Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 2167–2171, March 1990
Evolution

Concerted evolution of class I genes in the major histocompatibility complex of murine rodents

CRISTINA RADA*†, ROBERTO LORENZI*, SIMON J. POWIS‡, JOHAN VAN DEN BOGAERDE*§, PETER PARHAM¶, AND JONATHAN C. HOWARD*||

*Department of Immunology and ‡Monoclonal Antibody Centre, Agricultural and Food Research Council, Institute of Animal Physiology and Genetics Research, Cambridge Research Station, Babraham, Cambridge CB2 4AT, United Kingdom; and ¶Department of Cell Biology, Stanford University School of Medicine, Stanford, CA 94305

Communicated by Irving Weissman, November 7, 1989 (received for review August 29, 1989)

ABSTRACT **Full-length cDNA sequences of two class I major histocompatibility complex molecules from the DA strain of *Rattus norvegicus* are reported. One codes for the classical class I restriction element RT1.A^a, which maps to the locus in the rat major histocompatibility complex homologous to *H-2K* in the mouse. The other probably codes for a soluble nonclassical class I molecule present in DA rat serum; a short deletion in the fifth exon implies that the translated product will terminate in the membrane-spanning region. These sequences have been compared with mouse classical class I sequences as well as with three published rat class I cDNA partial sequences. The results show, first, that "locus-specific" substitutions from the H-2K, H-2D, and H-2L data set are scrambled in the RT1.A^a molecule; a majority of these substitutions have H-2D/L-specific features. Second, the data show that the four rat sequences are strikingly similar to one another regardless of locus or haplotype of origin; they share a number of apparently species-specific features that distinguish them all from mouse classical class I sequences, which likewise share distinctive features of their own. The results suggest that segmental sequence exchange plays a major role in determining the evolution of sequence in class I major histocompatibility complex molecules.**

Some multigene families display a degree of species specificity at the level of DNA sequence, as if the whole family were evolving within a species as a unit (1–3). Such concerted evolution has been suspected in mouse major histocompatibility complex (MHC) class I genes because of the lack of clear locus-specific sequence variation distinguishing the widely separated H-2K and H-2D end classical loci (4, 5). Short segmental exchanges of sequence are indeed occurring between widely separated class I MHC loci in laboratory mice (6). Nevertheless it remains controversial whether such processes contribute significantly to the pattern of evolution of class I genes (7), since sequences of class I genes from a closely related murine species such as the rat are not yet available.

Rattus norvegicus is in the same subfamily, the Murinae, as the mouse. The rat MHC, *RT1*, is like *H-2* in structure, with a classical class I gene, *RT1.A*, at the left end of the MHC in the same location as *H-2K* and a long array of other class I genes to the right (8). In the middle is a region specifying class II genes that closely resembles the homologous region of *H-2* (9). It is most likely that the translocation of class I genes that formed the H-2K end was a unique event occurring shortly before the separation of the rat and mouse lineages (7). Rat class I gene sequences thus provide a stringent test of the concerted evolution hypothesis. With simple divergent evolution alone, the sequence of an *RT1.A* allele should have greater similarity to *H-2K* on average than to other rat class I sequences, since *RT1.A* and *H-2K* share a more recent common ancestor than either does with genes at the other end of the MHC. If, on the other hand, concerted evolution is at work in this system, *RT1.A* should resemble other rat class I genes more closely than it does *H-2K*.

Here we report the full-length sequence of an RT1.A classical class I rat MHC molecule, as well as that of a nonclassical class I molecule from the same haplotype.** Using these and other published rat and mouse class I sequences, we show concerted evolution operating on many class I genes from these two murine species.

MATERIALS AND METHODS

Preparation and Screening of a DA Strain Rat cDNA Library. A cDNA library in λgt10, constructed essentially as described (10), was prepared from DA strain male rat Con A blasts and screened with the mouse 3′ class I cDNA probe, pH-2IIa (11). Clones containing the classical class I sequence, RT1.A^a, were identified by a second screen with an 18-mer oligonucleotide probe, TATTTCTACAC$^{A}_{C}$GCNGT$^{G}_{T}$ (where N = G, A, T, or C), representing amino acids 7–12 (YFYTAV) of the published N-terminal sequence (12). Thirteen clones were identified as positive with both the cDNA and oligonucleotide probes. Clones 3.3/1 and 3.6 were selected for analysis.

