# ORIGINAL PAPER

Juan F. Parra-Cuadrado · Manuel Gómez del Moral Pablo García-Pavía · Fernando Setién Eduardo Martínez-Naves

# Characterization of the MHC class I-related *MR1* locus in nonhuman primates

Received: 9 July 2001 / Accepted: 18 September 2001 / Published online: 6 November 2001 © Springer-Verlag 2001

Abstract We characterized the MHC-related 1 (MR1) locus in two nonhuman primates species, Pongo pygmaeus and Pan troglodytes. MR1 cDNA sequences encoding several isoforms generated through alternative splicing were observed in both species. Amino acid alignment between the five species in which MR1 has been characterized to date – human, chimpanzee, orangutan, mouse, and rat - reveals a very high degree of conservation specially in the  $\alpha 1$  and  $\alpha 2$  domains of the molecule. The main differences concentrate in the transmembrane and cytoplasmic domains. In the three primates species there is a lysine residue inside the putative transmembrane domain which is not present in rodents. Furthermore, the MR1 cytoplasmic region is longer in rodents, with a conserved serine-containing motif that could be involved in endocytosis; remarkably, this motif is absent in the three primate species. We also describe the presence in the chimpanzee of a sequence homologous to the MR1P1 pseudogene previously found in humans.

**Keywords** MR1 · HLALS · MR1P1 · MHC · Chimpanzee · Orangutan

## Introduction

Human classical MHC class I genes (*HLA-A*, -*B*, -*C*) map to MHC loci on Chromosome (Chr) 6. Among human nonclassical genes, *HLA-E*, -*F*, -*G* (O'Callaghan and Bell 1998), *HFE* (Feder et al. 1996), and *MIC* (Bahram 2001) map to MHC loci, whereas *AZGP1* (Pendás et al. 1994),

J.F. Parra-Cuadrado · M.G. del Moral · P. García-Pavía

E. Martínez-Naves (🖂)

Unidad de Inmunología, Facultad de Medicina, Universidad Complutense, 28040 Madrid, Spain e-mail: emnaves@med.ucm.es

Tel.: +34-91-3941644, Fax: +34-91-3941641

F. Setién

CAI Técnicas Inmunológicas, Facultad de Medicina, Universidad Complutense, 28040 Madrid, Spain *PROCR* (Fukudome and Esmon 1994), *ULBP* (Cosman et al. 2001), *FCGRT* (Ghetie and Ward 2000), *CD1* (Porcelli and Modlin 1999), and *HLALS* (*MR1*) (Hashimoto et al. 1995) are sited outside the MHC complex.

*MR1* was originally identified as a cDNA sequence by Hashimoto and co-workers (1995). The gene for human and mouse MR1 (Riegert et al. 1998; Yamaguchi et al. 1998) as well as a rat cDNA encoding MR1 (Walter and Günther 1998) have been described.

*MR1* is not polymorphic (Parra-Cuadrado et al. 2000) and is ubiquitously expressed, as revealed in various Northern blot and RT-PCR experiments. Several mRNA isoforms generated through alternative splicing have been found in human, mouse, and rat (Hashimoto et al. 1995; Riegert et al. 1998; Walter and Günther 1998; Yamaguchi et al. 1997).

MR1 function is unknown, and of the nonclassical MHC class I proteins it shows the highest level of amino acid identity to classical molecules, specially in the  $\alpha$ 1 and  $\alpha$ 2 domains. The sequence conservation between human, mouse and rat is also very high. In humans, a *MR1*-related pseudogene (*MR1P1*) has been found containing exon 2, intron2, and exon 3 sequences (Parra-Cuadrado et al. 2000).

Here we present the isolation and characterization of *MR1* homologues in chimpanzee and orangutan.

# **Materials and methods**

Cell lines

The following Epstein-Barr virus-transformed B-cell lines purchased from the ATCC (Manassas, Va.) were used: CARL (ATCC CRL 1857) from chimpanzee (*Pan troglodytes*) and PUTI (ATCC CRL 1850) from orangutan (*Pongo pygmaeus*).

cDNA amplification

*MR1* cDNA was obtained after RT-PCR amplification using oligonucleotides homologous to human *MR1* as previously described (Parra-Cuadrado et al. 2000). Genomic DNA amplification

*MR1* genomic DNA encompassing exons 2 and 3 and the intervening intron was amplified using primers homologous to the human *MR1* gene as previously described (Parra-Cuadrado et al. 2000).

