# IDENTIFICATION OF THREE NEW Ig λ-LIKE GENES IN MAN

By HENRY CHANG,[*‡] ETHAN DMITROVSKY,[*] PHILIP A. HIETER,[§]
KENNETH MITCHELL,[‖] PHILIP LEDER,[‖] LESTER TUROCZI,[¶]
ILAN R. KIRSCH,[*] AND GREGORY F. HOLLIS[*]

*From the *National Cancer Institute–Navy Medical Oncology Branch, National Cancer
Institute and Naval Hospital, and the ‡Transplantation Research Program Center, Naval
Medical Research Institute, Bethesda, Maryland 20814; §Laboratory of Molecular Genetics,
National Institute of Child Health and Human Development, Bethesda, Maryland 20805; the
‖Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115; and the
Department of Biology, Wilkes College, Wilkes-Barre, Pennsylvania 18766*

An antibody molecule is constructed from four polypeptides, two identical L chains, ($\kappa$ or $\lambda$), and two identical H chains ($\mu$, $\delta$, $\gamma$, $\epsilon$, or $\alpha$). In its germline state, Ig $\kappa$ L chain DNA exists in three discrete groups of noncontiguous segments: V, J, and C. H chain–related DNA is slightly more complex, and includes a fourth group of discrete D segments. To form a functional L chain gene, one V region undergoes a site-specific recombination event and becomes contiguous with one J segment. H chain gene formation is a slightly more complicated variation on this theme, which includes the rearrangement of a D segment to a J segment before V-D recombination (for review see 1).

Antibody diversity can be accounted for by combinations obtained from this V-J (or V-D-J) joining and the association within a cell of such recombined L and H chain genes. Toward the expansion of antibody diversity, higher organisms have evolved hundreds of V regions capable of recombining with several separate J segments. In the case of the human $\kappa$ or H chain gene systems, gene duplication has increased the number of J regions to about six, located just 5' of the single $\kappa$ or $\mu$ C region (2, 3). Because these J regions encode different amino acid sequences, and each of these J regions is capable of recombining with any appropriate V region, the number of different polypeptide chains that can be made is increased manyfold.

The human $\lambda$ gene system also has increased the number of available J regions by gene duplication. However, in the $\lambda$ locus, the unit of duplication includes

the C region as well as the J region. Thus, in humans, there are at least six λ C regions, each apparently associated with its own J region, as was reported earlier (4).

In this paper, we describe a new gene system that provides a potential third class of Ig L chain polypeptides. While most closely related to the previously reported λ L chain system, these genes are clearly distinct. We have characterized two genes from this new class, which have open reading frames and have J regions associated with them. A third λ-like gene is described that appears to be a pseudogene. These λ-like genes have not been described at the protein level, but may represent other L chain alleles that can increase L chain repertoire by providing additional J region coding sequences. Alternatively, these genes may be incorporated into some other gene system requiring a V-J type of activation. Such analogous systems, for example, have already been described as mediating T cell receptor diversity (5–11).

## Materials and Methods

*DNA Preparation.* High molecular weight DNA was prepared according to Hieter, et al. (4). Intact cellular nuclei were isolated from PBS-washed leukocytes by Dounce homogenization and subsequent centrifugation through a sucrose cushion. The nuclei were then incubated overnight in proteinase K and SDS. DNA was purified from the resultant solution by phenol/chloroform extraction, precipitated in ethanol, and resuspended in deionized water.

*Restriction Enzyme Digestion and Southern Analysis.* All restriction enzyme digests were performed according to manufacturer's suggested conditions. Hybridizations were performed on genomic or cloned DNA cut by restriction endonucleases, size-fractionated on 0.8% agarose gels, and blotted onto nitrocellulose filters as described by Southern (12). Filters were then hybridized to $^{32}$P-labeled probes overnight at 42°C in a 10% Dextran sulfate, 4× SSC, 40% formamide, 0.8% Denhardt's Tris-buffered solution. After hybridization, filters were washed with 2 × SSC, 0.1% SDS three times at room temperature, and with 0.1 × SSC, 0.1% SDS twice at 52°C before autoradiography.

*Genomic Cloning.* High molecular weight DNA was digested with Eco RI endonuclease and size-fractionated by preparative agarose gel electrophoresis. Appropriately sized fractions were ligated into λ phage vector Charon 4A and packaged. The resulting recombinant phage library was screened as described by Benton and Davis (13). Selected phage DNAs were mapped and areas of particular interest were subcloned into the plasmid pBR327.

*Nucleotide Sequencing Procedure.* Dideoxy sequencing was performed using reagents and instructions prepared by Bethesda Research Laboratories, Gaithersburg, MD.

