



Published in final edited form as:

Tissue Antigens. 2008 July ; 72(1): 29–38. doi:10.1111/j.1399-0039.2008.01057.x.

Allelic diversity within the high frequency *Mamu-A2*05/Mane-A2*05 (Mane-A*06)/Mafa-A2*05* family of macaque *MHC-A* loci

Jin Wu, Sue Bassinger, George D. Montoya, Leonard Chavez, Carrie E. Jones, Brigitte Holder-Lockyer, Barbara Masten, Thomas M. Williams, and Kiley R. Prilliman*

Department of Pathology, University of New Mexico School of Medicine, Albuquerque, NM, USA

Abstract

Macaque species serve as important animal models of human infection and immunity. To more fully scrutinize their potential in both the analysis of disease pathogenesis and vaccine development, it is necessary to characterize the Major Histocompatibility Complex (MHC) class I loci of *Macaca mulatta* (*Mamu*), *Macaca nemestrina* (*Mane*), and *Macaca fascicularis* (*Mafa*) at the genomic level. The oligomorphic *Mamu-A2*05/Mane-A2*05* (previously known as *Mane-A*06*) family of macaque *MHC-A* alleles has recently been shown to be present at high frequency in both Indian rhesus and pig-tailed macaque populations. Using a locus-specific amplification and direct DNA typing methodology, we have additionally found that the locus encoding this family is very prevalent (75%) among a sampling of 182 Chinese rhesus macaques and has a high prevalence (80%) within a larger, independent cohort of 309 pig-tailed macaques. Interestingly, among the Chinese rhesus macaques, only 6 alleles previously identified in Indian-origin animals were observed, while 3 recently identified in Chinese-origin animals and 25 new alleles were characterized. Among the pig-tailed macaques, we observed one previously known (*Mane-A*06*) and 19 new alleles. Examination of the orthologous locus in a preliminary sampling of 30 cynomolgus macaques revealed an even higher presence (87%) of *Mafa-A2*05* family alleles, with 5 previously identified and 15 new ones characterized. The continued discovery of novel alleles and thus further diversity within the *Mamu-A2*05/Mane-A2*05/Mafa-A2*05* family indicates that this *MHC-A* locus, although highly conserved across the three species of macaques, has remained a dynamic entity during evolution.

Keywords

Macaque; MHC-A; alleles

INTRODUCTION

In order to effectively use non-human primates such as macaques as research models in immunological studies of human disease and transplantation, it is important to analyze their Major Histocompatibility Complex (MHC) genetics as thoroughly as possible. Fully

Correspondence Thomas M. Williams, Department of Pathology, University of New Mexico School of Medicine, MSC08 4640, 1 University of New Mexico, Albuquerque, NM 87131, USA, e-mail: TWilliams@salud.unm.edu.

*Current address: Receptor Logic, Ltd., 1718 Pine Street, Abilene, TX, 79601.

characterizing the genes of the MHC will allow for a better understanding of the cellular immune responses of these animals, which underlie their capacities to respond to pathogens as well as to moderate tissue engraftment. The MHC of macaques, for which the most is currently known through studies employing Indian-origin rhesus macaques (*Macaca mulatta*) (1, 2), is comparatively more complicated than that of humans. While humans express molecules from three different classical MHC class I loci, macaques express *MHC-A* (*Mamu-A*) and *MHC-B* (*Mamu-B*) molecules but bear no evidence of an *MHC-C* locus (3). These *A* and *B* genes are furthermore known to be duplicated, with approximately 17 distinct class I genes (two *Mamu-A* and up to 15 *Mamu-B*) observed per haplotype when the full MHC regions of two rhesus macaques were independently sequenced (4, 5). Additional studies have indicated via haplotype analysis that Indian rhesus macaques have two to three *MHC-A* loci per haplotype (6,7,8). The implications of possessing so many MHC loci upon the overall immunological fitness of macaques remain unclear at this time (4, 9).

In addition to the rhesus macaque, two other species used extensively in biomedical research include the cynomolgus (crab-eating or long-tailed) macaque (*Macaca fascicularis*) and the pig-tailed macaque (*Macaca nemestrina*). Whereas apes and humans belong to the Hominoidea superfamily of Old World primates, the rhesus, cynomolgus, and pig-tailed macaques belong to the Cercopithecoidea superfamily. Pig-tailed macaques within the *Silensus* group diverged from the *Fascicularis* group, which includes rhesus and cynomolgus macaques, nearly 5 million years ago, while rhesus and cynomolgus macaques separated from one another 2.4 million years later (10,11). Though it has been demonstrated that both the *MHC-A* and *MHC-B* loci are duplicated in pig-tailed and cynomolgus macaques and that neither has a functional *MHC-C* locus (12–16), considerably more is known about the genomic organization of rhesus and cynomolgus macaques than for pig-tailed macaques. Microsatellite analysis of a feral population of Mauritian animals and BAC library-based contig mapping of the cynomolgus macaque MHC region have established an arrangement not unlike that of rhesus macaques (17,18), though lacking in as many *MHC-B* loci. Analysis thus far indicates that cynomolgus macaques have one to three distinct *MHC-A* loci per haplotype (7). In contrast, little is currently known of the genomic organization or loci of the pig-tailed macaque MHC other than that the gene which expresses the *Mane-A*06* allele is an orthologue, or common ancestral gene present among different species, of the locus encoding the *Mamu-A2*05* allele family in rhesus macaques and that this gene also exists in cynomolgus macaques (9).