#### Cloning and sequencing

PCR-amplified products were cloned and automatically sequenced. Sequence analysis was performed using GCG (Genetics Computer Group) software. Multiple sequence alignment was obtained with ClustalW software at the EBI (European Bioinformatics Institute) server.

## **Results**

PCR amplification of the chimpanzee and orangutan *MR1* cDNA

To characterize *MR1* cDNA sequences, RT-PCR experiments were carried out with primers previously used in our laboratory to amplify full-length human *MR1* cDNA



**Fig. 1** RT-PCR analysis of *MR1* expression in orangutan and chimpanzee EBV B-cell lines (*C* chimpanzee sample, *O* orangutan sample, – negative control, *Mw* molecular-weight marker)

(Parra-Cuadrado et al. 2000). Amplification was successful and we obtained several bands in samples from both species, as shown in Fig. 1.

## Orangutan MR1 cDNA

In a total 12 clones we obtained three types of sequence in the orangutan samples (see Fig. 2). A 1066-nucleotide sequence containing information for the leader peptide,  $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 3$  transmembrane and cytoplasmic domains was found (*MR1B1*). This sequence was 97.8% homologous to the human *MR1* cDNA clone reported by Hashimoto and co-workers (1995). A second kind of clone, 790 nucleotides in length, named *MR1B2* was found that lacks exon4 encoding the  $\alpha 3$  domain and is very likely generated through alternative splicing. The third *MR1* transcript (*MR1B3*) identified was 710 nucleotides long and identical to *MR1B2* except for a deletion of 79 nucleotides in the 3' end of the  $\alpha 1$  domain. A cryptic splice donor site located 77 bp from the 3' end of the  $\alpha 1$  domain generates this gap.

### Chimpanzee MR1 cDNA

In chimpanzee, we identified a 1066-nucleotide cDNA clone (*MR1B1*), 99.7% homologous to human *MR1* and 97.9% homologous to orangutan *MR1*. In addition, a chimpanzee *MR1* transcript lacking the  $\alpha$ 3 domain (*MR1B2*) was found (see Fig. 2).

Six out of 15 *MR1* chimpanzee clones show a point mutation in codon 197 (ACC $\rightarrow$ ATC) resulting in a substitution of threonine (accession numbers AJ275982, AJ275983) by isoleucine (accession number AJ275984, AJ275985). This indicates that CARL is a heterozygous cell line for the *MR1* gene. Orangutan and human MR1 sequences have T197, thus I197 is a chimpanzee-

Fig. 2 Diagram representing the different cDNA isoforms found in chimpanzee and orangutan. EMBL/GenBank accession numbers are indicated for each isoform. The *asterisk* in the orangutan *MR1B3* isoform indicates a cryptic splice donor site



Fig. 3 Multiple alignment of the predicted MR1 amino acid sequences from primates (human, chimpanzee, and orangutan) and rodents (mouse and rat). Asterisk marks position 303: lysine in the three primate species. Identities with the MR1 consensus sequence are indicated by dashes, deletions/insertions are shown by dots. GenBank accession numbers are: human MR1, U22963; chimpanzee MR1, AJ275984; orangutan MR1, AJ271828; mouse MR1, U944989; rat MR1, Y13972

consensus	▼Leader MGELMAFLLP	LIIVLMVKHS	▼ <i>α1 domai</i> DSRTHSLRYF	n RLGVSDPI-G	VPEFISVGYV					
HUMAN				Н-		50				
CHIMPANZEE				H-						
MOUSE	-ML	-LA-FLR-	НТ	AGPV						
RAT			-	AGP-	·					
consensus	DSHPITTYDS	VTRQKEPRAP	WMAENLAPDH	WERYTQLLRG	WQQMFKVELK					
HUMAN						100				
ORANGUTAN		Q								
MOUSE		K			TAR					
RAT					RT-QTR					
conconcilo	DI ODUVNUSC		FITEDCOMMC		די דדאניסייד פ					
HUMAN						150				
CHIMPANZEE										
ORANGUTAN	н									
RAT	H	L			-ID					
	MT ATTIMUT UT	TROVAEVNOR	ET OVORNWI E	FFCTANTEDF	ΙΕΥΛΥΠΤΙΟΡ					
HUMAN			L			200				
CHIMPANZEE			L							
ORANGUTAN	T	R			F-					
RAT	MI	T-RL-			S-A-E-					
	<b>▼</b> α <b>3 domai</b>	'n								
consensus	TEPPLVRVNR	KETFPGVTTL	FCKAHGFYPP	EIYMTWMKNG	EEIVQE-DYG	0.5.0				
HUMAN CHIMPANZEE		A_			I	250				
ORANGUTAN					M					
MOUSE	H-VTT-	F	R	S						
RAT	H-V11-		K	2-1-1	v					
consensus		OTWASVELDP	OSSNLYSCHV	EHCGVHMVLO	▼ <i>Trans-</i> VPOESETIPL					
HUMAN		-A				300				
CHIMPANZEE					 D					
MOUSE	G	LN	NDV	RQE	A-RGD-LR					
RAT	G	-M-VD	-TKDI	E	AGNTL-					
	membrane	domain	▼Cytoplasmic tail							
consensus	VM <b>K</b> AVSGSIV	LVIVLAGVGV	LVWRRRPREQ	NGAIYLPTPD	-EGSSPS	7				
HUMAN CHIMPANZEE					R	·				
ORANGUTAN		F								
MOUSE	STTT-	A n	SQ-L	KEVM-QQV	N					
KAT	-ANTT	A	<b>_</b> 3 <b>_</b> _E	VPANA - AA	14					



