Immunogenetics 38: 141–145, 1993

Immuno-genetics

Characterization of the rhesus macaque (*Macaca mulatta*) equivalent of *HLA-F*

Nel Otting, Ronald E. Bontrop

Department of Chronic and Infectious Diseases, ITRI-TNO, P.O. Box 5815, 2280 HV Rijswijk, The Netherlands

Received December 28, 1992

Abstract. Nucleotide sequence analysis of rhesus macaque major histocompatibility complex class I cDNAs allowed the identification of the orthologue of *HLA-F*, designated *Mamu-F*. Comparison of *Mamu-F* with earlier published human and chimpanzee orthologues demonstrated that these sequences share a high degree of similarity, both at the nucleotide and amino acid level, whereas a New World monkey (cotton-top tamarin) equivalent is more distantly related. Exon 7, encoding one of the cytoplasmatic domains, is absent for all primate *Mhc-F* cDNA sequences analyzed so far. In contrast to the human, chimpanzee, and rhesus macaque equivalents, the cotton-top tamarin *Saoe-F* gene seems to have accumulated far more nonsynomynous than synonymous differences.

Introduction

The major histocompatibility complex (*Mhc*) can be divided into two major segments, encoding transmembrane structures, designated the class I and II regions. In humans, the class I region contains the loci for the classical transplantation antigens; these are *HLA-A*, *-B*, and *-C*, which are expressed on a wide diversity of tissues. These class I molecules present foreign peptides from intracellular origin to cytotoxic T cells, which subsequently may lyse the target cell. To execute this function, Mhc molecules are equipped with a peptide binding site (Bjorkman et al. 1987). Transplantation antigens display abundant polymorphism in the population (Parham et al. 1989), and as a consequence different alleles may select distinct peptides for T-cell activation. The extreme degree of allelic diversity is maintained by overdominant selection (Hughes and Nei 1988). Equivalents of *HLA-A*, *-B*, or *-C* loci have been detected in chimpanzees (Balner et al. 1974; Lawlor et al. 1988; Mayer et al. 1988); gorillas (Lawlor et al. 1992), orangutans, gibbons (Lawlor et al. 1990; Chen et al. 1992), and rhesus macaques (van Vreeswijk et al. 1977; Miller et al. 1991).

Apart from these classical transplantation antigens, the *HLA* class I region contains several nonclassical genes named *HLA-E*, *-F*, *-G*, and *-H* (Koller et al. 1989; Orr 1989; WHO 1992). At least one of them, *HLA-H*, is clearly a pseudogene (Orr 1989; Zemmour et al. 1990). As was found for the classical transplantation antigens, the HLA-E, -F, and -G gene products complex with β2-microglobulin. In contrast to the classical transplantation loci, however, the nonclassical *HLA* or *H2* genes are mono- or oligopolymorphic, show a limited tissue distribution, and may display specialized functions (Wei and Orr 1990; Hedrick 1992). Equivalents of the *HLA-E* locus have been detected in the orangutan (Watkins et al. 1992), whereas ortholoques of the *HLA-F* locus have been identified in the chimpanzee (Lawlor et al. 1988, 1990) and cotton-top tamarin, a New World primate species (Watkins et al. 1990). Remarkably, at least 11 alleles have been identified for the *Saoe-G* locus in the cotton-top tamarin (Watkins et al. 1990, 1991a, b).

Resistance to collagen type-II-induced arthritis in rhesus monkeys is controlled by a particular *Mamu-A* allele, or alternatively by a closely linked gene (Bakker et al. 1992). To gain insight into the mechanisms underlying this association, we began to analyze the class I region. In the process of sequencing *Mamu* class I cDNAs an ortholoque of the *HLA-F* locus was identi-

The nucleotide sequence data reported in this paper have been submitted to the Genbank nucleotide sequence database and have been assigned the accession number Z 21819.

Correspondence to: R. E. Bontrop.

fied. The nucleotide and deduced amino acid sequence of the *Mamu-F* gene is reported in this communication and compared to available homologues of other primate species.

Materials and methods

Cells. Transformed B cells of rhesus macaque KM (Mamu-A26/B23) were grown in RPMI 1640 medium supplemented with 10% (vol/vol) fetal calf serum, glutamine, penicillin, and streptomycin.

Preparation and amplification of cDNA. RNA was isolated according to standard procedures, whereas the commercially available Riboclone kit (Promega, Madison, WI) was used to prepare cDNA. The protocol and primers used for amplification of *Mhc* class I cDNAs have been described previously (Ennis et al. 1990).

