Immunogenetics (1996) 45: 97–107 © Springer-Verlag 1996

ORIGINAL PAPER

Wenyi Cai · Wei Cao · Lizi Wu · Ginger E. Exley
Gerald L. Waneck · Barry L. Karger
Carol M. Warner

Sequence and transcription of Qa-2-encoding genes in mouse lymphocytes and blastocysts

Received: 29 May 1996 / Revised: 5 August 1996

Abstract The protein product of the mouse preimplantation embryo development (*Ped*) gene, which controls the rate of preimplantation embryonic cleavage division and subsequent embryo survival, is the Qa-2 antigen. This major histocompatibility complex (MHC) class Ib protein is encoded by four genes, *Q6, Q7, Q8*, and *Q9*. The present study was undertaken to begin to elucidate which of the four Qa-2-encoding genes are responsible for the *Ped* gene phenotype in the C57BL/6 mouse ($H2^b$). First, restriction maps of the four genes, using 25 restriction enzymes, were created. The RE maps confirmed that *Q6* is similar to *Q8* and *Q7* is similar to *Q9*, but that the *Q6/Q8* gene pair differs from the *Q7/Q9* gene pair. The genomic DNA sequences of *Q6* and *Q8* were determined, as well as the DNA sequences of exons 4–8 of *Q9*, and the 5′ regulatory regions of *Q6, Q8*, and *Q9*. This DNA sequence information, combined with the published DNA sequence information for the entire *Q7* gene and exons 1–3 of *Q9*, allowed us to design primers for reverse transcription-polymerase chain reaction that could distinguish which of the four genes were transcribed in mouse lymphocytes and embryos. It was found that all four genes are transcribed in lymphocytes, but only *Q7* and *Q9* are transcribed in mouse embryos. Thus, both *Q7* and *Q9* are candidates for the genes responsible for the *Ped* gene phenotype.

The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers U57392, U57393, U57395, U57396, U57397, U57398, U57399, and U57400

W. Cai · W. Cao · L. Wu · G. E. Exly · C. M. Warner (✉)
Department of Biology, Northeastern University, Boston, MA 02115, USA

G. L. Waneck
Department of Surgery, Massachusetts General Hospital, Charlestown, MA 02129, USA

W. Cai · B. L. Karger
Department of Chemistry and the Barnett Institute of Chemical Analysis and Materials Science, Northeastern University, Boston, MA 02115, USA

Introduction

The mouse major histocompatibility complex (MHC) encodes three classes of proteins, class I, class II, and class III. The genes for these proteins reside in approximately 2600 kilobases (kb) of DNA encompassing six subregions, *K, I, S, D, Q*, and *TL*. The class I genes are of two types, class Ia-encoded in the *K* and *D* subregions, and class Ib-encoded in the *Q* and *TL* subregions. The class Ia proteins are highly polymorphic and ubiquitously expressed on virtually all cells and organs of an individual. The function of the class Ia proteins is to present 8–10 amino acid peptides derived from intracellular degradation of foreign antigens, such as viral antigens, to cytotoxic T cells, which can then destroy the infected cells. The class Ib proteins, on the other hand, are much less polymorphic and of much more limited tissue distribution. Possible functions of the class Ib proteins are just beginning to be elucidated (reviewed in Shawar et al. 1994; Beckman and Brenner 1995). The class Ib proteins may function in cell-cell interactions during development and differentiation.

The class Ib protein of particular interest to our laboratory is the Qa-2 antigen, first described by Flaherty and coworkers (1976). The reason is that the Qa-2 protein has been shown to be the product of the preimplantation embryo development (*Ped*) gene that controls the rate at which preimplantation embryos cleave and subsequent embryo survival (Warner et al. 1987a, b, 1991, 1993; Tian et al. 1992; Xu et al. 1993, 1994).

The Qa-2 protein of $H2^b$ (C57BL/10 and C57BL/6) mice is encoded by four genes in the *Q* region, *Q6, Q7, Q8*, and *Q9* (reviewed in Flaherty et al. 1990; Stroynowski 1990; Shawar et al. 1994). These genes have the typical class I *MHC* gene structure: exon 1 (leader peptide), exon 2 (α1 domain), exon 3 (α2 domain), exon 4 (α3 domain), exon 5 (transmembrane domain), exons 6, 7, and 8 (cytoplasmic domains). Based on restriction enzyme mapping, the *Q6* and *Q8* genes are seen to be very similar, as are the *Q7* and *Q9* genes (Weiss et al. 1984). The sequences of part of the *Q6* gene [intron 6, exon 7, intron 7, exon 8 and 3′

Table 1 Primers for analyzing Q6/Q8 and Q7/Q9 gene expression by RT-PCR

Primer (length)	Sequence	Location	Size of product
Q6/Q8 upstream (21mer)	5′GTGGAGCCCCGGTTCATTATC 3′	exon 2	364 bp
Q6/Q8 downstream (18mer)	5′ATGTCCGCCGCCGTCCAG 3′	exon 3	
Q7/Q9 upstream (21mer)	5′TGGTATTGCAGAGAAAGACCA 3′	exon 3	380 bp
Q7/Q9 downstream (20mer)	5′ATCTCCCCCATCTCAGGGTA 3′	exon 4	

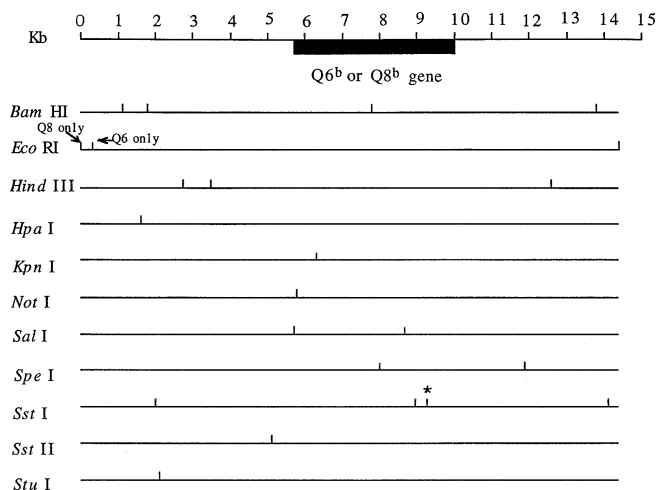


Fig. 1 Restriction enzyme map of the *Q6* and *Q8* genes. The area encompassing the exons is shown as a *solid block*. For *Sst* I, the orientation of the small fragment indicated by the *asterisk* could not be determined. It may be at the other end of the larger fragment

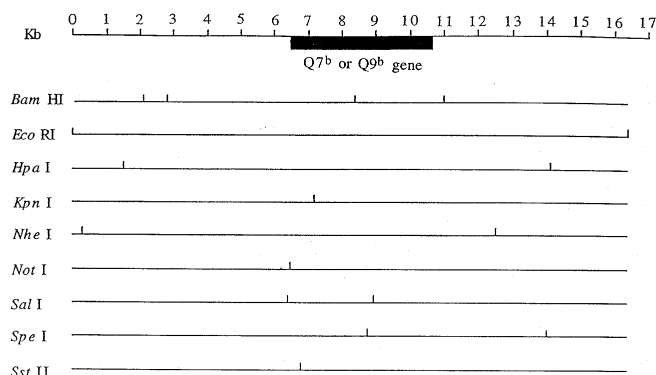


Fig. 2 Restriction enzyme map of the *Q7* and *Q9* genes. The area encompassing the exons is shown as a *solid block*

untranslated (UT) (Tine et al. 1990)], the complete *Q7* gene except an 1100 base pair (bp) gap in intron 3 (Devlin et al. 1985), the complete *Q8* gene except most of intron 3 and a few other gaps (Devlin et al. 1985; Elliott et al. 1989), and part of the *Q9* gene [exon 1, intron 1, exon 2, intron 2, exon 3 (Devlin et al. 1985) and intron 6, exon 7, intron 7, exon 8 and 3′UT (Tine et al. 1990)] have been reported.

