

# SOMATIC MUTATION CREATES DIVERSITY IN THE MAJOR GROUP OF MOUSE IMMUNOGLOBULIN $\kappa$ LIGHT CHAINS

BY GÜNTHER HEINRICH,\*\* ANDRÉ TRAUNECKER,\* AND  
SUSUMU TONEGAWA<sup>§</sup>

*From the \*Basel Institute for Immunology, Postfach, 4005 Basel, Switzerland; †Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138; §Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

The V and C regions of an immunoglobulin light chain are encoded by different gene segments. The C-region segment encodes the carboxy-terminal part, while two segments—V and J—are somatically rearranged to create the gene segment that encodes a V region (1–4). The V regions of light and heavy chains are characterized by great sequence variability. This variability is especially high in the three parts of the V region that primarily form the antigen-binding site and that are therefore referred to as “complementarity determining regions”—CDR1, CDR2, and CDR3—or “hypervariable regions” (5–7). The parts of the V region outside of CDR1, CDR2, and CDR3 are called “framework regions.”

It could be shown that somatic mutation in a single V gene segment generates antibody diversity in the  $\lambda_1$  light-chain system of BALB/c and NZB mice (2, 8–10). It could also be demonstrated that somatic mutation is responsible for antibody diversity in the  $V_H$  regions of specific mouse antibodies (11, 12).

Evidence has also been produced that suggests that somatic mutations also play a role in the amplification of the germ line–encoded genetic information in the  $\kappa$  chain genes (13, 14). However, confirmation of this evidence by direct analysis of cloned DNA fragments has been limited (15, 16).

Diversity in myeloma protein  $V_\kappa$  regions has only been extensively studied in BALB/c and NZB mice (17, 18). Based on amino acid sequence homology in the amino terminal 23 residues the variable regions of these myeloma  $\kappa$  light chains were first classified into different groups, of which the  $V_\kappa$ -21 group is one (19). Based on more extensive sequence data for more than 30 BALB/c and NZB myeloma  $\kappa$  light chains of the  $V_\kappa$ -21 group, they were further subdivided into seven major subgroups, designated  $V_\kappa$ -21A through  $V_\kappa$ -21G (20, 21).

Even though  $V_\kappa$ -21 is the best defined group, amino acid sequence analysis allows only indirect conclusions about the mechanisms that generate antibody diversity. Therefore, we cloned most of the germ line V genes of the  $V_\kappa$ -21 group using cDNA of MOPC321 as a probe.

The Basel Institute for Immunology was founded and is supported by F. Hoffman-La Roche and Co., Limited Company, Basel, Switzerland.

Related V<sub>H</sub> genes have been shown to be linked at an average distance of 10–12 kb (11, 12, 22). The analysis of the clones containing overlapping chromosomal segments enabled us to determine the linkage arrangement of the germ line genes that encode the V $\kappa$ -21 group. This also supports our assumption that we have cloned almost all germ line genes of the V $\kappa$ -21 group. We determined the complete nucleic acid sequence of five different germ line V genes and showed that three of these sequences encode the "prototype sequence" of three different subgroups, namely V $\kappa$ -21B, C, and E. Therefore, we conclude that antibody diversity in the  $\kappa$  system, as in the  $\lambda_1$  system, is generated predominantly by somatic mutation of few germ line V genes. Somatic mutation is not restricted to the hypervariable regions; it also occurs in the framework regions.

### Materials and Methods

*Bacteria, Phages, and DNA.* Phage Charon 4A was obtained from F. Blattner (23). The  $\lambda$ gtWES: $\lambda$ B was originally from P. Leder (24). *E. coli* 803 ( $r_c^-$ ,  $m_c^-$ , Su III<sup>+</sup>) was obtained from K. and N. E. Murray. Lysogens used for preparation of packaging mixtures, BHB 2688 (N205  $recA^-$  [ $\lambda$ imm434 b2 red3 Eam4 Sam7]/ $\lambda$ ) and BHB 2690 (N205  $recA^-$  [ $\lambda$ imm434 cIts b2 red3 Dam15 Sam7]/ $\lambda$ ) were obtained from B. Hohn (25). MPC11 DNAs LF and L+LF were obtained from H. Zachau (26).

*Mice.* We obtained pregnant BALB/c mice from Tierfarm, CH-4414 Füllinsdorf, Switzerland and Bomholtgard, 8680 Ry, Denmark.

*Enzymes.* Restriction enzymes were purchased from Boehringer Mannheim Biochemicals (Mannheim, Federal Republic of Germany), New England Biolabs (Beverly, MA) and BRL.

*DNA Blots.* Cellular Southern blot hybridization was carried out as described by Southern (27) and modified by Wahl et al. (28). 10  $\mu$ g of embryo BALB/c DNA was digested with the appropriate restriction endonuclease and separated on a vertical agarose slab gel. After partial depurination and denaturation the DNA was transferred to a nitrocellulose membrane filter (Schleicher and Schuell BA85) and hybridized with  $3 \times 10^7$  cpm of the nick-translated 850-bp-long 5D10 HhaI fragment (3).

Southern blot hybridization for phage DNA under low stringent washing conditions was carried out as described by Matthyssens et al. (29).

*DNA Cloning.* Two recombinant phage libraries previously constructed in this laboratory (30, 31) were screened using the HhaI fragment of plasmid 5D10 as a probe.  $6 \times 10^5$  plaques from an EcoRI partial library and  $6 \times 10^5$  plaques from an EcoRI\* partial library were screened by the method of Benton and Davis (32). The 2.3-kb EcoRI clone was isolated from a complete EcoRI digest of BALB/c embryo DNA using the preparative agarose gel method (33). All cloning experiments were carried out under P3-EK2 conditions, in accordance with the National Institutes of Health guidelines for recombinant DNA research issued in 1976.

*Electron Microscopy.* Electron microscopic analysis of double-stranded DNA-mRNA hybrids, DNA heteroduplexes, and R hybrids was carried out as described by Brack (34).

*DNA Sequencing.* The 3' end of DNA fragments were labeled with four deoxyribonucleoside triphosphates and radiolabeled with  $\alpha^{32}$ P (Radiochemical Centre, Amersham) and Klenow enzyme (Boehringer Mannheim), as described previously (33). The 5' ends of DNA fragments were labeled as described by Sakano et al. (35). The strands of labeled fragments were separated and sequenced as described by Maxam and Gilbert (36).

