Immunogenetics (2004) 56:118–121
DOI 10.1007/s00251-004-0671-4

BRIEF COMMUNICATION

**Doris Lambracht-Washington · Kirsten Fischer Lindahl**

# Active MHC class Ib genes in rat are pseudogenes in the mouse

Received: 2 January 2004 / Revised: 16 March 2004 / Published online: 7 April 2004


**Abstract** The most telomeric class I region of the MHC in rat and mouse is the *M* region, which contains about 20 class I genes or gene fragments. The central part carries three class I genes—*M4, M5*, and *M6*—which are orthologous between the two species. *M4* and *M6* are pseudogenes in the mouse but transcribed, intact genes in the rat. To analyze the pseudogene status for the mouse genes in more detail, we have sequenced the respective exons in multiple representative haplotypes. The stop codons are conserved in all mouse strains analyzed, and, consistent with the pseudogene status, all strains show additional insertions and deletions, taking the genes further away from functionality. Thus, *M4* and *M6* indeed have a split status. They are silent in the mouse but intact in the closely related rodent, the rat.

**Keywords** *H2-M* region · *RT1.M* region · Orthologous genes · Interspecies comparison · Pseudogenes

The least studied class I region of the MHC in rat and mouse is the 1-Mb *M* region (Takada et al. 2003). Its central 30-kb part encodes three class I genes: *M6*, *M4,* and *M5*. Based on the conserved genes flanking this region, the map position of these genes is homologous to the 380-kb stretch around *HLA-A*, *-G*, and *-F* in the human MHC (Lambracht et al. 1995; Yoshino et al. 1998; Jones et al. 1999). Recently, new functions have been discovered for class Ib proteins encoded in the proximal *M* region, which associate with the V2R pheromone receptors of the vomeronasal organ (Loconto et al. 2003).

The class I genes of the central *H2-M* region in the BALB/c mouse ($H2^d$) were considered to be silent because *M4* and *M6* are pseudogenes, and no transcripts were detected for *M5*, which has an open reading frame (ORF) (Wang and Fischer Lindahl 1993). The analysis of this region in the closely related rodent, the rat, shows us a different status for these genes. *RT1.M4*, *M5*, and *M6* have ORFs, and transcripts were detected in several tissues (D. Lambracht-Washington, Y.F. Moore, K. Wonigeit, and K. Fischer Lindahl, manuscript in preparation). *H2-M4* and *-M6* are pseudogenes due to single nucleotide changes; for verification of the pseudogene status, we have analyzed *M4* exon 3 in nine mouse and ten rat strains and *M6* exon 4 in 14 mouse strains.

*H2-M4*$^d$ carries an early stop codon in exon 3 (Wang and Fischer Lindahl 1993). Generally, exons 2 and 3 of class Ia genes exhibit the most nucleotide differences, yet these exons of *M4* show a high degree of similarity, even between species (Fig. 1). To see whether the stop codon is conserved in the mouse, exon 3 was analyzed in nine strains and seven haplotypes (Fig. 1). The mouse *M4* alleles show only minor differences, with conservation of the in-frame stop codon at the beginning of exon 3 in all strains analyzed, even in the three haplotypes from wild mice of different species: *cas3*, *sh1*, and *sp2*. The *RT1.M4* exon 3 sequences all showed a change from the in-frame stop codon of the mouse to TGG (tryptophan). Due to a nucleotide insertion, *RT1* haplotypes *l* and *lv3* carry a different stop codon at the end of exon 3, whereas all other analyzed *RT1* haplotypes possess an ORF for this exon. Exon 2 was also sequenced in haplotypes *c* and *n* and found to be an ORF as well and identical to exon 2 of *l*.

