Mamu-I: A Novel Primate MHC Class I B
-Related Locus with Unusually Low Variability

Julie A. Urvater, Nel Otting, Jamie H. Loehrke, Richard Rudersdorf, Igor I. Slukvin, Marian S. Piekarczyk, Thaddeus G. Golos, Austin L. Hughes, Ronald E. Bontrop and David I. Watkins

*J Immunol* 2000; 164:1386-1398; doi: 10.4049/jimmunol.164.3.1386
http://www.jimmunol.org/content/164/3/1386

References

This article cites 59 articles, 27 of which you can access for free at:
http://www.jimmunol.org/content/164/3/1386.full#ref-list-1

Subscription

Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

Permissions

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Mamu-I: A Novel Primate MHC Class I B-Related Locus with Unusually Low Variability

Julie A. Urvater, Nel Otting, Jamie H. Loehrke, Richard Rüdersdorf, Igor I. Slukvin, Marian S. Pickarczyk, Thaddeus G. Golos, Austin L. Hughes, Ronald E. Bontrop, and David I. Watkins

The rhesus macaque is an important animal model for several human diseases and organ transplantation. Therefore, definition of the MHC of this species is crucial to the development of these models. Unfortunately, unlike humans, lymphocytes from a single rhesus macaque express up to 12 different MHC class I cDNAs. From which locus these various alleles are derived is unclear. In our attempts to define the MHC class I loci of the rhesus macaque, we have identified an unusual MHC class I locus, Mamu-I. We isolated 26 I locus alleles from three different macaque species but not from three other Cercopithecine genera, suggesting that the I locus is the result of a recent duplication of the B locus occurring after the divergence of macaques from the ancestor of the other extant Cercopithecine genera. Mamu-I mRNA transcripts were detected in all tissues examined and Mamu-I protein was produced in rhesus B lymphoblastoid cell lines. Furthermore, Mamu-I protein was detected by flow cytometry on the surface of human 721.221 cells transfected with Mamu-I. In contrast to the polymorphism present at this locus, there is unusually low sequence variability, with the mean number of nucleotide differences between alleles being only 3.6 nt. Therefore, Mamu-I is less variable than any other polymorphic MHC class I locus described to date. Additionally, no evidence for positive selection on the peptide binding region was observed. Together, these results suggest that Mamu-I is an MHC class I locus in primates that has features of both classical and nonclassical loci. The Journal of Immunology, 2000, 164: 1386–1398.

Because rhesus macaques and humans have very similar immune systems, the macaque has been widely used as a model for many human diseases and organ transplantation. SIV infection of macaques has been perhaps the best animal model for the study of human AIDS (1, 2). Additionally, macaques are becoming increasingly important for preclinical evaluation of antirejection therapies in transplantation biology (3–7). Many genes that encode proteins important in the immune system are similar in humans and rhesus monkeys. In particular, homologues of the human MHC class I and II and TCR genes are found in macaques (8–13). However, to investigate the cellular immune responses to pathogens and allografts it is crucial to define the MHC of this species.

Classical MHC class I glycoproteins are expressed in almost every cell type and play a central role in immune recognition of self and nonself by binding pathogen-derived peptides and presenting them to CD8-positive CTLs. The function of nonclassical MHC class I glycoproteins is just beginning to be understood. MHC class II glycoproteins are expressed largely on cells of the immune system and bind exogenously derived peptides and present them to CD4-positive T helper cells. MHC class II genes have been extensively studied in the rhesus monkey. These class II genes are very well conserved and are stable over very long evolutionary periods, such that allelic lineages are preserved even between humans and macaques (9, 14, 15). Methods for molecular typing of the MHC class II alleles in the macaque have now been developed and nearly 100 rhesus DRB alleles have been identified (16).

Although analysis of the MHC class II loci in the rhesus macaque has been relatively straightforward, the MHC class I loci of this species have been more difficult to define. We and others have identified orthologs of the human HLA-A, -B, -E, -F, and -G genes in the rhesus monkey (8, 17–19). The orthologs of HLA-G in the rhesus monkey are pseudogenes, but the recently identified nonclassical Mamu-AG (Macaca mulatta-AG) locus may have evolved to serve the function of HLA-G in this primate (17, 20–22). We have also shown that duplication of the rhesus MHC class I A and B loci has occurred, making the study of the MHC class I loci in this species exceptionally complex (our unpublished observations and Ref. 8). In particular, the presence of multiple MHC class I A and B loci makes the development of molecular technologies for MHC class I typing of this species even more challenging.

Therefore, definition of the MHC class I loci and alleles in the rhesus macaque is central to the development of this animal model for understanding human diseases. In this paper we describe a new MHC class I locus in rhesus macaques, Mamu-I (Macaca mulatta-I), which is likely the result of a recent duplication of a classical MHC class I B locus. These results have implications for the development of the rhesus macaque as an animal model for human...
diseases and organ transplantation and, in addition, for the elicitation of cellular immune responses in this species.

