Immunogenetics (2005) 56: 818–832
DOI 10.1007/s00251-004-0751-5

ORIGINAL PAPER

Paula P. Cardenas · Carlos F. Suarez · Pilar Martinez · Manuel E. Patarroyo · Manuel A. Patarroyo

MHC class I genes in the owl monkey: mosaic organisation, convergence and loci diversity

Received: 29 July 2004 / Revised: 9 November 2004 / Accepted: 9 November 2004 / Published online: 15 January 2005

Abstract The MHC class I molecule plays an important role in immune response, pathogen recognition and response against vaccines and self- versus non-self-recognition. Studying MHC class I characteristics thus became a priority when dealing with *Aotus* to ensure its use as an animal model for biomedical research. Isolation, cloning and sequencing of exons 1–8 from 27 MHC class I alleles obtained from 13 individuals classified as belonging to three owl monkey species (*A. nancymaae*, *A. nigriceps* and *A. vociferans*) were carried out to establish similarities between *Aotus* MHC class I genes and those expressed by other New and Old World primates. Six *Aotus* MHC class I sequence groups (*Ao-g1*, *Ao-g2*, *Ao-g3*, *Ao-g4*, *Ao-g5* and *Ao-g6*) weakly related to non-classical Catarrhini MHC were identified. An allelic lineage was also identified in one *A. nancymaae* and two *A. vociferans* monkeys, exhibiting a high degree of conservation, negative selection along the molecule and premature termination of the open reading frame at exon 5 (*Ao-g5*). These sequences' high conservation suggests that they more likely correspond to a soluble form of *Aotus* MHC class I molecules than to a new group of processed pseudogenes. Another group, named *Ao-g6*, exhibited a strong relationship with Catarrhini's classical *MHC-B-C* loci. Sequence evolution and variability analysis indicated that *Aotus* MHC class I molecules experience inter-locus gene conversion phenomena, contributing towards their high variability.

Keywords *Aotus* · MHC class I molecules · Gene conversion · Evolution

P. P. Cardenas · P. Martinez · M. E. Patarroyo · M. A. Patarroyo (✉)
Molecular Biology Department,
Fundacion Instituto de Inmunologia de Colombia,
Bogota, Colombia
e-mail: mapatarr@fidic.org.co
Tel.: +57-1-3244671
Fax: +57-1-3244672

C. F. Suarez
Biomathemathics Department,
Fundacion Instituto de Inmunologia de Colombia,
Bogota, Colombia

M. E. Patarroyo · M. A. Patarroyo
Universidad Nacional de Colombia,
Bogota, Colombia

Introduction

MHC class I genes encode surface glycoproteins, playing an important role in the immune response to endogenous pathogens. Peptides derived from these pathogens are bound to these molecules and presented to $CD8^+$ cytotoxic T lymphocytes, triggering the destruction of the infected target cell and thus preventing infection of the host (Kundig et al. 1996).

Classical MHC class I (or class Ia) *HLA-A*, *HLA-B* and *HLA-C* loci products are expressed on the surface of almost all cell types in humans. These three class Ia genes are characterised by a high degree of allelic variation, mainly in the second and third exons encoding the MHC class Ia molecule's $\alpha 1$ and $\alpha 2$ domains (respectively), constituting the peptide-binding region (PBR). Variations are maintained at this point by positive selection pressure, as observed by the number of non-synonymous substitution patterns, resulting in amino acid replacements when compared with non-PBR sites in the same molecule (Hughes and Nei 1988).

In addition to MHC class Ia molecules, there are non-classical MHC class I loci encoding class Ib molecules, named *HLA-E*, *HLA-F* and *HLA-G* in humans. The products of these loci show very limited polymorphism and restricted tissue expression. Both HLA-E and HLA-G bind peptides and are involved in natural killer (NK) cell recognition; however, the role of HLA-F still remains unclear, although the conservation of these molecules suggests that they have been conserved to perform an important specialised biological function (O'Callaghan and Bell 1998).

Despite the wide diversity of MHC class I molecules, some human *HLA-A*, *HLA-B*, *HLA-E*, *HLA-F* and *HLA-G* loci orthologues have been preserved throughout the Catarrhini infra-order [Old World monkeys (OWM), apes and humans] for at least 35 million years by duplication and alternative gene divergence mechanisms (Cadavid and Watkins 1997; Knapp et al. 1998).

Previous New World primate MHC class I gene studies have reported human non-classical *HLA-G* and *HLA-E* loci orthologues (Adams and Parham 2001; Cadavid et al. 1997, 1999; Knapp et al. 1998).

A great number of these analyses have been focused on the cotton-top tamarin *Saguinus oedipus*, because it constitutes an exception to the diversity patterns ruling other primate MHC class I loci. Their MHC class I genes show low polymorphism and limited inter-and intra-locus variability, making it apparently susceptible to lethal infection by a wide variety of different viruses which are not fatal in humans (i.e. the Epstein-Barr virus), although this topic remains to be formally proved (Watkins et al. 1988, 1990, 1991).

Our institute has used *Aotus* monkeys as a model for human malaria vaccine development (i.e. S*Pf*66, Patarroyo et al. 1987), and an analysis of their immune system had become necessary for extrapolating the results obtained in the animal model to humans. Several *Aotus* immune molecules have been characterised in our institute, such as TCRα (Favre et al. 1998), TCRβ (Vecino et al. 1999), TCRγδ (Daubenberger et al. 2001), MHC class II DR (Nino-Vasquez et al. 2000), DQ (Diaz et al. 2000a) and immunoglobulin (Diaz et al. 2000b). Since immunity to malarial hepatic stage infection is mediated by MHC class I molecules, a detailed study of these molecules in *Aotus* is necessary to gain a deeper knowledge about peptide presentation in this context during vaccine development.

A previous study has analysed the PBR of these three *Aotus* monkey species (Suarez et al. 2003), finding a level of variability in *Aotus* non-classical molecules similar to that exhibited by classical Catarrhini molecules. However, no clear phylogenetic relationships could be established amongst the species analysed because of the different types of selective pressures acting on the PBR. The present work thus describes the complete *Aotus* monkey MHC class I molecule study, looking for regions such as exons 4–8, which are not under positive selection pressure and are therefore suitable for assigning loci to the different sequences.

The cDNA from three different owl monkey species (*A. nancymaae*, *A. vociferans* and *A. nigriceps*) was analysed and searched for similarity relationships between *Aotus* and other primates. An evolutionary approach was used for delimiting the main groups in the *Aotus* species considered and analysing the origin of variability. Such analysis included all the exons encoding the MHC class I molecule.

Materials and methods

Animals

Whole blood was obtained by venipuncture from 13 owl monkeys classified as belonging to three different species: *A. vociferans* (four monkeys), *A. nancymaae* (six monkeys) and *A. nigriceps* (three monkeys).

These monkeys were classified based on both karyological and morphological characteristics (Torres et al. 1998). Peripheral blood lymphocytes were separated from whole blood, using Ficoll gradient.

RNA extraction, cDNA synthesis, cloning and sequencing

Total cellular RNA was extracted from 5×10^6 lymphocytes, using TRIzol reagent (Gibco-BRL, Rockvillle, Md., USA). One microgram total RNA was used for synthesising cDNA. The cDNA synthesis reaction contained 50 mM Tris (pH 8.3), 5 mM $MgCl_2$, 1 mM dNTPs (Promega, Madison, Wis., USA). Then, 0.75 μM oligo-dT, 50 U Moloney Murine Leukaemia Virus reverse transcriptase (Promega) and 20 U RNAsin (Promega) were added to give a final 20-μl volume. cDNA was synthesised at 42°C for 1 h, 95°C for 5 min and 4°C for 5 min.

PCR used exon 1 (5′-CTC CTC CTG CTG CTC TCG GG-3′) and 3UTR (Watkins et al. 1991) primers at a final 0.8-μM concentration. The PCR mixture contained 1.5 μM $MgCl_2$, 50 mM Tris (pH 8.3) and 2.5 U *Taq* DNA polymerase (Promega). Five microlitres cDNA were added to each reaction, for a 25-μl final volume. The reactions were heated to 95°C for 5 min and then amplified for 40 cycles as follows: denaturing for 30 s at 94°C, annealing for 1 min at 60°C and extension for 2 min at 68°C. A final extension cycle was run at 60°C for 1 min and 68°C for 5 min.

A WIZARD PCR Preps Purification kit (Promega) was used for purifying PCR products which were then ligated into pGEM T vector (Promega). MiniPreps Purification Kit (Mo Bio, Carlsbad, Calif., USA) was used for isolating double-strand plasmid DNA.

On average, three clones from each cDNA were chosen and sequenced using fluorescent dye-labelled dideoxy terminators (Applied Biosystems, Foster City, Calif., USA) in an ABI Prism 310 Genetic Analyzer (Applied Biosystems). The Sequence Navigator program (PerkinElmer) was used for collecting sequence data. Internal primers called MHC class I Forward (5′-CGG CTA CTA CAA CCA GAC) and MHC class I Reverse (5′-TGT ATC TCT GCT CYT CTC CC-3′) were designed, as the whole molecule cannot be sequenced at once.

Aotus allele nomenclature

The alleles have been named according to rules designated for different species' MHC nomenclature (Klein et al. 1990). The first two letters of the genus and species names are

combined to produce the taxa designation; these are followed by a dash and then the letter of the locus type and the allele number.

Sequence and evolutionary analysis

Clustal X software (Thompson et al. 1994) was used for aligning representative MHC class I gene sequences from primates (*A*, *B*, *C*, *E*, *F*, *G* and *I* functional loci and also *H*, *AP*, *J*, *PS1*, *PS2* pseudogenic loci). The new *Aotus* sequences were included, and amino acid alignment was also performed. Only non-identical sequences were used in a later analysis. This alignment included the Catarrhini suborders [Hominoidea apes and Cercopithecoidea (OWM), Platyrrhini [New World monkeys (NWM)] and Strepsirhini (lemurs)]. The resulting alignment had a total of 1,074/358 nucleotide/amino acid positions.

