Structural and Functional Analysis of Three D/L-like Class I Molecules from *H-2ᵛ*: Indications of an Ancestral Family of *D/L* Genes

By Zeling Cai and Larry R. Pease

From the Department of Immunology, Mayo Clinic, Rochester, Minnesota 55905

Summary

Three cDNA with D region gene features have been identified from the *H-2ᵛ* haplotype. Provisionally, the sequences have been designated as D/L^v1, D/L^v2, and D/L^v3. The coding segments for the antigen binding domain (ABD) of all three D/L^v genes were engineered into a class I genomic expression vector and expressed in L cells. FACS® analysis of the three D/L^v-L^d gene transfectants revealed that the D/L^v1 molecules were recognized by both monoclonal antibodies (mAbs) 141 and 142, and the D/L^v2 molecules were recognized by mAb 143. In addition to the D/L^v1 molecules, the mAb 141 also recognized the D/L^v3 molecules. Both the D/L^v1-L^d and D/L^v2-L^d transfectants were killed efficiently by *H-2Dᵖ* region–specific alloreactive CTL. The D/L^v3 gene is the first identified D region gene other than D and L that is transcribed abundantly in spleen and the D/L^v3 RNA is present as two alternatively spliced forms. Structural analysis of the D/L^v3 hybrid molecules showed that it was susceptible to proteolysis and thermolabile at 37°C, suggesting D/L^v3 is a transcribed pseudogene. A parsimony tree analysis of three D/L^v sequences with a set of class I gene sequences revealed that the *H-2ᵛ* sequences clustered with D region genes. The presence of a third gene with D/L-like features in *H-2ᵛ*, yet structurally different from the known D/L alleles, raises the possibility that the current D/L genes evolved from a family of D/L-like genes, some of which are no longer represented among many of the mouse major histocompatibility complex haplotypes. The observation that D region alleles cluster into subgroups suggests that the alleles are not all related to each other by linear descent through a single locus. We propose that current alleles are derived from more than one ancestral locus in a manner similar to the origin of the γ2a immunoglobulin constant region alleles.

The molecules encoded by the K and D/L genes are highly polymorphic (1, 2) and function as presenters of endogenous or viral peptide antigens to class I–restricted T cells (3). In addition to the extensive polymorphism that exists among the K and D/L class I genes, the number of the D region class I genes varies among the $H-2$ haplotypes (4, 5). Using serological and peptide mapping techniques, heterogeneity in the number of expressed D/L class I genes has been demonstrated in both inbred laboratory and wild-derived mouse strains (6, 7). Molecular approaches have been used to dissect the D region heterogeneity in four haplotypes, $H-2^{b,d,k,q}$ (4–6, 8–10). Only one class I gene has been found in the D regions of the $H-2^b$ (C57BL/10) and the $H-2^k$ haplotypes (AKR and C3H), whereas five class I genes (D, $D2$, $D3$, $D4$, and L) have been identified in the D regions of the $H-2^d$ and $H-2^q$ haplotypes. However, only two D region products, D and L, have been shown to be expressed on the cell surface in $H-2^d$ and $H-2^q$.

An evolutionary model of the D region has been proposed by Rubocki et al. (9). According to this model, the primor-dial D region contained a single locus. The D region of the $H-2^b$ haplotype represents this prototype, containing a single class I gene (D^b). A D region containing two class I genes, $H-2D$ and L, appears to have arisen by duplication of an ancestral gene such as D^b, and evolved subsequently through interchanging the duplicated ancestral sequences with alleles from different haplotypes.

The structural characterization of the D regions of other haplotypes is incomplete. It is not known whether other D region structures differ from the $H-2^b$ and the $H-2^d$ prototypes. DNA hybridization with probes derived from the $D3^d$ gene and the 5' flanking sequence of L^d indicated that $H-2^v$ has a similar D region structure to $H-2^d$ and $H-2^q$ (9, 11). However, RFLP analysis with a probe from the 3' region of the L^d gene revealed differences between $H-2^v$ and $H-2^{k,q}$ (9). Genomic Southern blot analysis with an oligo probe specific for the SINE 2 insertion in the 3' untranslated region of all known D/L alleles also suggested the presence of multiple D/L loci in the the $H-2^v$ haplotype (12).

In this paper we describe the isolation of three cDNA

clones representing D/L-like genes from B10.SM mice (H-2ᵛ) using a locus-specific PCR approach. Structural and functional analysis of the three cDNA clones revealed that at least two D/L gene products in the H-2ᵛ haplotype are expressed on the cell surface and can be recognized by D/Lᵛ specific mAbs and CTLs. Our results indicate that the D region structure of H-2ᵛ differs from the H-2ᵇ and H-2ᵈ prototypes and raise the possibility that the D and L loci are descendants from a family of D/L-like genes, which includes loci that are no longer represented among many of the mouse MHC haplotypes.

Materials and Methods

Mice, Oligonucleotides, and mAbs. Inbred mouse strains B10.SM/ SgDV and B10.SM22R/SnDV were bred and maintained in the immunogenetics mouse colony of Dr. Chella S. David at the Mayo Clinic. C3H/HeJ mice were obtained from The Jackson Laboratory (Bar Harbor, ME). Oligonucleotides were synthesized on a DNA synthesizer (380A; Applied Biosystems, Foster City, CA) using phosphoramidite chemistry and purified with sepharose G25 columns (Pharmacia LKB Biotechnology, Inc., Piscataway, NJ). The oligos used in this study are listed in Table I. Hybridoma cell lines expressing mAb 28-14-8, recognizing the α3 domain of Lᵈ and mAb 16-1-11N specific for Kᵏ, were obtained from American Type Culture Collection, Rockville, MD. The D/Lᵛ-specific mAbs, 141, 142, and 143, were described previously (13).

Isolation of RNA and PCR Amplification of D/L-specific cDNA. Total cellular RNA was isolated from B10.SM mouse spleen by guanadinium isothiocyanate extraction followed by CsCl purification (14). 20 μg of total cellular RNA was used to synthesize a single-stranded cDNA with oligo dT primers. A pair of locus-specific primers (334 and 335) were used for PCR amplification. Thermal cycling was done in a thermal cycler (Perkin-Elmer Corp., Norwalk, CT) for 30 cycles of 1-min denaturation at 94°C, 2-min annealing at 60°C, and 1.5-min elongation at 72°C. Details of the method were described previously (15, 16).

cDNA Cloning and Sequencing. The PCR products were gel purified and digested with BamHI restriction enzyme (International Biotechnologies, Inc., New Haven, CT) and cloned into the BamHI site of pUC 18 vector. DH5α library efficiency competent cells (Bethesda Research Laboratories, Gaithersburg, MD) were transformed with the ligation reaction and were plated onto LB plates containing ampicillin and x-gal. After overnight growth at 37°C, the white colonies were selected and probed with a ³²P-end-labeled oligo 33 (33E2.3), which recognizes a conserved region within exon 2 of all known K and D/L class I sequences. Both strands of the cloned, double-stranded PCR products were subjected to dideoxy sequencing using a set of class I-specific primers (Table 1), 5'-α-³⁵S dATP (Amersham Corp., Arlington Heights, IL) and Sequenase (sequencing kit from US Biochemical, Cleveland, OH). The sequencing reactions were loaded on a 6% polyacrylamide, 100-cm sequencing gel and run for 7 h at 60 W. The gels were dried and then exposed to X-Omat AR film for 10–48 h.

DNA Blot. PCR-amplified DNA was resolved by electrophoresis in 3% NuSieve agarose gels containing ethidium bromide at 1 μg/ml and then transferred to a nylon membrane (MSI membrane; Fisher Scientific Co., Pittsburgh, PA). The blots were prehybridized for at least 3 h in 5 × SSPE, 0.2% SDS, and heat-denatured salmon sperm DNA (100 μg/ml) at the hybridization temperature. The hybridization temperature was calculated according to the formula, $T_w = [2(A+T)+4(G+C)]$ (17). Hybridi-

zations were performed overnight with ³²P-end-labeled oligo probes. After hybridization, the blots were washed in 5 × SSPE for 45 min at room temperature, followed by a wash for 1 min at the hybridization temperature. The blots were then exposed to X-Omat AR film with intensifying screens at -70°C.

DNA-mediated Transfer of Genes into Mouse L Cells. The cDNA sequences coding the antigen binding domains (ABD)[1] of the three D/Lᵛ proteins were engineered into a class I expression vector (18) using a gene splicing technique based on overlap extension (19, 20). The DNA fragments corresponding to introns 1 and 3 of the class I gene Kᵇ were individually amplified and fused to the PCR fragments encoding the ABD of three D/Lᵛ molecules. The presence of engineered restriction enzyme sites for SalI and XbaI within introns 1 and 3 allowed us to clone the fused DNA fragments containing the coding segments for the ABD of the D/Lᵛ molecules into the class I-genomic expression vector. The three D/Lᵛ hybrid genes were cotransfected with a pHSV thymidine kinase gene into mouse thymidine kinase-negative L cells using a calcium phosphate precipitation method. Transfectants expressing individual constructs were selected with HAT medium and cloned by limiting dilution (18).

Serological Analyses by Fluorocytometry. The L cell transfectants or PBL isolated through Ficoll-Paque (Pharmacia LKB Biotechnology Inc.) were incubated with specific mAbs or isotype control antibodies on ice for 30 min. After washing, the cells were incubated with the FITC-conjugated goat anti-mouse Ig (Tago, Inc., Burlingame, CA) for 30 min. Fluorescence staining was analyzed using FACS® IV (Becton Dickinson Immunocytometry Systems, Mountain View, CA).

Immunoprecipitation. Immunoprecipitation of class I molecules with specific mAbs was performed as previously described (21). The cloned L cell transfectants were intrinsically radiolabeled at 5 × 10⁶/ml with [³⁵S]methionine (50 μCi/ml, Amersham Corp.) in Met⁻ RPMI 1640 (ICN Biomedicals, Inc., Costa Mesa, CA) containing 5% FCS for 4 h at 37°C or 6 h at 27°C. The labeled cells were washed twice with Met⁻ RPMI and solubilized at 2 × 10⁷ cells/ml in 0.1 M Tris, 0.15 M NaCl, 0.5% Triton X-100 (Sigma Chemical Co.) and 1% aprotinin (Sigma Chemical Co.), pH 7.4, for 30 min on ice. The insoluble material was removed by centrifugation at 15,000 g for 30 min. The detergent extracts were precleared with 5% protein A-Sepharose 4B (Sigma Chemical Co.) for 30 min at 4°C and centrifuged at 3,000 g for 5 min to remove the protein A-Sepharose 4B. For each immunoprecipitate, a 500-μl aliquot of extract was incubated with 10 μg protein A-purified mAb (28-14-8) or 100 μl of ascites (16-1-11N) for 30 min at 4°C. The class I-mAb complex was precipitated in 5% protein A-Sepharose overnight at 4°C. The precipitate was pelleted and washed (50 vol buffer/1 vol precipitate) three times with TBS, 0.2% Triton X-100, pH 7.4; twice with 0.1 M Tris, 0.4 M NaCl, 0.5% Triton X-100, pH 8.0; and twice with 2 mM Hepes, 0.2% Triton X-100, pH 7.4. The pellet was eluted with 50 μl sample buffer (21) and the samples were analyzed by one-dimensional SDS-PAGE (12% polyacrylamide).