Materials and Methods

Isolation and identification of full-length Mamu-I cDNAs

We used two different methods for isolation and identification of full-length Mamu-I cDNAs. The first was described previously (23). The second method was as follows. Total cellular RNA was extracted from 2–7 × 10^10 rhesus monkey lymphocytes using RNAzol (Tel-Test, Friendswood, TX). cDNA was synthesized from 0.1 to 1 μg of RNA in a 20-μl reaction containing 50 mM Tris (pH 8.3), 5 mM MgCl₂, 1 mM each of dATP, dGTP, dCTP, and dTTP (Gene AMP, Perkin-Elmer, Foster City, CA), 0.5 μM random primers (Promega, Madison, WI), 50 U of SuperScript II Reverse Transcriptase (Life Technologies, Gaithersburg, MD), and 20 U of RNase inhibitor (Gene AMP-Perkin-Elmer). cDNA was synthesized at room temperature for 10 min, 42°C for 15 min, 99°C for 5 min, and 5°C for 5 min in a Perkin-Elmer Cetus 9600 thermocycler (Norwalk, CT). PCR was then performed in a Perkin-Elmer Cetus 9600 using the previously described rhesus macaque B locus-speciﬁc primers, 5′MBS and 3′MBS (8), each at a final concentration of 0.25 mM. The PCR mixture contained 2 mM MgCl₂, 50 mM Tris (pH 8.3), and 2.5 U AmpliTag DNA polymerase (Perkin-Elmer Cetus) in a final volume of 100 μl. The reactions were denatured initially for 2 min at 94°C before being subjected to 30 cycles of 94°C for 1 min, 60°C for 1 min, 72°C for 1.5 min, and a single final extension at 72°C for 10 min. Ampliﬁed, B locus-speciﬁc products were gel puriﬁed using QIAEX II suspension (Qiagen, Santa Clarita, CA) and were subcloned into the pCR2.1 vector using the TA cloning kit from Invitrogen (Carlsbad, CA). A total of 500 ng of plasmid DNA was sequenced using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq DNA polymerase, FS, (Perkin-Elmer-Applied Biosystems, Foster City, CA) according to the manufacturer’s instructions. At least three copies of each cDNA were sequenced to ensure that clones did not contain PCR-generated artifacts. Full-length sequences were generated using previously published primers (8), assembled using the AutoAssembler DNA sequence assembly software (Perkin-Elmer-Applied Biosystems), and analyzed with MacVector sequence analysis software (Oxford Molecular Group, Campbell, CA).

Mamu-I sequence-speciﬁc PCR (PCR-SSP), RT-PCR-SSP, and direct sequencing of PCR products

Mamu-I-speciﬁc primers were designed based on a rhesus macaque MHC class I sequence database generated by ourselves and others (MIMIF2, 5′-CTCCCCACCTCGGATTTGCTG-3′; MMIR, 5′-GTGGACCGTCCTTCCCC-3′). PCR-SSP was performed either on diluted 5′MBS/3′MBS PCR product or on genomic DNA using the QiAamp Blood kit (Qiagen).

Mamu-I PCR-SSP on genomic DNA was performed in 26 μl of 1× PCR buffer F from the Invitrogen PCR Optimizer kit, 2 mM MgCl₂, 2.5 mM of each of the four dNTPs (Invitrogen), 1.9 mM of each specific primer, and 0.4 μM each Mamu-DRB internal control primer (24) and 1.25 U of AmpliTag DNA polymerase (Perkin-Elmer, Foster City, CA). The reactions were denatured initially for 2 min at 96°C before being subjected to the following cycling conditions: 10 cycles of 96°C for 25 s, 70°C for 50 s, and 72°C for 45 s; and a final four cycles of 96°C for 25 s, 65°C for 50 s, and 72°C for 45 s; and a final four cycles of 96°C for 25 s, 55°C for 1 min, and 72°C for 1 min 20 s. A total of 6 μl of the PCR mix was electrophoresed in 1% agarose, stained with ethidium bromide, and visualized under UV light. A total of 5 μl of PCR product from positive reactions was then incubated for 15 min at 37°C with 2 U shrimp alkaline phosphatase and 10 U exonuclease 1 to remove unincorporated dNTPs and residual single-stranded PCR primers, respectively (United States Biochemical, Cleveland, OH). This reaction mixture was then diluted with 7–18 μl of H₂O and 3 μl were sequenced directly with the same primers used for PCR-SSP using the ABI PRISM cycle sequencing ready reaction kit as described above. The forward and reverse sequences were assembled using the MacVector sequence assembly software (Perkin-Elmer-Applied Biosystems). A second set of Mamu-I-specific primers was designed to amplify from a few samples that failed to amplify with the MIMIF2/MMIR primer set (MIMIF3, 5′-CTGGAGGTATTCCCCGACCCA-3′; MMIR2, 5′-TTGTCTGAGCCCGCTCAC-3′), the same conditions were used for amplification with this set of primers.

For Mamu-I PCR-SSP on diluted PCR product, full-length, locus-specific PCR products amplified with the MBS primers were diluted 1:250 in 500 μl H₂O and were subjected to a second round of nested, Mamu-I-specific PCR using buffer F from the Invitrogen PCR Optimizer kit. Mamu-I primers were used at a final concentration of 0.5 μM. The reactions were denatured initially for 2 min at 96°C before being subjected to the following cycling conditions: five cycles of 96°C for 25 s, 70°C for 50 s, and 72°C for 45 s; and 21 cycles of 96°C for 25 s, 69°C for 50 s, and 72°C for 45 s; and a final four cycles of 96°C for 25 s, 60°C–65°C for 1 min, and 72°C for 1 min 20 s.

For Mamu-I-specific RT-PCR, 1 μl of total RNA was reverse transcribed to cDNA using an oligo(dT) primer and Maloney murine leukemia virus reverse transcriptase (Perkin-Elmer). The cDNA was amplified with MIMIF2/MMIR primers for 36 cycles, with each cycle consisting of 94°C for 45 s, 66°C for 45 s, and 72°C for 2 min, in a reaction mixture containing 10 μM Tris-HCl, 1.5 mM MgCl₂, 50 mM KCl, 0.2 mM of each dNTP, and 1 U AmpliTag DNA polymerase (Perkin-Elmer). RT-PCR products were directly sequenced as described above.