### Results

*Gene Segments that Cross-hybridize with MOPC321.* To detect gene segments encoding the variable regions of the V $\kappa$ -21 group, we performed a series of Southern blotting experiments using, as a hybridization probe, a nick-translated

850-bp HhaI fragment from the plasmid 5D10, which carries the nearly complete cDNA of the  $\kappa$  light chain from the myeloma MOPC321. With this probe we would expect to be able to detect the germ line V gene segment of MOPC321 ( $V_{\kappa-21}$  C subgroup) as well as any other germ line  $V_{\kappa-21}$  gene segment sufficiently homologous to the V region of MOPC321.

When EcoRI-digested BALB/c DNA of 10–12-d-old embryos was separated on a 0.8% agarose gel, transferred to nitrocellulose paper, and hybridized to the  $\kappa$  probe, we detected seven bands of the following sizes: 1.5 kb, 2.6 kb, 4 kb, 6 kb, 8.5 kb, 15 kb, and 18 kb (Fig. 1A). We previously reported part of this experimental result (3). The 15-kb fragment had been previously cloned and a partial nucleotide sequence determined. By sequencing and R-loop mapping, it had been shown that it contains all of the  $J_{\kappa}$ 's and the  $C_{\kappa}$  but no MOPC321 or MOPC321-like  $V_{\kappa}$  sequence (33). The HpaI digest shows at least 8, BglII 10, HindIII 10, and BamHI 9 bands (Fig. 1A). In all of these digests some of the bands are more intense than others. In the EcoRI digest, the 4-kb, 6-kb, 8.5-kb, and 18-kb bands are stronger than the others, and as the DNA was depurinated before transfer, this is unlikely to be due to different transfer efficiencies. This view is supported by the fact that variation in the intensity of the bands was not dependent upon the size of the fragments. Several factors can be considered. First, these fragments might bear sequences with varying homology to the MOPC321  $V_{\kappa}$  sequences. Second, the intensity difference may be due to variation

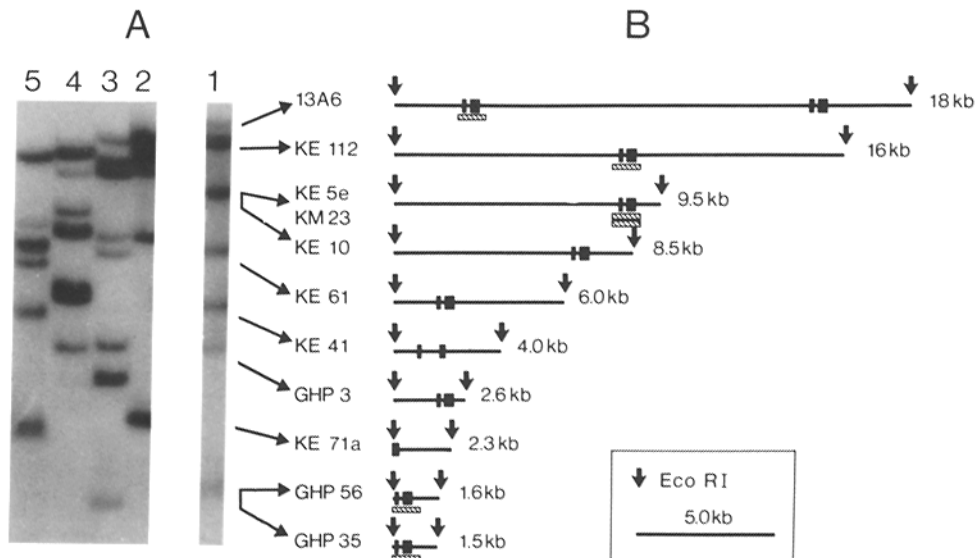


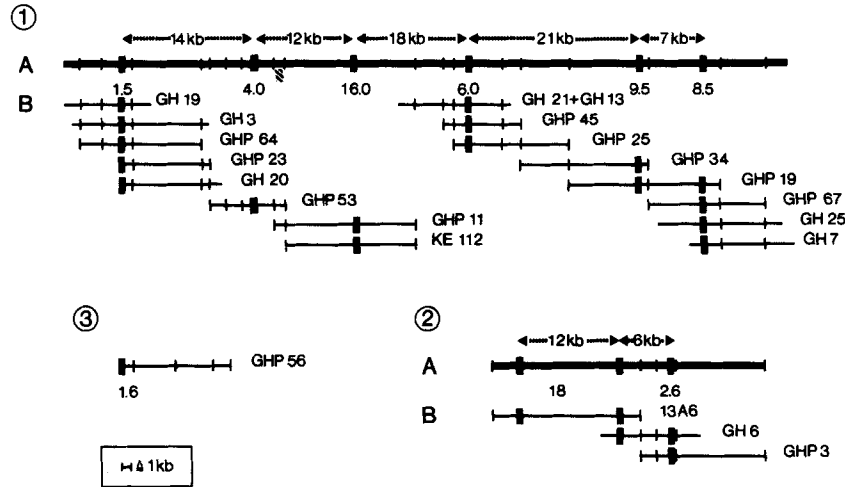
FIGURE 1. (A) Southern blot analysis of BALB/c embryo DNA.  $\sim 10 \mu\text{g}$  BALB/c DNA was digested with different restriction endonucleases (number 1, EcoRI; number 2, HpaI; number 3, BglII; number 4, HindIII; number 5, BamHI) fractionated on a 0.8% agarose gel, transferred to nitrocellulose paper, hybridized with  $5 \times 10^7$  cpm nick-translated 850-bp-long HhaI fragment of plasmid 5D10. (B) 10 embryo  $V_{\kappa}$  gene clones that cross-hybridized with the cDNA probe of MOPC321. The numbers indicate the length of the DNA between the two EcoRI sites ( $\downarrow$ ). The location of the exons which encode the variable-region polypeptide chain was determined by R-loop analysis. The shaded bars under certain V genes indicate which genes have been sequenced.

in the number of V genes on a single fragment. Third, some bands might contain two or more hybridizable fragments of the same or very similar size. For instance, when we separated the fragments over a larger distance we found that the 1.5-kb and 8.5-kb bands contain two fragments of slightly different size (see Fig. 4). Each of these points are dealt with in the subsequent section. Based on the overall results we conclude that there are ~8–10 germ line genes that cross-hybridize with MOPC321.