In order to evaluate which of the Qa-2 antigen-encoding genes are expressed in mouse embryos, and therefore to contribute to the *Ped* gene phenotype, it is necessary to have complete DNA sequence information from each of the four genes encoding the Qa-2 antigen. In this paper we report new DNA sequence information for the *Q6* gene, the *Q8* gene, exons 4, 5, 6, 7, and 8 of the *Q9* gene, and the 5′ regulatory regions of the *Q6, Q8*, and *Q9* genes. This new sequence information, in conjunction with published sequence information, was used to design primers for RT-PCR to test which of the *Q6, Q7, Q8*, and *Q9* genes are transcribed in mouse splenic lymphocytes and blastocyst stage embryos.

Materials and methods

Mice

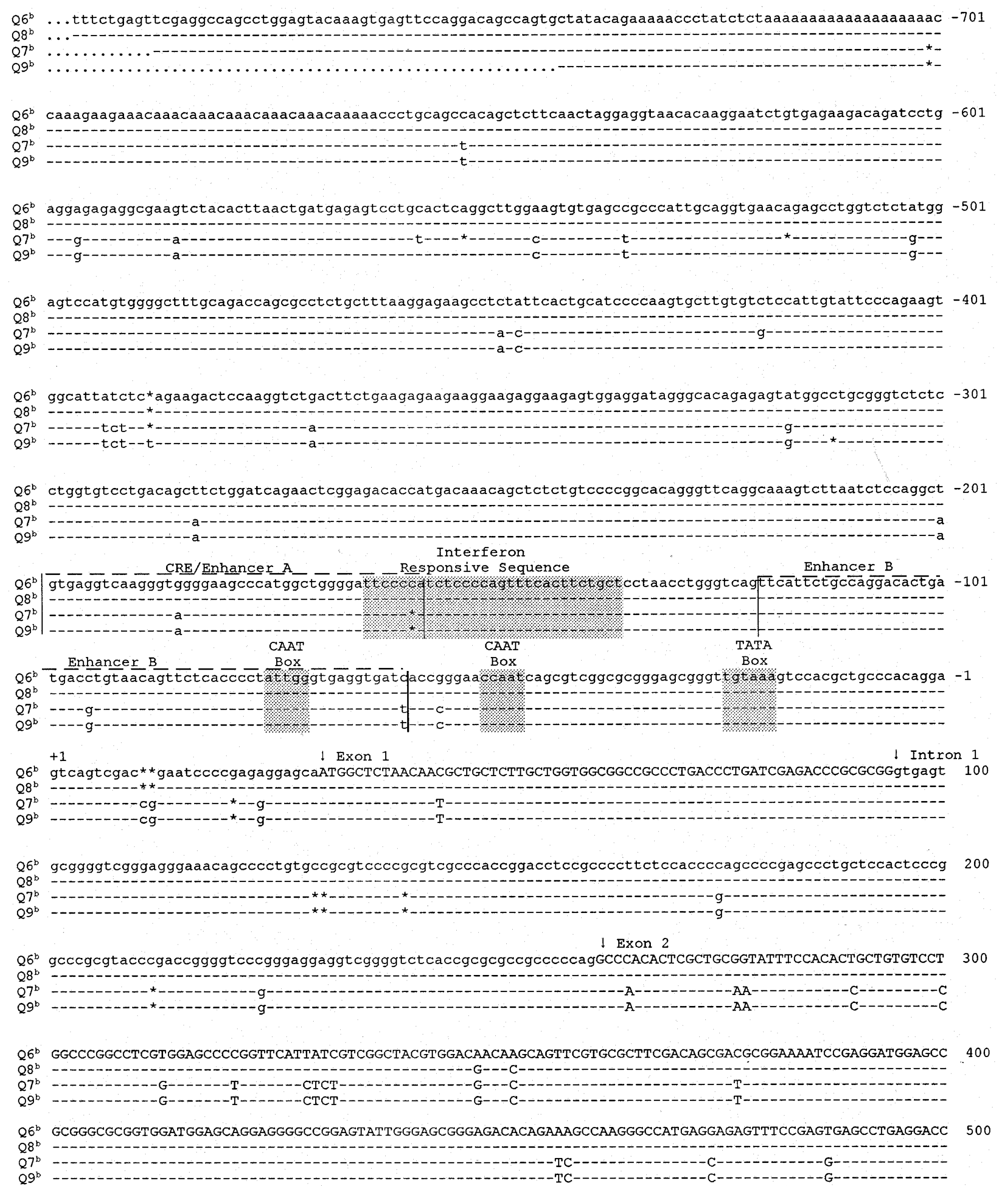
C57BL/6 mice, which possess the *Q6, Q7, Q8*, and *Q9* genes, and CBA/Ca mice, which have a deletion for these genes, were purchased as mating pairs from the Jackson Laboratory (Bar Harbor, ME) and then bred in our own laboratory. Mice were housed in an American

Table 2 Restriction enzyme polymorphisms in Q7 and Q9

Fragment sizes (kb)			
Pst I		*Sst* I	
Q7	Q9	Q7	Q9
2.9	2.9	7.6	–
2.6*	2.6	5.1	5.1
2.0	2.0	–	4.3
1.8	1.8	3.7	3.7
1.75	1.75	–	3.2
1.5	1.5	1.25	1.25
1.3	1.3	0.35	0.35
0.75	0.75	0.3	0.3
0.7	0.7		
0.6	0.6		
0.58	–		
0.55	0.55		
–	0.5		
–	0.35		
0.3	–		
0.2	0.2		
0.18	0.18		

* Underlined fragments are from the vector

Fig. 3 (For legend see p. 103)



Association for the Accreditation of Laboratory Animal Care (AAALAC)-approved facility. Lights were on from 4 a.m. to 6 p.m. eastern standard time, an Old Guilford Mouse Breeder Diet #911R (Emory Mouse Company, Guilford, CT) was provided, and water was supplied ad libitum. All experimental protocols followed the NIH guidelines. Mice were superovulated by injection of 10 IU eCG (Sigma, St. Louis, MO) at 3 p.m. e.s.t. followed by 5 IU hCG (Sigma) 48 h later. Embryos (blastocysts) were collected at 89 h post-hCG injection. Splenic lymphocytes were isolated using a modification of the procedure described by Parish and co-workers (1974) using Ficoll-Hypaque (density = 1.084) and a Beckman microfuge. The splenic lymphocytes were used as a control for the experiments on blastocysts.

