Immunogenetics (1994) 40: 339–351

ORIGINAL PAPER

Chanvit Leelayuwat · David C. Townend
Mariapia A. Degli-Esposti · Lawrence J. Abraham
Roger L. Dawkins

A new polymorphic and multicopy MHC gene family related to nonmammalian class I

Received: 31 May 1994 / Revised: 8 July 1994

Abstract We have used genomic analysis to characterize a region of the central major histocompatibility complex (MHC) spanning ~300 kilobases (kb) between *TNF* and *HLA-B*. This region has been suggested to carry genetic factors relevant to the development of autoimmune diseases such as myasthenia gravis (MG) and insulin dependent diabetes mellitus (IDDM). Genomic sequence was analyzed for coding potential, using two neural network programs, GRAIL and GeneParser. A genomic probe, JAB, containing putative coding sequences (PERB11) located 60 kb centromeric of *HLA-B*, was used for northern analysis of human tissues. Multiple transcripts were detected. Southern analysis of genomic DNA and overlapping YAC clones, covering the region from *BAT1* to *HLA-F*, indicated that there are at least five copies of PERB11, four of which are located within this region of the MHC. The partial cDNA sequence of *PERB11* was obtained from poly-A RNA derived from skeletal muscle. The putative amino acid sequence of *PERB11* shares ~30% identity to MHC class I molecules from various species, including reptiles, chickens, and frogs, as well as to other MHC class I-like molecules, such as the IgG FcR of the mouse and rat and the human Zn-α2-glycoprotein. From direct comparison of amino acid sequences, it is concluded that *PERB11* is a distinct molecule more closely related to nonmammalian than known mammalian MHC class I molecules. Genomic sequence analysis of *PERB11* from five MHC ancestral haplotypes (AH) indicated that the gene is polymorphic at both DNA and protein level. The results suggest that we have identified a novel polymorphic gene family with multiple copies within the MHC.

The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers L29406–L29411

C. Leelayuwat · M. A. Degli-Esposti · L. J. Abraham[1]
R. L. Dawkins (✉)
The Centre for Molecular Immunology and Instrumentation of The University of Western Australia, GPO Box F298, Perth, Western Australia 6001

D. C. Townend · R. L. Dawkins
Department of Clinical Immunology, Sir Charles Gairdner Hospital, Verdun St., Nedlands, Perth, Western Australia 6009

R. L. Dawkins
Department of Clinical Immunology, Royal Perth Hospital, Wellington St., GPO Box X2213, Perth, Western Australia 6001

Present address:
[1] Department of Biochemistry, The University of Western Australia, Nedlands, Perth, Western Australia 6009

Introduction

The MHC contains susceptibility genes for numerous autoimmune diseases. The best associations, especially in diseases such as insulin dependent diabetes mellitus (IDDM) and myasthenia gravis [(MG) (Dawkins et al. 1983; Kelly et al. 1985; Segurado et al. 1992; Durinovic-Bello et al. 1993; Degli-Esposti et al. 1992a, b)], are with MHC ancestral haplotypes (AH) rather than individual alleles. These haplotype associations suggest that disease susceptibility may be the result of *cis* interactions between multiple genes on a haplotype and/or the effect of unidentified genes carried by particular haplotypes. The genes responsible must be polymorphic to explain the haplotype associations, and must be involved in the process that leads to an autoimmune response.

Surprisingly, the conventional approach to gene identification (Spies et al. 1989a, b) has failed to recover coding sequences in the region between *BAT1* and *HLA-B*, where we have mapped susceptibility to both IDDM and MG (Degli-Esposti et al. 1992a, b). The approach we have taken to identify the gene content of this region involves the following steps: 1) identify disease and non-disease associated AHs; 2) compare the genomic sequence from relevant regions of these haplotypes; 3) search for polymorphic coding regions using neural network analysis; and 4) determine expression in relevant tissues. This strategy should identify genes critical to disease development rather than housekeeping genes. We have employed this strategy to determine the genetic content of the MHC region centromeric of the *HLA-B* locus.

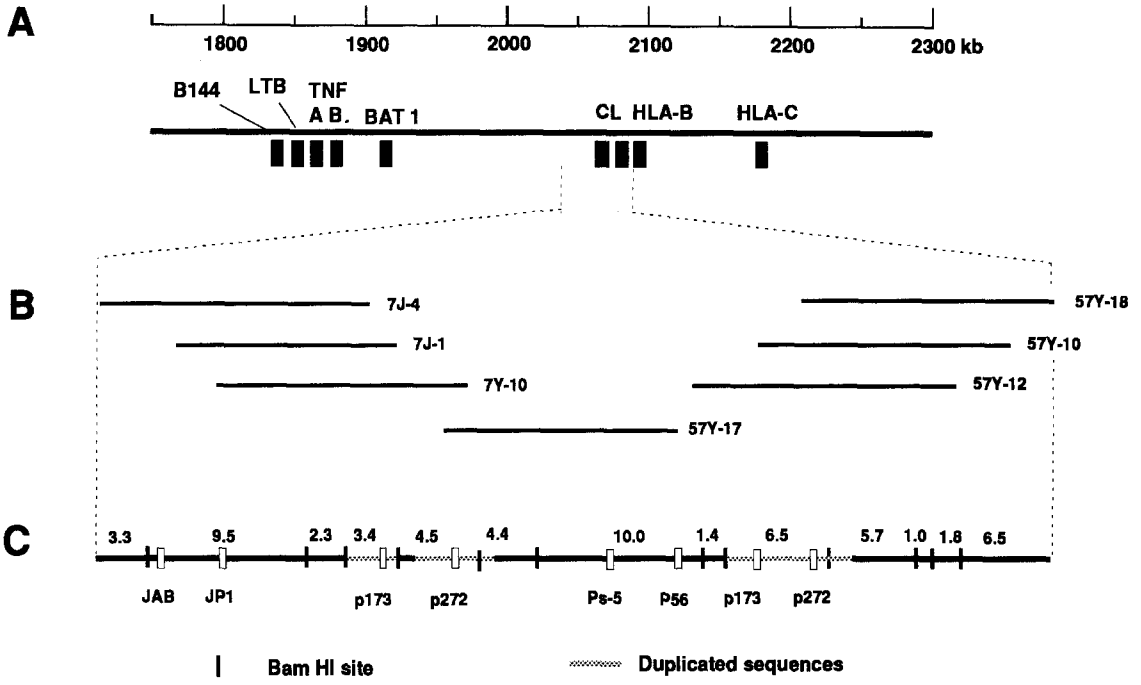


Fig. 1A–C *PERB11* (*JAB*) is located 60 kb centromeric of *HLA-B*. **A** The MHC region between B144 and HLA-C is depicted. A region approximately 60 kb centromeric of *HLA-B* was cloned. **B** Some of the genomic lambda clones isolated from two AH specific libraries (7.1 and 57.1) are shown. The clones were mapped by *Bam* HI digestion and were subsequently characterized by sequencing. **C** The combined *Bam* HI map of the clones is shown. The position of the genomic probes used for analysis of the region is shown on the map (*open boxes*). All probes except *JP1* (Chimini et al. 1990), are in-house. Ps-5 is a *Pst* I subclone of the 10 kb *Bam* HI fragment from the 7.1 library. The other probes were generated by PCR amplification of the appropriate clones. The lambda clone used for identification of *PERB11* is 7J-4

Materials and methods

Genomic cloning and sequencing

The ancestral haplotype specific genomic libraries used in this study (8.1, 57.1, 18.2, and 7.1) have been described previously (Abraham et al. 1991; Leelayuwat et al. 1992). The libraries were screened with multiple probes in order to obtain clones centromeric to the *CL* region (Leelayuwat et al. 1992). The JP1 probe (Chimini et al. 1990) was found to hybridize approximately 3 kb centromeric of the second copy of *CL* (Leelayuwat et al. 1992), and was used to isolate genomic clones from the 7.1 library. These clones were restriction mapped and compared with the clones covering the *CL* region in order to generate a continuous contig. A selected lambda clone, 7J-4 (Fig. 1), was cleaved with *Bam* H I and subcloned into pBC KS+ (Stratagene, La Jolla, CA). A plasmid subclone, 7J-4-9, carrying a 9.5 kb insert, was used for shotgun sequencing. Subclones were generated by the sonication method (Sambrook et al. 1989). The subclones were sequenced using a 373A DNA sequencer (Applied Biosystems, Foster City, CA) and the sequences were edited (SeqEd v1.03; Applied Biosystems) and aligned (AssemblyLIGN™; International Biotechnologies, New Haven, CT). Gaps and isolated contigs were filled in and joined, using an oligo-walking strategy (Leelayuwat et al. 1993).

Genomic sequence analysis to determine coding potential

An 8892 base pair (bp) genomic sequence was analyzed for coding potential, using two computer programs, GRAIL (Uberbacher et al. 1991) and GeneParser (Synder et al. 1993). The results were graphically displayed, using a recently developed program, Geneplot (Pinelli and co-workers, manuscript in preparation).