The identification of this orthologue suggests that the locus encoding the oligomorphic *Mamu-A2*05/Mane-A2*05 (Mane-A*06)/Mafa-A2*05* family of *MHC-A* alleles in these animals was present in their common ancestor. Considering the complexity of the macaque MHC as briefly outlined above in conjunction with the fact that rhesus, pig-tailed, and cynomolgus macaques all serve as animal models for studying HIV pathogenesis and vaccine development, we believe that a rational approach to comparatively understanding their MHC genetics is to first select the most highly conserved elements, such as this recently identified orthologue, and begin analyzing them in detail at the genomic level. To this end, we have developed a locus-specific typing strategy for characterizing this locus and implemented it in typing Chinese-origin rhesus macaques (a potential alternative model to

the Indian-origin rhesus macaque), pig-tailed macaques, and cynomolgus macaques. A total of 74 different alleles were characterized at the *A2*05* locus in this study; 59 of them are novel. Our results indicate that, despite being highly conserved across all three species examined, there is still marked polymorphism present in this region that is concentrated throughout the peptide-binding domains of the molecules.

MATERIALS AND METHODS

Animal specimens

Blood samples from Chinese rhesus macaques (*Macaca mulatta*) were obtained from 62 animals maintained at the Lovelace Respiratory Research Institute (Albuquerque, NM) and 120 animals at Alpha Genesis Incorporated (Yemassee, SC). Blood samples from pig-tailed macaques (*Macaca nemestrina*) were obtained from 309 animals maintained at the Tulane Regional Primate Research Center (Covington, LA). Blood samples from Vietnamese and Indonesian cynomolgus macaques (*Macaca fascicularis*) were obtained from 30 animals at the Lovelace Respiratory Research Institute. All animals were housed and treated in accordance with Institutional Animal Care and Use Committee guidelines. Fresh blood (5 ml) was collected from animals in ethylenediamine tetraacetic acid (EDTA) for subsequent isolation of genomic DNA.

Genomic DNA isolation

Genomic DNA was extracted from peripheral blood leukocytes by salting out using Gentra's Puregene™ DNA isolation reagents (Minneapolis, MN). EDTA-treated whole blood was centrifuged 6 min at 3,000 rpm. The buffy coat was then collected, subjected to red blood cell lysis, and centrifuged for 6 minutes at 3,000 rpm. The recovered leukocytes were then lysed with 1–2 mL of Cell Lysis solution. 0.5ml of cell lysis was added to protein precipitation solution and spun. The resulting supernatant containing genomic DNA was drawn off and precipitated in 0.5 mL of isopropanol. The DNA pellet was then washed in 0.5 mL of 70% ethanol and dried in a 65°C heat block for 6 min; 100 uL of DNA Hydration solution was then added to the dried pellet and incubated for 1 h at 65°C. The DNA was checked for purity and concentration on a NanoDrop 1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

Locus-specific amplification and sequencing

We employed a strategy based on the locus-specific amplification of genomic DNA followed by bidirectional sequencing of exons 2, 3, and 4 with intronic primers. Primers P000044 and P000023 (Table 1) were designed to specifically amplify the *A2* locus of rhesus, pig-tailed, and cynomolgus macaques. Genomic DNA (0.6 µg) was amplified in 100 µL reaction volumes containing 1.75 mM MgCl₂, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 200 µM each dNTP, 0.25 µM each primer, and 2.5 U Taq Gold Polymerase. The PCR reaction mixture was heated in an MJ Research Tetrad 2 thermal cycler (Bio-Rad Laboratories, Hercules, CA) at 95°C for 10 min and followed by 35 cycles of 20 sec at 95°C, 30 sec at 62°C, and 3 min at 72°C. A final extension was carried out for 10 min at 72°C. Direct sequencing of exons 2, 3, and 4 was performed with intronic primers (Table 1) and the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City,

CA). Sequences were then obtained on an Applied Biosystems 3130XL genetic analyzer and assembled with Assign software (Conexio Genomics, Perth, Australia) using published and in-house primate allele sequences as libraries.

Subcloning and sequencing

TA cloning and sequencing were performed for heterozygous sequences that involved novel alleles. Heterozygous PCR products were subjected to subcloning in the DH5 α strain of *E. coli* after ligation into the pCR 2.1 TOPO vector (Invitrogen, Carlsbad, CA). Plasmids were isolated with the QIAprep Spin Miniprep kit (QIAGEN Inc., Valencia, CA). Sequencing reactions for exons 2, 3, and 4 were performed and analyzed and the sequences assembled as described above. Novel allele sequences were determined by at least two independent identical clones.

Phylogenetic analyses

Nucleotide and amino acid sequences were aligned using MacClade 4 software (Sinauer Associates, Inc., Sunderland, MA), and phylogenetic trees were generated with PHYLIP 3.66 software. Consensus neighbor-joining trees (19) were constructed using Kimura's two-parameter distance matrices (20) and bootstrapping (1,000 replicates) to assign confidence to tree nodes (21). Values higher than 50% are indicated.

Nomenclature

Sequences were submitted first to the GenBank database for assignment of accession numbers and then to the IMGT/MHC-NHP (non-human primate) database for naming of new alleles according to the nomenclature proposal guidelines currently in use (22,23). As previously described (7), lineage is indicated by the first two digits after the asterisk, and the allele number is indicated by the third and fourth digits; the fifth and sixth digits indicate synonymous basepair differences between otherwise identical sequences. Alleles detected in this study are summarized in Tables 2, 3, and 4.

RESULTS AND DISCUSSION

Sequencing of *Mamu-A2*05/Mane-A2*05/Mafa-A2*05* family alleles

The relationship of *Mamu-A2*05/Mane-A2*05* as common ancestral genes, or orthologues, has been recently described in Indian-origin rhesus and pig-tailed macaques (9). Our objective in this study was to examine the MHC in a locus-specific manner at the genomic level to more fully characterize this locus among Chinese-origin rhesus, pig-tailed, and cynomolgus macaques. This approach provides important information about the composition of the MHC and includes discovery of alleles potentially expressed at all transcriptional levels, in contrast to the majority of other studies to date, which have relied largely upon the examination of cDNAs. Given that *A2*05* alleles are known to be expressed by rhesus and pig-tailed macaques *in vitro*, albeit at low levels, our approach was based upon genomic amplification primers (Table 1) that were developed in accordance with published sequence information for the MHC region of an Indian-origin rhesus macaque bearing the *A2* gene locus (5). We directly sequence exons 2–4 encoding the extracellular domains of the molecules.