Nomenclature of rhesus MHC class I alleles

The DNA and cDNA sequences were compared with a database of macaque MHC class I sequences generated by ourselves and others. Unique sequences were named in the order isolated according to established guidelines (25). In the nomenclature of the MHC class I genes, alleles that differ by even a single, non-amino acid-changing nucleotide substitution (synonymously identical) are given separate designations. For example, the alleles designated Mamu-I*01011, Mamu-I*01012, Mamu-I*01013 and Mamu-I*01014 are identical at the amino acid level but differ by one or more nucleotides. The alleles designated Mamu-I*01011, Mamu-I*02011, Mamu-I*04, Mamu-I*05, etc. differ by one or more amino acids.

Transfection of Mamu-I molecules into an MHC class I null cell line

Clones containing the consensus cDNA for full-length Mamu-I*02011 were subcloned into the pKG5 expression vector (a gift from Andrew McMichael, Oxford University, Oxford, U.K.). This vector was then electroporated into the 721.221 cell line, a cloned EBV-transformed B lymphoblastoid cell line (BCLL) (4) with homozygous deletions of the MHC class I loci (26). A total of 5 × 10⁶ 721.221 cells were transfected in a 1-cm electroporation cuvette with 5 μg of plasmid DNA. Electroporation was conducted at 1000 V and a capacitance of 600 μF. The cells were then put in RPMI 1640 culture medium supplemented with 50 U/ml penicillin, 50 μg/ml streptomycin, 2 mM l-glutamine, 5% defined FBS (HyClone, Logan, UT), and 10% defined supplemented bovine calf serum (HyClone) and were plated onto four 24-well plates. The cells were incubated for 4 days at 37°C. On day 5, the cells were placed under selection by adding 1 ml of culture medium containing G418 (Life Technologies) for a final concentration of 500 μg/ml. About 4 wk later, viable transfectants were tested for MHC class I surface expression by flow cytometry with W6/32, B7.7, or BBM.1 mAbs (American Type Culture Collection, Manassas, VA). The transfectant with the highest level of MHC class I expression was selected to be grown up for one-dimensional isoelectric focusing (1-D IEF) analysis.

Flow cytometry

A total of 2 × 10⁵–1 × 10⁶ cells were washed twice in PBS with 2% FBS and 0.1% sodium azide, and 50 μl of supernatant from the BBM.1 hybridomas were added to each sample and incubated on ice for 30 min. After additional washes, FITC-conjugated goat anti-mouse Abs (Dako, Carpinteria, CA) were added. After 30 min on ice, samples were washed twice, and 250 μl 2% paraformaldehyde was added. Sample data were acquired on a Becton Dickinson FACS-Calibur instrument and analyzed using CellQuest software (Becton Dickinson Immunocytometry Systems, San Jose, CA).

Radiolabeling and electrophoretic analysis

MHC class I glycoproteins were translated with [35S]methionine and immunoprecipitated with W6/32 and B7.7 mouse mAbs directed against human MHC class I proteins, as described previously (27).

Phylogenetic analysis

Gene trees were constructed using the neighbor-joining method (28). Gene trees were constructed using the neighbor-joining method (28). Gene trees were constructed using the neighbor-joining method (28).
At sites in which the alignment indicated a gap, nucleotides at this position in all the sequences were deleted. The reliability of the tree topology was tested by the bootstrap method (30). One thousand replications was used for bootstrap analysis, the results of which are reported as the strength of support for a given branch topology. Calculations of the rates of synonymous and nonsynonymous substitutions were performed using the method of Nei and Gojobori (31), and variances were calculated using the method of Nei and Jin (32). Phylogenetic analyses were performed using MEGA software (33). Tests of the significances of these differences were performed using a two-tailed test. A binomial test was used to determine the significance of differences in %GC content.

GenBank accession numbers

The sequences described in this manuscript have been deposited with GenBank and were assigned accession numbers AF161860–AF161882. Mamu-I*02011, Mamu-I*05, and Mamu-I*06 were previously submitted with the names and accession numbers Mamu-B*0902 AF157403, Mamu-B*0903 AF157404, and Mamu-B*0904 AF157405, respectively. The GenBank accession numbers for all other sequences used in Figs. 2, 3, and 4 are: Gogo-A*0101 X60258, Gogo-B*0101 X60255, Gogo-B*0102 X60693, Gogo-C*0201 X60251, HLA-A*0101 M24043, HLA-A*0201 K02883, HLA-A*1101 X13111, HLA-B*0702 M32317, HLA-B*2702 X03664 and X03667, HLA-B*3506 M84381, HLA-B*3802 L25028, HLA-C*1203 U06695, HLA-C*1402 U06487, HLA-C*1601 M24097, Hyla-A*0109 X60489, Hyla-A*0109 U50890, Hyla-A*0109 U50896, Hyla-A*0109 U50893, Hyla-B*02 U50837, Hyla-Mamu-I*03 U41379, Hyla-Mamu-I*04 U41880, Hyla-Mamu-I*05 U41831, Hyla-Mamu-I*06 U1834, Hyla-Mamu-I*07 U1832, Hyla-Mamu-I*08 (34), Hyla-Mamu-I*12 AF157398, Hyla-Mamu-I*1301 AF157399, Hyla-Mamu-I*01 U42837, Hyla-Mamu-I*02 U41833, Hyla-Mamu-I*03 U41825, Hyla-Mamu-I*04 U41826, Hyla-Mamu-I*05 U41828, Hyla-Mamu-I*07 U41829, Hyla-Mamu-I*08 U41830, Hyla-Mamu-I*12 (34), Paan-A*01 U35625, Paan-A*02 U35625, Paan-A*01 L39093, Papa-A*02 L39094, Papa-B*02 U05578, Papa-B*03 U05578, Papa-B*04 U05577, Papa-A*01 L39094, Papa-B*02 X13116, Papa-B*01 U50582, Papa-A*01 M30680, Papa-A*02 U50084, Papa-B*01 U50086, Papa-B*02 U50087, and Papa-B*03 U50088.