Direct genomic sequencing

Primers JP1A1 (5′ AGG ACA CGA TGT GCC AAC AG 3′) and JP1B2 (5′ CTT CCA CCA CCT TCA TGC TC 3′) were used to amplify a genomic fragment of 1270 bp from five B lymphoblastoid cell lines, each homozygous for an ancestral haplotype [(*Q6-3975: 7.1 = HLA-A3, -Cw7, -B7, TNF L, BAT3 L, C2 C, BF S, C4A3, C4B1, HLA-DR2, -DQ6; R6-12337: 57.1 = HLA-A1, -Cw6, -B57, TNF L, BAT3 L, C2 C, BF S, C4A6, C4B1, HLA-DR7, -DQ9; R6-12373: 8.1 = HLA-A1, -Cw7, -B8, TNF S, BAT3 S, C2 C, BF S, C4AQ0, C4B1, HLA-DR3, -DQ2; R5-5054: 18.2 = HLA-A30, -Cw5, -B18, TNF L, BAT3 S, C2 C, BF F1, C4A3, C4BQ0, HLA-DR3, -DQ2,* and *R6-12361: 46.1 = HLA-A2, -Cw1, -B46, TNF L, BAT3 L, C2 C, BF S, C4A4, C4B2, HLA-DR9, -DQ9*) (Wu et al. 1992; Degli-Esposti et al. 1993)]. 250 ng of DNA extracted from these cell lines was used for PCR amplification. Each reaction contained 200 μM each of dATP, dCTP, dGTP, and dTTP, 2 mM Tris-HCl (pH 8.3), 3 mM magnesium chloride, 50 mM KCl, 50 pmol of the primers, and 1 unit of *Taq* DNA polymerase (Amplitaq; Cetus, Emeryville, CA) in a total volume of 50 μl. The reaction was overlaid with light mineral oil (Sigma, St. Louis, MO) and subjected to 95° C for 2 min then thermocycling (35 cycles of 95° C for 45 s, 55° C for 45 s, 72° C for 2 min) followed by a final extension at 72° C for 10 min [(Gene Amplimer; Perkin Elmer (Norwalk, CT) and Cetus)]. The 1270 bp products obtained were fractionated through 1% low melting temperature agarose (Seaplaque; FMC Bioproducts, Rockland, ME) and further purified by Centricon column centrifugation (Centricon-100 microconcentrator; Amicon, Division of W. R. Grace & Co., Danvers, MA) according to the manufacturer's directions. Both strands of the products were sequenced, using JP1A1, JP1B2, and five internal primers: JA3 (5′ GTT CAT GGC CAA GGT CTG AG 3′), JP1A2 (5′ TGG GCT GAG TTC CTC ACT TG 3′) JP1B1 (5′ GGT CCT TGA TAT GAG CCA GG 3′), JB3 (5′ GTC AGG GTT TCT TGC TGA GG 3′), and JB4 (5′ ACA GAT CCA TCC CAG GAC AG 3′) in fluorescent-labeled dideoxy termination reactions analyzed on an automated 373A DNA sequencer (Applied Biosystems).

Exon connection and direct sequencing of cDNA

Poly-A RNA extracted from the skeletal muscle tissue of a 20-year-old Caucasian female (Clontech, Palo Alto, CA) was used for cDNA synthesis. The 3′ downstream primers, JP1B2 and JA3, were used separately in reactions containing 2 μg of poly-A RNA, 64 units of RNase inhibitor (Clontech), 700 μM each of dATP dCTP, dGTP, and dTTP, 12 units of AMV reverse transcriptase (Clontech), 0.3 μM of the JP1B2 or JA3 primers and 1 × reverse transcriptase buffer (Clontech) in a total volume of 30 μl. The RNA-primer mixture was preheated at 65° C for 5 min and then added to the reactions. *PERB11* cDNA was synthesised at 52° C for 30 min. Each reaction was terminated by adding 1 μl of 0.5 M ethylenediaminetetraacetate. The RNA in DNA-RNA hybrids was hydrolyzed by adding 2 μl of 6 N NaOH followed by incubation at 65° C for 30 min. The reaction was then neutralized by adding 2 μl of 6 N acetic acid and the cDNA purified by using GENO-BIND™ (Clontech) according to the manufacturer's specifications. The purified cDNA was used as template for exon connection (Fearon et al. 1990) and amplification. The JP1B2 or JA3 primer was used in the amplification reaction in combination with the JB3 primer, the latter expected to be located on a different exon. The conditions used for polymerase chain reaction (PCR) were as described above. The PCR product was sequenced, using the JA3, JP1B1, and JB3 primers.

Southern and northern analyses

High relative molecular mass genomic DNA was extracted from lymphoblastoid B-cell lines as described previously (Degli-Esposti et al. 1992c). The cell lines used are homozygous for the 7.1 (Q6-3975), 8.1 (R6-12373), 57.1 (R6-12337), 18.2 (R5-5054), 46.1 (R6-12361), and 62.1 (R6-12316) AHs (Wu et al. 1992). Nine YAC clones, covering the region from *BAT1* to *HLA-F*, were selected for analysis. DNA was extracted from the YAC clones according to a standard method (Sambrook et al. 1989). The DNA samples were then digested with *Bam* HI (Promega, Madison, WI), according to the manufacturer's specifications, fractionated through a 0.8% agarose gel (LE, Seakem, FMC Bioproducts) and transfered to a nylon membrane (GeneScreen-Plus; Dupont, Wilmington, DE). Pulsed field gel electrophoresis (PFGE) of *Mlu* I-digested DNA was performed as previously described (Zhang et al. 1990). The genomic JAB probe, carrying potential coding sequences, was labeled with 5′ ^{32}P-dCTP (Amersham, Arlington Heights, IL), using the random priming method (Feinberg et al. 1983) and used for hybridization analysis, as described previously (Degli-Esposti et al. 1992c). Northern analysis was performed on a multiple human tissue blot [(MTN2) (Clontech)], using the JAB probe as described above. The MTN2 blot contains poly-A RNA from human heart, brain, placenta, lung, skeletal muscle, kidney, and pancreas.

Results

Identification of coding potential from genomic sequence analysis

A 60 kb region centromeric of *HLA-B* was cloned and sequenced in order to identify new genes (Fig. 1). The 7J-4 clone was selected for genomic sequencing and analysis. A 9.5 kb subclone of 7J-4 (7J-4-9) was analyzed for coding probability and exon-intron structure, using the coding recognition module (CRM) of GRAIL (Uberbacher et al. 1991) and GeneParser (Synder et al. 1993). Both programs identified a 1 kb region containing sequence with high coding probabilities [(0.8–0.95) (Fig. 2)]. This region, designated *JAB*, was synthesized by PCR of the plasmid clone, 7J-4-9, using JP1A1 and JP1B2 as primers. The JAB product was used as a probe for northern and Southern analyses.

Further analysis of the 7J-4-9 sequence indicated the presence of a putative 3.7 kb intron containing L1-, Alu-, and MER-related sequences (Fig. 2). It should be noted that GeneParser predicted two small exons (20–30 bp) within the same region.

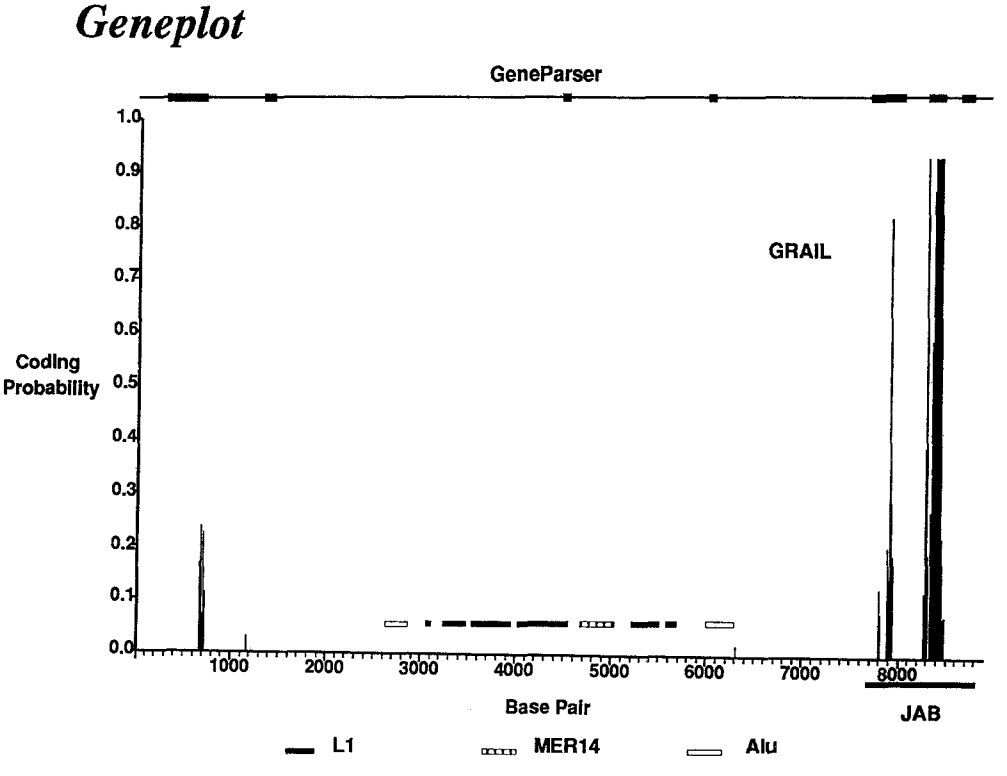


Fig. 2 Computer analysis of anonymous genomic sequence reveals the existence of potential coding sequences. Geneplot represents coding potential analysis of an 8892 bp genomic sequence located 60 kb centromeric of *HLA-B*. The coding potential information was obtained from two neural network programs, GRAIL and GeneParser. The results of the GRAIL analysis are presented as coding probability (Y-axis) along the region (X-axis). The result of the GeneParser analysis showing putative gene structure, with respect to exons and introns, is shown at the *top*. The position of sequences related to L1, Alu, and MER is also shown. A region of approximately 1 kb was predicted to contain potential coding sequences by Grail and GeneParser. This region, synthesized from the 7J-4-9 plasmid clone by PCR amplification, was used as a probe (*JAB*) for northern analysis