*Nucleotide Alignments.* Comparisons were made on a DEC-10 computer with the DNA:NUCALN program developed at the National Institutes of Health by Lipman and Wilbur (14).

## Results

*Identification and Cloning of Three New λ Genes.* Previously we have shown that at least six λ C region genes (each with its associated J segment) are clustered on an ~40 kb stretch of DNA on chromosome 22 (4). Numbered consecutively, the most 5' gene, $λ_{C1}$, resides on a 14 kb Eco RI fragment; $λ_{C2}$ and $λ_{C3}$ on the central 8 kb Eco RI fragment (polymorphic in the human population as 8, 13, 18, and 23 kb pieces, containing from 2–5 gene copies) (15); and the last three genes ($λ_{C4,5,6}$) on a 16 kb Eco RI fragment. At the time of that analysis it became clear that these six genes represented only a part of the λ-like genes found in the
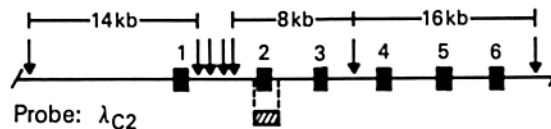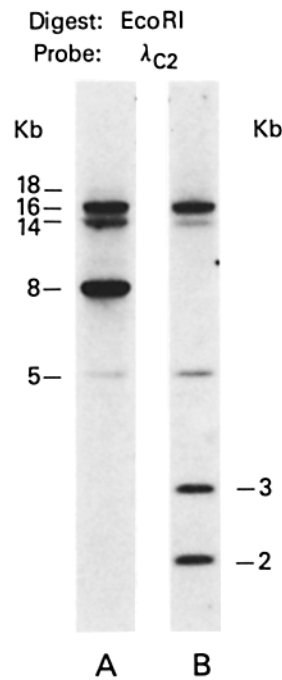
FIGURE 1. Demonstration of the existence of new λ-like genes. DNA from EBV-transformed B cell lines from the same patient (lane A, a κ producer; and lane B, a λ producer) was digested with the restriction enzyme Eco RI, electrophoresed, blotted, and hybridized with a $^{32}$P-labeled probe by the Southern technique (12). As shown in the diagram below, where arrows indicate Eco RI sites, the probe was derived from the C region of the second λ gene in the major functional cluster as described by Hieter et al. (4).

human genome. The abbreviations, such as $\lambda_{C1}$ and $\lambda_{J1}$, define specific coding areas. The abbreviation $\lambda_1$ is used as a more general term referring to coding and flanking regions around lambda constant gene one, including its associated J region.

The existence of additional λ-like genes is confirmed by the experiment shown in Fig. 1. In lane A of this Southern blot, germline DNA from a karyotypically normal EBV-transformed κ-producing B cell line has been digested with Eco RI and hybridized to a λ C region probe. It shows the germline 14, 8, and 16 kb bands associated with λ C genes 1–6, plus a faint 18 kb band and the 5 kb processed λ pseudogene (16). In addition there is a faintly hybridizing 10 kb band that remains uncharacterized at this time. In Fig. 1, lane B, DNA from a different EBV-transformed λ-producing B cell line from the same individual, digested with Eco RI, shows the loss of the central 8 kb Eco RI fragment, and two rearranged bands, 2 and 3 kb in size. These correspond to rearrangements
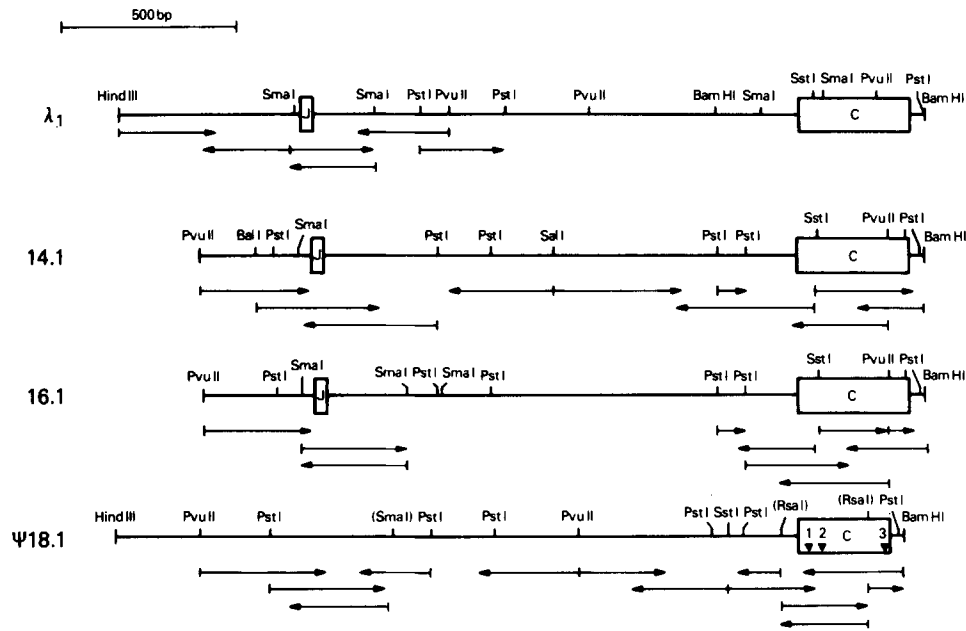
FIGURE 2. Comparison of the restriction enzyme maps of the genes λ₁, 14.1, 16.1, and pseudogene 18.1. The J and C regions are designated by boxes with the letters J and C. Characteristic restriction sites are shown; those in parentheses were useful for subcloning, but represent only some of the sites in the gene. Arrows indicate the direction and extent of dideoxy sequencing, except for λ_{Cl} whose C region sequence has already been reported (4). The position of deletions in the C portion of 18.1 are given by the numbers.