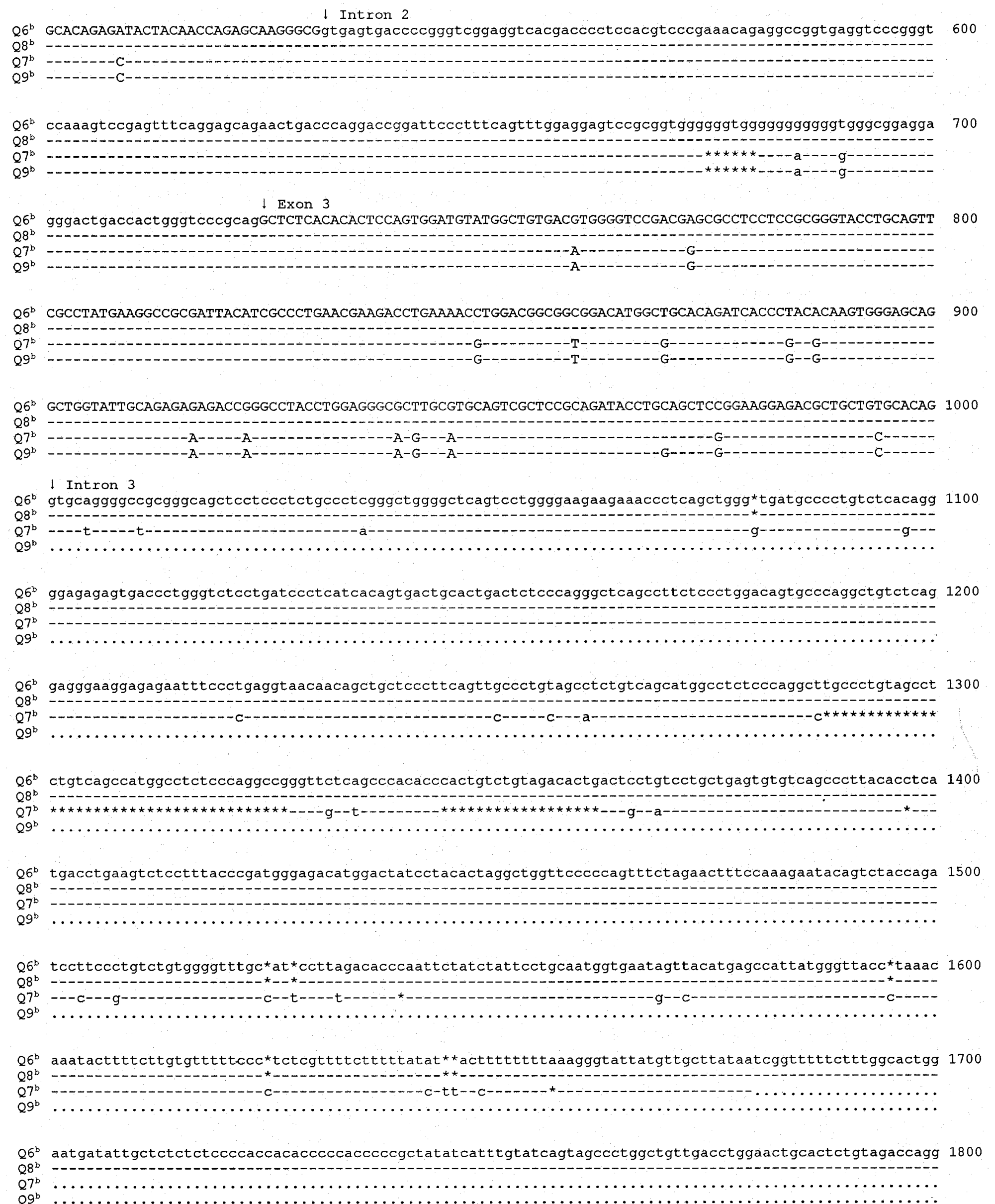
Q region genes

Four plasmid clones containing the *Q6, Q7, Q8*, and *Q9* genes were derived from cosmids originally isolated from a C57BL/10 (*H2b*) mouse cosmid library as described by Weiss and co-workers (1984). The 14–16 kb plasmid clones (*Q6*: 14.1 kb; *Q7*: 16.4 kb; *Q8*: 14.4 kb; *Q9*: 16.4 kb) were prepared by ligation into a pBR 327 vector (3.293 kb) at the *Eco* RI site. Bacterial cultures were prepared by streaking the frozen glycerol stock on a plate, picking a single colony from the plate, and growing it in 500 ml LB/Amp with shaking at 37 °C. The plasmid DNA was isolated by using a QIAGEN plasmid kit (QIAGEN, Chadsworth, CA) with a Maxi Qiagen-tip 500 following the protocol supplied by the manufacturer.

Restriction enzyme maps

The four plasmid DNAs were digested with restriction enzymes obtained from Gibco-BRL (Grand Island, NY) and New England BioLabs (Beverly, MA), according to the supplier's instructions. Twenty-five different restriction enzymes were used: *Aat* II, *Bam* HI, *Cla* I, *Eco* RI, *Hin* dIII, *Hpa* I, *Kpn* I, *Mlu* I, *Msp* I, *Nae* I, *Nar* I, *Nhe* I, *Not* I, *Nru* I, *Pst* I, *Pvu* II, *Sac* II, *Sal* I, *Sfi* I, *Spe* I, *Sst* I, *Sst* II, *Stu* I, *Taq* I, and *Xho* I. All enzyme digests were coupled with *Eco* RI digestion because the DNA was ligated into the vector at the *Eco* RI site. Double digestions with pairs of enzymes were used to create the restriction maps.

Fig. 3 (Continued, for legend see p. 103)



DNA sequences

DNA sequencing was performed using a Sequenase Quick-Denature plasmid sequencing kit (US Biochemicals, Cleveland, OH) and α-^{35}S-dATP (specific activity 1000 µCi/mmole (Amersham Life Science, Arlington Heights, IL). About 4 µg of DNA were denatured at 100 °C in the presence of glycols and the sequencing reaction performed according to the manufacturer's protocol. The fragments were analyzed on a 6% polyacrylamide/urea gel. Synthetic oligonucleotides were purchased from Tufts University to use as primers. The results were analyzed using the DNASIS software program for the PC (Hitachi Software, San Bruno, CA).

Reverse transcription-polymerase chain reaction (RT-PCR)

RT-PCR was performed on splenic lymphocytes and embryos by using the reagents in a kit from Perkin Elmer (Branchburg, NJ) and a modification of the protocols of Jin and co-workers (1992) and Pal and co-workers (1993). Briefly, splenic lymphocytes and embryos were washed three times in phosphate buffered saline (PBS) and then transferred in a minimum volume of PBS to a 0.5 ml thin-walled reaction tube. Two µl of lysis buffer containing 1.0 unit RNase inhibitor, 10 mM DTT, and 0.5% NP-40 were added. Then 1 µl of 50 µM random hexamer primers and RNase-free water to bring the total volume to 8 µl were added. In some experiments splenic lymphocyte RNA was isolated using TRISOLV reagent (Biotecx Lab, Houston, TX) and an 8 µl sample used directly in the subsequent procedures. The 8 µl sample was heated to 70 °C for 5 min and cooled to 30 °C for another 5 min. To obtain cDNA, 4 µl $MgCl_2$ (25 mM), 2 µl 10 × PCR buffer II, 4 µl dNTPs (10 mM), 1 µl (50 units) MuLV reverse transcriptase, 0.5 units RNase inhibitor and RNAse-free water to bring the reaction volume to 20 µl were added and the reaction incubated for 3 h at 37 °C. At the end of the incubation time the reaction was heated at 99 °C for 5 min to inactivate the reverse transcriptase. The amplification of the cDNA was performed in a 100 µl reaction mixture containing an aliquot (usually 5 µl) of the 20 µl of the cDNA product described above, 4.5 µl $MgCl_2$ (25 mM), 10 µl 10 × PCR buffer II, 8 µl dNTPs (10 mM), 2 µl upstream primer, and 2 µl downstream primer (10 µM each), 5 units AmpliTaq DNA polymerase (Perkin Elmer), and sterile water. Upstream and downstream primers for the *Q6, Q7, Q8*, and *Q9* genes were designed based

