Sequence Convergence in the Peptide-Binding Region of Primate and Rodent MHC Class Ib Molecules

Meredith Yeager, Sudhir Kumar, and Austin L. Hughes

Department of Biology and Institute of Molecular Evolutionary Genetics, The Pennsylvania State University

In addition to the universally expressed and highly polymorphic class Ia genes, the major histocompatibility complex (MHC) of placental mammals includes class Ib genes that are characterized by restricted expression and low levels of sequence polymorphism. The functional importance of class Ib genes as well as their actual function has long been controversial. Phylogenetic analyses have suggested that there are no orthologous relationships among class Ib loci of mammals belonging to different orders, suggesting that these loci have evolved independently since the placental mammals diverged. Here, we present evidence of convergent evolution at the molecular sequence level in the putative peptide-binding regions (PBRs) of human and mouse class Ib genes. So far, there are few if any convincing examples of convergent evolution at the amino acid sequence level, and such evolution is believed to be likely to occur only as a result of strong positive selection. Because the present case involves the functionally important PBR and because the primate and rodent molecules are known to bind similar peptides, this study represents both a convincing case of molecular-level convergence and evidence that MHC class Ib molecules, although not orthologous, may evolve similar functions convergently.

Introduction

The major histocompatibility complex (MHC) of vertebrates is a multigene family whose products encode cell surface glycoproteins responsible for presenting peptides to T cells (Klein 1986). There are two major subfamilies: class I and class II. MHC class I genes are further divided into the classical (class Ia) and the nonclassical (class Ib) gene categories. The class Ia molecules serve to present peptides to cytotoxic T cells. They have a near-universal pattern of tissue expression and are known to be highly polymorphic in their amino acid sequences. This polymorphism is maintained by a form of balancing selection (such as overdominant selection) that enhances the rate of nonsynonymous (amino-acid-altering) nucleotide substitution in the codons encoding the peptide-binding region of the molecule (Hughes and Nei 1988; Hughes and Hughes 1995). The products of the class Ib genes show high amino acid sequence similarity to the class I genes. However, class Ib genes are characterized by a much narrower tissue expression and much lower sequence polymorphism (Klein and O=hUigin 1994). There is no evidence of balancing selection at class Ib loci (Hughes and Nei 1989), and, unlike the class Ia gene products, it is not clear whether all class Ib gene products are functional.

Klein and Figueroa (1986) and Howard (1987) suggested that class Ib genes represent a kind of "expressed pseudogenes"—that is, genes that are on their way to becoming pseudogenes. Some recent studies have suggested that class Ib products have the potential to present peptides (Pamer et al. 1992; Joyce et al. 1994). In fact, one mouse class Ib gene product (H2-Qa-2), although monomorphic, has been shown to bind a diverse array of self-peptides (Joyce et al. 1994). Therefore, at least some class Ib gene products may be functional.

In humans, class Ia genes include *HLA-A*, *-B*, and *-C*, while in the mouse, class Ia gene products are encoded by *H2-K*, *-D*, and *-L*. Class Ib genes in humans include *HLA-E*, *-G*, *-F*, *MICA*, and *MICB*. Class Ib genes in mouse include *H2-M*, *-Q*, and *-T*. Orthologous relationships are sometimes found among class I loci of mammals of the same order but never among mammals of different orders (Hughes and Nei 1989). That is, class I genes of human (primate), mouse (rodent), and cow (artiodactyl) are not orthologous. Further, in phylogenetic analyses, the class Ib genes of one order cluster with the class Ia genes of that order (Hughes and Nei 1989). This type of clustering pattern suggests that the class Ib genes have arisen independently by gene duplication in different orders of mammals and is consistent with the hypothesis that nonclassical genes originate from classical genes but lose classical class I expression levels due to deleterious mutation (Hughes and Nei 1989; Hughes 1995).

Previous analysis has shown that a class Ib histocompatibility antigen of the mouse (H2-Qa-1^a) has strong sequence similarity with the product of the mouse *T23^d* gene (H2-Qa-1; Connolly et al. 1993). Although nomenclature of these two proteins suggests that they are allelic products of the same gene, they appear not to be, although they are encoded within the same *TL* region (Connolly et al. 1993). Fragments of the mycobacterial stress protein hsp65 have been observed to stabilize the cell surface expression of H2-Qa-1^b (Imani and Soloski 1991). Since stress proteins of the hsp65 family elicit a strong T-cell response (Lamb et al. 1989) and an element of the response includes γδ TCR molecules (Janeway, Jones, and Hayday 1988; O'Brien et al. 1989), it is possible that Qa-1 molecules play a role in the immune response in tissues where γδ TCR molecules are commonly found (Connolly et al. 1993). Interestingly, Connolly et al. (1993) observed that the Qa-1 molecules share two residues in the peptide-bind-

Key words: convergent evolution, HLA-E, MHC class Ib, peptide binding, positive selection.