Sequencing Strategy and Sequence Analysis. Inserts were subcloned into Bluescript KS(+) (Stratagene) and sequenced on both strands from exonuclease III/exonuclease VII nested deletion sets (13). Sequences were assembled using Staden programs (14) and analysis was with the University of Wisconsin Genetics Computer Group programs (15). Published class I sequences were assembled from the following Genbank files: K^b, MUSMHKB1 and MUSMHKB2; K^d, MUSMHKDA; K^k, MUSMHKK; K^{w28}, MUSMHKTA; K^q, MUSMHKQM; D^k, MUSMHH2DK; D^p, MUSMHH2DP; D^d, MUSMHDD; L^d, MUSMHLDA3 and MUSMHLDB1; D^b, MUSMHH2DB; pRTI.A-1, RATMHC2; pRTI.A-2, RATMHC1. The rat class I cDNA sequence pARI.5 was taken from ref. 16.

Expression of Rat Class I cDNAs in Mouse Cells. Inserts from pBS3.3/1 and pBS3.6 were subcloned into the expression vector pMSD (17). pMSD3.3/1 and pMSD3.6 were transfected into C3HAn/Boy mouse Ltk$^-$ fibroblasts and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MHC, major histocompatibility complex; CTL, cytotoxic T lymphocyte.

†Present address: Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, U.K.
§Present Address: Department of Surgery, Addenbrookes Hospital, University of Cambridge, Hills Road, Cambridge CB2 2QH, U.K.
||To whom reprint requests should be addressed.

**The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M31018 and M31038).

AKR mouse hypoxanthine phosphoribosyltransferase-deficient BW5147 thymoma cells by using calcium phosphate and the selection plasmid pBR322/ASV-2LTR/tkA (Ltk$^-$ cells) (18) or electroporation and the selection plasmid pMSD-HGPRT (BW5147). Surface expression of rat class I molecules on transfected mouse cells was sought by using the mouse anti-rat class I monoclonal antibodies MRCOX18 (19) and F16.4.4 (20). RT1.A^a was identified by using four rat alloantibodies against distinct epitopes (21). Antibody binding was detected with a fluorescent anti-immunoglobulin antibody by using a model 420 fluorescence-activated cell sorter (Becton Dickinson). Rat cytotoxic T cells specific for RT1.A^a were prepared and assayed as described (22). Mouse cytotoxic T lymphocytes (CTLs) with a similar specificity were prepared by stimulating C3H/AnBoy mouse spleen cells with PVG.R1 rat lymph node cells. Inhibition of cytotoxicity by the RT1.A^a-specific alloantibody R2/15S was as described (22).

Rat RT1.A^a class I molecules were precipitated from detergent lysates of surface-iodinated Ltk3.3/1 cells by using MRCOX18; mouse H-2K^k and H-2D^k class I molecules were precipitated from the same cells by using the rat anti-mouse class I antibody R1-21.2 (23). Both antibodies were conjugated directly to Sepharose CL-4B.

RESULTS

Sequence and Structure of Clones 3.3/1 and 3.6. Fig. 1 shows the nucleotide sequences and translation of two DA rat class I cDNA clones. The translation product of 3.3/1 corresponds to a typical eight-exon class I gene. The translation product of 3.6 has a deletion of 13 nucleotides (931–943) in the presumed fifth exon, causing premature termination in the transmembrane domain. The second deletion of 33 nucleotides (1040–1072) in 3.6 corresponds to the entire sixth exon.

3.3/1 contains the 18-mer probe sequence used for library screening and is consistent with the DA rat classical class I molecule RT1.A^a (12). Clone 3.6 departs from the probe sequence by one nucleotide: histidine is substituted for tyrosine at position 9 of the first domain. The termination of the 3.6 translated product in the transmembrane domain and the histidine at position 9 of the first domain are consistent with a nonclassical class I molecule present in DA rat serum (12).

3.3/1 Specifies RT1.A^a. The 3.3/1 product expressed on a BW5147 transfectant carried six epitopes typical of RT1.A^a (Fig. 2*A*). The epitopes identified by the alloantibodies R3/13, YR1/100, MAC210, and JY1/232 represent an antigenic signature found only on the RT1.A^a classical class I molecule (21). The 3.3/1 product is a typical class I molecule with a heavy chain of slightly slower mobility than H-2K^k and H-2D^k (Fig. 2*B*). Rat and mouse CTLs generated against RT1.A^a in mixed lymphocyte cultures are specific for the rat class I product expressed on pMSD3.3/1 transfectant mouse cells (Fig. 2*C*). Furthermore, target lysis is blocked by the RT1.A^a-specific alloantibody, R2/15S (Fig. 2*C Inset*). These experiments establish the 3.3/1 sequence as representing the *RT1.A^a* gene.