Fig. 3 (Continued, for legend see p. 103)

```
Q6b ctggccttgaactcagaaatcagcctggctctgcctctgcctctgcctctgtctctgcctctgcctcccaagtgtctgggattaaaggcttgggccacca 1900
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b ccagtgggcagaagaaaggttcctgcgagcttaaaatgttttctggcagaattaaccatccagatcacacctgatatccctgtgccccaccaagttacag 2000
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b tgctccccctggtgaatcagaacttggactctgagagacagggtcttctgcaatccaggcctgagtgagagggaagaccacacaccctgtgagcccactg 2100
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b tgttccagtgagtgctgcactggggtccacagcacactccagggatcctgtgtgacacatctgtaccttgtcccccagagtcaggggctgggagtcattt 2200
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b tctctggctgagtgtcagaggttcaccacatttctgctacacactccctgatggctgtttacttggactgacagttaatgttggtcagcaagatgaccac 2300
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b agtggtttagtctcaatggtgtcactcttccagtagcatatggtcctgatttctaatttagatacgaactcaaacacatatgaaatttcttattttccat 2400
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b tccatcttccattatatagctacctatctcgtgctattgaacatcacataaggatgaccatgtttacccactggctcatgtggattccctcttagcttct 2500
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b gagtcccctcaggaaaatgtgcagtcctgtgctgaggggaccagctctgcctgcaggtcactagtgccatgacagttaaagtgttcatacagacacatag 2600
Q8b ----------------------------------------------------------------------------------------------------
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b ttcattgtaattactgatttaacgttgtcttggcagttttcagtttgcatttatttatttatttatttatta*************************** 2700
Q8b -------------------------------------------------------------------------tttatttatttatttatttatttattt
Q7b ....................................................................................................
Q9b ....................................................................................................

Q6b *atgcatggaagtacactgttgctgtactgatggttgtttgcctttgtgtggttgttgggaattgaatttttttttag*acctctctttgctctggtcga 2800
Q8b a-----------------------------------------------------------------------------*---------------------
Q7b ............................................................................--g---------------------
Q9b ....................................................................................................

Q6b ccctgctcactccggtcaactcctatgggtcaactctgctcactcagtccctgcttgttctggcccaaagatttattttatttattttatttattacata 2900
Q8b ----------------------------------------------------------------------------------------------------
Q7b ------------------------------------------------------------------------------*---*----*---*--*-----
Q9b ....................................................................................................

Q6b aatacactgtagctgacttcagatgcaccagaagagggcgtcagatctcattacagatggttgtgagtcaccatgtggttgctggggtttgaactcagga 3000
Q8b ----------------------------------------------------------------------------------------------------
Q7b ----------------------------------------------------------------------------------------------------
Q9b ....................................................................................................

                                                                                                  ↓Exon 4
Q6b ccttcaaaagagcagtcagtgctcttaccctctgagccatctccccagtcctcagtttgtcttcttaattatgcgatttcttgaatcttccaaacagATC 3100
Q8b ----------------------------------------------------------------------------------------------------
Q7b ----------------------------------------------------------------------------------------------------
Q9b .................................................................................................---
```

on published sequence information and the data reported in this paper using the Oligo 5.0 program for the PC (National Biosciences, Plymouth, MN). The sequences of the exon-spanning primers to detect the *Q6/Q8* gene pair or the *Q7/Q9* gene pair are listed in Table 1.

The mixture was placed in a Perkin-Elmer Cetus Thermo Cycler (Norwalk, CT) and heated to 96 °C for 1 min, followed by two cycles using the settings of denaturation at 96 °C for 1 min, primer annealing at 58 °C for 45 s, and extension at 72 °C for 45 s. Another 40 cycles were performed using the settings of 94 °C for 1 min for denaturation, 58 °C for 45 s for annealing, and 72 °C for 45 s for extension followed by a final incubation at 72 °C for 5 min. The PCR product was then analyzed by electrophoresis on a 6% polyacrylamide gel followed by staining with ethidium bromide. A 100 bp DNA ladder (Gibco-BRL) was used as a marker to determine the size of the PCR product. The identity of the RT-PCR products was confirmed by cloning and sequencing the products using a TA cloning kit (Invitrogen, San Diego, CA).

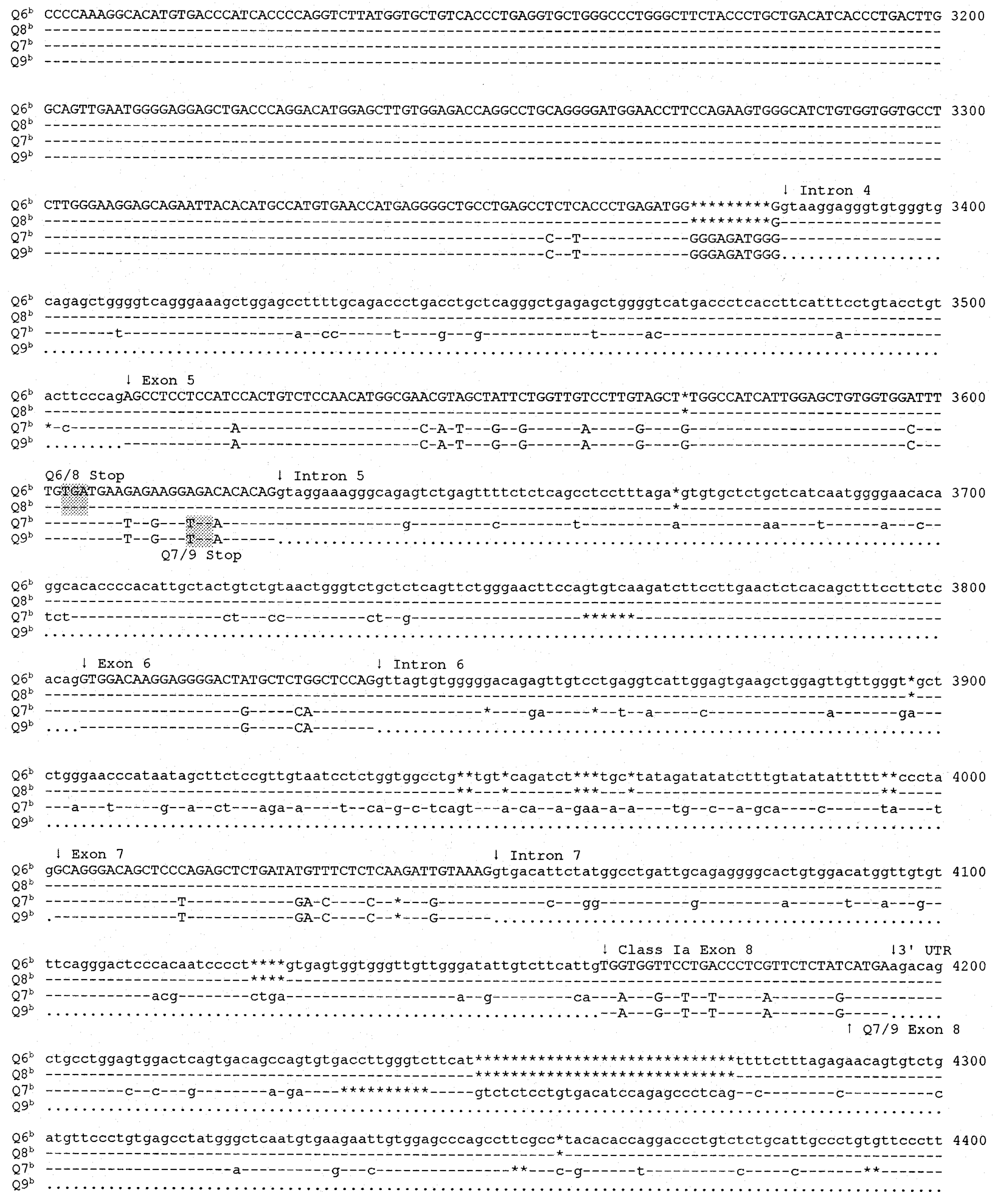
To distinguish *Q7* from *Q9* transcription, the cDNA product was electrophoresed on a 1.2% agarose gel, purified using a QIAquick Gel DNA extraction kit (QIAGEN), digested with *Pst* I, and the resulting fragments were separated on a 6% polyacrylamide gel.

