ed DNA. Since ligation of the digested fragments is not possible, these results demonstrate that λ vector sequences are not required for the expression of this gene. Furthermore, Southern blot analysis demonstrated that the L^d-transformed mouse L cells contain vector sequences, presumably flanking the L^d gene. It is not possible to directly identify the L^d gene in these transformed cells with cDNA probes because of the cross-hybridization of multiple H-2^k haplotype class I genes of the recipient mouse L cells.

The level of L^d expression in mouse L cells is comparable to that of K^k for the transformants (Fig. 1C) and approximately equal to that of L^d for BALB/c spleen cells (Fig. 1D). In contrast, the relative level of L^d expression by normal BALB/c fibroblast and spleen cells is lower than the level of K^d expression as determined by fluorescent labeling and analysis by the fluorescence-activated cell sorter (*12*). Therefore, relative to K molecules, the L^d gene products are expressed at a higher level on the transformed cells than on normal fibroblasts. Transformants expressing L^d molecules at an even higher level than that of 8-5 cells have been isolated, an indication that expression of the transferred L^d gene at a higher level than for K^k molecules is not a lethal event. Moreover, the expression of L^d molecules by the cloned 8-5 transformants is a stable phenomenon. The initial level of L^d expression by these cells has remained constant throughout the maintenance of the cell line in culture for over 49 passages.

To compare the charge and size distributions of the L^d polypeptides synthesized in transformed and normal cells, we analyzed L^d molecules immunoprecipitated from BALB/c spleen and 8-5 cells by two-dimensional gel electrophoresis (Fig. 2). The size and charge heterogeneity of the L^d antigens appear virtually identical for polypeptides isolated from BALB/c lymphocytes (Fig. 2A) and the 8-5 cloned transformed L cells (Fig. 2B). In general the two-dimensional gel patterns of these antigens are characteristic of transplantation antigens isolated by immunoprecipitation (*13*). Mouse Ltk^+ transformants analyzed in a similar manner do not express L^d polypeptides (Fig. 2C). These results demonstrate that the 8-5 cells express L^d antigens that are virtually indistinguishable from L^d antigens of normal spleen cells. The tryptic peptides of these proteins are presently being analyzed by high-performance liquid chromatography to confirm these structural similarities.

Gene transfer and immunoselection provide a powerful approach for identifying and characterizing the genes encoding transplantation antigens. The correlation between gene 27.5 and its serologically defined L^d gene products made on the basis of transformation has been substantiated by the DNA sequence of this cloned L^d gene (*10*). DNA-mediated gene transfer thus provides a general approach to both the phenotypic identification of individual members of a multigene family and the analysis of the functions of cell-surface recognition molecules.

ROBERT S. GOODENOW
MINNIE MCMILLAN
ANDERS ÖRN
Division of Biology, California Institute of Technology, Pasadena 91125
MARGERY NICOLSON
NORMAN DAVIDSON
Division of Chemistry and Chemical Engineering, California Institute of Technology
JEFFREY A. FRELINGER
Department of Microbiology, University of Southern California Medical School, Los Angeles 90033
LEROY HOOD
Division of Biology, California Institute of Technology

References and Notes

1. P. A. Gorer, *J. Pathol. Bacteriol.* **44**, 691 (1937).
2. ———, S. Lyman, G. D. Snell, *Proc. R. Soc. London Ser. B* **135**, 449 (1948).
3. J. Klein, *Science* **203**, 516 (1979).
4. R. M. Zinkernagel and P. C. Doherty, *Adv. Immunol.* **27**, 51 (1980).
5. M. Steinmetz *et al.*, *Cell* **24**, 125 (1981).
6. F. Bregegere *et al.*, *Nature (London)* **292**, 78 (1981).
7. H. L. Ploegh, H. T. Orr, J. L. Strominger, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6081 (1980).
8. A. K. Sood, D. Pereira, S. M. Weissman, *ibid.* **78**, 616 (1981).
9. M. Steinmetz, K. W. Moore, J. G. Frelinger, B. Taylor Sher, F. W. Shen, E. A. Boyse, L. Hood, *Cell* **25**, 683 (1981).
10. K. W. Moore, B. Taylor Sher, Y. H. Sun, K. A. Eakle, L. Hood, *Science* **215**, 679 (1982).
11. M. Wigler, A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, L. Chasin, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1973 (1979).
12. R. S. Goodenow, unpublished data.
13. T. Krakauer *et al.*, *J. Immunol.* **124**, 2149 (1980).
14. K. Ozato, T. H. Hansen, D. H. Sachs, *ibid.* **125**, 2473 (1980).
15. D. H. Sachs *et al.*, in preparation.
16. V. T. Oi, P. P. Jones, J. W. Goding, L. A. Herzenberg, L. A. Herzenberg, *Curr. Top. Microbiol. Immunol.* **7**, 115 (1978).
17. M. McMillan, J. A. Frelinger, P. P. Jones, D. B. Murphy, H. O. McDevitt, L. Hood, *J. Exp. Med.* **153**, 936 (1981).
18. S. E. Cullen and B. D. Schwartz, *J. Immunol.* **117**, 136 (1976).
19. P. H. O'Farrell, *J. Biol. Chem.* **250**, 4007 (1975).
20. D. Z. O'Farrell, H. M. Goodman, P. H. O'Farrell, *Cell* **12**, 1133 (1977).
21. Supported by NIH grants GM 06965, CA 22662, CA 26199, and CA 25911. R.S.G. is a senior Lievre fellow of the California Division of the American Cancer Society. J.A.F. is the recipient of an American Cancer Society faculty research award. We thank D. Sachs and T. Hansen for supplying the monoclonal hybridoma antibodies used in these studies.

