys. Hum Immunol 35:29–39

Trowsdale J (1996) Molecular genetics of HLA class I and class II regions. In: Browning MJ, McMichael AJ (eds) HLA and MHC: genes, molecules and function. BIOS Scientific, Oxford, pp 23–38

Wickings EJ, Dixson AF (1992) Application of DNA fingerprinting to familial studies of Gabonese primates In: Martin RD, Dixson AF, Wickings EJ (eds) Paternity in primates: genetic tests and theories. Karger, Basel, pp 113–130

Zhu ZF, Vincek V, Figueroa F, Schönbach C, Klein J (1991) Mhc-DRB genes of the pigtail macaque (*Macaca nemestrina*): implications for the evolution of human DRB genes. Mol Biol Evol 8(5):563–578