that have occurred within the middle 8 kb Eco RI fragment. In these two recombinations, the more 5' 14 kb Eco RI fragment containing the λ_{Cl} gene is deleted from both chromosomes, but surprisingly, a 14 kb band persists. Thus, this band must represent a distinct λ C homologous region found on another 14 kb Eco RI fragment. Cloning experiments confirmed the presence of a 14 kb band that is distinct from λ₁. Interestingly, when similar experiments were performed on the 16 kb band, they revealed another λ-like gene that was distinct from the 16 kb Eco RI fragment containing λ_{C4}–λ_{C6}. Therefore, Southern analysis and cloning experiments revealed two previously uncharacterized genes of man that share homology with a λ Ig C region.

Fig. 2 shows a comparison of the restriction enzyme maps of λ₁ and the new genes, called 14.1 and 16.1, plus that for a portion of a faintly hybridizing 18 kb Eco RI fragment (18.1) that we also cloned. These genes are homologous to λ C regions (as represented by the large black boxes), yet their restriction maps show that they are clearly different from the previously described major λ gene cluster. In the flanking region, however, the new genes appeared most similar to λ₁. Because the J region for λ₁ is within 2 kb 5' of the C region, we examined the equivalent area of 14.1 and 16.1 for a J region. This search was successful, as demonstrated by nucleotide sequence analysis below.

*14.1 and 16.1 are Nonallelic.* The restriction maps of 14.1 and 16.1 (Fig. 2) are very similar, and it was necessary to determine whether the two clones were

Digest: Eco RI + Bam HI
Probe: 16.1$_J$



A          B

16.1$_J$                    16.1$_C$
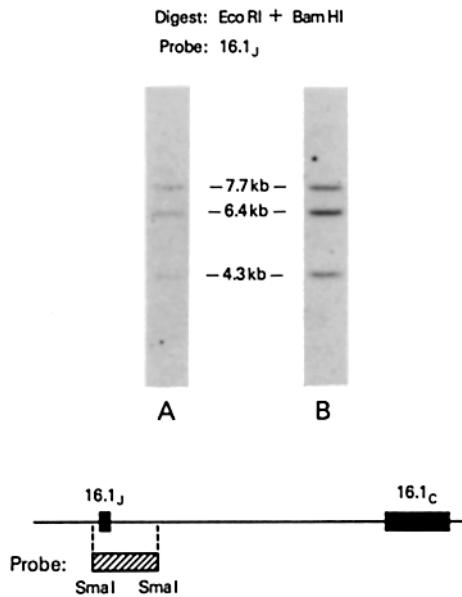
Probe:

Sma I   Sma I

FIGURE 3. Nonallelic nature of the new genes. DNA from the same cell lines as in Fig. 1 was digested with both restriction enzymes Eco RI and Bam HI, and analyzed by Southern blotting. The probe used was specifically derived from the J region of 16.1, as shown in the diagram below. The band at 4.3 kb corresponds to 14.1, the band at 6.4 kb corresponds to 16.1, and the remaining band at 7.7 kb remains uncharacterized at this time.

polymorphic alleles or different genes. Therefore, genomic DNA from 14 normal individuals was analyzed with a probe containing the J region from 16.1 (16.1$_J$), which also hybridizes to 14.1, but does not hybridize with $\lambda_1$ under stringent wash conditions (55°C, 0.1 × SSC). The typical result of this analysis was identical to Fig. 3, lane A. In all 14 samples, three bands were found, two of which correspond to 14.1 (4.3 kb) and 16.1 (6.4 kb), and a third band (7.7 kb), which remains unidentified at this time. Because all persons contained 14.1 and 16.1 bands, it seems unlikely that the two genes are polymorphic alleles.

