Immunogenetics (2002) 54:479–489
DOI 10.1007/s00251-002-0494-0

ORIGINAL PAPER

**Attila Kumánovics · Anup Madan · Shizhen Qin
Lee Rowen · Leroy Hood · Kirsten Fischer Lindahl**

Quod erat faciendum: sequence analysis of the *H2-D* and *H2-Q* regions of 129/SvJ mice

Received: 3 June 2002 / Accepted: 24 July 2002 / Published online: 20 September 2002

Abstract The *H2-D* and *-Q* regions of the mouse major histocompatibility complex (*Mhc* or *H2*) have been sequenced from strain 129/SvJ (haplotype *bc*), revealing a *D/Q* region different from all other investigated haplotypes, including the closely related *b* haplotype. The 300-kb class I-rich region consists of the classical class I, *H2-D,* and 11 non-classical class I genes. The *Q* region was formed by two series of tandem duplications. Comparison of the segment between the *D* and *Q1* genes with the *H2-K* region provides evidence that class I genes were translocated from the *K* region to the *D* region, and gives a new explanation for the weak locus specificity of the *H-K* and *H2-D* alleles.

Keywords Major histocompatibility complex · Class I genes · *H2-K* region · Evolution · Gene duplication

Introduction

The *H2* complex on Chromosome 17 covers about 4 Mb and is traditionally divided into three regions according to their gene content, designated as class I, II, and III. In mouse, the *Mhc* class I genes are further divided into five subregions, *K, D, Q*, *T*, and *M*, by their location on Chr 17, as defined by recombination. The *H2-K* is centromeric; the *H2-D*, *-Q*, *-T*, and *-M* regions are telomeric to the class II and class III regions. The *H2-K* and *D* loci encode the class Ia (or classical class I) molecules, whereas the *H2-Q*, *-T*, and *-M* regions encode the less-understood class Ib (or non-classical class I) molecules. The class Ia molecules (K, D, and L in some haplotypes) are found on virtually all cells and are highly polymorphic. Their function is to present short peptides (8–10 amino acids) derived mainly from the inside of the cell. The class Ib genes are in general characterized by more-limited expression and by a low level of polymorphism. The function of most of the class Ib gene products is still an open question.

None of the class I genes is orthologous between man and mouse (Amadou et al. 1999). *K*, *T*, and *M* region orthologues can be found in rat, but the *H2-D/Q* region is mouse specific (Günther and Walter 2001; Hughes 1991; Lambracht-Washington and Fischer Lindahl 2002). Six out of the ten class Ib genes (*Q1-Q10*) described in the *H2-Q* region of strain C57BL/10 (haplotype *b*) have been shown to encode protein products (Flaherty et al. 1990); none of these class Ib proteins has a well defined function, but there are many intriguing properties.

The class Ia proteins are type I transmembrane proteins, whereas Q4 and Q10 are secreted proteins (Kress et al. 1983; Robinson et al. 1988); nevertheless Q10 binds peptides 8–9 amino acids long, much like the class Ia proteins (Zappacosta et al. 2000). *Q6*, *Q7*, *Q8*, and *Q9* encode the widely expressed Qa2 antigen. A membrane-bound and two soluble forms of the Qa2 antigen have been identified (Soloski et al. 1986; Stroynowski et al. 1987; Tabaczewski et al. 1994; Waneck et al. 1987). Qa2 is anchored to the membrane via glycophosphatidylinositol (GPI) linkage (Ulker et al. 1990a; Waneck et al. 1988). Soluble forms are generated by a posttranslational cleavage of the GPI-linked isoform, or by splicing out the fifth exon (Tabaczewski et al. 1994). Soluble forms are upregulated upon stimulation (Soloski et al. 1986;

A. Kumánovics and A. Madan have contributed equally to this work.

A. Kumánovics (✉) · K. Fischer Lindahl
Center for Immunology,
University of Texas Southwestern Medical Center, Dallas,
TX 75390–9050, USA
e-mail: Attila.Kumanovics@UTSouthwestern.edu
Tel.: +1-214-6487554, Fax: +1-214-6485453

A. Madan · S. Qin · L. Rowen · L. Hood
The Institute for Systems Biology, Seattle, WA 98105–6099, USA

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

A. Kumánovics
University of Texas Southwestern Medical Center,
5323 Harry Hines Boulevard, Dallas, TX 75390–9050, USA

Tabaczewski et al. 1994; Ulker et al. 1990b). Qa2 binds peptide antigen and can serve as a restriction element for allogeneic T cells (Joyce et al. 1994; Rötzschke et al. 1993; Tabaczewski et al. 1997). The three-dimensional structure of Q9 determined by X-ray crystallography shows a peptide-binding groove significantly shallower and less polar than other investigated class I molecules (He et al. 2001). Qa2-deficient mice are susceptible to *Taenia crassiceps* infection, and the Qa2 expression level correlates with the resistance to the infection (Fragoso et al. 1998) and with the level of CD8α/α TCRα/β intraepithelial lymphocytes (Das et al. 2000). Q9 transfection can protect natural killer (NK)-sensitive tumor cells from lysis mediated by lymphokine-activated killer (LAK) and NK cells (Chiang et al. 2002). Qa2 is also implicated in early embryonic development. *Q7* and *Q9* were identified as the loci encoding the *Ped* gene (preimplantation embryo development), whose presence accelerates the preimplantational cleavage rate of the embryo (Goldbard et al. 1982; Wu et al. 1999; Xu et al. 1994). *Q6* and *Q8* (and sometimes *Q4*) are referred to as *Q* "even" genes, whereas *Q7* and *Q9* (and *Q5*) are referred to as the *Q* "odd" genes (O'Neill et al. 1986).

The class Ia *H2-K* and *-D* loci are highly polymorphic. In contrast, *Q4* and *Q10*, and the Qa2 encoding genes are non-polymorphic. What is variable in the *Q* region is the number of genes (Flaherty et al. 1990; Mellor et al. 1984; Tine et al. 1990).

Whereas the class Ia and Qa2 proteins are ubiquitously expressed (Ungchusri et al. 2001), *Q1* transcripts are restricted to the thymus and the intestinal epithelium, *Q4* and *Q10* are expressed in the liver, and *Q5* in AKR mice is expressed only in the thymus and in the uterus of pregnant adult mice (Robinson et al. 1988; Schwemmle et al. 1991).

The *H2-D* and *-Q* regions have been mapped previously with cosmids from different *H2* haplotypes (*b,d, k, p,* and *q*), and all of them were found to be different (Litaker et al. 1996; Stephan et al. 1986; Watts et al. 1989; Weiss et al. 1984, 1989). We have chosen strain 129/SvJ, for which high-quality bacterial artificial chromosome (BAC) libraries were available. 129/SvJ is haplotype $H2^{bc}$ (Fischer Lindahl 1997), which is different from the other investigated strains. The extensively studied $D^{-/-}K^{-/-}$ knock-out mice are also $H2^{bc}$ (Pascolo et al. 1997; Vugmeister et al. 1998). Here we present the complete sequence covering the *H2-*D^{bc} and *H2-*Q^{bc} regions encoding nine class I genes, one apparent class I pseudogene, four archain pseudogenes, one fertilization antigen-1 pseudogene, and two class I gene fragments.

Materials and methods

Clones and sequencing

The presented BAC assembly, derived from 129/SvJ embryonic stem cells, was purchased from Incyte Genomics (Palo Alto, Calif., USA). The BAC library arrayed on ten high-density filters was screened by Southern hybridization. The probes for library screening were pooled polymerase chain reaction (PCR) products labeled with non-radioactive DIG-11-dUTP (Roche, Indianapolis, Ind., USA). PCR primers were selected from known mouse *Mhc* gene sequences for initial library screens and from BAC end sequences for gap filling. At regions where mouse *Mhc* sequences were not available, primers were selected from the human counterpart. The BAC clones identified by library screening were verified by single-strand conformational polymorphism (SSCP) analysis, fingerprinting of contigs, and fluorescent in situ hybridization (FISH) of selected BACs from each contig to metaphase chromosomes. The BAC ends were sequenced and the end sequences used to generate probes for BAC end walking to fill the gaps.