*H2-M6*$^d$ is a pseudogene due to a single nucleotide deletion in exon 4 (Wang and Fischer Lindahl 1993). We sequenced that exon in 13 other strains. As the nucleotide

GenBank accession numbers: AF057065 to AF057072 (exon 3 of *H2-M4* of reported mouse strains), AF057976 to AF057985 (exon 3 of *RT1.M4* of reported rat strains), AF058923 and AF058924 (exon 2 of *RT1.M4* of strains PVG and BN), AY286080 to AY286092 (exon 4 of *H2-M6* of reported mouse stains), and AY303772 (full-length genomic sequence of *RT1.M6-1*$^l$)

D. Lambracht-Washington (✉) · K. Fischer Lindahl
Center for Immunology,
University of Texas Southwestern Medical Center,
Dallas, TX 75390-9050, USA
e-mail: doriswashington@yahoo.com

K. Fischer Lindahl
Howard Hughes Medical Institute,
University of Texas Southwestern Medical Center,
Dallas, TX, 75390-9050, USA

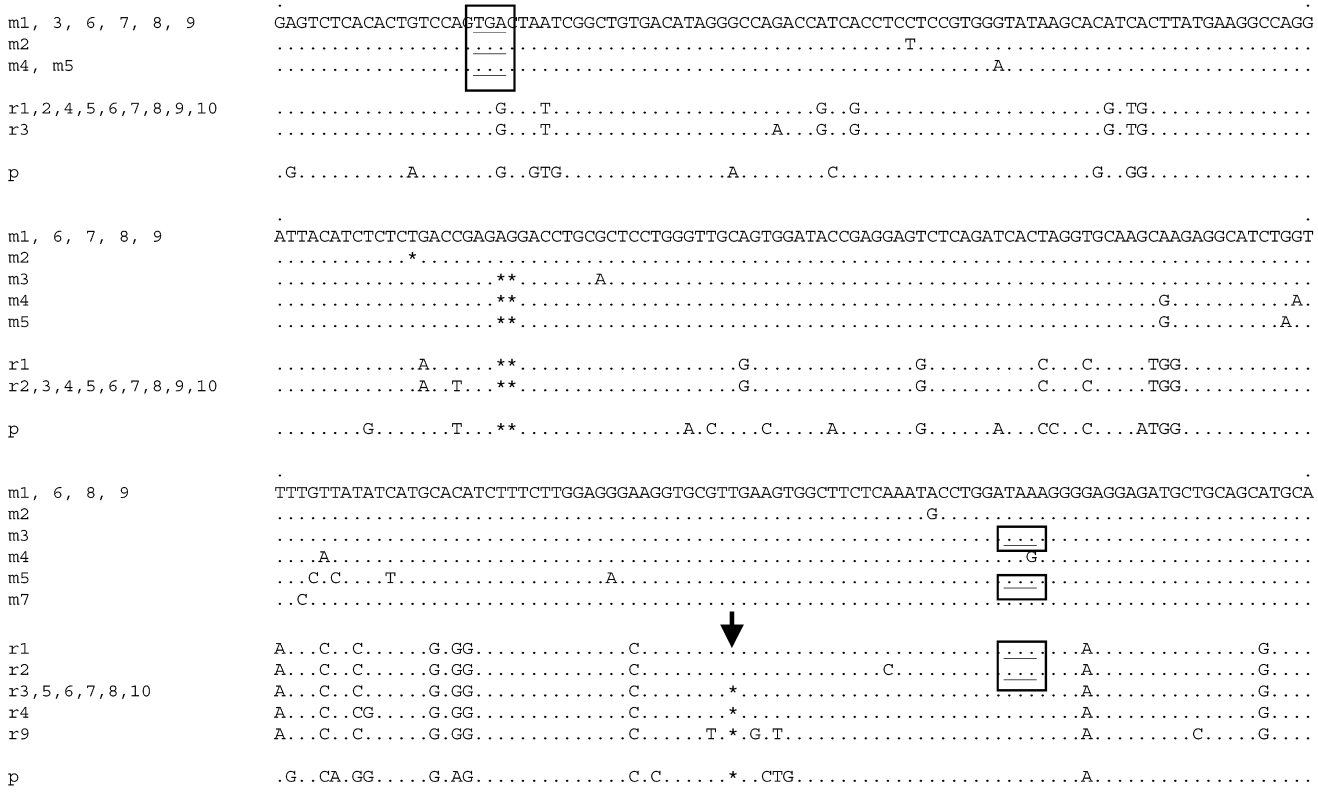


**Fig. 1** Comparison of exon 3 of *M4* in various *H2* and *RT1* haplotypes and in *Peromyscus leucopus* (p). The exon 3 sequence is divided into three fragments, and for each fragment the strains carrying a given sequence are indicated on the *left*. Only the exon sequence is shown. Potential stop codons are *underlined* and *boxed*. An *arrow* marks the nucleotide insertion in *RT1.M4* causing the stop codon in haplotypes *l* (LEW) and *lv3* (LEW.1LV3). *Asterisks* indicate alignment gaps. Analyzed were nine mouse strains: *m1* BALB/c ($H2^d$), *m2* DBA/2 J ($H2^d$), *m3* BALB-*B2m*w3/Kfl ($H2^d$), *m4* C3H/HeJ ($H2^k$), *m5* B10.SH1(R27)/Kfl ($H2^{sh1}$), *m6* C57BL/10SnJ ($H2^b$), *m7* A.CA/J ($H2^f$), *m8* B10.SP2(R40)/Kfl ($H2^{sp2}$), and *m9* B10.CAS3/Kfl ($H2^{cas3}$); and ten rat strains: *r1* LEW ($RT1^l$), *r2* LEW.LV3 ($RT1^{lv3}$), *r3* F344 ($RT1^{lv1}$), *r4* LEW.1 N ($RT1^n$), *r5* BN ($RT1^n$), *r6* BN.1B ($RT1^{r37}$), *r7* LEW.1C ($RT1^c$), *r8* PVG ($RT1^c$), *r9* DA ($RT1^{av1}$), and *r10* BDE ($RT1^u$). For the exon 2 and 3 sequences of LEW, a cosmid clone (Lambracht et al. 1995) and genomic DNA were analyzed in parallel. The published *H2-M4* sequence (L14278) is