**Fig. 4** PCR amplification of MR1 and MR1P1 fragments encompassing exon2, intron 2, and exon 3 from human (H), chimpanzee (C), and orangutan (O) genomic DNA. MR1P1 sequences were not amplified in orangutan

specific substitution that alters the hydrophobicity in the  $\alpha 2$  domain located outside the putative peptidebinding region.

A comparison of the MR1 predicted amino acid sequences from human, chimpanzee, orangutan, mouse, and rat reveals some interesting features as shown in Fig. 3. The highest similarities between species are found in the  $\alpha$ 1 and  $\alpha$ 2 domains,  $\alpha$ 3 and, in particular, the transmembrane and cytoplasmic domains being more divergent. In primates, there is a positively charged lysine residue (number 303 in Fig. 3) inside the putative transmembrane domain which corresponds to serine in mouse and asparagine in rat. The cytoplasmic domain of MR1 in primates is shorter than in rodents which display three serine residues absent in primates. Fig. 5 Comparison between MR1 sequences from chimpanzee (cMRI) and human (hMRI)genomic DNA encompassing exon 2, intron 2, and exon 3 (a) and comparison between MR1P1 sequences from chimpanzee ( $c\hat{M}RPI$ ) and human (hMRP1) genomic DNA encompassing exon 2, intron 2, and exon 3 (b). Identities between sequences are indicated by dashes, deletions/insertions are shown by dots, single point substitutions are indicated by diamonds. MR1 and MR1P1 exons are boxed. Database accession numbers are: cMR1, AJ315655; hMR1, AF073484; cMR1P1, AJ315654; hMR1P1, AJ132011

## A) MR1

- cMR1 cccatttcca ctcaggcttc tgatatgcat ctccttttcc attctaaatt tgcccattct gtgtctcact tcctggtagc 560 hMR1 ------t

- cMR1 ctttgacagg cagtagttac ttcacagata cctgctgaca ggtggggcttc cataaggggg acctcctggg gactcagcga 800 hMR1 -----g--- -----g--- ------
- cMR1 tgccacggca ggcctggggg gtgacataat gtagaccaaa ggatgcttcc agccatgccc ctgctcccag cacttgagag 880 hMR1 -----t---- ------
- EXCH 5 (COMAIN UZ) GTGAGGTGGCT GGAGGATGGA AGCACCACAG GATTTCTGCA GTATGCATAT GACGGGCAGG ATTTCCTGAT CTTCAATAAA 1040 hmri ------
- CMR1 GACACCCTTT CCTGGCTGGC TGTAGATAAT GTGGCTCACA CCATCAAGCA GGCATGGGAG GCCAATCAGC ATGAGTTGCT 1120
- hMR1 -----C- -----C- ------
- CMR1 GTATCAAAAG AATTGGCTGG AAGAAGAATG TATTGC 1156