Results

Given the confusing nature of the MHC class I loci in the rhesus macaque, we undertook an extensive analysis of cDNAs expressed in a large number of different rhesus macaques. While investigating a large number of different rhesus macaques, we undertook an extensive analysis of cDNAs expressed in a large number of different rhesus macaques, while investigating the MHC class I loci already exist in primates (8, 17–19, 35), we have called this a “new” locus. Alternatively, this animal could have two haplotypes that do not contain a Mamu-I locus.

Alleles at this new locus are inherited in a Mendelian fashion

To investigate whether the Mamu-B*09 variants that we had observed in our cohort of Shigella-infected rhesus macaques were indeed the products of a distinct B-like locus, we examined the inheritance of these alleles in a family of rhesus macaques. Given the confusing nature of the MHC class I loci in the rhesus macaque, we undertook an extensive analysis of cDNAs expressed in a large number of different rhesus macaques. While investigating the MHC class I loci already exist in primates (8, 17–19, 35), we have called this a “new” locus. Alternatively, this animal could have two haplotypes that do not contain a Mamu-I locus.

Cloning and sequencing revealed that the Mamu-I cDNAs were 1098 nt in length from the start codon in exon 1 to the end of exon 8 (Fig. 2). Exon 8 is untranslated in Mamu-I molecules due to a stop codon just before the end of exon 7, a characteristic of classical MHC class I B molecules. We showed previously that some human A and B locus-specific nucleotides (8) were conserved in rhesus monkey class I cDNAs (37). Macaque I molecules shared several B locus-specific nucleotide substitutions with Mamu-B alleles (Fig. 2, see nucleotides marked with *). Additionally, there were two areas of similarity between I alleles and Mamu-B alleles that suggested that interlocus recombination may have occurred (Figs. 2 and 3, boxed areas). Based on a comparison of I alleles and a database of over 50 rhesus A and B alleles, we noted only four nucleotide substitutions that were specific for Mamu-I alleles (Fig. 2, nucleotides marked with ●).

We are currently screening a cDNA library from an additional offspring in this family that has inherited the a haplotype to address this question (H. Horton, manuscript in preparation).

Mamu-I is polymorphic and its full-length products have the characteristics of functional, classical class I molecules

We wish to investigate whether this new locus was classical or nonclassical. Generally, the difference between these two categories of MHC molecules is the level of polymorphism and cell-surface expression (36). Nonclassical molecules tend to have limited polymorphism and low cell-surface expression in restricted tissues, whereas classical molecules are highly polymorphic and are readily detected on the surface of nearly all cell types.

PCR-SSP on DNA or diluted Mamu-B-specific PCR product from 60 rhesus macaques revealed that, with one exception, all animals were positive for at least one, often two, Mamu-I alleles (data not shown). Mamu-I alleles were as few as one allele per animal in the family of rhesus macaques. An additional 88090, which was found in each individual. Mamu-B typing will be described elsewhere (H. Horton, manuscript in preparation). We were unable to detect Mamu-I on the a haplotype.

FIGURE 1. Inheritance of Mamu-I in a family of rhesus macaques. Animals were typed by PCR-SSP and cloning and sequencing. Haplotypes are designated by lowercase letters. Mamu-I alleles on each haplotype are shown. At least 2 and as many as 10 Mamu-B alleles in addition to Mamu-I were found in each individual. Mamu-B typing will be described elsewhere (H. Horton, manuscript in preparation). *, We were unable to detect Mamu-I on the a haplotype.
<table>
<thead>
<tr>
<th>Exon 1</th>
<th>Exon 2</th>
<th>Exon 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>-70</td>
<td>-60</td>
<td>-50</td>
</tr>
<tr>
<td>-40</td>
<td>-30</td>
<td>-20</td>
</tr>
<tr>
<td>-10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Exon 1** consensus
...

**Exon 2** consensus
...

**Exon 3** consensus
...

---

The Journal of Immunology
The predicted amino acid sequence of MHC class I alleles is 362 aa in length, including the leader peptide. Macaque I proteins contained several amino acid residues that are conserved in all known mammalian and avian classical alleles (Fig. 3 and Ref. 38). They also contained the conserved cysteine residues at positions 101 and 164, and 203 and 259 that form the structurally important disulfide bonds. Furthermore, the N-linked glycosylation site at position 86 in the α1 domain was also present. There were only two residues that were specific to and conserved among all macaque I molecules, cysteine at position 99 and glycine at position 137 (Fig. 3).