Each position's variation for a given set of amino acid aligned sequences was represented by using the WebLogo software (http://weblogo.berkeley.edu//logo.cgi; Crooks et al. 2004; Schneider and Stephens 1990). Each amino acid is represented by one letter code in its respective position within the sequence. All the amino acids occupying each position are indicated, where the height of every amino acid letter represents its relative frequency at that position. The logo also allows conservative and non-conservative substitutions for each position to be determined where variation in the amino acid symbol's colour indicates non-conservative changes and maintaining it represents conservative changes based on PAM 250 substitution matrix parameters (Dayhoff et al. 1978).

Genetic trees were constructed by MEGA, version 2.1, software (http://www.megasoftware.net; Kumar et al. 2001), using the minimum evolution (ME) distance method (Rzhetsky and Nei 1993). Genetic distances were estimated by using the P, Kimura two-parameter and Jukes-Cantor (Kimura 1980) distances for nucleotide sequences and just P distance for the amino acid sequences. Bootstrap analysis (Hillis 1993) and the internal branch test were used (1,000 replicates for each) for assigning confidence levels to tree-branch nodes. Synonymous and non-synonymous substitutions and associated variance (assessed by the bootstrap method, using 1,000 replicates) rates were calculated according to Nei and Gojobori's method (Nei and Gojobori 1986), using MEGA.

Mosaic sequence structure was detected by two methods; the first used detection of changes in local estimated phylogenies through a 20-bp slide window clustering all along the gene's total length (ME with P distance). Topology results were interpreted by inspecting gene alignment. The second methodology followed the method proposed by Sawyer (Sawyer 1989), implemented in the GENECONV program (Sawyer 1999). This program performs statistical procedures to rank the possible gene conversion events in an alignment and provides multiple-comparison-corrected *P*-values for each instance.

MHC sequences' GenBank accession numbers

Aotus sequences have been registered in the GenBank under accession numbers AY659826 to AY659852. The GenBank accession numbers of the other sequences employed in this study are *Atbe-B01* (U59648), *Aotr-B1sw* (AB113204), *Aotr-B2sw* (AB113205), *Aotr-Fsw* (AB113203), *AotrG*01cd* (U59644), *Aotr-G1sw* (AB113092), *AotrG*02cd* (U59645), *Aotr-G2sw* (AB113202), *Aotr-PS2* (U52114), *Caja-G*04* (U59640), *Caja-G*02* (U59638), *Caja-G*01* (U59637), *Caja-G*05* (U59641), *Caja-G*03* (U59639), *Caja-PS2* (U52115), *Gogo-A0101* (X60258), *Gogo-B*0301* (AF157406), *Gogo-C*0103* (AF157410), *Gogo-H*01* (AF157392), *HLA-A*3301* (U83416), *HLA-B15* (L32862), *HLA-Cw*0103* (D64145), *HLA-E*01031* (L78934), *HLA-Fnn1* (XM_041595), *HLA-G*01041* (AF436089), *HLA-H* (NG_002398), *HLA-J* (NG_002398), *HLA-K* (NG_002398), *HLA-L* (NG_002771), *Hyla-A*01* (U50089), *Hyla-B*01* (U50091), *Lero-G*02* (U59643), *Lero-G*01* (U59642), *Mamu-A* (AB113108), *Mamu-B*01* (AJ556874), *Mamu-E*05* (U41837), *Mamu-I*01011* (AF161865), *Mane-A*09* (AY204730), *Mane-B*01*(AY204731), *Mane-E*0501* (AY204722), *Mane-I*0101* (AY204739), *Mimu-W01* (AJ297588), *Mimu-W02* (AJ297589), *Mimu-W03* (AJ297590), *Mimu-W04* (AJ302085), *Pacy-A*01* (AF288698), *Pacy-B*01* (AF288702), *Patr-A*0601* (AF500290), *Patr-B*0302* (AF500291), *Patr-C*1001* (AF179673), *Patr-E* (AF338354), *Patr-F1* (M30685), *Patr-H*01* (AF157393), *Patr-J* (AF338358), *Pipi-G*01* (U59651), *Pipi-G*02* (U59652), *Pipi-G*03* (U59653), *Pipi-G*04* (U59655), *Pipi-G*05* (U59656), *Pipi-B01* (U59654), *Pipi-G*04* (U59655), *Popy-A*0402* (AY034117), *Popy B*0401* (AF118892), *Popy-C*0301* (AF470379), *Safu-PS1* (U52117), *Safu-PS2* (U52116), *Sage-PS1* (U52118), *Sage-PS2* (U52119), *Samy-PS2* (U52121), *Saoe-B* (M33476), *Saoe-E*01* (AF004918), *Saoe-G*12* (AF020686), *Sasc-25* (AJ438577), *Sasc-31* (AJ438576), *Sasc-G*01* (U59657), *Sasc-G*02* (U59658), *Sasc-G*03* (AY282760), *Sasc-G*04* (AY28276) and *Sasc-G*05* (AY282762).

Results

Cloning *Aotus* MHC class I alleles

The MHC class I molecule (from exon 1 to exon 8 in *Aotus*) was amplified by PCR, using primers designed on the reported sequences available in the GenBank database (exon 1) or that had been previously reported (Watkins et al. 1991).

The amplification product had a size of around 1,074 bp; however, the size of the different amplification products varied depending on whether the sequences had insertions or deletions. The open reading frame (ORF) encoded a protein having around 355 amino acids, including the leader peptide.

A total of 27 *Aotus* MHC class I alleles were identified by employing this method; 18 were isolated from *A. nancymaae*, six from *A. vociferans* and three from *A. nigriceps*. The number of different sequences identified in one in-

dividual varied from one (*A. nigriceps–A. vociferans*) to four (*A. nancymaae*). The sequences obtained from each monkey are shown in Table 1.

All *Aotus* alleles encoded proteins having typical MHC class I molecule characteristics. These included a leader peptide (only 16 out of the 24 encoded amino acids were obtained for *Aotus* due to exon 1 primer annealing), two highly variable 90-amino acid length domains [α_1 and α_2, which might be participating in peptide and T-cell receptor (TCR) binding as occurs in humans], a less variable 90-amino acid length domain (α_3 possibly involved in interactions with CD8 and β_2-microglobulin), a 39-amino acid trans-membrane region and a 25- to 26-amino acid cytoplasmic domain. An *N*-linked glycosylation site was conserved near the α_1 domain carboxy-terminus (Asn^{86}) as well as two cysteines, two in the α_2 domain (Cys^{101} and Cys^{104}) and two in the α_3 domain (Cys^{203} and Cys^{259}) forming α_2 and α_3 domain disulfide bonds (Lafont et al. 2003).

Six *Aotus* MHC class I sequence groups can be distinguished

Evolutionary trees for each exon and several of them in combination were constructed by the ME method, using P distance, Jukes Cantor and Kimura two-parameter for nucleotides and P distance for amino acid sequences to see whether the *Aotus* MHC class I sequence clustering pattern varied depending on the exon being analysed (not shown).

These results show that previously described *Aotus* MHC class I g-like sequences (Suarez et al. 2003) can be subdivided into *Ao-g1*, *Ao-g2*, *Ao-g3*, *Ao-g4*, *Ao-g5* and *Ao-g6* sequence groups, independently of whether amino acid or nucleotide sequences were being employed. These groups can be delimited for each exon in most cases (depending on the degree of conservation exhibited by each one of them) and can be recovered in the supposedly most reliable combination for determining evolutionary relationships in MHC class I exons 1 and 4–7 (the non-PBR exons) (Cadavid et al. 1997; Fig. 1b). They can also be retrieved from analysing exons 2 and 3 (Fig. 1a). Several sequences displayed sequence instability, depending on the exon being considered (implying that sequences showed a different clustering pattern for each exon considered), which might have been caused by a recombination process (not shown). The statistical support for each sequence group increased significantly when unstable sequences were removed from the tree for both topologies derived from exons 1 and 4–7 and that derived from exons 2 and 3 (data not shown). The relationships were also well supported for complete sequences (Fig. 3).

Comparing *Aotus* exons 2 and 3 sequences described in this work, and those previously reported (Cadavid et al. 1997; Sawai et al. 2004; Suarez et al. 2003), showed that *Ao-g2* (one of the two MHC groups primarily described, the other being *Ao-g1*) was subdivided into new groups, while *Ao-g1* displayed stable behaviour (Fig. 1a). A trans-specific evolutionary pattern occurred in all putative loci in all cases, but no clear bias occurred in the studied species (since the differential sampling between *Aotus* species) for a determined loci. *Ao-g2* displayed four more related groups, *Ao-g3*, *Ao-g4*, *Ao-g5* and *Ao-g6*, within the context of complete sequences, being well supported in most cases. *Ao-g3* and *Ao-g6* showed close association at exons 2 and 3; *Ao-g3* showed the main concentration of unstable sequences, hampering its clear classification and forming a polyphyletic group. *Aog2b–Aog2d* groups of sequences were found in the first study (Suarez et al. 2003) but not in this one and their status (such as *Ao-g2* subgroups or new groups) requires more research for clarifying this issue (Fig. 1a).