#### B) MR1P1

DARIPI GACTUTTOGO ATCCCATCA TGGGGTCCT GAATTATTT CGGTTGGGTA TGTGGACTG CACCCTATCA CCACATATGA 80 ACCTOCTAGO ATCCCATCA CAGAAGG AGCCACGGGC CCCATGGATG GCAGAGAAC TGGGGCTG CACCCTATCA CCACATATGA 80 CAGTGTGACT CCGCAGAAGG AGCCACGGGC CCCATGGATG GCAGAGAAC TGGCGCCGA AGAGGCACT ACAATCACTC AGgtgtgggtgg AGCTGTCAGA GGGCTGGCAG CAGATGTTCA AGGTGGAACT GAAGGGCCGG AAGAGGCACT ACAATCACTC AGgtgtggggtgg AGCTGCTGAA GGGCTGGCAG CAGATGTTCA AGGTGGAACT GAAGGGCCCG GAAGAGGCACT ACAATCACTC AGgtgtgggg AGCTGCTGAA GGGCTGGCAG CAGATGTTCA AGGTGGAACT GAAGGGCCCG GAAGAGGCACT ACAATCACTC AGgtgtgggg CAGRIPI Intron 2 Intron 2 Intr		lavon 2 (dor	main all							
MNR1P1	CMR1P1	GGCGTTTCGG	ATCCCATCCA	TGGGGTCCCT	GAATTTATTT	CGGTTGGGTA	TGTGGACTCG	CACCCTATCA	CCACATATGA	80
CARIPI CANTETCACT COGCAGAAGG AGCCACGGGC CCCATGGATG GCAGAGAACC TGCGCCCTGA TCACTGGGAG AGGTACACTC 16 HMRIPI 	hMR1P1					-A				
<pre>cMRIPI cMRIPI cMRIPI cMRIPI cMRIPI cmm.cmm.cmm.cmm.cmm.cmm.cmm.cmm.cmm.cmm</pre>			•				•			
<pre>MMRIPI AGCTGCTGAN GGGCTGGCAG CAGATGTTCA AGGTGGAACT GAAGCGCCAG AAGAGGCACT ACACTTACCA AGGTqtgtgat 24 AGCTGCTGAAT CGAGAGGGG CAGATGTATGA TGAAGGGCACA AGAGGCACT ACACTTACCA GGGATGAT CACATTACCA TAGAGAGCACT ACACTTACCA AGAGAGCACT ACACTACCA GAGAATGATT 1 AGGCTGCTGAAC TAGAGATGA TAATGTGGCT CACACTATACCA AGAGGCACA CAGATATCC AGATATGC 1320 AGGGCTGACA TAGAATACCA TAAGAATAGC CTGGAGAGAA AATGTATGCC 1A20 AGGGCACA AGAGGCACA CAGAAGAAG AATGTATGCC 1A20 AGGGCACA AAGAGGCACA CAGAAGAAG AATGTATGCC 1A20 AGGGCACA AAGAGGCACA CAGAAGAAG AATGTATGCC 1A20 AGGACATCA TAAGAATAGC CTGGAAGAAG AATGTATGCC 1A20 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATGCC 1A20 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATGC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACCA TAAGAATAG CTGGAAGAAG AATGTATTC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACCA TAAGAATGG CTGGAAGAAG AATGTATTCC 1320 AGGALTACCA TAAGAAGAACA CAGGAATGTCC CGAATGTACCA CACCTACA TAAGAAGAACAC CCCACCAA AAGAGGCACA CACCTACACCAGAGAACACACACACACCACAAAGAACACACAC</pre>	cMR1P1	CAGTGTCACT	CCGCAGAAGG	AGCCACGGGC	CCCATGGATG	GCAGAGAACC	TCGCGCCTGA	TCACTGGGAG	AGGTACACTC	160
AGCTOCTGAA GGOCTGGCAG CAGATGTTCA ÁGGGAACT GAAGGGCCAA AAGAGGCACT ACAATCACTC AGgtgtgaca 24 INTRON 2 INTRON 2 INTRON 2 G	hMR1P1		-A				T			
AGCTGCTCAA GGGCTGGCAG CAGATGTTCA AGGTGGAACT GAAGCGCCAG AGAAGCACT ACAATCACT AGTGGACT AGGTGGAACT GAAGCGCCAG AGAAGGGCAG ACAATCGAT CAGAAGGGCAG GGAGGCCAA CAGAAAGGAT 22         INTRIP1					•		• •			
<pre>MARIPI intron 2 common provide addition of the second second</pre>	cMR1P1	AGCTGCTGAA	GGGCTGGCAG	CAGATGTTCA	AGGTGGAACT	GAAGCGCCAG	AAGAGGCACT	ACAATCACTC	AGgtgtgcat	240
<pre>intron 2 intron intrecon intron intron intron intron intron intro intron intr</pre>	hMR1P1				G		CA			
cMRIP1       gragcagaga       cagagagate       conservation       conservat		intron 2							•	
<pre>MARIPI</pre>	cMR1P1	gcagcagaga	cagacgette	cctcacccca	cagaggeeee	tgggctgacc	tccaatgage	aatgctgact	tgcacctact	320
