

Original Synthesis Article

Major Histocompatibility complex-DMB allelic diversity in old and new world nonhuman primates: Intraspecies pattern of evolution

Maria José Recio¹, Ester Muniz¹, Jose Palacio-Grüber¹, Jorge Martinez-Laso², Eduardo Gomez-Casado³, Manuel Martín-Villa¹, Narcisa Martinez-Quiles¹, Cristina Campos¹, Antonio Arnaiz-Villena^{1*}.

¹Departament of Immunology, University Complutense, School of Medicine, Madrid Regional Blood Center, Madrid, Spain. ²Immunogenetics and Molecular Diagnosis Unit. Centro Nacional de Microbiología. Instituto de Salud Carlos III, Madrid, Spain. ³Dept. Inmunología Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Autopista A6, Hipódromo, Madrid.

*Corresponding author: E. mail: arnaizantonio@gmail.com. Web page: <http://chopo.pntic.mec.es/biolmol/>

Abstract - Nineteen different new MHC-DMB complete cDNA sequences have been obtained in thirteen different individuals belonging to the following primate species/families: *Hylobates lar*, *Papio hamadryas*, *Macaca mulatta*, *Macaca fascicularis*, *Cercopithecus aethiops* and *Saguinus oedipus*. Exonic allelism has been recorded all along the DM molecule domains and analyses of the critical residues in the conformation of the MHC-DR peptide-binding site were done; it was found an evolutionary pressure over the putative peptide-binding region of the DMB molecule that favours synonymous changes. These results are in contrast with the ones found in the MHC class I and class II genes, where non-synonymous DNA base substitutions are favoured. The immunoreceptor inhibition motif Tyr230-X231-X232-Leu233 (ITIM) is invariantly present in all extant studied primates since 40 million years ago. It confirms the important function for this molecule, directing DR molecules towards the endosomal/ lysosomal class II compartment and sending inhibitory signals to cells in order to stop synthesis of unnecessary MHC-DR molecules. Some *Macaca* individuals DMB molecules (appear on Earth more than ten million years ago) do bear both short (without ITIM) and long cytoplasmic tails (with ITIM), similarly to what has been found in human individuals. These differences may have important functional implications. Other molecules, like NK-cell receptors and Fc receptors, bear this type of tyrosine-based inhibitory motifs in order to switch off specific cell functions. MHC-DMB variations occurring within each species suggest that their polymorphism may have an intraspecific evolution, since alleles of the same species cluster together, as it occurs in other MHC related genes (Bf, C4d). Other MHC class I and class II molecules, phylogenetic trees show a trans-species pattern of evolution. Finally, a cluster grouping human and gorilla DMB cDNA sequences is obtained using a dendrogram (for the MHC genes, i.e.: C4d trees); this is in contrast to others' results that obtain a human/chimpanzee cluster using different DNA sequences.

Keywords: MHC (Major Histocompatibility Complex), MHC-DMB, HLA-DMB, primates, evolution, ITIM, MHC Class II metabolism, gibbon, macaque, New World Monkeys.

Introduction

Class II major histocompatibility complex (MHC) molecules are glycoproteins that play an essential role in the immune system. Five families of class II MHC genes have been described in humans (HLA-DP, -DR, -DQ, -DM, and -DO; Horton et al., 2004). Each one of them are constituted by two genes named A and B that encode α and β chains, respectively. DR, DQ and DP are cell-surface glycoproteins with a high degree of polymorphism with a presenting antigenic peptides to CD4 T lymphocytes function. HLA-DM proteins are almost absent from the surface of antigen presenting cells and they accumulate in certain intracellular compartments, including the MIIC (MHC class II compartment; Kelly et al., 1991; Pierre et al., 1996); this is caused by a tyrosine-based signal that is placed in the cytoplasmic tail of the DM β chain and corresponds to the residue 230 at exon 5 (Marks et al., 1995).

The role of DM molecules (Kelly et al., 1991; Sanderson et al., 1994a) has been described as a peptide exchange factor required to help removal of endogenous invariant chain (class II-associated invariant chain peptides or CLIP) from antigen presenting class II (DR, DQ, and DP) $\alpha\beta$ -CLIP complexes in the MIIC compartment, facilitating the generation of class II complexes; these can accommodate antigenic peptides that are later exported to the plasmatic membrane (Sloan et al., 1995; Denzin and Cresswell, 1995; Sherman et al., 1995). This export is modulated by DO molecules that inhibit DM-mediated catalysis of MHC-II peptide exchange. It has also been postulated that DM may play a role in the release of peptides other than CLIP, having a potential function as a peptide editor that selects a certain high-stability subpopulation of peptides for presentation at the cell surface (Kropshofer et al., 1997).

Following this postulated function, DM molecules do not necessarily have to bind the peptide itself in order to promote peptide exchange on class II molecules. Rather, they would act as chaperones interacting with the class II $\alpha\beta$ heterodimers promoting the release of low-stability peptides, thereafter stabilizing empty class II heterodimers in the acidic pH of the MIIC and facilitating a subsequent loading until a high-stability peptide occupies the groove (Kropshofer et al., 1997). These findings led to different authors to consider susceptibility to editing by DM is a key aspect for T cell epitope selection (Mellins and Stern, 2014). P1 pocket of the peptide binding groove has been recently described as a crucial factor for resistance to DM editing since peptides that retain this pocket occupancy are resistant to DM edition (Pos et al., 2013; Schulze et al., 2013). However, MHC Class II conformational changes could be considered as an important feature determining DM susceptibility (Mellins and Stern, 2014). This function could explain the limited polymorphism of DM molecules, in contrast to classical HLA class II molecules, found in humans (Sanderson et al., 1994b; Robinson et al., 2013).