Figure 2. Continued
The Mamu-I locus is derived from a recent duplication of an MHC class I B locus

Sequence comparisons showed a high degree of similarity between Mamu-I alleles and Mamu-B alleles. Therefore, we wished to determine whether the Mamu-I locus was derived from the ancestor of a B locus and used gene tree analysis to address this question.

Gene tree analysis of full-length Mamu-I sequences revealed that this locus is derived from a classical MHC class I B locus as is shown by the significant clustering of the Mamu-I alleles with other primate B alleles (Fig. 4). Because exons 4–8 contain the majority of the locus-specific nucleotide substitutions, we conducted gene tree analysis of this region. Gene trees constructed with exons 4–8 confirmed the clustering seen in the full-length trees (data not shown). Mamu-I alleles did not consistently cluster with any particular Mamu-B alleles in gene trees of exon 2 (data not shown).

To investigate when this duplication event may have occurred, we conducted PCR-SSP using two different sets of Mamu-I-specific primers on DNA samples from several other Cercopithecine and Hominoid species: MHC class I DNA amplified from 10 of 10 stump-tailed macaques (Macaca arctoides) and 6 of 9 cynomolgus macaques (Macaca fascicularis) (Figs. 1 and 2, Maar and Mafa sequences). We were unable to amplify DNA using these primers from one diana monkey (Cercopithecus diana), two sooty mangabeys (Cercocebus torquatus), four baboons (Papio anubis), one orangutan (Pongo pygmaeus), and one gibbon (Hylobates lar). These data suggest that the I locus is the result of a very recent duplication event that occurred after macaques diverged from other Cercopithecine genera. Furthermore, this also suggests that Mamu-I is not orthologous to HLA-C. It is possible that the absence of amplified bands is due to a sequence difference that may exist between these species. However, it would be imprudent to rule out the possibility that the I locus is present in these species without doing extensive screening of genomic libraries. Nonetheless, given its low variability, it is unlikely that we would not have detected this locus using two different sets of primers.

FIGURE 2. Nucleotide sequence of macaque I alleles. Nucleotide sequences of rhesus, cynomolgus, and stump-tailed macaque I alleles were aligned to a macaque I consensus sequence. Mamu-A and -B sequences were included for comparison. Note that many alleles differ by non-amino acid-changing nucleotide substitutions (see Materials and Methods for further explanation of Mamu-I nomenclature). * Locus-specific nucleotides shared between Mamu-B and Mamu-I cDNAs; •, nucleotides specific to and conserved in all Mamu-I alleles. Possible regions of interlocus recombination are boxed.
Directly sequenced PCR-SSP products and sequences of full-length clones revealed the presence of at least 21 different Mamu-I variants (Fig. 2). In some cases, mixed sequences were apparent in PCR-SSP products, suggesting heterozygosity at this locus. In several cases, new alleles were identified in these heterozygous mixes, but exact sequence of these alleles could not be resolved without cloning. Therefore, these sequences were not reported.

Nonclassical MHC class I alleles and pseudogenes have a lower %GC content at the third codon positions (38). Thus, the longer the time between divergence of a locus from its classical ancestor, the lower the %GC content at the third codon position. Analysis of the %GC content of the third position of codons in the peptide binding region (PBR) and the codons of the remainder of exons 2 and 3 not involved in Ag recognition further suggested a recent origin of the I locus (Table I). Table I shows that the %GC content of the I alleles is more characteristic of classical MHC class I loci, suggesting a recent divergence from a classical MHC class I locus.

Mamu-I mRNA transcripts are detected in a wide range of tissues. Classical MHC class I mRNA transcripts are readily found in nearly all tissues. Conversely, one characteristic of some nonclassical MHC class I genes is their restricted tissue expression. Like the nonclassical gene, HLA-G, mRNA transcripts from the recently identified nonclassical Mamu-AG locus were detected at high levels in placenta and amniotic membranes but at lower levels in other tissues such as kidney, heart, and spleen. Mamu-AG mRNA transcripts were absent from peripheral blood lymphocytes and lymph nodes (39). We used RT-PCR with Mamu-I-specific primers to investigate the expression pattern of Mamu-I mRNA transcripts. Fig. 5 shows that Mamu-I mRNA transcripts were found in PBL.

**FIGURE 3.** Predicted amino acid sequence of macaque I molecules. Predicted amino acid sequence of rhesus, cynomolgus, and stump-tailed macaque I molecules were aligned to a macaque I consensus glycoprotein sequence. Mamu-A and -B glycoprotein sequences were included for comparison. Only those alleles that differ from one another in amino acid sequence are shown (see Materials and Methods for further explanation of nomenclature). , Amino acid residues of the PBR that are conserved in mammalian and avian classical alleles; , residues that are specific to and conserved among Mamu-I molecules; , N-linked glycosylation site at position 86 in the α1 domain and the conserved cysteine residues that form the disulfide bonds in the α2 and α3 domains. Possible areas of interlocus recombination are boxed.
placenta, thymus, lung, spleen, heart, kidney, liver, and small intestine. Directly sequenced PCR products from PBL1 and PBL2 confirmed that the reaction was specific for Mamu-I. Thus, unlike the restricted distribution of the nonclassical Mamu-AG mRNA, Mamu-I mRNA transcripts are found in most tissues.