The locus encoding A2*05 alleles was present in a majority, but not all, of the samples tested, confirming that this particular locus is not always present on the haplotypes carried by a particular animal. Specifically, 75% of the 182 Chinese rhesus macaques had at least one A2 allele, as did 80% of 309 pig-tailed macaques and 87% of 30 cynomolgus macaques. We employed a single set of primers for amplification in this study, so we can not be certain that animals lacking A2 alleles did not have unknown polymorphisms within the primer annealing sites. The orthologous relationship of *Mamu-A2*05* with *Mane-A2*05* alleles was previously demonstrated through phylogenetic analysis (9). Accordingly, alleles described here from both of these groups were likewise observed to cluster together, regardless of species (Figure 1 and data not shown). Not surprisingly, we additionally found this to be true for alleles of the *Mafa-A2*05* family. As shown in the phylogenetic tree for 73 of the 74 alleles detected in this study (*Mafa-A2*0528* bears an 8 nucleotide deletion at position 413 of exon 2 that causes a frameshift and was therefore excluded from this analysis), alleles from the three species are freely intermingled. Bootstrap values were generally low due to the shallow phylogeny involved in examining all of the closely related alleles simultaneously. Among the 20 *Mane* alleles, despite the significant evolutionary distance of pig-tailed from rhesus and cynomolgus macaques, close relationships were observed for these alleles in the context of *Mamu* and *Mafa* genes, with the greatest differences being a 7 nucleotide mismatch between *Mane-A2*0509* and *Mamu-A2*0512* and an 8 nucleotide mismatch between *Mane-A2*0513* and *Mafa-A2*0530* (Table 5). In fact, 5 sub-clusters of identical interspecies alleles were identified as follows (Figure 1 and Table 5): *Mane-A2*050101* and *Mafa-A2*0504*; *Mane-A2*050501* and *Mafa-A2*0507*; *Mane-A2*050502*, *Mamu-A2*050403*, and *Mafa-A2*050602*; *Mane-A2*0514* and *Mamu-A2*050301*; and *Mane-A*06* and *Mamu-A2*0522*. This is particularly striking given that the pig-tailed macaques have undergone 2.4 million years of evolution apart from the rhesus and cynomolgus macaques and additionally questions the notion that sharing of MHC alleles among primate species is a rare event (7).

Location and nature of *Mamu-A2*05/Mane-A2*05/Mafa-A2*05* polymorphisms

All of the proteins predicted to be expressed by the alleles described here are fully intact MHC proteins except for two; the majority of distinguishing substitutions are localized to the peptide binding domains, alpha 1 and alpha 2 (Figure 2). We found little diversity among alleles in the exon 4 sequences (data not shown). As mentioned previously, *Mafa-A2*0528* bears a frameshift deletion, and *Mane-A2*0507* is expected to be expressed as a truncated protein due to a stop at codon 75. The *Mamu* translated proteins exhibit 31 polymorphic positions throughout the extracellular portions of the molecule, 24 of which are located in the alpha 1 and alpha 2 domains. Of these, 12 are among the peptide binding pocket residues designated to environmentally interact with bound ligands (24). The *Mane* translated proteins are polymorphic at 19 extracellular positions, 13 of which are located in the alpha 1 and alpha 2 domains; 8 of these are capable of interacting with bound ligands. The *Mafa* translated proteins are polymorphic at 27 extracellular positions, 21 of which are located in the alpha 1 and alpha 2 domains; 11 can interact with bound ligands.

The proteins encoded by *Mamu-A2*05/Mane-A2*05/Mafa-A2*05* family alleles are characterized by polymorphic site substitutions of 6 categories, as highlighted in Figure 3:

those observed in *Mafa* only; *Mane* only; *Mamu* only; *Mane* and *Mafa*; *Mamu* and *Mafa*; or in all three species. No sites of substitution are commonly observed between pig-tailed and rhesus macaques alone. The sites of polymorphism shared by all three species encompass residues within the A and B (positions 62, 63, 163, 167, 171) as well as F (position 95) peptide binding pockets, in addition to a site on an external connecting loop (position 91). While the substitution at position 95 of leucine for isoleucine is fairly conservative, the additional substitutions observed at the other pocket positions are of higher physicochemical impact and likely influence the N-terminal preferences of peptides bound from one allele to another. The impact of the arginine for glycine substitution at position 91 is intriguing but unknown at this point; we hypothesize that it could differentially affect interactions with various immune effector cells and their receptors.

The *Mane* specific polymorphisms involve positions 75 and 76 in the truncated protein (*Mane-A2*0507*), a conservative glutamate for aspartate substitution at C, D, E, and F pocket position 114 (*Mane-A2*0511*), and positions 106 and 107 on an external connecting loop (*Mane-A2*0515* and *Mane-A2*0502*, respectively). *Mane* and *Mafa* share a site of polymorphism at position 5 with a conservative methionine for leucine substitution in the A pocket. The *Mamu* specific polymorphisms involve numerous positions throughout the alpha 1 helix, including among them 67 and 70 which contribute to the B pocket (*Mane-A2*0526*), as well as connecting loop positions 89 and 121 and positions 134, 138, and 158 within the alpha 2 domain. *Mamu* and *Mafa* share sites of polymorphism at various positions, most notably including alleles with predominantly non-conservative substitutions at 151, 152, and 156 (the latter two of which interact directly with the C, D, and E peptide binding pockets). With the exception of a non-conservative aspartate for alanine substitution at position 69 in the C pocket (*Mafa-A2*0518*), the *Mafa* specific polymorphisms are isolated to the alpha 2 domain.