DM proteins are structurally very distinct from conventional class II molecules, the deduced amino acid sequence of the $\beta 2$ domain (Kelly et al., 1991) shows only a 35% homology with DR. It has been suggested that *DM* genes could have split from the primordial class I/II genes several hundred million years ago, when class I and class II diverged from each other (Sanderson and Trowsdale, 1995). DM $\alpha 2$ and $\beta 2$ domains (and their paralogous mouse molecules, H-2M; Cho et al., 1991) are homologous to the membrane proximal domains of class I and II molecules (Sanderson and Trowsdale, 1995); DM $\alpha 1$ and $\beta 1$ domains, which in DR molecules form the peptide-binding groove, can be modelled into a tertiary structure similar to the DR archetype, although very few primary sequence homologies would be found and the peptide-binding groove would be closed.

In the present work, the complete *MHC-DMB* cDNA belonging to different species from different old world monkey families, *Hylobatidae* (which includes gibbon) and *Cercopithecidae* (which includes hamadryas baboon, rhesus macaque, crab-eating macaque and vervet monkey) and new world monkeys family *Callitrichidae* (which includes cotton-top tamarin), has been sequenced in order to compare the structure, polymorphism and evolution of this gene in apes and humans.

Material and Methods

Ape cell lines

Epstein-Barr Virus (EBV) cell lines from one *Hylobates lar* (Gibbon: H-861; *Hylobatidae* family), one *Papio hamadryas* (hamadryas baboon: 26CB1; *Cercopithecidae* family) and three *Saguinus oedipus* (cotton-top tamarin: 3F1, 5H5, and B95-8; *Callitrichidae* family) and renal embryonic cell culture from members of *Cercopithecidae* family: three *Macaca mulatta* (rhesus macaque; identified by codes Rh-3, Rh-6, and Rh-8), two *Macaca fascicularis* (crab-eating macaque; identified by codes Cy-9 and Cy-11) and three *Cercopithecus aethiops* (vervet monkey; identified by codes Ve-6, Ve-7 and Ve-9), were used to sequence the complete cDNA of the *DMB* locus. The cell lines from *Hylobates lar* and the cell lines from *Papio hamadryas* and *Saguinus oedipus* were obtained from American Tissue Cell Collection (ATCC) repository and renal embryonic cell culture from *Cercopithecidae* family members were obtained from the Whyttaker Bioproducts manufacturer (Walkersville, Maryland, USA). For the comparison analyses, *MHC-DMB* sequences of apes belonging to *Hominidae* family (chimpanzee, gorilla and orangutan; Alvarez et al., 1998) have also been used.

mRNA extraction/cDNA synthesis

RNA was extracted from EBV-cell lines (10^7 cells per line) using a NP-40 protocol (Sambrook et al., 1989). Cells were pelleted and resuspended in 250 μ l of lysis buffer (10 mM Tris pH 8.6, 140 mM NaCl, 1.5 mM MgCl₂, 0.5% NP-40) with RNAsin 1000 U/ml (final concentration) (Promega, Madison, WI, USA). After 1 min incubation at 4°C, lysates were spun down and supernatants were transferred into 250 μ l of proteinase K buffer 2X (200 mM Tris pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% SDS) with 200 μ g of proteinase K. These samples were vortexed and incubated for 30 min at room temperature. A standard phenol-chloroform extraction was done to obtain whole RNA. cDNA synthesis using a dT₁₆ oligonucleotide was performed using a Reverse Transcription System (Promega, Madison, WI, USA) according to the manufacturer's protocols.

Table 1. Distribution of the MHC-DMB alleles and number of clones sequenced in each species.

Species	Individual	Allele	No. clones
<i>Hylobates lar</i> (Gibbon)	H-861	<i>Hyla</i> -DMB*01	3
<i>Papio hamadryas</i> (Hamadryas baboon)	26CB1	<i>Paha</i> -DMB*01	3
	26CB1	<i>Paha</i> -DMB*02	3
<i>Macaca mulatta</i> (Rhesus macaque)	Rh-8	<i>Mamu</i> -DMB*01	8
	Rh-3, Rh-6, Rh-8	<i>Mamu</i> -DMB*02	14
	Rh-6	<i>Mamu</i> -DMB*03	9
	Rh-3	<i>Mamu</i> -DMB*04	4
<i>Macaca fascicularis</i> (Crab-eating macaque)	Rh-3	<i>Mamu</i> -DMB*05	3
	Cy-9	<i>Mafa</i> -DMB*01	4
	Cy-11	<i>Mafa</i> -DMB*02	8
	Cy-9, Cy-11	<i>Mafa</i> -DMB*03	3
<i>Cercopithecus aethiops</i> (Vervet monkey)	Ve-6	<i>Ceae</i> -DMB*01	3
	Ve-9	<i>Ceae</i> -DMB*02	7
	Ve-7	<i>Ceae</i> -DMB*03	6
	Ve-7	<i>Ceae</i> -DMB*04	4
	Ve-9	<i>Ceae</i> -DMB*05	3
<i>Saguinus oedipus</i> (Cotton-top tamarin)	3F1, 5H5	<i>Saoe</i> -DMB*01	5
	3F1, 5H5, B95-8	<i>Saoe</i> -DMB*02	4
	B95-8	<i>Saoe</i> -DMB*03	3

5 μ l of cDNA was amplified in a PCR reaction mixture that consisted of 50 mM KCl, 10 mM Tris-HCl, 1.5 mM MgCl₂, 0.01% gelatin, 0.2 mM of each dNTP and 2.5 Units of AmpliTaq (Perkin Elmer, Nonvank, CT). 100 pmoles of each of the DM specific primers 5' (5'-GGGGACTGAACTCCCGGCATC) and 3' (5'-GTGGAAGTTGTAGGATTCTGCCTC) were used. In order to obtain the complete sequence of cDNA, PCR was performed in a 9600 Thermocycler (Perkin Elmer, Nonvank, CT) under the following conditions: 94°C 20 s, 55°C 60 s, 72°C 1 min 30 s

for 40 cycles. MHC-DMB has 6 exons, adding up to 792 bp. Amplified cDNA fragment contains 783 bp.

cDNA sequencing

Random sequencing of the amplified products was carried out PCR products were purified and inserted into the pGEM-T vector (Promega, Madison, WI, USA).