Address for correspondence and reprints: Austin L. Hughes, Department of Biology, 208 Mueller Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802. E-mail: austin@hugaus3.bio.psu.edu.

Mol. Biol. Evol. 14(10):1035–1041. 1997

Table 1
Class Ia and Ib Sequences Analyzed in this Study

Sequence	Accession Number
Class Ia	
Rodent	
Mouse: H2-K^k	U47330
Mouse: H2-K^d	U47329
Mouse: H2-D^d	U47326
Mouse: H2-D^p	M12381
Primate	
Human: HLA-B*3501	U11265
Human: HLA-B*5701	M32318
Human: HLA-Cw*1401	X58536
Human: HLA-Cw*0801	M84174
Artiodactyl	
Bovine	M21044
Sheep	M34676
Marsupial	
Wallaby: Maru-UA01	L04950
Wallaby: Maru-UB01	L04952
Aves	
Quail	D29813
Chicken	M31012
Class Ib	
Rodent	
Mouse: H2-Q4	M18837
Mouse: H2-Q10^k	X16426
Mouse: H2-Q8/9^d	M73272
Mouse: H2-Q1^k	X16424
Mouse: H2-Qa-1^a	L00606
Mouse: H2-T23^d	Y00629
Primate	
Macaque: Mafa-E1	U02976
Macaque: Mafa-E2	U02977
Human: HLA-E	M21533
Orangutan: Popy-E	M30681

ing region (PBR) with human HLA-E and orangutan Popy-E class Ib molecules.

The reason and origin for the existence of these shared PBR residues between rodent and primate nonorthologous class Ib molecules are not known. The PBR of the class I MHC molecule includes six "pockets" (A–F) into which side chains of the residues of the bound peptide (usually a nonamer) fit (Saper, Bjorkman, and Wiley 1991). Differences in residues in these pockets are believed to affect peptide binding, particularly in the B pocket, into which the second residue of the peptide fits, and the F pocket, into which the C-terminal residue of the peptide fits. The B and F pockets serve as the main anchor residues in most class Ia molecules, and they bind positions P2 and P9 of the peptide, respectively (Rammensee, Friede, and Stevanovic 1995). The P2 and P9 "preferred residues" for a given allele tend to have similar properties (e.g., positively charged). The other pockets appear to allow a wide range of different residues of the peptide for binding; the positions bound by the remainder of the pocket vary, and it does not appear that similar properties of the residues at a given position are necessary. Since the similarities between the primate E and rodent Qa-1 gene products are located within the F pocket, these genes may have similar functions. Indeed, recent analyses have shown that the HLA-E and H2-Qa1^b molecules can bind similar peptides that are derived from class I leader sequences (Kurepa and Forman 1997; Braud, Jones, and McMichael 1997; DeCloux et al. 1997). The aim of the present study was to address whether this sharing was the result of convergent evolution using phylogenetic analysis and a method of ancestral-state reconstruction.

Methods

Class Ia and Ib sequences from mice, humans, and other mammals were obtained from GenBank. Names and database accession numbers of sequences used in phylogenetic analyses are given in table 1. Sequences were aligned using the CLUSTAL program (Higgins, Bleasby, and Fuchs 1992). Because the similarity of Qa-1 and primate E molecules involves residues in the PBR, we constructed phylogenetic trees from class Ia and class Ib genes from mouse and primates for the 63 residues involved in the PBR (Bjorkman et al. 1987*a*, 1987*b*; Saper, Bjorkman, and Wiley 1991) and for the remainder of the three extracellular domains (α1–α3). Among the 63 residues included in the PBR, there are 35 that make up the six pockets involved in peptide binding. The remainder of the 63 residues are involved in T-cell receptor binding and PBR structure. The following methods of phylogenetic reconstruction were used to estimate the topologies of the trees and produced similar topologies: (1) neighbor joining (Saitou and Nei 1987) based on both the proportion of amino acid difference (p) and amino acid distance based on the Poisson correction (Nei 1987); (2) maximum likelihood (Felsenstein 1981); and (3) maximum parsimony (Swofford 1993). We chose to present trees based on the proportion of amino acid difference (p) because the variance is expected to be smaller, which should lead to a more accurate distance estimate when the number of sites is small and the number of substitutions per site is large (Kumar, Tamura, and Nei 1993), as is true in some comparisons in this case. Statistical significance of internal branches was tested by Rzhetsky and Nei's (1992) standard error test.

In order to examine to what extent residues are shared and to ascertain whether the observed sequence similarity between the E (primate) and E-like (mouse) groups was due to conservation of ancient sequence or convergent (or parallel) evolution, ancestral sequences for internal nodes of the neighbor-joining tree were inferred by maximum parsimony and by the maximum-likelihood approach under an empirical model of amino acid substitution (Yang, Kumar, and Nei 1995). This method's output assigns each internal node of a given tree "posterior probabilities" for each site in the data set. The posterior probability is the probability of the amino acid residue reconstructed at a given site based on the given tree, the substitution model, and the observed residue frequencies.