cdmR1P1       gtggttttg gcaaaacctg ggaaatctgg ttggtggagt tcagagcatc ccatotgoct gtgcatcttc tggactgtcc 40         cdmR1P1       ctototcocc caggagcoct ttatoctctg cogcatoccc cttccccata tcoccatct cttcccatt ctaaatttgc 48         cdmR1P1       ctototcocc caggagcoct ttatoctctg cogcatoccc cttccccata tcoccatct cttcccatt ctaaatttgc 48         cdmR1P1       ccattctgtg tctcacttct tggtagcoag goctagtgac ctgottgott atgcataatg acattaactt tcoccaggagt 56         cdmR1P1	hMR1P1								g	
charlp1       dfggdtttg gcaaaacetg ggaatetgg ggaatetgg ttgggtggg ttgggatet cagaggate cetetget gtggatetee 4         charlp1       cetetetge cagaggeet ttateceteg eggaatetgg eggatetee eggatetee 4         charlp1       cetetetgg teteaette tggtageeag geetagtgae eggetgete atgeataat eteecate ettegeatetee tggategee 4         charlp1       cetetetgg teteaette tggtageeag geetagtgae etgetgete atgeataatg acattaaett ecceates ettegeatetee teteaettee ende         charlp1       attetetaget eetgeatete teteeete ceaetgeaet ecceates ettegeategea etteaettee ende         charlp1       attetetaget eetgeatete teteeete ecceates ettegeat teteaettee ende         charlp1       attetetaget eetgeatete teteeete ecceates ettegeat teteaete ecceates					•	•			4 h h h h	
<pre>MARIPI</pre>	cMR1P1	gtggcttttg	gcaaaacctg	ggaaatctgg	ttggtggagt	tcagagcate	ccatctgcct	gtgcatcttc	tggactgtcc	400
cdtretcoccc       caggagcoct       ttatctctg       cogcatoccc       cttccccata       tccccatot       ctttcccatf       cttcccatf       cttcccaff       cccattccaff       cccatccaff       cccatccaff       cccatccaff       cccatccaff       cccatccaff       cccatccaff       cccatccaff       cccatccaff       ccatccaff       ccacaff       ccacaff       ccacacaff       <	hMR1P1				£	g				
CMRIPI CORRIPI CONTROL OF CONTROL OF CONTR	->/0101			ttotaatata		attaaaata	toccatoto	atttaccatt	ctasatttac	480
<pre>MRRIPI</pre>	CMRIPI	CTCTCTCCCC	caggageeet	clatectety	cogcatecce	cttccccata				
control to the term of term of the term of	UWKIPI						•	1		
CARLEPIg	-MP1 P1	*	tatasattat	taataaccaa	acctagtgac	ctacttactt	atgcataatg	acattaactt	teccagggat	560
<pre>eMRIP1 atttctagct cotgoatoto totocotoct coactgoact caastagooa cagtttatoa atcaatcatt cocattacag 64 hMRIP1</pre>	LMD1D1									
cdMR1P1       atttctagct cctgcatct ctcccctct ccactgcact caatagcca cagtttatca atcatcatt cccattacag 64         hMR1P1	IMANIFI	y								
MARIPI	CMR1P1	atttctaget	cetecatete	tetectect	ccactocact	caaatagcca	cagtttatca	atcaatcatt	cccattacag	640
CMRIP1 gataactee caagaaggga attateacae teettggea tgeagtagga tetteacaaa taeetgtty caagatggta 72 	hMR1P1									
cMR1P1       gataactocc       caagagggg attateacac       toctttggca       tgcagtagga       tottcacaaa       tactgttgt       cagagtggta       72         cMR1P1							•		•	
hMR1P1	cMR1P1	gataactccc	caagaaggga	attatcacac	teetttggca	tgcagtagga	tcttcacaaa	tacctgttgt	cagagtggta	720
cMRIP1 taatggacac tggggacta ggaaggggaa aggtggaagg tgggtgagg gaaaaaaacc acctactga tacaacatac 80 MRIP1	hMR1P1									