Double-stranded DNA was automatically sequenced by using M13 universal primers and an internal one, DMBINT (5'-CAAGACACCCTGATGCAG) in an Applied Biosystems machine 3130XL (Foster City, CA, USA), as previously described (Arnaiz-Villena et al., 1992; Paz-Artal et al., 1994).

Rate comparison between synonymous and non-synonymous substitutions at peptide-binding sites

Comparison of the mean number of nucleotide substitutions *per* non-synonymous sites and the rate of substitutions *per* synonymous sites between the antigen-binding sites and the remaining sequence of exon 2 was made by using the Nei and Gojobori's method included in the NAG package of programs (Nei and Gojobori, 1986). Standard errors were calculated using Nei and Jin's method (Nei and Jin, 1989). Statistical significance was calculated with the chi-square test with Yates' correction. This comparison was done according to (Bjorkman et al. 1987a, b) results, which defined MHC protein binding and not-binding sites.

Dendrogram construction

Phylogenetic trees based on exon sequences were made using the neighbour-joining method (Saitou and Nei, 1987) and the computer programs MEGA 6 (<http://www.megasoftware.net/>). Phylogenetic trees were based on the genetic distance matrix of nucleotide substitutions per site obtained with the Kimura's biparametric method. The bootstrapping analysis for tree topology was used to evaluate the statistical significance of each internal branch of the tree. Bayesian and Maximum-likelihood methodologies were used, as controls.

A total of nineteen different new MHC-DMB alleles were found (1 in *Hylobates lar*, 2 in *Papio hamadryas*, 5 in *Macaca mulatta*, 3 in *Macaca fascicularis*, 5 in *Cercopithecus aethiops*, and 3 in *Saguinus Oedipus*; see Appendix and Table 1). All of these non-human primate alleles also showed the high number of cysteines (Cys) observed in human DMB molecules (positions 11, 25, 35, 46, 79, 117 and 174; Fig. 1) and also *Hylobates lar* show a possible glycosylation site described for humans (Asn at position 92) (Fig. 1).

Specific nucleotide changes in Hylobatidae family

The polymorphism detected in *Hylobates lar* is defined by three changes at positions 144 (AAG→AAA; Lys), 232 (CAG→AAT; Gln→Asn), and 470 (ATG→AGC; Met→Ser) (Fig. 1 and Appendix).

Specific nucleotide changes in Cercopithecidae family

Shared in all alleles of the family. DMB alleles of species of *Cercopithecidae* family show common specific nucleotide changes at positions 8 (ACA→ATA; Thr→Ile), 49 (GGA→AGA; Gly→Arg), 90 (CTG→CTA, Leu), 91 (TTG→CTG; Leu), 109 (CCA→GCA; Pro→Ala), 129 (TGC→TGT, Cys), 185 (GCC→GTC; Ala →Val), 292 (GCC→ACC; Ala→Thr), 462 (AAG→AAT; Lys→Asn), 478 (AGC→GGC; Ser→Gly, except in *Paha*-DMB*02), 519 (ACA→ACG, Thr), and 535 (CAT→TAT; His→Tyr) (Fig. 1 and Appendix).

Shared between Macaca and Paha alleles. DMB alleles of *Macaca* and *Papio hamadryas* show certain characteristics not shared by *Cercopithecus aetiops*: changes at positions 63 (TTC→TTT, Phe), 226 (CTC→ATC, Leu →Ile), and 495 (ACT→ACG, Thr, except in *Paha*-DMB*02) (Fig. 1 and Appendix).

Specific of Macaca alleles. DMB alleles in *Macaca* genus show changes that differentiate them from the other *Cercopithecidae* family alleles: substitutions at position 18 (CCG→CCA, Pro; exclusive of *Mamu*-DMB*02 allele and shared with *Saguinus oedipus* alleles); at position 111 (CCA→GCG, Pro→Ala) present in *Mamu*-DMB*02, -DMB*03, -DMB*04 and -DMB*05 alleles, at position 214 (TTG→GTG, Leu→Val) present in *Mamu*-DMB*02, -DMB*03, -DMB*05 and *Mafa*-DMB*01 alleles; at position 423 (TAT→TAC, Tyr) exclusive of *Mamu*-DMB*01 and *Mafa*-

DMB*02; and at position 780 together with non specific change at position 778 (TGG→CGA, Trp→Arg) present in *Mamu*-DMB*03, DMB*04, *Mafa*-DMB*02 and -DMB*03 alleles. In *Mamu*-DMB*01 this change is not combined with the substitution at position 778 generating a stop codon (TGG→TGA) (Fig. 1 and Appendix). Three different alleles have been found in one individual named Rh3, DMB1*02, *04, and *05 (see Table 1) postulating the possibility of a duplication in a DMB locus, at least in one chromosome. This event has been previously observed in other MHC loci from the same species (Paz-Artal et al., 1994).

Specific of Paha alleles. Baboon DMB alleles (*Paha*-DMB*01 and DMB*02) share synonymous changes at position 339 (CGG→CGT, Arg) and at position 744 (TAC→TAT, Tyr) that make them different from the rest of primates. However each allele show specific changes not shared with the rest of alleles described: *Paha*-DMB*01 presents substitutions at positions 145 and 146 (GAT→TGT, Asp→Cys), 193 and 194 (GAA→ACA, Glu→Thr), 397 (CTG→TTG, Leu), and 426 (CCA→CCC, Pro). *Paha*-DMB*02 presents a specific change at position 590 (ATT→ACT, Ile→Thr) (Fig. 1 and Appendix).