5 November 1981; revised 16 December 1981

DNA Sequence of a Gene Encoding a BALB/c Mouse L^d Transplantation Antigen

Abstract. *The sequence of a gene, denoted 27.5, encoding a transplantation antigen for the BALB/c mouse has been determined. Gene transfer studies and comparison of the translated sequence with the partial amino acid sequence of the L^d transplantation antigen establish that gene 27.5 encodes an L^d polypeptide. A comparison of the gene 27.5 sequence with several complementary DNA sequences suggests that the BALB/c mouse may contain a number of closely related L-like genes. Gene 27.5 has eight exons that correlate with the structural domains of the transplantation antigen.*

The transplantation or class I antigens of the major histocompatibility complex (MHC) are present on the surface of all mammalian somatic cells and play a key role in T cell immunosurveillance (*1, 2*). In the inbred BALB/c mouse, three genes for transplantation antigens, D, L, and R, are closely linked and these are separated from a fourth class I gene, K, by approximately 0.3 centimorgan (*3*). We cloned 30 to 40 genes of the BALB/c mouse (H-2^d haplotype) that are homologous to complementary DNA (cDNA) probes for transplantation antigens (*4*) and have demonstrated by gene transfer experiments that one of these clones (27.5) encodes an L^d transplantation antigen (*5*). We have determined the nucleotide sequence of the L^d gene present in clone 27.5.

The genomic clone 27.5 was isolated as previously described (*4*) from an amplified library of BALB/c sperm DNA cloned in the λ vector Charon 4A. The nucleotide sequence of gene 27.5 was determined by the dideoxy sequencing technique with M13mp2 as cloning vector (*4, 6, 7*). The sequencing strategy is shown in Fig. 1. The exons of gene 27.5 were defined on the basis of homology to available amino acid sequences for transplantation antigens (*8, 9*), to several DNA sequences of cDNA's encoding transplantation antigens (*10, 11*), and to a genomic class I clone 27.1, which bears a pseudogene mapping to the Qa-2,3 re-

0036-8075/82/0205-0679$01.00/0

gion (*4*). The DNA sequence of gene 27.5 is given in Fig. 2.

Gene 27.5 has eight exons whose boundaries correspond precisely to those determined earlier for the class I pseudogene 27.1 (Fig. 1c) (*4*). Apart from introns 1 and 3, the lengths of the introns in genes 27.5 and 27.1 are similar. Exon 1 encodes the signal peptide; exons 2 to 4, the three external domains; exon 5, the transmembrane segment; and exons 6 to 8, the cytoplasmic domain. There is a striking correlation between the discrete exon boundaries and the structural domains of the transplantation antigen (*4*). All of the exon-intron boundaries have the consensus upstream or downstream RNA splicing signals (*12*). Gene 27.5 appears to be a functional gene by sequence analysis in that it lacks any obvious elements that would render it a pseudogene (for example, termination codons or inappropriate reading frame shifts). This conclusion is supported by our gene transfer studies (*5*).

The translated sequence of gene 27.5 is identical to the amino acid sequence of the L^d molecule at 77 of 77 positions that can be compared (Table 1) (*13*). These comparisons include residues in the first, second, and third external domains. Although the paucity of amino acid sequence data permits us to compare only 21 percent (77 of 358 positions) of the gene and protein sequences, these comparisons support the conclusion reached in gene transfer studies that clone 27.5 contains an L^d gene (*5*).