*Cloning of the MOPC321-type V Fragments.* Clones of Southern-positive EcoRI fragments were obtained from two sources; from EcoRI fragments enriched by preparative agarose gel electrophoresis, and from a partial EcoRI phage library. Most of the isolated clones were characterized by restriction site mapping, Southern gel blotting with the 5D10 probe, and electron microscopy of R-loops formed with MOPC321  $\kappa$  mRNA. On the basis of this characterization we could place each V gene clone into 1 of 10 groups (Fig. 1B). For each clone group, several independent clones were isolated. R-loop and Southern blot analysis showed that all but the 18-kb EcoRI fragment carry a single copy of a MOPC321-type V sequence; the 18-kb EcoRI fragment having two V copies separated by 12 kb. Clones containing the 2.3-kb fragment have a partial V gene copy that is cleaved by EcoRI. This explains why the 2.3-kb band is undetectable in Fig. 1A but can be seen in Fig. 4, in which the hybridization conditions were different. The 4-kb clone seems to have a short insertion in the middle of a single-copy V sequence. This was shown by R-loop mapping and Southern blotting (data not shown). R-loops of some groups of clones, like 16 kb, 9.5 kb, and 6 kb, were well opened, while those on others were smaller or closed. These morphological differences of R-loops can be attributed to varying homology between the V sequences of the clones and that of the MOPC321 cDNA probe. This strengthens the argument that the variations in band intensity, demonstrated in our Southern blot experiments, are at least in part due to differences in sequence homology.

*V $\kappa$ -21 Germ Line Genes are Clustered in the Genome.* To determine the organization of V $\kappa$ -21 germ line genes, we cloned a set of overlapping chromosomal DNA fragments by screening a partial EcoRI and EcoRI\* library with the nick-translated plasmid 5D10 as a probe. In all, 87 clones were isolated and the overlaps between these clones were determined by "R-hybrid" (see Materials and Methods), heteroduplex, Southern blot analysis, and restriction mapping. In two cases (GH20 and GHP53; KE112 and GH21), DNA sequencing was applied to determine whether segments in two different clones were in fact identical. The conclusions drawn from this work are shown in Fig. 2. Within 90 kb of chromosomal DNA there is a cluster of six V $\kappa$ -21 genes, all in the same transcriptional orientation. The order and distances of these genes are 5' - "1.5-kb V $\kappa$ " - 14 kb - "4-kb V $\kappa$ " - 12 kb - "16-kb V $\kappa$ " - 18 kb - "6-kb V $\kappa$ " - 21 kb - "9.5-kb V $\kappa$ " - 7 kb - "8.5-kb V $\kappa$ " - 3'. Below are some of the characteristics of the clones that led to this conclusion.

Several clones, such as clone GH20, bear the 1.5-kb V-containing EcoRI fragment and three other fragments of 9 kb, 1 kb, and >1 kb, the last of which is connected with the short phage arm (Fig. 2). This last fragment partially overlaps the 2.6-kb fragment of clone GHP53. The GHP53 clone carries a V gene segment on a 4-kb fragment. Nucleotide sequences of the overlapping



**FIGURE 2.** Linkage arrangement of cross-hybridizable  $V_{\kappa-21}$  V gene clones. (1A) The linkage map of six MOPC321 cross-hybridizable V genes. The complete EcoRI cleavage map is shown. EcoRI sites are indicated with vertical lines. The V genes are indicated by black boxes. The 5'-3' orientation from each gene is from left to right. The distance between two V genes is indicated above the linkage map. Below the linkage map the different V-gene clone types are indicated. (1B) Cloned BALB/c mouse DNA fragment. The inserts arranged to show their overlapping regions. EcoRI sites and genes are indicated as in A. EcoRI partial clones are named GHP, partial EcoRI\* clones GH, and the complete EcoRI clone KE. (2A) Linkage map of three MOPC321 cross-hybridizable V genes. The complete EcoRI cleavage map is shown. V genes, V gene orientation, V gene distances, and V gene clone types are indicated as in 1A. (2B) Cloned BALB/c mouse DNA fragments. Inserts are arranged to show their overlapping region. EcoRI sites and V genes are indicated as in 1A. EcoRI partial and EcoRI\* clones are indicated as in 1B. The clone 13A6 is a complete EcoRI clone. (3) Partial EcoRI clone containing a 1.6-kb V-containing fragment that could not be linked to either of the two  $V_{\kappa-21}$  gene clusters. The V gene and EcoRI sites are indicated as in 1A.

segments from both clones were partially determined, and it was found that the sequences are not completely identical. Nevertheless, we believe that the 1.5-kb and 4-kb type V genes are only 14 kb apart from each other; the reasons for this will be discussed below.

The clones GHP11 and KE112 bear the 16-kb V-containing EcoRI fragment and GHP11 shares with GHP53 a 1.2-kb EcoRI fragment. This was shown by a Southern blotting experiment using the nick-translated 1.2-kb fragment from GHP11 as a probe (Fig. 3). The probe hybridized to a lesser extent with other fragments of other clones derived from different regions of our linkage map. This is due to sequence homologies in the flanking regions of related V genes. Clone GH21 and GH13 includes a V gene-containing a 6-kb fragment and four other fragments at 1.3, 3.4, 1.1, and 0.8 kb. The 1.3-kb and 0.8-kb fragments must have been created from the EcoRI\* site by joining with the phage arms because they could not be detected in EcoRI partial clones. The 1.3-kb fragment of clone GH21 and GH13 overlaps with clone KE112 because the nucleotide sequence around the common EcoRI site was completely identical.

Clone GHP25 also bears a 6-kb, V-containing fragment and two other fragments of 2.5 kb and 5.8 kb. Clone GHP34 contains this 5.8-kb fragment and another 9.5-kb fragment carrying a V gene. This was shown by 5.8-kb hetero-

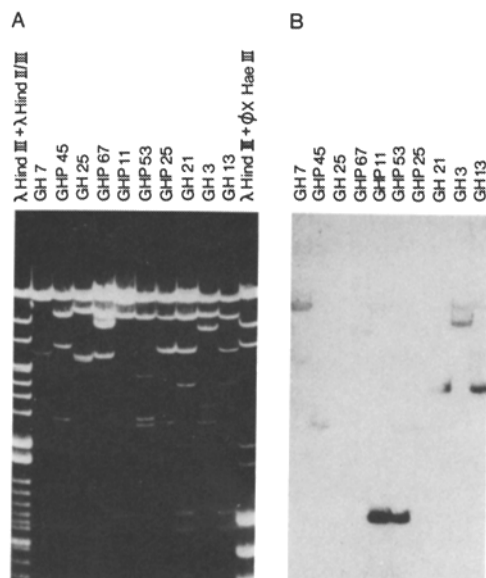


FIGURE 3. Detection of overlapping clones by Southern blot hybridization. (A) 0.5  $\mu$ g phage DNA was digested with EcoRI electrophoresed on a 1% agarose gel and stained with ethidium bromide (1  $\mu$ g/ml). (B) DNA fragments were blotted to a nitrocellulose membrane filter and hybridized with a nick-translated 1.2-kb large EcoRI fragment isolated from clone GHP11 and indicated in Fig. 2 by a shaded box.

duplexes formed between both clones.