CTL Assays. D/Lᵛ-specific CTL lines were generated by in vitro stimulation of C3H (Kᵏ, Dᵏ) splenocytes with irradiated B10.SM22R (Kᵏ, Dᵛ) splenocytes for 5 d as described previously (22). The D/Lᵛ-specific CTL lines were maintained by restimulation once a week with irradiated B10.SM22R splenocytes in the presence of IL-2 (10 U/ml). CTL clones were generated by limiting

---

[1] *Abbreviation used in this paper: ABD, antigen-binding domain.*

dilution from the D/L^v-specific CTL lines. Cytotoxicity was assayed in duplicates in an 8-h ^{51}Cr release assay with a range of E/T ratios. L cell transfectants expressing comparable levels of D/L^v 1, 2, or 3, and L^d were labeled with ^{51}Cr and used as targets. Specific lysis was calculated with the formula: percent specific lysis = 100 × [(the mean of the experimental lysis − the mean of the spontaneous lysis)/(maximum lysis − spontaneous lysis)].

Results

Analysis of the Heterogeneity of D/L^v-specific cDNA Clones. Spleen cDNA from B10.SM mice was amplified with D/L-specific primers. The resulting PCR product was about 1.5 kb. The fragments were cloned into pUC 18. From the first ligation, 19 clones were obtained that hybridized with oligo

probes 33 and 381 (33E2.3 and 201E4.5). These probes recognize DNA sequences conserved in all known *K*, *D*, and *L* class I sequences. The plasmid DNA from seven randomly chosen clones (13.1, 15.1, 16.1, 18.1, 19.1, 24.1, and 25.1) were prepared and sequenced. Clones 13.1 and 19.1 had distinct sequences and both contained full-length cDNA. Clones 16.1, 18.1, 24.1, and 25.1 represented a third sequence. Two alternative splicing forms were found among these four clones. Exon 7 was deleted in clones 18.1 and 24.1, whereas clones 16.1 and 25.1 were missing both exons 5 and 7. One clone, 15.1 appeared to be a PCR-generated recombinant between the 13.1- and the 16.1-like sequences. We have provisionally designated the sequences as *D/L^{v}1* (13.1-like), *D/L^{v}2* (19.1-like), and *D/L^{v}3* (16.1-like).

Table 1. *The Oligos Used in This Study*

Oligo	Description	Sequence
334*	D/L 5'UT	5 ' - AGACACCCGG**GGATCCC**AG<u>ATG</u> - 3 '
335†	D/L 3'UT	5 ' - CGC**GGATCC**GAAGAGCAGTC^AGCGCTGAG - 3 '
101	LDNTJ.5§	5 ' - AATTTGAATGT<u>GGGTGGCT</u> - 3 '
33	33E2.3‖	5 ' - CTGTCGAAGCGCACGAA - 3 '
29	33E2.5	5 ' - GTTCGTGCGCTTCGACA - 3 '
19	123E3.3	5 ' - CTTCGTTCAGGGCGATG - 3 '
84	123E3.5	5 ' - CATCGCCCTGAACGAAG - 3 '
23	201E4.3	5 ' - CCAGGGCCCAGCACCTCAG - 3 '
381	201E4.5	5 ' - TGGGCCCTGGGCTTCTA - 3 '
382	272E4.3	5 ' - CCATCTCAGGGTGAG - 3 '
383	272E4.5	5 ' - CTCACCCTGAGATGG - 3 '
391	NC1.3	5 ' - TCTGGATGTCACAGGAG - 3 '
390	NC1.5	5 ' - CTCCTGTCACATCCAGA - 3 '
404	D/L^{v}3E5.3	5 ' - ACTGCAAAAACGATGAGAAG - 3 '
321	K^b I1.5¶	5 ' - AGGGAAACGGCCTCTGA - 3 '
174	D^b I3.3	5 ' - TGGGAGAGGCCATGGCT - 3 '
435	K^b I1.3	5 ' - CTGGGGGCGGCGCGCG - 3 '
437	K^b I3.5	5 ' - GTGCAGGGGCCGCGGG - 3 '
434**	D/L^v E3.3	5 ' - CCCGCGGCCCCTGCACCTGTGCGCAGCAGCGT - 3 '
436	D/L^{v}1E2.5	5 ' - CGCGCGCCGCCCCCAGGCCCACGCTCGCTGAG - 3 '
444	D/L^{v}3E2.5	5 ' - CGCGCGCCGCCCCCAGGCTCACATYCGATGCG - 3 '
445	D/L^{v}2E2.5	5 ' - CGCGCGCCGCCCCCAGGCCCACACTCCATGAG - 3 '

* The oligo 334 (D/L 5'UT), which recognizes the 5' untranslated region of the *D/L* genes, was used as the 5' primer for amplification of the *D/L^v* genes. The restriction enzyme site is shown in bold and the ATG site is underlined.
† The oligo 335 (D/L 3'UT), which recognizes the 3' untranslated region of the *D/L* genes, was used as the 3' primer for amplification of the *D/L^v* genes. The SINE 2 sequence is underlined and the ^ indicates the 11-bp deletion. The restriction enzyme site is shown in bold.
§ The oligo 101 (LDNTJ.5) recognizes the joining region between the noncoding region I (NC1) and the B2 SINE insertion in the *D/L* genes. The SINE 2 sequence is underlined.
‖ The numbers in the description of oligos correspond to the positions of the amino acid residues in the primary sequence of the mature class I glycoproteins. The 33E2 indicates that this oligo recognizes the exon 2 sequences encoding the amino acid residues beginning at the 33rd residue of the mature protein. The .3 or the .5 indicates that the oligo is complementary to the sense or antisense strand, respectively.
¶ The I1.5 indicates that the oligo recognizes the sequence in the intron 1 and is complimentary to the antisense sequence.
** The oligos 434, 436, 444, and 445 were the oligoes used to fuse the junctions of intron and coding sequences in the construction of the *D/L^v-L^d* hybrid genes. The sequences complimentary to the intron sequences of class I gene are shaded.
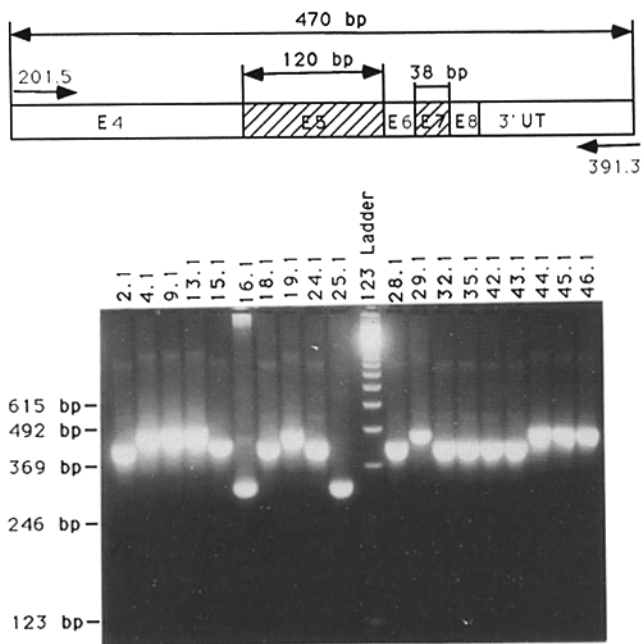
Figure 1. PCR analysis of D/Lᵛ cDNA clones. Plasmid DNA from 19 clones were amplified separately with a pair of oligo primers, 381 (201.5) and 391 (NC1.3). (*Top*) The strategy of this analysis. Exons are designated E. Shaded exons may be deleted in some clones. PCR products were separated in 3% Nusieve agarose gel. (*Bottom*) Ethidium bromide staining of the gel. Each lane represents the PCR product from each independent clone with the clone number indicated on top of each lane. A 123-bp ladder was included as a marker.

In an attempt to identify the full-length cDNA containing the $D/L^{v}3$ sequence, the original 19 clones were PCR amplified with a pair of primers (381 and 391) specific for exon 4 and 3′ untranslated region sequences of the class I K and D genes. The PCR fragments amplified from the 19

clones were resolved into three groups by agarose gel electrophoresis (Fig. 1). The longest fragment represents the PCR products from the clones with full-length cDNA, and the intermediate band represents the products from the cDNA clones without exon 7. The shortest band was derived from the clones missing both exons 5 and 7. In addition to clones 13.1 ($D/L^{v}1$) and 19.1 ($D/L^{v}2$), six clones with full-length cDNA were identified (clones 4.1, 9.1, 29.1, 44.1, 45.1, and 46.1). Sequence analysis revealed that clones 9.1 and 44.1 contained $D/L^{v}1$ sequence, and clones 4.1, 29.1, 45.1, and 46.1 $D/L^{v}2$ sequence. No full-length cDNA having $D/L^{v}3$ sequence was identified. Spleen cDNA from a second mouse was PCR amplified, cloned into pUC 18, and analyzed. 26 clones hybridized with the class I gene–specific probe 33, and also hybridized to the D/L gene–specific probe 101. Oligo probe 404, specific for an exon 5 sequence of the $D/L^{v}3$ cDNA, was synthesized and used to probe the 26 clones. 10 of 26 clones hybridized with oligo 404. All 26 clones were subjected to the same PCR analysis used to analyze the first 19 clones. None of the 404⁺ clones contained full-length cDNA and all the clones with full-length cDNA were 404⁻, suggesting that the $D/L^{v}3$ transcripts are present mainly as truncated forms, whereas the $D/L^{v}1$ and $D/L^{v}2$ transcripts are present primarily as the full-length form.