There is one striking observation that can be made in comparing the sequences of gene 27.5 and of two cDNA clones, pH-2II (*10*) and pH-2^d-3 (*11*), whose translated sequences are identical with the available partial L^d amino acid sequence. Gene 27.5 is identical to cDNA clone pH-2^d-3 at 467 of 470 positions compared. Two of the sequence differences are in the first 60 nucleotides of the fourth exon and lead to one codon substitution. The other difference, which also leads to a codon substitution, is in the fifth exon. Likewise, the exons of gene 27.5 are identical to clone pH-2II at 514 of 520 nucleotide positions. The differences are clustered in the first 45 nucleotides of the fourth exon and lead to five codon substitutions. There are also six nucleotide differences between gene 27.5 and clone pH-2II of the 3′ untranslated region of 587 nucleotides compared. It is interesting to note that the differences in coding sequence are clustered at the beginning of the fourth exon, because this exon encodes the third external domain of the L^d polypeptide, and this domain is the most conserved region of the genes encoding transplantation antigens (*14*).

The genomic clone 27.5 was derived from BALB/c (H-2^d haplotype) sperm DNA. The cDNA clone pH-2II was derived from a BALB/c tumor cell line (*10*), and the cDNA clone pH-2^d-3 was derived from a DBA/2 (H-2^d haplotype) lymphoma line (*11*). It is unlikely that all of the differences in these class I sequences arose from cloning or sequencing artifacts. Accordingly, mice of the H-2^d haplotype may have three closely related class I genes that encode L^d-like polypeptides. The three L-like genes might arise from genetic polymorphism at a single locus within mice of the H-2^d haplotype or from duplicated genes.

The possibility that these different clones represent duplicated genes is attractive in view of historical precedent and of recent evidence. By serological analysis, the D end of the H-2 complex initially appeared to have a single D gene (*2*). Subsequently, more refined serological analyses have suggested that this region encodes three closely linked genes for transplantation antigens—D, L, and R (*3*). Cosmid clones that have been isolated from BALB/c DNA contain three L-like genes according to restriction map analyses (*14*). Gene transfer studies have demonstrated that at least one of these cosmid genes encodes an L^d polypeptide (*15*). It will be interesting to determine whether the remaining two of these putative L-like genes also are expressed as class I polypeptides reacting with the monoclonal antibodies to the L^d antigen.

A comparison of the translated sequences of pseudogene 27.1 and the L^d gene 27.5 with the amino acid sequence of the mouse K^b molecule (Table 1) shows that pseudogene 27.1 appears to have diverged significantly more from the translated L^d gene than the K^b molecule has (Table 1). The difference between 27.1 and the other class I genes might arise because the genes encoding Qa antigens diverged from those encoding the classical transplantation antigens before the divergence of the individual class I genes. Alternatively, this divergence may reflect changes accumulating in the pseudogene, which has presumably been released from selective pressure. In addition, the exons of pseudo-

Table 1. Homology of the exons of the translated L^d gene (27.5) to the corresponding regions of the translated pseudogene 27.1 (*4*) and to the K^b (*12*) and L^d (*13*) molecules. The homology of the translated portion of each exon to the corresponding portion of the K^b and L^d sequences, as well as to the translated 27.1 pseudogene exons (*4*), is shown.

Exon	Homology (%)		
	K^b	L^{d*}	27.1
Leader (exon 1)			76
First domain (exon 2)	84	100	71
Second domain (exon 3)	80	100	76
Third domain (exon 4)	88	100	89
Transmembrane domain (exon 5)	73		69
Cytoplasmic exon			
First (exon 6)	100		73
Second (exon 7)	64		69
Third (exon 8)	0		100

*Thirty-four of 90 positions were compared in exon 2, 28 of 92 in exon 3, and 15 of 92 in exon 4.