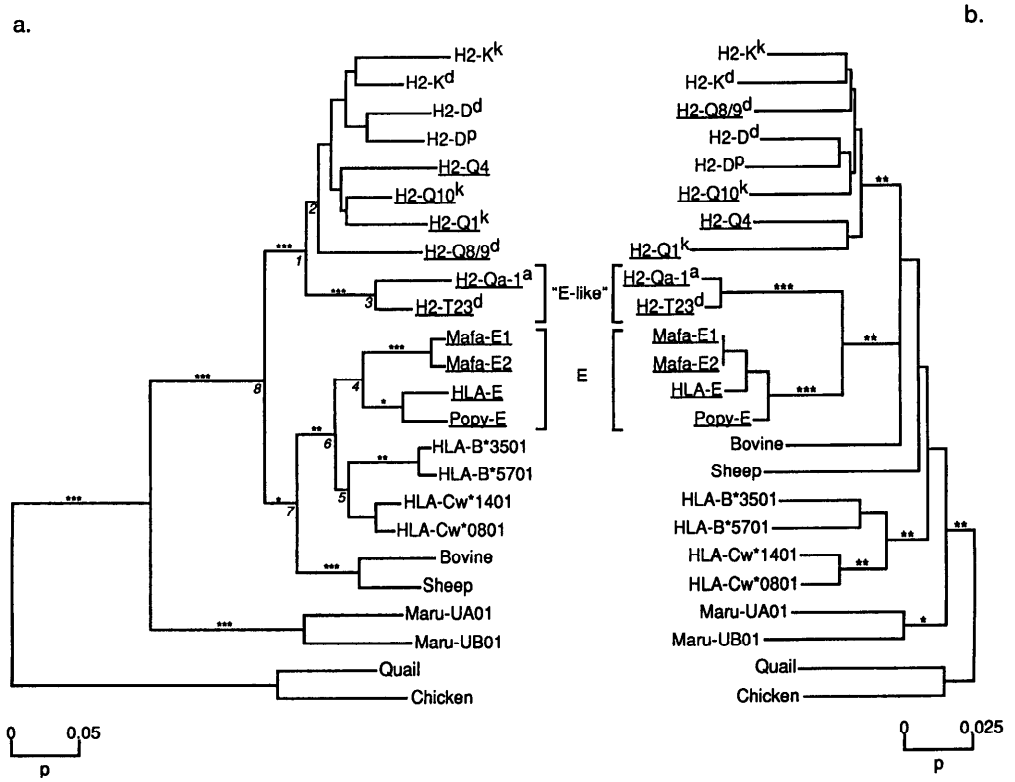

FIG. 1.—Convergence of amino acid sequence in MHC genes. In phylogenetic trees (*a* and *b*), allele designations are as follows: HLA-, human; H-2, mouse; Mafa-, crab-eating macaque; Popy-, orangutan; Maru-, red-necked wallaby. Class Ib sequences are underlined. Neighbor-joining (Saitou and Nei 1987) trees are shown. *a,* Tree based on proportion of amino acid differences (p) in the 211 amino acids from exons 2–4, excluding the 63 residues which are involved in antigen presentation (Bjorkman et al. 1987*a,* 1987*b*; Saper, Bjorkman, and Wiley 1991). The groups in which convergent evolution is hypothesized to have occurred are bracketed. *b,* Tree based on proportion of amino acid difference (p) in only the 63 PBR residues. Tests of the significance of internal branches (Rzhetsky and Nei 1992): * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. All sequences were obtained from GenBank.

Results

In the tree constructed using the amino acid sequence excluding the PBR residues, the mouse E-like (H2-Qa-1^a and H2-T23^d) molecules form a cluster with the remaining mouse genes, and the primate E-locus genes cluster with the same significance with the remaining human genes ($P < 0.001$) (fig. 1*a*). The topology of this tree is consistent with previous phylogenetic analysis. Class Ib genes cluster with the class Ia genes of the same order of mammals (Hughes and Nei 1989), and the topology is consistent with the one predicted for the different mammalian orders (Li et al. 1990). By contrast, in the tree based on only the 63 PBR residues, the primate E-locus molecules were clustered significantly ($P < 0.01$) with the rodent E-like molecules (fig. 1*b*). This is inconsistent with previous phylogenetic analysis, in that class Ib genes of different orders are more closely related to each other than to their corresponding class Ia genes. Also, the artiodactyl sequences group with the mice sequences, although not significantly; therefore, this could be due to stochastic error.