Comparison of RT1.A^a with Mouse and Rat Class I Sequences. A comparison between RT1.A^a and mouse H-2K, H-2D, and H-2L sequences was undertaken to identify locus-specific features. As frequently noted, the three long exons

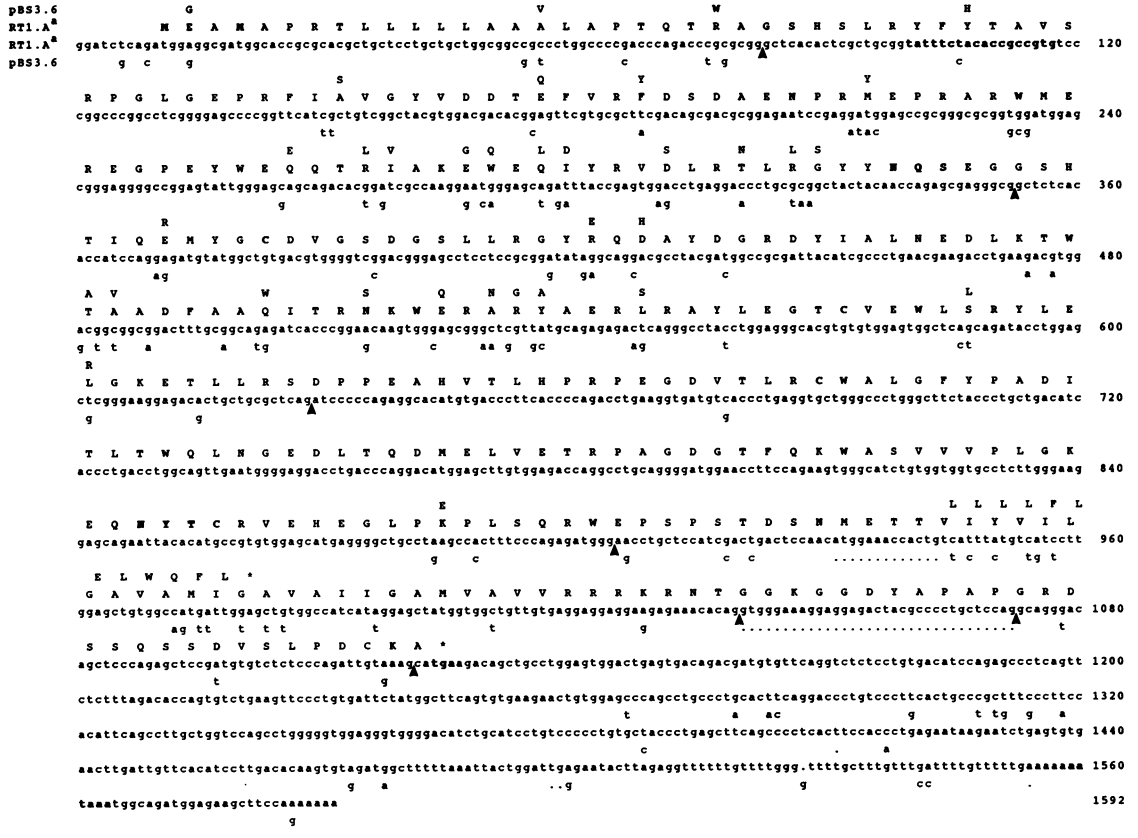

FIG. 1. Nucleotide sequence and translation of cDNA clones pBS3.3/1 (RT1.A^a) and pBS3.6. Sequences are aligned to optimize colinearity; gaps are indicated by dots. Putative exon boundaries are indicated by arrowheads. The following are highlighted in boldface letters: the oligonucleotide probe sequence tatttctacaccgccgtg at bases 100–117, putative N-linked glycan attachment signals at bases 347 and 847, conserved cysteine residues associated with intradomain disulfide bridges at bases 382 and 571 and bases 688 and 856, and a putative aataaa polyadenylylation signal starting at base 1558. The tag and tga termination codons at bases 980–982 (pBS3.6) and 1123–1125 (RT1.A^a), respectively, are indicated by an asterisk.