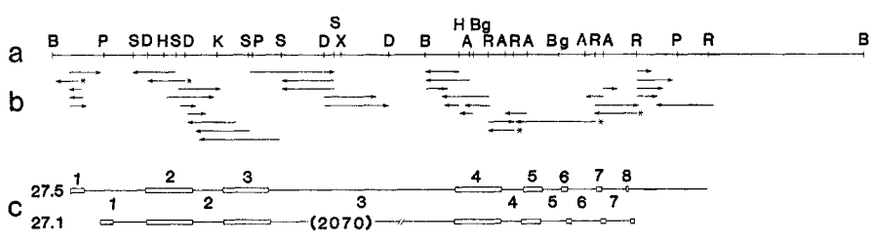

Fig. 1. The organization and sequencing strategy for gene 27.5. (a) A restriction map of clone 27.5. This map was generated by Southern blot analyses of clone 27.5 cleaved with various restriction enzymes and probed with selected M13 clones into which fragments of gene 27.5 had been inserted (see below). The restriction sites are designated as follows: *B*, Bam HI; *P*, Pst I; *S*, Sau 3A; *D*, Dde I; *K*, Kpn I; *Bg*, Bgl II; *R*, Rsa I; *H*, Hinf I; *X*, Xba I; and *A*, Alu I. (b) Sequence strategy for gene 27.5. Each arrow represents the sequence of an M13 clone. The M13 clones indicated by asterisks were used as probes in Southern blot analyses to generate the restriction map. At positions where clones do not overlap, alignment was determined by restriction mapping or by homology with cDNA clone pH-2II. (c) Organization of genes 27.5 and 27.1. Exons are represented by boxes and introns by lines. Both exons (above boxes) and introns (below lines) are numbered.

gene 27.1 are homologous to those of the L^d gene 27.5. This homology suggests a common evolutionary origin. If gene 27.1 is a Qa-2,3 pseudogene [see (*4*) for discussion], then this homology suggests that the Qa antigens are class I molecules and that the Tla complex should be considered a part of the H-2 complex. This conclusion is supported by observations that the TL and Qa antigens resemble the classical transplantation antigens in size, peptide map profiles, and their association with β_2-microglobulin (*16–19*).