Subcloning and sequencing. Polymerase chain reaction products were digested with the restriction enzymes *Sal* I and *Hin* dIII and cloned into the M13 derivatives tg130 and tg131 (Kieny et al. 1983). Inserts were sequenced by the dideoxy chain termination method (Sanger et al. 1977) using ^{35}Sa-α-thio-ATP and modified T7 DNA polymerase (Promega). The reported *Mamu-F* sequence represents the consensus of five independent clones.

Results and discussion

As can be seen, the nucleotide sequence of the *Mamu-F* cDNA has been aligned with its human (*HLA-F*), chimpanzee (*Patr-F*), and cotton-top tamarin (*Saoe-F*) equivalents (Fig. 1). Apart from diagnostic nucleotide substitutions, some features distinguish the *Mhc-F* locus from other types of *Mhc* class I genes. For example, the *HLA-F* nucleotide sequence, originally obtained from genomic DNA, was found to possess a mutated 3′ splice site in intron 6, resulting in the fact that the corresponding messenger lacks exon 7 (Orr 1989). This phenomenon seems to be more or less conserved in primates, since cDNA sequences obtained from chimpanzees (Lawlor et al. 1990), rhesus macaques (this communication), and tamarins (Watkins et al. 1991a, 1993) also lack exon 7 (Fig. 1). If the absence of exon 7 is caused by one single genetic event, then the mutation of the 3′ splice site in question is likely to have taken place prior to speciation of hominoids, Old-, and New World monkeys, starting about 35 million years ago (Pilbeam 1984).

It is noted that the *Mhc-F* locus gene products have a somewhat shorter leader peptide than most other *Mhc* class I sequences due to a mutation affecting the first ATG initiation codon. As a consequence, an ATG initiation codon lying more upsteam is being employed (Fig. 1). Such observations were not only done for rhesus macaques but also in humans (Orr 1989) and chimpanzees (Lawlor et al. 1990), whereas for the tamarin (Watkins et al. 1993) this type of information is not available, since the corresponding segment of exon 1 has not been sequenced (Fig. 1).

The rhesus macaque and tamarin sequences show a 6-nucleotide insertion in exon 2, which genetically seems to be unstable, since it differs between these two species (Fig. 1). The *HLA-* and *Patr-F* genes lack these 6 extra nucleotides (Fig. 1). It seems conceivable that this particular insert originates from one primordial event affecting an ancestral form of the *Mhc-F* gene. Since both the Old- and New World monkey representatives share the insert, the original integration must predate their speciation and must have taken place more than 35 million years ago. If this scenario is true, than this insert must have been lost somewhere along hominoid speciation.

The deduced amino acid sequences of the various primate *Mhc-F* genes are depicted in Figure 2. Differences between the *Mhc-F* genes of different species are mainly explained by point mutations. Some of these mutations result in amino acid replacements, whereas others are silent (Figs. 1, 2). The degree of similarity of various sets of *Mhc-F* amino acid sequences has been calculated (Table 1). As can be seen, the *HLA-* and *Patr-F* sequences share a degree of 98.4% similarity. Of the 17 nucleotide differences only six resulted in amino acid replacements (Lawlor et al. 1991). The *Mamu-F* sequence shares a degree of similarity of about 94% with its hominoid equivalents. For this comparison, of 54 differences 35 (human) or 34 (chimpanzee) turned out to be of a synonymous character. This comparison shows that the structure of the *HLA-*, *Patr-F*, and *Mamu-F* sequences has been maintained over long evolutionary spans of time. No evidence could be found that these genes have mutations that would render them pseudogenes. In this light it is possible that there may be a biological role for the corresponding gene products. For the New World monkey sequence a disparate situation is observed. Comparison with Hominoid and Old World monkey ortholoques demonstrates that degrees of similarity reach levels of about 81–82% (Table 1). As expected, in these cases more non-synomynous than synonymous differences are observed. This may indicate that the *Saoe-F* sequence has diversified to a considerable extent and may be less subject to constraint than its homoloques. In the extreme situation the *Saoe-F* gene is, or may be, on the way to become a pseudogene. This situation is plausible, as is reflected by the nucleotide composition of exon 5 encoding the transmembrane region (Figs. 1, 2). Exon 5 is affected by at least one deletion and a high number of non-synonymous differences (Figs. 1, 2). Possibly some of these alterations influence the ability of the tamarin gene product to stay in a membrane-bound configuration. This is in agreement with the observation that transfection of the *Saoe-F* gene in COS cells did not result in expression of the corresponding gene product (Watkins et al. 1991).