The sequences constituting *Ao-g2* for exons 1 and 4–7 (Fig. 1b) belonged to *Ao-g2a* for exons 2 and 3 (Fig. 1a). *Ao-g1* remained as a stable group, and the differences regarding exons 2 and 3 corresponded to the instable (mosaic) *Aovo-g10301* sequence (see next section). *Ao-g5* displayed a higher internal divergence caused by *Aona-g50101* mosaic sequence. Although *Ao-g6* and *Ao-g3* showed significant clustering using exons 1 and 4–7, *Ao-g6* sequences completely lacked exon 6, displaying clear sequence differentiation in exons 1 and 4–7, sustaining their categorisation within an independent sequence group. The loss of that exon does not necessarily imply functional inactivation, since previous reports have shown that signal

Table 1 Sequences identified in each one of the analysed monkeys. Numbers correspond to the amount of sequenced clones

		A. nancymaae					*A. vociferans*				*A. nigriceps*		
		1	2	3	4	5	6	1	2	3	4	1	2	3
Sequence	Aoni-g10102													3
	Aoni-g10103												3	
	Aovo-g10301								2					
	Aona-g10302						2							
	Aona-g10303	2												
	Aona-g10304			3										
	Aoni-g20104											2		
	Aona-g20105						3							
	Aona-g20106					4								
	Aona-g20107		3											
	Aona-g20108		2											
	Aovo-g20108									3				
	Aona-g20109					2								
	Aona-g20110				2									
	Aona-g20111				3									
	Aona-g20113			2										
	Aona-g20114			2										
	Aona-g20115	4												
	Aona-g30102				2									
	Aona-g30201		3											
	Aona-g30202					2								
	Aovo-g40101									2				
	Aona-g50101		3											
	Aovo-g50101										2			
	Aovo-g50102							2			3			
	Aona-g60101		3											
	Aovo-g60101							2		3				

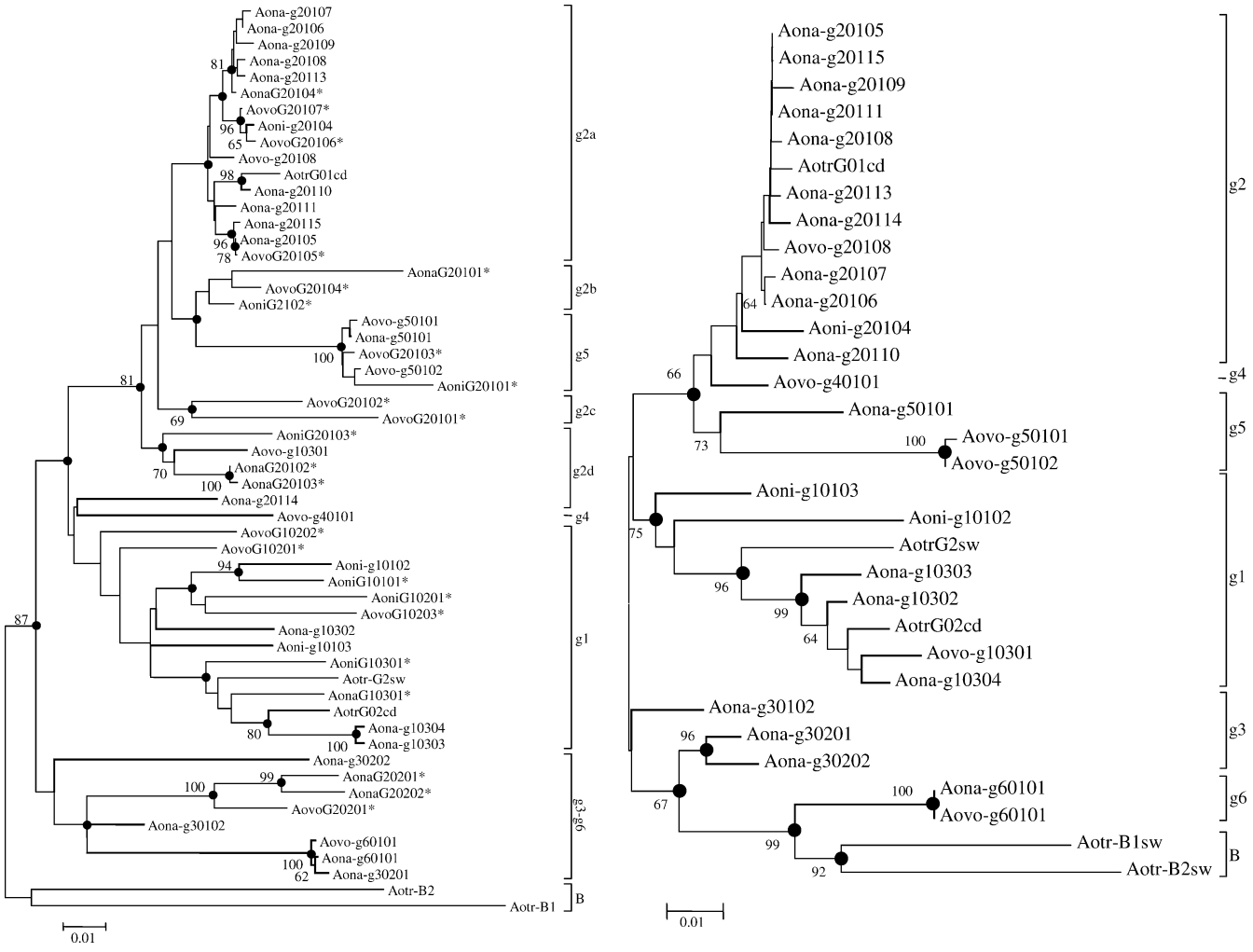

Fig. 1 *Aotus* MHC class I gene genetic trees. Trees were constructed by the minimum evolution (ME) method, P distance, with percentage recovery at each node in 1,000 bootstrap replications. **a** Exons 2 and 3 nucleotide sequences, where (*) indicates those sequences obtained in previous work (Suarez et al. 2003) and **b** exons 1 and 4–7 nucleotide sequences. Recombinant sequences are indicated by *boldface* and *thicker lines*, the branches. Only those branches having ≥60% bootstrap values are indicated. Branches having significant internal branch test (IBT) values (≥90%) are indicated by (●), *sw* indicates Sawai et al. (2004) sequences used, and *cd* indicates Cadavid et al. (1997) sequences used

transduction may occur independently of the presence of cytoplasmic domains which can be mediated by the MHC class I molecule extra-cellular domain (Gur et al. 1997, 1999). *Ao-g3* sequences constitute a polyphyletic group because of their sequence instability, as in exons 2 and 3.

The *Aovo-g10301* sequence is characterised by a 12-bp insertion between nucleotides 366–379 in exon 3 derived from the TTGGGGCCGGAC duplicate sequence (data not shown). This does not appear to alter MHC class I molecule structure, because this insertion lies in a loop zone within the molecule (Parham et al. 1988). Previous reports have reinforced the idea of this zone being duplication prone, because the same event has been identified in one pig-tailed macaque's sequence (*Mane-B*06*) possessing an residue insertion in this region (Lafont et al. 2003).

The range and number of pairwise nucleotide differences were calculated within each of the three human HLA classical loci and compared with *Aotus* sequence groups. The maximum values for alleles belonging to *HLA-A*, *HLA-B* and *HLA-C* loci were 73, 85 and 72, respectively. The maximum values in *Aotus* 'classical' sequence groups were found to be 84 and 73 for *Ao-g1* and *Ao-g2*, respectively. This analysis could not be made for *Ao-g4*, because only one sequence had been identified; however, based on the clustering pattern and the presence of sequence motifs distinguishing it from other sequence groups, it was decided to classify it as a new sequence group. *Ao-g5* was characterised by exhibiting a behaviour resembling non-classical loci, displaying less variation than the classical groups (pairwise differences ranged from 3 to 34). *Ao-g3* displayed medium variability; however, its high degree of sequence instability hampered similar evaluation. *Ao-g6* displayed a high degree of conservation, similar to the *Ao-g5* group.

The extent of internal divergence in human loci was very similar to that found in *Aotus* sequence groups,

suggesting that these sequence groups might reflect a loci division in *Aotus* MHC class I.

Aotus MHC class I in the primate context

Aotus displayed different clustering patterns within the primate MHC I context, depending on the molecular sector chosen for evolutionary reconstruction. The topologies constructed by employing exons 2 and 3 showed that *Aotus* (excluding recombinant sequences) and other NWM constitute a separate branch from OWM, apes and lemurs. Each NWM genera almost clustered in a genus-specific way: two major groups could be observed, the first containing *Ao-g1* sequences and the second including *Ao-g2*, *g3*, *g4*, *g5* and *g6* sequences (named *g2*-like group). These NWM MHC class I sequences were distantly related to non-classical loci *MHC-G–MHC-F*. Some *B-like* sequences described by Cadavid et al. (1997) and Sawai et al. (2004) cluster independently from other NWM sequences. These were more closely related to *MHC-E* locus (*Aotr-B2*), *MHC-F* locus (*Aotr-B1*) and *MHC-A* locus (*AtbeB01* and the partial *Aotr-B3* exon 2 sequence available).

The topology constructed using complete sequences resembles the evolutionary pattern described for these exons due to great exons 2 and 3 variability (data not shown). Statistical support was found for Catarrhini loci in both analyses, contrasting with NWM branches that did not show statistical support in most cases, like the deep branch nodes (indicating inter-loci relationships). This reflects the great uncertainty of exons 2 and 3 sequences in adequately resolving a MHC evolutionary pattern, due to the strong evolutionary pressure and recombinant phenomena described previously (Cadavid et al. 1996, 1997; Cadavid and Watkins 1997).