When the 16.1$_J$ probe was used against the $\lambda$-producing cell line that had rearranged both $\lambda$ loci in the major functional cluster, both the 14.1 and 16.1 genes persist in the germline state (Fig. 3, lane B). This result confirms that these genes are not alleles of $\lambda_{C1}$, which is deleted in these cells (data not shown). Further, it supports the idea that these genes are not located between the $\lambda$ V regions and the major $\lambda$ cluster, a region which is also deleted during V-J recombination.

*Sequence Analysis of 14.1 and 16.1.* To characterize the new genes further, we determined the nucleotide sequence of 14.1 and 16.1 in the areas that correspond to the J and C regions (Figs. 4 and 5). The C regions of 14.1 and 16.1 are encoded as single exons that are 96% homologous to each other and 86–89% homologous to the known $\lambda$ genes. The number of nucleotides altered, and percentage differences between the genes is given in Table I. Compared to $\lambda_{C1}$, 14.1 and 16.1 have base changes in a total of 42 positions. Of these, 31 alterations are identical in both genes, 3 occur in 14.1 alone, and 8 in 16.1 alone.

```
                                                30                                        60
λc1   GGT CAG CCC AAG GCC AAC CCC ACG GTC ACT CTG TTC CCG CCC TCC TCT GAG GAG CTC CAA
14.1  --- --- --- --- --- -C- --- T-- --- --- --- --- --- --G --- --- --- --- --- ---
16.1  --- --- --- --- --- -C- --- T-- --- --- --- --- -T- --G --- --- --- --- --- ---

                               112     114             119
λc1   GLY GLN PRO LYS ALA ASN PRO THR VAL THR LEU PHE PRO PRO SER SER GLU GLU LEU GLN
14.1   -   -   -   -   -  ASN  -  SER  -   -   -   -   -   -   -   -   -   -   -   -
16.1   -   -   -   -   -  ASN  -  SER  -   -   -   -  LEU  -   -   -   -   -   -   -

                                                90                                       120
λc1   GCC AAC AAG GCC ACA CTA GTG TGT CTG ATC AGT GAC TTC TAC CCG GGA GCT GTG ACA GTG
14.1  --- --- --- --T --- --G --- --- --C --G -A- --- --T --T --- --- ATC T-- --G ---
16.1  --- --- --- --- --- --G --- --- --C --G -A- --- --T -T- --- --- ATC T-- --G ---

                                        136 137         141     143 144
λc1   ALA ASN LYS ALA THR LEU VAL CYS LEU ILE SER ASP PHE TYR PRO GLY ALA VAL THR VAL
14.1   -   -   -   -   -   -   -   -   -  MET ASN  -   -   -   -   -  ILE LEU THR VAL
16.1   -   -   -   -   -   -   -   -   -  MET ASN  -   -   -  LEU  -  ILE LEU  -   -

                                                150                                      180
λc1   GCT TGG AAG GCA GAT GGC AGC CCC GTC AAG GCG GGA GTG GAG ACG ACC AAA CCC TCC AAA
14.1  A-C --- --- --- --- --T -C- --- A-- -CC CA- --C --- --- -T- --- -CG --- --- ---
16.1  A-C --- --- --- --T -C- --- A-- -CC CA- --C --- --- -T- --- -CG --- --- ---

      147                     153         155 156 157         161     163
λc1   ALA TRP LYS ALA ASP GLY SER PRO VAL LYS ALA GLY VAL GLU THR THR LYS PRO SER LYS
14.1  THR  -   -   -   -   -  THR  -  ILE THR GLN  -   -   -  MET  -  THR  -   -   -
16.1  THR  -   -   -   -   -  THR  -  ILE THR GLN  -   -   -  MET  -  THR  -   -   -

                                                210                                      240
λc1   CAG AGC AAC AAC AAG TAC GCG GCC AGC AGC TAC CTG AGC CTG ACG CCC GAG CAG TGG AAG
14.1  --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G-
16.1  --- --- --- -G- --- --- AT- --- --- --- --- --- --- --- --- --- --- --- --- -G-

                      171         174                                                    187
λc1   GLN SER ASN ASN LYS TYR ALA ALA SER SER TYR LEU SER LEU THR PRO GLU GLN TRP LYS
14.1   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -  ARG
16.1   -   -   -  SER  -   -  MET  -   -   -   -   -   -   -   -   -   -   -   -  ARG

                                                270                                      300
λc1   TCC CAC AGA AGC TAC AGC TGC CAG GTC ACG CAT GAA GGG AGC ACC GTG GAG AAG ACA GTG
14.1  --- -G- --- --- --- --- --- --- --- -T- --C --- --- --- --- --- --- --- --G ---
16.1  --- -G- --- --- --- --- --- --- --- -T- --- --- --- --T -CA --- --- --G ---

          189                             197                     205
λc1   SER HIS ARG SER TYR SER CYS GLN VAL THR HIS GLU GLY SER THR VAL GLU LYS THR VAL
14.1   -  ARG  -   -   -   -   -   -   -  MET  -   -   -   -   -   -   -   -   -   -
16.1   -  ARG  -   -   -   -   -   -   -  MET  -   -   -   -   -  ALA  -   -   -   -

λc1   GCC CCT ACA GAA TGT TCA TAG
14.1  --- --- G-- --- --- --- ---
16.1  --- --- G-- --- --- --- ---

              212
λc1   ALA PRO THR GLU CYS SER TER
14.1   -   -  ALA  -   -   -   -
16.1   -   -  ALA '-   -   -   -
```