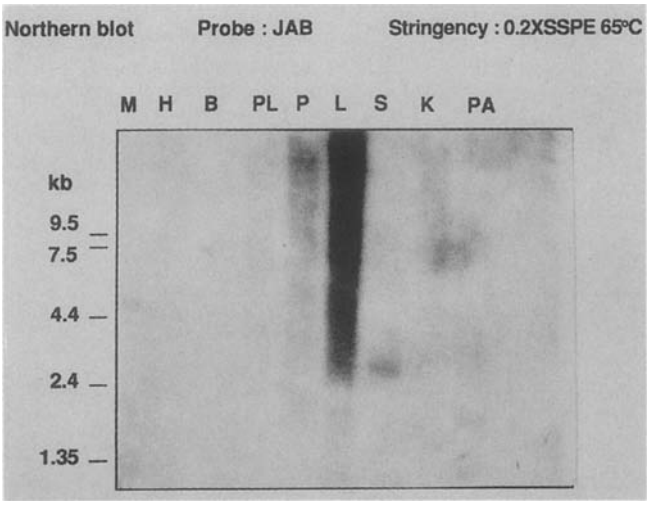


Fig. 3 Northern analysis of *PERB11* on multiple human tissues, using the JAB probe. A transcript of ~3.0 kb is detected in skeletal muscle (S). A large transcript (>9.5 kb) is detected in lung (P). The size of the transcripts in the liver (L) is not assessable because of the high background. No transcripts are detected in heart (H), brain (B), placenta (PL), kidney (K), and pancreas (PA). There is an artefact between the K and PA lanes. M refers to the RNA size marker

Expression of PERB11

To investigate the expression profile of the putative coding sequences, the *JAB* probe was used for nothern analysis of multiple human tissues. As shown in Figure 3, the probe detects a 3 kb transcript in skeletal muscle (designated *PERB11*) as well as a large transcript (>15 kb) in lung tissue. Shorter exposures (not shown) revealed transcripts of 10 kb and 8 kb in liver. Similar transcripts (10 and 8 kb) were also detected in thymus, prostrate, testis, small intestine, and peripheral blood leukocytes (data not shown). It is unclear whether the multiple transcripts are the result of alternative splicing or pre-processing of *PERB11* RNAs or whether they represent the products of multiple genes. No transcripts were observed in heart, brain, placenta, kidney, and pancreas (Fig. 3).

Localization of multiple copies of PERB11 within the MHC

Pulsed field gel electrophoresis of human DNA digested with *Mlu* I was performed to confirm the physical location of the *JAB* sequence used for northern analysis. *JAB* hybridized to at least two different *Mlu* I fragments, as shown in Table 1. To determine whether there are multiple copies of JAB within the MHC, we probed the same blot with MHC probes adjacent to JAB. JP1 is located ~5–9 kb telomeric of *JAB* and a single copy probe (Ps-5) ~15 kb telomeric of *JP1* between the two *CL* loci (Fig. 1). As shown in Table 1, genomic DNA from the 7.1 ancestral haplotype carries two *JP1*-positive fragments (~30 and 40 kb) and two *JAB*-positive fragments (~40 and 160 kb). The 40 kb fragment hybridizes to both *JP1* and *JAB* and therefore it should arise from the cloned region (Fig. 1). It

Table 1 Southern analysis (*Mlu* I) of *PERB11* suggests the existence of multiple copies of PERB11 in the genome. The density of each band is scored semi-quantitively as: 3 = +, 4 = ++, etc to 9 = 7+

			Fragment sizes (kb)		
Lab No.	Name	Probe	30	100	170
Q6-3975	7.1	JP1	4 2		
		JAB	3		4
		Ps-5	4		
R5-5054	18.2	JP1	4 3		
		JAB	3		4
		Ps-5	4		
Q6-8187	57.1,8.1	JP1	4 4	5 5	
		JAB	3	3	4
		Ps-5	4	5	
R6-12373	8.1	JP1	3	4	
		JAB	3		4
		Ps-5		4	
R6-12337	57.1	JP1	7	4	
		JAB		4	5
		Ps-5	7		

should be noted that there is no *Mlu* I site between *JAB* and *JP1*. The shortest (JP1 and Ps-5) and largest (JAB) fragments have not been localized, but the data suggest that there are at least two copies of *JP1* and *JAB*. The data in Table 1 also indicate that there are differences between the haplotypes in this region (see below).

In order to map the additional *PERB11*-related sequences in the MHC, YAC clones covering the region between *BAT1* and *HLA-F* were investigated. This region was selected because we have shown sequence similarities between this region and the *TNF* to *HLA-B* interval (Leelayuwat and co-workers, manuscript submitted). As shown in Figure 4, *JAB* hybridizes to six of nine YAC clones: 318G11, 152G3, 368D7, 225B1, 188A4, and 421F7. Apparently, *PERB11* is replicated in at least three regions of the MHC, close to *HLA-B, HLA-E,* and *HLA-A,* respectively. Since the clones were isolated from a YAC library generated from the 62.1 ancestral haplotype (*R6-12316*), the *Bam* H I RFLP of *R6-12316* is also shown in Figure 4. Four of the five fragments found on 62.1 were detected in six YAC clones. YAC 318G11 carries two fragments, 11 kb and 10 kb. The 10 kb fragment is shared by 152G3. This result maps two copies of *PERB11* in the region centromeric to *HLA-B*. One copy (*PERB11.1*), on the 10 kb *Bam* H I fragment, is shared by 318G11 and 152G3. This result was expected from the cloning and sequencing data. The second copy (*PERB11.2*) carried by the 11 kb fragment (Fig. 4) must be close to the *BAT1* locus, since it is present in 318G11, but absent in 152G3. YAC 225B1 carries two fragments, 8 kb and 6 kb, the former being shared by 368D7 and 188A4. The 6 kb *Bam* H I fragment carried by 225B1 is shared by 421F7. Sharing of the 8 kb fragment by 225B1, 368D7, and 188A4 indicates that the third copy of *PERB11* (*PERB11.3*) is located in close proximity to *HLA-E*. The 6 kb *Bam* H I fragment shared by 225B1 and 421F7 suggests that the fourth copy of *PERB11* (*PERB11.4*) is

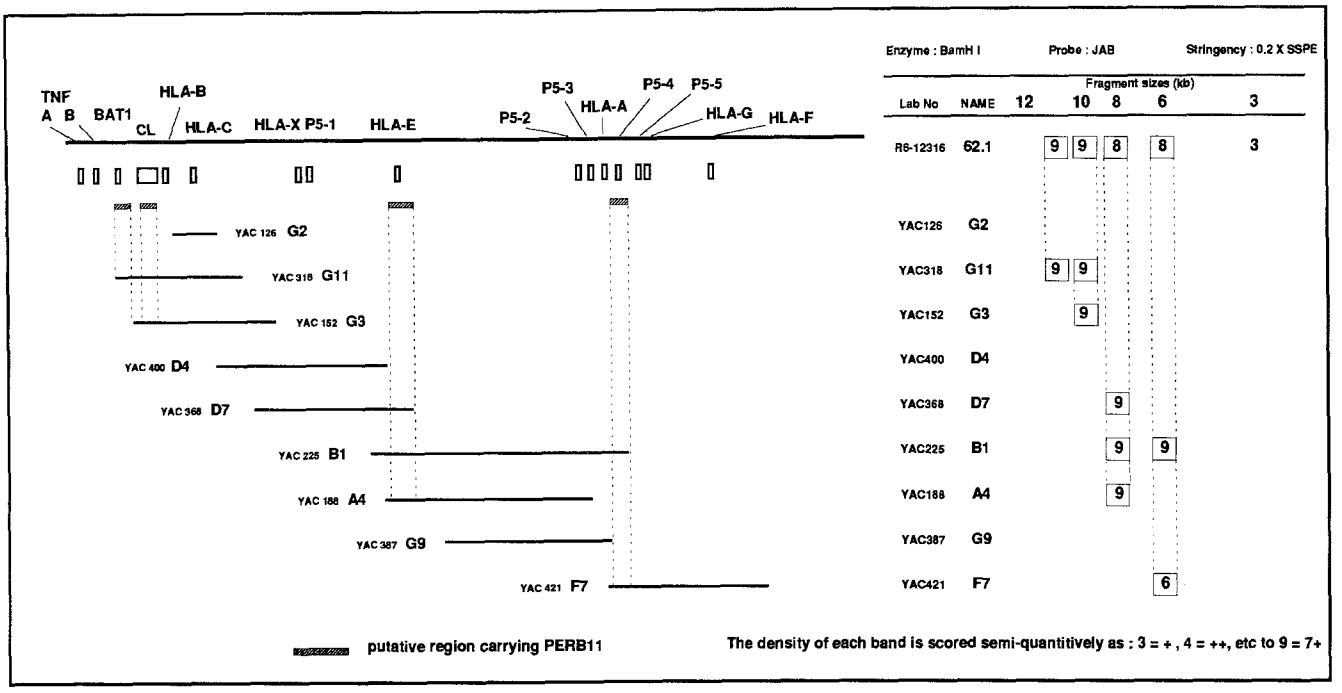


Fig. 4 The genome contains at least five copies of *PERB11*, four of these are in the MHC region telomeric of the *TNF* loci. Nine YAC clones covering the region from *BAT1* to *HLA-F* were selected to map the location of *PERB11* sequences using the JAB probe. Southern analysis (*Bam*H I) of R6-12316 (62.1 AH) was also performed, since the YAC clones were isolated from a library constructed from this individual. *Dotted lines* indicate regions shared by the positive clones. *Hatched boxes* represent the putative location of *PERB11* sequences. *Bam* HI fragments found in the clones as well as the genomic DNA control (62.1) are shown on the *right*. Four of the five fragments detected in the genomic DNA correspond to fragments carried by six of the nine YAC clones studied. Clearly, the *JAB* sequence is replicated and confined to three locations; two copies are in the region centromeric of *HLA-B*, one copy is in the region in close proximity to *HLA-E* and one copy is in the region telomeric of *HLA-A*

Table 2 Southern analysis (*Bam* HI) of *PERB11* reveals sequence polymorphism and suggests differences in gene copy number between different AHs. The density of each band is scored semi-quantitively as: 3 = +, 4 = ++, etc to 9 = 7

Enzyme: *Bam* H I		Probe: JAB			Stringency: 0.2×SSPE		
		Fragment sizes (kb)					
Lab No.	Name	12	10		8	6	3
Q6-3975	7.1	9	9		9	8	3
R6-12337	57.1		9	9	9	8	
R6-12373	8.1	8	8		6	6	4
R6-12361	46.1	8	8		7	7 4	9
R5-5054	18.2	8	8		7	7	5
R1-3319X	B75	8	9		9	8	
R6-12316	62.1	9	9		8	8	3

located telomeric to *HLA-A*. It should be noted that *JAB* has no significant DNA sequence similarity to *HLA* class I or P5 sequences. This is confirmed by the fact that YAC clones which carry either class I (126G2) or P5 (400D4 and 387G9) sequences do not hybridize to the *JAB* probe. The above results indicate that four of the five *PERB11* copies carried by R6-12316 (62.1) are located within the region between *BAT1* and *HLA-F*.