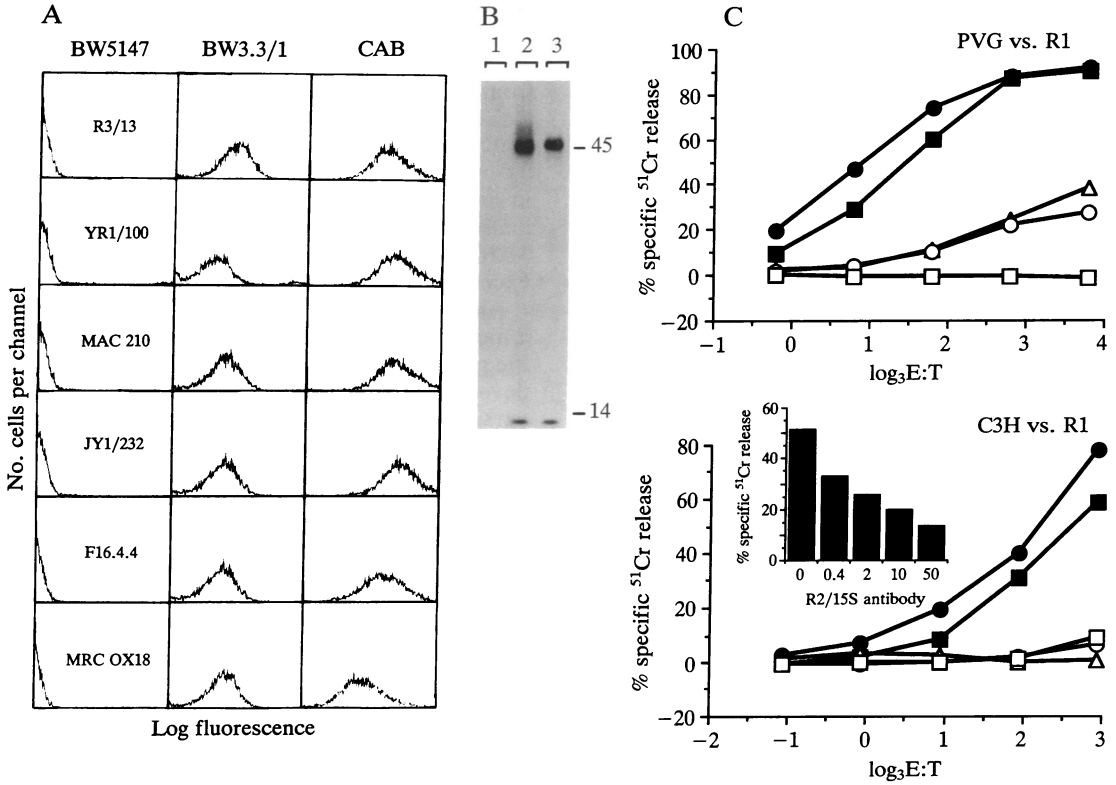

FIG. 2. The 3.3/1 sequence codes for the RT1.A^a classical class I glycoprotein. (*A*) Expression of RT1.A^a on the surface of BW5147 transfected with pMSD3.3/1 (BW3.3/1) assayed by monoclonal allo- and xenoantibodies in the fluorescence-activated cell sorter. Nontransfected BW5147 and DA strain rat Con A blasts (CAB) were negative and positive controls, respectively. (*B*) Immunoprecipitation of RT1.A^a by MRC OX18 from surface-iodinated Ltk$^-$ cells transfected with pMSD3.3/1 (lane 3). Immunoabsorbents conjugated with an irrelevant monoclonal antibody (lane 1) and R1-21.2 (lane 2), specific for mouse H-2D^k and H-2K^k class I molecules, were negative and positive controls, respectively. (*C*) Lysis of BW3.3/1 targets (●) by PVG rat anti-PVG.R1 (*Upper*) and C3H mouse anti-PVG.R1 (*Lower*) cytotoxic T-cell populations. PVG.R1 (■), PVG (□), and C3H mouse (△) Con A blast targets, as well as nontransfected BW5147 (○), were also used to indicate the specificity of the CTL populations. (*Inset*) Inhibition of the lysis of BW3.3/1 by C3H anti-PVG.R1 CTLs with the RT1.A^a-specific monoclonal alloantibody R2/15S at concentrations of 0–50 μg/ml. E:T, effector to target ratio.

of *H-2* classical class I alleles showed no evidence of locus-specific divergence (Fig. 3 and data not shown). Only a single position out of 822 compared showed complete locus specificity. However, in the short exons and untranslated regions up to the beginning of the short interspersed nuclear element (SINE) insertion (25), 24 positions were fully locus-specific out of 660 compared (Fig. 4). Out of the total of 25 apparently locus-specific positions, RT1.A^a showed the H-2D/H-2L-specific nucleotide at 15 positions, the H-2K-specific nucleotide at 5 positions, and a third nucleotide at 5 positions.