Variable sites of the reconstructed ancestral sequences as inferred by the maximum-likelihood approach (Yang, Kumar, and Nei 1995) are shown in figure 2. The maximum-parsimony method of reconstruction gave similar results; however, in many cases, more than one amino acid state was inferred. In such cases, the maximum-likelihood method was used to decide among the possibilities. Individual residues shared by E and E-like sequences were categorized as follows: (1) A residue was convergent (indicated by striped bars in fig. 2) if a change to that residue occurred both from the primate ancestor (node 1) to the ancestor of the E sequences (node 3) and from the rodent ancestor (node 6) to the ancestor of the E-like sequences (node 4). (2) A residue was retained (indicated by black bars in fig. 2) if it occurred in the E (node 3) and E-like (node 4) ancestors and in the ancestral placental mammal sequence (node 8) but not at both node 2 (other primate sequences besides E) and node 5 (other mouse sequences besides E-like).

Six cases were observed in which residues that were shared between the primate E and mouse E-like ancestral sequences were found to be the result of convergent evolution. Nine other residues were observed to be retained in the E and E-like sequences from the placental mammal ancestral sequence. Only one residue was found to be convergent between the other mouse and human ancestral sequences, whereas 24 were found to be retained. The proportion of convergent-shared residues within the E and E-like sequences is significantly different ($P < 0.007$, Fisher's Exact Test) from that of

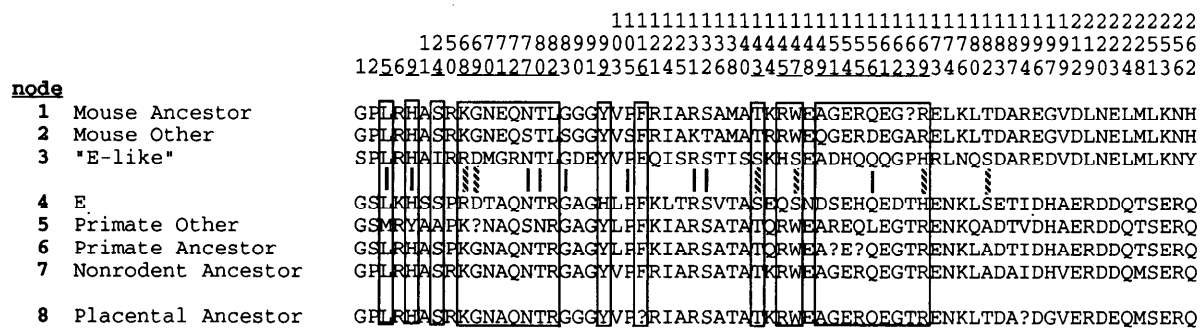

FIG. 2.—Estimated ancestral sequences for selected internal nodes as inferred by the maximum-likelihood approach under an empirical model of amino acid substitution (Yang, Kumar, and Nei 1995). The maximum-parsimony method of reconstruction gave similar results; however, in many cases, more than one amino acid state was inferred. In such cases, the maximum-likelihood method was used to discern among the possibilities. Only variable sites within these reconstructed sequences are shown. Sites are numbered according to coding amino acid sequence. Residues included within the PBR of the molecule are boxed. "?" denotes instances where the ancestral sequence could not be reconstructed with a posterior probability of greater than 50%. Black bars between the *E*-like and *E* ancestral sequences designate cases of retention of ancestral sequence within these lineages. Striped bars represent those residues in which sequence convergence is inferred. The number of convergent sites (6) is greater than would be likely to occur by chance alone, since, at most, two such sites are expected to be found in sequences of this length compared between mammals from different orders (unpublished simulations).

the proportion shared by the other primate and mouse sequences.

Five of the six sites experiencing convergent change within the E and E-like sequences (residues 68, 69, 143, 147, and 169) are located within the PBR of the molecule (table 2). The posterior probabilities of the reconstructions for these sites were, in most cases, very high (table 3).

Discussion

These results are significant for several reasons. First, they indicate that the sequence similarity observed within the primate E and mouse E-like sequences has arisen due to convergent evolution at different amino acid sites. Although there are some cases in which convergence has been speculated, it is a phenomenon that has previously been difficult to prove (Doolittle 1994; Kreitman and Akashi 1995; but see Stewart, Schilling, and Wilson 1987; Yokoyama and Yokoyama 1990; Kornegay, Schilling, and Wilson 1994). The residues found to be due to parallel/convergent evolution between the E and E-like sequences were very rare among other class I sequences. The E and E-like extant sequences are the only known class I sequences containing arginine at position 68 among 244 class I vertebrate sequences surveyed, and aspartic acid at position 69 is found in the E and E-like sequences but in only 6.3% of these other sequences. The crystal structure of the complete human T-cell receptor (TCR), bound peptide, and MHC molecule has recently been elucidated, and residues 68 and 69 were shown to be involved in TCR binding and are buried by the variable loop CDRα (Garboczi et al. 1996). Since arginine at position 68 and aspartic acid at position 69 are both charged residues (+ and −, respectively), they may indeed be extremely important for TCR binding to E and E-like molecules. Residues 143 and 147 are located within the F pocket of the PBR, which serves to anchor the C-terminal end of the presented peptide (Saper, Bjorkman, and Wiley 1991; Ram-

Table 2
Five of the Six Residues Found to be Convergent Between E and E-like Sequences are Located Within the PBR

Residue	Amino Acid	Position[a]	Frequency[b]
68.....	Arg	TCR-accessible	Unique
69.....	Asp	TCR-accessible	6.3%
143.....	Ser	F pocket	2.7%
147.....	Ser	F pocket	0.4%
169.....	His	TCR-accessible	3.4%

NOTE.—The proportion of convergent shared residues within the *E* and *E*-like sequences is significantly different ($P < 0.007$, Fisher's Exact Test) from the proportion shared by the other primate and mouse sequences.