FIGURE 4. Comparison of the nucleotide and deduced amino acid sequences of the C regions of λc1, and 14.1 and 16.1. Dashed lines indicate identity. The site for splicing with the J region is indicated with an arrow.

Thus, the two genes are more closely related to each other than to the previously sequenced λ genes.

In addition, the most striking feature of the sequences of 14.1 and 16.1 is that they predict open reading frames. Their deduced amino acid sequences are 95% identical (Table I). When compared to Mcg (the myeloma protein product of λc1) (17) and Kern⁻ (the protein made by λc2 and λc3) (18), 14.1 differs by 17–18 amino acids (16–17%) and 16.1 by 22–23 amino acids (21–22%). All of the replacement amino acids caused by the changes can be found at these positions in other human or mouse λ chains that have been reported (19), except methionines 161 and 197, and the alanine in position 205 of 16.1, which are unique.

A comparison of λc1–λc3 reveals a greater divergence of nucleotide sequence than amino acid sequence, as is expected in genes where selective pressure maintains specific protein sequences. In contrast, the percentage of nucleotide sequence differences between 14.1 and 16.1 is less than their percentage of amino acid sequence differences. Similarly, 14.1 and 16.1 are more closely related to λc1–λc3 in nucleotide than amino acid sequence on a percentage basis.

```
λJ1   GGG TTT TGG TCT GAG CCT CAG TCA CTG TGT TAT GTC TTC GGA ACT GGG ACC AAG GTC ACC GTC CTA GGT
14.1  --- --- CAA --C A-- -A- A-C --- G-- ACG C-- --G --T --C -GC --- --- C-- C-- --- --T T-- A--
16.1  --- --- CAA --C A-- -A- A-T --- G-- AAG C-- --G --T --C -GC --- --- C-- C-- --- --T T-- G--
                                                              100                    105
λJ1            Y   V F G   T G T   K V   T V L   G
14.1         T H   V F G   S G T   Q L   T V L   S
16.1         K H   V F G   S G T   Q L   T V L   G
```
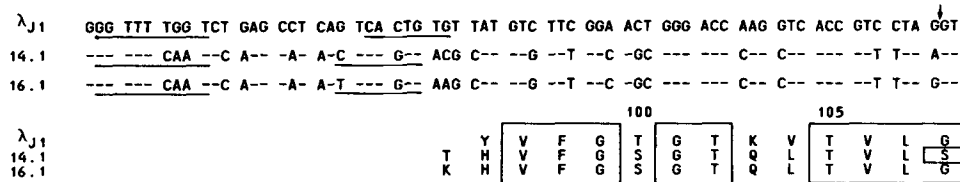
FIGURE 5. Comparison of the nucleotide and deduced amino acid sequences of the J region of λ_C1, and 14.1 and 16.1. The nonameric and heptameric signals are underlined, the splice site to the C region is indicated with an arrow, and the homologous amino acids, using the single letter amino acid code, are boxed.

TABLE I

*Interrelationships between C Regions of Genes*

| C region genes | λC1 | λC2 | λC3 | 14.1 | 16.1 |
|---|---|---|---|---|---|
| λC1 | | 17 (5.3) | 22 (6.9) | 34 (10.7) | 39 (12.2) |
| λC2 | 5 (4.7) | | 9 (2.8) | 38 (11.9) | 43 (13.5) |
| λC3 | 6 (5.7) | 1 (0.9) | | 39 (12.2) | 44 (13.8) |
| 14.1 | 17 (16.0) | 17 (16.0) | 18 (17.0) | | 11 (3.5) |
| 16.1 | 22 (20.8) | 22 (20.8) | 23 (21.7) | 5 (4.7) | |

Total bases, 318. Total amino acids, 106. Nucleotide differences are shown above the diagonal; amino acid differences below. Numbers in parentheses are percent differences.