Specific of Cercopithecus aetiops alleles. Finally, vervet monkey alleles differ from each others in substitutions at position 36 (AGC→AGT, Ser) for *Ceae*-DMB*01, -DMB*04 and -DMB*05 alleles; at position 214 (TTG→ATG, Leu→Met, for *Ceae*-DMB*03, and TTG→GTG, Leu→Val, for the rest of *Ceae*-DMB alleles); and substitution at position 381 (ACA→ACG, Thr) exclusive of *Ceae*-DMB*01 allele (Fig. 1 and Appendix).

Specific nucleotide changes in Callitrichidae family

Three new MHC-DMB alleles for the *Sanguinus oedipus* were found and share common characteristics that differ them from other primate groups studied: changes at positions 50 (GGA→GAA, Gly→Glu), 108 (ACT→ACG, Thr), 122 (ACA→AAA, Thr→Lys), 232 (CAG→GAT, Gln→Asp), 286 (AAT→GAT, Asn→Asp), 362 (AAA→CGA, Lys→Arg), 408 (TAT→TAC, Tyr), 437 (ACT→ATC, Thr→Ile), 464 (CTT→CCT, Leu→Pro), 471 (ATG→ATT, Met→Ile), 488 (CAC→CGG, His→Arg), 498 (GCC→GCT, Ala), 529 (CTC→GTC, Leu→Val), 534 (TCC→TCT, Ser), 566 (ACT→ATC, Thr→Ile), 614 (CGG→CAG, Arg→Gln), 648 (ACC→ACG, Thr), 649 (CTG→GTG, Leu→Va), 675 (ACT→ACG, Thr), 676 (CTG→TTG, Leu), 702 (TCT→TCA, Ser), 726 (AGA→AGG, Arg), 741 (AGT→AGC, Ser), 746 (ACT→ATT,

Thr→Ile), 771 (TCA→CCG, Ser→Pro), and 776 (GGA→GCA, Gly→Ala). Also, substitutions in three nucleotides for the same codon (at positions 211-213 (AGC→CTG, Ser→Leu) and in two nucleotides for the same codon (at positions 228 and 228 (CTC→TTA, Leu) are characteristic for the three *Saoe*-DMB alleles. However, each allele shows specific changes not shared with the rest of *Saoe*-DMB alleles described with synonymous changes: *Saoe*-DMB*03 at position 45 (TGC→TGT, Cys) and *Saoe*-DMB*01 and -DMB*03 at position 462 (AAG→AAA, Lys) (Fig. 1 and Appendix).

These results do not support the possibility that MHC-DM would act as an antigen-presenting molecule in these species; this is suggested by the limited allelic polymorphism found and also by the fact that amino acid changes are not accumulated in the three hypervariable regions of exon 2, as described for the MHC-DRB, DPB or DQB molecules.

Analysis of rates of synonymous and non-synonymous substitutions

The rate of synonymous DNA base substitutions (d_s) and the rate of non-synonymous substitutions (d_n) have been obtained in order to determine the evolution of the different domains and functional regions of the MHC-DMB molecule. The value of the ratio between these two rates (d_n/d_s) is always lower than 1, both in total DMB and in different parts of the molecule.

From the results obtained (Table 2), an evolutionary pressure over the putative peptide-binding region favouring synonymous changes have been detected; these results are in contrast with the ones found in the MHC class I and class II peptide binding site, where non-synonymous DNA base substitutions are favoured. However, in these genes, the regions that do not participate in peptide binding site have similar values to DMB molecule. The region less conserved in MHC-DMB is the cytoplasmic domain ($d_n/d_s=0.92$). This fact indicates that both synonymous and non-synonymous substitutions are favoured, therefore, this region is subjected to neutral evolution and it could suggest that is not crucial for the molecule function.

However, when this ratio is calculated taking into account *Hominidae* (Alvarez et al., 1998) and *Cercopithecidae* (this study) as two different groups, it can be observed that in the first group, the cytoplasmic domain suffers more non-synonymous ($d_n=4.3$) than synonymous changes ($d_s=0$) which implies a no conservative pattern of evolution

while the second group shows the opposite situation ($d_n=0$, $d_s=5.7$), a conservative evolution, as in the whole DMB molecule.

Table 2. Mean number of nucleotide substitutions *per* 100 synonymous sites and *per* 100 non-synonymous sites of the *MHC-DMB* gene.

	Total DMB	LP vs. non-LP	$\beta 1$ domain vs. non- $\beta 1$ domain	$\beta 2$ domain vs. non- $\beta 2$ domain	CP vs. non-CP	TM vs. non-TM	CyP vs. non-CyP
$d_s \pm SE^*$	6.5 ± 1.1	5.1 ± 2.7 vs. 6.6 ± 1.1	8.9 ± 2.6 vs. 5.5 ± 1.1^a	6.7 ± 1.8 vs. 6.4 ± 1.3	2.2 ± 2.2 vs. 6.7 ± 1.1^b	3.8 ± 2.0 vs. 6.8 ± 1.2^a	5.4 ± 2.6 vs. 6.6 ± 1.1
$d_n \pm SE^*$	3.1 ± 0.5	3.4 ± 2.1 vs. 3.1 ± 0.5	3.8 ± 1.0 vs. 2.8 ± 0.5	2.8 ± 0.7 vs. 3.3 ± 0.6	0.8 ± 0.8 vs. 3.2 ± 0.5^b	0.8 ± 0.8 vs. 3.3 ± 0.5^b	5.0 ± 2.0 vs. 3.0 ± 0.5^a
d_n/d_s	0.47	0.66 vs. 0.46	0.42 vs. 0.50	0.41 vs. 0.50	0.36 vs. 0.47	0.21 vs. 0.48 ^a	0.92 vs. 0.45 ^b