[a] Position within the peptide-binding region (PBR) as defined by Bjorkman et al. (1987*a*, 1987*b*), Saper, Bjorkman, and Wiley (1991), and Garboczi et al. (1996). "TCR-accessible" indicates that the residue is located on one of the two α-helices and is accessible to the T-cell receptor.

[b] Frequency among 244 class Ia and Ib sequences representing the mammalian orders Rodentia, Artiodactyla, Carnivora, Lagomorpha, Primates, and Perissodactyla. "Unique" refers to an instance where the observed amino acid is unique to the *E* and *E*-like sequences.

Table 3
Posterior Probabilities for Retained and Convergent Residues Within the PBRs of E and E-like Ancestral Sequences

Site	Amino Acid (*E*-like/*E*)[a]
5.....	L (0.602/0.602)
9.....	H (0.906/0.892)
68.....	**R (0.957/0.957)**
69.....	**D (0.620/0.620)**
77.....	N (0.725/0.725)
80.....	T (0.921/0.921)
83.....	G (0.793/0.793)
105.....	P (0.936/0.936)
131.....	R (0.830/0.830)
132.....	S (0.852/0.852)
143.....	**S (0.973/0.973)**
147.....	**S (0.999/0.999)**
156.....	Q (0.676/0.676)
169.....	**H (0.529/0.692)**

[a] Posterior probability of occurrence of residue at given site as inferred by Yang, Kumar, and Nei's (1995) method of ancestral sequence reconstruction. Reconstrutions with <50% posterior probabilities are not included and were not considered in the results. The residues found to be convergent are shown in bold type.

Table 4
F-Pocket Residues in E and E-like Sequences

	Amino Acid Position								
Sequence	77*	80*	81	84	116	123	143***	146	147***
H2-Qa1[a].....	N	T	L	Y	*E*	Y	S	K	S
H2-T23[d].....	N	T	L	Y	*E*	Y	S	K	S
HLA-E.....	N	T	L	Y	*F*	Y	S	K	S
MAFA-E1....	N	T	L	Y	*F*	Y	S	K	S
MAFA-E2....	N	T	L	Y	*F*	Y	S	K	S
POPY-E.....	N	T	L	Y	*F*	Y	S	K	S

Note.—The F pocket serves to anchor the C-terminal end of the peptide into the binding site. * = sites found to be retained from ancestral sequences for both *E* and *E*-like sequences; *** = sites found to be the result of convergent evolution. Residues which differ between rodent and primate are italicized.

mensee, Friede, and Stevanovic 1995). Serine at these positions is quite rare; it is found at position 143 in less than 3% of the 244 sequences, while only one other sequence, a rabbit class I gene which may also be nonclassical (Tykocinski et al. 1984), contained serine at position 147. In fact, both of these positions are conserved in most other class I alleles known (Thr143 and Trp147). In this respect, the presence of serine in the rabbit sequence at both position 143 and position 147 is interesting. Two of the residues that were retained within the E and E-like sequences from the placental mammal ancestor (residues 77 and 80) are also located within the F pocket. The sharing of these two residues, along with the two found to be convergent (143 and 147), and other residues that are identical between the primate E and mouse E-like sequences strongly suggests similarity of function for the E and E-like molecules at the F pocket (table 4). There is only one difference among the nine residues comprising F pockets of primate E and rodent E-like sequences (Asp116 → Phe116). The histidine at position 169 is present in only 3.4% (excluding the E and E-like sequences) of the 244 class Ia and Ib sequences surveyed. Thus, all of the convergent residues within the PBRs of the E and E-like sequences are rarely found in other mammalian class I sequences. Table 5 includes all 63 PBR-related residues. Residues are marked designating location in each of the six pockets, as well as those that are TCR-directed. Between mouse H2-Qa-1[a] and human HLA-E, there are 41 identical sites. Of the nonidentical sites, 9 do not cause a charge change. There are 13 differing sites that do have charge changes; however, of these sites, only 2 are located in the B pocket (positions 66 and 99) and 1 in the F pocket (position 116). Since the repertoire of residues found to be located at P1 and P3–P8 is generally very broad (Rammensee, Friede, and Stevanovic 1995), the presence of charge changes outside the B and F pockets is not surprising.