In an attempt to identify rare cDNA representing full-length $D/L^{v}3$ mRNA, or short forms of $D/L^{v}1$ and $D/L^{v}2$ mRNA, cDNA from two independent mice were analyzed using a Southern blot strategy. Spleen cDNA from two mice were amplified independently using the D/L-specific primers and then subjected to a nested PCR analysis with the same internal primers used to distinguish full-length and truncated sequences among the cDNA clones. Three major bands were obtained which corresponded to the DNA fragments observed among the three groups of cloned cDNA analyzed previously (Fig. 2 A). The gel was transferred to nylon membrane and
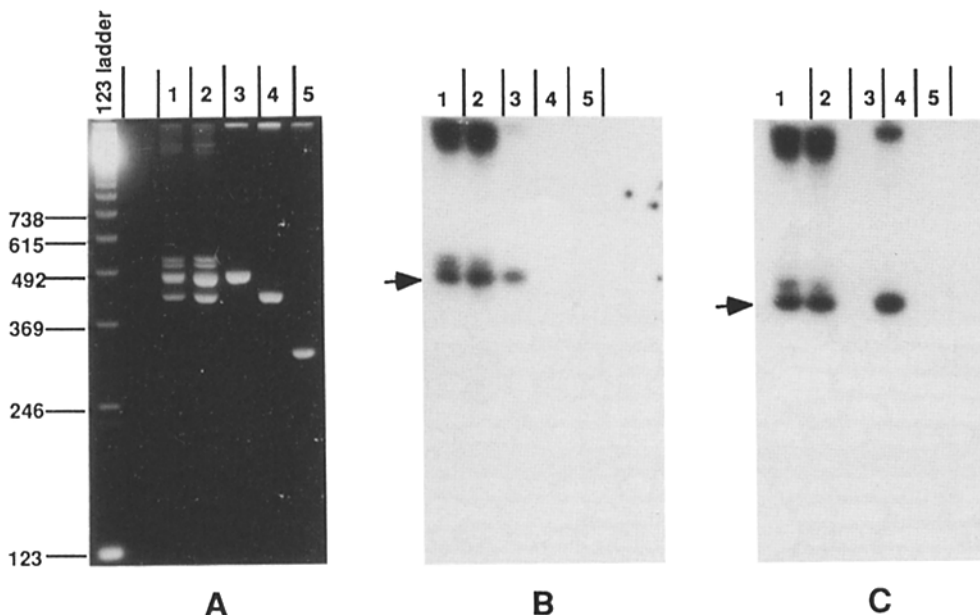
Figure 2. PCR analysis of D/Lᵛ cDNA. Spleen cDNA from two B10.SM mice were amplified using the D/L-specific primers, 334 and 335, and then subjected to PCR analysis using the same strategy illustrated in Fig. 1. Lanes *1* and *2* represent the PCR product of D/Lᵛ cDNA from two mice. Lane *3* represents a clone with full-length cDNA (D/Lᵛ1). Lane *4* represents a D/Lᵛ3 cDNA clone with exon 7 missing, and lane *5* represents a D/Lᵛ3 clone with both exons 5 and 7 missing. (*A*) Ethidium bromide staining of the gel. (*B* and *C*) Southern blot analysis of the gel with two oligo probes, 383 (272E4.5) and 404 (D/Lᵛ3E5.3), respectively.

probed with oligo 382 (272E4.5), which does not hybridize with the D/L^v3 sequence at the T_w hybridization temperature. Only the band corresponding to the full-length cDNA hybridized to this probe, whereas the bands corresponding to the D/L^v3 sequence (both intermediate and short form) did not (Fig. 2 B). The blot was stripped and reprobed with the oligo 404, which is specific for the exon 5 of D/L^v3 sequence. Only the band corresponding to the intermediate length fragment hybridized to this probe (Fig. 2 C). Analyses with both the cloned and uncloned cDNA from the $H-2^v$ haplotype revealed no detectable full-length D/L^v3 transcript. Both approaches indicated that D/L^v1 and D/L^v2 mRNA was present mainly as full-length form. D/L-specific cDNAs from other haplotypes ($H-2^{d,\ dx,\ f,\ k,\ q,\ r,\ s,\ u}$) were also analyzed and no bands corresponding to D/L^v3 sequences (intermediate and short) were identified (data not shown).

Sequence Analysis of Three D/L^v Genes. The cDNA sequences of the three D/L^v genes were determined. Two to four clones representing mRNA from each gene were sequenced completely. Clones from independent PCR reactions were analyzed in each case. In situations where two clones differed by single nucleotide substitutions, a third or fourth clone was sequenced through that region. The cDNA sequences of three D/L^v genes along with deduced amino acid sequences are shown in Fig. 3. All three D/L^v genes have classic D/L gene features: homology at the 5' ends of the genes, a short form of exon 8, and a B2 SINE sequence at the 3' end. Three potential N-linked glycosylation sites (at amino acid residues 86, 176, and 256) are present in the D/L^v1 and D/L^v2 genes, only two in D/L^v3 (amino acid residues 86 and 176). As in the D^d and D^f genes, there is a 9-bp insertion at the 3' end of exon 4 of the D/L^v1 gene. There is an unusual amino acid (Thr) at position 9 which is shared by all three D/L^v genes, but not found in any previously characterized K or D/L alleles. In comparison with D/L^v1 and D/L^v2, D/L^v3 gene is more divergent from the other D/L genes in exon 5 and the 3' untranslated region. There is a 1-bp deletion between the two ATG sites of the D/L^v3 and an unusual proline substitution at the amino acid position 169 of the mature protein.

The relationship of the three D/L^v genes to a set of mouse class I sequences from the K, D, and Q regions was assessed using a genetic tree analysis. A consensus sequence representing known HLA-A, B, and C sequences was used as an outgroup (23). The rat RT1.A sequence also was included (24). The tree was generated using a maximum parsimony method and was assembled by the Phylogenic Analysis Using Parsimony (PAUP) computer program (25). A tree, based on the entire coding sequences represented in the panel, showed all three D/L^v genes clustered with other D region sequences and separated from the Q and K genes (Fig. 4 A). Inclusion of the $D2^d$ sequence (26) in the analysis separated the D/L^v3 sequence from the D/L alleles forming a separate subgroup (Fig. 4 B). This suggests that D/L^v3 is related to other D/L sequences, but may not be a member of the allelic family of D region antigen-presenting molecules.

Expression of Three D/L^v Genes. To determine whether

all three D/L^v genes are expressed normally on the cell surface and function as alloantigens, the three D/L^v cDNA were expressed individually in L cells. The strategy selected for expressing the three D/L^v genes is illustrated in Fig. 5. The sequences representing exons 2 and 3 encoding the functional domains of the three D/L^v gene products were engineered respectively into a class I gene expression vector using the splicing by overlap extension method (19, 20). This approach was used because the class I expression vector contains exon 4 sequences encoding the $\alpha3$ domain of the L^d (18). This domain appears to fold independently from the antigen-binding domain ($\alpha1$ and $\alpha2$). Using a mAb (28-14-8) that recognizes the $\alpha3$ domain of L^d, the expression of the chimeric class I molecules could be detected on the cell surface and normalized by fluorocytometry.

The three D/L^v-L^d hybrid genes were transfected independently into L cells along with an HSV thymidine kinase gene. L cells expressing each gene were cloned. Clones with comparable surface expression of each gene product, as judged by the mAb 28-14-8, were analyzed with three $H-2D^v$ region–specific mAbs, 141, 142, and 143. The histograms from the FACS® analysis are displayed in Fig. 6. The clone expressing the D/L^v1-L^d gene was recognized by two mAbs, 141 and 142. The D/L^v2-L^d gene transfectant was recognized by mAb 143. The clone transfected with D/L^v3-L^d weakly expressed the epitope recognized by the mAb 141 and even more weakly that recognized by mAb 143. The D/L^v2-specific mAb 143 also cross-reacted with an L^d gene transfectant. All three D/L^v-L^d gene transfectants were recognized by a D/L^v-specific alloantiserum that was produced by immunizing (B10.A × A/SN)F_1 mice with B10.SM22R cells (data not shown).

Functional Analysis of Three D/L^v Genes with Alloreactive CTL. To test whether the three D/L^v molecules can function as alloantigens, $H-2D^v$ region–specific CTL lines were generated by stimulating spleen cells from C3H (K^k, D^k) mice with irradiated spleen cells from B10.SM 22R (K^k, D^v) mice in vitro. The three D/L^v-L^d gene transfectants and an L^d gene transfectant were labeled with ^{51}Cr and used as targets in a chromium release assay. L cells transfected only with a TK gene were used as a negative control. As shown in Fig. 7 A, all targets used in this study were lysed similarly by an $H-2^k$ specific CTL line (C3H.SW anti C3H). The D/L^v1-L^d gene transfectant was killed most efficiently by the D/L^v-specific CTL lines (C12491 and C2691). The D/L^v2-L^d target was killed by one of three bulk CTL lines at levels comparable with the D/L^v1-L^d targets (data not shown). This line also killed L^d-transfected targets and reacted to a much lesser degree with D/L^v3-L^d. The D/L^v3-L^d gene transfectant was always a poor target compared with the D/L^v1-L^d gene transfectant. Three CTL clones were isolated from the CTL line C12491 by limiting dilution. Two of the clones (2A5 and 2A7) were specific for the D/L^v1-L^d target and one (2B3) was specific for the D/L^v2-L^d target (Fig. 7 B).