PERB11 is polymorphic

PERB11 polymorphism was demonstrated using *Bam* HI (Table 2). At least four restriction fragment length polymorphism (RFLP) patterns were detected in the seven individuals tested. The RFLP patterns are haplotypic, i.e., unrelated individuals carrying the same MHC ancestral haplotype have the same pattern (data not shown). Comparison of RFLP patterns from different haplotypes (Table 2) shows variations in the number as well as intensity of fragments, suggesting that *PERB11* may be polymorphic in gene copy number as well as sequence.

cDNA and genomic sequence analysis of PERB11

Skeletal muscle was selected as a source of cDNA, since only one transcript (3 kb) is detected in this tissue. *PERB11* cDNA was obtained, using an exon connection procedure (Fearon et al. 1990). cDNA synthesis using JP1B2 followed by amplification with JP1B2 and JB3 failed to yield a product when poly-A RNA from skeletal muscle was used as substrate. However, using JA3 to synthesise cDNA and JA3 and JB3 for amplification, a PCR product of 350 bp was obtained. The same primer pair amplified a 640 bp product from both 7.1 genomic DNA (Q6-3975) and the 7J-4-9 plasmid DNA. Comparison of cDNA, genomic DNA, and plasmid DNA sequences indicated the presence of a 274 bp intronic sequence between the JA3 and JB3 primers (Fig. 5.). Although there are several nucleotide

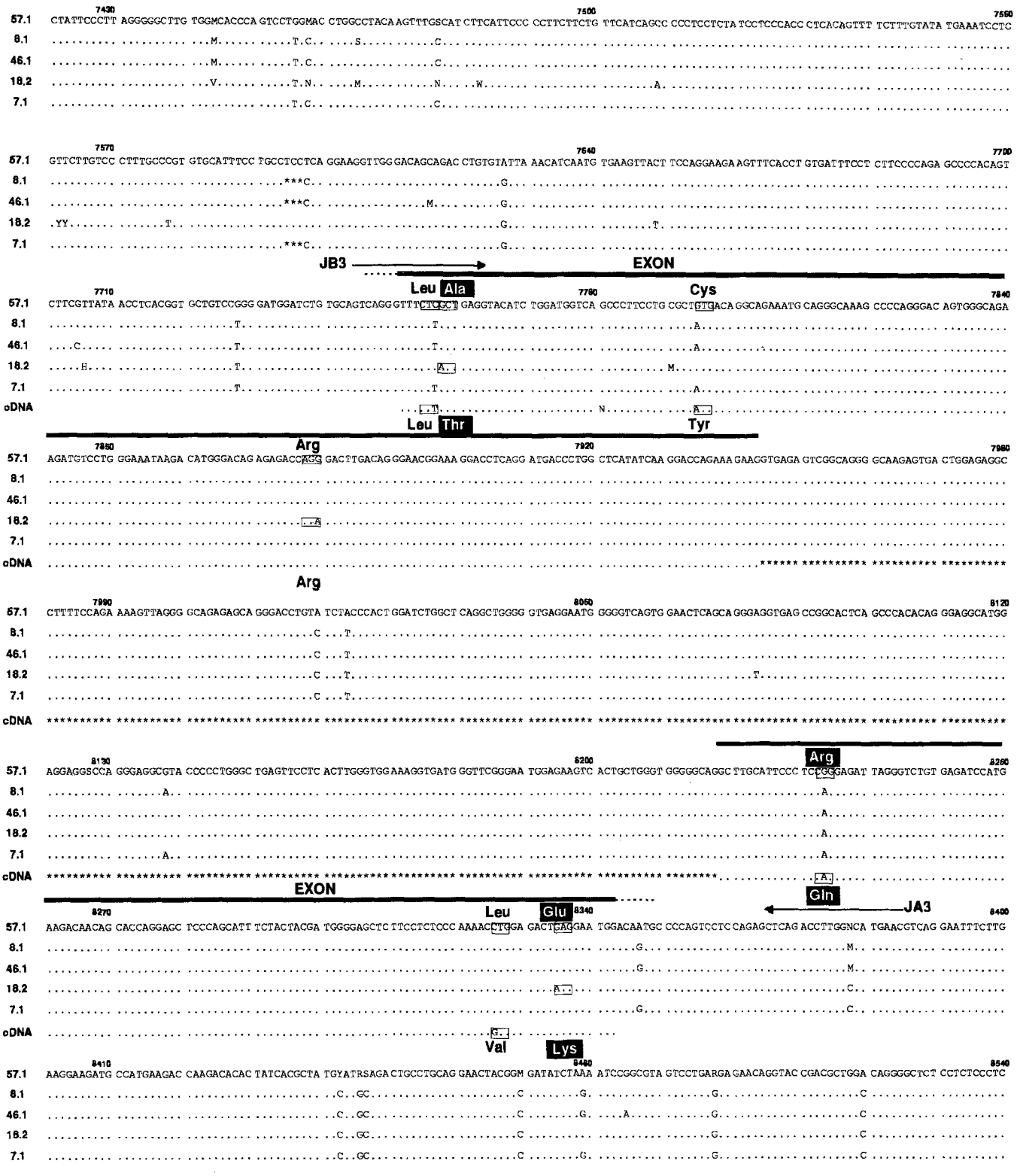


Fig. 5 Genomic analysis of *PERB11.1* reveals polymorphism in both intronic and exonic sequences. Stretches of 1120 bp of *PERB11.1* genomic sequence from 5 MHC AHs are compared. The sequence of 57.1 is used as a reference and shown at the *top*. Nucleotide positions correspond to the positions in the 8892 bp sequence. *Dots* represent the same nucleotide as in the reference sequence (57.1). *Asterisks* indicate deletions. Standard ambiguity codes are used where: S = C or G, W = A or T, Y = C or T, M = A or C, R = A or G, V = A or C or G, H = A or C or T and N = A or T or C or G. The cDNA sequence was obtained from poly-A RNA derived from skeletal muscle, using JA3 and JB3 as primers (see text). *Asterisks* within the cDNA sequence represent nucleotides not found in the cDNA i.e., putative *PERB11.1* intron. Nucleotides that are *boxed* represent exonic differences; amino acids are shown *adjacent* to the polymorphic codons. Nonsynonymous amino acid changes, which are not conservative, are shown in *white characters* on a black background. Leu = leucine, Ala = alanine, Thr = Threonine, Cys = Cysteine, Tyr = Tyrosine, Arg = Arginine, Gln = Glutamine, Val = valine, Glu = glutamic, Lys = lysine. Nucleotide differences are distributed along the region and can be found in both intron and exons. Five of the seven nucleotide changes in the exons lead to non-synonymous amino acid substitutions