The clones were sequenced by a high-redundancy shotgun method (Rowen et al. 1999). Briefly, BAC DNA was prepared in AutoGen740 and sheared using sonication. After end repair and size selection, the insert DNA was subcloned into either M13mp9 or pUC18. Sequences were resolved on Applied Biosystems (Foster City, Calif., USA) 373, 377, or 3700 sequencers using a mixture of dye-primer and dye-terminator chemistries. After obtaining enough reads for about eightfold coverage, the sequence data were assembled with Phrap. In the case of BAC 214O16, the sequence was first assembled with Sequencher then with Phrap. Finishing was done by either resequencing with alternative chemistry, or directed sequencing with custom oligonucleotide primers, or by subcloning PCR products or restriction digest fragments. The sequence was determined to accuracy of about one error per 35,000 base pairs.

Sequence analysis

The sequences were analyzed with the Wisconsin Package v10 (Genetics Computer Group, Madison, Wis., USA) and the EMBOSS (http://www.emboss.org) program packages. Blast programs (Altschul et al. 1997) were either run locally or using http://www.ncbi.nlm.nih.gov. Multiple alignments were generated by ClustalW (Thompson et al. 1994). Repetitive sequences were identified by RepeatMasker (AFA Smit and P Green, unpublished data; http://ftp.genome.washington.edu/cgi-bin/RepeatMasker). Dot matrices were either created by Dottup (from the EMBOSS package) or PipMaker. Dottup makes a simple word match comparison, and it was used for displaying regions of substantial similarity. We used a word size of 15 (15 nucleotide-long perfect match) for comparing the sequence with itself. Word size of 10 was used for locating the class I and archain genes in the *D/Q* region. More distantly related sequences (i.e., sequences from different organisms) were compared by the Blast-based PipMaker (Schwartz et al. 2000), which was run at http://bio.cse.psu.edu/PipMaker. PipMaker can incorporate the results from RepeatMasker; we used this function to decrease background on some of the dot matrices and for color coding the interspersed-repeat regions. The gene structures were determined by using the Wise2 package (version 2.1.17b, http://www.sanger.ac.uk/Software/Wise2).

Results

General features of the *H2-D/Q*^{bc} region sequence

We have sequenced five BAC clones: 205A4 (GenBank: AC007080; 39,909 bp), 11M9 (GenBank: AC087216; 226,594 bp), 384N12 (GenBank: AC087217; 77,366 bp), 214O16 (GenBank: AF111102; 158,852 bp), and 322F16 (GenBank: AF111103; 159,179). The 661,900 bp of sequence result in a 589-kb contiguous non-overlapping sequence from *Nfkbil1* (or *IκBL*) in the class III region to *Pou5f1*, covering the entire *H2-D/Q* region. Here we analyze the segment between *Bat1* and *Pou5f1* covering the entire *H2-D* and *-Q* regions (332,201 bp). *Bat1* is the

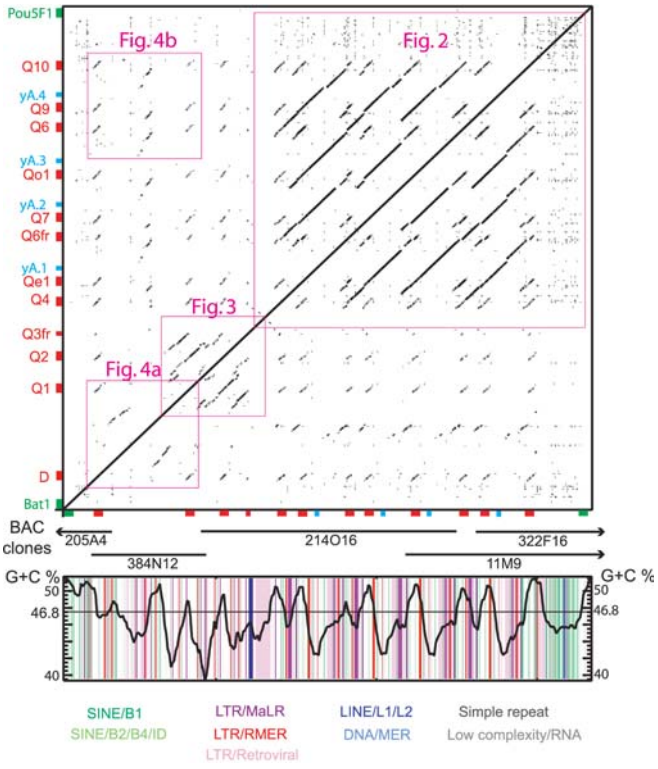


Fig. 1 Dot-matrix analysis of the *H2-D/Q* region. The sequence of 344 kb from *Bat1* to *Pou5f1* has been compared with itself without repeat-masking. The gene content is labeled by *colored boxes* on both axes: the class I genes are *red*, the archain-like pseudogenes are *blue*, and *Bat1* and *Pou5f1,* the genes bordering the region, are *green*. The *diagonal line* represents the sequence matching to itself. *Lines parallel to the diagonal* identify internally repeated regions. The shorter matches are class I genes. Longer matches are the repeat units including one or more class I genes. *Magenta rectangles* highlight regions shown in more detail in subsequent figures. The *middle part of the figure* shows the sequenced BAC clones. *Arrowheads* on clones 205A4, 11M9, and 322F16 indicate that these clones extend beyond the analyzed region. The *bottom part of the figure* shows the GC and genome-wide repeat content of the region. The percentage of guanine plus cytosine was determined in a sliding window of 10^4 bp with a shift of 2,000 bp (Isochore from Emboss). All the genes are associated with GC content peaks. In the case of the two class I fragments (*Q3fr* and *Q6fr*) the peaks are lower. The segment corresponding to the peak between *D* and *Q1* is discussed in Fig. 4. The genome-wide repeats are specified by coloring the background as indicated. The class I-containing region is dominated by various long terminal repeat-like (LTR) sequences, labeled by *purple*, *red*, and *pink*. The only LINE repeat (*dark blue*; followed by a full-length IAP retroviral sequence, *pink*) in the region replaced the second half of the *Q3* gene, making it a gene fragment. SINE repeats (various shades of *green*) are dominant outside of the class I-rich region. Simple repeats and low-complexity regions are *gray*

last gene of the *Mhc* class III region, and *Pou5f1* marks the beginning of the >600 kb segment between *H2-Q* and *H2-T* class I genes (Amadou et al. 1999).

The average G+C content of the sequence is 46.8%. Much of the *D/Q* region (37.8%) is interspersed repeats (Fig. 1). The whole genome average of the interspersed repeat content for regions with similar G+C level (Smit 1999) is14.4% for short interspersed nuclear elements (SINE) but 8.1% in our sequence; the genome average for long interspersed nuclear elements (LINE) is 3.4% and only 1.3% in our sequence, and the average for long terminal repeat elements (LTR) is 9.9% but 28.2% in our sequence. Low complexity and simple repeat/satellite sequences occupy 1.6% of the *D/Q* region. In conclusion, the LTR sequences are over- and the SINE sequences under-represented in the *D/Q* region. The *D/Q* region values are also in contrast to the neighboring *Mhc* class III region: 51% G+C content, 28% interspersed repeats, half of which are SINE.

The most remarkable feature of the sequence is revealed by a dot-matrix comparison of the sequence with itself: all the sequence is present two or more times within the region (Fig, 1), rendering most of the sequence formally a "region-specific repeat". For detailed analysis, we follow these clusters of repeated segments as labeled on Fig. 1.

The cluster of *Q4-Q10*

The largest of these clusters represents two-thirds of the *H2-D/Q* region (Fig. 1). The comparison of the region with itself shows four duplications (Fig. 2). These duplications generated three copies of a ~15-kb segment (yellow box) and four copies of a ~28-kb (green box) segment. The ~15-kb and the ~28-kb segments are also a result of an ancient duplication, as they retain sequence similarity to each other outside of the class I genes (Motif1 – filled black circle, Fig. 2).

The most parsimonious development of the region is depicted on Fig. 2. After the initial duplication of the primordial Qa2-containing segment into two, the present 28-kb unit gained not only additional interspersed repeats but also a retrointegrated pseudogene. Then, the whole segment duplicated resulting in two ~15-kb and two ~28-kb units. The most recent duplication included two ~28-kb units and one ~15 kb-unit.