R-loop analysis shows that clone GHP19 carries two V genes located 7 kb from each other. These genes are carried on 8.5-kb and 9.5-kb EcoRI fragments. Clone GHP67 bears the 8.5-kb, V-containing EcoRI fragment and another fragment of 7 kb.

As mentioned above the distance between the "6-kb V $\kappa$ " and the "9.5-kb V $\kappa$ " is 21 kb, the longest among the five spacers in this gene cluster. To exclude the possibility that another cross-hybridizable V gene is located between these two V $\kappa$ -21 genes, we performed a series of Southern blotting experiments under low stringency conditions (32) using 5D10 as a probe and found none (38).

It was previously shown by R-loop analysis that clone 13A6 carries two V genes, located 12 kb from each other on an 18-kb EcoRI fragment (3). Clone GH6 contains a 2.4-kb fragment and two Southern positive fragments at 2.6 kb and >4 kb, the last of which is connected with a short phage arm. R-loop mapping showed that both V genes are 6 kb apart. The V gene encoded by the >4-kb fragment corresponds to the 3' V gene of clone 13A6, as shown by restriction mapping and Southern blotting. Clone GHP3 bears the 2.6-kb V-containing EcoRI fragment and two other fragments of 2.4 and 10 kb. Their order was determined by restriction enzyme analysis. Thus, the second cluster extends over 30 kb of chromosomal DNA and contains three V $\kappa$ -21 genes in the same transcriptional orientation (Fig. 2).

Clone GHP56 bears the 1.6-kb V-containing EcoRI fragment and three other fragments of 1.8 kb, 5.0 kb, and 5.3 kb. The order of these fragments was

determined by restriction mapping. This is a third "cluster" with only one  $V_{\kappa}$ -21 gene segment.

*Determination of the Relative Order of  $V_{\kappa}$ -21 Gene Clusters.* To determine the relative order of the  $V_{\kappa}$ -21 gene clusters, we performed a series of Southern blotting experiments with EcoRI-digested DNA from BALB/c embryo DNA and from two variant lines of myeloma MPC11: MPC11 LF and MPC11 LF+L. Both MPC11 variants have one abnormally rearranged "9.5 kb  $V_{\kappa}$ " (39, 40). On the other homologue MPC11 LF+L has a normally rearranged MPC11 V gene (belongs to  $V_{\kappa}$ -19 group) and as a consequence the entire  $V_{\kappa}$ -21 group V genes are deleted. The other variant of MPC11, MPC11 LF, has lost this homologue altogether. We subjected the digested DNA to prolonged electrophoretic separation and analyzed it using the whole 5D10 plasmid as a probe under washing conditions of low stringency. Under these hybridization conditions, we could detect three additional bands at a size of 1.8, 2.3, and 2.5 kb in the BALB/c embryo DNA digests (Fig. 4) that were not apparent in previous experiments performed under hybridization conditions of higher stringency (Fig. 1 and other data not shown). The 9.5-, 8.5-, 2.6-, and 18-kb Southern-positive bands were not detectable in the DNA digest of the two MPC 11 variants, but the 1.5-, 1.6-, 4-, 6-, 16-, and 17-kb fragments were detectable. (In Fig. 4, the faint bands in the MPC11 digests are due to DNA from non-MPC11 cells contaminating the solid tumor). The 17-kb band, which contains two unresolved fragments, encodes the normally and abnormally rearranged V gene and the C gene (26). These results led to three conclusions. First, the order of the  $V_{\kappa}$ -21 genes in the 90-kb cluster is consistent with the deletion pattern. Second, the "1.6-kb  $V_{\kappa}$ " is probably located to the 5' side of the 90-kb cluster. Third, the 30-kb cluster as well as the

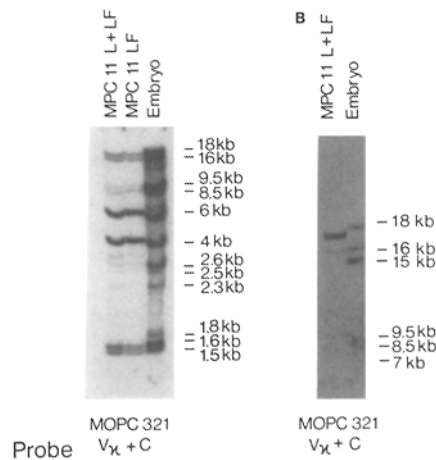


FIGURE 4. Southern blot analysis of MPC11 and BALB/c embryo DNA. (A) 10  $\mu$ g of MPC11 L+LF, MPC11 LF, and BALB/c embryo DNA were digested to completion with EcoRI, fractionated on a 1% agarose gel by electrophoresis, transferred to a nitrocellulose membrane filter, and hybridized with  $5 \times 10^7$  cpm of nick-translated 5D10 probe. Sizes of the V-gene-containing fragments are indicated in kilo bases (kb). (B) 10  $\mu$ g MPC11 L+LF and BALB/c embryo DNA were digested with EcoRI and separated on a 0.8% agarose gel over a longer distance. The transfer of the DNA fragments and the hybridization with the probe was performed as in A. Fragment size is indicated in kilo bases (kb).

"1.8 V $\kappa$ ", "2.3 V $\kappa$ ", and "2.5 kb V $\kappa$ " are located to the 3' side of the 90-kb V gene cluster.

*The Germ Line V Genes of the V $\kappa$ -21 Group.* Fig. 7 shows the nucleotide sequences of five cloned V $\kappa$  genes. The "9.5-kb V $\kappa$ " and the "16-kb V $\kappa$ " showed the greatest homology to the cDNA of MOPC321 and the rearranged V+C gene of MOPC321 of clone Ig25 $\kappa$  (3). Ig25 $\kappa$  could form well-associated heteroduplexes with only KE5e (9.5 kb, Fig. 1B) and KE112 (16 kb, Fig. 1B). The heteroduplexes were Y-shaped and the homology extended from the 5' EcoRI end to a position corresponding to the 3' end of the V $\kappa$  sequences (C. Brack, unpublished results). Restriction-map analysis of the three clones gave further indication that one of the two embryo clones contains the germ line V gene segment for MOPC321. Clones Ig25 $\kappa$ , KE5e, and KM23 contain one BamHI site in their V-coding part which is absent in KE112 (Fig. 5, A and B). Thus, the "9.5-kb V $\kappa$ " was the best candidate for the germ line V gene for MOPC321. This notion was confirmed by nucleotide sequencing of clone KE5e and KM23 using the sequencing strategy indicated in Fig. 5A. The protein sequence encoded by the sequenced V genes is exactly the same as those of five myeloma  $\kappa$  chains of subgroup C in the amino terminal 95 residues (1a in Fig. 6). The sequence of this "9.5-kb V $\kappa$ " is different from the cDNA sequence of MOPC321 (O. Bernard, unpublished results) in amino acid positions 27 (*Lys* instead of *Glu*), 27D (*Thr* instead of *Ser*), and 60 (*G* instead of *C*, with no amino acid change). The sequencing data support the view that the V region of MOPC321 is encoded by a gene that is derived by somatic mutation from the germ line "9.5-kb V $\kappa$ ".