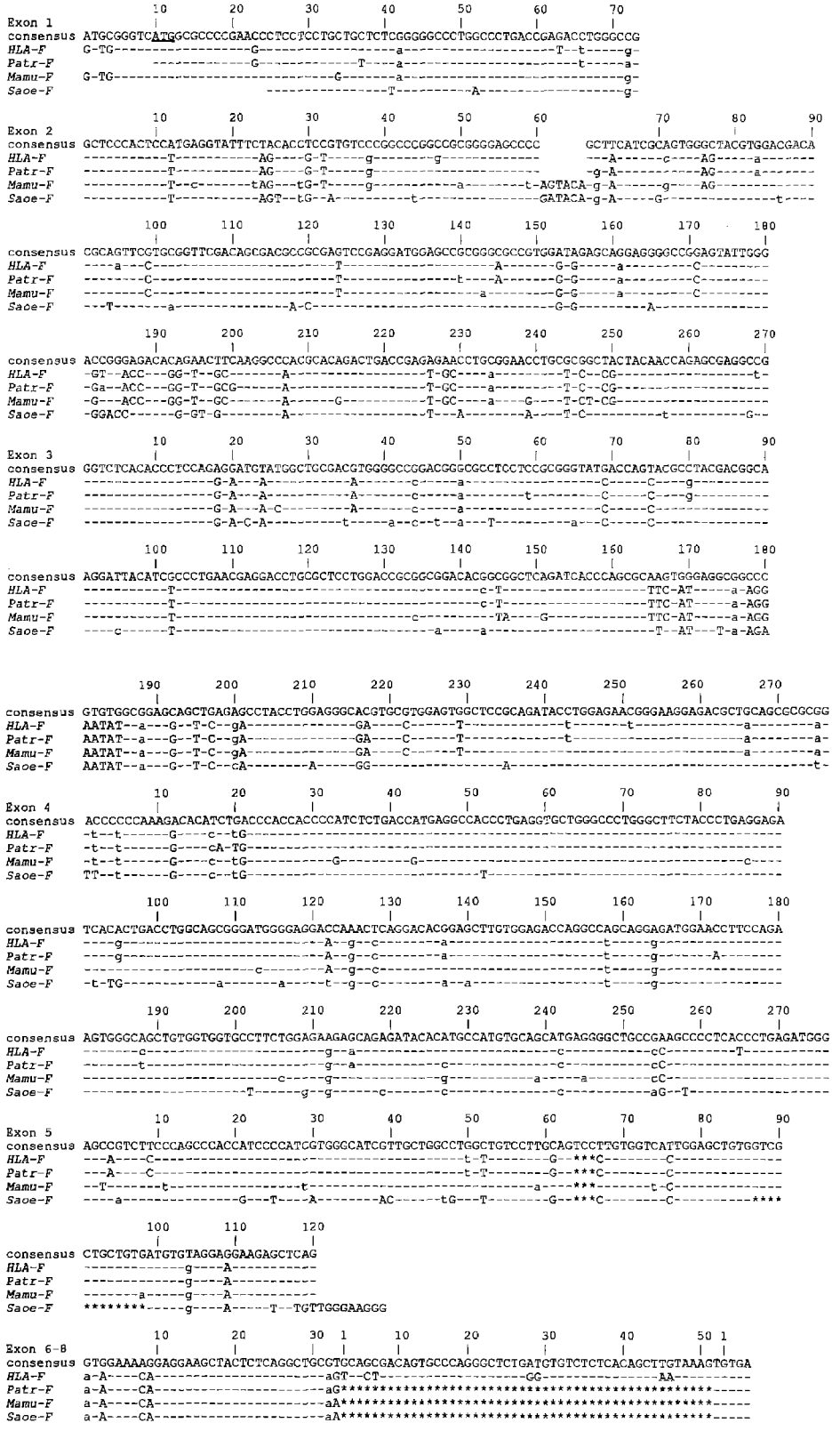

Fig. 1. Nucleotide sequence alignment of human (*HLA*), chimpanzee (*Patr*), rhesus macaque (*Mamu*), and cotton-top tamarin (*Saoe*) *Mhc-F* alleles. The *top line* represents a consensus *Mhc* class I sequence. Identity to this sequence is indicated by a *dash*, non-synonymous and synonymous substitutions by *capitals* and *lower-case* characters, respectively, whereas gaps to the consensus are depicted by an *asterisk*. The ATG initiation codon has been *underlined*. The *HLA-F* (Orr 1989), *Patr-F* (Lawlor et al. 1990), and *Saoe-F* (Watkins et al. 1993) sequences were taken from the literature.