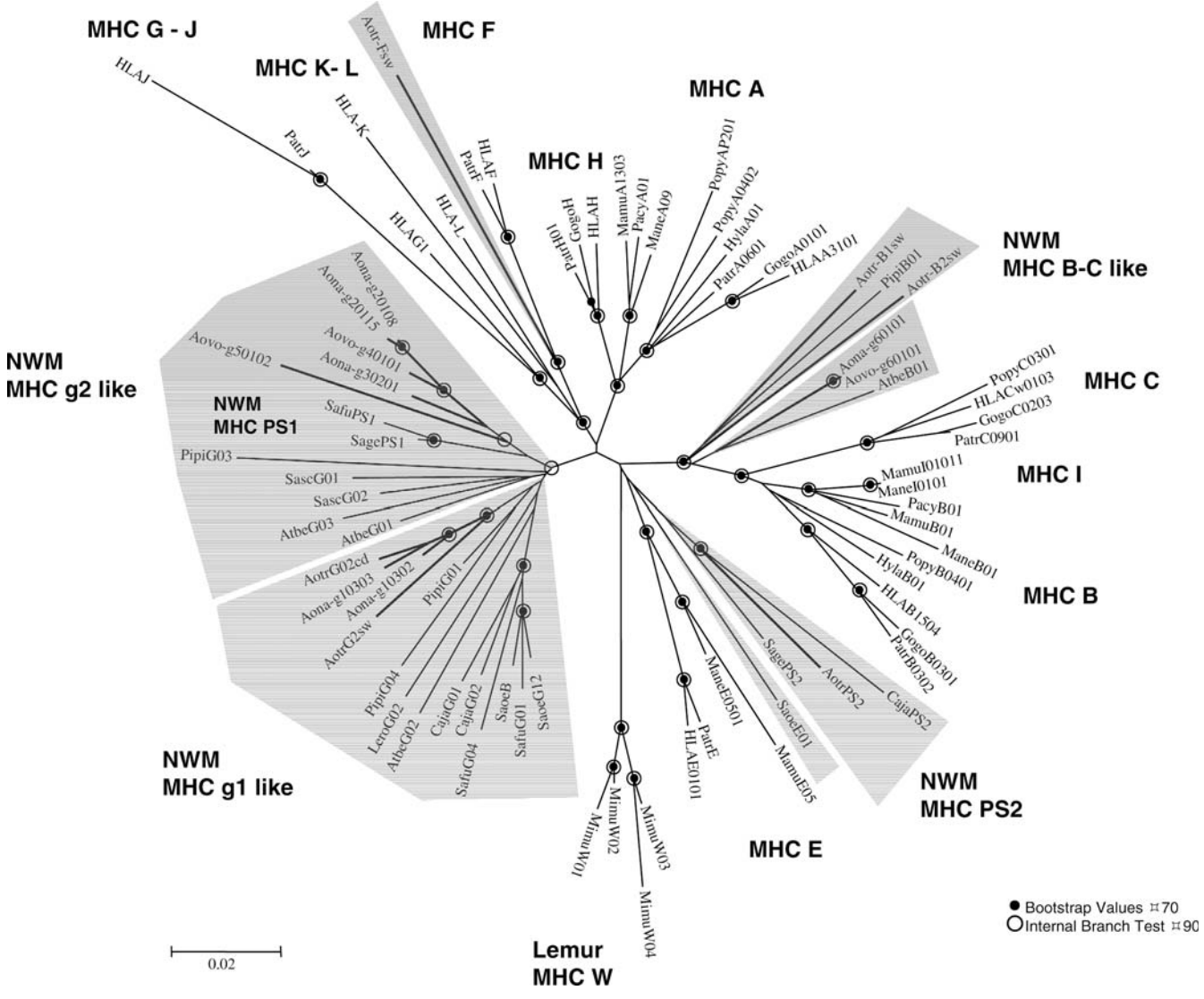

Fig. 2 Primate MHC class I gene ME genetic tree. The tree was constructed based on P-distance matrix, excluding all unstable *Aotus* sequences. Node significance was calculated by the bootstrap test and IBT

Calculations were thus made using exons 1 and 4–7 to determine a more accurate evolutionary reconstruction. The actual data indicate that NWM possess seven major MHC class I divisions (Fig. 2). Six *Aotus* sequence groups can be distinguished within primate MHC class I. *g1-like* and *g2-like* clustered together weakly related to *MHC-G*, *MHC-F* and *MHC-A*; *B-like* group clustered together with Catarrhini classical *MHC-B-C*; *PS2*-processed pseudogene clustered with *MHC-E* locus and *Ao-F* clustered with other *MHC-F* sequences (Fig. 2). *Aotus MHC-E* has been previously observed and partially characterised (Knapp et al. 1998; Sawai et al. 2004). No *PS1* NWM pseudogene had been detected in *Aotus* until now. The overall clustering pattern among primate MHC class I loci is in agreement with a previous analysis by Sawai et al. (2004).

The first '*g1*-like' group (containing *Ao-g1*, all Callitrichidae genera studied and other Cebidae genera distinct to *Aotus* such as *Ateles* and *Pithecia*) appeared to be related to the '*g2*-like' group (containing only Cebidae such as *Ao-g2* and derived sequences, as well as *Saimiri* and *Ateles*). Statistical support clearly became improved with these exons, showing significant values for *Aotus* and Callitrichidae nodes. A genus-specific pattern was less evident in this topology. When analysing exons 1 and 4–7 only, several NWM sequences, such as the *Ao-g6* and *Aotr-B1-B2* sequences (Sawai et al. 2004) and *Ateles* and *Pithecia* '*B*' sequences (reported by Cadavid et al. 1997), were significantly associated with Catarrhini *MHC-B-C* conforming an 'NWM *B-C*-like group'. This pattern suggests either a convergent pressure or ancient recombination processes at exons 2 and 3 for these *B-C*-like sequences driving their PBR towards an NWM g-like PBR in the case of *PipiB01* and *Ao-g6*, or towards a *MHC-E*, *MHC-F* and *MHC-A* for *Aotr-B02*, *Aotr-B01* and *Aotr-B03–AtbeB01*, respectively. This pattern is very similar to the topology deduced from amino acid sequences; the same groups were conformed but with minor statistical support. *Ao-g6* and *Atbe-B01* sequences were included within Catarrhini *MHC-B-C*, being more closely related to *MHC-C* sequences than *MHC-B* sequences. *Pipi-B01*, *Aotr-B01* and *Aotr-B02* clustered as a sister group of the former.

Aotus MHC class I molecule mosaic organisation

A mosaic sequence organisation for *Aotus* MHC class I molecules was observed when comparing the different trees for each exon. Unstable sequences are shown by bold branches in Fig. 1. Seventeen out of 32 sequences showed a mosaic constitution (Fig. 3). In spite of this behaviour, it was possible to distinguish sequences having no evidence of that process and define them as diagnostic sequences for the sequence groups in *Aotus* MHC class I genes.

Evidence of known recombination motifs was only found within exon 5 in positions 928–937 (GCTGGCCTGG) and in positions 967–973 (GCTGTGG) resembling the λχ recombination (GCTGGTGG) (Crew et al. 1991). Some recombination events were found in these sites (Fig. 3), constituting a recombinational hot spot in other species as has been previously shown (Wheeler et al. 1990) when sequences similar to the χ site were found in the endpoints of recombinational events in mouse and human MHC class I genes. The presence of these chi sites apparently makes MHC class I genes more prone to experience gene conversion, duplication and deletion events (Crew et al. 1991).

Figure 3 shows the schematic delimitation of the interchanged sectors amongst different *Aotus* MHC class I loci. Gene conversion events were evaluated by GENECONV software (Sawyer 1999) and using topological analysis. Most recombinant tracts agree using two methods; some tracts detected by topological analysis were not identified by GENECONV and were denoted as pale colours in Fig. 3. It was noted that the sectors for segment exchange were not only located in exons 2 and 3 (where variability is located), but also in exons 1 and 4–7 (Fig. 3). These recombination events do not preferentially show a sector, as they occur almost equally in both variable and conserved regions.

Although we cannot exclude the possibility of chimera formation explaining some recombination events, several reports indicate that this occurs in a 5–8% of cases in similar conditions as those used by us (Ennis et al. 1990; Meyerhans et al. 1990; Qiu et al. 2001); so, only a very few proportion of the sequences analysed might be artifacts.

The finding of segment exchange in *Aotus* sequences led us to examine whether this interlocus recombination phenomenon was shared by different primate genera. A phylogenetic analysis of MHC class I complete sequences available in the GenBank database for apes, NWM and OWM has thus been carried out using the previously described approach.

This analysis showed that the *HLA-A*, *HLA-B* and *HLA-C* sequences and *Macaca mulata* classical genes exhibited evidence that recombination was limited to a given locus's sequences (intra-locus recombination) in virtually any sector of the molecule. The results obtained for NWM different to *Aotus* (*S. oedipus*, *C. jacchus*, *P. pithecia*) showed that recombination was a common phenomena extending throughout the whole molecule for these primate genera (data not shown). Although definitive loci assignation remains undetermined in these primates (Cadavid et al. 1997), these sequences' evolutionary pattern suggests interloci recombination events in NWM. This behaviour also occurs for Lemurs (Go et al. 2003) and mouse MHC class I (Gu and Nei 1999; Yun et al. 1997).

The unstable sequences coincided by possessing recombination segments (Fig. 3). All groups showed sequences having recombination tracts. All *Ao-g3* sequences presented recombinant sectors followed by *Ao-g1* and *Ao-g2*. *Ao-g2* constitutes the major donor of recombination sectors for the other putative loci; however, minor recombination events were shown with just *Ao-g1* as donor group. *Ao-g6* showed high similarity with *Ao-g3* in exons 2 and 3; however, it exhibited significant differences allowing its characterisation. *Aona-g30201* presented a recombination segment comprising exon 2, 3 and a substantial part of exon 4 which was derived from the *Ao-g6* group.