Mamu-A2*05/Mane-A2*05/Mafa-A2*05 diversity

Greater variation both in terms of percentages of new alleles discovered per number of animals sampled and positions of polymorphism were observed for Chinese rhesus and cynomolgus macaques than for pig-tailed macaques. For example, 20 distinct A2 alleles were detected in 30 cynomolgus animals and 20 were found in 309 tested pig-tailed macaques, suggesting that there may be less extensive diversity present at this locus in the latter species. Among Chinese-origin rhesus macaques, we observed little allelic similarity with Indian-origin animals from which the majority of *Mamu-A2*05* alleles have previously been described. This low degree of similarity was likewise seen between our Chinese macaque cohort and that of another group (7), with only 3 alleles (*Mamu-A2*050301*, *Mamu-A2*050402*, and *Mamu-A2*0526*) in common. We also noticed variation in allelic distributions and frequencies between the two sources from which we received our Chinese rhesus macaques (data not shown). This indicates that preliminary typing of animals from a chosen distributor will be mandatory prior to selecting Chinese-origin animals as experimental models heavily dependent upon defined MHC backgrounds.

Acknowledgments

This research was supported by NIH contract N01-AI-40087. We would like to acknowledge the support for DNA sequence analysis provided to this project via Assign software from Conexio Genomics of Perth, Australia.

References

1. Adams EJ, Parham P. Species-specific evolution of MHC class I genes in the higher primates. *Immunol Rev.* 2001; 183:41–64. [PubMed: 11782246]
2. Bontrop RE. Comparative genetics of MHC polymorphisms in different primate species: duplications and deletions. *Hum Immunol.* 2006; 67:388–97. [PubMed: 16728259]
3. Boyson JE, Shufflebotham C, Cadavid LF. The MHC class I genes of the rhesus monkey. Different evolutionary histories of MHC class I and II genes in primates. *J Immunol.* 1996; 156:4656–65. [PubMed: 8648109]
4. Daza-Vamenta R, Glusman G, Rowen L, Guthrie B, Geraghty DE. Genetic divergence of the rhesus macaque major histocompatibility complex. *Genome Res.* 2004; 14:1501–15. [PubMed: 15289473]
5. Kulski JK, Anzai T, Shiina T, Inoko H. Rhesus macaque class I duplicon structures, organization, and evolution within the alpha block of the major histocompatibility complex. *Mol Biol Evol.* 2004; 21:2079–91. [PubMed: 15269276]
6. Otting N, Heijmans CM, Noort RC. Unparalleled complexity of the MHC class I region in rhesus macaques. *Proc Natl Acad Sci U S A.* 2005; 102:1626–31. [PubMed: 15665097]
7. Otting NA, de Vos-Rouweler AJ, Heijmans CM, de Groot NG, Doxiadis GG, Bontrop RE. MHC class I A region diversity and polymorphism in macaque species. *Immunogenetics.* 2007; 59:367–75. [PubMed: 17334754]
8. Sauermann U, Siddiqui R, Suh Y-S, et al. Mhc class I haplotypes associated with survival time in simian immunodeficiency virus (SIV)-infected rhesus macaques. *Genes Immun.* 2008; 9:69–80. [PubMed: 18094710]
9. Lafont BA, McGraw CM, Stukes SA. The locus encoding an oligomorphic family of MHC-A alleles (Mane-A*06/Mamu-A*05) is present at high frequency in several macaque species. *Immunogenetics.* 2007; 59:211–23. [PubMed: 17256149]
10. Morales JC, Melnick DJ. Phylogenetic relationships of the macaques (Cercopithecidae: Macaca), as revealed by high resolution restriction site mapping of mitochondrial ribosomal genes. *J Hum Evol.* 1998; 34:1–23. [PubMed: 9467779]
11. Tosi AJ, Morales JC, Melnick DJ. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. *Mol Phylogenet Evol.* 2000; 17:133–44. [PubMed: 11083929]
12. Lafont BA, Buckler-White A, Plishka R, Buckler C, Martin MA. Characterization of pig-tailed macaque classical MHC class I genes: implications for MHC evolution and antigen presentation in macaques. *J Immunol.* 2003; 171:875–85. [PubMed: 12847257]
13. Smith MZ, Dale CJ, De Rose R, et al. Analysis of pigtail macaque major histocompatibility complex class I molecules presenting immunodominant simian immunodeficiency virus epitopes. *J Virol.* 2005; 79:684–95. [PubMed: 15613296]
14. Krebs KC, Jin Z, Rudersdorf R, Hughes AL, O'Connor DH. Unusually high frequency MHC class I alleles in Mauritian origin cynomolgus macaques. *J Immunol.* 2005; 175:5230–9. [PubMed: 16210628]
15. Uda A, Tanabayashi K, Yamada YK, et al. Detection of 14 alleles derived from the MHC class I A locus in cynomolgus monkeys. *Immunogenetics.* 2004; 56:155–63. [PubMed: 15164234]
16. Uda A, Tanabayashi K, Fujita O, Hotta A, Terao K, Yamada A. Identification of the MHC class I B locus in cynomolgus monkeys. *Immunogenetics.* 2005; 57:189–97. [PubMed: 15900490]
17. Wiseman RW, Wojcechowskyj JA, Greene JM, et al. Simian immunodeficiency virus SIVmac239 infection of major histocompatibility complex-identical cynomolgus macaques from Mauritius. *J Virol.* 2007; 81:349–61. [PubMed: 17035320]

18. Watanabe A, Shiina T, Shimizu S, et al. A BAC-based contig map of the cynomolgus macaque (*Macaca fascicularis*) major histocompatibility complex genomic region. *Genomics*. 2007; 89:402–12. [PubMed: 17174065]
19. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987; 4:406–25. [PubMed: 3447015]
20. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980; 16:111–20. [PubMed: 7463489]
21. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985; 39:783–791.
22. Klein J, Bontrop RE, Dawkins RL, et al. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics*. 1990; 31:217–9. [PubMed: 2329006]
23. Robinson JM, Waller J, Parham P, et al. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res*. 2003; 31:311–4. [PubMed: 12520010]
24. Chelvanayagam G. A roadmap for HLA-A, HLA-B, and HLA-C peptide binding specificities. *Immunogenetics*. 1996; 45:15–26. [PubMed: 8881033]
25. Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J Mol Biol*. 1991; 219:277–319. [PubMed: 2038058]