Results

Restriction enzyme maps

Restriction enzyme maps for the *Q6/Q8* and the *Q7/Q9* gene pair are shown in Figures 1 and 2, respectively. Of the 25 restriction enzymes tested, eight (*Aat* II, *Cla* I, *Mlu* I, *Nae* I, *Nhe* I, *Nru* I, *Sac* II, and *Sfi* I) did not cut either the *Q6* or the *Q8* genes. Likewise, eight enzymes (*Aat* II, *Cla* I, *Mlu* I, *Nae* I, *Nru* I, *Sac* II, *Sfi* I, and *Xho* I) did not cut either the *Q7* or the *Q9* genes. Enzymes that produced too many fragments to place on the restriction maps shown in Figures 1 and 2 are *Msp* I, *Nar* I, *Pst* I, *Pvu* II, and *Taq* I for the *Q6/Q8* gene pair and *Hin* dIII, *Msp* I, *Nar* I, *Pst* I, *Pvu* II, *Sst* I, *Stu* I, and *Taq* I for the *Q7/Q9* gene pair. The only restriction fragment length polymorphism (RFLP) between

Fig. 3 (Continued, for legend see p. 103)



the *Q6* and *Q8* genes was detected using *Eco* RI (shown in Figure 1). Two RFLPs were found for the *Q7/Q9* gene pair, using *Pst* I and *Sst* I, but both of these enzymes gave too many fragments to place them in order on the map shown in Figure 2. A list of the restriction fragments from the *Q7* and *Q9* genes digested with *Pst* I and *Sst* I is shown in Table 2.

DNA and protein sequences

The DNA sequences of the *Q6, Q7, Q8*, and *Q9* genes are shown in Figure 3. The similarity of the *Q6* and *Q8* and the *Q7* and *Q9* restriction maps allowed us to design primers for the sequencing of *Q6* based on the known sequence of *Q8* and primers for the sequencing of *Q9* based on the known sequence of *Q7*. As a control, the *Q8* sequence was run alongside the *Q6* sequence. For completeness and to enable comparisons to be made, the previously published *Q7* (Devlin et al. 1985) and partial Q9 sequence (Devlin et al. 1985; Tine et al. 1990) are included in Figure 3. New DNA sequence information reported in this paper and the corresponding GenBank accession numbers are: *Q6* genomic sequence (U57393), *Q8* genomic sequence (U57392), *Q9* 5′ regulatory region (U57395), *Q9* exon 4 (U57396), *Q9* exon 5 (U57397), *Q9* exon 6 (U57398), *Q9* exon 7 (U57399), and *Q9* exon 8 (U57400).

The amino acid sequences encoded by the exons of the *Q6, Q7, Q8*, and *Q9* genes are depicted in Figure 4.

Transcription of the Q6, Q7, Q8, and Q9 genes

Knowledge of the sequence of all four Qa-2-encoding genes allowed us to design primers (Table 1) for the detection of

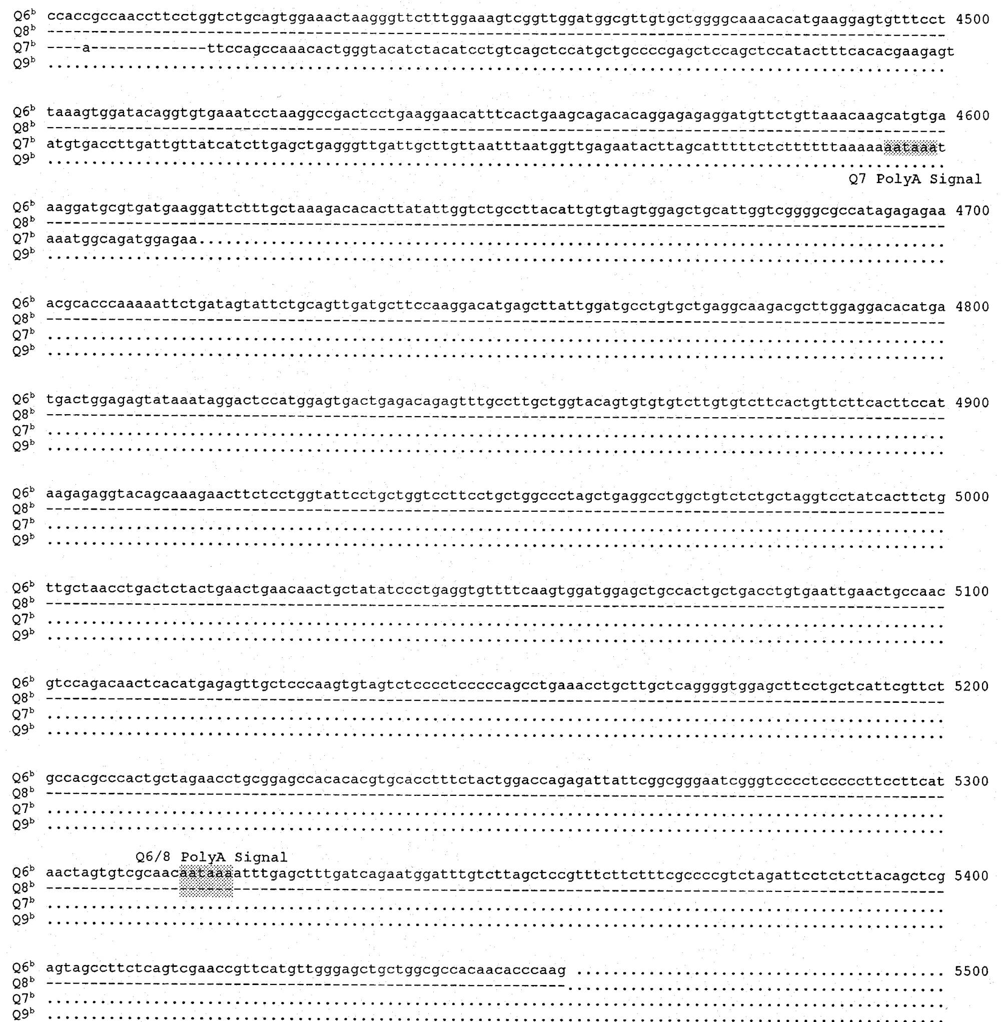


Fig. 3 DNA sequence alignment of the *Q6, Q7, Q8*, and *Q9* genes. All sequences were generated in our laboratory except for the *Q7* gene and exons 1–3 of the *Q9* gene which were published by Devlin and co-workers (1985). Similarity to the *Q6* gene is denoted by a *dash*. Missing sequence information is denoted by a *dot*. Gaps inserted to facilitate alignment are represented by *asterisks*. Exon sequences are presented as *uppercase* letters and intron sequences are presented as *lowercase* letters. Exon/intron boundaries are indicated by *arrows*. Bases are numbered relative to the start of transcription. In the 5′ regulatory region CRE/Enhancer A and Enhancer B are indicated by a *dashed line* over the sequence and *vertical lines* indicating the boundaries. The IRS, two CAAT boxes, the TATA box, the stop codons in exon 5, and the PolyA signals are indicated as *shaded areas*. It should be noted that CRE/Enhancer A and IRS overlap. The last 200 bp of the 3′UTR of the *Q7* gene is so dissimilar from that of the *Q6/Q8* gene pair that they could not be aligned based on similarity. Therefore both the *Q6* and *Q7* sequences are shown in this region

transcription of the *Q6/Q8* gene pair and the *Q7/Q9* gene pair. The results of RT-PCR on splenic lymphocytes and embryos are shown in Figure 5. It is seen that although both the *Q6/Q8* and the *Q7/Q9* gene pairs are transcribed in splenic lymphocytes, only the *Q7/Q9* gene pair is transcribed in blastocysts. In order to distinguish whether *Q7* or *Q9* or both are transcribed in blastocysts, the cDNA product was digested with *Pst* I and the fragments separated on a polyacrylamide gel. The results are shown in Figure 6. It is seen that both the *Q7* and *Q9* genes are transcribed in C57BL/6 mouse blastocysts.

Discussion

The differential expression of the Qa-2 antigen-encoding genes *Q6, Q7, Q8*, and *Q9* was studied in mouse splenic lymphocytes and blastocyst stage embryos from the C57BL/6 (*H2*^{b}) mouse by using RT-PCR. In order to distinguish transcription of the four genes from each other, complete sequence information of the four genes was required. Since sequence information for the *Q6, Q7, Q8*, and *Q9* genes was incomplete, we undertook a project to sequence these genes. Based on the *Q6, Q7, Q8*, and *Q9* gene restriction maps (Figs. 1, 2) we confirmed the original observation of Devlin and co-workers (1985) that *Q6* is very similar to *Q8* (Fig. 1) and *Q7* is very similar to *Q9*. Only one of 25 enzymes tested, *Eco* RI, resulted in a polymorphism between *Q6* and *Q8*. Only two of 25 enzymes tested, *Pst* I and *Sst* I, resulted in polymorphisms between *Q7* and *Q9* (Table 2).