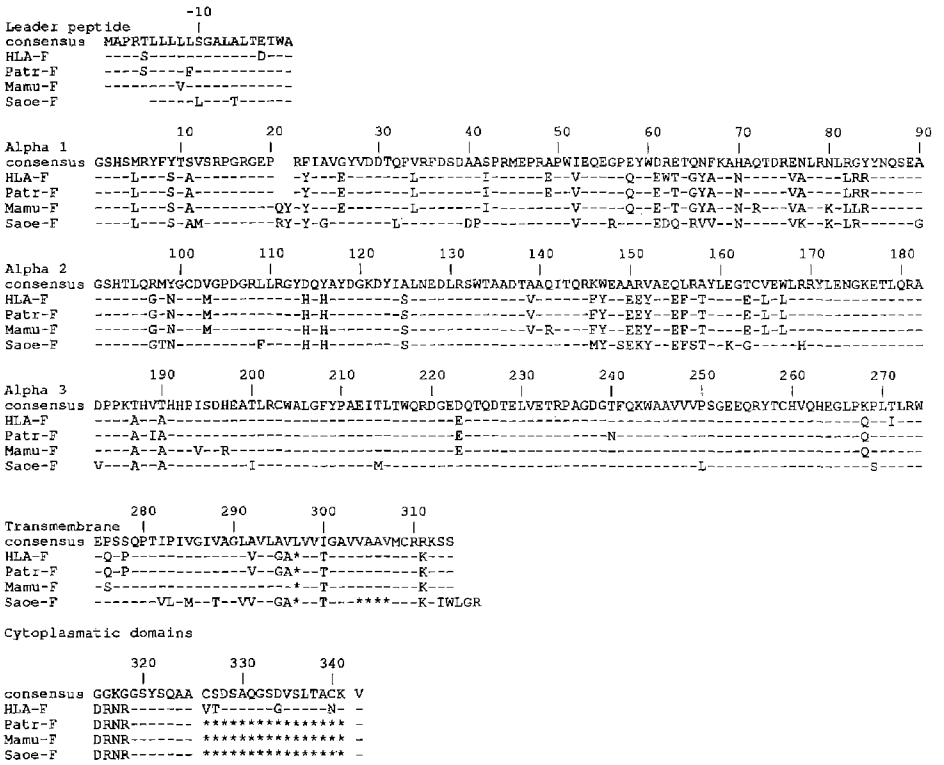

Fig. 2. Alignment of deduced amino acid sequences of various primate *Mhc-F* sequences.

Table 1. Percentages of amino acid sequence similarity observed between various sets of primate *Mhc-F* sequences.

	Patr	*Mamu*	*Saoe*
HLA	98.4%	94.5%	82%
Patr	–	94.2%	81.7%
Mamu	–	–	81.0%

In conclusion, these data show that the *Mhc-F* gene has been conserved in hominoid and Old World primate species, whereas its equivalent diversified considerably in a New World primate species and possibly may constitute a pseudogene in tamarins.

References

Balner, H., Gabb, B. W., D'Amaro, J., van Vreeswijk, W., and Visser, T. P. Evidence for two linked loci controlling the serologically defined leucocyte antigens of chimpanzees (ChL-A). *Tissue Antigens 4:* 313–328, 1974

Bakker, N. P. M., van Erck, M. G. M., Otting, N., Lardy, N. M., Noort, R. C., 't Hart, B. A., Jonker, M., and Bontrop, R. E. Resistance to collagen-induced arthritis in a nonhuman primate species maps to the major histocompatibility complex class I region. *J Exp Med 175:* 933–937, 1992

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennet, W. S., Strominger, J. L., and Wiley, D. D. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature 329:* 512–518, 1987

Chen, Z. W., McAdam, S. N., Hughes, A. L., Dogon, A. L., Letvin, N. L., and Watkins, D. I. Molecular cloning of orangutan and gibbon MHC class I cDNAs: the *HLA-A* and *-B* loci diverged over 30 million years ago. *J Immunol 148:* 2547–2554, 1992