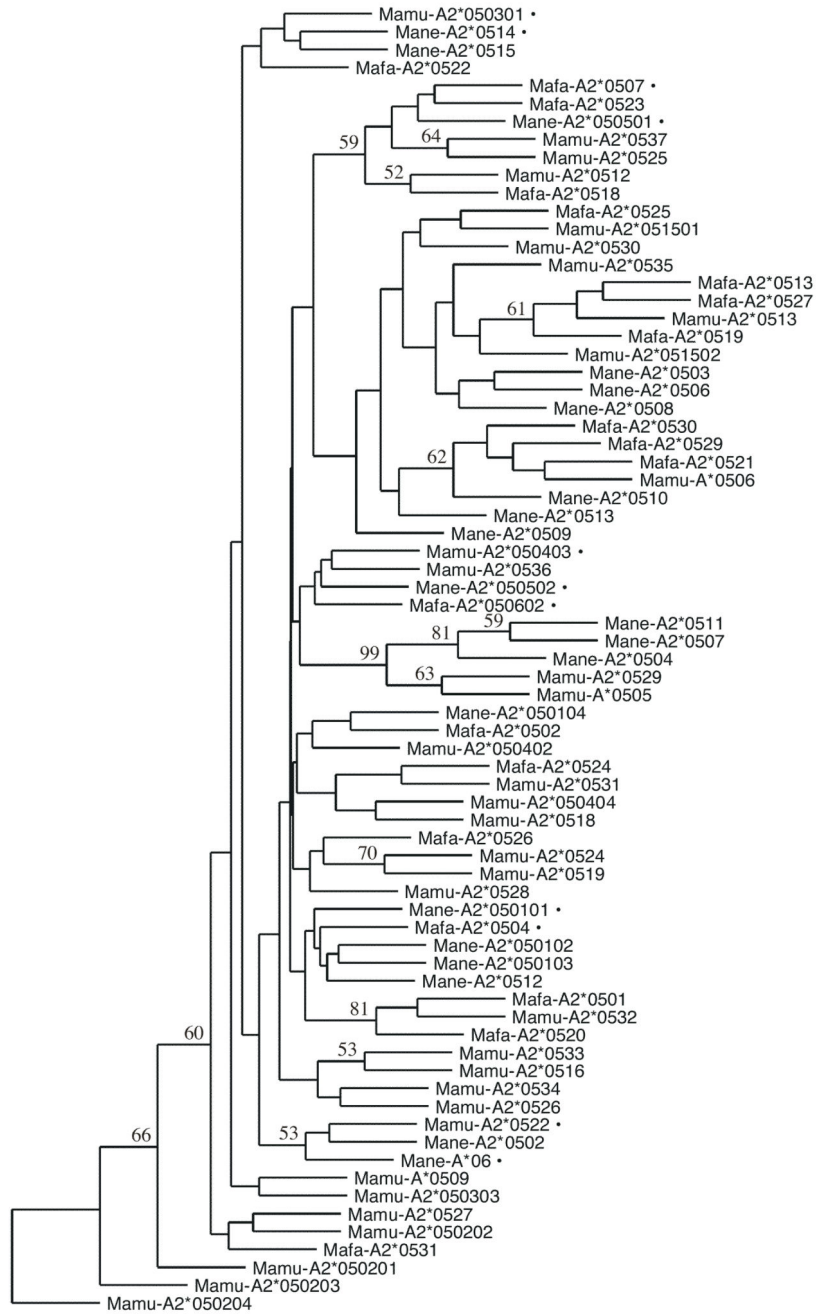


Figure 1. Consensus neighbor-joining tree of 73 *Mamu-A2*05*, *Mane-A2*05*, and *Mafa-A2*05* alleles detected in this study. This phylogenetic tree was drawn from the sequences of exons 2, 3, and 4 using a Kimura's two-parameter distance matrix after ascertaining that the *A2*05* family members grouped together on a more extensive tree (not shown) which contained additional *Mamu*, *Mane*, and *Mafa* *MHC-A* and *MHC-B* alleles. Bootstrap values greater than 50% for 1,000 replicates are indicated. Vertical lines to the right indicate clusters of

alleles that are identical between different species, which are specifically marked within these regions by bullets and listed in Table 5.

Figure 2a

Alpha 1

	10	20	30	40	50	60	70	80	90
Consensus	-SHSLRYFYT	SMSRPGRGQP	RFISVGYVDD	TQFVRFSDA	ESPREEPRAP	WMEQEGPEYW	DQNTTRICKAD	TQTLRENLRT	LLRYNQSEA
Mamu-A*0502						V			
Mamu-A2*050202						V			
Mamu-A2*050203						V			
Mamu-A2*050204						V			
Mamu-A*0509						V			
Mamu-A2*0516						V			
Mamu-A2*0533						V			
Mamu-A2*0527						V			
Mamu-A2*050301									
Mamu-A2*050303									
Mamu-A2*0522									
Mamu-A2*0532		TV							
Mamu-A2*0528									
Mamu-A2*050402									
Mamu-A2*050403									
Mamu-A2*050404									
Mamu-A2*0536						Q			
Mamu-A*0505									
Mamu-A2*0529									
Mamu-A2*0531									
Mamu-A2*0525									D
Mamu-A2*0537									
Mamu-A2*0518									
Mamu-A2*0512									
Mamu-A*0506							RE		
Mamu-A2*0535							RE		
Mamu-A2*051502							RE		
Mamu-A2*051501							RE		
Mamu-A2*0513		S					RE		
Mamu-A2*0530							RE		
Mamu-A2*0519									
Mamu-A2*0524									
Mamu-A2*0534									
Mamu-A2*0526							A	N	A
Consensus	-SHSLRYFYT	SMSRPGRGQP	RFISVGYVDD	TQFVRFSDA	ESPREEPRAP	WMEQEGPEYW	DQNTTRICKAD	TQTLRENLRT	LLRYNQSEA