Amino acid sequence data showed that subgroup B and subgroup C are over 94% homologous (41). Since the 9.5-kb and 16-kb fragments are very highly

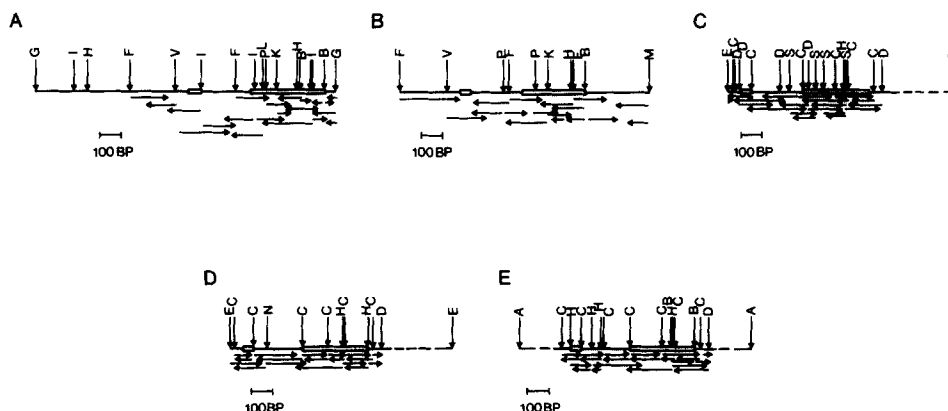


FIGURE 5. Sequencing strategy for five different V $\kappa$ -21 type V genes. The restriction sites for KM23 and KE5e were identical and the nucleotide sequence was determined in the same way for both clones. The exon that encodes most of the leader peptide and the other exon that encodes the other part of the signal peptide and the V region are indicated by an open box. Horizontal arrows indicate the direction and extent of the determined nucleotide sequence. The restriction sites in the five different clones were determined by single or double digestions. Restriction endonucleases are abbreviated as follows: A, AccI; B, BamHI; C, BstNI; D, DdeI; E, EcoRI; F, AvaII; G, BglII; H, HinfI; I, HphI; K, KpnI; L, HaeIII; M, HpaI; N, Sau3AI; P, PstI; S, Sau96I; V, PvuI. 5A: KE5e, KM23; 5B: KE112; 5C: GHP53; 5D: GHP56; 5E: 5' 13A6.



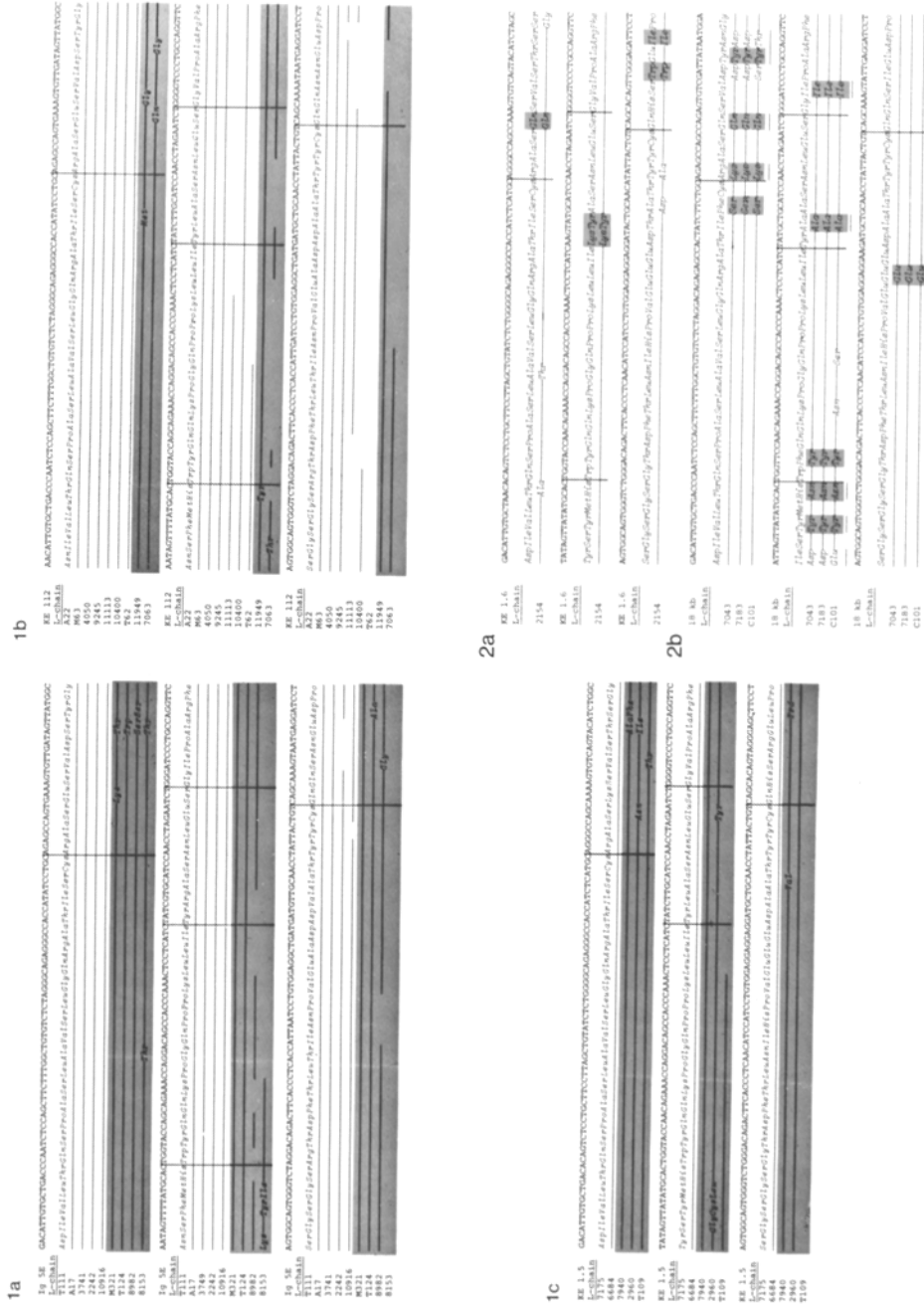


FIGURE 6. Comparison of germ line V sequences with protein sequences of different subgroups. (1) The translated amino acid sequences of KE5c, KM23 (a), KE112 (b), and GHP53 (c) were compared with the protein sequence of myeloma  $\kappa$  light-chain regions from subgroup C (a), subgroup B (b), and myeloma E (c). Variant protein sequences of subgroups C, B, and E are colored gray. The framework regions and hypervariable regions are separated by vertical lines. (2A) The amino acid sequence of clone GHP56 was compared with the NZB mouse  $\kappa$  light-chain protein sequence of myeloma 2154. Characteristic amino acid residues shared by the sequence encoded by the 1.6-kb clone type and 2154 are colored gray. (2B) The amino acid sequence