Ennis, P. D., Zemmour, J., Salter, R., and Parham, P. 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, 1990

Hedrick, S. M. Dawn of the hunt for nonclassical MHC function. *Cell 70:* 177–180, 1992

Hughes, A. L. and Nei, M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature 335:* 167–180, 1988

Kieny, M. P., Lathe, R., and Lecocq, J. P. New versatile and sequencing vectors based on bacteriophage M13. *Gene 26:* 91–96, 1983

Koller, B. H., Geraghty, D. E., DeMars, R., Duvick, L., Rich, S. S., and Orr, H. T. Chromosomal organization of the human major histocompatibility complex class I gene family. *J Exp Med 169:* 469–480, 1989

Lawlor, D. A., Ward, F. E., Ennis, P. D., Jackson, A. P., and Parham, P. HLA-A and B polymorphism predate the divergence of humans and chimpanzees. *Nature 326:* 268–271, 1988

Lawlor, D. A., Warren, E., Ward, F. E., and Parham, P. Comparison of class I MHC alleles in humans and apes. *Immunol Rev 113:* 147–185, 1990

Lawlor, D. A., Warren, E., Taylor, P., and Parham, P. Gorilla class I alleles comparison to human and chimpanzee class I. *J Exp Med 174:* 1491–1509, 1992

Mayer, W. E., Jonker, M., Klein, D., Ivanyi, P., van Seventer, G., and Klein, J. Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-species mode of evolution. *EMBO J 7:* 2765–2774, 1988

Miller, M. D., Yamamoto, H., Hughes, A. L., Watkins, D. I., and Letvin, N. L. Definition of an epitope and MHC class I molecule recognized by Gag-specific cytotoxic T lymphocytes in SIVmac-infected rhesus monkeys. *J Immunol 147:* 320–329, 1991

Orr, H. T. HLA class I gene family: Characterization of genes encoding non-HLA-A,B,C proteins. *In* B. Dupont (ed.): *Immunobiology of HLA, vol. II.*, pp. 33–41, 1989

Parham, P., Lawlor, D. A., Lomen, C. E., and Ennis, P. D. Diversity and diversification of *HLA-A, B, C* alleles. *J Immunol 142:* 3937–3950, 1989

Pilbeam, D. The descent of homonoids and hominids. *Sci Am 250:* 60–69, 1984

Sanger, F., Nicklen, S., and Carlson, R. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA 77:* 5463–5467, 1977

Van Vreeswijk, W., Roger, J. H., D'Amaro, J., and Balner, H. The major histocompatibility complex of rhesus monkeys, RhL-A. VII. Identification of five new seriologically defined antigens. *Tissue Antigens 9:* 17–30, 1977

Watkins, D. I., Chen, Z. W., Hughes, A. L., Evans, M. G., Tedder, T. F., and Letvin, N. L. Evolution of the MHC class I genes of a New World primate from ancestral homologues of human non-classical genes. *Nature 346:* 60–63, 1990

Watkins, D. I., Chen, Z. W., Garber, T. L., Hughes, A. L., and Letvin, N. L. Segmental exchange between MHC class I genes in a higher primate. Recombination in the gorilla between the ancestor of a human non-functional gene and an *A* locus gene. *Immunogenetics 34:* 185–191, 1991a

Watkins, D. I., Chen, Z. W., Garber, T. L., Hughes, A. L., and Letvin, N. L. Unusually limited nucleotide variation of the expressed major histocompatibility complex class I genes of a New World primate species (*Sanguinus oedipus*). *Immunogenetics 33:* 79–89, 1991b

Watkins, D. I., Zemmour, J., and Parham, P. Non-human primate MHC class I sequences, 1992. *Immunogenetics 37: 317–330, 1993*

Wei, X. and Orr, H. T. Differential expression of HLA-E, HLA-F and HLA-G transcripts in human tissue. *Hum Immunol 29:* 131–142, 1990

WHO Nomenclature Committee for factors of the *HLA* system: nomenclature for factors of the *HLA* system, 1991. *Immunogenetics 36:* 135–148, 1992

Zemmour, J., Koller, B. H., Ennis, P. D., Geraghty, D. E., Lawlor, D. A., Orr, H. T., and Parham, P. HLA-AR, an inactivated antigen-presenting locus related to HLA-A: implications for the evolution of the MHC. *J Immunol 144:* 3619–3629, 1990