Alpha 2

	100	110	120	130	140	150	160	170	180
Consensus	GSHTIQTMYG	CDLGPDRLL	RGYDQSAIDG	RDYIALNEDI	RSWTAADMAA	QNTQRKWEAA	GVAEQWRAYL	EGECLEWLRR	HLENGKETLQ RA
Mamu-A*0502	L							Q	S
Mamu-A2*050202	L							Q	S
Mamu-A2*050203	L							Q	S
Mamu-A2*050204	L							Q	S
Mamu-A*0509	L							Q	S
Mamu-A2*0516	L							Q	S
Mamu-A2*0533	L						RA	H	T
Mamu-A2*0527	L							Q	S
Mamu-A2*050301	L							Q	S
Mamu-A2*050303	L							Q	S
Mamu-A2*0522	L							S	
Mamu-A2*0532								S	
Mamu-A2*0528								S	
Mamu-A2*050402									
Mamu-A2*050403									
Mamu-A2*050404									
Mamu-A2*0536									
Mamu-A*0505						T			Y
Mamu-A2*0529						T			
Mamu-A2*0531									
Mamu-A2*0525									
Mamu-A2*0537									
Mamu-A2*0518									
Mamu-A2*0512									
Mamu-A*0506									
Mamu-A2*0535									
Mamu-A2*051502	R								
Mamu-A2*051501	R								
Mamu-A2*0513	R								
Mamu-A2*0530	R							S	
Mamu-A2*0519									
Mamu-A2*0524								T	
Mamu-A2*0534							RA	H	T
Mamu-A2*0526				K			A	H	T
Consensus	GSHTIQTMYG	CDLGPDRLL	RGYDQSAIDG	RDYIALNEDI	RSWTAADMAA	QNTQRKWEAA	GVAEQWRAYL	EGECLEWLRR	HLENGKETLQ RA

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Figure 2b

Alpha 1	10	20	30	40	50	60	70	80	90
Consensus	XSHSLRYFYT	SMSRPGRGQP	RFISVGYVDD	TQFVRFDSDA	ESPREEPRAP	WMEQEGPEYW	DQNTTRICKAD	TQTLRENLRT	LLRYYNQSEA
Mane-A2*0514	-	-	-	-	-	-	-	-	-
Mane-A2*0515	-	-	-	-	-	-	-	-	-
Mane-A2*0502	-	-	-	-	-	-	-	-	-
Mane-A*06	G	-	-	-	-	-	-	-	-
Mane-A2*0512	-	-	-	-	-	-	-	-	-
Mane-A2*050104	-	-	-	-	-	-	-	-	-
Mane-A2*050101	-	-	-	-	-	-	-	-	-
Mane-A2*050103	-	-	-	-	-	-	-	-	-
Mane-A2*050102	-	-	-	-	-	-	-	-	-
Mane-A2*0504	-	-	-	-	-	-	-	-	-
Mane-A2*0507	-	-	-	-	-	-	-	-	-
Mane-A2*0511	-	-	-	-	-	-	-	-	-
Mane-A2*0505	-	-	-	-	-	-	-	-	-
Mane-A2*050502	-	-	-	-	-	-	-	-	-
Mane-A2*0503	-	-	-	-	-	-	.RE	-	-
Mane-A2*0508	-	-	-	-	-	-	.RE	-	-
Mane-A2*0506	-	-	-	-	-	-	.RE	-	-
Mane-A2*0509	-	-	-	-	-	-	.RE	-	-
Mane-A2*0510	-	-	-	-	-	-	.RE	-	-
Mane-A2*0513	-	.M	-	-	-	-	.RE	-	-
Consensus	XSHSLRYFYT	SMSRPGRGQP	RFISVGYVDD	TQFVRFDSDA	ESPREEPRAP	WMEQEGPEYW	DQNTTRICKAD	TQTLRENLRT	LLRYYNQSEA

Alpha 2	100	110	120	130	140	150	160	170	180
Consensus	GSHTIQTMYG	CDLGPDGRLL	RGYDQSAYDG	RDYIALNEDL	RSWTAADMAA	QNTQRKWEAA	GVAEQWRAYL	EGECLEXLRR	HLENGKETLQ RA
Mane-A2*0514	. . . L Q . . S . . Y
Mane-A2*0515	. . . L E Q . . S . . Y
Mane-A2*0502	. . . L R S Y
Mane-A*06	. . . L S Y
Mane-A2*0512 S
Mane-A2*050104 S
Mane-A2*050101 S
Mane-A2*050103 S
Mane-A2*050102 S
Mane-A2*0504 W
Mane-A2*0507 W
Mane-A2*0511 E W
Mane-A2*0505 W
Mane-A2*050502 W
Mane-A2*0503	R S . . . Y
Mane-A2*0508	R S
Mane-A2*0506 S . . . Y
Mane-A2*0509 W
Mane-A2*0510 W
Mane-A2*0513 W
Consensus	GSHTIQTMYG	CDLGPDGRLL	RGYDQSAYDG	RDYIALNEDL	RSWTAADMAA	QNTQRKWEAA	GVAEQWRAYL	EGECLEXLRR	HLENGKETLQ RA