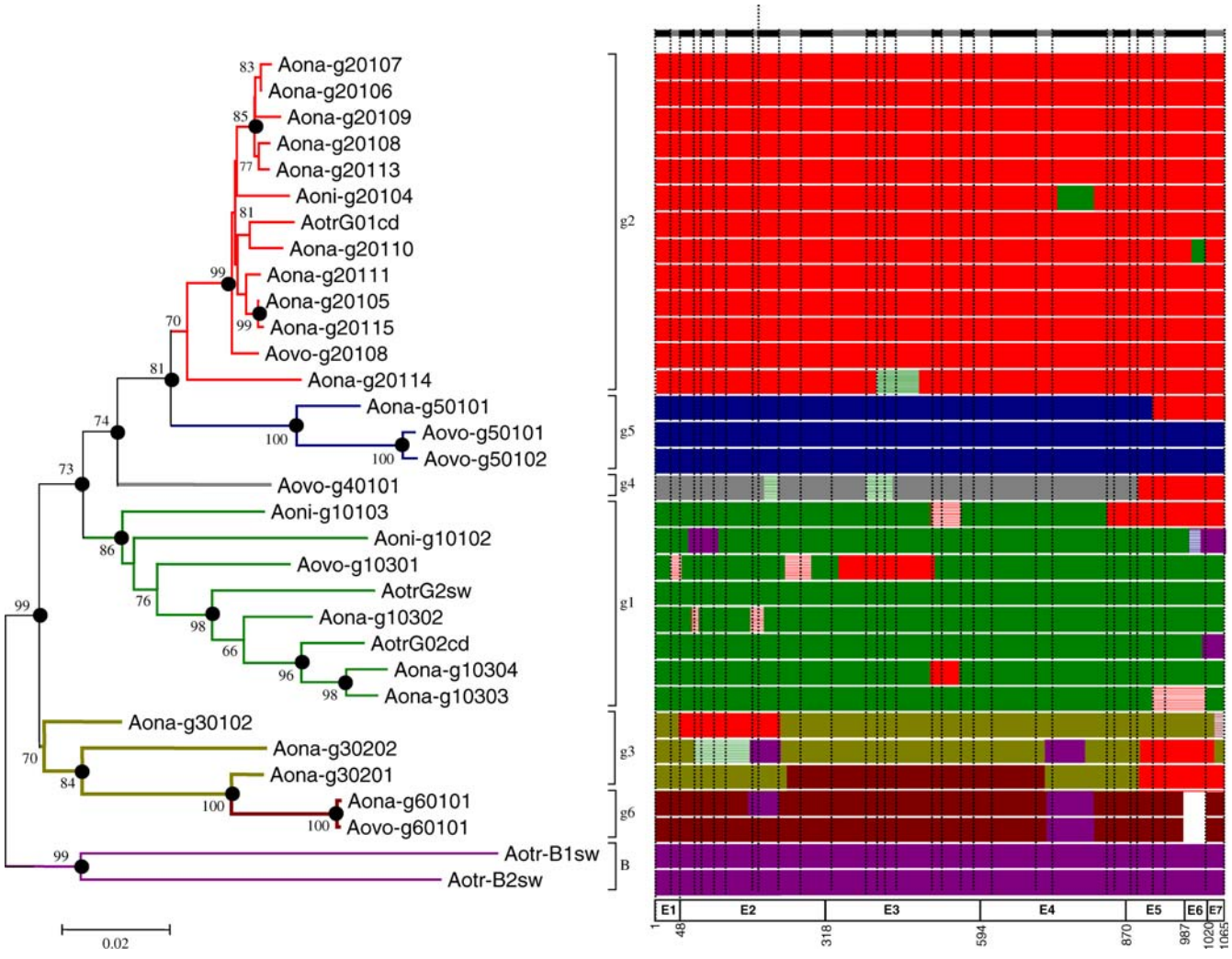

Fig. 3 Schematic representation of recombination tracts in *Aotus* MHC class I molecules. The *left-hand side* of the diagram shows a phylogenetic tree obtained by the ME method for exons 1–7 nucleotide sequences, where the main phylogenetic groups are represented by a *characteristic branch colour* for each one of them. Exon division for the whole MHC class I molecule is shown at the *bottom right-hand side* of the diagram. A representation of variable and conservative regions is shown at the *top right-hand side* of the diagram in a *dark grey bar* and *light grey bar* (respectively); prolongations from this bar pass down through the schemes *besides the names of the sequences*, indicating whether the recombination phenomenon occurs in a variable or conserved region. Recombination fragments are represented within the schematic sequence by the *branch colour* assigned to the locus donating the tract. In the scheme, *pale colours* indicate those recombination tracts only detected by the topological method

Ao-g5 presents a premature stop codon in exon 5

Ao-g5 sequences were characterised by possessing similar 8-bp AGCCCATCA insertions for *Aovo-g50101* and *Aovo-g50102* and AGCCTACCA for *Aona-g50101* in positions 895–903, causing a stop codon in exon 5. This fact implies that these molecules do not possess a trans-membrane domain, suggesting that they represent a soluble form, as observed in soluble HLA-G isoforms (Fujii et al. 1994), or constitute a pseudogene lineage (Fig. 4). Nevertheless, the high degree of conservation (maximum 2.2% divergence) at exons 2 and 3 (containing the PBR) and other exons, not only at intra-species level (as in *A. vociferans* sequences) but at inter-species level (*A. vociferans* versus *A. nancymaae*) (Fig. 1), contrasts with the higher mutation rate exhibited by other NWM pseudogenic lineages such as *PS2* (10.8% divergence at exons 2 and 3) and *PS1* (4.4% divergence at exons 2 and 3) (Cadavid et al. 1996). This might suggest a functional role for these molecules and their low variability at PBR level might imply a non-classical status for *Ao-g5*.

There was evidence that the *Ao-g5* lineage is present in the three *Aotus* species, because the previously obtained partial *Aoni-g20101* and *Aovo-g20103 Aotus* sequences (Suarez et al. 2003) grouped with the complete *Ao-g5* sequences in exons 2 and 3, presenting significant bootstrap values (Fig. 1a).

Aona-g50101 and *Aovo-g5* stop codons were located in different positions in exon 5. *Aona-g50101* exhibited a recombination tract derived from the *Ao-g2* group (Fig. 3) that produced a stop codon between nucleotides 946–948, 12 base pairs before the premature break of *Aovo-g5* (placed

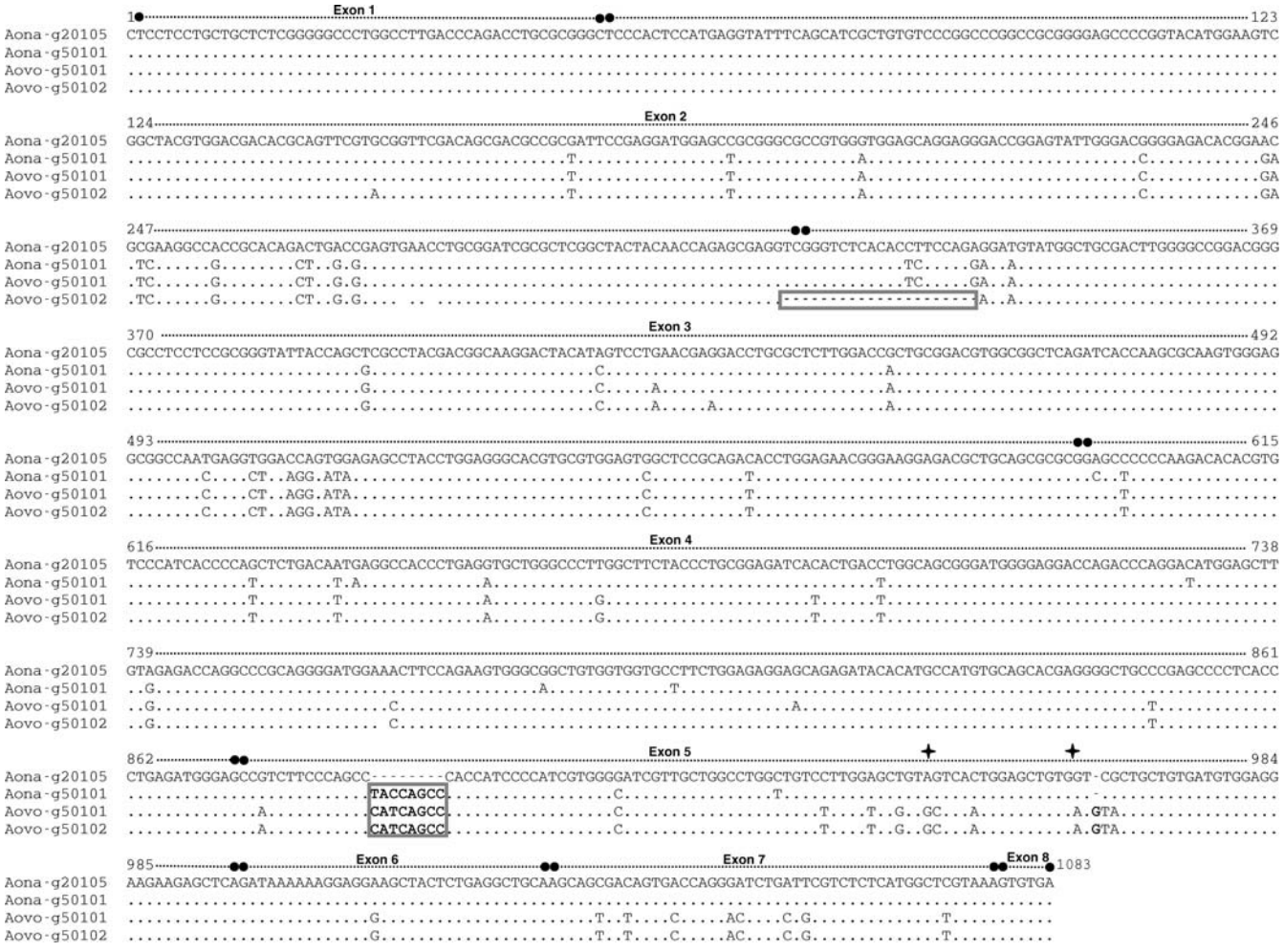

Fig. 4 *Ao-g5* sequence nucleotide alignment. The alignment shows the three MHC class I sequences identified in *A. nancymaae* and *A. vociferans*, which were classified as *Ao-g5*. Identity with *Aona-g20105* sequence is indicated by (●). Insertions and deletions are marked inside *boxes*. Open reading frame ruptures are indicated by (✦)

between nucleotides 961–963). The *Aovo-g50102* sequence showed a 21-bp deletion in exon 3 (nucleotides 317–337) that did not affect the ORF (Fig. 4), corresponding to a secondary structure region having turns and β strands (Parham et al. 1988). Other MHC I molecules also showed deletions at PBR level, like soluble HLA-G2 (sHLA-G2) which lacks an α2 domain (Fujii et al. 1994; Hunt et al. 2003).

Aotus MHC class I variability pattern at protein level

Analysing the *Aotus* MHC class I molecule leader peptide sequence showed that the substitutions found within this region were non-conservative; for example, *Ao-g2* sequences were characterised by most of them expressing a cysteine residue in position 2 which could be tryptophan or arginine in the other groups (Fig. 5).

Most variable positions in *Aotus* loci corresponded to residues located in α1 and α2 domains interacting with the peptide. The higher variable positions were: 65 (TCR), 66 (peptide), 67 (peptide), 70 (peptide), 73 (peptide), 74 (peptide), 80 (peptide), 82 (TCR), 95 (peptide), 114 (TCR), 116 (peptide), 152 (peptide-TCR), 156 (peptide) and 163 (peptide-TCR) (Fig. 5).

An exception to the high variability inside this region was constituted by positions 63 (residue forming pocket A, one of the most conserved pockets), 124, 141, 145, 147 (forming pocket E and F), 154 (interacting with TCR), 165 and 166 (interacting with TCR) (Fig. 5).

There were seven variable positions in the α3 domain, involving five conservative changes. The number of variable positions in this conserved domain increased to 12 when recombining sequences were included, *Ao-g2* contributing most non-conservative variant positions (data not shown).

The intracellular domains exhibited a high degree of conservation. For example, the trans-membrane domain only showed three variable positions (285, 307 and 310), whilst the cytoplasmic domain only had four non-conservative variable positions (317, 323, 324, 330) (Fig. 5).