which was deduced from the nucleotide sequence of one of the two V regions of clone 13A6 (indicated in Fig. 1B by a shaded box) was compared with the NZB mouse  $\kappa$  light-chain V-region sequences, which represent the prototype sequence of subgroup D and the  $\kappa$  light-chain V-region protein sequence of BALB/c myeloma protein C101. Subgroup-specific amino acid residues are underlined and residues common to the NZB and BALB/c myeloma proteins are colored gray. The protein sequences were determined by Weigert et al. (21), McKean et al. (20), Kabat et al. (58), and by M. Weigert (personal communication).

homologous, it seemed likely that the "16-kb V $\kappa$ " encodes a member or members of subgroup B. In Fig. 6 (1*b*), this is shown to be the case: the nucleotide sequence corresponds exactly to the amino acid sequences of seven  $\kappa$  chains of subgroup B.

The "1.5-kb V $\kappa$ " shows less homology to the cDNA of MOPC321 in the Southern blot experiments and the mRNA of MOPC321 in the R-loop experiments. The nucleotide sequence indicates that this V gene segment encodes the V regions of myeloma proteins 6684 and 7175, which belong to subgroup E (Fig. 6, 1*c*).

The nucleotide sequence of the "1.6-kb V $\kappa$ " is very homologous to that of the "1.5-kb V $\kappa$ ". This was shown by heteroduplex analysis and confirmed by nucleotide sequencing (the sequencing strategy is indicated in Fig. 5*D*). The protein sequence encoded by this gene does not correspond to any of the known V $\kappa$ -21  $\kappa$  chains, but it shows high homology to the V region 2154  $\kappa$  chain (Fig. 6, 2*a*). This myeloma does not define an independent V $\kappa$ -21 subgroup because no other closely related sequence is known to date. It is likely that the sequence for this myeloma protein represents a variant of the protein sequence encoded by the "1.6-kb V $\kappa$ ".

The amino acid sequence encoded by the "5' 18-kb V $\kappa$ " (Fig. 6, 2*b*) corresponds to none of the known protein sequences. It shows similarities to the recurring sequence of subgroup D, but also differences. This germ line gene segment probably encodes a new subgroup.

### Discussion

*Genes for V $\kappa$ -21 Variable Regions.* The V regions of more than 80 myeloma  $\kappa$  chains of BALB/c and NZB origins have been classified into 49 different groups based on the amino acid sequence of the amino terminal 23 residues. Of these about 30, belonging to the V $\kappa$ -21 group, have been completely sequenced and classified into seven subgroups based on the entire V region sequence. Members of the V $\kappa$ -21 group are closely homologous by amino acid sequence and nucleic acid hybridization (41, 42). Using a probe for a given V $\kappa$ -21 subgroup i.e. V $\kappa$ -21C (MOPC 321) we should detect all genes coding for the V $\kappa$ -21 group. Indeed, EcoRI-digested embryo DNA exhibited 12 Southern positive bands with the MOPC321 cDNA probe (Fig. 4). We isolated several independent clones from the nine most strongly hybridizing of these bands. We did not try to clone the three weakly hybridizing fragments, but nevertheless one was found in a clone library. The analysis of multiple clones isolated from a single band indicated that in all cases one band represents an EcoRI fragment of a single size. All fragments contained one V $\kappa$  sequence, except the 18 kb, which encoded two. Thus, we have identified 13 gene segments homologous to MOPC321. This value is substantially higher than that determined by saturation-hybridization experiments (4-6 germ line genes [42]) or the number predicted based on the amino acid sequence analysis (7 genes for 7 subgroups [21]). The number of known subgroups is not exhaustive and further amino-acid sequencing might reveal new subgroups. On the other hand, numerous examples for inactive genes or pseudogenes have been reported in mouse and human heavy-chain V genes as well as in other gene families (11, 12, 43, 44). It is possible that one of our gene

segments, the "4-kb  $V_{\kappa}$ ," contains an inactive V gene, because this gene seems to be split atypically in the mid region. Further nucleotide sequence analysis might show that more  $V_{\kappa}$ -21 genes are inactive similar to  $V_H$  genes. Furthermore, as already shown for the  $\lambda_1$  system, saturation hybridization does not allow an exact determination of the gene number (45). In view of the above consideration, the correlation of subgroups as defined by Potter (19) with germ line V genes seems to hold well in the  $V_{\kappa}$ -21 group. This correlation was previously suggested by the analysis of the  $\lambda_1$  system.

*V-region Sequences Arise by Somatic Mutation.* The amino acid sequence variability within one  $V_{\kappa}$ -21 subgroup is similar to the variability previously observed in the V regions of  $\lambda_1$  chains (8, 9). For instance, in subgroup C, at least five independently induced myelomas, T111, A17, 3741, 2242, and 10916 produce  $\kappa$  chains with identical V regions. At least four other myelomas including MOPC321, also independently induced, were shown to synthesize  $\kappa$  chains with V regions that differ from each other and from the first type by two or more residues, mostly but not exclusively localized in the three CDR regions. We will refer to the former, recurring type V region, as the prototype and the latter uniquely occurring type V regions as variants.

In  $\lambda_1$  chains, it was previously demonstrated that the prototype V region is encoded by the germ line V gene segment and the variant V regions are encoded by the genes that arose in lymphocyte clones by introduction of somatic mutations in the germ line V gene segment (2, 8, 10). The present study extends this concept to  $V_{\kappa}$ -21 chains. Heteroduplex and restriction enzyme analysis indicated that the "9.5-kb  $V_{\kappa}$ " is the germ line counterpart of the rearranged  $\kappa$  gene active in myeloma MOPC321 belonging to subgroup C of the  $V_{\kappa}$ -21 group. However, the nucleotide sequence of two independent isolated "9.5-kb  $V_{\kappa}$ " differ from that of the corresponding region of MOPC321 cDNA in three nucleotides, one each in codons for residues 27, 27D, and 60 (O. Bernard, unpublished results). On the other hand, the germ line "9.5-kb  $V_{\kappa}$ " sequence corresponds exactly to the amino-acid sequence of the prototype V region of subgroup C (Fig. 6, 1a). This result indicates that the V gene of MOPC321 arises by somatic mutation from the germ line "9.5-kb  $V_{\kappa}$ ".