The sequence used in these comparisons were: HLA-DMB*01:01 (2), *Patr*-DMB*01, *Patr*-DMB*02, *Papa*-DMB*01, *Papa*-DMB*02, *Gogo*-DMB*01, *Gogo*-DMB*02, *Gogo*-DMB*03, *Popy*-DMB*01, *Popy*-DMB*02, *Popy*-DMB*03 (17), *Hyla*-DMB*01, *Paha*-DMB*01, *Paha*-DMB*02, *Mamu*-DMB*01, *Mamu*-DMB*02, *Mamu*-DMB*03, *Mamu*-DMB*04, *Mamu*-DMB*05, *Mafa*-DMB*01, *Mafa*-DMB*02, *Mafa*-DMB*03, *Ceae*-DMB*01, *Ceae*-DMB*02, *Ceae*-DMB*03, *Ceae*-DMB*04, *Ceae*-DMB*05, *Saoe*-DMB*01, *Saoe*-DMB*02, and *Saoe*-DMB*03 (present paper)

LP: leader peptide; CP: connecting peptide; TM: transmembrane; CyP: cytoplasmic

d_s : mean number of nucleotide substitutions *per* 100 synonymous sites

d_n : mean number of nucleotide substitutions *per* 100 non-synonymous sites

*: d_s and d_n are estimated by Nei and Gojobori's method (Nei and Gojobori, 1986)

SE: standard error of mean calculated by Nei and Jin's method (Nei and Jin, 1989)

a: statistical significance $p < 0.01$

b: statistical significance $p < 0.001$

Conserved residues at exons 4, 5 and 6: an immunoreceptor tyrosine-based inhibition motif (ITIM) preserved at least 40 million years in DM molecules

MHC-DMB exons 4, 5 and 6 encode the corresponding classical class II cytoplasmic tail. The motif Tyr-Thr-Pro-Leu at positions 230-231-232-233 is found in this region in both man and mouse; this motif is responsible for targeting the heterodimer DM $\alpha\beta$ to lysosomal or MIIC compartments. However, several non-synonymous substitutions have been found in the presently described new ape alleles (see Table 3). The two most important positions are Tyr 230 and Leu 233 (Marks et al.,

1995) and these are conserved in human and nonhuman primate DMB alleles studied (Alvarez et al., 1998 and this study); this is consistent with the postulated functional importance of residues 230 and 233. Interestingly, Tyr-X-X-Leu is one motif present in inhibitory immunoreceptors, which exert their function through this type of intracytoplasmic tyrosine motifs (ITIMs); they are used for the negative regulation of many cell activation programs and are present in Fc receptors, NK-cell inhibitory receptors, CD22 β , CTLA-4, gp49B1 and MAFA lectins (Daeron et al., 2008; del Rio et al., 2013). It is feasible that DM molecules may use this motif not only for heading to the MIIC lysosomal/endosomal compartment in order to meet DR/CLIP molecules, but also to transmit an inhibitory signal to the cell once they have edited and loaded DR molecules with the corresponding antigenic peptides. This signal may stop DR synthesis once exogenous antigenic peptides are all loaded. However, residue 232 is also invariant in studied primates and could also be important in the transport of DMB molecules through cell compartments. In contrast, residue 231 is changed by Ile in the three cotton-top tamarin alleles. An ITIM motif is also found in mouse MHC-DM molecules but in another cytoplasmic position (Cho et al., 1991); this suggests the possible functional importance of this motif. On the other hand, other variations are observed in DMB ape residues, close to the cytoplasmic carboxil terminus (see Table 3); these are quite dramatic, non-conservative variations, i.e. Ser \rightarrow Pro at codon 239 and Trp \rightarrow Arg at codon 242 (except in *Mamu*-DMB*01 in which Trp changes to a stop codon) in all the new alleles described and Gly \rightarrow Ala, at codon 241, in *Saoe*-DMB*01, -DMB*02 and -DMB*03. This may imply that this part of the cytoplasmic tail is not important for the function of DMB molecules; however, a different functionality for different alleles may not be discarded, since the changes are not randomized and only two possibilities are found for each variable residue (Table 3) in human and ape species, which were separated at least 7 million years ago (Goodman et al., 1990). More ape sequences and targeted mutation experiments would clarify this point.

On the other hand, the presence of two variants of cDNA, one long (with exon 5) and another short (without exon 5) is observed in MHC-DMB alleles of *Macaca* species analysed in this study. Two hypotheses can be raised, one in which the DMB gene in macaques undergoes alternative splicing resulting in two types of different RNA messengers, and therefore two different protein isoforms, one normal and one with a shorter cytoplasmic region lacking the ITIM motif as previously described in humans

(Modrek et al., 2001). The other one is based in the DMB gene duplication in these species, thus, one copy of the gene could introduce a mutation in exon 4 that eliminates the signal processing and therefore originate messenger RNA without exon 5. However this is unlikely since the nucleotide sequences of the variants without exon 5, identified in these individuals, are identical to the sequences with exon 5. Thus, it seems feasible that these cDNAs represent different transcripts of the same gene. Given the importance of HLA-DM in antigen processing, this regulation is interesting from a functional point of view. The fact that the two DMB protein isoforms in macaques are present, probably, implies that arrival of the DM molecule to class II compartments takes place through two different routes, perhaps increasing efficiency in antigen presentation. Thus, these two DMB isoforms existed unless ten million years ago. The arrival of HLA-DM to endosomal compartments, including or deleting the ITIM motif is regulated by alternative splicing of human DMB gene (Modrek et al., 2001). However, functional studies in these individuals might help to clarify this point.

Topologies of MHC-DMB evolutionary trees (Fig. 2)

Neighbor-joining trees were constructed using both the complete cDNA (Fig. 2). Only one HLA-DMB cDNA sequence is available for constructing Fig. 2 (Kelly et al., 1991; Robinson et al., 2013), and an exon 3 DMB dendrogram was also used since seven different HLA-DMB sequences were available. Results in both dendrograms, using full cDNA and exon 3 DMB sequences, are concordant.