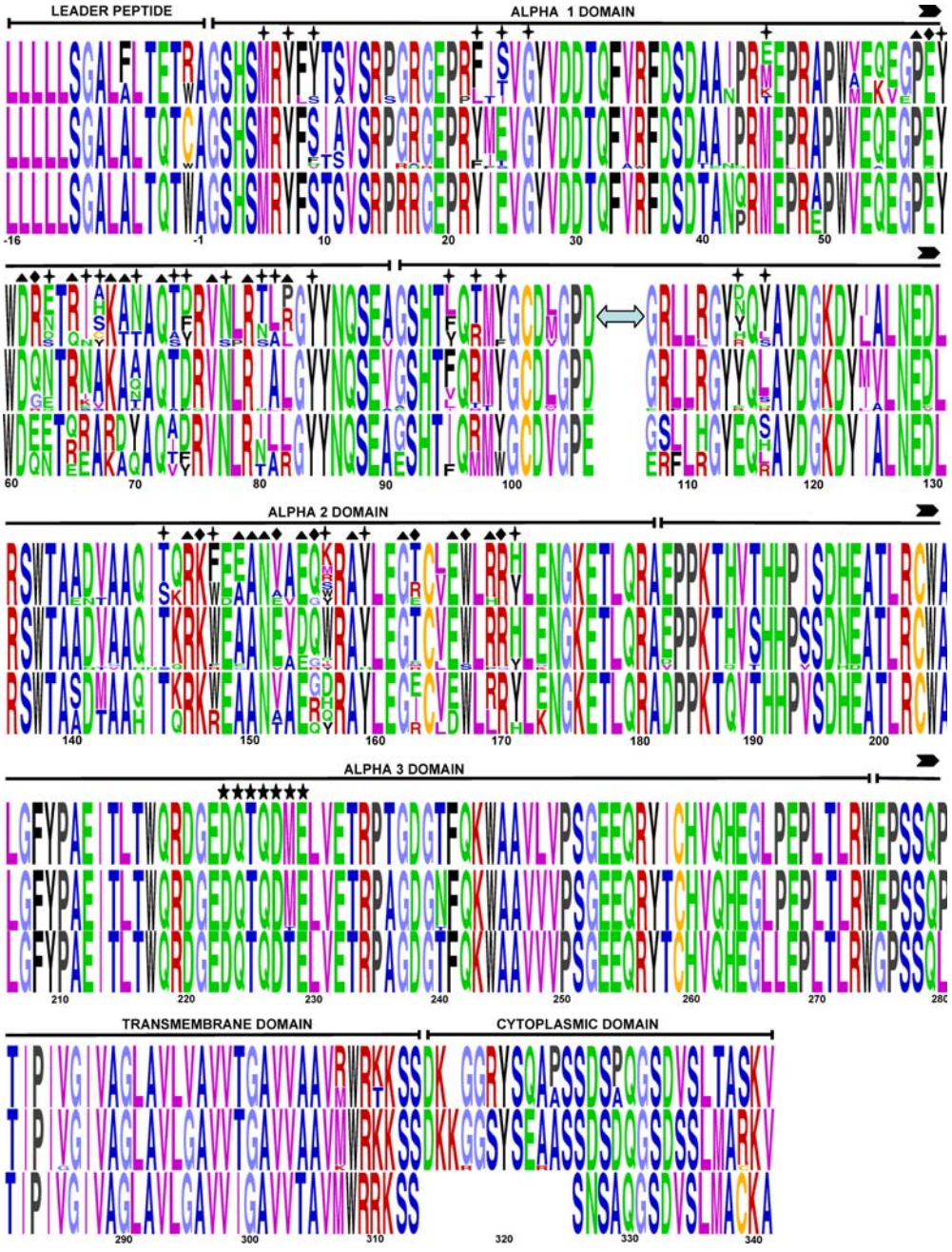

Fig. 5 *Aotus* MHC class I molecule amino acid sequence logo. The amino acid sequences corresponding to *Ao-g1*, *Ao-g2* and *Ao-g6* are located in the *first*, *second* and *third lanes*, respectively. The mature HLA-A02011 sequence is used for reference. Arrows represent one insertion found in allele *Aovog10301*. Symbols indicate HLA molecule residues interacting with the peptide (✦), the T-cell receptor (▲), or both (♦) and with β2-microglobulin (★). The division by domains is represented *above the amino acid sequences*

The CD8 glycoprotein is necessary for stabilising the interaction that exists between the TCR and the MHC class I molecule during antigen presentation and recognition processes (Gao et al. 1997; Salter et al. 1990). Conserved critical residues which are essential for interaction with $CD8^+$ T cells were found in *Aotus* MHC class I sequences' α_3 domains (Fig. 5), suggesting that their interaction with $CD8^+$ T cells occurs in the same way as in human molecules. Although the non-conservative substitution of residue Thr^{228} for Met^{228} was found in several *Aotus* MHC class I molecules, complete conservation of residues previously shown to be implied in hydrogen bond formation between HLA-A2 and CD8 (Gao et al. 1997) was also found in *Aotus* as follows: Gln^{115}, Asp^{122}, Glu^{128}, Glu^{232}, Lys^{243}, Thr^{225}, Gln^{262}, Leu^{230}, Gln^{226} and Asp^{227}, helping stabilise MHC class I and CD8 molecule interaction (Fig. 5).

Diversity and natural selection in *Aotus* MHC class I molecules

The results of Nei–Gojobori's selection test for the *Aotus* MHC class I sequences analysed showed that these molecules exhibited two tendencies. The first occurred when only variable positions in exons 1 and 4–7 were evaluated, showing a high degree of conservation inside each putative lineage; no significant selective pressure was detected due

Table 2 Synonymous (d_S) vs non-synonymous (d_N) changes in *Aotus* MHC class I molecules

Exons 1,4-7	Ao-MHC I[a]		Ao-g1[a]		Ao-g2[a]		Ao-g3[a]		Ao-g5[a]		Ao-g6[a]	
	All	Var	All	Var	All	Var	All	Var	All	Var	All	Var
d_S	0.0863	0.2213	0.0690	0.1678	0.0129	0.0300	0.0701	0.1663	0.0628	0.6044	0.0000	0.0000
d_N	0.0381	0.0937	0.0333	0.0798	0.0068	0.0158	0.0197	0.0472	0.0286	0.2296	0.0000	0.0000
d_N-d_S	−0.0482	−0.1276	−0.0358	−0.0880	−0.0061	−0.0142	−0.0504	−0.1191	−0.0342	−0.3748	0.0000	0.0000
Sel	–	–	–	–	ns	ns	–	–	–	ns	nc	nc
Exons 2-3	PBR	NonPBR	PBR	NonPBR	PBR	NonPBR	PBR	NonPBR	PBR	NonPBR	PBR	NonPBR
d_S	0.1277	0.0985	0.0655	0.0697	0.0291	0.0185	0.3071	0.1188	0.0000	0.0221	0.0000	0.0108
d_N	0.3457	0.0405	0.3091	0.0302	0.0964	0.0110	0.3485	0.0436	0.0000	0.0000	0.0000	0.0000
d_N-d_S	0.2180	−0.0581	0.2535	−0.0395	0.0673	−0.0075	0.0414	−0.0752	0.0000	−0.0221	0.0000	−0.0108
Sel	+	−	+	−	+	ns	ns	−	nc	−	nc	ns

[a]The number of d_S and d_N changes are shown for the seven exons encoding each *Aotus* MHC class I molecule. *All* Total nucleotides *Var* variable sites. Selection was calculated for peptide-binding regions (*PBR*s), and non-PBR (*NonPBR*) regions in exons 2 and 3. The result of applying the Nei–Gojobori test is shown in *row Sel*, where a *plus sign* indicates significant positive selection, a *minus sign* represents significant negative selection, *ns* means non-significant selection, and *nc* indicates non-calculable selection test

to such a high level of conservation. This tendency also occurred when all putative loci were considered as one group; in this case, a significant negative selective pressure was detected in exon 4, whilst no significant selective pattern was found in the other exons considered (Table 2).

The second tendency was presented in PBR positions, as shown previously (Cadavid et al. 1997; Parham et al. 1988; Suarez et al. 2003). *Aotus* MHC class I in exons 2 and 3 PBR positions showed strong positive selection pressure when all sequences were considered as one group. A consistent negative selection pressure was detected in exons 2 and 3 non-PBR positions, as has been reported in other primates, including humans (Cadavid et al. 1997; Hughes and Nei 1988; Parham et al. 1988). On the other hand, when each sequence group was considered individually, only *Ao-g1* and *Ao-g2* showed the previously described tendency; the remaining groups showed a high degree of conservation an/or non-significant selection (Table 2).

The major variability in *Aotus* MHC class I was concentrated in exons 2 and 3, whilst the remaining exons showed lesser variation; however, together they showed similar variation to exons 2 and 3 (Fig. 1). The recombinant tracts significantly increased variation inside each sequence group; for example, *Ao-g3* (the most recombinant group) exhibited great variability in exons involved in recombination events. A considerable decrease in non-synonymous changes was presented when recombinant sequences were excluded from computation, suggesting that recombination phenomena increased loci variability (data not shown).

The extreme conservation of *Ao-g5* and *Ao-g6* at PBR encoding exons, as well as in the remaining part of the gene, involved sequences from different *Aotus* species (*A. vociferans* and *A. nancymaae*); such active conservation pressure may indicate non-classical behaviour for these putative loci.

Discussion

Evidence was found of *HLA-B* or *HLA-C* orthologues in *Aotus*; *Aona-g60101* and *Aovo-g60101* sequences were related to *MHC-B* and *MHC-C* loci from Catarrhini primates. Cadavid et al. (1997) has reported *Atbe-B01* and *Pipi-B01* sequences as *MHC-B* locus orthologues and suggested that the Atelidae family represented an intermediate stage between Platyrrhini and Catarrhini MHC class I molecule diversification. However, Adams (Adams and Parham 2001) has argued that this relationship was not well supported and that these New World primate sequences do not segregate as typical *MHC-B* sequences do. Sawai et al. (2004), analysing the 5′ flanking region of several NWM MHC class I genes, have shown the existence of four *MHC*-B related loci in NWM, proposing, as Cadavid et al. (1997) have done, the existence of a true *MHC-B* orthology relationship in NWM.