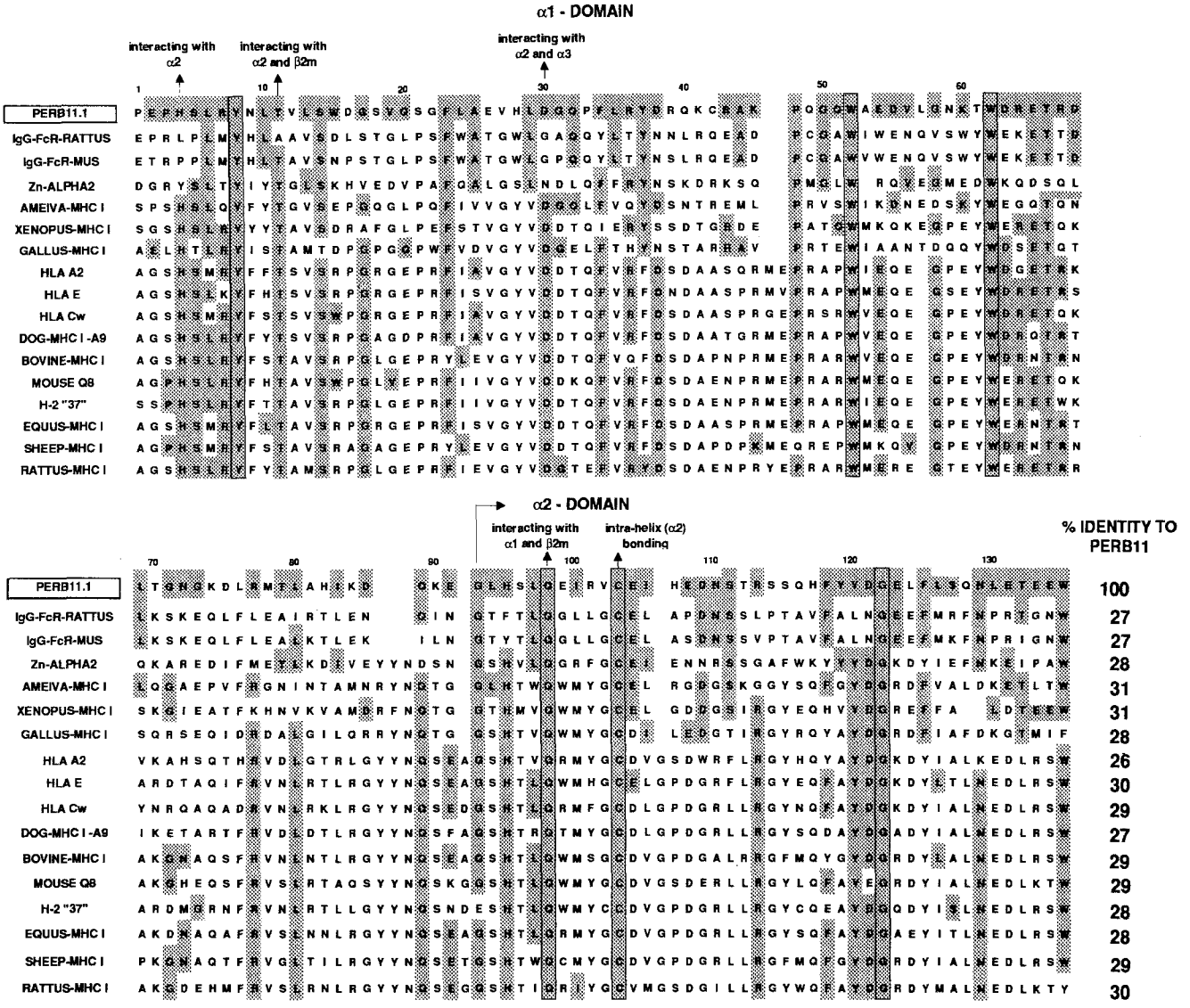


Fig. 6 *PERB11.1* is related to MHC class I molecules. The putative protein sequence of *PERB11.1* was deduced from the cDNA sequence, using a coding frame not interrupted by non-sense codons. The amino acid sequence of *PERB11.1* including predicted amino acids upstream of the cDNA (see text) was compared with known proteins in the Gene peptide data base version 78, using the BlastP program. The top matches were selected for direct comparison. Amino acids identical to those of *PERB11.1* are *shaded. Blank spaces* represent gaps introduced to obtain the best match between the sequences. Residues 1-92 correspond to the α1 domain and residues 93-135 to the amino terminal part of the α2 domain of MHC class I molecules. Residues conserved in MHC class I molecules from different species (Grossberger et al. 1992) are shown by *vertical arrows* and their interactions with other parts of the molecule or an associated molecule (β_2m) are listed. All the top matches are with MHC class I molecules or MHC class I-like molecules, such as the Zn-α2-glycoprotein and the MHC class I-like IgG Fc-receptor from mouse and rat. The overall similarity of *PERB11.1* to these molecules is approximately 30% and conserved residues range from 66% to 72%. Six residues (*boxed*) are shared by all the molecules compared

differences in this region when five different AHs are compared, the cDNA sequence is identical to the genomic sequences of 7.1, 8.1, and 46.1, except for a G → C transversion at position 8326 leading to a conservative amino acid change (valine in the cDNA to leucine in 7.1, 8.1 and 46.1). This result suggests that the cDNA is derived from the *PERB11.1* locus carried by the 7J-4-9 clone.

To investigate polymorphism of *PERB11.1* at the DNA level, PCR products generated from the genomic DNA of five different AHs were directly sequenced. The 1120 bp amplified region is polymorphic with up to 1.4% nucleotide variation including both exonic and intronic differences (Fig. 5.). A three nucleotide difference at position 7595 to 7597 represents a deletion in 8.1, 46.1, and 7.1 or an insertion in 57.1 and 18.2. The *PERB11.1* sequences of 57.1 and 18.2 are quite unique, whereas 7.1 and 8.1 are rather similar in this region. Seven of the observed nucleotide differences are in the exons deduced from the cDNA sequence. Of these, five lead to non-synonymous amino acid changes. Two of five changes are conservative, e.g., cysteine at position 7795–7797 in 57.1 and 18.2 is replaced by tyrosine in 8.1, 46.1, 7.1; valine at position 8336–8338 in the skeletal muscle cDNA sequence is replaced by leucine in all haplotypes tested. The remaining three

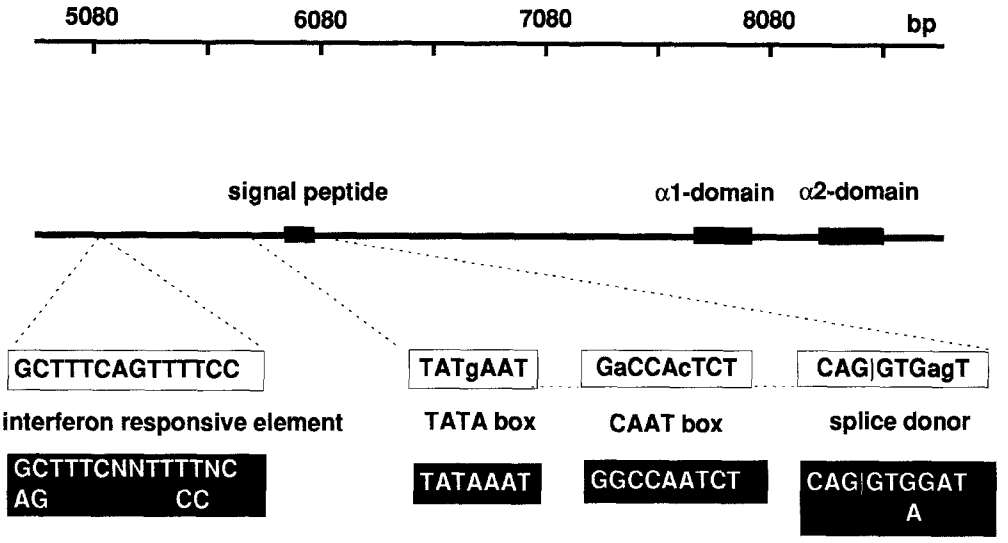


Fig. 7 Putative structure of the 5′ region of the *PERB11.1* gene as deduced from genomic sequence analysis. Relevant nucleic acid subsequences (e.g., TATA box) are shown in *boxes*. Their position relative to the 8892 bp sequence of 7J-4-9 is shown. Nucleotides in DNA motifs that are different from those of the consensus sequences are shown in *lower-case* letters. Consensus DNA motifs are shown in *white characters* on a *black background*. A ~1.7 kb intron separates the signal peptide and α1 domain

amino acid changes lead to alterations in charge or hydrophobicity, for example lysine (positive) at position 8335–8337 in 18.2 is replaced by glutamic acid (negative) in 57.1, 8.1, 46.1, 7.1 and the skeletal muscle cDNA.

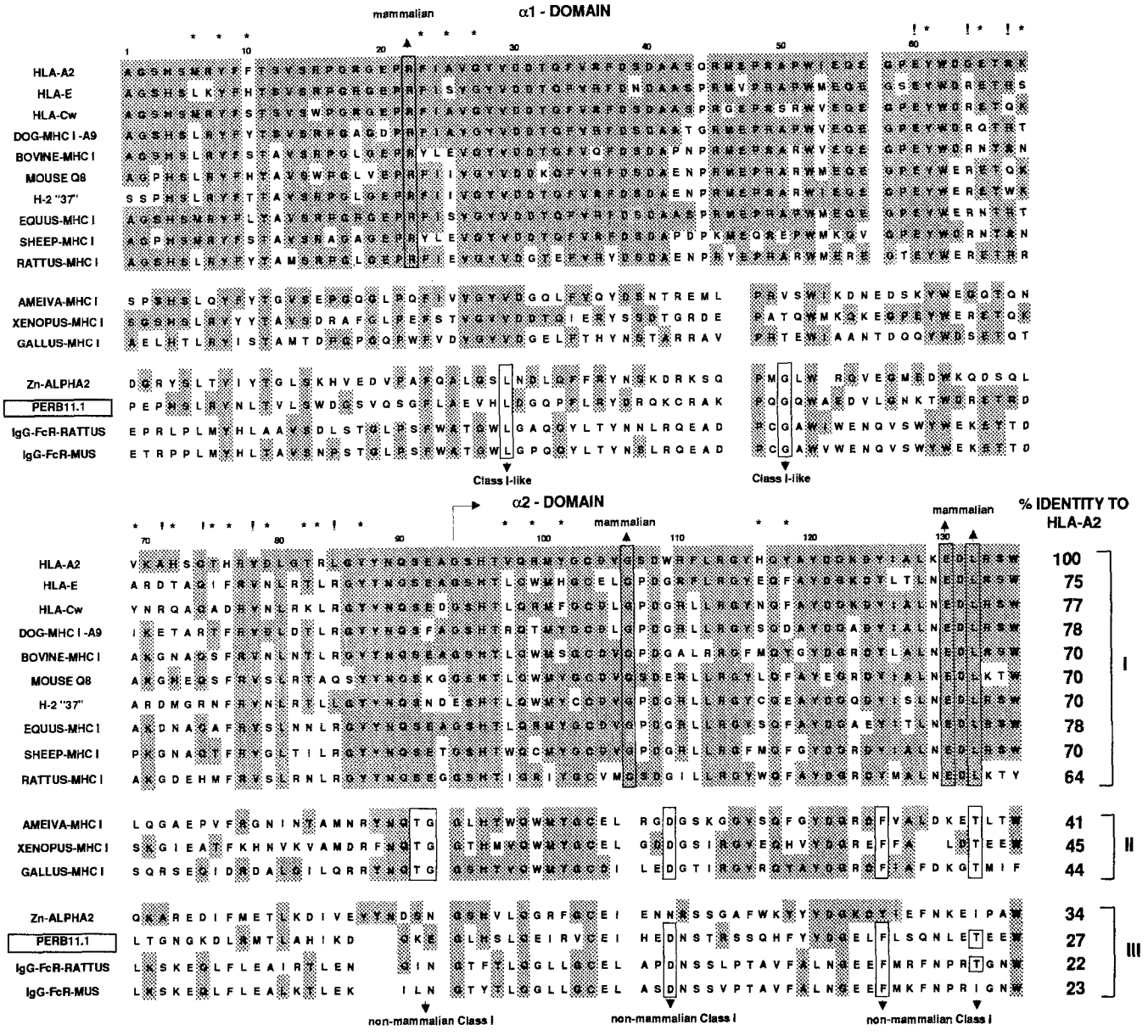
Analysis of the predicted PERB11.1 protein sequence