derived from cosmid clones of a BALB/c subline (Wang and Fischer Lindahl 1993). The *sp2* haplotype comes from *Mus spretus*, the *cas3* haplotype from *M. m. castaneus*, and the *sh1* haplotype from a wild mouse from Shanghai (Fischer Lindahl 1994). Exon 3 of *M4* was amplified by PCR. The forward exon 3 primer (from intron 2), 5′CTCAA*GGATCC*ATAGAACTACCC3′, was identical for mouse and rat; the reverse primers (from intron 3) were mouse, exon 3 reverse: 5′GGACATG*GAATTC*ACCACTTTGGC3′; and rat, exon 3 reverse: 5′GGACAC-G*GAATTC*ACCTCTTTGG3′. Primers were designed with recognition sites for restriction enzymes (*italics*) to facilitate cloning of the PCR products into M13 in both directions. The PCR cycle protocol was as follows: 5 min denaturation at 94°C, followed by 30 cycles of 3 min annealing and polymerization at 65°C and 1 min denaturation at 94°C. To minimize PCR errors, five to ten clones were pooled for DNA isolation, and at least two independent PCRs were done and sequenced for each *M4* allele

deletion is conserved in all haplotypes analyzed, we confirmed the pseudogene status of *H2-M6* (Fig. 2). The overall sequence for the exon encoding the $\alpha$3 domain showed variability among the 12 haplotypes (*a*, *b*, *bac1*, *cas2*, *cas3*, *d*, *f*, *k*, *k2*, *r*, *sh1*, *sp2*), which was not seen in the two *H2-M6* sequences (*d*, *bc*) in the database. Only the sequences of strains B6 ($M6^b$) and DBA/2 ($M6^d$) were identical to the database sequences. Strains A.CA ($M6^f$) and the Asian haplotypes of B10.SH1(R27) ($M6^{sh1}$) and B10.BAC1 ($M6^{bac1}$) had identical sequences, which showed a number of nucleotide changes relative to the database sequences. In B10.SP2(R40) ($M6^{sp2}$), the consensus splice site in the beginning of exon 4 was missing. B10.BR ($M6^{k2}$), C3H ($M6^k$), LP.RIII ($M6^r$), B10.CAS2 ($M6^{cas2}$), B10.CAS3 ($M6^{cas3}$), and A/J ($M6^a$) carried an early stop codon after 8, 14, 15, or 16 amino acids. This variability in the generally conserved exon 4 is consistent with the pseudogene status of *M6* in the mouse.