cdR1P1       taatggacac       tggacggcacca       ggacggggaa       aggtggaggt       tggacggtggat       accacctactgat       tcacccactgat       tcaccactactgat       tcaccacctactgat       tcaccacctactgat       tcaccacctactgat       tcaccaccactac       tcaccactactactgat       tcaccaccactactgat       tcaccactacacgat       tcaccactactactgat       tcaccactactgat       tcaccactactgat       tcaccactacactactactactactactactactactact										
hMR1P1	cMR1P1	taatggacac	tggcgactca	ggaaggggaa	aggtggaagg	tgggtgaggt	gaaaaaaacc	acctactgag	tacaacatac	800
cMRIP1       acgattcgg tgacagatac accaaaagcc cagagttcac cactacacg ttcatccat caatacaaa ttcatcatg caatcaaaaa ctacttgtat 86         cMRIP1      t	hMR1P1									
ckR1P1       acgattgggg tgacagatac accaaaagcc cagagttac cactacacg ttoatcatg cattaaaaa ctattgtat 86         hMR1P1      t		•				•	•			
hNR1P1      t	cMR1P1	acgattcggg	tgacagatac	accaaaagcc	cagagttcac	cactacacga	ttcatccatg	caatcaaaaa	ctacttgtat	880
cMRIP1       ccctaaagct atatgaacca       caaaaaattt aaaaacaac cctgctgtca ggtggggttc caaagaggga cctcctgggg 96         hMRIP1	hMR1P1	t				g				
ccctaaagct atatgaacca caaaaattt aaaaacaac cctgctgtca ggtgggcttc caaagaggg cctcctgggg 96         cMRIP1				•						
hMRIP1	cMR1P1	ccctaaagct	atatgaacca	caaaaaattt	aaaaaacaac	cctgctgtca	ggtgggette	caaagaggga	cctcctgggg	960
cMR1P1       acttagcgat gccacagcag gcctgtgggg tgacgtagac caaaggatgc ttccagccat gcccctgct ccagcacttg 10         hMR1P1	hMR1P1			a						
cMRIP1     acttaggat gccacagaag gcctgtgggg taggtgacgtagac caaaggatge ttecageat gccocuput tcagtacteg tcaggacteg trecageat gccocuput tcagtacteg tcaggacteg trecageat gccocuput tcagtacteg trecageat gccacagaag acttgtgggcttg tgtgtgtgtt caadgGTCTC ACACTTACCA GAGAATGATT 11       cMRIP1					• <b>•</b> • • • • •		********		aasaasatta	1040
ANRIP1	cMR1P1	acttagcgat	gccacagcag	geetgtgggg	tgacgtagac	caaaggatge	ttecagecat	geeeetgete	ccaycactcy	1040
eMR1P1       agageceace tetytetety tytytyacce tetyggetty tytytytt cead@GTCTC ACACTTACCA GAGAATGATT 11         hMR1P1	hMR1P1									
cmxR1P1	-100101		******	tataasaaa	tetagaetta	tatatatat	CCAGGTCTC	ACACTTACCA	GAGAATGATT	1120
exon 3 (domain 0.2)       •         central GGCTETGAGC TGCTGGAGGA TGGAAGCACT ACAGGATTTC TGCAGTATGC ATATGATGGG CAGAATTTCC TGATTTTCAA 12	CMRIPI	agageeeace		cycygacece						
exemple       GGCTGTGAGGC TGCTGGAGGA TGGAAGCACT ACAGGATTTC TGCAGTATGC ATATGATGGG CAGAATTTCC TGATTTCAA 12         hMR1P1	HERIFI	area 3 /de	math (72)						•	
CHARLP1      C         •      C         •      C         •          •          •          •          •		GGCTGTGAGC	TGCTGGAGGA	TGGAAGCACT	ACAGGATTTC	TGCAGTATGC	ATATGATGGG	CAGAATTTCC	TGATTTTCAA	1200
eMRIPI       TAAAGACACC CTCTCCTGGC TGGCTGAGA TAATGTGGCT CACACCATCA AGAGGGCACG GGAGGCCAAT CAGCATGAGT 12         hMRIPI	hMR1P1								C	
eMR1P1 TAAAGACACC CTCTCCTGGC TGGCTGTAGA TAATGTGGCT CACACCATCA AGAGGGCACG GGAGGCCAAT CAGCATGAGT 12 hMR1P1A			•							
hMR1P1	cMR1P1	TAAAGACACC	CTCTCCTGGC	TGGCTGTAGA	TAATGTGGCT	CACACCATCA	AGAGGGCACG	GGAGGCCAAT	CAGCATGAGT	1280
CMR1P1 TGCAATATCA TAAGAATTGG CTGGAAGAAG AATGTATTGC 1320	hMR1P1		A-							
CMRIPI TGCAATATCA TAAGAATTGG CTGGAAGAAG AATGTATTGC 1320 hMRIPI A			•			[				]
hmR1P1 A	cMR1P1	TGCAATATCA	TAAGAATTGG	CTGGAAGAAG	AATGTATTGC	1320				
	hMR1P1		A							