NIH-PA Author Manuscript

Figure 2c

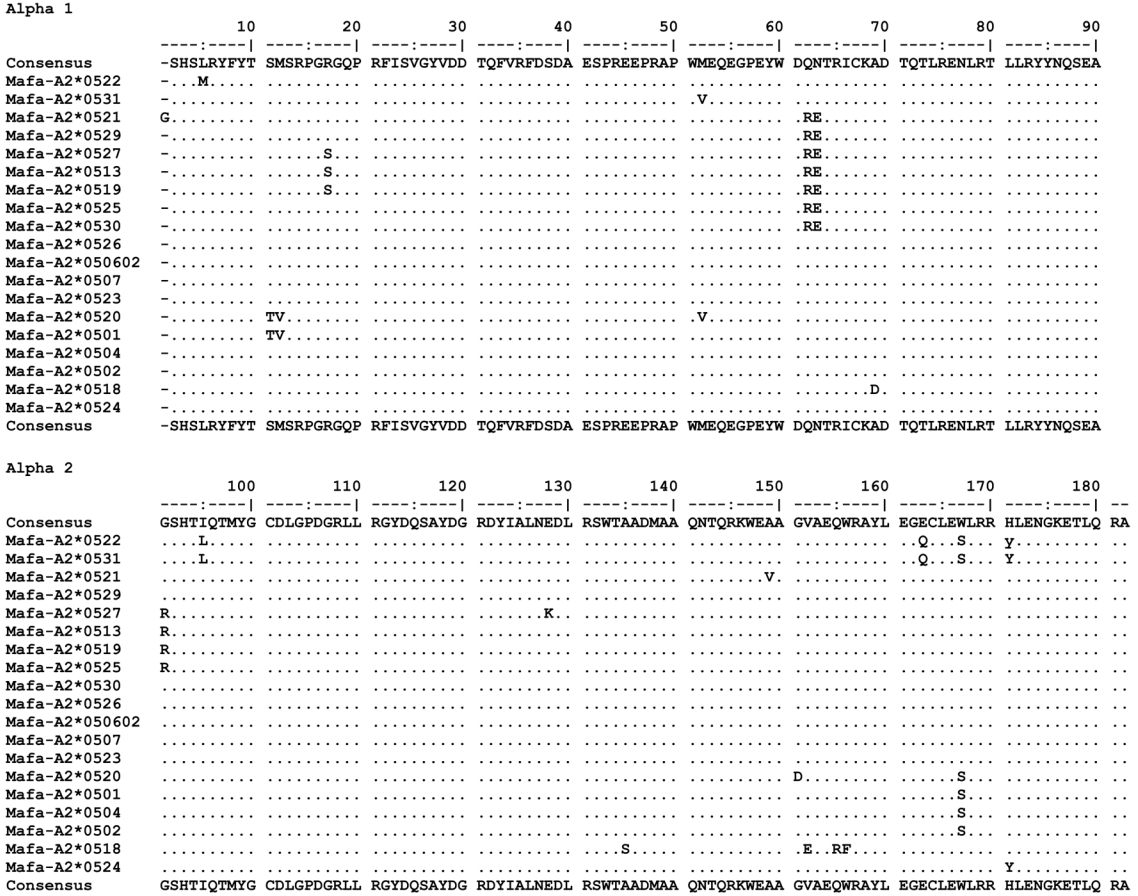


Figure 2. MHC class I amino acid alignments of the predicted alpha 1 and alpha 2 domain sequences for (a) *Mamu-A2*05*, (b) *Mane-A2*05*, and (c) *Mafa-A2*05* alleles detected in this study. Identity with the consensus is indicated by periods. Gaps in sequence are indicated by dashes, and stop codons are indicated by an asterisk.

NIH-PA Author Manuscript

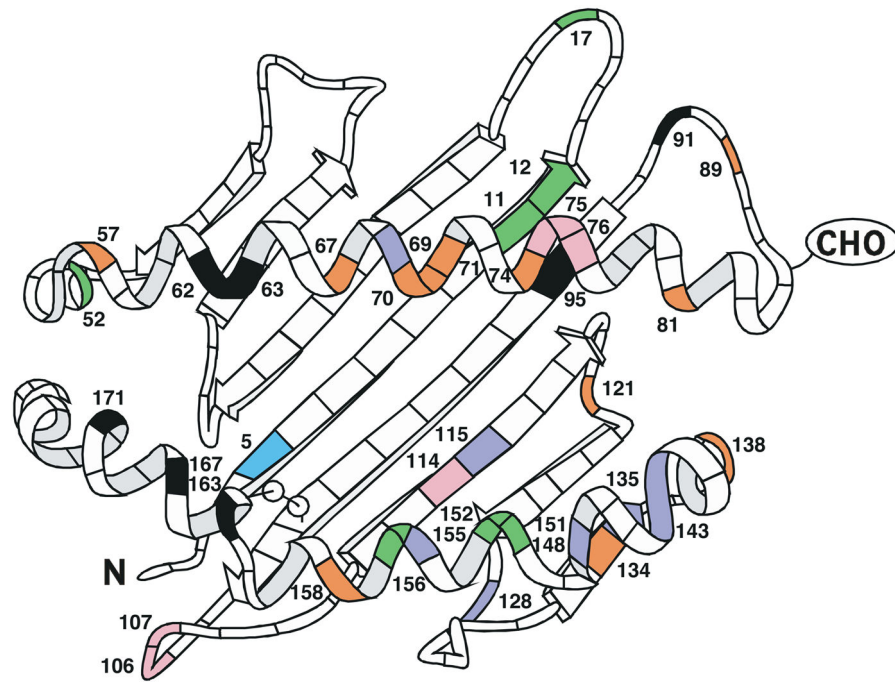


Figure 3.

A ribbon diagram structurally demonstrating the positions of alpha 1 and alpha 2 domain amino acid substitutions observed among all of the *Mamu*-, *Mane*-, and *Mafa-A2*05* family alleles previously published and described herein. Sites of polymorphism are indicated for the species as follows: *Mafa* only (violet); *Mane* only (pink); *Mamu* only (orange); *Mane* and *Mafa* (indigo); *Mamu* and *Mafa* (green); or all three species (black). The diagram is adapted from Figure 2 of Saper et al. (25).