Structural Analysis of Three Expressed D/L^v Genes. The three D/L^v-L^d hybrid molecules were immunoprecipitated

```
                                                                                                    /EXON 2-ALPHA1
       /EXON 1-LEADER PEPTIDE
D/L^V1 ATG GGG GCG ATG GCT CCG CGC ACG CTG CTC CTG CTG CTG GCG GCC GCC CTG GCC CCG ACT CAG ACC CGC GCA GGG GGC GGC CCA CAC TCG CTG AGG 90
       (M) (G) (A) (M) (A) (P) (R) (T) (L) (L) (L) (L) (L) (A) (A) (A) (L) (A) (P) (T) (Q) (T) (R) (A) (G) (G) (P) (H) (S) (L) (R) 6
D/L^V2 --- --- --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C A-- ---
                                                                                                                              (M)
D/L^V3 --- --- .-- --- A-- -T- A-- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --T --- A-- C--
              (?)         (T) (L) (S)                                                                              (S)         (M)

D/L^V1 TAT TTC ACC ACC GCC GTG TCC CGG CCC GGC CTC GGG GAG CCC CGG TTC ATC TCT GTC GGC TAC GTG GAC AAC ACG GAG TTC GTG CGC TTC 180
       (Y) (F) (T) (T) (A) (V) (S) (R) (P) (G) (L) (G) (E) (P) (R) (F) (I) (S) (V) (G) (Y) (V) (D) (N) (T) (E) (F) (V) (R) (F) 36
D/L^V2 --- --- --- --- --- --- --- --- --- --- -A- --- A-- --- --- --- --- --- --- --T --- --- --- --- --- --- --- --- --- ---
                                               (H)     (K)
D/L^V3 --- --- --- --- --- --- --- --- --- --- --G --- --- --- --- --T ATC --- --T --- A-A --- G-- --- C-- --- --- --- --- ---
                                                              (I)         (I)     (D)     (Q)

D/L^V1 GAC AGC GAC GCG GAG AAT CCG AGA GAT GAG CCG CGG GCG CGG TGG ATG GAG CAG GTG GAG CCC GAG TAT TGG GAG CGG AAC ACA CGG AGA 270
       (D) (S) (D) (A) (E) (N) (P) (R) (D) (E) (P) (R) (A) (R) (W) (M) (E) (Q) (V) (E) (P) (E) (Y) (W) (E) (R) (N) (T) (R) (R) 66
D/L^V2 --- --- --- --- --- --- --- --- T-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -TC
                                       (Y)                                                                                   (I)
D/L^V3 --- --- --- A-- --- --- --- --- TT- --- --- --- --- --- --- --- --- --- --- --- --- --A -G- --G --- --- --- --- --- ---
                   (T)                 (F)                                                     (E) (G)
                                                                                  /EXON 3-ALPHA2
D/L^V1 GCC AAG GGC CAT GAG CAG TGG TTC CGA GTG AGC CTG AGG AAC CTG CTC GGC TAC TAC AAC CAG AGC GCG GGC GGC TCT CAC ACA CTC CAG 360
       (A) (K) (G) (H) (E) (Q) (W) (F) (R) (V) (S) (L) (R) (N) (L) (L) (G) (Y) (Y) [N] [Q] [S] (A) (G) (G) (S) (H) (T) (L) (Q) 96
D/L^V2 --- --- --- -A- A-- --- --- --- A-T --- --- -A- --- --- --- -C- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
                    (D) (N)           (S)         (N)                 (T)
D/L^V3 --- --- --- --- --- --- --- --- A-T --- --- --- --- --- --- -C- --- --- --- --- --- --- AA- --- --- --- --- --- --- ---
                                       (S)                           (T)                       (K)

D/L^V1 TGG ATG TCT GGC TGT GAC GTG GGG TCG GAC GGG CGC CTC CTC CGC GGG TAC GAG CAG TAC GCA TAC GAC GGC TGC GAT TAC ATC GCC CTG 450
       (W) (M) (S) (G) (C) (D) (V) (G) (S) (D) (G) (R) (L) (L) (R) (G) (Y) (E) (Q) (Y) (A) (Y) (D) (G) (C) (D) (Y) (I) (A) (L) 126
D/L^V2 C-- T-- --C --- --- --- --- --- --- --- --- --- T-- --- --- --- --- CG- --- -C- --T --- --- --- C-- --- --- C-- --- ---
       (R) (L)                                         (F)                 (R)     (S)                 (L)
D/L^V3 CA- --- -A- --- --- --G --- --- --- --- --- --- --- --- --A --T --- --- --- --- --T --- CT- --- --- --T --- ---
       (Q)     (Y)     (E)                                                                 (L)

D/L^V1 AAC GAA GAC CTG AAA ACG TGG ACG GCG GCG GAC ATG GCG GCG CTG ATC ACC AAA CAC AAG TGG GAG CAG GCT GGT GCT GCA GAG TAT TAC 540
       (N) (E) (D) (L) (K) (T) (W) (T) (A) (A) (D) (M) (A) (A) (L) (I) (T) (K) (H) (K) (W) (E) (Q) (A) (G) (A) (A) (E) (Y) (Y) 156
D/L^V2 --- --- --- --- --- --- --- --- --- --- GA- --- --- --- CG- -G- --- --- --- --- --- -G- --- -AA --- --- AGA CT-
                                               (E)             (R) (R)                     (G)     (E)         (R) (L)
D/L^V3 --- --- --T --- --- --- --- --- --- --- C-- --C --A -A- --- --- CG- -G- --- --- --- --- --- --- --- --- AGG CGT
                                               (L)         (Q)         (R) (R)                                 (R) (R)
                                                                             /EXON 4-ALPHA3
D/L^V1 AGG GCC TAC CTG GAG GGC GAG TGC GTG GAG TGG CTC CTC AGA CAC CTG AAG AAC GGG AAT GCG ACG CTG CTG CGC ACA GAT TCC CCA AAG 630
       (R) (A) (Y) (L) (E) (G) (E) (C) (V) (E) (W) (L) (L) (R) (H) (L) (K) (N) (G) [N] [A] [T] (L) (L) (R) (T) (D) (S) (P) (K) 186
D/L^V2 --- --- --- --A --- --- ACC --- --- --- --- --- -G- --- --- --- --- --- --C ---
                      (A)     (T)                       (R)
D/L^V3 --- --- --- --- --- --G -C- --- --- --- --- --- -C- --- --- --- --- --- ---
                       (A)                             (P)


D/L^V1 GCA CAT GTG ACC CAT CAC CCC AGA TCT AAA GGT GAA GTC ACC CTG AAG TGC TGG GCC CTG GGC TTC TAC CCT GCT GAC ATC ACC CTG ACC 720
       (A) (H) (V) (T) (H) (H) (P) (R) (S) (K) (G) (E) (V) (T) (L) (K) (C) (W) (A) (L) (G) (F) (Y) (P) (A) (D) (I) (T) (L) (T) 216
D/L^V2 --C --- --- --- T-- --- --- --- --- --- --- --- --- --- --- --- -G- --- --- T-- --- --- --- --- --- --- --- -G- --- ---
                      (Y)                                               (R)                                         (S)
D/L^V3 --C --- --- G-- --- --- -T- --- --- G-- -A- A-- --- --- --- -G- --- --- --- --- --- --- --- --- --- --- --- -G- --- ---
                   (A)         (L)         (E) (D) (K)                 (R)                                           (R)

D/L^V1 TGG CAG TTG AAT GGG GAG GAG CTG ACC CAG GAC ATG GAG CTT GTG GAG ACC AGG CCT GCA GGG GAT GGA ACC TTC CAG AAG TGG GCA TCT 810
       (W) (Q) (L) (N) (G) (E) (E) (L) (T) (Q) (D) (M) (E) (L) (V) (E) (T) (R) (P) (A) (G) (D) (G) (T) (F) (Q) (K) (W) (A) (S) 246
D/L^V2 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- G--
                                                                                                                      (A)
D/L^V3 --- --- --- --- --- --- --A G-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
                              (V)
                                                                                                   /EXON 5-TM
D/L^V1 GTG GTG GTG CCT CTT GGG AAG GAG CAG AAT TAC ACA TGC CAT GTG CAC CAT GAG GGG CTG CCT GAG CCC CTC ACC CTG AGA TGG GGG AGA 900
       (V) (V) (V) (P) (L) (G) (K) (E) (Q) [N] [Y] [T] (C) (H) (V) (H) (H) (E) (G) (L) (P) (E) (P) (L) (T) (L) (R) (W) (G) (R) 276
D/L^V2 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --T --- --- --- --- --- --- --- --- --- --- ... ...
                                                                             (K)                                     (.) (.)
D/L^V3 --- --- --- --- --- --- --- --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A- --- ... ...
                                             (K)                                                         (K)       (.) (.)

D/L^V1 TGG GAG CCT CCT CCG TCC ACT GAC TCT TAC ATG GTG ATC GTT GCT GTT CTG GGT GTC CTT GGA GCT ATG GCC ATC ATT GGA GCT GTG GTG 990
       (W) (E) (P) (P) (P) (S) (T) (D) (S) (Y) (M) (V) (I) (V) (A) (V) (L) (G) (V) (L) (G) (A) (M) (A) (I) (I) (G) (A) (V) (V) 306
D/L^V2 ...
       (.)
D/L^V3 ... --- --- --- --A --- --- --- --- --- --- --- A-- -T- --- -T- --- --- --- ... C-T CT- --- G-- TTT --A --- ---
       (.)                 (I)                 (V)         (V)                 (.) (L) (L)     (V) (F)
                                                    /EXON 6-CY                        /EXON 7-CY
D/L^V1 GCT TTT GTG ATG AAG AGA AGG AGA AAC ACA GGT GGA AAA GGA GGG GAC TAT GCT CTG GCT CCA GGC TCC CAG AGC TCT GAA ATG TCT CTC 1080
       (A) (F) (V) (M) (K) (R) (R) (R) (N) (T) (G) (G) (K) (G) (G) (D) (Y) (A) (L) (A) (P) (G) (S) (Q) (S) (S) (E) (M) (S) (L) 336
D/L^V2 --- --- --- --- --- --- --- -T- -TG --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
                                    (M) (M)
D/L^V3 --- --- --- --- -G- ... --- --- --- TTG --- --- C-- --- --- --- --- -CT --- --- ... ... ... ... ... ... ... ... ...
                        (R) (.)           (L)         (Q)                 (P)         (.) (.) (.) (.) (.) (.) (.) (.) (.)
       /EXON 8-CY  /NC1
D/L^V1 CGA GAT TGT AAA GCG TGA AGACAGCTGCCTGGACTGTACTGAGTGACAGACGTGTGTTCAGGTGTCTCCTGTGACATCCAGAGCCCTCAGTTCTCTTTACACAACATTGTCTGATGTTCC 1201
       (R) (D) (C) (K) (A) (*)                                                                                                    441
D/L^V2 --- --- --- --- --- --- -------------------------------------C---------------------------C-------------------------------G---------------
D/L^V3 ... ... ... ... --- --- ----------------G--G---TG---------A----C----CACA------------------------------GT---..G-------------
       (.) (.) (.) (.)

D/L^V1 CTGTGAGCTT..GGGTTCAGTGTGAAGAACTGTGGAGCCCAG..CCTGCCCTGCACACCAGGACCCTATCCCTGCACTGCCCTGTGTTCCCTTCCATAGCCAACCTTGCTGCTCCAGCCAAACACTGGGG 1327
D/L^V2 --------------------------------------C--------------------------------------------------------------------------
D/L^V3 --------C-AT--AC---AA------------------T--AC-----------------G--------------------------------G-----G-T-------
                                     /NC2(D/L)-SINE2
D/L^V1 GACATCTGCATCCTGTAAGCTCCATGCTACCCTGAGCTGCAGCTCCTCACTTCCACACTGAGAATAATAATTTGAATGTGGGTGGCTGGAGAGA...TGG                       1424
D/L^V2 ------------------------------------------------------------------------------------------------
D/L^V3 ----------G-----C--------------------------------------T-G---------------------G-T---------TGA---
```
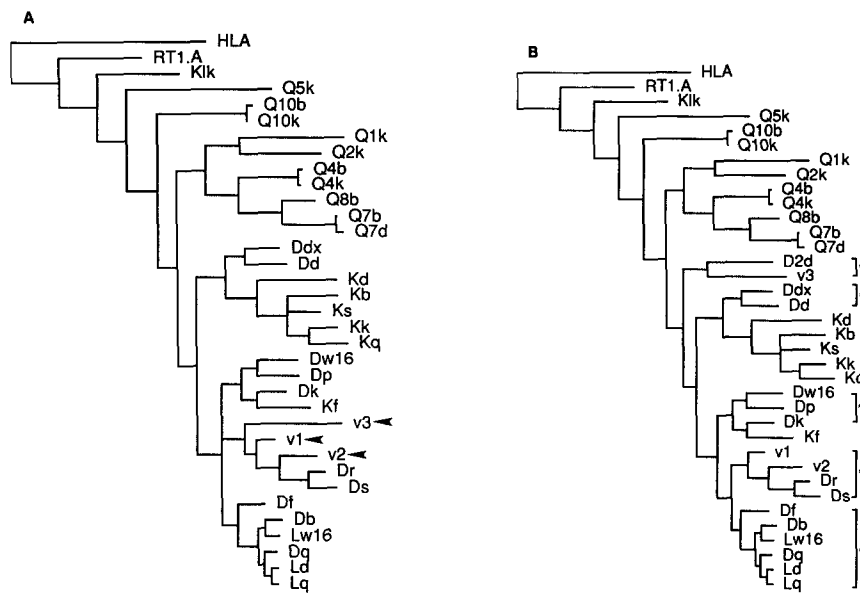
Figure 4. Genetic trees of class I sequences. The trees were generated by the PAUP computer program (25) using a maximum parsimony method. A consensus sequence representing known *HLA-A*, *-B*, and *-C* sequences was used as an outgroup (23). The tree in *A* was based on the entire coding sequences of 35 genes, while the tree in *B* was generated with the addition of *D2ᵈ*. The three *D/Lᵛ* genes are marked by arrows, and five distinct clustered *D/L* gene subgroups are indicated. The references for all the sequences are as followed: *HLA* consensus (23), *RT1.A* (24), *D2ᵈ* (26), *Dᶠ* (44), *Dʳ* and *Dˢ* (15), *Dʷ¹⁶*, and *Lʷ¹⁶* (33), *Dᵈˣ* (40), and the rest of the sequences (45).