When all available rat class I sequences were compared with mouse classical sequences, a striking degree of species-specific homogenization was apparent. A matrix of comparisons in the three long exons (Fig. 3) showed that the rat sequences departed significantly from the interallelic or interlocus comparisons within the mouse species group. However, despite representing the products of different loci, the three available rat sequences, including RT1.A^a, were strikingly more similar to each other than to the mouse classical class I sequences.

The short exons and untranslated regions also displayed extensive species-specific homogenization (Fig. 4); this was conspicuous in a series of deletions and insertions by which the two sequence groups differed. Only completely homogenized positions are indicated in Fig. 4, but many other partially homogenized positions and motifs are apparent.

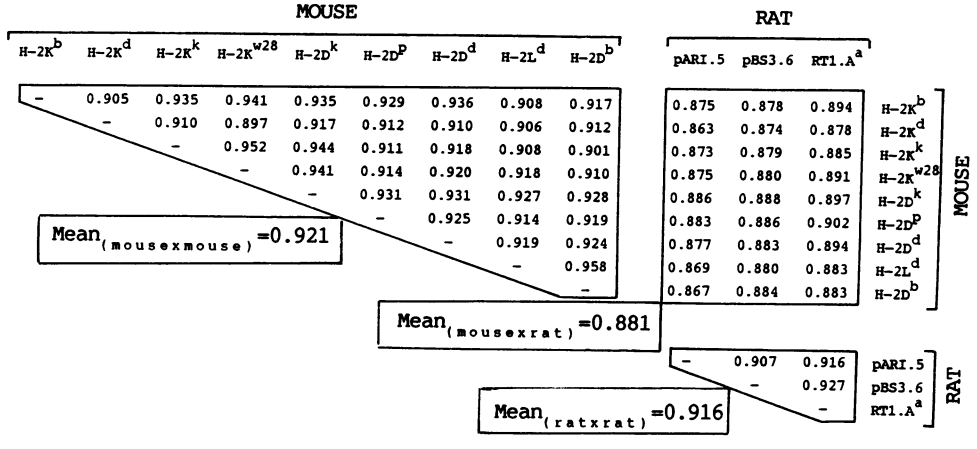

FIG. 3. Similarity matrix for exons 2–4 (822 nucleotides compared) of mouse and rat class I sequences. By a simple randomization argument, the rat sequences differ significantly from the mouse group (P = 0.0045), since the observed arrangement of the data gives rise to the most extreme values of the intra- and interspecific means out of the 220 different possible arrangements. The sources of the sequences are given in *Materials and Methods*.

DISCUSSION

This paper documents species-specific sequence homogenization in the classical and nonclassical class I genes of mice and rats. This result demonstrates beyond reasonable doubt that unconventional mechanisms of sequence exchange, such as gene conversion, are significant in determining the pattern of murine class I sequence evolution.

The outstanding problem now is to understand the interplay between mutation and segmental exchange, selection, and drift in originating and sustaining polymorphism in class I MHC molecules. Allelic diversity is apparently accumulating in the population faster than single-base replacements. Thus RT1.A^a differs at about 90 positions on average from any one mouse classical allele in the three long exons. Yet only 11 nucleotides in these three exons are unique to RT1.A^a—that is, not found in any of the nine mouse alleles sequenced so far. We therefore conclude that segmental exchange between alleles and between loci has been the principal mode of sequence evolution in classical class I genes in the murine rodents.

The scrambling in RT1.A^a of locus-specific markers for H-2K and H-2D/L suggests that the apparent locus specificity arises from stochastic properties of the homogenization process. Perhaps only the B2-SINE insertion in the 3′ untranslated region (25, 26) is truly a locus-specific marker in these sequences.

Segmental exchange between HLA loci has been proposed from patterns of sequence shared between alleles at different loci (27–29). However, clear-cut divergence of sequence between loci in HLA suggests that intralocus exchanges play a more important evolutionary role than interlocus exchanges in generating polymorphism in humans and the higher primates (29, 30). Extensive recent duplication in the murine MHCs may favor nonhomologous pairing and thus a relative excess of interlocus exchanges in this taxonomic group.

An explicit genetic mechanism in the MHC that generates a high frequency of short segmental sequence exchanges in the germ line has been proposed (31). The exploitation of a developmentally regulated, tissue-specific mitotic gene conversion process to generate somatic diversity in chicken immunoglobulin V genes (32) may be relevant since it is reminiscent of previous evidence that germ-line conversion in class I genes may also be a mitotic rather than a meiotic event, presumably occurring during germ cell multiplication in fetal life (24).