We have used a computer-generated dot matrix to analyze the homology relationships of the DNA sequences from genes 27.1 and 27.5 (Fig. 3). This analysis compares every hexamer of gene 27.1 (vertical axis) against every hexamer of gene 27.5 (horizontal axis) and places a dot in the two-dimensional matrix at positions where at least five of six nucleotides are identical. Homologies are displayed as diagonal lines in the matrix; nucleotide divergences are represented as gaps in the diagonal lines; and sequence insertions or deletions offset the diagonal lines (*20*). Four important points emerge from this analysis. (i) Genes 27.1 and 27.5 are quite homologous to one another, as can be seen by the strong diagonal lines. (ii) Five of seven introns (2, 4, 5, 6, and 7) are almost as highly conserved in both genes as the exons are; the exons exhibit 87 percent homology, and these introns ex-

```
                                                                                                                 EXON 1
                                                                                           Met Ala Pro Arg Thr Leu Leu Leu Leu Leu
CCAGGGCGGATTGAGAGGGAAGACCACACACCCTGTGAGCTCACTGTGTCCAGTGAGTAGCTGCACTGGGGTCCACAGCTCACTCCAGGGATCACTCCCAGATGGGGGCG ATG GCT CCG CGC ACG CTG CTC CTG CTG CTG    140

Ala Ala Ala Trp Pro Asp Ser Asp Pro Arg G
GCG GCC GCC TGG CCC GAC TCA GAC CCG CGG G GTGAGTGSGGGTGCGGAGGGAAACGGCCTCTGCGGGGAGAGGGGCGGCAGGGGGAAGCGC ...GATCCCCTCGCCTCGCAGCCGCCCGGGGTTTGGTAGGAGGTCGG    278

                                                EXON 2
                         ly Pro His Ser Met Arg Tyr Phe Glu Thr Val Ser Arg Arg Gly Leu Gly Glu Pro Arg Tyr Ile Ser Val Gly Tyr Val Asp Asn Lys Glu Phe
GGTCTCACCGCGCGCCCGCCCCAG GC CCA CAC TCG ATG CGG TAT TTC GAG ACC GTG TCC CGG CGC GGC CTC GGG GAG CCC CGG TAC ATC TCT GTC GGC TAT GTG GAC AAC AAG GAG TTC    397

Val Arg Phe Asp Ser Asp Ala Glu Asn Pro Arg Tyr Glu Pro Arg Ala Pro Trp Met Glu Gln Glu Gly Pro Glu Tyr Trp Glu Arg Ile Thr Gln Ile Ala Lys Gly Gln Glu
GTG CGC TTC GAC AGC GAC GCG GAG AAT CCG AGA TAT GAG CCG AGG GCG CCG TGG ATG GAG CAG GAG GGG CCG GAG TAT TGG GAG CGG ATC ACG CAG ATC GCC AAG GGC CAG GAG    511

Gln Trp Phe Arg Val Asn Leu Arg Thr Leu Leu Gly Tyr Tyr Asn Gln Ser Ala Gly G
CAG TGG TTC CGA GTG AAC CTG AGG ACC CTG CTC GGC TAC TAC AAC CAG AGC GCG GGC G GTGAGTGACCCCGCGTCGGAGGTCACGGCCCCTCCCTTTCCCGACACAGGGACGCTGACTTCGTACCCAAGT    641

                                                                                                        ly Thr His Thr Leu Gln Trp Met Tyr
CCGAGGTTCGGGAACAGAACGGACCCGGAACCAGTTTCCCTTTCAGTTTGGAGGAGTCGCGGCCGGGCGGCGGGGCCGGGGCGGGTGAGCGGGGCTGAACCGGGGTACCGCAG GC ACT CAC ACA CTC CAG TGG ATG TAC    780

                                                EXON 3
Gly Cys Asp Val Gly Ser Asp Gly Arg Leu Leu Arg Gly Tyr Glu Gln Phe Ala Tyr Asp Gly Cys Asp Tyr Ile Ala Leu Asn Glu Asp Leu Lys Thr Trp Thr Phe Ala Asp
GGC TGT GAC GTG GGG TCG GAC GGG CGC CTC CTC CGC GGG TAC GAG CAG TTC GCC TAC GAC GGC TGC GAT TAC ATC GCC CTG AAC GAA GAC CTG AAA ACG TGG ACG TTC GCG GAC    894

Met Ser Ser Gln Ile Thr Arg Arg Lys Trp Glu Gln Ala Gly Ala Ala Glu Tyr Tyr Arg Ala Tyr Leu Glu Gly Glu Cys Val Glu Trp Leu His Arg Tyr Leu Lys Asn Gly
ATG TCG TCG CAG ATC ACC CGA CGC AAG TGG GAG CAG GCT GGT GCT GCA GAG TAT TAC AGG GCC TAC CTG GAG GGC GAG TGC GTG GAG TGG CTC CAC AGA TAC CTG AAG AAC GGG   1008

Asn Ala Thr Leu Leu Arg Thr A
AAT GCT ACG CTG CTG CGC ACA G GTGCAGGGGCCGCGGCGGGCAGGCTCCTCCCTCTCGCCTCGGGCTGGGGGCTCAGTCCTGGGGAAGAAGAAACCCTCAGCTGGGGTGATGCCCTGTCTCAGAGGGGAGAGAGTGACCCTG   1150

GGTTCCTGATGCCTCAGCACAGTGACTGCACTGACTCTCCCAGGGCTCAGCTTCTCCCTGGACAGTCCCAGCCTGTCTCAGGAGGGAAGGAGAGAATTTCCCTGAGGTAACAACAGCTGCTCCCTTCAGTTCCCCTGTAGCCTCTGTCAG   1300