The *Q6, Q7, Q8*, and *Q9* gene sequences are shown in Figure 3. The complete sequence of the *Q6* gene was determined and the complete sequence of the *Q8* gene was analyzed at the same time. The results show that the only differences in *Q6* compared with *Q8* are two bp changes in exon 2 and a 28 bp deletion in intron 3 of *Q6* compared with *Q8*. The difference between *Q6* and *Q8* in intron 3 is in a simple repetitive repeat unit, attt, the *Q6* gene has six copies and the *Q8* gene has thirteen. The two base differences in *Q6* compared with Q8 create a restriction site for an unusual restriction enzyme, *Tth111* II, which should be able to cut *Q6* but not *Q8* (Shinomiya et al. 1980).

Fig. 4 Comparison of the predicted amino acid sequences of the Qa-2 proteins encoded by the *Q6, Q7, Q8*, and *Q9* genes. Differences among the proteins are shown in *boxes*

```
               ↓(LP)                    ↓+1(α1)  9         19         29
Q6.AMI   -21  MALTTLLLLV AAALTLIETR AGPHSILRYFH TAVSWPGLVE PRFIIVGYVD
Q8.AMI   -21  MALTTLLLLV AAALTLIETR AGPHSILRYFH TAVSWPGLVE PRFIIVGYVD
Q7.AMI   -21  MALTMLLLLV AAALTLIETR AGQHSLQYFH TAVSRPGLGE PWFISVGYVD
Q9.AMI   -21  MALTMLLLLV AAALTLIETR AGQHSLQYFH TAVSRPGLGE PWFISVGYVD

                      39         49         59         69         79
Q6.AMI    30  NKQFVRFDSD AENPRMEPRA RWMEQEGPEY WERETQKAKG HEESFRVSLR
Q8.AMI    30  DTQFVRFDSD AENPRMEPRA RWMEQEGPEY WERETQKAKG HEESFRVSLR
Q7.AMI    30  DTQFVRFDSD AENPRMEPRA RWMEQEGPEY WERETQIAKG HEQSFRGSLR
Q9.AMI    30  DTQFVRFDSD AENPRMEPRA RWMEQEGPEY WERETQIAKG HEQSFRGSLR

                      89 ↓(α2)   99        109        119        129
Q6.AMI    80  TAQRYYNQSK GGSHTLQWMY GCDVGSDERL LRGYLQFAYE GRDYIALNED
Q8.AMI    80  TAQRYYNQSK GGSHTLQWMY GCDVGSDERL LRGYLQFAYE GRDYIALNED
Q7.AMI    80  TAQSYYNQSK GGSHTLQWMY GCDMGSDGRL LRGYLQFAYE GRDYIALNED
Q9.AMI    80  TAQSYYNQSK GGSHTLQWMY GCDMGSDGRL LRGYLQFAYE GRDYIALNED

                     139        149        159        169        179
Q6.AMI   130  LKTWTAADMA AQITLHKWEQ AGIAERDRAY LEGACVQSLR RYLQLRKETL
Q8.AMI   130  LKTWTAADMA AQITLHKWEQ AGIAERDRAY LEGACVQSLR RYLQLRKETL
Q7.AMI   130  LKTWTAVDMA AQITRRKWEQ AGIAEKDQAY LEGTCMQSLR RYLQLGKETL
Q9.AMI   130  LKTWTAVDMA AQITRRKWEQ AGIAEKDQAY LEGTCMQSLR RYLELGKETL

                 ↓(α3)189        199        209        219        229
Q6.AMI   180  LCTDPPKAHV THHPRSYGAV TLRCWALGFY PADITLTWQL NGEELTQDME
Q8.AMI   180  LCTDPPKAHV THHPRSYGAV TLRCWALGFY PADITLTWQL NGEELTQDME
Q7.AMI   180  LRTDPPKAHV THHPRSYGAV TLRCWALGFY PADITLTWQL NGEELTQDME
Q9.AMI   180  LRTDPPKAHV THHPRSYGAV TLRCWALGFY PADITLTWQL NGEELTQDME

                                                                ↓(TM)
                     239        249        259        269        279
Q6.AMI   251  LVETRPAGDG TFQKWASVVV PLGKEQNYTC HVNHEGLPEP LTLRW---EP
Q8.AMI   251  LVETRPAGDG TFQKWASVVV PLGKEQNYTC HVNHEGLPEP LTLRW---EP
Q7.AMI   251  LVETRPAGDG TFQKWASVVV PLGKEQNYTC HVNHEGLPEP LTLRWGRWEP
Q9.AMI   251  LVETRPAGDG TFQKWASVVV PLGKEQNYTC HVNHEGLPEP LTLRWGRWEP

                     289        299        309
Q6.AMI   280  PPSTVSNMAN VAIIVVLVAW PSLELWWIL*
Q8.AMI   280  PPSTVSNMAN VAIIVVLVAW PSLELWWIL*
Q7.AMI   280  PPYTVSNMAT IAVVVDLGAV AIIGAVVAFV MNRR*
Q9.AMI   280  PPYTVSNMAT IAVVVDLGAV AIIGAVVAFV MNRR*
```

Our *Q8* gene sequence (Fig. 3) was compared with the published *Q8* gene sequence (Devlin et al. 1985). We found 15 differences in the published sequence (Table 3). Only two of the 15 differences were in exons and only one of these resulted in an amino acid change and this was in the leader peptide (exon 1).

The deduced amino acid sequences of the protein products from the *Q6, Q7, Q8*, and *Q9* genes are shown in Figure 4. The two nucleotide differences in exon 2 of *Q6* compared with *Q8* result in the change of two amino acids at positions 30 and 31 in the α1 domain of the Qa-2 molecule. *Q6* has an asparagine and a lysine at these two positions, whereas *Q8* has an aspartic acid and a threonine at these two positions. The functional significance of these changes is unknwon. These two polymorphic amino acids are not in the peptide-binding pocket formed by the α1 and α2 domains of Qa-2 (Shawar et al. 1994). The only differences between the Q7 and Q9 proteins is at amino acid position 173 in the α2 domain. Position 173, at which Q7 has a glutamine and Q9 has a glutamic acid, is also not part of the peptide binding pocket.

The amino acid sequences shown in Figure 4 display three other interesting facets of Qa-2 protein structure. First, both the Q6/Q8 protein and the Q7/Q9 protein have stop signals in exon 5, which lead to truncated proteins missing the three cytoplasmic domains. The Q6/Q8 protein is five amino acids shorter than the Q7/Q9 protein. Second, the amino acid that determines whether a class I MHC protein will be linked to the cell surface by a glycosylphosphatidylinositol (GPI) linkage, an aspartic acid at amino acid position 295 (Waneck et al. 1988 a, b), is present in the *Q7/Q9* protein, but this amino acid is a valine in the *Q6/Q8* protein. The implication is that only the Qa-2 proteins from the *Q7/Q9* gene pair could be linked to the cell surface by a GPI linkage. This is important because only GPI-linked Qa-2 antigen can participate in cell activation after cross-linking with antibody in the presence of PMA (phorbol myristate acetate); (Robinson et al. 1989). Third, due to a deletion in exon 5 in the *Q6/Q8* gene pair (Fig. 3), there are major differences in the amino acid sequences of the transmembrane portion of the Qa-2 antigen encoded by the *Q6/Q8* gene pair compared with the *Q7/Q9* gene pair.

The *Q6, Q7, Q8*, and *Q9* genes are part of the class I *MHC* gene family and therefore are highly homologous to one another. A summary of the similarities among the *Q6, Q7, Q8*, and *Q9* genes and their predicted proteins is given in Table 4.