The presence of a significant proportion of shared residues found within the functionally important PBR, particularly within the F pocket, is consistent with the findings that the E and E-like sequences bind similar peptides. HLA-E and H2-Qa1[b] both bind conserved leader sequences from class Ia molecules (table 6). In the case of H2-Qa1[b], leader sequences from H2-D and H-2L were shown to be bound by the molecule (Kurepa and Forman 1997; DeCloux et al. 1997). Likewise, HLA-E can bind the same leader peptide from HLA-A and HLA-B molecules.

The site changes within the E and E-like sequences include residues that are very rarely found in other class I sequences. The B and F pockets are most important for determining the range of peptides that a given class I molecule can bind, and the F pocket seems to be important for peptide-binding specificity in far more cases than the B pocket. Rammensee, Friede, and Stevanovic (1995) summarize information on peptide motifs for 42 class I products from human and mouse. In 13 (31%) of these, the residue at P2 of the peptide (bound by the B pocket) and at P9 of the peptide (bound by the F pocket) are highly constrained. In 4 (9%), the P2 residue is highly constrained but P9 is not. However, in 25 cases (60%), P9 is highly constrained but P2 is not. Thus, in a substantial majority of class I molecules, the sequence of the F pocket is the main determinant of peptide-binding specificity. The residues that are conserved in most other class I molecules have not been conserved in E

Table 5
Similarity Between E-like (mouse) and E (primate) Sequences Among the 63 PBR (Bjorkman et al. 1987*a*, 1987*b*; Saper, Bjorkman, and Wiley 1991; Garboczi et al. 1996) Residues

		* * x *
		** xx 1111111111111111111111111
		x x22223455566666666677777777888899911234444455555555666666667
		57924564578912345678901234567012457946333567901245678901235679 1
E-like.....	H2-Qa-1[a]	LYHFVVGVMPEYERETWKARDMGRNFRVNTLLYLWYCEYWSHKSMVDEHQQRAYLQGPVEWHY
	H2-T23[d]	..T.I.....A.....
E.....	HLA-E	....S....S..D...RS...TAQI.....R...HEF...Q..DAS.EH.....EDT.....
	Mafa-E1	....S.....DQ..RS...TAQT.....R...HEF...Q..DGS.EH.....EDT.....
	Mafa-E2	....S.....DQ..RS...TAQT.....R...HEF...Q..DGS.EH.....EDT.....
	Popy-E	....S.....D...RS...TAQT.....R...HEF...R..DAC.EH.....EDT...R.
Position†		aab bb bb ta a tabttb tcc fff f cadffef fettte dd tad a ta a
		bc b b c ebe f te d t t
		t d t

Note.—* = site found to be shared as the result of convergent/parallel evolution. x = site found to be retained from ancestral sequences for both E and E-like sequences. † = location of residue within the PBR. a, b, c, d, e, f = pockets A, B, C, D, E, F (Saper, Bjorkman, and Wiley 1991); t = T-cell-receptor-directed (Garboczi et al. 1996).

Table 6
HLA-E and H2-Qa1^b Are Known to Bind Similar Conserved Class Ia Leader Sequences

	Sequence	Source (reference)
HLA-E	VMAPRTVLL	HLA-A2, B8 (Braud et al. 1997)
H2-Qa1^b	AMAPRTLLL	H2-D^d, D^b, L^d (DeCloux et al. 1977; Kurepa and Forman 1997)

NOTE.—Residues shown in boldface are anchor residues (P2 and P9).

and E-like molecules. This is also consistent with the E and E-like genes= specialized peptide-binding function (Connolly et al. 1993), although it is not clear whether these genes share a biological function.

Similar cases of functionality of class Ib molecules have been reported. HLA-G is thought to be involved in maternal tolerance to the fetus (Schmidt and Orr 1995; Diehl et al. 1996). It is expressed in the human placenta and is known to bind a large array of peptides. The *H2-M3* mouse class Ib gene product (Hmt) presents a maternally transmitted antigen (Mta) which is a hydrophobic, N-formylated mitochondrial peptide (Wang, Loveland, and Lindahl 1991). In Hmt, two residues that are conserved in class Ia sequences (Val34 and Tyr171) are replaced with glutamine and phenylalanine, respectively. These changes of normally conserved residues, particularly the Val34 → Gln34 since it is included within the B pocket, may allow Hmt to bind Mta and other bacterial and mitochondrial peptides (Wang, Loveland, and Lindahl 1991; Pamer et al. 1992; Shawar et al. 1990). It has been proposed that the biological function of Hmt is in the immune defense against intracellular parasites (Fischer Lindahl, Hausmann, and Chapman 1989). In the present case, in all class I sequences (out of 244) examined, all alleles that did not contain Thr143 also did not have tryptophan at 147 (n = 16 including E and E-like sequences), and, likewise, most alleles that did not contain tryptophan at position 147 also did not have threonine at position 143. Therefore, it appears that for most class I molecules, it may be crucial functionally to conserve the Thr143/Trp147 combination, and any deviation from this combination could facilitate the binding of a novel type of peptide.