The amino acid sequence homology between subgroups B and C prototype sequences is >95% (41). Our analysis of the 9.5-kb clone and the 16-kb clone showed that the two are highly homologous. Therefore, it is likely that the 16-kb clone encodes the prototype sequence of subgroup B. The nucleotide sequence of the "16-kb  $V_{\kappa}$ " proved the validity of this assumption (Fig. 6, 1b). It encodes the V region of seven identical but independently isolated myeloma  $\kappa$  chains from BALB/c and NZB mice. The other  $\kappa$  chains of this subgroup would be encoded by somatic variants of this germ line V gene.

The third example of the correlation between a germ line V gene and a prototype V region can be seen with "1.5-kb  $V_{\kappa}$ " and subgroup E. Again the nucleotide sequence of this germ line gene corresponds exactly to the amino acid sequences of two independently induced myelomas, 7175 and 6684. In addition to the two prototype V regions at least three variant V regions are known for subgroup E (Fig. 6, 1c).

The prototype sequences of subgroups E and F are about as homologous to

each other as are those of the subgroups B and C (41). Since the nucleotide sequence of "1.5-kb V $\kappa$ " (which encodes subgroup E) and "1.6-kb V $\kappa$ " are also very homologous, we expected the V gene of the latter clone to encode subgroup F. However, the nucleotide sequence of the "1.6-kb V $\kappa$ " did not correspond to the protein sequences of subgroup F or of any other known V $\kappa$ -21 V regions. Nevertheless, the protein sequence predicted from this germ line V gene is very similar to the V region of the NZB myeloma 2154. It differs from it in framework amino-acid residues 5, 12, 84, and 86 as well as in the first hypervariable region in residue 29, but it shares characteristic amino acids at residues 27, 49, 50, 92, and 94 (Fig. 6, 2a). We believe that "1.6-kb V $\kappa$ " is the germ line V gene from which the V $\kappa$  gene of myeloma 2154 arose by somatic mutations.

The nucleotide sequence of one of the 18-kb type V genes ("5' 18-kb V $\kappa$  gene") also does not encode any of the known V $\kappa$ -21 V regions. However, the protein sequence predicted by this germ line V gene contains several amino acid residues at defined positions that are characteristic for subgroup D (Fig. 6, 2b). Despite this, it is unlikely that the members of this subgroup arose from this germ line V gene by somatic mutation. There exist two independently induced NZB myelomas, 7043 and 7183, that synthesize  $\kappa$  chains with identical V regions that differ from the amino acid sequence predicted from the "5' 18-kb V gene". Since it is unlikely that two  $\kappa$  chains sharing exactly the same V region arise by somatic mutations from a common germ line V gene, 7043 and 7183 V regions define the subgroup D prototype sequence, for NZB at least. The V region of BALB/c myeloma C101 shares the characteristic amino acids with the prototype V-region of subgroup D at residues 22, 24, 34, and 36. On the other hand, all three differ from the "5' 18-kb" residues in the same positions (Fig. 6, 2b). Perhaps the subgroup defined by the "5' 18-kb V gene" has not yet been discovered, or perhaps this V gene is not expressed.

In summary, the results show that the three sequenced germ line V genes of the 9.5-kb, 16-kb, and 1.5-kb types encode the prototype sequences of subgroups V $\kappa$ -21C, V $\kappa$ -21B, and V $\kappa$ -21E. All other protein sequences in these three subgroups differ by one or two base pair changes from these three germ line genes. Somatic mutation is not restricted to the hypervariable regions; it also occurs in the framework regions. The generation of different V regions by somatic mutation also occurs in the  $\lambda_1$  V regions as well as in the heavy chain V regions (46). Transitions and transversions occur with approximately equal frequency. Among the transitions, purine exchanges are three times more frequent than pyrimidine exchanges, and among transversions, purine to pyrimidine transitions are seven times more frequent than those in the opposite direction. These statistics more likely reflect idiosyncrasies of the DNA repair enzymes involved in the mutation process than the mutant selection process.

*V Genes for V $\kappa$ -21 Are Linked in the Genome.* Although the linkage arrangement of heavy-chain constant-region immunoglobulin genes has been determined (31, 47), and limited data are available on how V<sub>H</sub> genes are organized (11, 12, 22), there is no data in the literature on the arrangement of V $\kappa$  genes. We attempted to determine the arrangement of the germ line V genes for the V $\kappa$ -21 group by molecular cloning of overlapping fragments. These fragments were shown to define regions of 90 kb and 30 kb, containing six and three V genes, respectively.

The linkage arrangement supports our contentions that most of the Southern-positive bands contain a DNA fragment of one size and that the  $V_{\kappa}$ -21 group is probably encoded by not more than 13 germ line V genes.

The nucleotide sequences of two partially overlapping EcoRI and EcoRI\* clones containing the "4-kb  $V_{\kappa}$ " and the "1.5-kb  $V_{\kappa}$ ", respectively, revealed six differences in a stretch of 130 base pairs. Nevertheless, we placed the latter immediately 5' to the former. These sequence differences can best be explained by a lack of inbreeding or by substrain differences; the two EcoRI libraries were constructed from different BALB/c embryo DNA preparations. The fact that we could isolate only one 4-kb clone type out of more than 80 EcoRI partial clones indirectly supports the gene linkage. Also, our linkage map shows that at the 5' and 3' ends of the 4-kb-type V gene are two large EcoRI fragments that cannot be cloned together with this V gene in phage Charon 4A.

It has been suggested that rearrangement of V genes leads to the deletion of DNA that is located at the 3' side of the rearranged V gene and the 5' side of the rearranged J segment (33, 40). The Southern gel blot analysis of the two variant cell lines of MPC11, MPC11 L+LF, and MPC11 LF provided some supporting evidence for the order of the various  $V_{\kappa}$ -21 genes in the 90-kb cluster and, in addition, allowed us to produce the tentative order of  $V_{\kappa}$ -21 genes. According to these results, the probable order is, from 5' to 3', "1.6-kb  $V_{\kappa}$ ", "1.5-kb  $V_{\kappa}$ ", "4.0-kb  $V_{\kappa}$ ", "16-kb  $V_{\kappa}$ ", "6.0-kb  $V_{\kappa}$ ", "9.5-kb  $V_{\kappa}$ ", and "8.5-kb  $V_{\kappa}$ ". In addition, the data suggests that the 30-kb cluster containing "5' 18-kb  $V_{\kappa}$ ", "3' 18-kb  $V_{\kappa}$ ", and "2.6-kb  $V_{\kappa}$ " is located further downstream (i.e., 3' side). "2.5-kb  $V_{\kappa}$ ", "2.3-kb  $V_{\kappa}$ ", and "1.8-kb  $V_{\kappa}$ " also appear to be located further downstream. We believe that the "18-kb  $V_{\kappa}$ " are at some distance from the "8.5-kb  $V_{\kappa}$  gene" because none of our EcoRI\* clones containing the latter V gene overlapped with the 18-kb clone. Perhaps the "1.8-kb  $V_{\kappa}$ ", "2.3-kb  $V_{\kappa}$ ", and/or "2.5-kb  $V_{\kappa}$ " are located between the 90-kb and 30-kb V gene clusters.