To gain some insight into the function of the PERB11.1 molecule, the amino acid sequence was deduced from the cDNA sequence. The translated sequence was compared with other known proteins in the Gene peptide data base version 78, using the BlastP program (Karlin et al. 1990). All the best matches represent MHC class I molecules from various species, including lizard, frog, and chicken, or MHC class I-like molecules, such as IgG Fc receptor from mouse and rat. A comparative sequence alignment is presented in Figure 6. Overall amino acid identity between these molecules and PERB11.1 ranges from 26 to 31% over 135 residues. The percentage of conserved residues found in PERB11.1 and homologous proteins ranges from 72 to 66 (72% with *HLA-E*; 70% with Zn-α2-glycoprotein; 69% with Xenopus and Gallus MHC class I; 67% with Ameiva MHC class I and mouse IgG FcR, and 66% with IgG FcR of the rat). The amino acid sequence of PERB11.1 deduced from the available cDNA sequence covers most of the α1 domain of MHC class I molecules (residue 20 onwards) and about half of the α2 domain. In PERB11.1 these two domains are encoded by different exons, as in other MHC class I genes. The conserved residues, which characterize MHC class I molecules from different species (Grossberger et al. 1992), are also present in the *PERB11.1* sequence. These are histidine, threonine, aspartic acid, glutamine, and cysteine at positions 4, 11, 30, 98, and 103, respectively (Fig. 6.). Interestingly, there are six residues conserved amongst PERB11.1, MHC class I molecules and MHC class I-like proteins. They are tyrosine, tryptophan, tryptophan, glutamine, cysteine and glycine at positions 8, 52, 62, 98, 103, and 122, respectively (Fig. 6.). It should be noted that the sequence similarity between PERB11.1 and these molecules is significant at the protein but not at the DNA level (<50% identity). Surprisingly, the CD1 molecule, which is one of the MHC class I-like molecules, was not identified by the above search. However, when we compared PERB11.1 with CD1 (Balk et al. 1989), 17% identity was observed over 81 amino acids (data not shown).

Analysis of the genomic sequences 5′ of the cDNA sequence shows that the sequence similarity to class I continues through to the beginning of the α1 domain. Furthermore, a consensus interferon responsive element GCTTTCAGTTTTCC identical to the consensus (A/G)(C/G)TTTCNNTT(T/C)(C/T)NC (Cohen et al. 1988; Fan et al. 1989) was identified at position 5086 of the 8892 bp DNA sequence (Fig. 7.). An imperfect TATA box (TATgAATt), CAAT box (GaCCAcTCT), and splice donor sequence (CAGGTGgaT) were identified at positions 5769, 5778, and 5918, respectively. The reading frame compatible with the described splice donor sequence translates to 60 amino acids starting with a methionine at position 5770–5772. A second methionine is encoded at position 5815–5817 translating to a highly hydrophobic 35 amino acid sequence (MFVFLXLAYFTYNIFEVRWAQWLTPGFPALWEAEA). This sequence has a potential to code for the signal peptide and, therefore, represents the first exon of PERB11.1. The coding sequence is then interrupted by an intron of ~1.7 kb.

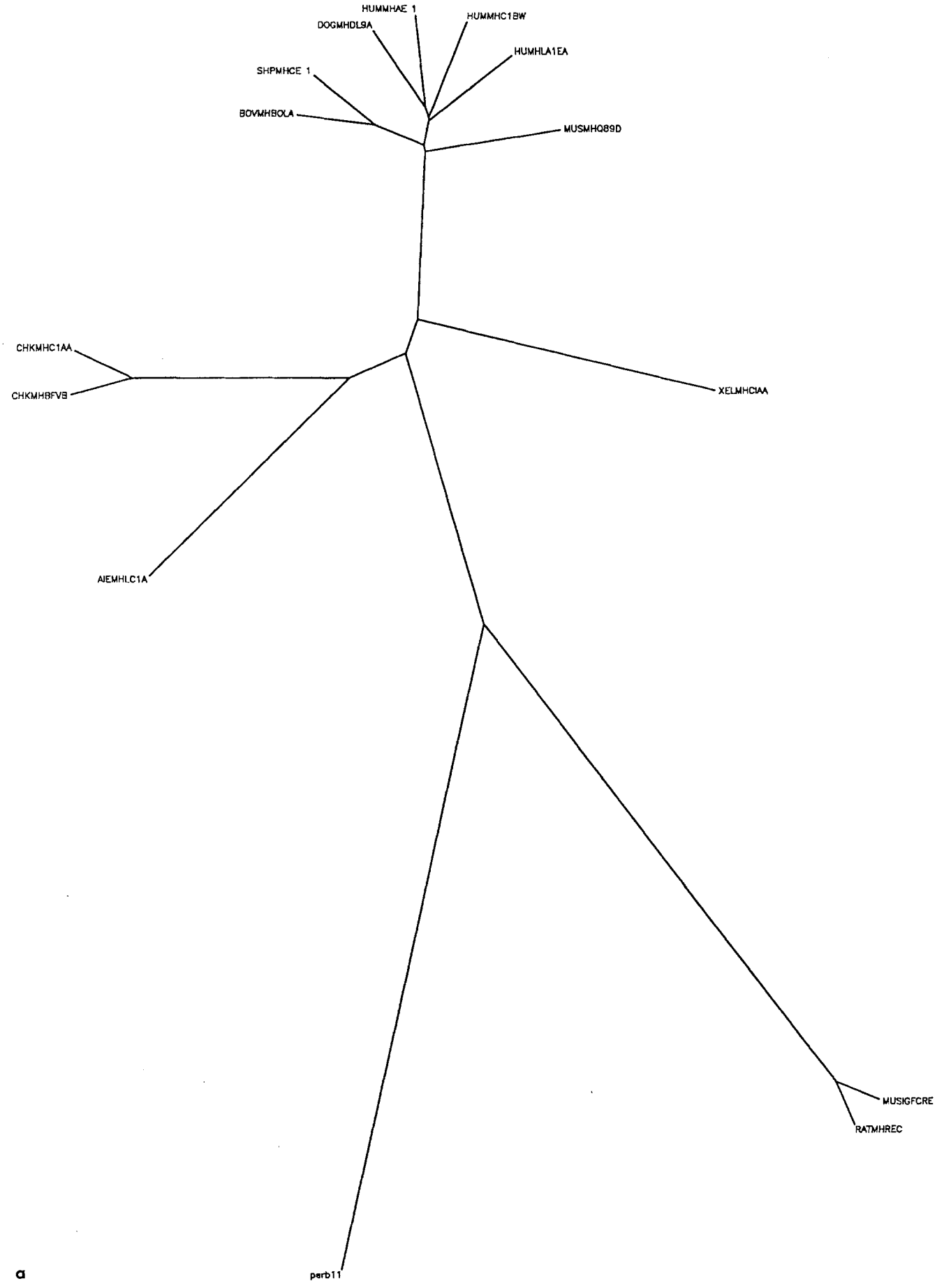
Comparison of PERB11.1 to MHC class I molecules

The amino acid sequences of PERB11.1 and other MHC class I and MHC class I-like molecules were compared with *HLA-A2* (Ennis et al. 1990). These include: HLA-C (Ellis et al. 1989); HLA-E (Mizuno et al. 1988); dog MHC class I (DLA A9) (Sarmiento et al. 1990); cattle MHC class I (Ennis et al. 1988); sheep MHC class I (Grossberger et al. 1990); horse MHC class I (A. J. Martin, submitted), mouse H-2 and Q8/9d (Matsuura et al. 1989); rat MHC class I



(Rothermel et al. 1993); frog MHC class I (Shum et al. 1993); chicken MHC class I [(BF) (Kaufman et al. 1992)]; lizard MHC class I (Grossberger et al. 1992); the neonatal IgG Fc receptors of the rat (Simister et al. 1989) and mouse (Ahouse et al. 1993) and human Zn-alpha2-glycoprotein (Araki et al. 1988; Freije et al. 1991). The PERB11.1 amino acid sequence used in this analysis includes the predicted sequence upstream of the cDNA sequence. The sequences analyzed can be grouped into 3 categories: (1) mammalian MHC class I molecules; (2) nonmammalian MHC class I molecules and (3) MHC class I-like, FcR-like molecules (Fig. 8.). These 3 groups show different levels of amino acid identity (64%–78%, 41%–45%, and 22%–34%) when compared with the reference *HLA-A2* sequence. The level of similarity within and between the three groups, however, is underestimated as a result of the polymorphism which exists in molecules involved in peptide binding. Some residues appear to be specific to each of the three groups,