We have aligned a class I coding sequence (*H2-K^{bc}* derived from GenBank: AF100956) to the entire region to locate all the class I molecules. The alignment revealed eight class I genes, all in the same centromere-to-telomere orientation (Fig. 2). BLAST searches of the *D/Q* region sequence against the non-redundant protein and expressed sequence tag (EST) databases showed significant similarity to the human *archain* (*ARCN1*) gene. *ARCN1* gene encodes the δCOP coatamer protein, which is highly conserved among eukaryotes, hence the name (Radice et al. 1995; Tunnacliffe et al. 1996). The alignment of the human *ARCN1* mRNA to the *Q4-Q10* segment showed four copies of archain-like sequences (Fig. 2, blue boxes). These *ARCN1*-like sequences in the *Q* region are pseudogenes, as they all harbor three characteristic insertions disrupting the coding sequence (between amino acid residues 322–323, 336–338, and 410–412 of NP_001646). The retrointegration occurred prior to the duplication of the larger Qa2 segment (Fig. 2, green box), as all four pseudogenes share the

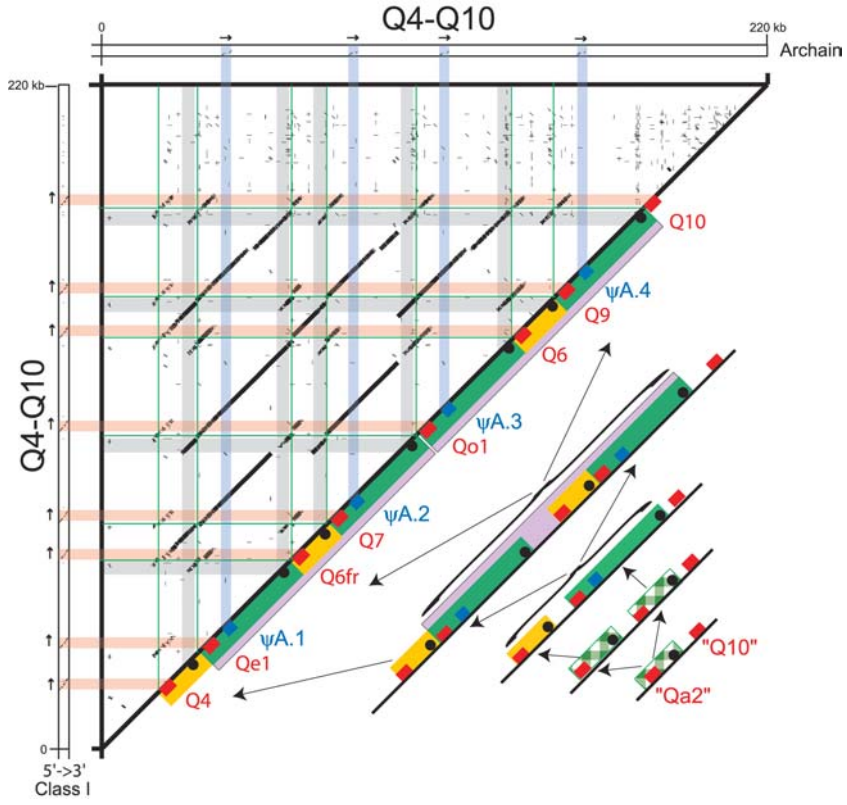


Fig. 2 Dot-matrix analysis and predicted origin of the cluster between *Q4* and *Q10*. Word length of 15 was used for comparison by Dottup (without repeat-masking). Because the sequence is compared with itself the half-plot contains the whole information. To locate the class I and archain-like sequences, the *Q4-Q10* segment was compared with the *H2-K*^{bc} genomic sequence and with the human *ARCH1* cDNA sequence by Dottup (word size of 10). These plots are aligned to the two axes of the main plot. The class I matches are highlighted in *pink* and the archain-like sequences highlighted in *blue* across the plots. The class I genes are *red*, and the archain-like pseudogenes are shown as *blue boxes*. *Arrows* next to the *red* and *blue* boxes represent the transcript orientation. The *gray* highlights represent Motif1 (shown as a *black-filled circle*), a short sequence present in both basic duplication units predicted in this region. Motif1 is defined based on the dot-matrix analysis presented (i.e., from 33,310 bp to 34,230 bp of AC087216). The duplication units are shown as *yellow* and *green* boxes at the diagonal of the main plot and shown by *green lines* across the plot. The most parsimonious evolution of the region is illustrated below the diagonal of the plot: an ancient segment containing a Qa2-like class I gene duplicated and evolved into the two basic duplication units identifiable today (*yellow* and *green*) by acquiring genome-wide repeats and an archain pseudogene. Then these two units duplicated together, generating two copies from each. The *green units* and the *yellow unit* in between them (linked by a *purple box*) duplicated again, generating the current arrangement. *Q6fr* was fragmented by a 1.5-kb deletion within the gene (not visible at this magnification). *Q10,* which diverged early from the rest of the *Q* genes, is not part of this expansion (Mellor et al. 1984)

same three insertions. Because these archain-like sequences have no introns and the human *ARCN1* genomic sequence (Accession number NT_009334) has nine, these sequences are likely to be retrointegrated from a processed *Arcn1* message into the *Q* region.

The cluster of *Q1-Q3*

The cluster of *Q1-Q3* is also the result of duplications of a segment containing a class I gene. Two full-length class I genes and a class I gene fragment were revealed by aligning a class I sequence (*H2-K*^{bc}) to the region. The comparison of the region to itself shows two duplications and a deletion (Fig. 3). The putative ancestral segment duplicated, giving rise to *Q1* and *Q2*, then this whole *Q1-Q2* segment duplicated (alternatively, only the *Q1* segment duplicated again) (Fig. 3). The telomeric half of the *Q3* (and the entire fourth segment, in case the whole *Q1–Q2* segment duplicated) was replaced by LINE1 and intracisternal A-particle (IAP) retrotransposition. LINE1 elements are present in up to 10^5 copies in the mouse genome, and IAPs are found in about 10^3 copies, but usually in short fragments (Kuff and Lueders 1988; Smit 1999). Almost 30% of the *D/Q* region is LTR/retroviral-like sequence (Fig. 1), but the IAP at *Q3* is the only full-length retrotransposon.

After the initial duplication, a fertilization antigen-1 (Ssfa1)-like processed mRNA retrointegrated into the *Q2*-segment (Fig. 3). *Ssfa1* is located on chromosome 15 and it has two introns, whereas the retrointegrated copy has no introns [GenBank NW_000106; (Zhu and Naz 1997)]. The retrointegrated *Ssfa1-like* pseudogene accumulated four insertions, two deletions, and numerous substitutions (data not shown).

The segment between *D* and *Q1*

The third and smallest duplicated region is found between the *H2-D* and the *H2-Q* regions (Fig. 1 and

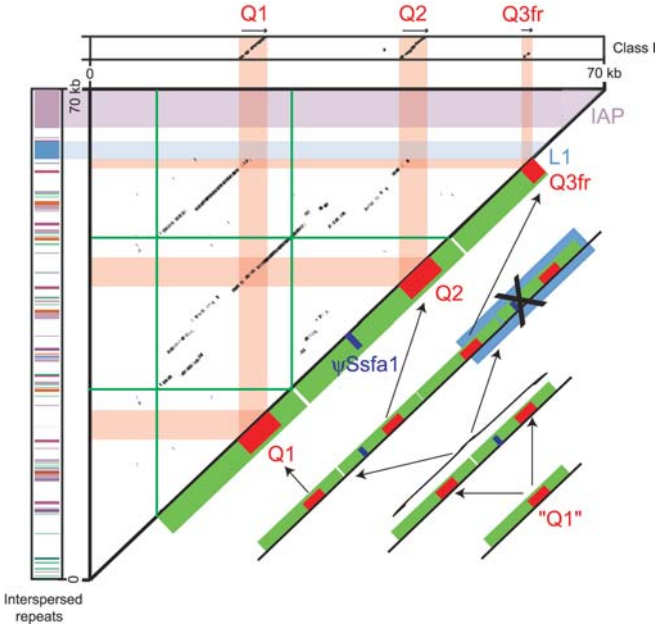


Fig. 3 Dot-matrix analysis and predicted origin of the *Q1,2,3* cluster. Because the sequence is compared with itself the half-plot contains the whole information. To locate the class I sequences, the segment was compared with the *H2-K*^{bc} genomic sequence by Dottup (word size of 10, no repeat-masking). The genome-wide repeats are shown to the *left* of the plot; the color code is the same as in Fig. 1. The class I matches are highlighted in *pink*, the L1 repeat sequence highlighted in *blue*, and the IAP retroviral sequence highlighted in *purple* across the plots. Insertion of mobile elements made *Q3* a gene fragment. The class I genes are shown as *red boxes*. *Arrows* next to the *red boxes* represent the transcript orientation. The *Ssfa1*-like pseudogene is the *blue box*. The duplication units are shown as *green boxes* at the diagonal of the main plot and by *green lines* across the plot. The matches within the repeat units are partly due to the genome-wide repeats, and partly due to further rearrangements within the units. The most parsimonious evolution of the region is illustrated below the diagonal of the plot: an ancient *Q1*-like class I gene-containing segment underwent a duplication and created *Q1* and *Q2*. In the next step either the whole *Q1-Q2* segment duplicated or the only the *Q1* segment, but the result of this second round of duplication was, for the most part, destroyed by a L1-IAP insertion leaving only a gene fragment (*Q3fr*) behind