Moreover, the relationship between NWM *MHC-B-C*-like and Catarrhini *MHC-B-C* was supported for exons 1 and 4–7, but not for exons 2 and 3 (PBR); at PBR, NWM B-like sequences cluster with NWM g groups (*Ao-g6*, *PipiB01*), and other MHC loci (*Aotr B1*, *B2*, *B3* and *AtbeB01*). These data indicate that an ancient recombination event or molecular convergence phenomena erased orthology between *MHC-B-C* Catarrhini lineage and its orthologues from PBR region in Platyrrhini, whilst this orthology was maintained in the remaining exons. The high variability of exons 2 and 3 masks this orthology when the entire molecule was used for evolutionary analysis, thus explaining Adams' postulate.

The existence of orthologous *MHC-B-C* lineages in *Aotus* shows that this locus predates the Platyrrhini–Catarrhini division. Cadavid (Cadavid et al. 1999; Watkins et al. 1988, 1990, 1991) did not find evidence of this lineage in *Saguinus*, indicating a loss of this lineage in that genus and, possibly, in all Callitrichidae. However, its presence in three divergent Platyrrhini genera indicates that this

locus constitutes a primitive state in NWM, and its loss was a derived characteristic (as has occurred in *Saguinus*).

The *Pipi-B01* sequence also shows a complete exon 6 deletion, just like *Ao-g6* sequences (data not shown). However, such deletion does not represent a specific characteristic of this lineage, since *Atbe-B01* shows a complete exon 6 sequence as *Aotr-B1* and *Aotr-B2*. Based on existing data, *Ao-g6* cannot be assigned to any of the loci proposed by Sawai et al. (2004). Clear differences exist with complete *Aotr-B1* and *Aotr-B2* sequences, and no association with *Aotr-B3* was demonstrated using its available partial sequence (complete exon 1 and partial exon 2, approximately 150 nucleotides) (not shown).

Low *Ao-g6* sequence variability led to us proposing that these sequences might constitute a non-classical type of MHC class I molecule. Besides, these types of molecules might be involved in specialised, important functions, thus explaining their presence and conservation in different Platyrrhini genera. Paradoxically, the NWM MHC I repertory appears to be a complementary image of the Catarrhini MHC I repertory: classical loci in the former appear as non-classical in the latter and vice versa.

Sawai et al. (2004) have found two *G* loci (*G1* and *G2*) in NWM (*Aotus* included). However, our analysis shows that these groups were included in the *g1* group (Fig. 2). Sawai's *Aotr-G2* complete sequence was included in all calculations in *Ao-g1*, and Sawai's *Aotr-G1* partial sequence (complete exon 1 and partial exon 2, 150 nucleotides) was also included in this group. This partial sector clearly and significantly differentiates *Ao-g2* from *Ao-g1* (not shown). This implies that classification based on 5′ flanking region might be incomplete or does not correspond to the sequence groups exhibited by coding regions.

Aotus MHC class I molecules are characterised by exhibiting a high variability level at PBR, as occurs in other New World primate genera like *Pithecia* (Suarez et al. 2003). This variability could be generated by two processes. The first concerns gene conversion in which new alleles are generated by non-reciprocal inter-gene exchange from another allele or from a different homologous gene; variability thus becomes increased within or between loci. This phenomenon has been implied in generating homogenisation or polymorphism, depending on the length of the exchanged segment. The transfer of long fragments can lead to homogenisation whilst exchanging short fragments produces polymorphism (Hogstrand and Bohme 1999a, b).

The second process concerns the birth-and-death model, proposing that new alleles are generated by repeated cycles of gene duplication and become inactive when they have accumulated deleterious mutations.

Evidence was found of inter-locus sequence exchange of long DNA fragments in MHC class I genes in Platyrrhini primates (including *Aotus*) in all gene sectors; this pattern was also shared by mice (Gu and Nei 1999). *In extremis*, this evolutionary pattern may lead to a high homogenisation process, producing a severe reduction in loci number and alleles. This is compatible with the nature of *Saguinus* MHC class I, where a great reduction of repertory has been observed (Watkins et al. 1988, 1990, 1991).

This contrasts with the evolutionary model found in HLA and *M. mulata* MHC class I genes, where most recombination is intra-locus. The gene conversion model could therefore play a more important role in MHC class I gene evolution in Platyrrhini and lemur monkeys than in Catarrhini.

This model of evolution agrees with the mouse MHC class I gene evolutionary model that also shows frequent inter-locus exchange, generating a lack of locus specificity in polymorphic alleles (Gu and Nei 1999; Yun et al. 1997). There have been reports of inter-locus recombination generating new alleles in humans; for example, *HLA-B0713* (Gu and Nei 1999), *HLA-B6702*, *HLA-B5401* and *HLA-B5402* originated from recombination events between *HLA-B* and *HLA-C* loci (Hildebrand et al. 1992; Steiner et al. 2002; Zemmour et al. 1992).

Another example of inter-locus recombination is exhibited by *MHC-A* in exon 4, which originated from a genetic exchange between an ancestral *MHC-A* locus and an *HLA-E*-like gene; all *A* locus alleles identified in humans and chimpanzees were apparently derived from this hybrid allele. The mosaic *Gogo-Oko* sequence (Lawlor et al. 1991) constituted another case in Catarrhini, showing a combined *MHC-H*- and *MHC-A*-derived organisation.

We believe that these two processes are not mutually exclusive, based on the evidence provided by *Aotus* sequences; the *Ao-g2* related sequences might be evidence of a duplication process (having subsequent differentiation) and the interchange of sequence tracts, together with point mutation, all driven by selective forces contributing towards increasing sequence diversity at PBR level, but also in other gene regions. An active duplication process also produces a significant number of pseudogenes; however, this was not detected in our observations. *Saguinus* shows a considerable number of pseudogenes, providing evidence of an active loci turnover process (Cadavid et al. 1999), contrasting with *Aotus* MHC class I. This might imply that the process undergone by *Aotus* had lower intensity than *Saguinus* or that the origin of *Ao-g2* related loci was recent. *Ao-g2*-related loci's supposed short time-span can explain their lower variability when compared to *Ao-g1*. However, the relationship of *Aotus* MHC I *g1*- and *g2*-like to other NWM and the relationship of *Ao-g6* to the *MHC B-C* lineage emphasizes the antiquity of the *Aotus* MHC class I repertory.

Ao-g5 could represent evidence of an inactivation process driving pseudogene sequence generation. However, its degree of conservation at interspecies level suggests a functional role as a soluble isoform, analogous to the non-classical MHC I molecules in Catarrhini which might be involved in developing a different function from that accomplished by membrane-bound molecules, although no specific function has been assigned to these soluble variants to date (Sebti et al. 2003).

Aotus shows a genus-specific clustering pattern for MHC class I (Fig. 2) that agrees with the idea that primate MHC class I genes are generated by a higher rate of evolutionary turnover when compared to MHC class II genes. This was evidenced by the low number of true orthologous

relationships in MHC class I genes amongst the different orders of primates analysed and by the genus-specific clustering pattern exhibited by them. This contrasts with MHC class II genes, where trans-specific polymorphism exhibits high frequency and high sequence conservation, even amongst different primate genera.

Two possible explanations for the high evolutionary variation showed by MHC class I genes have been proposed by Go (Go et al. 2003). First, the differences allowing MHC class I and class II molecules to be distinguished, peptide origin and processing, the types of cells expressing them and the type of T cells interacting with them might have a role in this variability. Peptide binding to MHC class I molecules seems to be more restricted because of the rigidity of the anchor residues, contrasting with the broadened peptide binding capability exhibited by MHC class II molecules. This causes MHC class I molecules to recognise peptides more specifically; so, if the peptide experiences subtle variations in anchor residues it cannot be recognised by the same MHC class I molecule, originating pressure for variation in the MHC class I that allows the peptide's recognition accompanied by the subsequent high rate of specialisation.

The second explanation refers to the higher gene duplication rates exhibited by MHC class I genes when compared to MHC class II genes (Hughes and Nei 1990; Hughes et al. 1990). The high number of gene duplication events generates diverse functional molecules but can also produce molecules having alterations that will be eliminated from the genome.

Aotus MHC class I molecules might interact in a similar way as humans with CD8 glycoprotein, based on the conservation of residues in the *Aotus* MHC class I sequences which are critical for interaction with CD8. This agrees with previous reports in humans establishing that mutations in these residues diminish or even abolish the binding between these two molecules. It is also interesting to note that the high degree of conservation between human and *Aotus* residues interacting with the TCR inside $\alpha 1$ and $\alpha 2$ domains suggests that, although the *Aotus* express molecules related to human's non-classical MHC class I loci, the interaction between their respective MHC and TCR might be occurring in a similar way.

Selective and diversity patterns indicate that the nature of *Aotus* putative loci varies from a clear MHC class I classic behaviour, with high variability and strong positive selection pressure in the PBR region (such as *Ao-g1*, *Ao-g2* and *Ao-g3*) and non-classical pattern, with low variation and strong conservation (such as *Ao-g5* and *Ao-g6*). The overall pattern follows that usually observed in MHC class I: the PBR is under strong positive selection pressure, whilst the rest of the molecule is under a conservative force that is evidenced as a trend towards negative selection.

The above results lead us to conclude that the *Aotus* MHC class I sequence groups found here (*Ao-g1* to *Ao-g6*) display a characteristic mechanism of possible inter-loci recombination that contributes towards generating variability; these characteristics distinguish them from MHC class I genes expressed by Catarrhini. The PBR divergence between NWM *B-C* sequences and Catarrhini *MHC-B-C* loci indicates a different peptide-binding specificity between them, suggesting functional divergence processes.

Acknowledgements We are greatly indebted to Johana Ochoa, Andrómeda Gómez, Carlos Barrero and José E. Guerrero for their valuable help in preparing this manuscript and Paola Barato for her help in obtaining the blood samples. We would like to thank Jason Garry for reading and correcting the manuscript. This work was financed by the Colombian Ministry of Public Health. All experiments and work done regarding this paper involving live animals was carried out in accord with existing pertinent law in Colombia. The first and second authors contributed equally.