*J Regions of 14.1 and 16.1.* The J segment of λ₁ is located 1.4 kb 5′ of the C region, and 14.1 and 16.1 each have a single J segment 1.4 kb and 1.3 kb 5′ of their respective C regions (Fig. 2). In addition, the detailed structure of these areas is quite similar, and consists of the J coding region flanked on its 5′ side by the recognition sequences that presumably mediate DNA rearrangements. At the nucleotide level (Fig. 5), $\lambda_{J1}$, $14.1_J$ and $16.1_J$ have a nonamer recognition sequence separated by 10–12 basepairs from a heptamer recognition sequence that abuts the 5′ end of the J coding sequence. The J regions of 14.1 and 16.1 are more homologous to each other than to $\lambda_{J1}$.

For both 14.1 and 16.1, the J coding sequence appears to be one amino acid longer than $\lambda_{J1}$. This prediction is based on the location of the heptamer recognition sequence. While the alignment of the heptanucleotide sequence appears to have been shifted by point mutations, compensating mutations have occurred that produce a good match with the heptamer consensus sequence. In fact, 14.1 and 16.1 match six and five basepairs, respectively, out of seven with the heptamer CACTGTG sequence of $\lambda_{J1}$. Similarly, six of nine nucleotides of 14.1 and 16.1 are identical to the nonamer sequence of $\lambda_{J1}$. This level of conservation of the heptamer and nonamer recognition sequences is comparable to that seen in the human κ and H chain J regions (2, 3), and the human T cell receptor α chain J regions (11).

The 3′ end of each presumptive J region sequence ends with a dinucleotide, –GT, which could be cleaved off with the first nucleotide of the C region to
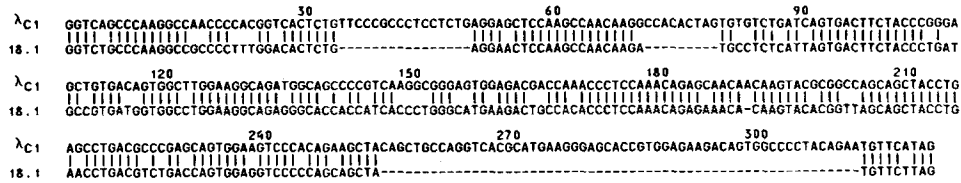
```
          30                              60                           90
λc₁  GGTCAGCCCAAGGCCAACCCCACGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTCCAAGCCAACAAGGCCACACTAGTGTGTCTGATCAGTGACTTCTACCCGGGA
     IIII IIIIIIIIIII IIII  II IIIIIII            IIII IIIIIIIIIIIIII      II III II IIIIIIIIIIIIIII I
18.1 GGTCTGCCCAAGGCCGCCCCTTTGGACACTCTG----------------AGGAACTCCAAGCCAACAAGA---------TGCCTCTCATTAGTGACTTCTACCCTGAT

          120                             150                          180                           210
λc₁  GCTGTGACAGTGGCTTGGAAGGCAGATGGCAGCCCCGTCAAGGCGGGAGTGGAGACGACCAAACCCTCCAAACAGAGCAACAACAAGTACGCGGCCAGCAGCTACCTG
     II IIIII  IIIII IIIIIIIIIII IIII I II III    III II IIII  III IIIIIIIIIIIIIII IIII IIIIIII III  IIIIIIIIIIII
18.1 GCCGTGATGGTGGCCTGGAAGGCAGAGGGCACCACCATCACCCTGGGCATGAAGACTGCCACACCCTCCAAACAGAGAAACA-CAAGTACACGGTTAGCAGCTACCTG

          240                             270                          300
λc₁  AGCCTGACGCCCGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGGCCCCTACAGAATGTTCATAG
     I IIIIIII I II IIIIIII IIIII III IIIII                                                       IIIII III
18.1 AACCTGACGTCTGACCAGTGGAGGTCCCCCAGCAGCTA------------------------------------------------------------TGTTCTTAG
```

FIGURE 6. Comparison of the nucleotide sequence of λc₁, with that of pseudogene 18.1. The vertical lines indicate homology and the dashed lines indicate the deletions.

allow a splice to form a nucleotide triplet, GGT, the codon for amino acid 108, glycine. Thus, the structure of 14.1ⱼ and 16.1ⱼ are consistent with J regions known to be functional in Ig and T cell receptor genes.

*18.1 Is a Pseudogene.* A clone of an 18 kb Eco RI fragment (18.1) also contained an area homologous to the λ C region, as shown in Fig. 2. Nucleotide sequence analysis of this segment revealed that, although it was most closely related to 16.1 and 14.1, it had undergone many more changes than these genes when compared to λc₁–λc₃. There are three large deletions within the putative coding region, as well as numerous point mutations, which introduce several frameshifts and stop codons (Fig. 6). We have not been able to find a sequence upstream from the 18.1 C region that encodes a J region. Thus, the structure of this gene, with many nucleotide changes and the absence of a J region, suggests that it is a nonfunctional λ pseudogene.