Fig. 4a). Although the segment does not contain a class I gene or an archain-like sequence, it shows strong similarities to portions of the *Q4-Q10* segment (Fig. 1 and Fig. 4b). These segments of similarities were termed Motif1 and Qa2-repeat (Fig. 4b). Although there is no class I coding sequence in the *D-Q1* region (Fig. 1 and Fig. 4a), the dot-matrix analysis identified *Mhc* class I 5′ and 3′ flanking sequences in the duplicated regions of the *D-Q1* interval (Fig. 4a and b).

Surprisingly, Blast analysis of the *D-Q1* region (BlastX against GenPept) detected similarities to *Sacm2l*, a gene immediately centromeric to the *H2-K* class I genes (Walter and Günther 1998). The *Sacm2l* gene in the *H2-K* region contains 20 exons (8970 bp, AF100956). The *Sacm2l* gene fragments present in the *D-Q1* duplication segment contain the *Sacm2l* genomic (exons and introns) sequence from exon 11 through 20, with a deletion between exon 14 and 18. To further analyze this, we compared the *D-Q1* segment with the *H2-K* region (Fig. 4c), and also compared the *H2-K* region with the *Q9-Q10* segment (Fig. 4d). Because the *Q4-Q10* region is the result of serial duplications, it is enough to use only one, non-redundant, segment, e.g., between *Q9* and *Q10* (Fig. 1). These comparisons clearly showed the similarities between the *K* and *D-Q1* regions, and recognized segments of similarities between the *H2-K* region and the *Q4-Q10* regions (Motif1 and Qa2-repeat). Similarities between these regions were noted before in studies using restriction mapping and Southern hybridization (Weiss et al. 1984).

Only rat and mouse are known to have class I genes next to the class II region. The human extended class II region (Stephens et al. 1999) between *DAXX* and *COL11A2* was compared with the orthologous region of mouse (Fig. 5a). The conservation disappears between the *SACM2L* and *RING1* genes. In rat, only the region between *Sacm2l* and *RT1.A1* has been sequenced [AJ276628; (Walter and Günther 2000)]. We compared the rat, mouse, and human sequences using Blast (Bl2seq), and identified three short sequences conserved among these species (light blue 550–830 bp, dark blue 8361–8531 bp, purple 8942–9165 bp of AJ276628). These segments of unknown function, if any, are also present in the *D-Q1* region (Fig. 5b).

Mhc class I genes

The *H2-D/Q* region contains 12 class I genes, two of which are gene fragments (Fig. 1). *H2-D*^{bc} and *H2-Q1*^{bc} are full-length class I genes and their proteins are identical to those of haplotype *b* (Q1^{b} – AAB31648, D^{b} – P01899). *H2-Q2*^{bc} is probably a pseudogene based on an insertion in exon 4 (at amino acid residue 266), leading to early termination after two codons, although a truncated soluble α1-α2-α3 class I protein product is also possible. *H2-Q2*^{k} (NP_034522) can encode a traditional (not truncated) class I protein, suggesting that *H2-Q2*^{bc} is a pseudogene. Having only the first three exons of a class I gene, *H2-Q3* is a gene fragment in both haplotypes *b* and *bc*. The deduced H2-Q4^{bc} and H2-Q10^{bc} proteins are identical to Q4^{d} (A32273) and Q10^{k} (NP_034521).

Six genes encode the Qa2 antigen in 129/SvJ and a Q6-like fragment is also present (Fig. 2). Three of them are identical matches to Q6^{b}, Q7^{b}, and Q9^{b}. There are no *Q8*^{b} and *Q5*^{b} genes in 129/SvJ. Both *Q5*^{b} and *Q6fr* are gene fragments, but *Q5*^{b} is an odd-like gene (Robinson et al. 1988), whereas *Q6fr* is the remnant of an even gene. Two genes present in 129/SvJ have no equivalent in haplotype *b* (Figs. 6 and 7). Qo1^{bc} (Q odd-like 1) has one amino acid difference from the odd genes (Q7, Q9) in the α1 domain. Qe1^{bc} (Q even-like1) is even like in the leader, α1 and α2 domains, whereas the transmembrane domain is more odd-like. The α3 domain of Qe1 is 96.7% identical (all the others on

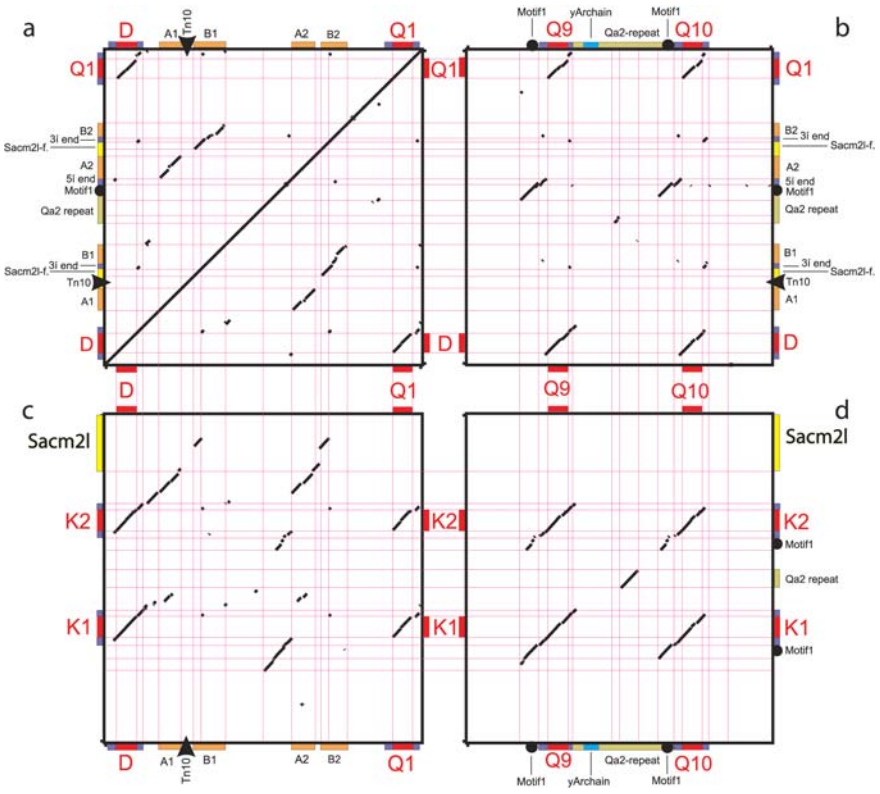


Fig. 4 **a** Dot-matrix analysis of the *D-Q1* interval. The plot was generated by PipMaker. The genome-wide repeats were masked. At the top of the plot we marked the class I genes with *red boxes*; the 5′ and 3′ class I sequences are colored *purple*. The duplicated sequences identified by this self-comparison are shown as *brown boxes*. The *D*-proximal duplication unit is interrupted by an artificial Tn*10* transposon integration (Hill et al. 2000); the corresponding pairs of the duplication units (*brown*) are labeled A1-A2 and B1-B2. At the left side of the plot the result of the four dot-matrix comparisons are shown together. *Magenta lines* are drawn across the plots to help the assessment. **b** Dot-matrix comparison of the *Q9-Q10* segment with *D-Q1* segment. The gene and repeat unit content of the region is shown at the *top of the plot*. These results are taken from Fig. 2; the class I genes are labeled as before, the archain-like sequence is *blue*, and the *green box* represents the repeat unit as on Fig. 2. At the *right side of the plot* the result of the four dot-matrix comparisons are shown together. **c** Dot-matrix comparison of the *D-Q1* interval and the *K* region. At the *left side of the plot* we marked the *K* class I genes with *red boxes*; the 5′ and 3′ class I sequences are colored *purple*. The duplicated sequences identified in **a** are shown as *brown-colored boxes on the bottom of the plot*. *Sacm2l* gene is shown as a *yellow box*. **d** Dot-matrix comparison of the *K* region and the *Q9-Q10* segment. The gene and repeat unit content of the *Q9-Q10* segment is shown at the *bottom of the plot*, as in **b**. The gene and repeat unit content of the *K* region is shown at the *right side of the plot*

Fig. 6 are 91.3% identical) to the intriguing Q8/9^{d} class I molecule from BALB/c (Matsuura et al. 1989). Both *Qe1* and *Qo1* potentially encode secreted proteins, as they have a stop codon in the transmembrane region at position 292. This not only deletes the membrane spanning portion of the protein but also removes Asp 292, where the GPI linkage is attached to the odd genes (Waneck et al. 1988).