with mAb 28-14-8, which recognizes the α3 domain of all three hybrid molecules. As shown in Fig. 8 *A*, a 45-kD heavy chain was precipitated from all three *D/Lᵛ-Lᵈ* gene transfectants with the mAb 28-14-8 (lanes *2*, *4*, and *6*, but not with an isotype control antibody (lanes *1*, *3*, and *5*). In addition to the full-length D/Lᵛ3-Lᵈ heavy chain, a truncated form (~27 kD) was precipitated from the *D/Lᵛ3-Lᵈ* gene transfectant (lane *6*).

Previously, we had engineered a hybrid *class II/class I* gene in which the sequences coding the ABD of the molecule were derived from *class II* genes, and the sequences encoding the rest of the molecule were derived from *class I* genes (27). The L cells transfected with the *class II/class I* hybrid gene only expressed the α3 domain of the class I molecule on the cell surface, whereas the ABD encoded by the class II genes was lost as judged by FACS® analysis and immunoprecipitation with mAb 28-14-8 and several class II–specific antibodies (unpublished observation). It is interesting that the size of the truncated D/Lᵛ3-Lᵈ molecule was identical to that of the truncated form of the class II/class I fusion product (Fig. 8 *B*), suggesting that the truncation of both molecules occurred at the same site, probably between the ABD and the α3 domain.

The presence of a truncated form of the D/Lᵛ3-Lᵈ molecules raised the question as to the folding properties and stability of the D/Lᵛ3 molecules. Since the mAb 141 recognized the ABD of both D/Lᵛ1-Lᵈ and D/Lᵛ3-Lᵈ molecules, and the α3 domains of the D/Lᵛ1-Lᵈ and D/Lᵛ3-Lᵈ mole-

cules were derived from Lᵈ, it was convenient to analyze the ABD and the α3 domain of the two molecules separately by fluorocytometry. The potential influence of temperature on the expression of 141 epitope on the D/Lᵛ3-Lᵈ molecule was investigated. As shown in Table 2, incubation of the L cell transfectants at 27°C increased the expression of the 141 epitope on the D/Lᵛ3-Lᵈ molecule, but had little effect on the D/Lᵛ1-Lᵈ molecule. This suggests that increased expression of the 141 epitope on D/Lᵛ3-Lᵈ at 27°C could be caused by minimizing the energy required for correct folding of D/Lᵛ3-Lᵈ or by slowing down the proteolysis of the D/Lᵛ3-Lᵈ molecules.

Folding abnormalities could result from nonassociation with β_2m and peptides (28, 29) or from point mutations at critical amino acid residues of the class I heavy chain (30). However, in molecules exhibiting these properties, no truncated forms have been reported. Immunoprecipitation of the D/Lᵛ3-Lᵈ molecules revealed that the ratio of the full-length D/Lᵛ3-Lᵈ molecules to truncated molecules increased after incubation at 27°C, while the quantity of the endogenous Kᵏ molecule was not influenced by the temperature treatment of the L cells expressing the D/Lᵛ3-Lᵈ gene (Fig. 8 *B*). These results support the hypothesis that the chimeric D/Lᵛ3-Lᵈ molecule is not stable at 37°C and is, therefore, more susceptible to cleavage by proteases.

Lᵈ molecules are poorly associated with β_2m and it has been suggested that the α3 domain of Lᵈ is responsible for the poor association (31). In spite of the α3 domain being

---

Figure 3. The cDNA sequences and translated amino acid sequences of three *D/Lᵛ* genes. Amino acid positions are numbered beginning with the first amino acid in the mature protein and nucleotide acid positions are numbered beginning with the first ATG. Identity among the sequences is indicated by dashes, and deletions are indicated by dots. *The stop codon. Exons and encoded domains as well as the 3' untranslated regions are indicated. The two potential translational initiation sites are in bold and are underlined. The unique residues of each gene are underlined. The possible N-linked glycosylation sites are shaded. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers M69067 (*D/Lᵛ1*), M69068 (*D/Lᵛ2*), and M69069 (*D/Lᵛ3*).

589   Cai and Pease

Class I expression vector

E1 —a— Sal I | Exon2 — Exon3 —e— Xba I — Exon4 — ☐ — ☐☐ ☐
    —b—   —f—

Rx.A   Rx.C

PCR

c (1,2,3)
Exon2|Exon3 | | | | | 3UT   cDNA
—d—   (D / L^v 1,2,3)

Rx.B

↓

Fusion   a Sal I ~~Exon2|Exon3~~ Xba I
(SOEING) —————   ————— f
              A   B (1,2,3)   C

↓

Cloning   Sal I Exon2|Exon3 Xba I
              —+——————————————+—

                              L^d (28-14-8)
                              ↓

E1 — Sal I | Exon2|Exon3 —Xba I— Exon4 — ☐ — ☐☐☐
     —+——————————————+—

              D/L^v 1,2,3 - L^d
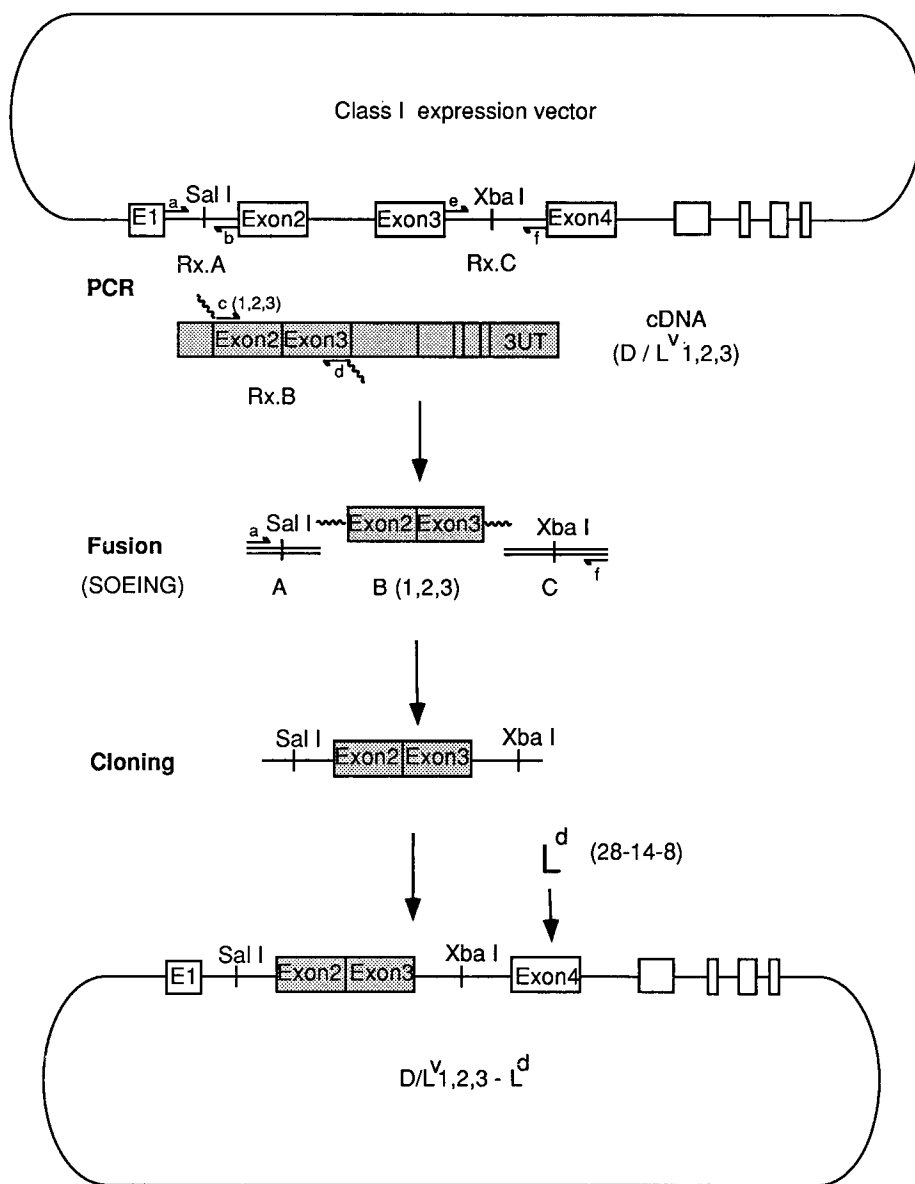
Figure 5. Strategy for expression of the ABD of the three D/L^v genes. (*PCR*) In reaction A (*Rx.A*), primer a (321) and primer b (435) were used to amplify the intron 1 sequences from a class I genomic expression vector (18). The intron 3 sequences were amplified by primers e (437) and f (174) in reaction C (*Rx.C*). In reaction (*Rx.B*), the cDNA sequences encoding ABD of the three D/L^v genes were PCR amplified with primers c (436, 444, or 445) and d (434). The 5' end of primers c (436, 444, and 445) are complementary to the intron 1 sequences of class I gene and the 5' end of primer d is complementary to the intron 3 sequence of class I gene (Table 1). (*Fusion*) Fragments (*B1–3*) containing exons 2 and 3 of $D/L^v1–3$ were separately fused to fragment A (intron 1) and fragment C (intron 3). (*Cloning*) The fused PCR products containing the exons of the three D/L^v genes were digested with the restriction enzymes SalI and XbaI, and then were cloned into the class I expression vector as indicated.

derived from L^d, the hybrid D/L^v1-L^d and D/L^v2-L^d molecules associated better with β_2m than did the L^d molecule. The full-length D/L^v3-L^d molecule appears to associate with β_2m. These results confirmed that the ABD of the L^d molecule was responsible for the poor association of L^d with β_2m as suggested previously (32).