We thank George Gutman for providing DA strain rats; Dan Denney for his gift of pMSD plasmids; David Lawlor, Tony Jackson, Steve Jameson, Gos Micklem, and David Brown for advice; Nigel

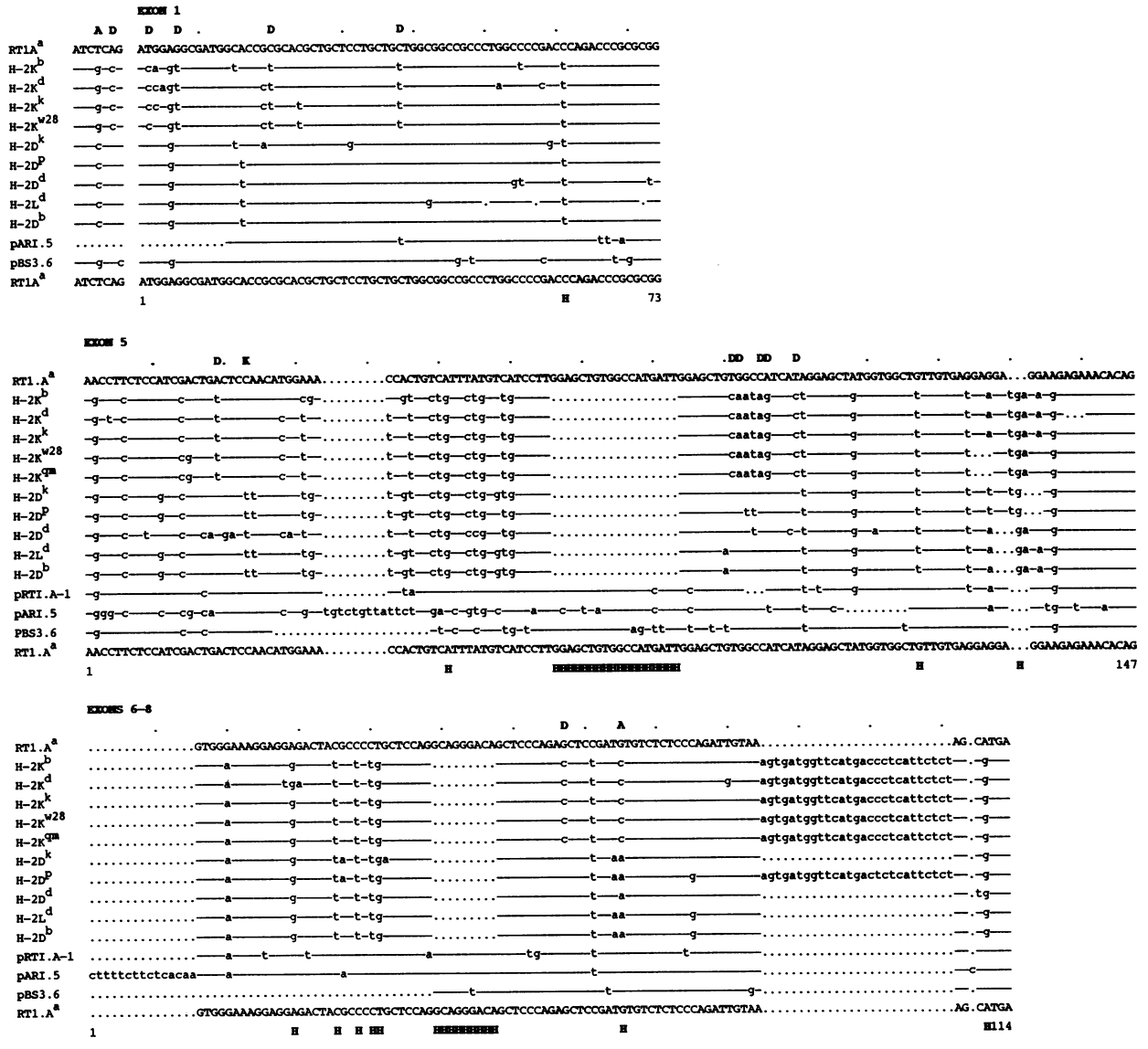

FIG. 4. (*Figure continues on opposite page.*)