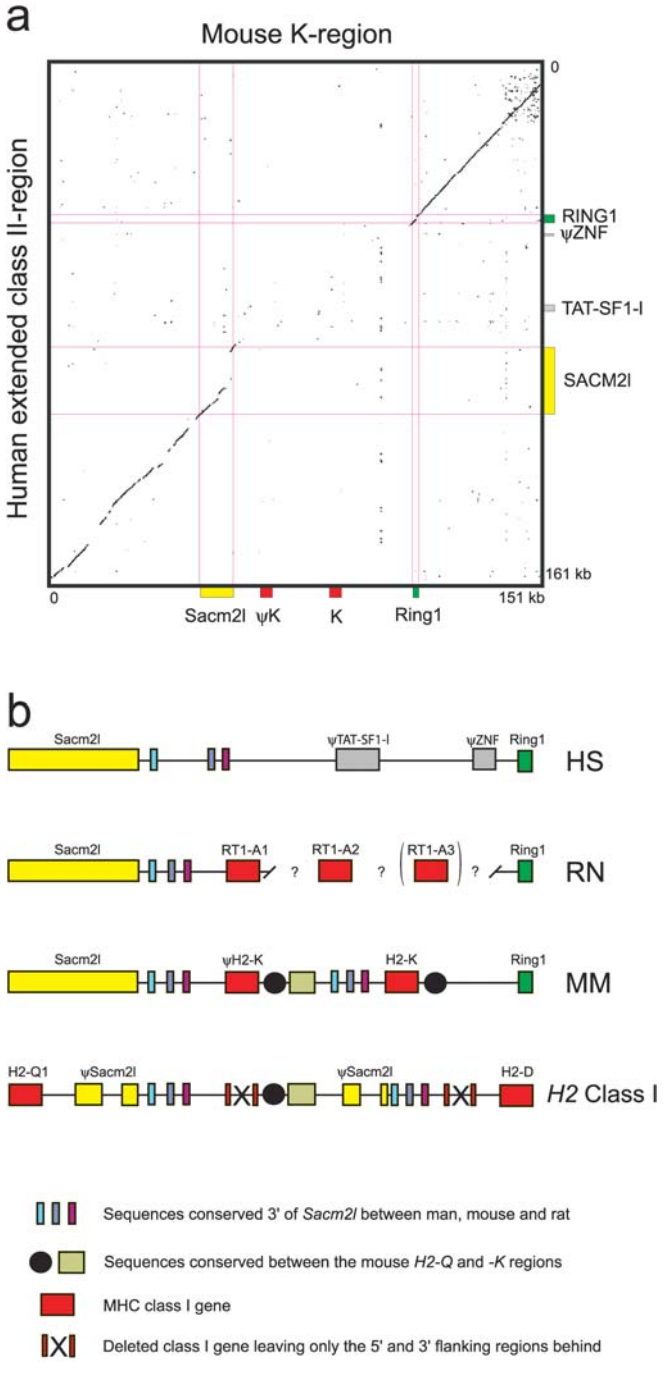
Discussion

The *Q* class I genes

The *H2-D* and *-Q* regions have been mapped previously with cosmids from five haplotypes (Fig. 7). C57BL/10 (*H2^{b}*) has 1 *D* and 10 *Q* class I genes (*Q1–10*, (Weiss et al. 1984). BALB/c (*H2^{d}*) contains 5 class I genes in the *D* region (*D, D2, D3, D4*, and *L*) and 9 class I genes in the *Q* region (*Q1–7*, *Q8/9*, and *Q10*) (Stephan et al. 1986). B10.AKM [*H2^{q},* (Weiss et al. 1989)] has a *D* region similar to BALB/c, but the *Q* region has not been investigated. One *D* gene and five *Q* region genes (*Q1*, *Q2*, *Q4*, *Q5/9*, and *Q10*) are found in C3H/HeN [*H2^{k},* (Watts et al. 1989)]. B10.P (*H2^{p}*) has a C57BL/10-like *D* region, and eight class I genes in the *Q* region [*Q1, Q2*, *Q4, Q5*, and the tentatively named *Q11–15*, (Litaker et al. 1996)]. By Southern blot analysis, *Q10* is the only *Q* region gene in B10.M [*H2^{f}*, (O'Neill et al. 1986)], and it was found not to be transcribed (Lew et al. 1986).

The picture that emerged from those studies was that there are a few prototypical arrangements in the *H2-D/Q* region, and these basic arrangements are modified by deletions, duplications, and unequal crossing-over. The two typical *D* region organizations are the "one gene" (*D* in haplotypes *b, k*), and the "five genes" (*D, D2, D3, D4*, *L* in haplotypes *d, q*). *H2-L* is derived from *H2-D* by a duplication taking part of the neighboring *Bat1* gene with it (Rubocki et al. 1990; Wroblewski et al. 1994). Because of restriction map homologies and cross-hybridization experiments, the *D2, D3,* and *D4* genes were

thought to originate from the *Q* region by unequal crossing-over (Rubocki et al. 1990; Stephan et al. 1986). In the *Q* region, haplotype *b* is thought to represent the basic organization with genes from *Q1* to *Q10* (Watts et al. 1989). This arrangement is modified in BALB/c by a putative deletion generating a fusion gene between *Q8* and *Q9* (Matsuura et al. 1989). Similarly, a deletion is thought to have created a fusion gene between *Q5* and *Q9* in C3H ($H2^k$), removing most of the Qa2-encoding genes (Watts et al. 1989). An even larger deletion removed most of the *Q* region in haplotype *f* leaving only *Q10* behind (O'Neill et al. 1986).



Sequence analysis of the *D* and *Q* regions from 129/SvJ now makes it possible to reevaluate those earlier schemes. The most revealing *Q* region genes in 129/SvJ are the two novel genes *Qe1* and *Qo1* (Fig. 6). $Q8/9^d$, a gene very similar to *Qe1* (97.9% amino acid identity), was identified earlier from BALB/c. $Q8/9^d$ was proposed to be the result of a fusion between an "even" ($Q8^d$) and an "odd" ($Q9^d$) gene after the deletion of the segment, including the 3′ half of the $Q8^d$ and the 5′ half of the $Q9^d$ gene (Matsuura et al. 1989). Because of additional amino acid differences in the α3 domain (Fig. 6), a simple fusion between $Q8^d$ and $Q9^d$ was questioned (Nakayama et al. 1991). 129/SvJ has *Qo1* in addition to *Qe1*, and these two genes share a unique transmembrane region, suggesting that these two genes derived from a duplication/diversification and not from a deletion/fusion. Probably the case of $Q5/9^k$ will be similar to that of $Q8/9^d$, simply representing a different class I gene rather than a fusion.

All the investigated haplotypes differ in the organization of the *Q* region (Fig. 7), yet the *Q* proteins themselves appear non-polymorphic among the wild-derived and inbred strains (Tine et al. 1990). Frequent expansion and contraction of the *Q* region could generate these features so markedly different from the *D* and *K* regions, as proposed by Flaherty (1988). The isotypic and allelic variations are lost during the successive gene duplications and deletions. Near-perfect repeats can make the region prone to deletions and unequal crossing-overs. Similarly, frequent sequence homogenization between the duplicated units was proposed to eliminate mutations in the human and mouse T-cell receptor β loci (Glusman et al. 2001).