Discussion

Three D/L-like cDNA from the H-2^v haplotype were identified using a D locus–specific approach. The presence of three D/L^v genes is consistent with early RFLP studies, which suggested the presence of multiple D region genes in the H-2^v haplotype (9, 11). All three D/L^v cDNA have characteristics of D/L genes: D-like 5' untranslated regions; a short form of exon 8 caused by alternative splicing; and a

SINE 2 insertion in the 3' untranslated region. Maximum parsimony tree analysis revealed that the three D/L^v genes were clustered among the D region sequences in a panel of class I sequences. The three D/L^v gene transfectants were recognized by the D/L^v-specific mAbs, and the D/L^v1 and D/L^v2 were recognized by D/L^v-specific alloreactive CTL. Taken together, we conclude that the three D/L^v genes are D region genes. Traditionally, the assignment of D or L loci has been based on the gene location on the chromosome 17. Without chromosome walking, it is impossible to assign the three genes to the D or L loci, although it was proposed previously (13) that the mAbs 141 and 142 recognized D^v (D/L^v1), and that mAb 143 recognized L^v (D/L^v2).

Two prototypes of D region structure have been identified among the inbred mouse strains with standard haplotypes (4, 5, 9). Only one gene is identified in the D region of the
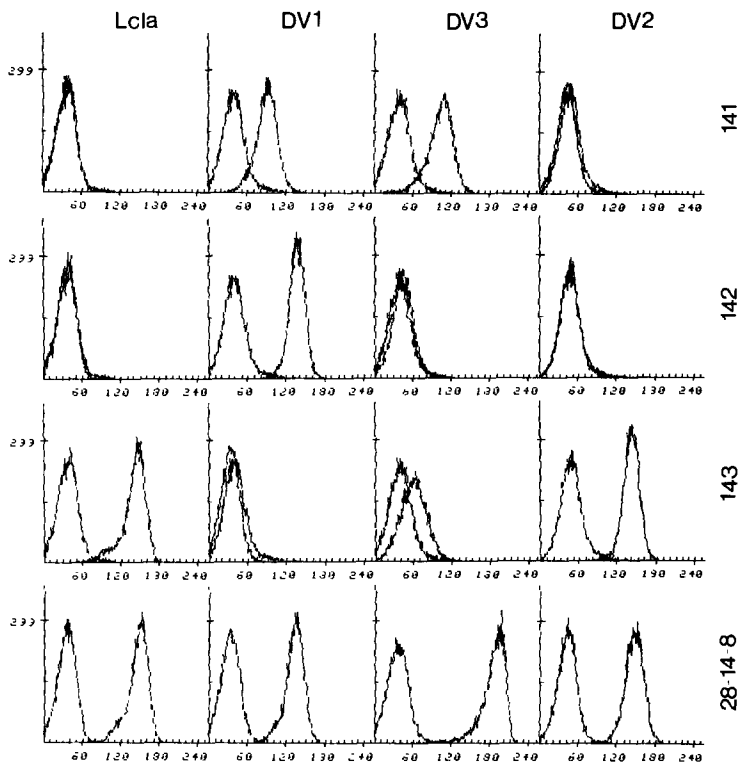
Figure 6. FACS® analysis of the three D/L^v-L^d genes in L cell transfectants. The x-axis represents the relative fluorescence intensity and the y-axis represents the relative cell numbers. Lc1a is an L cell clone that expresses the L^d molecule. DV1, DV2, and DV3 represent the L cell clones that express the D/L^v1-L^d, D/L^v2-L^d and D/L^v3-L^d chimeric genes respectively. The mAb 28-14-8 recognizes the $\alpha3$ domain of L^d. Mouse IgM was used as an isotype control for mAbs 141, 142, and 143, whereas mouse IgG2a was used as an isotype control for mAb 28-14-8. FITC-conjugated goat anti–mouse Ig was used as secondary antibody.

H-2^b prototype, whereas five D region genes are presented in the H-2^d prototype, with products from two of the genes expressed on the cell surface. The H-2^q haplotype has a similar D region structure to the H-2^d prototype (9). It has been postulated that the D and L genes in H-2^q arose by a recent gene duplication of a D^b-like ancestral gene, and subsequently an unequal crossover between the duplicated D^b-like genes (D^q and L^q) and the ancestral gene of D^d produced the D region of H-2^d (9). The presence of a family of highly homologous D/L genes (D^b, D^q, L^q, and L^d) referred to as the L^d-like subfamily is consistent with this hypothesis. A recent study of D^{w16} region gives another example where an L^d-like gene is also present in a wild haplotype with multiple D region genes (33). H-2^v represents a fourth haplotype with a complex D region. While sequence analysis of the three D/L^v genes showed that none were members of the L^d-like subgroup, the 3' coding sequences (exon-4–8) of the D/L^v1 gene differs from the members of the L^d subgroup by only 3 bp.

The D/L^v3 gene is the first identified D region gene other than D and L that is transcribed abundantly in the spleen. The D/L^v3 RNA is present as two alternatively spliced forms. Virtually all the D/L^v3 RNA in the spleen is missing exon 7 sequences. A minor component is missing both exon 5 and 7 sequences. Alternative splicing of class I RNA occurs in a number of instances in humans and mice (34–36). The functional significance of these observations, if any, is still unknown. It has been demonstrated that alternative splicing of exon 7 of D^d resulted in membrane expression of a H-2D^d protein lacking a major site of in vivo phosphorylation in

resting and PMA-stimulated spleen cells (34). At least one HLA class I allele, $A24$, produces an alternatively spliced form missing exon 5 sequences, which leads to the production of a secreted protein bound to β_2m (36). In both circumstances, the alternatively spliced forms occur along with the predominant production of the normally spliced mRNA, and both of these genes encode antigen-presenting molecules. The D/L^v3 gene predominantly produces two alternatively spliced forms with no detectable full-length RNA.

D/L genes have two potential ATG translational initiation sites, whereas K locus sequences have only one. The 3' ATG site in the D/L genes corresponds to the ATG site in K alleles. The 5' ATG site found in the D/L genes is 6 bp upstream of this consensus ATG site. According to the scanning model for translation, the first AUG role holds for 90–95% of 699 vertebrate mRNA sequences that have been analyzed (37, 38). However, the context of the AUG also affects the initiation. The sequence GCCGCC$_G^A$CCAUGG has emerged as the consensus sequence for initiation in higher eukaryotes. A purine in position −3 is the most highly conserved nucleotide in all eukaryotic mRNAs, and mutation in this position has more profound effects on translation than a point mutation anywhere else (37). The G+4 is also essential for efficient translation. Sequence analysis of K and D/L mRNA showed that the 3' AUG site shared by all the K and D/L mRNA lies in a more favorable context than does the 5' AUG site present in the D/L mRNA. However, it is still not known which AUG site in the D/L transcripts is used. The mRNAs of D/L^v1 and D/L^v2 genes have six nucleotides between the two AUG codons, therefore initia-
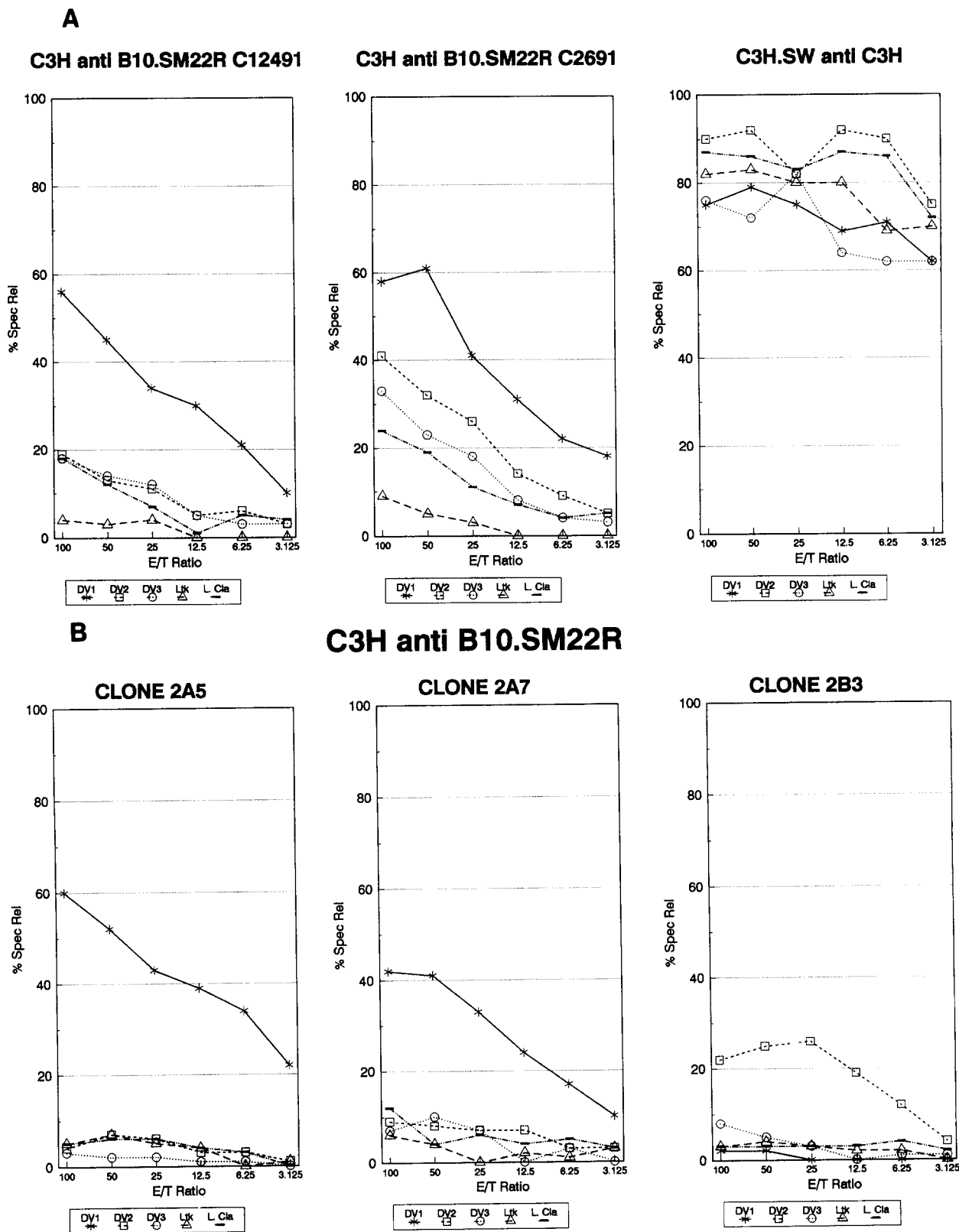
Figure 7. CTL assays. The x-axis represents different E/T ratios and the y-axis represents the percentage of specific ^{51}Cr release. CTL targets are indicated at the bottom of each panel. DV1, DV2, and DV3 represents L cell clones that express comparable levels of D/L^v1-L^d, D/L^v2-L^d and D/L^v3-L^d. Ltk is an L cell clone that was transfected with the TK gene only and Lcla is an L cell clone that expresses L^d. Fig. 7 A represents a CTL assay using CTL lines. C12491 and C2691 are two H-$2D^v$ region specific CTL lines that were generated independently. C3H.SW anti-C3H is an H-2^k specific CTL line, which was used as a control. Fig. 7 B represents a CTL assay using CTL clones. Clones 2A5, 2A7, and 2B3 were generated by limiting dilution.
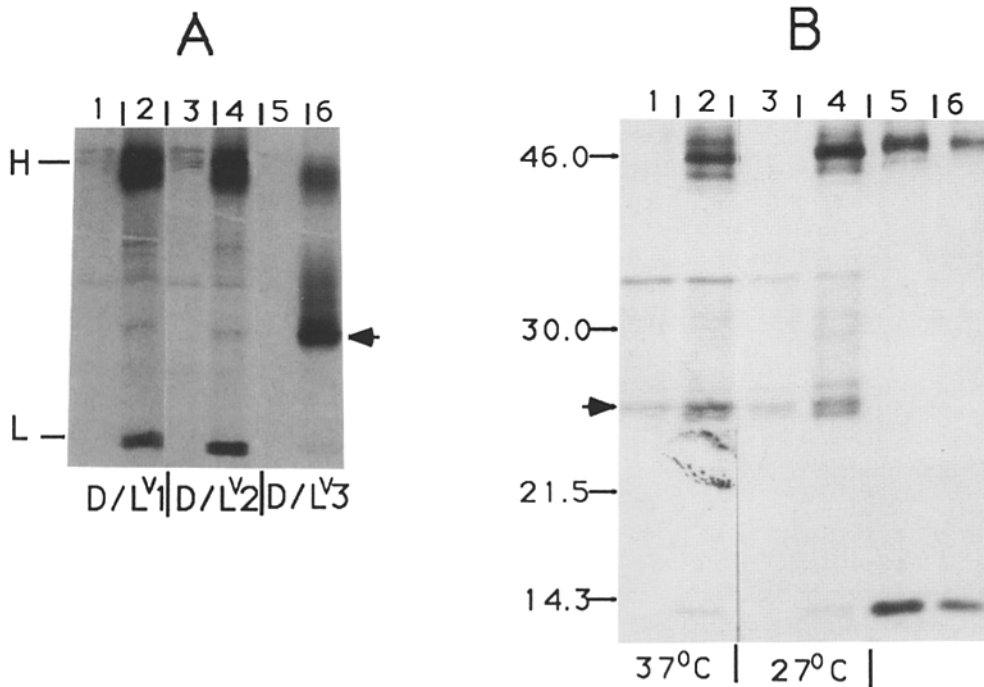
Figure 8. SDS-polyacrylamide gel analysis of immunoprecipitates of the three D/L^v-L^d hybrid molecules. ^35S-labeled L cell transfectants were lysed and immunoprecipitated with mAb. The precipitates were analyzed in a 12% polyacrylamide gel. (*A*) Samples of lanes *1*, *3*, and *5* were immunoprecipitated with an isotype control antibody, and samples for lanes *2*, *4*, and *6* were immunoprecipitated with mAb 28-14-8. (*B*) Samples in lanes *1*, *3*, and *5* were from a *class II/class I* hybrid gene transfectant (27), and the samples in lanes *2*, *4*, and *6* were from an D/L^v3 transfectant. The samples in lanes *1–4* were immunoprecipitated with mAb 28-14-8 (anti α3-L^d), and the samples in lanes *5* and *6* were immunoprecipitated with mAb 16-1-11N (anti K^k). Arrows indicate the position of common truncated fragments.