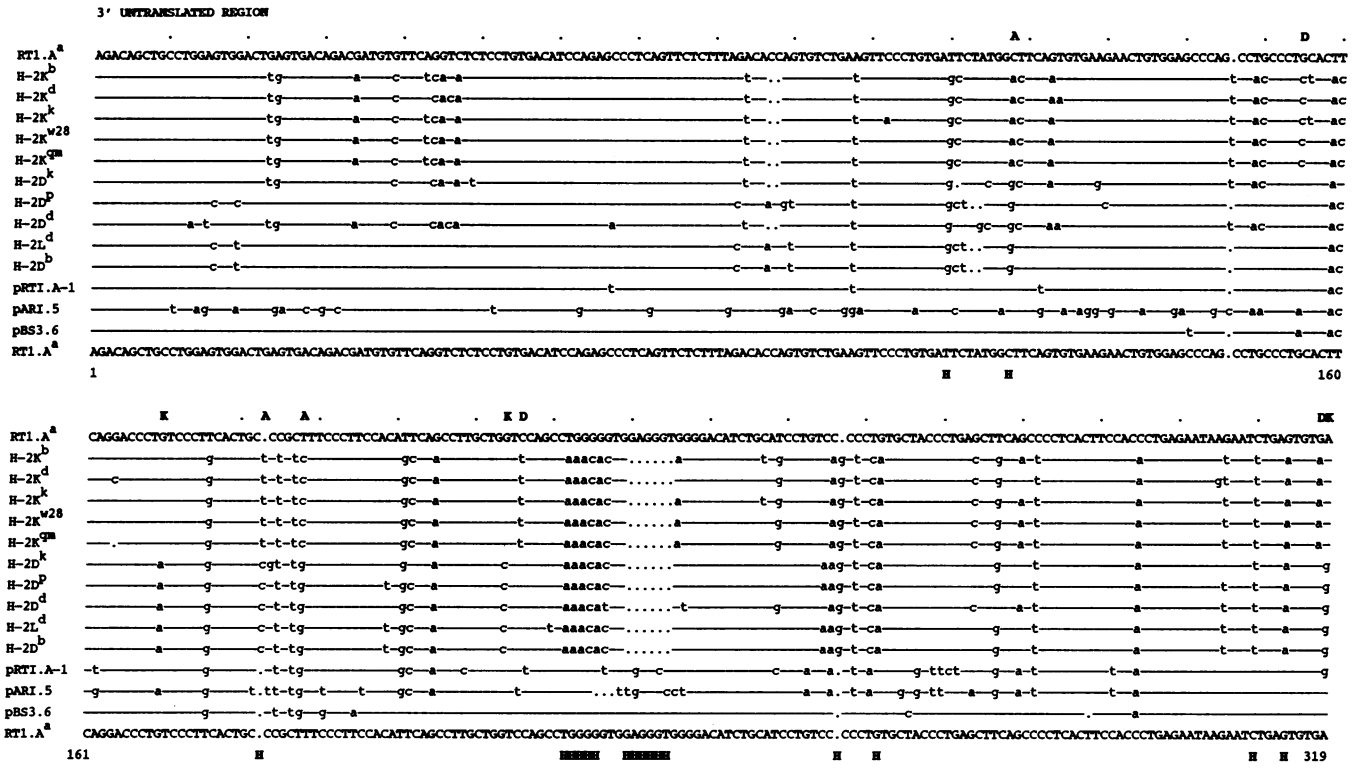

FIG. 4. Sequence alignments for the 5′ untranslated region, short exons, and 3′ untranslated region to nucleotide 319 [the H-2D/L-specific SINE insertion (25), also found in rat (26) but absent from 3.3/1, begins at nucleotide 320 of the 3′ untranslated region] of nine mouse H-2K and H-2D/L locus class I alleles and four rat class I sequences. The sequences are compared against the full sequence of RT1.A^a (top and bottom lines, clone 3.3/1). Locus-specific nucleotides among the mouse classical sequences are marked by letters above the alignment blocks. These letters indicate whether RT1.A^a has the H-2D/L (D) or H-2K (K) base at these positions. Positions where RT1.A^a follows neither H-2D/L nor H-2K are indicated by the letter A. Positions exhibiting species-specific homogenization are indicated by the letter H below the alignment blocks. In the case of single nucleotides, the letter H indicates complete homogenization in both species at each position. For insertions and deletions, the letter H refers only to the presence or absence of the differential segment. The sources of the sequences are given in *Materials and Methods*. Gaps are indicated by dots, and identity with the sequence of RT1.A^a is indicated by a dash.

Miller for fluorescence-activated cell sorter analysis; Evelyn Tredget and Cariel Taylor-Edwards for technical assistance; and Gabriel Dover and Nick Holmes for reading the manuscript. This work was supported in part by grants from the Agricultural Food Research Council Cell Signalling and Recognition Programme and the National Institutes of Health. S.J.P. was supported by a studentship from the Agricultural and Food Research Council; J.v.d.B. was supported by the Bradlow Foundation; and R.L. was supported by St. John's College, Cambridge. J.C.H. acknowledges the generous hospitality from Garry Fathman and his colleagues.