One of the most remarkable features of the *Q* region class I genes is that the polymorphism characteristic of the *Mhc* is not localized at the level of individual genes, but instead it manifests at the number of genes. Inbred and wild-derived mice can be grouped by Qa2 expression level as high, medium, low, and negative, where the Qa2 expression level is thought to reflect the number of genes encoding the Qa2 antigen (Tine et al. 1990). Are

Fig. 5 **a** Dot-matrix comparison of the mouse *H2-K* and the human extended class II regions between *Daxx* and *Col11a2* (GenBank accession number AF100956) and from 42,571 bp to 107,863 of the published human *MHC* sequence (MHC Sequencing Consortium 1999). The plot was generated by PipMaker using repeat-masked sequences. *Sacm2l* (*yellow*) and *Ring1* (*green*) border the region of dissimilarity. The human *SACM2L* genomic sequence is longer due to genome-wide repeat insertions. The region between *Sacm2l* and *Ring1* contains two *K* class I genes in mouse and two pseudogenes in human (*ZNF*, zinc-finger-like and *TAT-SF1-l*, transcription factor TAT-SF1-like). **b** Schematic representation of the region between *Sacm2l* and *Ring1*from human (*HS*), mouse (*MM*), and rat (*RN*). The region is also compared with the *D-Q1* interval from mouse (*H2* class I). The rat sequence is not complete. All investigated mouse strain have two class I genes in the *K* region, whereas the number of rat class I genes varies among strains. Additional short, shared sequences were identified among these four regions by pair-wise comparison using Blast (Bl2seq) (*light blue*, *dark blue*, and *purple boxes*)

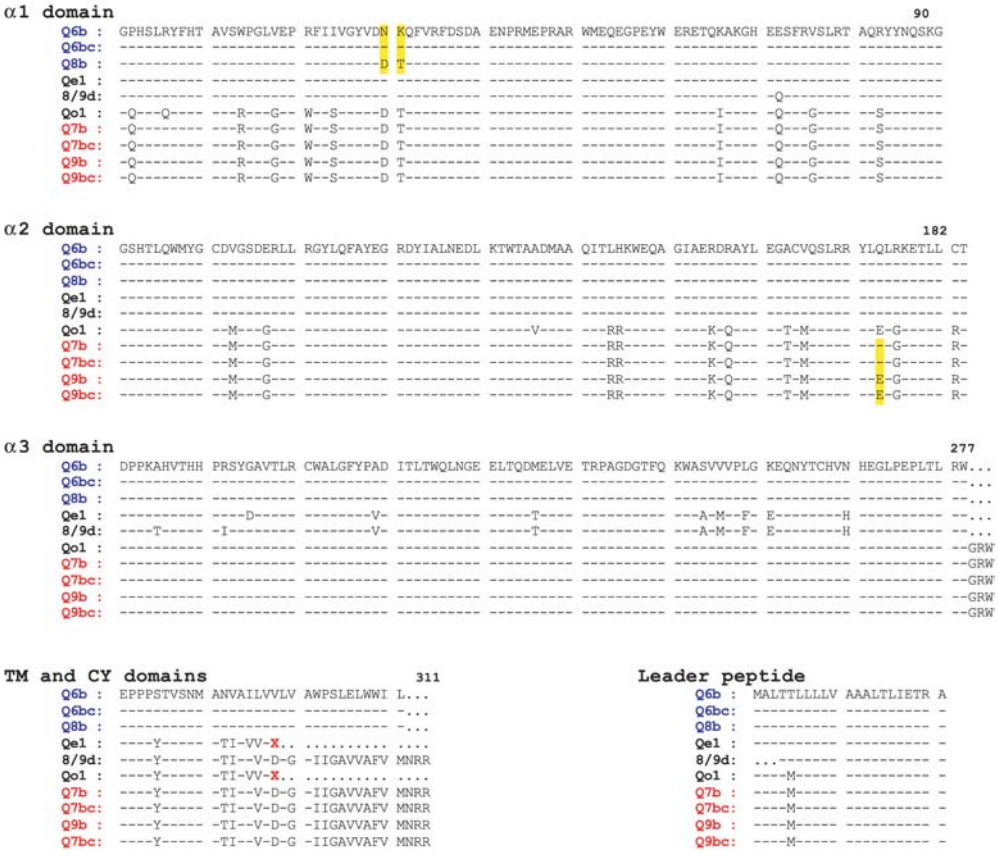


Fig. 6 The deduced amino acid sequences of the Qa2-encoding genes from 129/SvJ and the corresponding protein sequences from haplotype *b*; Q8/9^{d} is from BALB/c. The "even" (Q6 and Q8) sequence names are in *blue* and the odd (Q7 and Q9) are in *red*. The "mixed" proteins (Qe1^{bc}, Qo1^{bc}, Q8/9^{d}) are labeled in *black*. Identical residues are shown as "-". Residues that differentiate within the odd and even groups are highlighted in *yellow*. Stop codons are shown as a *red* "X"; gaps are labeled with ".". The protein sequences were derived from the following accession numbers: Q6^{b} – AAB41658, Q7^{b} – P14429, Q8^{b} – AAB41657, Q8/9^{d} – AAA39678. The complete Q9^{b} sequence was provided by Dr. Iwona Stroynowski

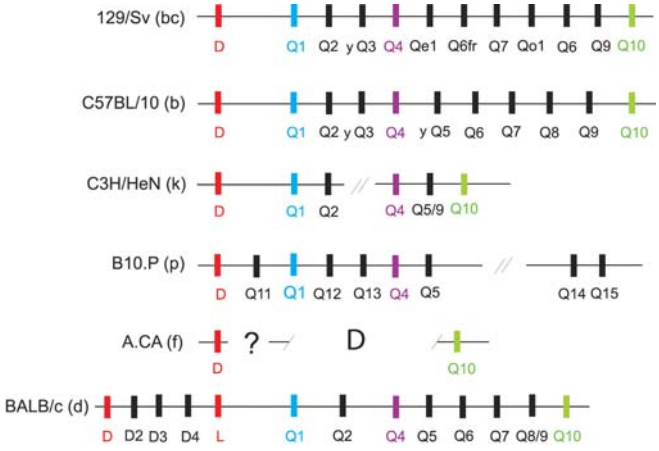


Fig. 7 Schematic comparison of the *H2-D/Q* regions from the investigated haplotypes. The class I genes present in most strains are colored: the class Ia *D* and *L* genes are *red*, *Q1* is *blue*, *Q4* is *purple,* and *Q10* is *green*. Haplotypes *b, k, d*, and *p* were mapped using cosmid clones, but not all of the class I sequences are known. In haplotype *p*, only *D, Q1, Q4*, and *Q5* were identified by gene-specific oligonucleotide hybridization, and the unidentified class I genes were tentatively named *Q11*, *12*, etc. (Litaker et al. 1996). In haplotype *d*, the *D3* and *D4* sequences are not known. Haplotype *f* was investigated only by genomic Southern hybridization (O'Neill et al. 1986)

the Qa2 level variants selected for (the region is plastic), or are the quantitative differences the result of the instability of the region (the region is fragile)? Increasing amount of experimental evidence suggests the *Q* region class Ib genes are functional and not a "junkyard" of the class I genes destined to disappear or an "experimental laboratory" packed with failures (Klein and Figueroa 1986). The quantitative variation of Qa2 among mouse strains taken together with the experimental evidence (Chiang et al. 2002; Das et al. 2000; Fragoso et al. 1998) suggests a function different from the polymorphic class Ia genes, most likely in the innate immune system. Intriguingly, the KIR/ILT NK receptor-complex analyzed in human, chimpanzee, and mouse presents a comparable genomic organization (Barten et al. 2001; Rajalingam et al. 2001; Wilson et al. 2000).

The *H2-K* region

The hypothesis of the *Q* region origin of the *K* region is based on cross-hybridization experiments. Weiss and colleagues (1984) suggested that an odd/even-like gene pair translocated from the *Q* region to a position centromeric to the class II region. They hybridized with probes derived 5′ and 3′ of *Q* class I genes and detected the similarities, which are also seen in our dot-matrix analysis (Fig. 4d). But the *H2-K* genes themselves are more closely related to *H2-D* genes than to the *Q* region genes using any criterion (Cai and Pease 1992). The contradiction of these two findings can be resolved by the sequence analysis of the *D-Q1* segment of 129/SvJ.

The segment between the *D* and *Q* region of 129 contains two copies of *Sacm2l* gene fragments and the

remnants of two class I genes (Fig. 4). *Sacm2l* is a gene of unknown function, related to the *Saccharomyces cerevisiae* Suppressor of Actin Mutation 2 (*SAC2* or YDR484w) gene (Walter and Günther 1998). *SAC2* may interact with actin as a component or controller of the assembly or stability of the actin cytoskeleton (Kölling et al. 1994; Winsor and Schiebel 1997).