Tree shows two differentiated clusters (Fig.2): one corresponding to *Saguinus oedipus* alleles belonging to the *Platyrrhini* clade, *Callitrichidae* family, and a second one clustering alleles belonging to the *Catarrhini* clade. This second group is divided in other two groups, one corresponding to the *Hominoidea* superfamily and the other one belonging to the *Cercopithecoidea* superfamily, *Cercopithecidae* family, and *Cercopithecinae* subfamily. The *Hominoidea* superfamily part of the dendrogram is formed by two clusters, one corresponding to the *Hylobates lar* (gibbon) sequences corresponding to the *Hylobatidae* family and the second one belonging to the *Hominidae* family. The latter is also divided in other two clusters, the first one with the Human, Gorilla, and Chimpanzees sequences of the *Homininae* subfamily and the other one belonging to the Orangutan sequences of the *Ponginae* family (see Fig. 2). Therefore, MHC-DMB gene could be a good genetic marker to study the classification of primates from the clade (parvorder) to the subfamilies as shown in Figure 2.

Table 3. Most important variable residues of ape MHC-DMB alleles at cytoplasmic domain (exon 4, 5 and 6). Existing ITIM (Immunoreceptor tyrosine-based inhibition motif) forming residues are indicated.

Allele	Position							
	ITIM Motif				239	241	242	245
	230	231	232	233				
HLA-DMB*01:01	Tyr	Thr	Pro	Leu	Ser	Gly	Trp	Ser
<i>Hyla</i> -DMB*01	—	—	—	—	Pro	—	Arg	—
<i>Paha</i> -DMB*01	—	—	—	—	Pro	—	Arg	—
<i>Paha</i> -DMB*02	—	—	—	—	Pro	—	Arg	—
<i>Ceae</i> -DMB*01	—	—	—	—	Pro	—	Arg	—
<i>Ceae</i> -DMB*02	—	—	—	—	Pro	—	Arg	—
<i>Ceae</i> -DMB*03	—	—	—	—	Pro	—	Arg	—
<i>Ceae</i> -DMB*04	—	—	—	—	Pro	—	Arg	—
<i>Ceae</i> -DMB*05	—	—	—	—	Pro	—	Arg	—
<i>Mamu</i> -DMB*01	—	—	—	—	Pro	—	***	—
<i>Mamu</i> -DMB*02	—	—	—	—	Pro	—	Arg	—
<i>Mamu</i> -DMB*03	—	—	—	—	Pro	—	Arg	—
<i>Mamu</i> -DMB*04	—	—	—	—	Pro	—	Arg	—
<i>Mamu</i> -DMB*05	—	—	—	—	Pro	—	Arg	—
<i>Mafa</i> -DMB*01	—	—	—	—	Pro	—	Arg	—
<i>Mafa</i> -DMB*02	—	—	—	—	Pro	—	Arg	—
<i>Mafa</i> -DMB*03	—	—	—	—	Pro	—	Arg	—
<i>Saoe</i> -DMB*01	—	Ile	—	—	Pro	Ala	Arg	—
<i>Saoe</i> -DMB*02	—	Ile	—	—	Pro	Ala	Arg	—
<i>Saoe</i> -DMB*03	—	Ile	—	—	Pro	Ala	Arg	—