CCATGGCCTCTCCCAGGCGGGTTCTCTGCCCACGCCACTGTCGGTGACACTGACTCCTGTCCTGGCTGATGTGTGTCAGCGCCTTACACCTCAGGACCGGAAGTCGCCTTACCTGATTGGAAACATGGACTCCTATACACTAGCCGTGTT   1450

GCCCCAGCTTCTAGAACTTTCCAGAGAATACATTCTCCCAGATCCCTCCCTGTCTGTGGGGTTTGCACCCCTTCGCACAACCTAATTCTCTCTATTCCTACAGTGGTGAATGGTCACATGAGCTCTTATGGGGTACSTGGAGGAATATAA   1600

ATCGAGAATTTTCTTTTTTGTTTTCTCTCTCTCTTTCTCTCTCTCTCTCTCTCTCTCTTTTCCGATTTTTTTGAGACCGGGTTTCTCTCTATAGCCCTSTGTCCTGACTCATTTTAGCCCTTTTCTAATTCACATCTTCTCT ...TT   1749

                                                                                                                          sp Ser Pro Lys Ala
TGGTCACTAGTGCAATGACAGTGTAGTGTCAAATAGACACATAGTTCACTCTCATCATTGATTTAACTGAGTCTTGTGTAGATTTCAGTTTGTCTTGTTAATTGTGGAATTTCTTAAATCTTCCACACAG AT TCC CCA AAG GCA   1893

                                                EXON 4
His Val Thr His His Pro Arg Ser Lys Gly Glu Val Thr Leu Arg Cys Trp Ala Leu Gly Phe Tyr Pro Ala Asp Ile Thr Leu Thr Trp Gln Leu Asn Gly Glu Glu Leu Thr
CAT GTG ACC CAT CAC CCC AGA TCT AAA GGT GAA GTC ACC CTG AGG TGC TGG GCC CTG GGC TTC TAC CCT GCT GAC ATC ACC CTG ACC TGG CAG TTG AAT GGG GAG GAG CTG ACC   2007

Gln Asp Met Glu Leu Val Glu Thr Arg Pro Ala Gly Asp Gly Thr Phe Gln Lys Trp Ala Ser Val Val Val Pro Leu Gly Lys Glu Gln Asn Tyr Thr Cys Arg Val Tyr His
CAG GAC ATG GAG CTT GTG GAG ACC AGG CCT GCA GGG GAT GGA ACC TTC CAG AAG TGG GCA TCT GTG GTG GTG CCT CTT GGG AAG GAG CAG AAT TAC ACA TGC CGT GTG TAC CAT   2121

Glu Gly Leu Pro Glu Pro Leu Thr Leu Arg Trp G
GAG GGG CTG CCT GAG CCC CTC ACC CTG AGA TGG G GTAAGGAGGGTGTGGGTGCAGAGCTGGGGTCAGGGAAAGCTGGAGCCTTCTGCAGACCCTGAGCTGGTCAGGGATGAGAGACTGGGGTCATAACCCTCACCTTC   2259

                                                EXON 5
                         lu Pro Pro Pro Ser Thr Asp Ser Tyr Met Val Ile Val Ala Val Leu Gly Val Leu Gly Ala Met Ala Ile Ile Gly Ala Val Val Ala Phe Val
ATTTCCTGTACCTGTCCTTCCCAG AG 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 GCT TTT GTG   2378

Met Lys Arg Arg Arg Asn Thr G
ATG AAG AGA AGG AGA AAC ACA G GTAAGAAAGGGNAGGGTCTGAGTTTTCTCTCAGCCTCCTTTAGAAGTGTCTCTGCTCATTAATGGGGAACACAGCCACACCCCACATTGCTACTGTCTCTAACTGGGTCTGCTGTCAGTT   2520

                                                EXON 6
                                                  ly Gly Lys Gly Gly Asp Tyr Ala Leu Ala Pro G
CTGGGAARTTCCAGTGTCAAGATCTTCCTTGAACTCTCACAGCTTTXCTTTTCACAG GT GGA AAA GGA GGG GAC TAT GCT CTG GCT CCA G GTTANTGTGGGGACAGGATNGTTCTGGGGGACATTGGAGTGAAGTTG   2657

                                                                                                                     EXON 7
                                                                                                               ly Ser Gln Ser Ser Glu
GAGATGATGGGAGCTCTGGGAATCCATAATAGCTCCTCCAGAGAAATCTTCTAGGGGCCTGAGTTGTGCCATGAAGTGAATACATTCATGTACATATGCATATACATTTGTTTTGTTTTACCCTAG GC TCC CAG AGC TCT GAA   2800

Met Ser Leu Arg Asp Cys Lys A
ATG TCT CTC CGA GAT TGT AAA G GTGACACTCTAGGGTCTGATTGGGGAGGGGCAATGTGGACATGATTGGGTTTCAGGGACTCCCAGAATCTCCTGAGAGTGAGTGGTGGGTTGCTGGAATGTTGTCTTCACAGTGATGGTT   2942

                    EXON 8
                    la Trm
CATGACTCTCATTCTCTAG CG TGA AGACAGCTGCCTGGACTGTACTGAGTGACAGACGATGTGTTCAGGTCTCTCCTGTGACATCCAGAGCCCTCAGTTCTCTTTACACAZCATTGTCTGATGTTCCCTGTGAGCTTGGGTTCAGTG   3089

TGAAGAACTGTGGAGCCCAGCCTGCCCTGCACACCAGGACCCTATCCCTGCACTGCCCTGTGTTCCCTTCCATAGCCAACCTTGCTGCTCCAGTCAAACACTGGGGGACATCTGCATCCTGTAAGCTCCATGCTACCCTGAGCTGCAGCT   3239

CCTCACTTCCACACTGAGAATAATAATTTGAATGTGGGTGGCTGGAGAGATGGCTCAGCGCTGACTGCTCTTCCAAAGGTCCTGAGTTCAAATCCCAGCAACACATGGTGGCTCACAACCATCTGTAATGGTATCTAACACCCTCTTCTG   3389

CAGTGTCTGAAGACAGCTACAGTGT                                                                                                                                  3414
```