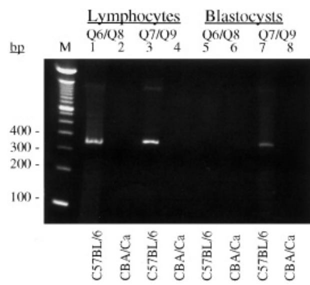


Fig. 5 RT-PCR assays for *Q6/Q8* and *Q7/Q9* gene expression in splenic lymphocytes and blastocyst stage embryos from C57BL/6 and CBA/Ca mice. CBA/Ca mice were used as negative controls because they have a deletion for the *Q6, Q7, Q8*, and *Q9* genes. Lanes 1–4 show RT-PCR performed on RNA purified from splenic lymphocytes, and lanes 5–8 show RT-PCR performed on blastocyst lysates. *Q6/Q8* primers were used in lanes 1, 2, 5 and 6, while *Q7/Q9* primers were used in lanes 3, 4, 7 and 8 (see Table 1)

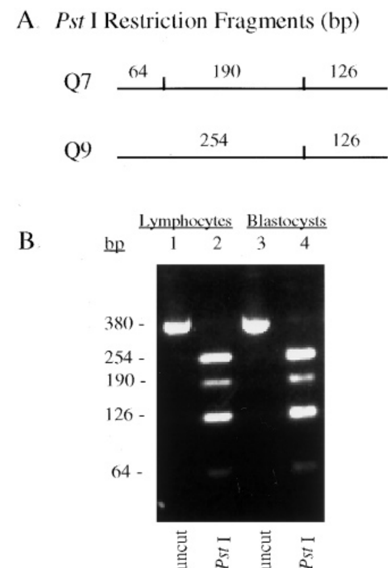


Fig. 6A, B Co-expression of the *Q7* and *Q9* genes in splenic lymphocytes and blastocyst stage embryos from C57BL/6 mice. **A** A diagram of the *Pst* I restriction fragments for both Q7 and Q9 RT-PCR products is shown. **B** Lanes 1 and 2 show RT-PCR for Q7/Q9 performed on splenic lymphocytes, and lanes 3 and 4 show RT-PCR for Q7/Q9 performed on blastocysts. Lanes 1 and 3 present RT-PCR products without *Pst* I digestion, and lanes 2 and 4 present RT-PCR products after *Pst* I digestion

Finally, the differential expression of the *Q6/Q8* and *Q7/Q9* gene pairs in mouse splenic lymphocytes compared with mouse blastocysts needs to be discussed. Regulation of class I MHC gene expression is controlled by *cis* and *trans*-acting regulatory factors (reviewed in Singer and Maguire 1990; Tatake and Zeff 1993). Three *cis*-acting regulatory elements in the 5′ regulatory region have been defined for class I *MHC* genes, class I regulatory element (CRE) or enhancer A, interferon responsive sequence (IRS), and Enhancer B. In Figure 3, the boundaries of the CRE/Enhancer A element are based on experimental evidence derived from the K^b (Kimura et al. 1986) and the L^d genes (Burke et al. 1989). The boundaries of the IRS element are based on experimental evidence obtained from the K^b gene (Kimura et al. 1986), the D^d gene (Korber et al. 1987), and the L^d gene (Shirayoshi et al. 1988). The boundaries of the Enhancer B element are based on experimental evidence derived from the K^b gene (Kimura et al. 1986).

In order to try to explain the lack of Q6/Q8 expression in mouse blastocysts, we compared the complete sequence of the 5′ regulatory region of the four genes (Fig. 3, Table 4). Overall, 27 differences appear between the *Q6/Q8* and the *Q7/Q9* gene pairs. Only four of the 27 differences between the *Q6/Q8* and the *Q7/Q9* gene pairs are in the three regulatory regions, CRE/Enhancer A, IRS, and Enhancer B. Within CRE/Enhancer A there are two nucleotide differences between *Q6/Q8* and *Q7/Q9*. The first is within region II of the CRE (position –186), which has been shown to play a significant role in the modulation of class I expression upon binding of a member of the nuclear hormone receptor family of transcription factors, H-2RIIBP. The second is a single bp deletion in region I of the CRE (position –160), which influences class I expression. In this region are binding sites for several nuclear transcription factors including NF-κB and H2TF1 (Tatake and Zeff 1993). Within the IRS region (and overlapping the CRE region) the single bp deletion at position –160 falls outside the core consensus sequence (CAGTTTCACT) required for IFN-mediated induction of class I genes (Tatake and Zeff 1993). Within Enhancer B there are two nucleotide differences between *Q6/Q8* and *Q7/Q9* (positions –96 and –61). The changes in the CRE/Enhancer A, IRS, and Enhancer B regulatory regions of the *Q6/Q8* vs the *Q7/Q9* gene pair probably account for some differences in expression of these genes. However, the details of the mechanisms of this differential gene expression will need to be the subject of future research.

The Ped gene product is the Qa-2 antigen. It has been shown that all Qa-2 antigen on the embryonic cell surface is subject to cleavage by phospholipase C, implying that all embryonic Qa-2 antigen is linked to the embryonic cell surface by a GPI linkage (Tian et al. 1992). The data reported in this paper, that only the *Q7* and *Q9* genes are transcribed in mouse blastocysts, are consistent with this finding. It has been shown that the *Q9* gene is translated into a functional gene product by converting a *Ped slow* mouse strain to a *Ped fast* mouse strain by the introduction of the *Q9* transgene (Xu et al. 1994). It remains to be determined whether the *Q7* gene is also translated and functional in preimplantation embryos.

The DNA sequencing reported in this paper was made possible by the Human Genome Project. B. L. Karger is currently being supported in order to develop high throughput DNA sequencing using capillary electrophoresis. As a pilot demonstration project, the laboratory of C. M. Warner is preparing plasmid clones of the mouse *MHC* for sequencing by the new technology. The present project represents an outcome of the preparation of samples and demonstrates the useful information possible from DNA sequence data. Work is continuing toward the pilot sequencing demonstra-

Table 3 *Q8* gene differences between the present and the GenBank data

Location	Source	Nucleotide changes (italics)						Amino acid changes
Exon								
			44					
1	Present data	5′	A*C*G	3′				Threonine
	GenBank		A*T*G					Methionine
			290					
2	Present data	5′	C*T*G	3′				No change
	GenBank		C*C*G					
Intron								
			130,31	140	175	212		
1	Present data	5′	G*CC*;	C*G*C;	C*C*A;	C*CG*	3′	
	GenBank		G– –;	C–C;	C*G*A;	C–G		
			216	224				
	Present data	5′	C*C*G	C*C*G	3′			
	GenBank		C–G	C*G*G				
			1041					
3	Present data	5′	C*T*G	3′				
	GenBank		C–G					
			3503					
4	Present data	5′	C*T*T	3′				
	GenBank		C*C*T					
			3636	3641				
5	Present data	5′	G*G*G;	G*A*G;	3′			
	GenBank		G*A*G;	G*G*G				
			3904					
6	Present data	5′	G*G*G	3′				
	GenBank	5′	G*T*G	3′				
3′ untranslated region								
			5086					
	Present data	5′	G*A*A	3′				
	Elliott et al. 1989		G–A					

Table 4 Summary of similarities among the *Q6, Q7, Q8*, and *Q9* genes and their predicted proteins

Region	No. nucleotides or a.a. s	No. diffs. *Q6* vs *Q8*	%Similarity	No. diffs. *Q7* vs *Q9*	%Similarity	No. diffs. *Q6/Q8* vs *Q7/Q9*	%Similarity
5′ Regulatory							
CRE/Enh. A	42	0	100	0	100	2	95
IRS	29	0	100	0	100	1*	97
Enhancer B	61	0	100	0	100	2	97
Total	**743**	**0**	**100**	**6**	**99.2**	**27**	**96**
Exons							
Exon 1	64	0	100	0	100	1	98
Exon 2	270	2	99.3	0	100	19	93
Exon 3	276	0	100	1	99.6	15	95
Exon 4	285	0	100	0	100	11	96
Exon 5	117	0	100	0	100	14	88
Total	**1012**	**2**	**99.8**	**1**	**99.9**	**60**	**94**
Protein							
LP	21	0	100	0	100	1	95
α1	90	2	97.8	0	100	12	87
α2	92	0	100	1	98.9	12	87
α3	95	0	100	0	100	3	97
TM	36†	0	100	0	100	22	39
Total	**334**	**2**	**99.4**	**1**	**99.7**	**50**	**85**

* Change is the same as that found in the overlapping region of CRE/Enhancer A
† Up to stop codon of Q7/Q9

tion. The long-range goal is to use the new high throughput DNA sequencing technology to sequence the entire *Q* region of the mouse MHC.