(*): stop codon; (—): identity

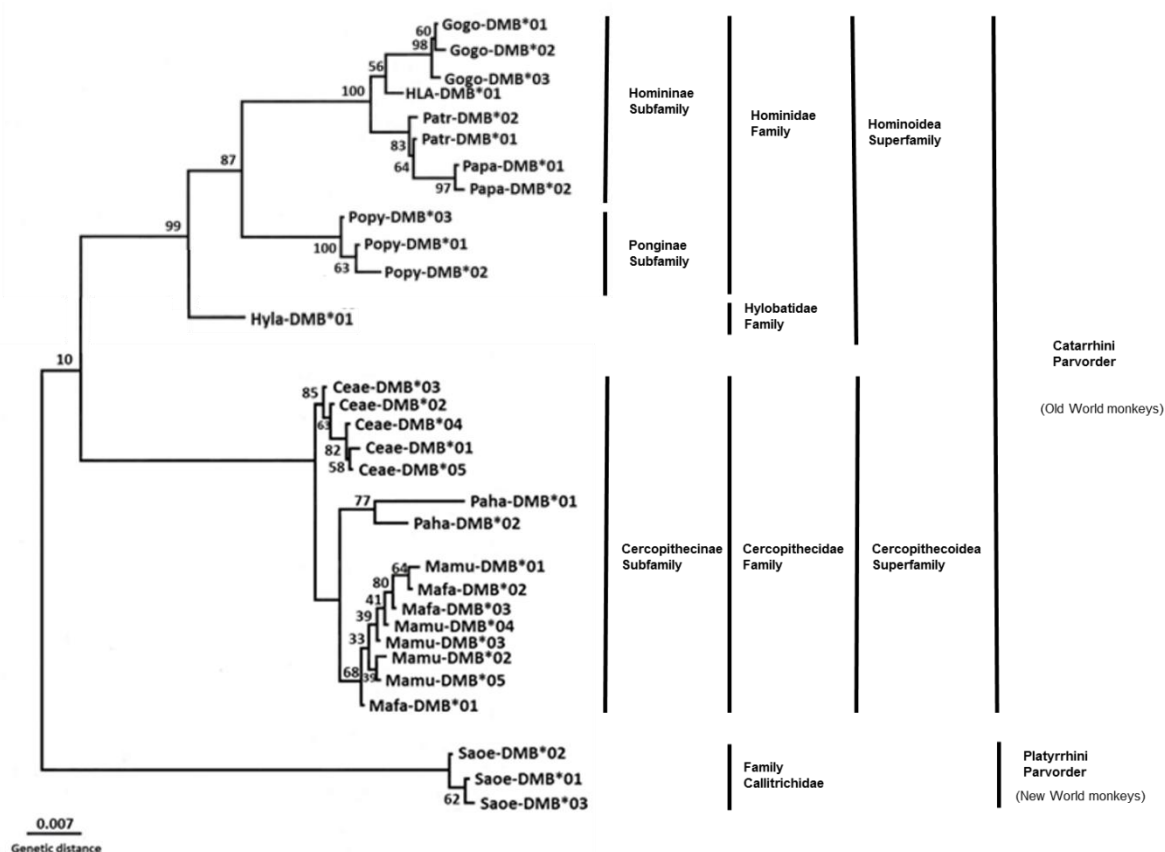


Figure 2. Phylogenetic tree of the whole cDNA sequences obtained. Neighbour-joining tree with 500 bootstrap methodology is used. Maximum-likelihood and Bayesian analyses showed the same tree topology and results. 19 MHC-DMB new sequences were obtained for this work: all *Hylobatidae*, *Cercopithecidae* and *Callitrichidae* families alleles.

It is striking that the one HLA-DMB allele sequenced for all exons (exon 1-6; Kelly et al., 1991) clusters together with *Gorilla gorilla* alleles (Alvarez et al., 1998); this is in contrast with other DNA comparisons that show a closer relationship between chimpanzees and humans (Goodman et al., 1990). However Bf, C4d and MICA exonic trees also clustered human and gorilla sequences (Paz-Artal et al., 1994; Meyer et al., 2014). It seems that MHC molecules that do not interact with the clonotypic T-cell receptor or NK-cell receptors (like DMB) do not support the postulated chimpanzee/human closer relationship, although this is not the case of MICA that acts as ligand of NKG2D receptor. This different clustering could be due to the fact that the divergence of the genes would have been previous to the divergence of the different species studied, although other possible theories, as convergent evolution, cannot be discarded. It is possible that phylogenetic relationships among chimpanzee, gorilla, orangutan, and humans could be more complex than postulated.

MHC-DMB variations occurring within each species suggest that this polymorphism may have undergone intraspecific evolution, since alleles of the same species cluster together (Fig. 2) as in other MHC related genes (Bf, C4d and MIC; Paz-Artal et al., 1994; Meyer et al., 2014); whereas for the rest of MHC class I and class II molecules, phylogenetic trees show a trans-species pattern of evolution (Corell et al., 1992; Slierendregt et al., 1992; Chen et al., 1992; Bergstrom and Gyllensten, 1995; Huchard et al., 2006). MHC-DMB genes may have also undergone an extensive homogenization within each ape species (Kawaguchi et al., 1992) or a recent generation of alleles may have occurred. This could be more graphically demonstrated with the description of identical MHC-E alleles in 2 different *Macaca* (Alvarez et al., 1997) and in two different *Pan* (Suarez et al., 1997) species.

Acknowledgments

This work was supported in part by grants from the Spanish Ministry of Health-Economy (PI11/000807 and PI14/01067), European FEDER funds and three different Mutua Madrileña Automovilista grants

References

- Alvarez, M., Martinez-Laso, J., Varela, P., Diaz-Campos, N., Gomez-Casado, E., Vargas-Alarcón, G., Garcia-Torre, C., Arnaiz-Villena, A. (1997). High polymorphism of Mhc-E locus in non-human primates: alleles with identical exon 2 and 3 are found in two different species. *Tissue Antigens* 49 (2): 160-167.
- Alvarez, M., Recio, M.J., Martinez-Laso, J., Perez-Blas, M., Garcia-de-la-Torre, C., Vargas-Alarcón, G., Alegre, R., Gomez-Casado, E., Arnaiz-Villena, A. (1998). Allelic diversity at the primate MHC-DMB locus: presence of a conserved tyrosine inhibitory motif in the cytoplasmic tail. *Tissue Antigens* 51 (2): 174-182.

- Arnaiz-Villena, A., Timon, M., Corell, A., Perez-Aciego, P., Martin-Villa, J.M., Rigueiro, J.R. (1992). Primary immunodeficiency caused by mutations in the gene encoding the CD3-gamma subunit of the T-lymphocyte receptor. *N. Engl. J. Med.* 327 (8): 529-533.
- 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.
- Bjorkman P.J., Saper M.A., Samraoui B., Bennet W.S., Strominger J.L. and Wiley A.C. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329: 506-512, 1987a
- Bjorkman P.J., Saper M.A., Samraoui B., Bennet W.S., Strominger J.L. and Wiley A.C. The foreign antigen binding site and T cell recognition region by class I histocompatibility antigens. *Nature* 329: 512-518, 1987b
- Chen, Z.W., McAdam, S.N., Hughes, A.L., Dogon, A.L., Letvin, N.L., Watkins, D.I. (1992). Molecular cloning of orangutan and gibbon MHC class I cDNA. The HLA-A and -B loci diverged over 30 million years ago. *J. Immunol.* 148 (8): 2547-2554.
- Cho, S.G., Attaya, M., Monaco, J.J. (1991). New class II-like genes in the murine MHC. *Nature* 353 (6344): 573-576.
- Corell, A., Morales, P., Varela, P., Paz-Artal, E., Martin-Villa, J.M., Martinez-Laso, J., Arnaiz-Villena, A. (1992). Allelic diversity at the primate major histocompatibility complex DRB6 locus. *Immunogenetics* 36 (1): 33-38.
- Daeron, M., Jaeger, S., Du Pasquier, L., Vivier, E. (2008). Immunoreceptor tyrosine-based inhibition motifs: a quest in the past and future. *Immunol. Rev.* 224: 11-43.
- del Rio, M.L., Seebach, J.D., Fernández-Renedo, C., Rodriguez-Barbosa, J.I. (2013). ITIM-dependent negative signaling pathways for the control of cell-mediated xenogeneic immune responses. *Xenotransplantation.* 20 (6): 397-406.
- Denzin, L.K., Cresswell, P. (1995). HLA-DM induces CLIP dissociation from MHC class II alpha beta dimers and facilitates peptide loading. *Cell* 82 (1): 155-165.
- Goodman, M., Tagle, D.A., Fitch, D.H., Bailey, W., Czelusniak, J., Koop, B.F., Benson, P., Slightom, J.L. (1990). Primate evolution at the DNA level and a classification of hominoids. *J. Mol. Evol.* 30 (3): 260-266.
- Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., Khodiyar, V.K., Lush, M.J., Povey, S., Talbot, C.C., Jr., Wright, M.W., Wain, H.M., Trowsdale, J., Ziegler, A., Beck, S. (2004). Gene map of the extended human MHC. *Nat. Rev. Genet.* 5 (12): 889-899.