Because there have been numerous reports of interallelic gene conversion within the MHC (She et al. 1991), it might be argued that the similarity between the E and E-like genes is the result of the ancestors of both sets of genes having been converted by similar donors. Since the affected residues in this study are not part of a contiguous tract but are scattered throughout the exons, such an explanation seems unlikely. Further, although gene conversion is expected to affect both synonymous and nonsynonymous nucleotide sites, in phylogenetic trees constructed based on the proportion of synonymous nucleotide substitution (Nei and Gojobori 1986) in the PBR, the E and E-like sequences did not cluster together (data not shown). Finally, the fact that the F-pocket sequences of E and E-like are quite rare suggests that a donor for gene conversion is unlikely to have been available.

The lack of orthologous relationships among class Ib loci of mammals of different orders may seem inconsistent with the functional importance of these molecules. However, our results are consistent with other evidence of functionality for class Ib molecules including (1) that these primate and rodent molecules bind the same peptide (Kurepa and Forman 1997; DeCloux et al. 1997); (2) evidence from nucleotide sequence comparisons that some of these loci, including *HLA-E,* are conserved at the amino acid level (Boyson et al. 1995; Hughes 1995); and (3) evidence that other class Ib molecules, including HLA-G, H2-Qa-2, and H2-M3, bind peptides (Diehl et al. 1996; Joyce et al. 1994; Wang, Loveland, and Lindahl 1991). Our results suggest that, although they are not orthologous, class Ib loci of different mammalian orders may evolve similar functions convergently.

LITERATURE CITED

BJORKMAN, P. J., M. A. SAPER, B. SAMRAOUI, W. S. BENNETT, J. L. STROMINGER, and D. C. Wiley. 1987*a*. Structure of the human class I histocompatibility antigen, HLA-A2. Nature **329**:506–512.

———. 1987*b*. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature **329**:512–518.

BOYSON, J. E., S. N. MCADAM, A. GALLIMORE, T. G. GOLOS, X. LI, F. M. GOTCH, A. L. HUGHES, and D. L. WATKINS. 1995. The MHC locus *E* in macaques is polymorphic and is conserved between macaques and humans. Immunogenetics **41**:59–68.

BRAUD, V., E. Y. JONES, and A. MCMICHAEL. 1997. The human major histocompatibility complex class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor residues at positions 2 and 9. Eur. J. Immunol. **27**:1164–1169.

CONNOLLY, D. J., L. A. COTTERILL, R. A. HEDERER, C. J. THORPE, P. J. TRAVERS, J. H. MCVEY, P. J. DYSON, and P. J. ROBINSON. 1993. A cDNA clone encoding the mouse Qa-1^a histocompatibility antigen and proposed structure of the putative peptide binding site. J. Immunol. **151**:6089–6098.

DECLOUX, A., A. S. WOODS, R. J. COTTER, M. J. SOLOSKI, and J. FORMAN. 1997. Dominance of a single peptide bound to the class Ib molecule, Qa1^b. J. Immunol. **158**:2183–2191.

DIEHL, M., C. MUNZ, W. KEIHOLZ, S. STEVANOVIC, N. HOLMES, Y. W. LOKE, and H. RAMMENSEE. 1996. Nonclassical HLA-G molecules are classical peptide presenters. Curr. Biol. **6**:305–314.

DOOLITTLE, R. F. 1994. Convergent evolution: the need to be explicit. Trends Biochem. Sci. **19**:15–18.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

FISCHER LINDAHL, K., B. HAUSMANN, and V. M. CHAPMAN. 1989. Molecular definition of a mitochondrially encoded

mouse minor histocompatibility antigen. Cold Spring Harb. Symp. Quant. Biol. **54**:563–569.

GARBOCZI, D. N., P. GHOSH, U. UTZ, Q. R. FAN, W. E. BIDDISON, and D. C. WILEY. 1996. Structure of the complex between human T-cell receptor, viral peptide, and HLA-A2. Nature **384**:134–141.

HIGGINS, D. G., A. J. BLEASBY, and R. FUCHS. 1992. CLUSTAL V: improved software for multiple sequence alignment. Comput. Appl. Biosci. **8**:189–191.

HOWARD, J. C. 1987. MHC organization in the rat: evolutionary considerations. Pp. 397–427 *in* G. KLESOE and D. H. SCHULZE, eds. Evolution and vertebrate immunity. University of Texas Press, Austin.

HUGHES, A. L. 1995. Origin and evolution of HLA class I pseudogenes. Mol. Biol. Evol. **12**:247–258.

HUGHES, A. L., and M. K. HUGHES. 1995. Natural selection on the peptide-binding regions of major histocompatibility complex molecules. Immunogenetics **42**:233–243.

HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitution at MHC class I loci reveals overdominant selection. Nature **335**:167–170.

———. 1989. Evolution of the major histocompatibility complex: independent origin of nonclassical class I genes in different groups of mammals. Mol. Biol. Evol. **6**:559–579.