Transfection of MHC class I null cells and 1-D IEF analysis demonstrates that Mamu-I protein is produced in rhesus BLCL

To ascertain whether Mamu-I protein is expressed, we used the BBM.1 mAb to stain the MHC class I null cell line, 721.221, transfected with Mamu-I*02011. We used BBM.1, which recognizes β2-microglobulin, to decrease allele to allele differences in Ab binding. Flow cytometry histograms showed a shift in fluorescence in 721-221-Mamu-I*02011 cells compared with 721 cells (Fig. 6). The level of staining that we observed for 721-Mamu-I*02011 was lower than that observed for other Mamu transfectant cell lines (Fig. 6). Whether this low surface expression is unique to the 721.221 transfectants or is also found on macaque lymphocytes remains to be shown. We also demonstrated using 1-D IEF analysis that the Mamu-I*02011 protein produced in the 721.221 transfectants has the same isoelectric point as a protein produced in the BLCL of the animal that is positive for the particular allele (data not shown). Flow cytometry histograms showed a shift in fluorescence in 221-Mamu-I*02011 cells compared with 221 cells (Fig. 6). The level of staining that we observed for 221-Mamu-I*02011 was lower than that observed for other Mamu transfectant cell lines (Fig. 6). Whether this low surface expression is unique to the 721.221 transfectants or is also found on macaque lymphocytes remains to be shown. We also demonstrated using 1-D IEF analysis that the Mamu-I*02011 protein produced in the 721.221 transfectants has the same isoelectric point as a protein produced in the BLCL of the animal that is positive for the particular allele (data not shown). Furthermore, we also translated Mamu-I*0207 in vitro and this product had the same isoelectric point as a protein produced in BLCL from the animal, 88090, from which this allele was isolated (data not shown). Taken together, these results suggest that the Mamu-I protein is produced. We have not shown cell-surface expression directly. However, given the close relationship of Mamu-I to other Mamu-B glycoproteins and based on sequence analysis, there are no striking differences that lead us to conclude that these molecules are not transported to the cell surface.

The Mamu-I locus has limited variability

One of the hallmarks of the MHC is the high polymorphism and variability of its loci. More than 149 alleles of HLA-B have now been sequenced (40), and in the 41 sequences that we analyzed for this report, there are between 1 and 49 aa differences over their entire length. The classical loci are much more polymorphic and diverse than nonclassical loci, which could be a result of their different functions in the immune system.

The macaque I locus is extremely limited in its variability compared with other classical and nonclassical loci (Figs. 2, 3, and 7 and Table II). There are 16 variable sites in exon 2, and 14 variable sites in exon 3 of I alleles from the three species of macaques. However, pairwise comparisons between the macaque I alleles show that the nucleotide variability is very low, with numbers of nucleotide differences ranging from only 0–9 in exons 2 and 3. In this same region, HLA-A and -B alleles differ by 1–51 and 0–37 nt, respectively (Fig. 7). Mamu-B alleles differ by 7–76 nt over this region. However, note that this number may be artificially high due to the fact that Mamu-B alleles probably derive from at least three different loci (H. Horton, manuscript in preparation, and Ref. 8). Pairwise comparisons of the predicted glycoprotein sequences revealed differences ranging only from 0 to 6 compared with 0–28 and 1–32 in HLA-A and -B, respectively, in the a1 and a2 domains. The predicted protein sequences of Mamu-B molecules differ by 2–46 aa in this same region. Again, this number may be misleading in that the comparison includes Mamu-B alleles from different loci.

Many of the alleles of the I locus differ by only one or two synonymous nucleotide changes. Therefore, many of the I alleles are identical at the amino acid level (Figs. 2 and 3). Moreover,
Mafa-I*1011 and Maar-I*01 are identical over the entire region of exons 2 and 3 that we sequenced. This was not due to cross-contamination of samples because there were differences in intron 2 (data not shown). However, these two alleles may differ in other exons not included in the region sequenced here. Although several groups have now identified identical alleles between species (41–43), the lack of variability of I alleles between species is striking.

The I locus has evolved under purifying selection

Classical MHC class I loci differ in their evolution from nonclassical loci. Positive selection acts on the PBR of classical loci to select for nonsynonymous changes. This is reflected by a higher rate of nonsynonymous substitution ($d_N$) (amino acid-altering substitutions) compared with the rate of synonymous substitution ($d_S$)
(non-amino acid-altering substitutions) in PBR codons of exons 2 and 3. The remaining codons in exons 2 and 3 and codons in exon 4 evolve under purifying selection which is characterized by a higher $d_S$ than $d_N$ (44). By contrast, there is no evidence for positive selection acting on the PBR of nonclassical loci (44, 45). To determine whether Mamu-I has evolved in a manner similar to classical or nonclassical loci, we compared the $d_S$ and $d_N$ in the PBR of macaque I alleles and found that evidence for positive selection in the PBR is absent (Table III). In fact, not only is $d_N$ not higher than $d_S$ in PBR codons, but $d_S$ is slightly higher than $d_N$ at these sites, though the difference is not statistically significant. However, in the remaining codons of exons 2 and 3, $d_S$ is significantly higher than $d_N$, indicating that the I locus has evolved under purifying selection. Thus, Mamu-I has evolved in a manner more characteristic of a nonclassical locus.

**Discussion**

In this paper we describe, for the first time, a duplicated B locus in the rhesus macaque that has features of both classical and nonclassical loci. The I locus is present in macaques but not in other Cercopithecine genera, suggesting that it has a recent origin. Gene trees suggest that the I locus was likely derived from the duplication of a classical MHC class I B locus. Mamu-I mRNA transcripts were detected in most tissues, and Mamu-I protein appears to be produced in lymphocytes, supporting the notion that Mamu-I is a classical MHC class I molecule. However, unlike classical MHC class I loci from humans and rhesus macaques, this locus has unusually low variability. Furthermore, similar to the nonclassical MHC class I locus, Mamu-AG, evidence for positive selection in the PBR is absent. The $d_S$ is greater than the $d_N$ in the remainder of the PBR, suggesting that this locus has evolved under purifying selection.