#### MR1 genomic DNA amplification

We wanted to investigate whether the *MR1P1* pseudogene found in humans is also present in orangutan and chimpanzee genomes. To do this, PCR reactions using genomic DNA were carried out with primers used previously in our laboratory to amplify DNA fragments of the human *MR1* gene as well as *MR1P1* pseudogene, containing exon 2, intron 2, and exon 3.

In orangutan, a single band of 1156 nucleotides in length was observed (Fig. 4), DNA sequencing revealed that this PCR product corresponds to the MR1 gene fragment (Accession number AJ315656). MR1P1-related sequences were not found.

In chimpanzee two PCR products of 1320 and 1156 nucleotides in length respectively were observed (Fig. 4). The short product corresponded to the expected fragment of the *MR1* gene and was 98.87% homologous to the human MR1 segment. The long product was 98.6% homologous to human MR1P1. This result indicates that there are two MR1 sequences in chimpanzee as in humans. Comparison between chimpanzee and human MR1 fragments (Fig. 5a) reveals a very high degree of conservation with just one point mutation in exon 2 and one in exon 3, both of them synonymous. In contrast, 11 substitutions can be seen in the intron sequence. This situation is very different from that observed in the MR1P1-like fragment (Fig. 5b). In the intron, there are nine point substitutions as well as five deletions that differentiate human and chimpanzee sequences. In exon 2, there are six point substitutions, only two of which are synonymous. In exon 3, there are three point substitutions (one synonymous). Interestingly, the stop codon (position 151) present in human *MR1* is tryptophan in chimpanzee.

## Discussion

We have found that *MR1* orthologues exist in chimpanzee and orangutan. *MR1* cDNA isoforms generated by alternative splicing were observed in both species. Similar isoforms have also been found previously in human, rat, and mouse. Thus, the pattern of multiple cDNA isoforms generated through alternative splicing mechanisms is extended to the species analyzed here and reinforces interest in the possible functional implications of these isoforms.

Our results show that MR1 is a highly conserved molecule between primate and rodent species, especially in the  $\alpha$ 1 and  $\alpha$ 2 domains, the  $\alpha$ 3 domain and, in particular, the transmembrane and cytoplasmic regions showing higher variability. This indicates strong evolutionary pressure to maintain the  $\alpha$ 1 and  $\alpha$ 2 domains, suggesting that this molecule binds a very conserved ligand.

This pressure can be appreciated by comparing the chimpanzee and human genomic sequences we have obtained from the MR1 gene and MR1P1-like sequences. In the MR1 sequences, the differences between species accumulate inside intron 2, with the exons being almost identical in both species. In contrast, the differences be-

tween both species appear in both intronic and exonic sequences in *MR1P1*.

Interestingly, the stop codon found at position 151 in human *MR1P1* encodes a tryptophan in chimpanzee, suggesting the locus could be a functional gene. However, in our experiments with cDNA, we found no transcripts corresponding to chimpanzee *MR1P1*.

We did not find *MR1P1*-like sequences in orangutan. This could indicate that this locus does not exist in orangutan or that orangutan *MR1P1* cannot be amplified with the primers we used.

Multiple alignment of MR1 amino acid sequences from human, chimpanzee, orangutan, rat, and mouse demonstrates some interesting features with possible evolutionary and functional implications.

As already mentioned, there is a remarkable conservation of the  $\alpha 1$  and  $\alpha 2$  domains, with the  $\alpha 3$  domain being more divergent. The most intriguing differences are located in the transmembrane and cytoplasmic domains. The three primate MR1 sequences exhibit a positively charged amino acid (lysine) inside the putative transmembrane region which is converted to serine in mouse and asparagine in rat. In contrast, both rodent species show longer cytoplasmic tails, with a conserved serine-containing motif (NEGSS) similar to that found in HLA class I (SD/EXSL) and pig CD1 (CDPSS) molecules (Chum et al. 1999). In HLA class I molecules, this motif is necessary for constitutive endocytosis (Vega and Strominger 1989) and the serine residue can be phosphorylated (Guild and Strominger 1984). Although the function of MR1 is currently unknown, these features invite some speculations. MR1 possibly binds to a ligand that has been highly conserved through evolution, such as a peptide or a membrane protein, through the  $\alpha 1$  and  $\alpha$ 2 domains. MR1 is also conceivably involved in endocytosis and/or signal transduction. If so, the signal could be transduced using different mechanisms in rodents and primates. Rodents may phosphorylate the conserved serine in the cytoplasmic tail. In primates, this signal would be transduced using a coupled protein associated with MR1 through saline bridges between the lysine in MR1 and a negatively charged amino acid present inside the transmembrane region of the second protein, in a manner resembling immune receptors such as the T-cell receptor, some FcRs, or some natural killer receptors.