1. Hood, L., Campbell, J. H. & Elgin, S. (1974) *Annu. Rev. Genet.* **9**, 305–353.
2. Dover, G. A. (1982) *Nature (London)* **299**, 111–117.
3. Dover, G. A. & Strachan, T. (1987) in *Evolution and Vertebrate Immunity*, eds. G. Kelsoe & D. H. Schulze (Univ. of Texas Press, Austin), pp. 15–33.
4. Lafay, F., Ewald, S. J., McMillan, M., Frelinger, J. A. & Hood, L. (1981) *Immunogenetics* **12**, 21–32.
5. Baltimore, D. (1981) *Cell* **24**, 592–594.
6. Melvold, R. W. & Kohn, H. I. (1989) in *Transgenic Mice and Mutants in Major Histocompatibility Complex Research*, eds. Egorov, I. & David, C. S. (Springer, New York), in press.
7. Klein, J. & Figueroa, F. (1986) *Crit. Rev. Immunol.* **6**, 295–386.
8. Gunther, E., Wurst, W., Wonigeit, K. & Epplen, J. T. (1985) *J. Immunol.* **124**, 1257–1261.
9. Diamond, A. G., Hood, L., Howard, J. C., Windle, M. & Winoto, A. (1989) *J. Immunol.* **142**, 3268–3274.
10. Ennis, P. D., Jackson, A. J. & Parham, P. (1988) *J. Immunol.* **141**, 642–651.
11. Steinmetz, M., Moore, K. W., Frelinger, J. G., Sher, B. T., Boyse, E. A. & Hood, L. (1981) *Cell* **25**, 683–692.
12. Spencer, S. C. & Fabre, J. W. (1987) *J. Exp. Med.* **165**, 1595–1608.
13. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103–109.
14. Staden, R. (1984) *Nucleic Acids Res.* **12**, 499–503.
15. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
16. Radojcic, A., Stranick, K. S., Locker, J., Kunz, H. W. & Gill, T. J. (1989) *Immunogenetics* **29**, 134–137.
17. Holmes, N., Ennis, P., Wan, A. N., Denney, D. W. & Parham, P. (1987) *J. Immunol.* **139**, 936–941.
18. Folger, K. R., Wong, E. A., Wahl, G. & Capecchi, M. R. (1982) *Mol. Cell. Biol.* **2**, 1372–1387.
19. Fukumoto, T., McMaster, W. R. & Williams, A. F. (1982) *Eur. J. Immunol.* **12**, 237–243.
20. Hart, D. N. J. & Fabre, J. W. (1981) *Transplant. Proc.* **13**, 1329–1332.
21. Diamond, A. G., Larkins, A. P., Wright, B., Ellis, S., Butcher, G. W. & Howard, J. C. (1984) *Eur. J. Immunol.* **14**, 405–412.
22. Livingstone, A. M., Powis, S. J., Diamond, A. G., Butcher, G. W. & Howard, J. C. (1989) *J. Exp. Med.* **170**, 777–795.
23. Koch, S., Koch, N., Robinson, P. & Haemmerling, G. (1983) *Transplantation* **36**, 177–180.
24. Geliebter, J., Zeff, R. A., Melvold, R. W. & Nathenson, S. G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3371–3375.
25. Kress, M., Barra, Y., Seidman, J. G., Khoury, G. & Jay, G. (1984) *Science* **226**, 974–977.
26. Mauxion, F., Sobczak, J. & Kress, M. (1989) *Immunogenetics* **29**, 397–401.
27. Seeman, G. H. A., Rein, R. S., Brown, C. S. & Ploegh, H. L. (1986) *EMBO J.* **5**, 547–552.
28. Strachan, T. (1987) *Br. Med. Bull.* **43**, 1–14.
29. Parham, P., Lawlor, D. A., Lomen, C. E. & Ennis, P. D. (1989) *J. Immunol.* **142**, 3937–3950.
30. Hughes, A. L. & Nei, M. (1989) *Genetics* **122**, 681–686.
31. Jaulin, C., Perrin, A., Abastado, J.-P., Dumas, B., Papamatheakis, J. & Kourilsky, P. (1985) *Immunogenetics* **22**, 453–470.
32. Reynaud, C. A., Anquez, V., Grimal, H. & Weill, J.-C. (1987) *Cell* **48**, 379–388.