## Discussion

We provide evidence for the existence of new λ L chain–like genes in man. Two of them, 14.1 and 16.1, have a structure similar to the known functional locus, with a C region associated with its own J segment. However, they map on distinctly different stretches of DNA, and are nonallelic. In the C region, the nucleotide sequences of 14.1 and 16.1 are most homologous to each other, and less homologous to pseudogene 18.1 and the known functional λ genes. Their most important features are conserved open reading frames associated with appropriate recognition signals for recombination and splicing in the J regions.

Although there may be considerable selective pressure to preserve the coding sequences of functional genes, the flanking sequences may have fewer constraints to vary (20). As a result, interrelationships between the genes can be surmised. We observed that a probe derived from an area flanking λⱼ₁ hybridizes well to 14.1, 16.1, and 18.1, but poorly to λ₂–λ₆. Conversely, probes obtained from 16.1ⱼ and the corresponding area of 18.1 hybridize best to λ₁, 14.1, 16.1, and 18.1. Probes that originate from λⱼ₂ preferentially hybridize to λ₂–λ₆, rather than to the other genes (data not shown).

These findings, in addition to nucleotide sequence analysis, suggest a possible evolutionary model for the human λ gene family (Fig. 7). In this scheme, a common λ progenitor would be the source of the λ genes present in human DNA through gene duplication. The first event would lead to two divergent arms of the λ family, one most closely related to λ₁, and the other, a progenitor of λ₂–λ₆. We separate the latter genes from λ₁ at this early stage because the DNA flanking λ₂–λ₆ hybridizes poorly with genes on the λ₁ side of the family.
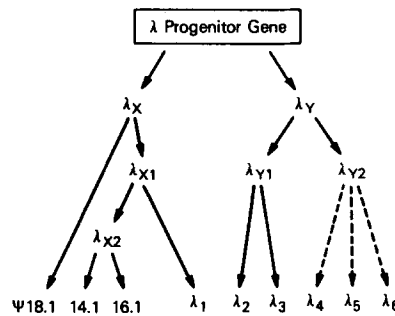
FIGURE 7. A model for the evolution of the $\lambda$ gene family. Intermediate precursors are designated by $X$ and $Y$. The length of the arrows are not meant to imply evolutionary distances. Dotted lines are used where the basic relationship is unknown.

The individual genes $\lambda_2-\lambda_6$ would arise from their progenitor gene by further gene duplication; the most recent event appears to have resulted in the generation of $\lambda_2$ and $\lambda_3$, which share extensive homology in both coding and flanking regions. Insufficient data are available at this time about the relationships of $\lambda_4-\lambda_6$.

On the other side of the family, the progenitor gene, $\lambda_X$, appears to have given rise to $\lambda_{X1}$ and pseudogene 18.1. $\lambda_{X1}$ is the precursor of $\lambda_1$, 14.1, and 16.1. 18.1 appears to have separated from $\lambda_1$, 14.1, and 16.1 before their duplication, because Southern analysis and nucleotide sequence analysis indicate it is approximately equally related to all three genes. Based on nucleotide sequence analysis of the C and flanking regions, 14.1 and 16.1 are most closely related to $\lambda_1$. The suggestion that 14.1 and 16.1 arose from a more recent duplication is supported by the fact that the nucleotide sequences of the C regions of 14.1 and 16.1 differ by ~4%, but both diverge 11–12% from $\lambda_{C1}$. In addition, most base substitutions in $14.1_C$ also occur in $16.1_C$, although a few mutations have occurred in each gene individually.

The chromosomal location of 14.1 and 16.1 appears to be consistent with the mechanism of gene duplication proposed in this model. While we have not formally assigned 14.1 and 16.1 to human chromosome 22, our analysis of somatic cell hybrids to map the $\lambda$ C regions $\lambda_1-\lambda_6$ might have shown a discordance for the 14.1 and 16.1 genes had they not been on this chromosome (21). To assign these genes to a chromosome more definitively, we have cloned unique pieces of DNA from the genomic clones and are examining their chromosomal location by in situ hybridization.

Two new genes have been found that can encode $\lambda$-like C regions, and expression of these potentially functional genes is being sought. It is also possible that these two genes, 14.1 and 16.1, may function in a role different from that of the antibody molecule. Recent studies have shown the T cell receptor of man to share homology with $\lambda$ L chains (5), but neither T cell receptor $\alpha$, $\beta$, nor $\gamma$ chains are identical to these genes (5–11). No teleological rationale is obvious to explain the conservation of these $\lambda$-like sequences in man, but selective pressure to maintain their open reading frames is noteworthy. We expect that eventually they will be shown to expand the power and versatility of the humoral immune response, although we cannot rule out their use in an as yet unreported parallel system similar to the Ig $\lambda$ or T cell receptor systems.