In a related rodent, *Peromyscus leucopus*, the *M4* gene is intact in all inbred lines analyzed and exhibits intraspecies genetic polymorphism (Crew and Bates 2003). *Peromyscus* and *Mus* separated 40–60 million years ago; *Mus* and *Rattus* separated 10–20 million years ago. The frame-shift mutation in exon 3 of all *H2-M4* alleles examined represents an insertion of a single nucleotide relative to the *Peromyscus* sequence. The presence of the same insertion in two of ten *RT1.M4* alleles suggests that *M4* was functional in primitive rodents that gave rise to *Mus*, *Rattus*, and *Peromyscus* (Crew and Bates 2003), and that it was already dimorphic for the frame-shift mutation in the *Mus*/*Rattus* precursor population. Subsequent to the split of *Rattus* and *Mus*, the *H2-M4* gene acquired addi-

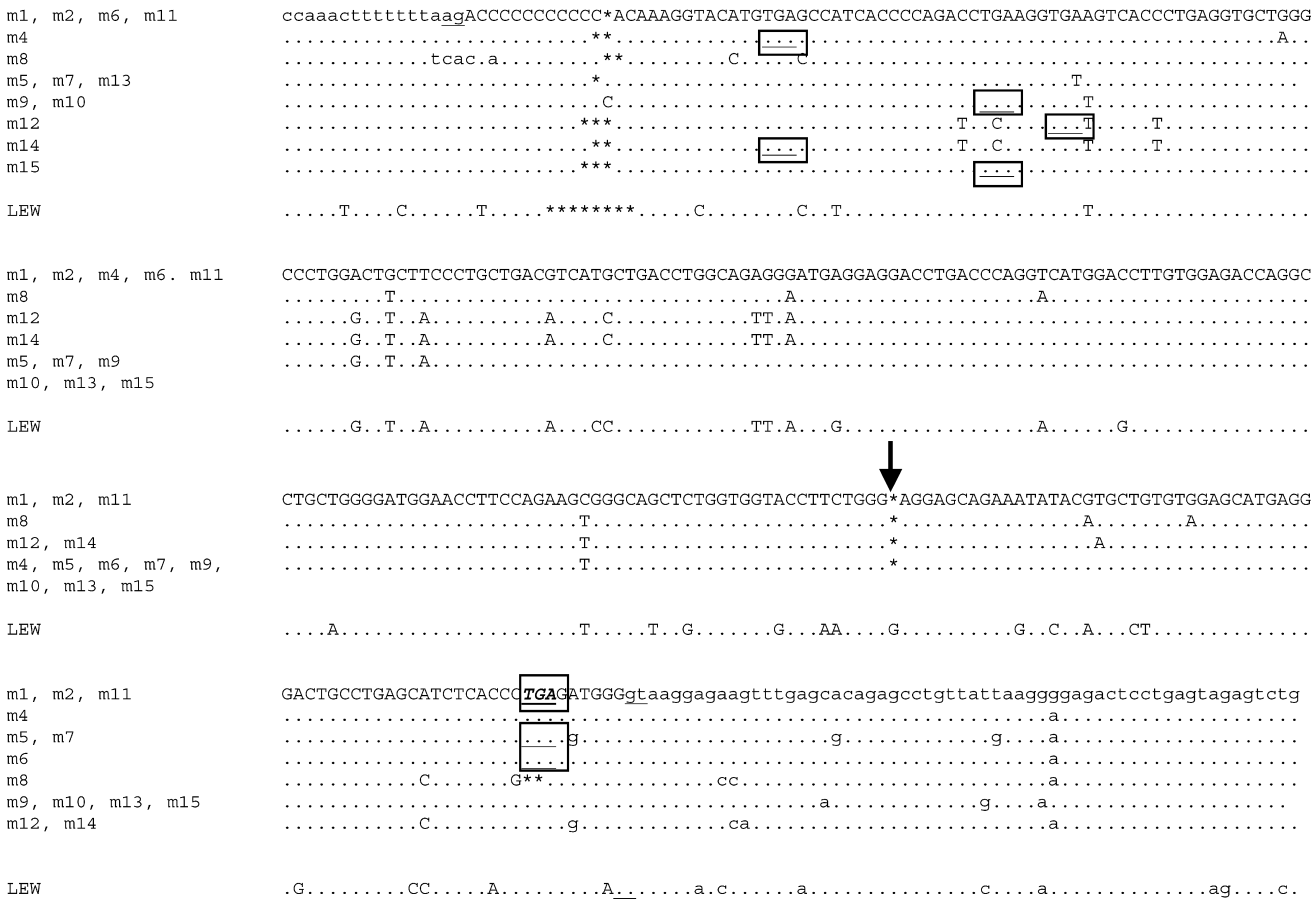


**Fig. 2** Comparison of exon 4 of *M6* in various *H2* haplotypes with *RT1.M6-1* of the LEW rat. The exon 4 (*capital letters*) and surrounding intron (*lower case letters*) sequences are divided into four fragments, and for each fragment the strains carrying a given sequence are indicated on the *left*. An *arrow* marks the nucleotide deletion in *H2-M6* that causes a stop codon. Potential stop codons are *underlined* and *boxed*. Analyzed were 14 mouse strains: those of Fig. 1, except BALB-$B2m^{w3}$/Kfl (*m3*), and in addition the following strains: *m10* B10.CAS2/Kfl ($H2^{cas2}$), *m11* C57BL/6 J ($H2^{b}$), *m12* A/J ($H2^{a}$), *m13* B10.BAC1($H2^{bac1}$), *m14* B10.BR ($H2^{k2}$), and *m15* LP.RIII/J ($H2^{r}$). The LEW sequence is derived from an *RT1.M6-1* genomic cosmid clone (Lambracht et al. 1995). For amplification of *H2-M6* exon 4, we used the primer pair *H2-M6* ex4 forward (5′CTCATCTTGATTCTCCTGTTCATT3′) and *H2-M6* ex4 reverse (5′CCTAGCACAGACTCTACT3′), located in introns 3 and 4, respectively, and the following PCR protocol: an initial DNA denaturation step for 3 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 54°C, and 1 min at 72°C, and a final amplification step for 10 min at 72°C. The PCR products were analyzed by gel electrophoresis and cloned by TA cloning into the pCR1 vector (Invitrogen, Carlsbad, Calif.). DNA was extracted from four to ten clones, and the individual clones were sequenced automatically

tional single-base