The I locus is similar to other MHC class I classical loci in its high polymorphism, wide tissue expression, and high %GC content at the third position of PBR codons. However, it is similar to the nonclassical Mamu-AG and HLA-E in its lack of positive selection in the PBR. Unlike the macaque I locus, the E locus is very old and so the purifying selection that has acted on it is evident from a significantly higher $d_N$ than $d_S$ in the PBR (46). The function, or lack thereof, of nonclassical MHC class I (class Ib) molecules has been greatly debated. It has been hypothesized that these genes are largely nonfunctional and are pseudogenes (38, 47). It is now known that HLA-E binds MHC class I-derived leader peptides and is the ligand for the CD94-NKG2A, B, and C receptors on NK cells. Recognition of HLA-E/leader peptide can either inhibit or trigger lysis by the NK cell, depending on the receptors involved (48–50). The function of HLA-G is still unknown but it is expressed in extravillous trophoblasts and is also an NK-inhibitory ligand. Thus, it is thought to play a role in maternal-fetal tolerance by protecting trophoblasts from NK cell killing (51–56). The recent discovery of a new nonclassical locus in the rhesus

### Table I.

The %GC content at the third codon position of the I locus is similar to that of classical MHC class I loci

<table>
<thead>
<tr>
<th>Loci</th>
<th>PBR Codons</th>
<th>Remaining Codons in Exons 2 and 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Third</td>
<td>All</td>
</tr>
<tr>
<td>I (10)</td>
<td>82.1</td>
<td>65.5</td>
</tr>
<tr>
<td>Classical (43)</td>
<td>82.3</td>
<td>60.7</td>
</tr>
<tr>
<td>Nonclassical (2)</td>
<td>77.2</td>
<td>59.0</td>
</tr>
</tbody>
</table>

*a* Shows percent GC content of 57 PBR codons and the 125 remaining codons in exons 2 and 3 for full-length rhesus monkey MHC class I sequences. The number of sequences used in each analysis is shown in parentheses.

**FIGURE 5.** Mamu-I mRNA transcripts are found in most tissues. RT-PCR on mRNA from various tissues was conducted with the I-specific primer set, MMIF2/MMIR, as described in Materials and Methods.

**FIGURE 6.** Expression of Mamu-I glycoprotein in 721.221 transfected cells compared with other Mamu transfectants. The 721.221 transfectants were stained with the BBM.1 Ab followed by a goat anti-mouse-FITC conjugated secondary Ab, and then were analyzed by flow cytometry. The transfected cell lines are indicated by the grey histograms. The .221 MHC class I-negative cell line is shown as an overlay in each panel.
monkey, *Mamu-AG*, that has similarities to *HLA-G* yet evolved independently further supports a functional role for these molecules. Thus, nonclassical loci appear to evolve from the duplication of classical loci and take on specific functions. *Mamu-I* may be at an intermediate step in this process, retaining some of its classical characteristics.

A striking feature of the MHC class I I locus is that it is highly polymorphic, and yet the numerous alleles of this locus demonstrate low variability. We isolated at least 21 *Mamu-I* alleles from 60 rhesus monkeys, yet the alleles differed from each other by only 0–9 nt. This lack of sequence variability is lower even than that observed in nonclassical loci such as *Mamu-AG* and *HLA-E* (Fig. 2 and Table II). Many more alleles of the *Mamu-I* locus have been isolated than of any other nonclassical locus. At the protein level, sequence diversity is even lower with many of the *I* alleles being identical at the amino acid level. The lack of sequence variation is most easily explained by the recent origin of this locus. This locus appears to be present in macaques but not in other *Cercopithecine* genera, suggesting that it is only 5–7 million years old (57, 58). However, it should be noted that without extensive screening of genomic libraries, we cannot say with certainty that this locus is absent in other *Cercopithecine* genera.

The MHC is of great interest not only because of the important immunologic function of the glycoproteins that it encodes but also for its unique evolutionary characteristics. Evolution of the MHC class I region is characterized by duplication and differential expansion of loci (38, 59, 60). In New World primates, the classical MHC class I Ags are more similar to the nonclassical *HLA-G* molecule than they are to any other human MHC class I molecule (61). Thus, in these primates, the ancestral *G* locus duplicated and its products functioned as the classical MHC class I molecule, presenting peptides to CTL in tamarins, whereas in humans the

---

**Table II. Mean numbers of synonymous (\(d_S\)) and nonsynonymous (\(d_N\)) nucleotide substitutions per 100 sites (with SE) in pairwise comparisons of exon 2–3 sequences at primate MHC class I loci**

<table>
<thead>
<tr>
<th>Locus</th>
<th>No. of Sequences Compared</th>
<th>(d_S)</th>
<th>(d_N)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Mamu-I</em></td>
<td>10</td>
<td>2.1 (0.8)</td>
<td>0.6 (0.2)</td>
</tr>
<tr>
<td><em>Mamu-A</em></td>
<td>22</td>
<td>7.3 (1.1)**</td>
<td>8.3 (0.6)**</td>
</tr>
<tr>
<td><em>Mamu-B</em></td>
<td>20</td>
<td>9.3 (1.2)**</td>
<td>10.5 (0.7)**</td>
</tr>
<tr>
<td><em>Mamu-AG</em></td>
<td>7</td>
<td>2.7 (0.9)</td>
<td>1.4 (0.4)</td>
</tr>
<tr>
<td>HLA-A</td>
<td>24</td>
<td>3.5 (0.9)</td>
<td>4.8 (0.5)**</td>
</tr>
<tr>
<td>HLA-B</td>
<td>40</td>
<td>5.0 (0.9)*</td>
<td>5.4 (0.5)**</td>
</tr>
</tbody>
</table>

* Tests of the hypothesis that \(d_S\) or \(d_N\) equals the corresponding value for *Mamu-I*:
  * \(p < 0.05\);
  ** \(p < 0.001\).