Fig. 8 *PERB11.1* is more closely related to nonmammalian than to mammalian MHC class I molecules. The amino acid sequence of MHC class I molecules from different species, including mammalian and nonmammalian vertebrates, are aligned. These include three human HLA molecules (HLA-A2, HLA-C, HLA-E), dog, cattle, horse, goat, rodent (mouse and rat), reptilian (Ameiva), amphibian (Xenopus), and avian (Gallus) MHC class I molecules. The sequences analyzed, including *PERB11.1*, Zn-α2 glycoprotein, and the MHC class I-like IgG Fc-receptors from mouse and rat, are organized into three groups: (1) mammalian MHC class I molecules; (2) nonmammalian MHC class I molecules; and (3) MHC class I-like molecules. Amino acids identical to HLA-A2 are *shaded*. *Asterisks* and *vertical bars* above amino acids indicate peptide binding residues and upward facing residues which may interact with the T-cell receptor (Grossberger et al. 1992). Residues shared by class I-like molecules and mammalian and nonmammalian MHC class I molecules are *boxed*. The mammalian molecules are grouped together and share several conserved regions with overall identities between 64% and 78%. In contrast, the molecules are less conserved in the lizard, frog, and chicken, with overall identities between 41% and 45%. Most of the divergent residues are confined to the peptide binding and TcR interaction regions. Although *PERB11.1* shares the MHC class I conserved residues (Fig. 6), the overall identity to HLA-A2 is only 27%


HUMMHAE 1
HUMMHC1BW
DOGMHDL9A
HUMHLA1EA
SHPMHCE 1
BOVMHBOLA
MUSMHQ89D
CHKMHC1AA
CHKMHBFVB
XELMHCIAA
AIEMHLC1A
MUSIGFCRE
RATMHREC
perb11

a

Fig. 9 Phylogenetic analysis indicates that *PERB11.1* is more closely related to nonmammalian than mammalian MHC class I molecules. Dendrograms display phylogenetic distances separating PERB11.1 from related molecules. These distances were calculated, using **A** the Neighbor-joining and **B** protein parsimony methods. Both trees are similar in that the molecules are organized roughly into three groups: mammalian MHC class I, nonmammalian MHC class I and MHC class I-like molecules including PERB11.1. HUMHAE1 = HLA-E, HUMMHC1BW = HLA-Cw, HUMHLA1EA = HLA-E, DOGMHDL9A = dog MHC class I, SHPMHCE1 = sheep MHC class I, BOVMHBOLA = cattle MHC class I, MUSMHQ89D = mouse Q8, XELMHC1AA = frog MHC class I, CHKMHC1AA = chicken BF, CHKMBFVB = chicken BF, AIEMHLC1A = lizard MHC class I, MUSIGFCRE = mouse IgG FcR, RATMHREC = Rat IgG Fc

whereas others are conserved throughout (e.g., tryptophan at positions 52 and 62 and cysteine at position 103). In several instances, putative insertions/deletions distinguish mammalian class I from the other two groups (e.g., at positions 46 and 47, 57, 93, 106). The relationship between these 3 groups of proteins was confirmed by phylogenetic analysis using the Neighbor-joining and protein parsimony methods (Fig. 9.). It is concluded that PERB11.1 is a member of a new class of proteins encoded in the human MHC. Based on protein similarities, the molecules may be functionally more closely related to FcR and nonmammalian class I than to HLA class I.

Discussion

Analysis of anonymous genomic DNA sequence from the region centromeric of the *HLA-B* locus has led to the identification of a new MHC gene, *PERB11.1*. The strategy used has identified a gene which is polymorphic and selectively expressed, two of the characteristics required by a gene which may be a candidate in the etiology of autoimmune diseases such as IDDM and MG.

Polymorphism of *PERB11.1* has been demonstrated by diverse approaches, ranging from RFLP analysis to DNA sequencing. At the protein level, 5 of 7 nucleotide differences observed lead to non-synonymous amino acid substitutions, suggesting that these polymorphisms may be functionally relevant. Variations in gene copy number, as

suggested by the *Bam* HI analysis, also appear to exist and have been confirmed by PFGE analysis (data not shown; Gaudieri and co-workers, manuscript in preparation).

Duplication is a feature of most polymorphic MHC genes. In the case of *PERB11*, multiple copies have been demonstrated by RFLP analysis and confirmed by mapping of YAC clones. Although a proportion of YAC clones have been reported to be unstable and/or unfaithful replicas of the genome (Garza et al. 1989; Neil et al. 1990), this was not the case for the selected clones used in this study, at least for the regions carrying *PERB11*. Adequate stability is indicated by the fact that the fragments from the YAC clones are present in the corresponding genomic DNA. Furthermore, several overlapping YAC clones carry the same size *PERB11* fragments (Fig. 4).

Northern analysis of *PERB11.1* indicates that the gene sequence detects multiple transcripts in different tissues (Fig. 3). It is unclear whether these transcripts reflect multiple genes and/or isoforms. In general, however, the expression level of the *PERB11* gene family is low when compared with highly expressed proteins, such as actin and *BAT1* (Spies et al. 1989a, b). Inducibility is suggested by the identification of a consensus interferon responsive element (Cohen et al. 1988; Fan et al. 1989) located ~2.5 kb upstream of the *PERB11.1* sequence.

The expression status of other copies of *PERB11* is unknown. Some of them could represent pseudogenes or gene fragments, similar to the ones described for the human MHC class I gene family (Geraghty et al. 1992a; Geraghty et al. 1992b).

A striking similarity in gene structure has been observed between *PERB11.1* and members of the MHC class I family. The α1 and α2 domains are encoded by separate exons and the intron-exon boundaries are always situated between the first and second nucleotide of the first codon (Fig. 5), features common to most members of the immunoglobulin supergene family (Kroemer et al. 1990). However, the first intron of *PERB11.1* does not conform to MHC class I structure, being 1.7 kb in size compared with the 200–300 bp size of the first intron of MHC class I genes (Kroemer et al. 1990). Of particular interest is the finding that *PERB11.1* codes for a molecule similar to MHC class I, but more closely related to nonmammalian MHC class I molecules. Thus, the *PERB11* gene family could share an ancestor with mammalian and nonmammalian class I molecules, which probably diverged some 350 million years ago.

Although the function of the *PERB11* gene family has yet to be elucidated, the predicted structure of *PERB11.1* suggests that this molecule could be a receptor. The function of MHC class I-like molecules is quite varied. Some molecules, such as mouse H2-M, a non-classical MHC class I protein, have very specialized functions, like the presentation of N-formylated peptides of bacterial or mitochondrial origin (Pamer et al. 1992). In contrast, the rat and mouse neonatal IgG Fc receptor, also associated with β_2-microglobulin, and expressed on intestinal epithelial cells, specifically binds IgG and mediates the uptake of this protein from milk (Simister et al. 1989; Ahouse et al. 1993). Thus, MHC class I-like molecules can be involved in very different functions. We propose that the *PERB11* gene family will encode receptors which are involved in the regulation of immune responses.

In relation to function, the tissue-specific expression of *PERB11* is potentially interesting. The 3 kb transcript in muscle may suggest that *PERB11* could be directly involved in the etiology of GMG. However, it is known that GMG results from the destruction of motor end plates caused by high titres of Anti-AchR (Garlepp et al. 1982). Thus, a more likely explanation is that *PERB11* may be an immune modulator capable of enhancing specific antibody production. In this context, it is compelling that skeletal muscle associated macrophages with a distinctive dendritic morphology have been identified in the rat (Honda et al. 1990).

Further characterization of the *PERB11* gene family, starting with the isolation of the full-length cDNA for *PERB11.1* will provide further insight into the function of this new family of molecules encoded in the MHC.

Acknowledgments We would like to thank L. Smith, G. Saueracker, S. Gaudieri, C. Causerano, R. Darovic, N. Economou, and D. Witt for technical assistance. We are grateful to Dr. P. Pontarotti for providing the JP1 probe, Dr. D. Le Paslier and Dr. H. Abderrahim for providing the YAC clones, and M. Pinelli who helped with the computer analysis. C. Leelayuwat is from the Department of Clinical Immunology, Faculty of Associated Medical Sciences, Khon Kaen University, Thailand and was an EMSS scholar sponsored by AIDAB. This work was supported by the National Health and Medical Research Council and the Immunogenetics Research Foundation. Publication 9351 of The Centre for Molecular Immunology and Instrumentation of the University of Western Australia, Royal Perth Hospital and Sir Charles Gairdner Hospital.