tion at either site would lead to the formation of mature protein. However, there is a 1-bp deletion between the two AUG sites in the D/L^v3 mRNA. Initiation at the first AUG site in the D/L^v3 transcript would lead to the termination of the peptide chain, whereas initiation at the second AUG site would produce the mature protein.

To study the functional properties of the putative D/L^v3 antigen-presenting molecule, we circumvented the potential problems associated with a frameshift mutation between the two ATG sites by expressing the sequences corresponding to exons 2 and 3 of the D/L^v3 gene in a class I expression vector. Structural analysis of the D/L^v3-L^d hybrid molecule

Table 2. *FACS® Analysis of Antigen Binding*

Exp.	Temperature	D/L^v1-L^d(ABD/α3)	D/L^v3-L^d(ABD/α3)
	°C		
1	37	0.81	0.54
2	37	0.76	0.56
	27*	0.85	0.75
3	37	0.74	0.51
	27	0.77	0.72

The cloned L cells expressing D/L^v1-L^d and D/L^v3-L^d were stained with mAbs 141 and 28-14-8 separately. The ratio of fluorescence peak channel for mAbs 141 vs. 28-14-8 was calculated as the ratio of ABD/α3 for each gene transfectant.
* The L cell transfectants were cultured in a CO_2 incubator at 27°C overnight and then subjected to FACS® analysis.

revealed that it was thermolabile at 37°C and was susceptible to the proteolysis. These structural defects of the ABD of D/L^v3, in conjunction with the unusual patterns of alternate splicing and a potential translational initiation problem, strongly suggest that *D/L^v3* is a transcribed pseudogene.

The presence of a third gene in *H-2^v*, with *D/L* gene features, raises questions about the evolution of *D* region genes and the mechanisms by which the heterogeneity of *D* regions was generated. The current view of the evolution of the *D* region in the mouse is that a single class I locus encoding an antigen-presenting molecule has given rise to the set of alleles currently present in mouse populations. The observation that all known members of the *D* region sequences encoding the classical antigen-presenting molecules contain specific molecular characteristics marking their common descent is in agreement with this view. The complexity observed in the number and organization of class I loci among the various haplotypes represented in captive mice has been attributed to an unequal recombination event(s) between two chromosomes. The recombination resulted in the duplication of the *D* locus along with a telomeric flanking region that included a few genes (*D2*, *D3*, and *D4*) from the *Q* region (5). The fact that the *D2*, *D3*, and *D4* genes are not closely related to known *Q* region sequences has not been addressed adequately. One possibility is that this discordance is related to heterogeneity among the loci at the centromeric (*D*-proximal) end of the *Q* gene family.

We propose an alternative view of the evolution of the *D* region genes. The central feature of this model is that the current family of *D* and *L* alleles that encode the *D* region antigen-presenting molecules may not be related to each other by continuous linear descent through a single locus. The ob-

servation that other loci (e.g., $D/L^{v}3$) share D/L-associated molecular markers such as the SINE 2 insertion in the 3' untranslated region, and have related sequences as judged from genetic tree analyses, suggests that the D/L sequences are derived from an ancestral family of sequences. According to this view, the ancestral gene was marked by the SINE 2 insertion sometime before the divergence of rats and mice, probably more than 10 million years ago (39). This sequence was duplicated, possibly more than one time, giving rise to several loci, some of which (e.g., $D/L^{v}3$) have been separated from the current D/L alleles for a very long time. Other duplication events have been much more recent as in the case of the H-$2^{d,q}$ haplotypes. Contraction events as evident in the $dm1$ and $dm2$ deletion mutations also play a continuing role in shaping the linear array of genes that comprise the various regions of the MHC and have most likely contributed to the shaping of the current D regions. We have argued previously that the D and L loci do not encode genetically separated allelic series (2). The intermixing of sequences from these two highly similar loci is the consequence of recombination events between chromosomes encoding two loci and chromosomes encoding one. The blending of alleles at the D and L loci provides insight into possible consequences after the contraction of a family of related sequences within a multigene family.

We propose that some of the current alleles of the D/L series are historically pseudoalleles that have collapsed in more recent times into the single or double locus systems represented in modern day haplotypes. The hypothesis is supported by the fact that the D/L genes form clusters in the genetic tree shown in Fig. 4, and that members of each cluster share identical or near identical sequences in their 3' exons (exons 4–8). All previously described members of the L^{d} subgene family share identical sequences in their 3' exons. Likewise, the D^{p} and D^{w16} alleles are also identical to each other in this same region. D^{d} and D^{dx} differ from each other by a single base pair, while 4 bp differentiate D^{r} and D^{s}. Although the analysis of the entire coding sequences did not place $D/L^{v}1$ in the L^{d}-subgene family, the sequences of the 3' exons differ by only 3 bp. $D/L^{v}2$ is not closely related to any group, differing by 9 bp from D^{r} and 12 bp from L^{d}. One attractive aspect of this hypothesis is that it explains the extreme range of differences observed among alleles. The cluster of closely related L^{d}-like alleles could be derived from one locus, while the more distant D^{d} and D^{dx} alleles may have descended from a second. More than two loci may have participated in this process. The separation of individual loci from the ancestral locus that gave rise to the L^{d} subgene family may have occurred at different times in the evolutionary history of the D-region, accounting for the spectrum of differences observed among the D/L clusters in Fig. 4. $D/L^{v}3$ may have become too divergent to be integrated in the newly forming allelic series, but appears to have participated in some intergenic recombinations as evidenced by the sharing of specific sequences such as the Thr encoded for position 9 in the $\alpha 1$ domain. The evolutionary relationship between D/L genes may be better preserved in these 3' exons due to the

lack of selection for diversity in the amino acids encoded by these sequences. In contrast, selection for diversity is evident for coding sequences in the 5' region of the genes that encode the ABD (2). A series of recent recombination and gene conversion events, particularly of the 5' exon sequences, have obscured the evolutionary history of the current alleles, some representing hybrids of others (40).

A similar mechanism has been postulated to explain the divergent mouse Ig $\gamma 2a$ alleles of the BALB/c and C57BL/6 mouse strains (41). Comparison of related mouse species has provided evidence that the structurally diverse $\gamma 2a$ sequences are derived from different loci, but segregate as alleles in *Mus domesticus*. We suggest that this phenomenon is more evident within the family of MHC antigen-presenting molecules as compared with other gene systems because natural selection favors the maintenance of diversity among antigen-presenting molecules. The diverse alleles generated by the contraction of multigene families by recombination are preserved in the MHC while comparable contractions in other gene families, as in the $\gamma 2a$ gene family, may have no functional significance and persist only by chance.