Acknowledgements This work was supported by grants from the Spanish Ministerio de Ciencia y Tecnología (BMC 2001-1382) and from Comunidad Autónoma de Madrid to E.M.-N. J.F.P.-C. is recipient of a predoctoral fellowship from Ministerio de Eduación y Ciencia. M.G.d.M. is recipient of a postdoctoral fellowship from Comunidad Autónoma de Madrid.

#### References

- Bahram S (2001) MIC genes: from genetics to biology. Adv Immunol 76:1–57
- Chum T, Wang K, Zuckermann FA, Gaskins HR (1999) Molecular cloning and characterization of a novel CD1 gene from the pig. J Immunol 162:6562–6571

- Cosman D, Müllberg J, Sutherland CL, Chin W, Armitage R, Fanslow W, Kubin M, Chalupny NJ (2001) ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytoxicity through the NKG2D receptor. Immunity 14:123–133
- Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, Domingo R, Ellis MC, Fullan A, Hinton LM, Jones NL, Kimmel BE, Kronmal GS, Lauer P, Lee VK, Loeb DB, Mapa FA, McClelland E, Meyer NC, et al (1996) A novel MHC class I-like gene is mutated in patients with hereditary hemochromatosis. Nat Genet 13:399–408
- Fukudome K, Esmon CT (1994) Identification, cloning and regulation of a novel endothelial cell protein C/activated protein C receptor. J Biol Chem 269:26486–26491
- Ghetie V, Ward WE (2000) Multiple roles of the major histocompatibility complex class I-related receptor FcRn. Annu Rev Immunol 18:739–766
- Guild BC, Strominger JL (1984) Human and murine class I MHC antigens share conserved serine 335, the site of HLA phosphorylation in vivo. J Biol Chem 259:9235–9240
- Hashimoto K, Hirai M, Kurosawa Y (1995) A gene outside the human MHC related to classical HLA class I genes. Science 269:693–695
- O'Callaghan CA, Bell JI (1998) Structure and function of the human MHC class Ib molecules HLA-E, HLA-F and HLA-G. Immunol Rev 163:129–138
- Parra-Cuadrado JF, Navarro P, Mirones I, Setién F, Oteo M, Martínez-Naves E (2000) A study on the polymorphism of

human MHC class I-related MR1 gene and identification of an MR1-like pseudogene. Tissue Antigens 56:170–172

- Pendás AM, Matilla JA, Uría JA, Freije JP, Fueyo A, Estivill X, López-Otin C (1994) Mapping of the human Zn-α2-glycoprotein gene (AZGP1) to Chr 7q22 by in situ hybridization. Cytogenet Cell Genet 66:263–266
- Porcelli SA, Modlin RL (1999) The CD1 system: antigen-presenting molecules for T cell recognition of lipids and glycolipids. Annu Rev Immunol 17:297–329
- Riegert P, Wanner V, Bahram S (1998) Genomics, isoforms, expression, and phylogeny of the MHC class I-related MR1 gene. J Immunol 161:4066–4077
- Vega MA, Strominger JL (1989) Constitutive endocytosis of HLA class I antigens requires a specific portion of the intracytoplasmic tail that shares structural features with other endocytosed molecules. Proc Natl Acad USA 86:2688–2692
- Walter L, Günther E (1998) Isolation and molecular characterization of the rat MR1 homologue, a non-MHC-linked class I-related gene. Immunogenetics 47:477–482
- Yamaguchi H, Hirai M, Kurosawa Y, Hashimoto K (1997) A highly conserved major histocompatibility complex class I-related gene in mammals. Biochem Biophys Res Commun 238: 697–702
- Yamaguchi H, Kurosawa Y, Hashimoto K (1998) Expanded genomic organization of conserved mammalian MHC class I-related genes, human MR1 and its murine ortholog. Biochem Biophys Res Commun 250:558–564