The intact *Sacm2l* gene is located next to the *H2-K* class I genes in mice and to the *RT1.A* class I genes in rat [Fig. 5; (Walter and Günther 1998)]. The human extended class II region aligns very well to the *H2-K* (Fig. 5a) and *RT1.A* regions. The genes and the gene order in all three species are highly conserved (Stephens et al. 1999). The only difference among these species is between *Sacm2l* and *Ring1*. Rodents have one, two or three class I genes in this segment, whereas the corresponding human region has two non-class I pseudogenes (Fig. 5a and b) (Hanson and Trowsdale 1991; Stephens et al. 1999).

The origin of these rodent-specific class I genes is a subject of speculation. Hanson and Trowsdale (1991) suggested first that the *H2-K* class I genes are the result of an insertion into the extended class II region. Whether this insertion happened independently in mouse and rat or before the speciation is debatable (Lambracht-Washington et al. 2000; Walter and Günther 2000). Nevertheless, the ancestral *H2-K* region contributed to the *H2-D* class I genes in mouse, and this *H2-K* to *H2-D* step is strongly supported by the presence of *Sacm2l* fragments in the *D-Q1* interval of the 129/SvJ strain.

The origin and interlocus relationships of the *H2-K/D/L* loci has proven notoriously difficult to define using the class I sequences alone. While the human class I molecules show strong locus specificity (the *A*, *B*, and *C* alleles form three clearly separated groups on a phylogenetic tree), the mouse *K/D/L* sequences have weak or no allelic association (Pullen et al. 1992). Gene conversion was proposed to explain the absence of locus specificity, although the evidence for gene conversion in the *Mhc* was always weak at best (Klein and Figueroa 1986; Martinsohn et al. 1999). Interlocus exchange among the mouse class I genes could lead to homogenization of the mouse class I gene products. Human and mouse class I genes have the same three-dimensional structure and bind their peptide ligand the same way. They have similar amino acid substitutions and the selection driving the diversity appears to be the same. It is therefore hard to explain why the rodent *Mhc* would evolve differently (Hughes 1991). The presented sequence provides an alternative to ectopic gene conversion. The *K* region is traditionally viewed as an insertion derived from the class I region telomeric to the class II and class III regions (Hanson and Trowsdale 1991). Our sequence suggests that the genetic exchange was bi-directional: a segment containing class I-gene(s) was translocated from *K* to the *D* locus. This *D* region now contained at least two different ancestral class I lineages: a *D* like and a *K* like. This blended region underwent duplications and deletions. Some haplotypes lost the *K*-like lineage and others lost the *D*-like lineage. This hypothesis would explain why some haplotypes (e.g., *d*) have *D* region genes equally or more similar to *K* genes than to other *D* genes.

Acknowledgements We thank Mary Ellen Ahearn, Stephen Lasky, and Carol Loretz for help with mapping and sequencing, and Dr. Iwona Stroynowski for critical reading of the manuscript and for the sequence of $Q9^b$. This work was supported by NIH grant AI 37818. The performed experiments comply with the current laws of the United States.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ(1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Amadou C, Kumánovics A, Jones EP, Lambracht-Washington D, Yoshino M, Fischer Lindahl K (1999) The mouse major histocompatibility complex: some assembly required. Immunol Rev 167:211–221

Barten R, Torkar M, Haude A, Trowsdale J, Wilson MJ (2001) Divergent and convergent evolution of NK-cell receptors. Trends Immunol 22:52–57

Cai Z, Pease LR (1992) Structural and functional analysis of three D/L-like class I molecules from $H\text{-}2^v$: indications of an ancestral family of D/L genes. J Exp Med 175:583–596

Chiang EY, Henson M, Stroynowski I (2002) The nonclassical major histocompatibility complex molecule Qa-2 protects tumor cells from NK cell- and lymphokine-activated killer cell-mediated cytolysis. J Immunol 168:2200–2211

Das G, Gould DS, Augustine MM, Fragoso G, Scitto E, Stroynowski I, Kaer LV, Schust DJ, Ploegh H, Janeway CA Jr (2000) Qa-2-dependent selection of CD8α/α T cell receptor α/β+ Cells in murine intestinal intraepithelial lymphocytes. J Exp Med 192:1521–1528

Fischer Lindahl K (1997) On naming *H2* haplotypes: functional significance of MHC class Ib alleles. Immunogenetics 46:53–62

Flaherty L (1988) Major histocompatibility complex polymorphism: a nonimmune theory for selection. Hum Immunol 21:3–13

Flaherty L, Elliott E, Tine JA, Walsh AC, Waters J B (1990) Immunogenetics of the Q and TL regions of the mouse. Crit Rev Immunol 10:131–175

Fragoso G, Lamoyi E, Mellor A, Lomelí C, Hernández M, Sciutto E (1998) Increased resistance to *Taenia crassiceps* murine cysticercosis in Qa-2 transgenic mice. Infect Immun 66:760–764

Glusman G, Rowen L, Lee I, Boysen C, Roach JC, Smit AFA, Wang K, Koop BF, Hood L (2001) Comparative genomics of the human and mouse T-cell receptor loci. Immunity 15:337–349

Goldbard SB, Verbenac KM, Warner CM (1982) Genetic analysis of *H-2* linked gene(s) affecting early mouse embryo development. J Immunogenet 9:77–82

Günther E, Walter L (2001) The major histocompatibility complex of the rat (*Rattus norvegicus*). Immunogenetics 53:520–542

Hanson IM, Trowsdale J (1991) Colinearity of novel genes in the class II regions of the MHC in mouse and human. Immunogenetics 34:5–11

He X, Tabaczewski P, Ho J, Stroynowski I, Garcia KC (2001) Promiscuous antigen presentation by the nonclassical MHC Ib Qa-2 is enabled by a shallow, hydrophobic groove and self-stabilized peptide conformation. Structure 9:1213–1224

Hill F, Gemünd C, Benes V, Ansorge W, Gibson T J (2000) An estimate of large-scale sequencing accuracy. EMBO Rep 1:29–31

Hughes AL (1991) Independent gene duplications, not concerted evolution, explain relationships among class I MHC genes of murine rodents. Immunogenetics 33:367–373

Joyce S, Tabaczewski P, Angeletti RH, Nathenson SG, Stroynowski I (1994) A nonpolymorphic major histocompatibility complex class Ib molecule binds a large array of diverse self-peptides. J Exp Med 179:579–588

Klein J, Figueroa F (1986) Evolution of the major histocompatibility complex. Crit Rev Immunol 6:295–386

Kölling R, Lee A, Chen EY, Botstein D (1994) Nucleotide sequence of the *SAC2* gene of *Saccharomyces cerevisiae*. Yeast 10:1211–1216

Kress M, Cosman D, Khoury G, Jay G (1983) Secretion of a transplantation-related antigen. Cell 34:189–196

Kuff EL, Lueders KK (1988) The intracisternal A-particle family: structure and functional aspects. Adv Cancer Res 51:183–276

Lambracht-Washington D, Fischer Lindahl K (2002) Does the rat have an *H2-D* ortholog next to *Bat1*? Immunogenetics 53:1039–1046

Lambracht-Washington D, Fischer Lindahl K, Wonigeit K (2000) Promoter structures suggest independent translocations of ancestral rat *RT1.A* and mouse *H2-K* class I genes. Immunogenetics 51:873–877

Lew AM, Maloy WL, Coligan JE (1986) Characteristics of the expression of the murine soluble class I molecule (Q10). J Immunol 136:254–258

Litaker W, Peace-Brewer A, Frelinger JA (1996) A physical map of the Q region of B10.P. Mamm Genome 7:200–205

Martinsohn JT, Sousa AB, Guethlein LA, Howard JC (1999) The gene conversion hypothesis of MHC evolution: a review. Immunogenetics 50:168–200

Matsuura A, Schloss R, Shen F-W, Tung J-S, Hunt SW 3rd, Fisher DA, Hood L, Boyse EA (1989) Expression of the *Q8/9^d* gene by T cells of the mouse. Immunogenetics 30:156–161

Mellor AL, Weiss EH, Kress M, Jay G, Flavell R (1984) A nonpolymorphic class I gene in the murine major histocompatibility complex. Cell 36:139–144