mutations, insertions, and deletions, which further distanced the gene from functionality. Several of these mutations appear in some haplotypes but not others; for example, a dinucleotide insertion is present in three of nine *H2* haplotypes (Fig. 1). For *H2-M6* we propose a similar scenario. The deletion of a single nucleotide in exon 4 caused the loss of functionality. Subsequently, additional deletions and nucleotide insertions occurred which are all, except two, located in the poly-C stretch in the beginning of the exon (Fig. 2). These alterations caused the appearance of additional stop codons in exon 4, further silencing the gene.

**Acknowledgement** This work was supported in part by NIH grant AI 37818.

## References

Crew MD, Bates LM (2003) Sequence, expression, and polymorphism of the *Peromyscus leucopus Mhc* class Ib gene, *M4*. Immunogenetics 55:95–99

Fischer Lindahl K (1994) Treasures of the Orient: Qa1, Mta, and $\beta$2m polymorphisms in Asian wild mice. Moriwaki K, Shiroishi T, Yonekawa H (eds) Genetics in wild mice: its application to biomedical research. Japan Scientific, Tokyo and Karger, Basel, pp 179–191

Jones EP, Kumánovics A, Yoshino M, Fischer Lindahl K (1999) MHC class I and non-class I gene organization in the proximal *H2-M* region of the mouse. Immunogenetics 49:183–195

Loconto J, Papes F, Chang E, Stowers L, Jones EP, Takada T, Kumánovics A, Fischer Lindahl K, Dulac C (2003) Functional expression of murine V2R pheromone receptors involves selective association with the M10 and M1 families of MHC class Ib molecules. Cell 112:607–618

Lambracht D, Prokop C, Hedrich HJ, Fischer Lindahl K, Wonigeit K (1995) Mapping of *H2-M* homolog and *MOG* genes in the rat MHC. Immunogenetics 42:418–421

Takada T, Kumánovics A, Amadou C, Yoshino M, Jones EP, Athanasiou M, Evans GA, Fischer Lindahl K (2003) Species-specific class I gene expansions formed the telomeric 1 Mb of the mouse major histocompatibility complex. Genome Res 13:589–600

Wang C-R, Fischer Lindahl K (1993) Organization and structure of the *H2-M4-M8* class I genes in the mouse major histocompatibility complex. Immunogenetics 38:258–271

Yoshino M, Xiao H, Jones EP, Fischer Lindahl K (1998) BAC/YAC contigs from the *H2-M* region of mouse chromosome 17 define gene order as *Znf173-Tctex5-Mog-D17Tu42-M3-M2*. Immunogenetics 47:371–380