- Huchard, E., Cowlshaw, G., Raymond, M., Weill, M., Knapp, L.A. (2006). Molecular study of Mhc-DRB in wild chacma baboons reveals high variability and evidence for trans-species inheritance. *Immunogenetics* 58 (10): 805-816.
- Kawaguchi, H., Zaleska-Rutczynska, Z., Figueroa, F., O'hUigin, C., Klein, J. (1992). C4 genes of the chimpanzee, gorilla, and orang-utan: evidence for extensive homogenization. *Immunogenetics* 35 (1): 16-23.
- Kelly, A.P., Monaco, J.J., Cho, S.G., Trowsdale, J. (1991). A new human HLA class II-related locus, DM. *Nature* 353 (6344): 571-573.
- Kropshofer, H., Hämmerling, G.J., Vogt, A.B. (1997). How HLA-DM edits the MHC class II peptide repertoire: survival of the fittest? *Immunol. Today* 18 (2): 77-82.
- Marks, M.S., Roche, P.A., van Donselaar, E., Woodruff, L., Peters, P.J., Bonifacio, J.S. (1995). A lysosomal targeting signal in the cytoplasmic tail of the beta chain directs HLA-DM to MHC class II compartments. *J. Cell. Biol.* 131 (2): 351-369.
- Mellins, E.D., Stern, L.J. (2014). HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. *Curr. Opin. Immunol.* 26: 115-122.
- Meyer, A., Carapito, R., Ott, L., Radosavljevic, M., Georgel, P., Adams, E.J., Parham, P., Bontrop, R.E., Blancher, A., Bahram, S. (2014). High diversity of MIC genes in non-human primates. *Immunogenetics* 66 (9-10): 581-587.
- Modrek B., Resch A., Grasso C., Lee C. 2001 Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res.* 29: 2850-2859.
- Nei, M., Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3 (5): 418-426.
- Nei, M., Jin, L. (1989). Variances of the average numbers of nucleotide substitutions within and between populations. *Mol. Biol. Evol.* 6 (3): 290-300.
- Paz-Artal, E., Corell, A., Alvarez, M., Varela, P., Allende, L., Madrono, A., Rosal, M., Arnaiz-Villena, A. (1994). C4 gene polymorphism in primates: evolution, generation, and Chido and Rodgers antigenicity. *Immunogenetics* 40 (6): 381-396.
- Pierre, P., Denzin, L.K., Hammond, C., Drake, J.R., Amigorena, S., Cresswell, P., Mellman, I. (1996). HLA-DM is localized to conventional and unconventional MHC class II-containing endocytic compartments. *Immunity.* 4 (3): 229-239.
- Pos, W., Sethi, D.K., Wucherpfennig, K.W. (2013). Mechanisms of peptide repertoire selection by HLA-DM. *Trends Immunol.* 34 (10): 495-501.

- Robinson, J., Halliwell, J.A., McWilliam, H., Lopez, R., Parham, P., Marsh, S.G. (2013). The IMGT/HLA database. *Nucleic Acids Res.* 41 (Database issue): D1222-D1227.
- Saitou, N., Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4 (4): 406-425.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
- Sanderson F., Kleijmeer M.J., Kelly A., Verwoerd D., Tulp A., Neefjes J.J., Geuze, H-J., Trowsdale, J. (1994a) Accumulation of HLA-DM, a regulator of antigen presentation, in MHC class II compartments. *Science* 266: 1566-1569.
- Sanderson, F., Powis, S.H., Kelly, A.P., Trowsdale, J. (1994b). Limited polymorphism in HLA-DM does not involve the peptide binding groove. *Immunogenetics* 39: 56-58.
- Sanderson, F., Trowsdale, J. (1995). Antigen presentation: Kissing cousins exchange CLIP. *Curr. Biol.* 5 (12): 1372-1376.
- Schulze, M.S., Anders, A.K., Sethi, D.K., Call, M.J. (2013). Disruption of hydrogen bonds between major histocompatibility complex class II and the peptide N-terminus is not sufficient to form a human leukocyte antigen-DM receptive state of major histocompatibility complex class II. *PLoS. One.* 8 (7): e69228.
- Sherman, M.A., Weber, D.A., Jensen, P.E. (1995). DM enhances peptide binding to class II MHC by release of invariant chain-derived peptide. *Immunity.* 3 (2): 197-205.
- Slierendregt, B.L., van Noort, J.T., Bakas, R.M., Otting, N., Jonker, M., Bontrop, R.E. (1992). Evolutionary stability of transspecies major histocompatibility complex class II DRB lineages in humans and rhesus monkeys. *Hum. Immunol.* 35 (1): 29-39.
- Sloan, V.S., Cameron, P., Porter, G., Gammon, M., Amaya, M., Mellins, E., Zaller, D.M. (1995). Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* 375 (6534): 802-806.
- Suárez, B., Morales, P., Castro, M.J., Fernandez-Soria, V., Recio, M.J., Perez-Blas, M., Alvarez, M., Diaz-Campos, N., Arnaiz-Villena, A. (1997). Mhc-E polymorphism in Pongidae primates: the same allele is found in two different species. *Tissue Antigens* 50 (6): 695-698.