MHC Sequencing Consortium T (1999) Complete sequence and gene map of a human major histocompatibility complex. Nature 401:921–923

Nakayama K, Tokito S, Pannetier C, Nakauchi H, Gachelin G (1991) *MHC* gene *Q8/9^d* of BALB/cJ mouse strain cannot encode a Qa-2,3 class I antigen. Immunogenetics 33:225–234

O'Neill AE, Reid K, Garberi JC, Karl M, Flaherty L (1986) Extensive deletions in the *Q* region of the mouse major histocompatibility complex. Immunogenetics 24:368–373

Pascolo S, Bervas N, Ure JM, Smith AG, Lemonnier FA, Pérarnau B (1997) HLA-A2.1-restricted education and cytolytic activity of CD8$^+$ T lymphocytes from β2 microglobulin (β2m) HLA-A2.1 monochain transgenic H-D^b β2m double knockout mice. J Exp Med 185:2043–2051

Pullen JK, Horton RM, Cai Z, Pease LR (1992) Structural diversity of the classical H-2 genes: *K*, *D*, and *L*. J Immunol 148:953–967

Radice P, Pensotti V, Jones C, Perry H, Pierotti MA, Tunnacliffe A (1995) The human archain gene, *ARCN1,* has highly conserved homologs in rice and *Drosophila*. Genomics 26:101–106

Rajalingam R, Hong M, Adams EJ, Shum BP, Guethlein LA, Parham P (2001) Short *KIR* haplotypes in pygmy chimpanzee (bonobo) resemble the conserved framework of diverse human KIR haplotypes. J Exp Med 193:135–146

Robinson PJ, Bevec D, Mellor AL, Weiss EH (1988) Sequence of the mouse Q4 class I gene and characterization of the gene product. Immunogenetics 27:79–86

Rötzschke O, Falk K, Stevanovic S, Grahovac B, Soloski MJ, Jung G, Rammensee HG (1993) Qa-2 molecules are peptide receptors of higher stringency than ordinary class I molecules. Nature 361:642–644

Rowen L, Lasky L, Hood L (1999) Deciphering genomes through automated large-scale sequencing. Methods Microbiol 28:155–192

Rubocki RJ, Lee DR, Lie W-R, Myers NB, Hansen TH (1990) Molecular evidence that the H-2D and H-2L genes arose by duplication. Differences between the evolution of the class I genes in mice and humans. J Exp Med 171:2043–2061

Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W (2000) PipMaker – a web server for aligning two genomic DNA sequences. Genome Res 10:577–586

Schwemmle S, Bevec D, Brem G, Urban MB, Baeuerle PA, Weiss EH (1991) Developmental and tissue-specific expression of the *Q5* k gene. Immunogenetics 34:28–38

Smit AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev 9:657–663

Soloski MJ, Vernachio J, Einhorn G, Lattimore A (1986) Qa gene expression: biosynthesis and secretion of Qa-2 molecules in activated T cells. Proc Natl Acad Sci U S A 83:2949–2953

Stephan D, Sun H, Fischer Lindahl K, Meyer E, Hämmerling G, Hood L, Steinmetz M (1986) Organization and evolution of D region class I genes in the mouse major histocompatibilty complex. J Exp Med 163:1227–1244

Stephens R, Horton R, Humphray S, Rowen L, Trowsdale J, Beck S (1999) Gene organization, sequence variation and isochore structure at the centomeric boundary of the human MHC. J Mol Biol 291:789–799

Stroynowski I, Soloski M, Low MG, Hood L (1987) A single gene encodes soluble and membrane-bound forms of the major histocompatibility Qa-2 antigen: anchoring of the product by a phospholipid tail. Cell 50:759–768

Tabaczewski P, Shirwan H, Lewis K, Stroynowski I (1994) Alternative splicing of class Ib major histocompatibility complex transcripts in vivo leads to the expression of soluble Qa-2 molecules in murine blood. Proc Natl Acad Sci USA 91:1883–1887

Tabaczewski P, Chiang E, Henson M, Stroynowski I (1997) Alternative peptide binding motifs of Qa-2 class Ib molecules define rules for binding of self and nonself peptides. J Immunol 159:2771–2781

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tine JA, Walsh A, Rathburn D, Leonard L, Wakeland EK, Dilwith R, Flaherty L (1990) Genetic polymorphisms of *Q* region genes from wild-derived mice: implications for *Q* region evolution. Immunogenetics 31:315–325

Tunnacliffe A, Vrugt H van de, Pensotti V, Radice P (1996) The coatomer protein δCOP, encoded by the archain gene, is conserved across diverse eukaryotes. Mamm Genome 7:784–786

Ulker N, Hood LE, Stroynowski I (1990a) Molecular signals for phosphatidylinositol modification of the Qa-2 antigen. J Immunol 145:2214–2219

Ulker N, Lewis KD, Hood LE, Stroynowski I (1990b) Activated T cells transcribe an alternatively spliced mRNA encoding a soluble form of Qa-2 antigen. EMBO J 9:3839–3847

Ungchusri T, Chiang EY, Brown G, Chen M, Tabaczewski P, Timares L, Stroynowski I (2001) Widespread expression of the nonclassical class I Qa-2 antigens in hemopoietic and nonhemopoietic cells. Immunogenetics 53:455–467

Vugmeister Y, Glas R, Pérarnau B, Lemonnier FA, Eisen H, Ploegh H (1998) Major histocompatibility complex (MHC) class I K^{b}D^b -/- deficient mice possess functional CD8+ T cells and natural killer cells. Proc Natl Acad Sci USA 95:12492–12497

Walter L, Günther E (1998) Identification of a novel highly conserved gene in the centromeric part of the major histocompatibility complex. Genomics 52:298–304

Walter L, Günther E (2000) Physical mapping and evolution of the centromeric class I gene-containing region of the rat MHC. Immunogenetics 51:829–837

Waneck GL, Sherman DH, Calvin S, Allen H, Flavell RA (1987) Tissue-specific expression of cell-surface Qa-2 antigen from a transfected Q7^b gene of C57BL/10 mice. J Exp Med 165:1358–1370

Waneck GL, Stein ME, Flavell RA (1988) Conversion of a PI-anchored protein to an integral membrane protein by a single amino acid mutation. Science 241:697–699

Watts S, Davis AC, Gaut B, Wheeler C, Hill L, Goodenow RS (1989) Organization and structure of the Qa genes of the major histocompatibility complex of the C3H mouse: implications for Qa function and class I evolution. EMBO J 8:1749–1759

Weiss EH, Golden L, Fahrner K, Mellor AL, Devlin JJ, Bullman H, Tiddens H, Bud H, Flavell RA (1984) Organization and evolution of the class I family in the major histocompatibility complex of the C57BL/10 mouse. Nature 310:650–655

Weiss EH, Bevec D, Messer G, Schwemmle S, Großhaus C, Steinmetz S, Schmidt W (1989) Organization of the AKR Qa region: structure of a divergent class I sequence, $Q5^k$. J Immunogenet 16:283–290

Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D, Beck S, Trowsdale J (2000) Plasticity in the organization and sequences of human KIR/ILT gene families. Proc Natl Acad Sci USA 97:4778–4783

Winsor B, Schiebel E (1997) Review: an overview of the *Saccharomyces cerevisiae* microtubule and microfilament cytoskeleton. Yeast 13:399–434

Wroblewski JM, Kaminsky SG, Nakamura I (1994) *Bat1* genes and the origin of multiple class I loci in the *H-2D* region. Immunogenetics 39:276–280

Wu L, Feng H, Warner CM (1999) Identification of two major histocompatibility complex class Ib genes, *Q7* and *Q9*, as the *Ped* gene in the Mouse. Biol Reprod 60:1114–1119

Xu Y, Jin P, Mellor AL, Warner CM (1994) Identification of the *Ped* gene at molecular level: the Q9 MHC class I transgene converts the *Ped slow* to *Ped fast* phenotype. Biol Reprod 51:695–699

Zappacosta F, Tabaczewski P, Parker KC, Coligan JE, Stroynowski I (2000) The murine liver-specific nonclassical MHC class I molecule Q10 binds a classical peptide repertoire. J Immunol 164:1906–1915

Zhu X, Naz RK (1997) Fertilization antigen-1: cDNA cloning, testing, testis-specific expression, and immunocontraceptive effects. Proc Natl Acad Sci U S A 94:4704–4709