References

Adams EJ, Parham P (2001) Species-specific evolution of MHC class I genes in the higher primates. Immunol Rev 183:41–64

Cadavid LF, Hughes AL, Watkins DI (1996) MHC class I-processed pseudogenes in New World primates provide evidence for rapid turnover of MHC class I genes. J Immunol 157:2403–2409

Cadavid LF, Mejia BE, Watkins DI (1999) MHC class I genes in a New World primate, the cotton-top tamarin (*Saguinus oedipus*), have evolved by an active process of loci turnover. Immunogenetics 49:196–205

Cadavid LF, Shufflebotham C, Ruiz FJ, Yeager M, Hughes AL, Watkins DI (1997) Evolutionary instability of the major histocompatibility complex class I loci in New World primates. Proc Natl Acad Sci USA 94:14536–14541

Cadavid LF, Watkins DI (1997) The duplicative nature of the MHC class I genes: an evolutionary perspective. Eur J Immunogenet 24:313–322

Crew MD, Filipowsky ME, Neshat MS, Smith GS, Walford RL (1991) Transmembrane domain length variation in the evolution of major histocompatibility complex class I genes. Proc Natl Acad Sci USA 88:4666–4670

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14:1188–1190

Daubenberger CA, Salomon M, Vecino W, Hubner B, Troll H, Rodriques R, Patarroyo ME, Pluschke G (2001) Functional and structural similarity of V gamma 9V delta 2 T cells in humans and *Aotus* monkeys, a primate infection model for *Plasmodium falciparum* malaria. J Immunol 167:6421–6430

Dayhoff MO, Schwartz RM, Orcutt BC (1978) A model of evolutionary change in proteins. In: Dayhoff M (ed) Atlas of protein sequence and structure. National Medical Research Foundation, pp 345–352

Diaz D, Naegeli M, Rodriguez R, Nino-Vasquez JJ, Moreno A, Patarroyo ME, Pluschke G, Daubenberger CA (2000a) Sequence and diversity of MHC *DQA* and *DQB* genes of the owl monkey *Aotus nancymaae*. Immunogenetics 51:528–537

Diaz OL, Daubenberger CA, Rodriguez R, Naegeli M, Moreno A, Patarroyo ME, Pluschke G (2000b) Immunoglobulin kappa light-chain *V*, *J*, and *C* gene sequences of the owl monkey *Aotus nancymaae*. Immunogenetics 51:212–218

Ennis PD, Zemmour J, Salter RD, Parham P (1990) Rapid cloning of HLA-A,B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. Proc Natl Acad Sci USA 87:2833–2837

Favre N, Daubenberger C, Marfurt J, Moreno A, Patarroyo M, Pluschke G (1998) Sequence and diversity of T-cell receptor alpha *V*, *J*, and *C* genes of the owl monkey *Aotus nancymaae*. Immunogenetics 48:253–259

Fujii T, Ishitani A, Geraghty DE (1994) A soluble form of the HLA-G antigen is encoded by a messenger ribonucleic acid containing intron 4. J Immunol 153:5516–5524

Gao GF, Tormo J, Gerth UC, Wyer JR, McMichael AJ, Stuart DI, Bell JI, Jones EY, Jakobsen BK (1997) Crystal structure of the complex between human CD8alpha(alpha) and *HLA-A2*. Nature 387:630–634

Go Y, Satta Y, Kawamoto Y, Rakotoarisoa G, Randrianjafy A, Koyama N, Hirai H (2003) Frequent segmental sequence exchanges and rapid gene duplication characterize the MHC class I genes in lemurs. Immunogenetics 55:450–461

Gu X, Nei M (1999) Locus specificity of polymorphic alleles and evolution by a birth-and-death process in mammalian MHC genes. Mol Biol Evol 16:147–156

Gur H, Geppert TD, Lipsky PE (1997) Structural analysis of class I MHC molecules: the cytoplasmic domain is not required for cytoskeletal association, aggregation and internalization. Mol Immunol 34:125–132

Gur H, Geppert TD, Wacholtz MC, Lipsky PE (1999) The cytoplasmic and the transmembrane domains are not sufficient for class I MHC signal transduction. Cell Immunol 191:105–116

Hildebrand WH, Madrigal JA, Little AM, Parham P (1992) HLA-Bw22: a family of molecules with identity to *HLA-B7* in the alpha 1-helix. J Immunol 148:1155–1162

Hillis D (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol 42:182–192

Hogstrand K, Bohme J (1999a) Gene conversion can create new MHC alleles. Immunol Rev 167:305–317

Hogstrand K, Bohme J (1999b) Gene conversion of major histocompatibility complex genes is associated with CpG-rich regions. Immunogenetics 49:446–455

Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167–170

Hughes AL, Nei M (1990) Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol Biol Evol 7:491–514

Hughes AL, Ota T, Nei M (1990) Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. Mol Biol Evol 7:515–524

Hunt JS, Pace JL, Morales PJ, Ober C (2003) Immunogenicity of the soluble isoforms of HLA-G. Mol Hum Reprod 9:729–735

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

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

Knapp LA, Cadavid LF, Watkins DI (1998) The MHC-E locus is the most well conserved of all known primate class I histocompatibility genes. J Immunol 160:189–196

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

Kundig TM, Bachmann MF, Oehen S, Hoffmann UW, Simard JJ, Kalberer CP, Pircher H, Ohashi PS, Hengartner H, Zinkernagel RM (1996) On the role of antigen in maintaining cytotoxic T-cell memory. Proc Natl Acad Sci USA 93:9716–9723

Lafont BA, Buckler-White A, Plishka R, Buckler C, Martin MA (2003) Characterization of pig-tailed macaque classical MHC class I genes: implications for MHC evolution and antigen presentation in macaques. J Immunol 171:875–885

Lawlor DA, Warren E, Taylor P, Parham P (1991) Gorilla class I major histocompatibility complex alleles: comparison to human and chimpanzee class I. J Exp Med 174:1491–1509

Meyerhans A, Vartanian JP, Wain-Hobson S (1990) DNA recombination during PCR. Nucleic Acids Res 18:1687–1691

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–4126

Nino-Vasquez JJ, Vogel D, Rodriguez R, Moreno A, Patarroyo ME, Pluschke G, Daubenberger CA (2000) Sequence and diversity of *DRB* genes of *Aotus nancymaae*, a primate model for human malaria parasites. Immunogenetics 51:219–230

O'Callaghan CA, Bell JI (1998) Structure and function of the human MHC class Ib molecules HLA-E, HLA-F and HLA-G. Immunol Rev 163:129–138

Parham P, Lomen CE, Lawlor DA, Ways JP, Holmes N, Coppin HL, Salter RD, Wan AM, Ennis PD (1988) Nature of polymorphism in HLA-A, -B, and -C molecules. Proc Natl Acad Sci USA 85:4005–4009

Patarroyo ME, Romero P, Torres ML, Clavijo P, Moreno A, Martinez A, Rodriguez R, Guzman F, Cabezas E (1987) Induction of protective immunity against experimental infection with malaria using synthetic peptides. Nature 328:629–632

Qiu X, Wu L, Huang H, McDonel PE, Palumbo AV, Tiedje JM, Zhou J (2001) Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. Appl Environ Microbiol 67:880–887

Rzhetsky A, Nei M (1993) Theoretical foundation of the minimum-evolution method of phylogenetic inference. Mol Biol Evol 10:1073–1095

Salter RD, Benjamin RJ, Wesley PK, Buxton SE, Garrett TP, Clayberger C, Krensky AM, Norment AM, Littman DR, Parham P (1990) A binding site for the T-cell co-receptor CD8 on the alpha 3 domain of HLA-A2. Nature 345:41–46

Sawai H, Kawamoto Y, Takahata N, Satta Y (2004) Evolutionary relationships of major histocompatibility complex class I genes in simian primates. Genetics 166:1897–1907

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

Sawyer SA (1999) GENECONV: a computer package for the statistical detection of gene conversion. Distributed by the author (http://www.math.wustl.edu/~sawyer). Department of Mathematics, Washington University, St. Louis

Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 18:6097–7100

Sebti Y, Le Friec G, Pangault C, Gros F, Drenou B, Guilloux V, Bernard M, Lamy T, Fauchet R, Amiot L (2003) Soluble HLA-G molecules are increased in lymphoproliferative disorders. Hum Immunol 64:1093–1101

Steiner NK, Baldassarre L, Koester R, Ng J, Hartzman RJ, Hurley CK (2002) Novel HLA-B alleles formed by an inter-locus recombination with HLA-C, HLA-B*0713 and B*6702. Tissue Antigens 59:148–150

Suarez CF, Cardenas PP, Llanos-Ballestas EJ, Martinez P, Obregon M, Patarroyo ME, Patarroyo MA (2003) Alpha1 and alpha2 domains of *Aotus* MHC class I and Catarrhini MHC class Ia share similar characteristics. Tissue Antigens 61:362–373

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Torres OM, Enciso S, Ruiz F, Silva E, Yunis I (1998) Chromosome diversity of the genus *Aotus* from Colombia. Am J Primatol 44:255–275

Vecino W, Daubenberger C, Rodriguez R, Moreno A, Patarroyo M, Pluschke G (1999) Sequence and diversity of T-cell receptor beta-chain *V* and *J* genes of the owl monkey *Aotus nancymaae*. Immunogenetics 49:792–799

Watkins DI, Garber TL, Chen ZW, Toukatly G, Hughes AL, Letvin NL (1991) Unusually limited nucleotide sequence variation of the expressed major histocompatibility complex class I genes of a New World primate species (*Saguinus oedipus*). Immunogenetics 33:79–89

Watkins DI, Hodi FS, Letvin NL (1988) A primate species with limited major histocompatibility complex class I polymorphism. Proc Natl Acad Sci USA 85:7714–7718

Watkins DI, Letvin NL, Hughes AL, Tedder TF (1990) Molecular cloning of cDNA that encode MHC class I molecules from a New World primate (*Saguinus oedipus*). Natural selection acts at positions that may affect peptide presentation to T cells. J Immunol 144:1136–1143

Wheeler CJ, Maloney D, Fogel S, Goodenow RS (1990) Microconversion between murine H-2 genes integrated into yeast. Nature 347:192–194

Yun TJ, Melvold RW, Pease LR (1997) A complex major histocompatibility complex D locus variant generated by an unusual recombination mechanism in mice. Proc Natl Acad Sci USA 94:1384–1389

Zemmour J, Gumperz JE, Hildebrand WH, Ward FE, Marsh SG, Williams RC, Parham P (1992) The molecular basis for reactivity of anti-Cw1 and anti-Cw3 alloantisera with HLA-B46 haplotypes. Tissue Antigens 39:249–257