## Summary

Three new human λ L chain–like Ig genes are identified by restriction enzyme and nucleotide sequence analysis. Two genes, 14.1 and 16.1, have intact J and C regions, and are potentially functional, with open reading frames. A third gene, 18.1, is a pseudogene. The evolutionary lineage of these genes compared to the known functional locus $\lambda_{CI}-\lambda_{C6}$ can be surmised from Southern blot and nucleotide homologies. This study demonstrates that the human λ gene family is more complex than previously recognized.

## References

1. Leder, P. 1982. The genetics of antibody diversity. *Sci. Am.* 246:102.
2. Hieter, P. A., J. V. Maizel, Jr., and P. Leder. 1982. Evolution of human immunoglobulin κ J region genes. *J. Biol. Chem.* 257:1516.
3. Ravetch, J. V., U. Siebenlist, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Structure of the human immunoglobulin μ locus: characterization of embryonic and rearranged J and D genes. *Cell.* 27:583.
4. Hieter, P. A., G. F. Hollis, S. J. Korsmeyer, T. A. Waldmann, and P. Leder. 1981. Clustered arrangement of immunoglobulin λ constant region genes in man. *Nature (Lond.).* 294:536.
5. Yanagi, Y., Y. Yoshikai, K. Leggett, S. P. Clark, I. Aleksander, and T. W. Mak. 1984. A human T cell–specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. *Nature (Lond.).* 308:145.
6. Hedrick, S. M., E. A. Nielsen, J. Kavaler, D. I. Cohen, and M. M. Davis. 1984. Sequence relationships between putative T-cell receptor polypeptides and immunoglobulins. *Nature (Lond.).* 308:153.
7. Hayday, A. C., H. Saito, S. D. Gillies, D. M. Kranz, G. Tanigawa, H. N. Eisen, and S. Tonegawa. 1985. Structure, organization, and somatic rearrangement of T cell gamma genes. *Cell.* 40:259.
8. Arden, B., J. L. Klotz, G. Siu, and L. E. Hood. 1985. Diversity and structure of genes of the α family of mouse T-cell antigen receptor. *Nature (Lond.).* 316:783.
9. Hayday, A. C., D. J. Diamond, G. Tanigawa, J. S. Heilig, V. Folsom, H. Saito, and S. Tonegawa. 1985. Unusual organization and diversity of T-cell receptor α-chain genes. *Nature (Lond.).* 316:828.
10. Winoto, A., S. Mjolsness, and L. Hood. 1985. Genomic organization of the genes encoding mouse T-cell receptor α-chain. *Nature (Lond.).* 316:832.
11. Yoshikai, Y., S. P. Clark, S. Taylor, U. Sohn, B. I. Wilson, M. D. Minden, and T. W. Mak. 1985. Organization and sequences of the variable, joining and constant region genes of the human T-cell receptor α-chain. *Nature (Lond.).* 316:837.
12. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503.
13. Benton, W. D., and R. W. Davis. 1977. Screening $\lambda_{gt}$ recombinant clones by hybridization to single plaques in situ. *Science (Wash. DC).* 196:180.
14. Wilbur, W. J., and D. J. Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA.* 80:726.
15. Taub, R. A., G. F. Hollis, P. A. Hieter, S. Korsmeyer, T. A. Waldmann, and P.

Leder. 1983. Variable amplification of immunoglobulin λ light-chain genes in human populations. *Nature (Lond.).* 304:172.

16. Hollis, G. F., P. A. Hieter, O. W. McBride, D. Swan, and P. Leder. 1982. Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. *Nature (Lond.).* 296:321.

17. Fett, J. W., and H. F. Deutsch. 1974. Primary structure of the Mcg λ chain. *Biochemistry.* 13:4102.

18. Ponstingl, H., M. Hess, and N. Hilschmamm. 1971. Die Primarstruktur einer monoklonalen Immunoglobulin-L-Kette vom Lambda-Typ, Subgruppe IV (Bence-Jones-Protein Kern) *Hoppe-Seyle's Z. Physiol. Chem.* 352:247.

19. Kabat, E. A., T. T. Wu, H. Bilofsky, M. Reid-Miller, and H. Perry. 1983. Sequences of proteins of immunological interest. U.S. Department of Health and Human Services. U.S. Govt. Printing Office Number 396-847. 323 pp.

20. Konkel, D. A., J. V. Maizel, Jr., and P. Leder. 1979. The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes. *Cell.* 18:865.

21. McBride, O. W., P. A. Hieter, G. F. Hollis, D. Swan, M. C. Otey, and P. Leder. 1982. Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J. Exp. Med.* 155:1480.