---

**Figure 7.** The MHC class I I locus has limited variability. Frequency distribution of numbers of pairwise nucleotide differences between rhesus, cynomolgus, and stump-tailed macaque *I* alleles are compared with rhesus and human classical and nonclassical MHC class I loci.
HLA-G molecule appears to have evolved to serve a much more specialized function. In some cases, after duplication, a locus may accumulate deleterious mutations and degenerate into a pseudogene. For example, HLA-H originated as a duplication of the A locus but is now a pseudogene (62). Therefore, the fate of a locus after duplication depends on the selective forces acting on it. Based on the analysis of synonymous and non-synonymous substitutions, evidence for positive selection in the PBR of Mamu-I is absent. One interpretation of these data is that the PBR of Mamu-I has evolved under neutral selection since its divergence from the B locus. However, an alternative hypothesis is that the PBR of Mamu-I has undergone purifying selection and, due to the short time since its divergence, \( d_\delta \) is not significantly higher than \( d_\delta \) in this region.

The rhesus monkey has proved to be one of the most important animal models for the study of AIDS and transplantation (1–6). The recent development of the MHC tetramer technology has led to new insight regarding T cell responses to HIV in humans and SIV in the rhesus macaque (63, 64). This technology has also led to the elucidation of the ligand for the nonclassical MHC class I molecule HLA-E (48). However, to utilize this new tetramer technology to further understand immune responses in the rhesus monkey, knowledge of the MHC class I molecules expressed by this species is critical. It is now becoming increasingly clear that the MHC class I locus of the rhesus macaque is much more complex than its human counterpart. Orthologs of the human HLA-A, -B, -C, -E, -F, and -G loci have now been identified in the rhesus monkey (8, 17–19). However, there is also evidence to suggest that the Mamu-A and Mamu-B loci have duplicated at least once (8). Furthermore, the nonclassical Mamu-AG locus is likely the result of a duplication of an A locus (45). In this paper we have described a new MHC class I locus in the rhesus macaque, Mamu-I, that shares features with both classical and nonclassical loci. These results have implications for the further development of the rhesus macaque animal model and for understanding cellular immune responses in this species.

Acknowledgments

We thank Dr. Paul Johnson for providing the sooty mangabey BLCLs, the Milwaukee Zoological Society (Milwaukee, WI) for the diana monkey blood sample, Dr Prabhat Sehgal (New England Regional Primate Research Center) for cynomolgus monkey samples, and Drs. Linda Klavins (Guy’s Hospital, London, U.K.) and Jeff Rowell (New Iberia Research Center, University of Southwestern, New Iberia, LA) for providing rhesus monkey samples. We also thank Jon Boyson for helpful discussions and review of this manuscript.

References


Table III. Mean numbers of nucleotide substitutions per 100 synonymous sites (\( d_\delta \)) and per 100 nonsynonymous sites (\( d_\delta \))

<table>
<thead>
<tr>
<th>Locus</th>
<th>PBR</th>
<th>Remainder of ( a_1 ) and ( a_2 )</th>
<th>Remainder of Coding Region (Exons 4–8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macaque ( P )</td>
<td>( 1.08 \pm 0.70 )</td>
<td>( 0.54 \pm 0.25 )</td>
<td>( 0.46 \pm 0.24^* )</td>
</tr>
<tr>
<td>Mamu-A</td>
<td>( 8.48 \pm 2.21 )</td>
<td>( 21.77 \pm 2.00^\dagger )</td>
<td>( 2.99 \pm 0.49^\dagger )</td>
</tr>
<tr>
<td>Mamu-B</td>
<td>( 11.33 \pm 2.60 )</td>
<td>( 24.19 \pm 1.84^\dagger )</td>
<td>( 5.12 \pm 0.60^* )</td>
</tr>
<tr>
<td>Mamu-AG</td>
<td>( 3.4 \pm 1.8 )</td>
<td>( 2.1 \pm 0.8 )</td>
<td>( 1.5 \pm 0.5 )</td>
</tr>
<tr>
<td>HLA-C</td>
<td>( 1.60 \pm 0.96 )</td>
<td>( 6.30 \pm 1.02^\dagger )</td>
<td>( 1.74 \pm 0.37^* )</td>
</tr>
</tbody>
</table>

\( ^* \) Only full-length sequences (10 sequences) were used in the remainder of the \( a_1 \) and \( a_2 \) and exons 4–8 analyses.

\( ^\dagger \) Only exon 4 was analyzed for Mamu-AG (see Ref. 45).

\( ^p \) \( d_\delta \) and \( d_\delta \) are significantly different.

\( ^\dagger \) \( p < 0.001 \), \( d_\delta \) and \( d_\delta \) are significantly different.

\( ^\dagger \) \( p < 0.01 \), \( d_\delta \) and \( d_\delta \) are significantly different.