Fig. 2. Nucleotide sequence of gene 27.5. The amino acid translation of each exon is given above the nucleotide sequence. Certain amino acid residues are encoded by two exons and are so indicated by splits in the three-letter amino acid code. Triple dots indicate gaps in intron sequences. Gaps in introns 1 and 3 are each approximately 200 nucleotides in length. The ambiguous base code is as follows: *P*, either A or G; *R*, either A or T; *Z*, A, T, or C; *Y*, either C or T; *S*, either G or C; *X*, A, T, or G.

hibit 80 percent homology. (iii) Two of the introns, 1 and 3, exhibit extensive regions of nonhomology in these two genes. (iv) The homology in the 3′ end of these genes terminates almost precisely at the end of the 3′ untranslated region as defined by comparison with the L^d-like cDNA clone pH-2II (*10*). Thus the exons and many of the introns of these class I genes are conserved. However, extensive divergences both in size (Fig. 1c) and sequence are seen in two introns and beyond the 3′ untranslated region of these genes. To what extent these divergences reflect a lack of selective pressure operating in these regions or the rapid changes that may occur in a pseudogene cannot be determined at this time.

Two highly repetitive sequence elements homologous to the human Alu sequence (*21*) have been found in the third intron of pseudogene 27.1 (*4*). Since Alu-like repetitive sequences appear to be ubiquitous in mammals and are transcribed in many cells, it has been postulated that they may play some undefined role in gene expression (*22*). The Alu-like sequences are not present in the third intron of gene 27.5. Hence the presence of these Alu-like sequences in intron 3 does not play a role in regulating the expression of gene 27.5. Furthermore, the Alu-like sequences occur in the third intron of gene 27.1 at precisely the boundaries of a large region of nonhomology (1000 nucleotides) in the comparison with the third intron in gene 27.5 (Fig. 3). Thus it is tempting to speculate that the Alu-like sequences, which have the characteristics of transposons (*22*), may be responsible for the insertion of a large region of foreign DNA in intron 3 of gene 27.1. This is consistent with the observation that the third intron of 27.1 is approximately 1000 nucleotides larger than the third intron of 27.5. It will be interesting to determine whether Alu-like sequences are present in other class I genes.

In summary, we have determined the coding sequence for an L^d gene of the mouse major histocompatibility complex. Thus, an H-2 gene corresponding to a serologically defined protein product has been identified and characterized. Cloning and DNA sequence studies suggest that there may be three (or more) L^d-like genes. It will be interesting to determine whether these putative L-like genes are codominantly expressed on all cells, or whether these genes are expressed in a tissue-specific manner such as is seen with the Qa and TL antigens. It appears that, in the future, other D-end functional genes will be defined by the strategies of gene cloning, gene transfer, and sequence analysis.