References

Abraham, L. J., Du, D. C., Zahedi, K., Dawkins, R. L., and Whitehead, A. S. Haplotypic polymorphisms of the *TNFB* gene. *Immunogenetics 33:* 50–53, 1991

Ahouse, J. J., Hagerman, C. L., Mittal, P., Gilbert, D. J., Copeland, N. G., Jenkins, N. A., and Simster, N. E. Mouse MHC class I-like Fc receptor encoded outside the MHC. *J Immunol 151:* 6076–6088, 1993

Araki, T., Gejyo, F., Takagaki, K., Haupt, H., Schwick, G., Burgi, W., Marti, T., Schaller, J., Rickli, E., Brossmer, R., Atkinson, P. H., Putnam, F. W., and Schmid, K. Complete amino acid sequence of human plasma Zn-α2-glycoprotein and its homology to histocompatibility antigens. *Biochemistry 85:* 679–683, 1988

Balk, S. P., Bleicher, P. A., and Terhost, C. Isolation and characterization of a cDNA and gene coding for a fourth CD1 molecule. *Immunology 86:* 252–256, 1989

Chimini, G., Boretto, J., Marguet, D., Lanau, F., Lauguin, G., and Pontarotti, P. Molecular analysis of the human MHC class I region using yeast artificial chromosome clones. *Immunogenetics 32:* 419–426, 1990

Cohen, B., Peretz, D., Vaiman, D., Benech, P., and Chebath, J. Enhancer-like interferon responsive sequences of the human and murine (2′–5′) oligoadenylate synthetase gene promoters. *EMBO J 7 (No 5):* 1411–1419, 1988

Dawkins, R. L., Christiansen, F. T., Kay, P. H., Garlepp, M. J., McCluskey, J., Hollingsworth, P. N., and Zilko, P. J. Disease associations with complotypes, supratypes and haplotypes. *Immunol Rev 70:* 5–22, 1983

Degli-Esposti, M. A., Abraham, L. J., McCann, V., Spies, T., Christiansen, F. T., and Dawkins, R. L. Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM. *Immunogenetics 36:* 345–356, 1992a

Degli-Esposti, M. A., Andreas, A., Christiansen, F. T., Schalke, B., Albert, E., and Dawkins, R. L. An approach to the localization of the susceptibility genes for generalized myasthenia gravis by mapping recombinant ancestral haplotypes. *Immunogenetics 35:* 355–364, 1992b

Degli-Esposti, M. A., Leelayuwat, C., and Dawkins, R. L. Ancestral haplotypes carry haplotypic and haplospecific polymorphisms of BAT1: possible relevance to autoimmune disease. *Eur J Immunogenet 19:* 121–127, 1992c

Degli-Esposti, M. A., Griffiths, M. S., Daly, L. N., Witt, C. S., Simons, M. J., Carcassi, C., Albert, E. D., Giphart, M. J., and Dawkins, R. L. Characterization of the 4AOH cell line panel including new data for the 10HIWS panel. *Hum Immunol 38:* 3–16, 1993

Durinovic-Bello, I., Schendel, D. J., Kastelan, A., and Segurado, O. G. A novel diabetes-susceptibility HLA haplotype is present in the Croatian population. *Tissue Antigens 41:* 107–109, 1993

Ellis, S. A., Strachan, T., Palmer, M. S., and McMichael, A. J. Complete nucleotide sequence of a unique HLA class I C locus product expressed on the human choriocarcinoma cell line bewo. *J Immunol 142:* 3281–3285, 1989

Ennis, P. D., Jackson, A. P., and Parham, P. Molecular cloning of bovine class I MHC cDNA. *J Immunol 141:* 642–651, 1988

Ennis, P. D., Zemmour, J., Slater, R. D., and Parham, P. Rapid cloning of HLA-A,B cDNA by using the epolymerase chain reaction: frequency and nature of errors produced in amplification. *Immunology 87:* 2833–2837, 1990

Fan, C. M. and Maniatis, T. Two different virus-inducible elements are required for human B-interferon gene regulation. *EMBO J 8 (No 1):* 101–110, 1989

Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. W., and Vogelstein, B. Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science 247:* 49–56, 1990

Feinberg, A. P. and Vogelstein, B. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem 132:* 6–10, 1983

Freije, J. P., Fueyo, A., Uria, J., and Lopez-Otin, C. Human Zn-α2-glycoprotein cDNA cloning and expression analysis in benign and malignant breast tissues. *FEBS Lett 290:* 247–249, 1991

Garlepp, M. J., Kay, P. H., and Dawkins, R. L. The diagnostic significance of autoantibodies to the acetylcholine receptor. *J Neuroimmunol 3:* 337–350, 1982

Garza, D., Ajioka, J. W., Burke, D. T., and Hartl, D. L. Mapping the Drosophila genome with yeast artificial chromosome. *Science 246:* 641–646, 1989

Geraghty, D. E., Koller, B. H., Hansen, J. A., and Orr, H. T. The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. *J Immunol 149:* 1934–1945, 1992a

Geraghty, D. E., Koller, B. H., Pei, J. I., and Hansen, J. A. Examination of four HLA class I pseudogenes: common events in the evolution of HLA genes and pseudogenes. *J Immunol 149:* 1947–1956, 1992b

Grossberger, D., Hein, W., and Marcus, A. Class I major histocompatibility complex cDNA clones from sheep thymus: alternative splicing could make a long cytoplasmic tail. *Immunogenetics 32:* 77–87, 1990

Grossberger, D. and Parham, P. Reptilian class I major histocompatibility complex genes reveal conserved elements in class I structure. *Immunogenetics 36:* 166–174, 1992

Honda, H., Kimura, H., and Rostami, A. Demonstration and phenotypic characterization of resident macrophages in rat skeletal muscle. *Immunology 70:* 272–277, 1990

Karlin, S. and Altschul, S. F. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc Natl Acad Sci USA 87:* 2264–2268, 1990

Kaufman, J., Andersen, R., Avial, D., Engberg, J., Lambris, J., Salomonsen, J., Welinder, K., and Skjodt, K. Different features of the MHC class 1 heterodimer have evolved at different rates. *J Immunol 148:* 1532–1546, 1992

Kelly, H., McCann, V. J., Kay, P. H., and Dawkins, R. L. Susceptibility to IDDM is marked by MHC supratypes rather than individual alleles. *Immunogenetics 22:* 643–651, 1985

Kroemer, G., Zoorob, R., and Auffray, C. Structure and expression of a chicken MHC class I gene. *Immunogenetics 31:* 405–409, 1990

Leelayuwat, C., Abraham, L. J., Tabarias, H., Christiansen, F. T., and Dawkins, R. L. Genomic organization of a polymorphic duplicated region centromeric of *HLA-B*. *Immunogenetics 36:* 208–212, 1992

Leelayuwat, C., Zhang, W. J., Townend, D. C., Gaudieri, S., and Dawkins, R. L. Differences in the central MHC between man and chimpanzee: implications for autoimmunity and acquired immune deficiency development. *Hum Immunol 38:* 30–41, 1993

Matsuura, A., Schloss, R., Shen, F., Tung, J., Hunt, S. W., Fisher, D. A., Hood, L. E., and Boyse, E. A. Expression of the *Q8\9* gene by T cells of the mouse. *Immunogenetics 30:* 156–161, 1989

Mizuno, S., Trapani, J. A., Koller, B. H., Dupont, B., and Yang, S. Y.: Isolation and nucleotide sequence of a cDNA clone encoding a novel HLA class 1 gene. *J Immunol 140:* 4024–4030, 1988

Neil, D. L., Villasante, A., Fisher, R. B., Vetrie, D., Cox, B., and Tyler-Smith, C. Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic Acids Res 18:* 1421–1428, 1990

Pamer, E. G., Wang, C. R., Flaherty, L., Fischer Lindahl, K., and Bevan, M. J. H-2M3 presents a listeria monocytogenes peptide to cytotoxic T lymphocytes. *Cell 70:* 215–223, 1992

Rothermel, E., Heine, L., Wurst, W., and Gunther, E. Characterization of a class I b gene of the rat major histocompatibility complex. *Immunogenetics 38:* 82–91, 1993

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1. *Molecular cloning: A Laboratory Manual (edn 2),* Cold Spring Harbor Laboratory, Cold Spring Harbor, 1989

Sarmiento, M. U. and Storb, R. Nucleotide sequence of a dog class I cDNA clone. *Immunogenetics 31:* 400–404, 1990

Segurado, O. G., Iglesias-Casarrubios, P., Morales, P., Martinez-Laso, J., Partanen, J., Campbell, R. D., and Villena, A. Genetic structure of the novel low-frequency haplotype *HLA-B49,SC01,DR4* and its contribution to insulin-dependent diabetes susceptibility. *Immunogenetics 37:* 67–72, 1992

Shum, B. P., Avila, D., Pasquier, L. D., Kasahara, M., and Flajnik, M. F. Isolation of a classical MHC class I cDNA from an amphibian evidence for only one class I locus in the xenopus MHC. *J Immunol 151:* 5376–5386, 1993

Simister, N. E. and Mostov, K. E. An Fc receptor structurally related of MHC class I antigens. *Nature 337:* 184–187, 1989

Spies, T., Blanck, G., Bresnahan, M., Sands, J., and Strominger, J. L. A new cluster of genes within the human major histocompatibility complex. *Science 243:* 214–217, 1989a

Spies, T., Bresnahan, M., and Strominger, J. L. Human major histocompatibility complex contains a minimum of 19 genes between the complement cluster and HLA-B. *Proc Natl Acad Sci USA 86: 8955–8958, 1989b*

Synder, E. E. and Stormo, G. D. Identification of coding regions in genomic DNA sequences: an application of dynamic programming and neural networks. *Nucleic Acids Res 21:* 607–613, 1993

Uberbacher, E. C. and Mural, R. J. Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc Natl Acad Sci USA 88:* 11261–11265, 1991

Wu, X., Zhang, W. J., Witt, C. S., Abraham, L. J., Christiansen, F. T., and Dawkins, R. L. Haplospecific polymorphism between HLA B and tumour necrosis factor. *Hum Immunol 33:* 89–97, 1992

Zhang, W. J., Degli-Esposti, M. A., Cobain, T. J., Cameron, P. U., Christiansen, F. T., and Dawkins, R. L. Differences in gene copy number carried by different MHC ancestral haplotypes: quantitation after physical separation of haplotypes by pulsed field gel electrophoresis. *J Exp Med 171: 2101–2114, 1990*

Note added in proof: Recent analysis of the genomic sequence downstream of the *PERB11.1* α2 domain has revealed the presence of an α3 domain with similar sequence relationships to those displayed by the sequences of the α1 and α2 domains. After acceptance of this manuscript, M. Pinelli identified a GenBank sequence (L14848, Bahram and co-workers), released on 7 July 1994, which shows nucleotide similarity of 94–97% and amino acid similarities of 94, 96, and 79% for the α1, 2