Acknowledgments This work was supported by grants from the NIH (HD31505 and GM46467) and the DOE (DE-FG02-90ER 60985). Support by the DOE does not constitute an endorsement of the views expressed in this article. Wei Cao was partially supported by a Matthews Research Co-op position. The expert technical assistance of Miriam Paschetto is gratefully acknowledged. This paper is contribution number 672 from the Barnett Institute.

References

Beckman, E. M. and Brenner, M. B. MHC class I-like, class II-like and CD1 molecules: distinct roles in immunity. *Immunol Today 16:* 349–352, 1995

Burke, P. A., Hirschfeld, S., Shirayoshi, Y., Kasik, J. W., Hamada, K., Appella, E., and Ozato, K. Developmental and tissue-specific expression of nuclear proteins that bind the regulatory element of the major histocompatibility complex class I gene. *J Exp Med 169:* 1309–1321, 1989

Devlin, J. J., Weiss, E. H., Paulson, M., and Flavell, R. A. Duplicated gene pairs and alleles of class I genes in the Qa-2 region of the murine major histocompatibility complex: a comparison. *EMBO J 4:* 3203–3207, 1985

Elliot, E., Rathbun, D., Ramsingh, A., Garberi, J., and Flaherty, L. Genetics and expression of the Q6 and Q8 genes. An LTR-like sequence in the 3′ untranslated region. *Immunogenetics 29:* 371–379, 1989

Flaherty, L. The Tla region of the mouse: identification of a new serologically defined locus, Qa-2. *Immunogenetics 3:* 533–539, 1976

Flaherty, L., Elliot, E., Tine, J. A., and Walsh, A. C. Immunogenetics of the Q and TL regions of the mouse. *Crit Rev Immunol 10:* 131–175, 1990

Jin, P., Meyer, T. E., and Warner, C. M. Control of embryo growth by the *Ped* gene: use of reverse transcriptase polymerase chain reaction (RT-PCR) to measure mRNA in preimplantation embryos. *Assisted Reprod Tech Androl 3:* 377–383, 1992

Kimura, A., Israël, A., Le Bail, O., and Kourilsky, P. Detailed analysis of the mouse H-2 K^b promoter: enhancer-like sequences and their role in the regulation of class I gene expression. *Cell 44:* 261–272, 1986

Korber, B., Hood, L., and Stroynowski, I. Regulation of murine class I genes by interferons is controlled by regions located both 5′ and 3′ to the transcription initiation site. *Proc Natl Acad Sci USA 84:* 3380–3384, 1987

Pal, S. K., Crowell, R., Kiessling, A. A., and Cooper, G. M. Expression of proto-oncogenes in mouse eggs and preimplantation embryos. *Mol Reprod Dev 35:* 8–15, 1993

Parish, C. R., Kirov, S. M., Bowern, N., Blanden, N., and Blanden, R. V. A one-step procedure for separating mouse T and B lymphocytes. *Eur J Immunol 4:* 808–815, 1974

Robinson, P. J., Millrain, M., Antoniou, J., Simpson, E., and Mellor, A. L. A glycophospholipid anchor is required for Qa-2 mediated T cell activation. *Nature 342:* 85–87, 1989

Shawar, S. M., Vyas, J. M., Rodgers, J. R., and Rich, R. R. Antigen presentation by major histocompatibility complex class I-B molecules. *Annu Rev Immunol 12:* 839–880, 1994

Shinomiya, T., Kobayashi, M., and Sato, S. A second site specific endonuclease from *Thermus thermophilus 111, Tth111* II. *Nucleic Acids Res 8:* 3275–3285, 1980

Shirayoshi, Y., Burke, P. A., Appella, E., and Ozato, K. Interferon-induced transcription of a major histocompatibility class I gene accompanies binding of inducible nuclear factors to the interferon consensus sequence. *Proc Natl Acad Sci USA 85:* 5884–5888, 1988

Singer, D. S. and Maguire, J. E. Regulation of the expression of class I MHC genes. *Crit Rev Immunol 10:* 235–257, 1990

Stroynowski, I. Molecules related to class I major histocompatibility complex antigens. *Annu Rev Immunol 8:* 501–530, 1990

Tatake, R. J. and Zeff, R. A. Regulated expression of the major histocompatibility complex class I genes. *Proc Soc Exp Biol Med 203:* 405–417, 1993

Tian, Z., Xu, Y., and Warner, C. M. Removal of Qa-2 antigen alters the *Ped* gene phenotype of preimplantation mouse embryos. *Biol Reprod 47:* 271–276, 1992

Tine, J. A., Walsh, A., Rathbun, D., Leonard, L., Wakeland, E. K., Dilwith, R., and Flaherty, L. Genetic polymorphisms of *Q* region genes from wild-derived mice: implications for *Q* region evolution. *Immunogenetics 31:* 315–325, 1990

Warner, C. M., Brownell, M. S., and Rothschild, M. F. Analysis of litter size and weight in mice differing in *Ped* gene phenotype and the Q-region of the H-2 complex. *J Reprod Immunol 19:* 303–313, 1991

Warner, C. M., Gollnick, S. O., and Goldbard, S. B. Linkage of the preimplantation embryo development (*Ped*) gene to the major histocompatibility complex (MHC). *Biol Reprod 36:* 606–610, 1987 a

Warner, C. M., Gollnick, S. O., Flaherty, L., and Goldbard, S. B. Analysis of Qa-2 antigen expression by preimplantation mouse embryos: possible relationship to the preimplantation embryo development (*Ped*) gene product. *Biol Reprod 36:* 611–616, 1987 b

Warner, C. M., Panda, P., Almquist, C. D., and Xu, Y. Preferential survival of mice expressing the Qa-2 antigen. *J Reprod Fertil 99:* 145–147, 1993

Waneck, G. L., Sherman, D. H., Kincade, P. W., and Low, M. G. Molecular mapping of signals on the Qa-2 antigen required for attachment of the phosphatidylinositol membrane anchor. *Proc Natl Acad Sci USA 85:* 577–581, 1988 a

Waneck, G. L., Stein, M. E., and Flavell, R. A. Conversion of a PI-anchored protein to an integral membrane protein by a single amino acid mutation. *Science 241:* 697–699, 1988 b

Weiss, E. H., Golden, L., Fahrner, K., Mellor, A. L., Devlin, J. J., Bullman, H., Tiddens, H., Bud, H., and Flavell, R. A. Organization and evolution of the class I gene family in the major histocompatibility complex of the C57BL/10 mouse. *Nature 310: 650–655, 1984*

Xu, Y., Jin, P., and Warner, C. M. Modulation of preimplantation embryonic development by antisense oligonucleotides to major histocompatibility complex genes. *Biol Reprod 48:* 1042–1046, 1993

Xu, Y., Jin, P., Mellor, A. L., and Warner, C. M. Identification of the *Ped* gene at the molecular level: the Q9 MHC class I transgene converts the *Ped slow* into the *Ped fast* phenotype. *Biol Reprod 51:* 695–699, 1994