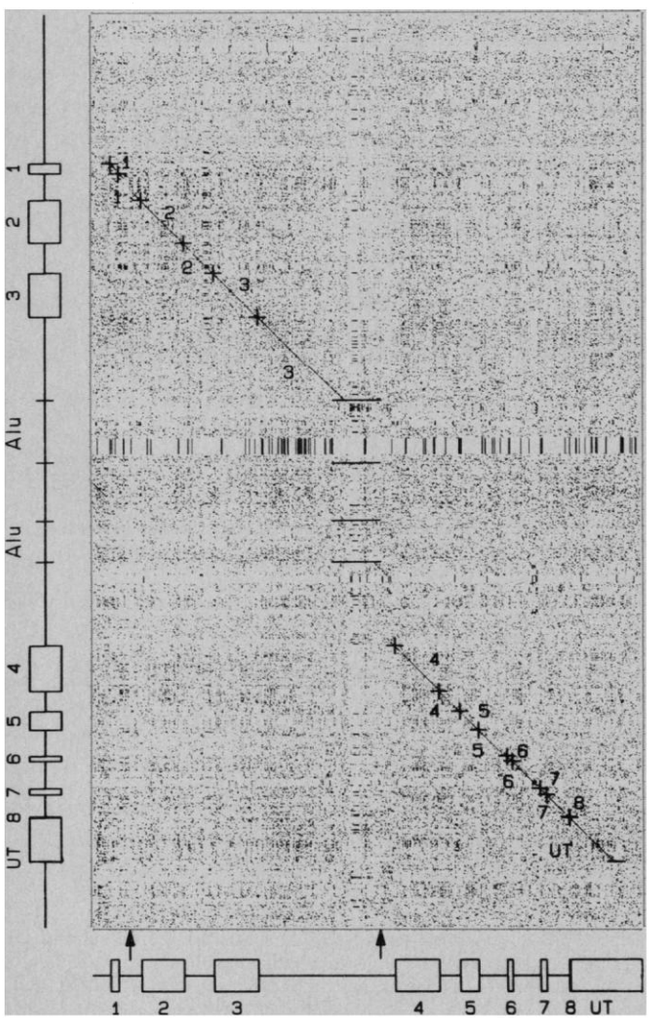

Fig. 3. Dot matrix homology analysis of gene 27.5 and pseudogene 27.1. The exon and intron structures of the two genes are shown along the sides of the figure. The positions of intron-exon boundaries are indicated by horizontal lines for 27.1 and vertical lines for 27.5 along the line of homology. Introns are numbered below the line of homology and exons are numbered above the line of homology. The positions of the sequence gaps in gene 27.5 are indicated by arrows. The positions of the Alu-like repetitive elements in gene 27.1 are indicated by horizontal lines along the line of homology and in intron 3 of pseudogene 27.1.

KEVIN W. MOORE
DNAX Research Institute,
Palo Alto, California 94304
BEVERLY TAYLOR SHER
Y. HENRY SUN
KURT ANDREW EAKLE
LEROY HOOD*
Division of Biology,
California Institute of Technology,
Pasadena 91125

References and Notes

1. J. Klein, *Science* **203**, 516 (1979).
2. G. D. Snell, J. Dausset, S. Nathenson, *Histocompatibility* (Academic Press, New York, 1976).
3. T. H. Hansen, K. Ozato, M. R. Melino, J. E. Coligan, T. J. Kindt, J. J. Jandinski, D. H. Sachs, *J. Immunol.* **126**, 1713 (1981).
4. M. Steinmetz, K. W. Moore, J. G. Frelinger, B. T. Sher, F.-W. Shen, E. A. Boyse, L. Hood, *Cell* **25**, 683 (1981).
5. R. S. Goodenow, M. McMillan, A. Örn, M. Nicolson, N. Davidson, J. A. Frelinger, L. Hood, *Science* **215**, 677 (1982).
6. F. Sanger, A. R. Coulson, B. G. Barrell, A. J. H. Smith, B. A. Roe, *J. Mol. Biol.* **143**, 161 (1980).
7. J. Messing, R. Crea, P. H. Seeburg, *Nucleic Acids Res.* **9**, 309 (1981).
8. J. A. Lopez de Castro, H. T. Orr, R. J. Robb, T. Kostyk, D. L. Mann, J. L. Strominger, *Biochemistry* **18**, 5704 (1979).
9. J. E. Coligan, T. J. Kindt, H. Uehara, J. Martinko, S. G. Nathenson, *Nature (London)* **291**, 35 (1981).
10. M. Steinmetz *et al.*, *Cell* **24**, 125 (1981).
11. F. Bregegere *et al.*, *Nature (London)* **292**, 78 (1981).
12. B. Lewin, *Cell* **22**, 324 (1980).
13. J. E. Coligan, T. J. Kindt, R. Nairn, S. G. Nathenson, D. H. Sachs, T. H. Hansen, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1134 (1980).
14. M. Steinmetz, A. Winoto, K. Minard, L. Hood, *Cell*, in press.
15. R. Goodenow, unpublished results.
16. J. Michaelson, L. Flaherty, E. Vitetta, M. D. Poulik, *J. Exp. Med.* **145**, 1066 (1977).
17. L. Flaherty, *The Role of the Major Histocompatibility Complex in Immunology* (Garland, New York, 1980).
18. T. H. Stanton and L. Hood, *Immunogenetics* **11**, 309 (1980).
19. M. J. Soloski, J. W. Uhr, L. Flaherty, E. S. Vitetta, *J. Exp. Med.* **153**, 1080 (1981).
20. T. Hunkapiller and L. Hood, in preparation.
21. A. S. Krayev, D. A. Kramerov, K. G. Skryabin, A. P. Ryskov, A. A. Baev, G. P. Georgiev, *Nucleic Acids Res.* **8**, 1201 (1980).
22. S. R. Haynes, T. P. Toomey, L. Leinwand, W. R. Jelinek, *Mol. Cell. Biol.* **1**, 573 (1981).

23. Supported by NIH grant GM 06965. K.W.M. was initially supported by a postdoctoral fellowship from the National Institutes of Health. B.T.S. was supported by a predoctoral fellowship from the National Science Foundation. Y.H.S. and K.A.E. were supported by a national research service award (1 T32 GM07616) from the National Institute of General Medical Sciences. We thank M. Douglas for computer graphics, T. Hunkapiller for computer programs, and B. Larsh for manuscript preparation. We thank W. R. Jelinek for pointing out the presence of several Alu-like sequences in gene 27.1.

* To whom correspondence should be addressed.

5 November 1981; revised 16 December 1981