IMANI, F., and M. J. SOLOSKI. 1991. Heat shock proteins can regulate expression of the *Tla* region-encoded class Ib molecule *Qa-1*. Proc. Natl. Acad. Sci. USA **88**:10475.

JANEWAY, C., B. JONES, and A. HAYDAY. 1988. Specificity and function of T cells bearing γδ receptors. Immunol. Today **9**:73.

JOYCE, S., P. TABACZEWSKI, R. H. ANGELETTI, S. G. NATHENSON, and I. STROYNOWSKI. 1994. A nonpolymorphic major histocompatibility complex class Ib molecule binds a large array of diverse self-peptides. J. Exp. Med. **179**:579–588.

KLEIN, J. 1986. Natural history of the major histocompatibility complex. Wiley, New York.

KLEIN, J., and F. FIGUEROA. 1986. Evolution of the major histocompatibility complex. CRC Crit. Rev. Immunol. **6**:395–386.

KLEIN, J., and C. O'HUIGIN. 1994. The conundrum of nonclassical major histocompatibility complex genes. Proc. Natl. Acad. Sci. USA **91**:6251–6252.

KORNEGAY, J. R., J. W. SCHILLING, and A. C. WILSON. 1994. Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. Mol. Biol. Evol. **11**:921–928.

KREITMAN, M., and H. AKASHI. 1995. Molecular evidence for natural selection. Annu. Rev. Ecol. Syst. **26**:403–422.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetic analysis. Version 2.0. Pennsylvania State University, University Park.

KUREPA, Z., and J. FORMAN. 1997. Peptide binding to the class Ib molecule, Qa-1^b. J. Immunol. **158**:3244–3251.

LAMB, J., V. BAL, P. SAMPERIO, A. SO, J. ROTHBARD, S. JINDAL, R. YOUNG, and D. YOUNG. 1989. Stress proteins may provide a link between the immune response to infection and autoimmunity. Int. Immunol. **2**:191.

LI, W.-H., M. GOUY, P. M. SHARP, C. O'HUIGIN, and Y.-W. YANG. 1990. Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. Proc. Natl. Acad. Sci. USA **87**:6703–6707.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418–426.

O'BRIEN, R., M. HAPP, A. DALLAS, E. PALMER, R. KUBO, and W. BORN. 1989. Stimulation of a major subset of lymphocytes expressing T cell receptor γδ by an antigen derived from Mycobacterium tuberculosis. Cell **57**:667.

PAMER, E. G., C.-R. WANG, L. FLAHERTY, K. FISCHER-LINDAHL, and M. J. BEVAN. 1992. H-2M3 presents a Listeria monocytogenes peptide to cytotoxic T lymphocytes. Cell **70**:215–223.

RAMMENSEE, H.-G., T. FRIEDE, and S. STEVANOVIC. 1995. MHC ligands and peptide motifs: first listing. Immunogenetics **41**:178–228.

RZHETSKY, A., and M. NEI. 1992. A simple method for estimating and testing minimum evolution trees. Mol. Biol. Evol. **9**:945–967.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SAPER, M. A., P. J. BJORKMAN, and D. C. WILEY. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. J. Mol. Biol. **219**:277–319.

SCHMIDT, C. M., and H. T. ORR. 1995. HLA-G transgenic mice: a model for studying expression and function at the maternal/fetal interface. Immunol. Rev. **147**:53–65.

SHAWAR, S. M., R. G. COOK, J. R. RODGERS, and R. R. RICH. 1990. Specialized functions of MHC class I molecules: I. An *N*-formyl peptide receptor is required for construction of the class I antigen Mta. J. Exp. Med. **171**:897.

SHE, J. X., S. A. BOEHME, T. W. WANG, F. BONHOMME, and E. K. WAKELAND. 1991. Amplification of major histocompatibility complex class II gene diversity by intraexonic recombination. Proc. Natl. Acad. Sci. USA **88**:453–457.

STEWART, C.-B., J. W. SCHILLING, and A. C. WILSON. 1987. Adaptive evolution in the lysozymes of foregut fermenters. Nature **330**:401–404.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Illinois Natural History Survey, Champaign.

TYKOCINSKI, M. L., P. N. MARCHE, E. E. MAX, and T. J. KINDT. 1984. Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expressed gene and a putative pseudogene. J. Immunol. **133**:2261–2269.

WANG, C.-R., B. E. LOVELAND, and K. F. LINDAHL. 1991. *H-2M3* encodes the MHC class I molecule presenting the maternally transmitted antigen of the mouse. Cell **66**:335–345.

YANG, Z., S. KUMAR, and M. NEI. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics **141**:1641–1650.

YOKOYAMA, R., and S. YOKOYAMA. 1990. Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus,* and human. Proc. Natl. Acad. Sci. USA **87**:9315–9318.

SHOZO YOKOYAMA, reviewing editor

Accepted June 26, 1997