One challenge to this model is that interlocus exchange, so evident among the current K and D/L alleles, would prevent recently duplicated loci from diverging from each other. However, it is clear from recent studies of the human class I homologues that such exchanges are not a necessary consequence of multigene families containing highly homologous sequences (2, 42, 43). Differences in interlocus recombination may be related to the regulation of key enzymes governing the recombinatory machinery expressed in the germlines of the two species. The expression of critical enzymes also could vary during the evolutionary history of a species, making the role of interlocus exchange more important during some periods than during others. Such a scenerio would provide for periods of genetic isolation and classical divergence of sequences from sister loci, followed by periods of intermingling of sequences by interlocus exchange and by the collapse of a multilocus complex into a single locus system by nonhomologous alignment and recombination. Such an evolutionary history would provide the appearance of an allelic series that contained individual sequences that had been diverging from each other in the absence of genetic barriers for many more million years than is actually the case.

Is the presence of multiple D region genes in mouse selected or just an evolution accident? The polygenic features of the class I genes in most species results in redundancy within the MHC system. From a functional point of view, an individual with multiple D/L genes has the potential to present an increased number of the pathogen-associated peptides. However, increased numbers of D/L molecules expressed in the thymus could lead to a decreased T cell repertoire through negative selection. Thus, a constant changing pathogenic environment and constraints on the development of the T cell repertoire could be two forces influencing the expansion and contraction of the D region class I genes. The presence of mouse haplotypes with one, two, or three D/L genes reflects compromises between increased capacity for antigen presen-

tation and the generation of holes in the T cell repertoire. Haplotypes with multiple D/L genes are presumably selected when it permits a better immune response to a prevailing pathogen. However, when the prevailing pathogen is no longer present, the corresponding haplotype will be lost because it contains excess class I genes. Occasionally, individual genes are functionally silenced and are either lost over time or persist as evolutionary vestiges by random drift. D/L^v3 may represent one such evolutionary vestige.

References

1. Klein, J., and F. Figueroa. 1986. Evolution of the major histocompatibility. *CRC Crit. Rev. Immunol.* 6:295.
2. Pease, L.R., R.M. Horton, J.K. Pullen, and Z. Cai. 1991. Structure and diversity of class I antigen presenting molecules in the mouse. *Crit. Rev. Immunol.* 11:1.
3. Townsend, A., and H. Bodmer. 1989. Antigen recognition by class I-restricted T lymphocytes. *Annu. Rev. Immunol.* 7:601.
4. Weiss, E.H., L. Golden, K. Fahrner, A.L. Mellor, J.J. Devlin, H. Bullman, H. Tiddens, H. Bud, and R.A. Flavell. 1984. Organization and evolution of the class I gene family in the major histocompatibility complex of the C57BL/10 mouse. *Nature (Lond.).* 310:650.
5. Stephan, D., H. Sun, K. Fischer Lindahl, E. Myer, G. Hämmerling, L. Hood, and M. Steinmetz. 1986. Organization and evolution of D region class I genes in the mouse major histocompatibility complex. *J. Exp. Med.* 163:1227.
6. Hansen, T.H., K. Ozato, and D.H. Sachs. 1983. Heterogeneity of H-2D region associated genes and gene products. *Adv. Immunol.* 34:39.
7. Lillehoj, E.P., N.B. Myers, D.R. Lee, T.H. Hansen, and J.E. Coligan. 1985. Structural definition of a family of L^d-like molecules distributed among four of seven haplotypes compared. *J. Immunol.* 135:1271.
8. Lee, D.R., R.J. Rubocki, W.-R. Lie, and T. Hansen. 1988. The murine MHC class I genes, *H-2D^q* and *H-2L^q* are strikingly homologous to each other, *H-2L^d*, and two genes reported to encode tumor-specific antigens. *J. Exp. Med.* 168:1719.
9. Rubocki, R.J., D.R. Lee, W.-R. Lie, N.B. Myers, and T.H. Hansen. 1990. Molecular evidence that the *H-2D* and *H-2L* genes arose by duplication. *J. Exp. Med.* 171:2043.
10. Watts, S., A.C. Davis, B. Gaut, C. Wheeler, L. Hill, and R.S. Goodenow. 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 (Eur. Mol. Biol. Organ.) J.* 8:1749.
11. Duran, L.W., R.M. Horton, C.W. Birschbach, A. Chang-

Miller, and L.R. Pease. 1989. Structural relationships among the H-2 D-regions of murine MHC haplotypes. *J. Immunol.* 142:288.
12. Duran, L.W., and L.R. Pease. 1988. Tracing the evolution of H-2 D regions genes using sequences associated with a repetitive element. *J. Immunol.* 141:295.
13. Hasenkrug, K.J., J.M. Cory, and J.H. Stimpfling. 1987. Monoclonal antibodies defining mouse tissue antigens encoded by the H-2 region. *Immunogenetics.* 25:136.
14. Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.
15. Cai, Z., and L.R. Pease. 1990. Locus-specific cDNA cloning in the class I multigene family: structure of H-2D^r and H-2D^s. *Immunogenetics.* 32:456.
16. Cai, Z., J.K. Pullen, R.M. Horton, and L.R. Pease. 1991. Specific amplification of cDNA from targeted members of multigene families. *Methods in Enzymol.* In press.
17. Suggs, S.V., T. Hirose, T. Miyake, E.H. Kawashima, M.J. Johnson, K. Itakura, and R.B. Wallace. 1981. Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences. *In* Developmental Biology Using Purified Genes. D.D. Brown and C.F. Fox, editors. Academic Press, New York. pg. 683.
18. Pullen, J.K., H.D. Hunt, R.M. Horton, and L.R. Pease. 1989. The functional significance of two amino acid polymorphisms in the antigen-presenting domain of class I MHC molecules. *J. Immunol.* 143:1674.
19. Ho, S.N., H.D. Hunt, R.M. Horton, J.K. Pullen, and L.R. Pease. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene (Amst.).* 77:51.
20. Horton, R.M., H.D. Hunt, S.N. Ho, J.K. Pullen, and L.R. Pease. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene (Amst.).* 77:61.
21. Moosic, J.P., E. Sung, A. Nilson, P.P. Jones, and D.J. McKean. 1982. The selective solubilization of different murine spleno-

cyte membrane fractions with Lubrol WX and Triton X-100 distinguishes two forms of Ia antigens. *J. Biol. Chem.* 257:9684.

22. Hunt, H.D., J.K. Pullen, R.F. Dick, J.A. Bluestone, and L.R. Pease. 1990. Structural basis of K^{bm8} alloreactivity: Amino acid substitutions on the β-pleated floor of the antigen recognition site. *J. Immunol.* 145:1456.

23. Zemmour, J., and P. Parham. 1991. HLA class I nucleotide sequences 1991. *Immunogenetics.* 33:310.

24. Kastern, W. 1985. Characterization of two class I major histocompatibility rat cDNA clones, one of which contains a premature termination codon. *Gene (Amst.).* 34:227.

25. Swofford, D.L. 1990. PAUP: phylogenetic analysis using parsimony, version 3.0. Computer program distributed by *Ill. Nat. Hist. Surv.*, Champaign, IL.

26. Hedley, M.L., S.W. Hunt, III, K.A. Brorson, J.S. Andris, L. Hood, J. Forman, and P.W. Tucker. 1989. Analysis of D2^d: a D-region class I gene. *Immunogenetics.* 29:359.

27. Horton, R.M., Z. Cai, S.N. Ho, and L.R. Pease. 1990. Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. *Biotechniques.* 8:528.

28. Townsend, A., C. Öhlén, J. Bastin, H.-G. Ljunggren, L. Foster, and K. Kärre. 1989. Association of class I major histocompatibility heavy and light chains induced by viral peptides. *Nature (Lond.).* 340:443.

29. Ljunggren, H.-G., N.J. Stam, C. Öhlén, J.J. Neefjes, P. Höglund, M.-T. Heemels, J. Bastin, T.N.M. Schumacher, A. Townsend, K. Kärre, and H.L. Ploegh. 1990. Empty MHC class I molecules come out in the cold. *Nature (Lond.).* 346:476.

30. Rubocki, R.J., J.M. Connolly, T.H. Hansen, R.W. Melvold, B.S. Kim, W.H. Hildebrand, and J. Martinko. 1991. Mutation at amino acid position 133 of H-2D^d prevents β_2m association and immune recognition but not surface expression. *J. Immunol.* 146:2352.

31. McCluskey, J., J.A. Bluestone, J.E. Coligan, W.L. Maloy, and D.H. Margulies. 1986. Serologic and T cell recognition of truncated transplantation antigens encoded by in vitro deleted class I major histocompatibility genes. *J. Immunol.* 136:1472.

32. Beck, J.C., T.H. Hansen, S.E. Cullen, and D.R. Lee. 1986. Slower processing, weaker β2M association, and lower surface expression of H-2L^d are influenced by its amino terminus. *J. Immunol.* 137:916.

33. Tummuru, U., M.K. Zubairu, and D.R. Lee. 1991. The multilocus H-2D^{w16} region represents a novel D region organization. *Immunogenetics.* In press.

34. McCluskey, J., L.F. Boyd, W.L. Maloy, J.E. Coligan, and D.H. Margulies. 1986. Alternative processing of H-2D^d pre-mRNAs results in membrane expression of differentially phosphorylated protein products. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2477.

35. Vogel, J.M., R.Y. Morse, and R.S. Goodenow. 1989. A novel H-2K splice form: predictions for other alternative H-2 splicing events. *Immunogenetics.* 29:33.

36. Krangel, M.S. 1986. Secretion of HLA-A and -B antigens via an alternative RNA splicing pathway. *J. Exp. Med.* 163:1173.

37. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229.

38. Cavener, D.R., and S.C. Ray. 1991. Eukaryotic start and stop translation sites. *Nucleic Acids. Res.* 19:3185.

39. Mauxion, F., J. Sobczek, and M. Kress. 1989. Characterization of five distinct cDNA clones encoding for class I RT1 antigens. *Immunogenetics.* 29:397.

40. Cai, Z., and L.R. Pease. 1991. An intragenic recombinant class I gene: H-2D^{dx}. *Immunogenetics.* 34:273.

41. Morgado, M.G., P. Cam, C. Gris-Liebe, P.-A. Cazenave and E. Jouvin-Marche. 1989. Further evidence that BALB/c and C57BL/6 γ2a genes originate from two distinct isotypes. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3245.

42. Pullen, J.K., R.M. Horton, Z. Cai, and L.R. Pease. 1992. Structural Diversity of the Classical *H-2* Genes: *K, D,* and *L. J. Immunol.* In press.

43. Lawlor, D.A., J. Zemmour, P.D. Ennis, and P. Parham. 1990. Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. *Annu. Rev. Immunol.* 8:23.

44. Hildebrand, W.H., R.M. Horton, L.R. Pease, and J.M. Martinko. 1991. Nucleotide sequence analysis of H-2D^f and the spontaneous in vivo H-2D^{fm2} mutation. *Mol. Immunol.* In press.

45. Kuhner, M.K., and R.S. Goodenow. 1989. DNA sequences of mouse H-2 and Qa genes. *Immunogenetics.* 30:458.