Table 1

Mamu/Mafa/Mane-A2 locus-specific amplification and sequencing primers

Primer ID	Annealing Site	Sequence (5'-3')	Application
P000044	Sense 5' UT	GATTCTCCGACGACGCCCA	Paired with P000023 for PCR
P000023	Antisense Exon 5	GGAGAACCAGGCCAGCAAT	Paired with P000044 for PCR
P000076	Sense Intron 1	GAGCAGCGACGGGACCGCA	Exon 2 sequencing (forward)
P000060	Antisense Intron 2	CCTGGGGCTCTCCGGGTCA	Exon 2 sequencing (reverse)
P000096	Sense Intron 2	TGTA CTGAGTCTCCCTGATGG	Exon 3 sequencing (forward)
P000098	Antisense Intron 3	TTCATTCCCTCAGAGATTTT	Exon 3 sequencing (reverse)
P000055	Sense Intron 3	CCCAGGTRCCTSTGTCCAGGA	Exon 4 sequencing (forward)
P000294	Antisense Intron 4	CTCTGGGAAAGGAGGGGA	Exon 4 sequencing (reverse); <i>Mamu</i>
P000281	Antisense Intron 4	AGAGGGGAAAGTGAGGGGT	Exon 4 sequencing (reverse); <i>Mane & Mafa</i>

Note: Unless otherwise specified, primers are useful for all three species

Table 2*Mamu-A2*05* alleles detected in this study

Allele	Number of Observations	Accession Number for Prior Report of Allele	Accession Number for this Report
A2*050201	11	AF157394	----
A2*050202	1	----	EF112552
A2*050203	1	----	EF112553
A2*050204	1	----	EF112554
A2*050301	13	AM295925	EF112530
A2*050303	2	----	EF112539
A2*050402	10	AM295926	EF057836
A2*050403	3	----	EF057838
A2*050404	2	----	EF057839
A*0505	2	AJ551315	----
A*0506	11	AJ551316	----
A*0509	3	AJ551318	----
A2*0512	7	----	EF112532
A2*0513	8	----	EF112540
A2*051501	9	----	EF112546
A2*051502	5	----	EF112543
A2*0516	2	----	EF112533
A2*0518	8	----	EF112536
A2*0519	12	----	EF112542
A2*0522	8	----	EF112531
A2*0524	10	----	EF112549
A2*0525	3	----	EF112537
A2*0526	2	AM295940	EF057837
A2*0527	4	----	EF057835
A2*0528	1	----	EF112534
A2*0529	18	----	EF112535
A2*0530	3	----	EF112538
A2*0531	1	----	EF112541
A2*0532	2	----	EF112544
A2*0533	1	----	EF112545
A2*0534	1	----	EF112547
A2*0535	2	----	EF112548
A2*0536	7	----	EF112550
A2*0537	1	----	EF112551

Table 3**Mane-A2*05 (A*06)**

alleles detected in this study

Allele	Number of Observations	Accession Number for Prior Report of Allele	Accession Number for this Report
A2*050101	135	-----	EF112575
A2*050102	30	-----	EF112578
A2*050103	13	-----	EF112579
A2*050104	6	-----	EF394347
A2*0502	26	-----	EF112576
A2*0503	62	-----	EF112577
A2*0504	5	-----	EF112580
A2*0505	12	-----	EF112581
A2*050502	11	-----	EF394341
A2*0506	8	-----	EF112582
A2*0507	1	-----	EF112583
A2*0508	6	-----	EF112584
A2*0509	3	-----	EF394342
A2*0510	20	-----	EF394343
A2*0511	2	-----	EF394344
A2*0512	5	-----	EF394345
A2*0513	2	-----	EF394346
A2*0514	4	-----	EF394348
A2*0515	3	-----	EF394349
A*06	5	EF010501, AY204727	-----

Table 4*Mafa*-A2*05 alleles detected in this study

Allele	Number of Observations	Accession Number for Prior Report of Allele	Accession Number for this Report
A2*0501	1	AM295861	----
A2*0502	1	AM295862	----
A2*0504	2	AM295864	EF5500527
A2*050602	2	----	EF550522
A2*0507	2	AM295867	EF550524
A2*0513	1	AM295873	----
A2*0518	2	----	EF550523
A2*0519	2	----	EF550525
A2*0520	2	----	EF550521
A2*0521	1	----	EF589356
A2*0522	1	----	EF589357
A2*0523	1	----	EF589358
A2*0524	1	----	EF589359
A2*0525	1	----	EF589360
A2*0526	1	----	EF589361
A2*0527	1	----	EF589362
A2*0528	1	----	EF589363
A2*0529	1	----	EF589364
A2*0530	1	----	EF589365
A2*0531	2	----	EF550520

Table 5Cross-species similarity of *Mane-A2*05* alleles

Mane alleles	Closest Mamu Allele	Number of Mismatches	Closest Mafa Allele	Number of Mismatches
Mane-A2*050101	Mamu-A2*0528	1	Mafa-A2*0504	0
Mane-A2*050102	Mamu-A2*0528	2	Mafa-A2*0525	3
Mane-A2*050103	Mamu-A2*0528	2	Mafa-A2*0504	1
Mane-A2*050104	Mamu-A2*050402	1	Mafa-A2*0504	1
Mane-A2*0502	Mamu-A2*0522	1	Mafa-A2*0504	3
Mane-A2*0503	Mamu-A2*051502	2	Mafa-A2*0525	3
Mane-A2*0504	Mamu-A2*0529	2	Mafa-A2*050602	6
Mane-A2*050501	Mamu-A2*0512	1	Mafa-A2*0507	0
Mane-A2*050502	Mamu-A2*050403	0	Mafa-A2*050602	0
Mane-A2*0506	Mamu-A2*051502	3	Mafa-A2*0525	4
Mane-A2*0507	Mamu-A2*0529	3	Mafa-A2*050602	7
Mane-A2*0508	Mamu-A2*051502	1	Mafa-A2*0525	2
Mane-A2*0509	Mamu-A2*0512	7	Mafa-A2*0525	4
Mane-A2*0510	Mamu-A2*051502	6	Mafa-A2*0529	3
Mane-A2*0511	Mamu-A2*0529	3	Mafa-A2*050602	7
Mane-A2*0512	Mamu-A2*0528	2	Mafa-A2*0504	1
Mane-A2*0513	Mamu-A2*051502	5	Mafa-A2*0530	8
Mane-A2*0514	Mamu-A2*050301	0	Mafa-A2*0531	2
Mane-A2*0515	Mamu-A2*050301	1	Mafa-A2*0531	3
Mane-A*06	Mamu-A2*0522	0	Mafa-A2*0504	2