*Sequence Comparison of Germ Line V Genes.* Comparison of the nucleotide sequences of the five germ line V genes reveals several common features (Fig. 7). All V regions are encoded by two exons that are separated by an intron. One exon encodes the majority of the leader peptide, while the other encodes the rest of the leader peptide and the variable region. In all V regions of heavy and light chains sequenced to date, the leader peptide is interrupted by an intron at amino acid position -4. The introns of the five sequenced  $V_{\kappa}$  genes are all ~240 base pairs long. However, the intron of the 18-kb clone is 12 base pairs shorter. In the relevant area (see Fig. 7) there are two short, direct repeats of the sequence AAGG separated by 12 bases. The deletion removes one of the repeats entirely but not the other. A similar small repeated sequence borders the deletions in introns of heavy-chain variable regions and  $\beta$ -like globin (11, 48). As proposed by Farabaugh et al. (49), these short, direct repeats probably promote deletions by a slippage mispairing during DNA replication.

We compared the nucleotide sequence of the  $V_{\kappa}$ -21 germ line genes with each other and calculated the percentage of silent and replacement substitutions according to the procedures of Lomedico et al. (50). Table I shows the results of these calculations from which several conclusions can be drawn. First, the nucleotide sequence of the leader exon is highly preserved; only two silent

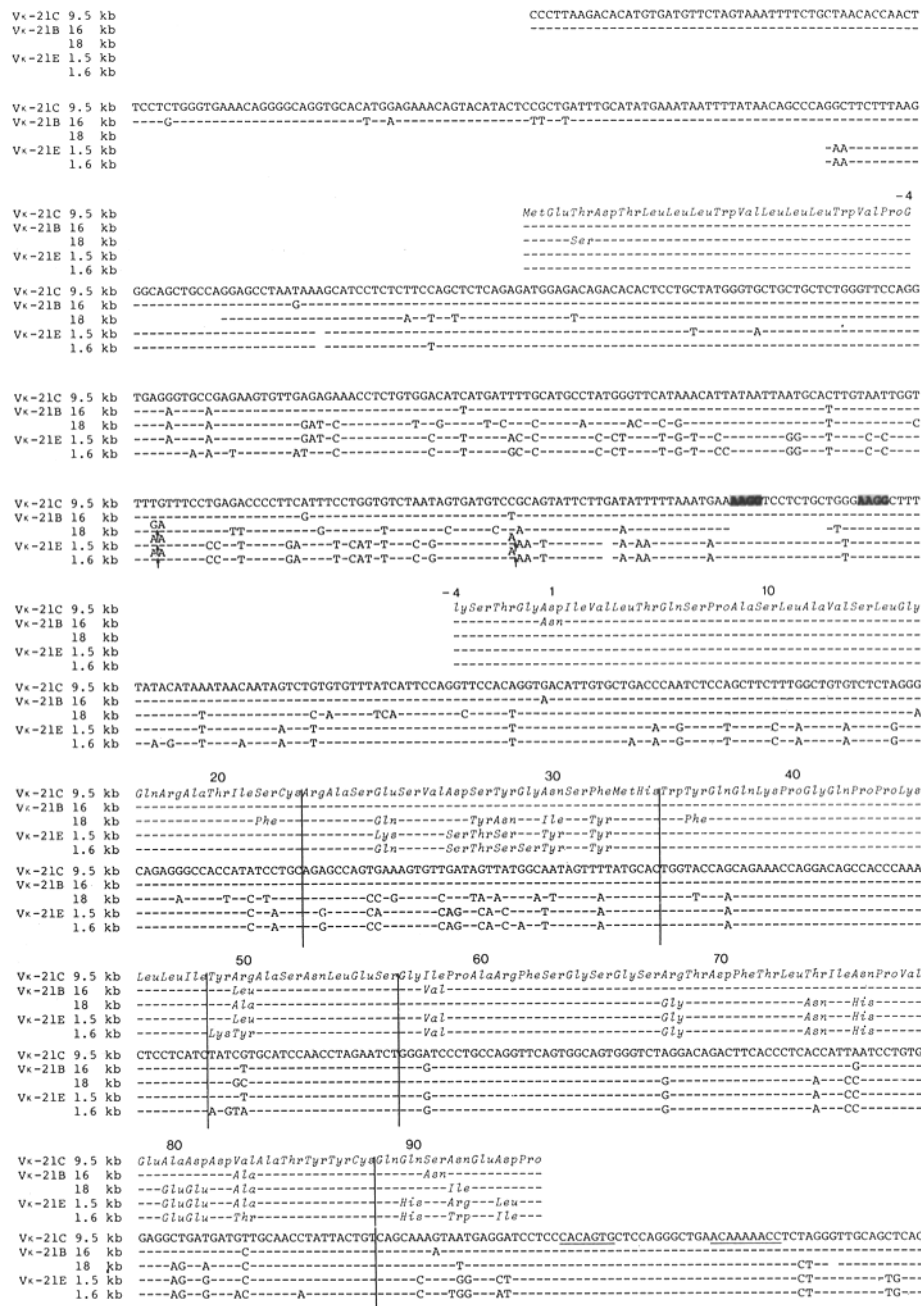


FIGURE 7. Nucleotide sequence of germ line Vκ-21 regions. The nucleotide sequences were determined by the method of Maxam and Gilbert. The sequence of the 9.5-kb clone type V gene was used as a reference sequence for comparison with the others. A dash (—) indicates sequence identity with the 9.5-kb clone type V gene. Amino acids translated from the nucleotide sequence are numbered from the beginning of the mature light chain according to the nomenclature of Kabat et al. (58). The order of the amino acid sequences corresponds to the order of the nucleotide sequence, beginning with the amino acid sequence for the 9.5-kb type V gene. The intron begins with the nucleotides GT and ends with AG so that the first nucleotide of the glycine codon is encoded by a different exon from the second and third nucleotides. The framework regions and hypervariable regions are separated by vertical lines. The recognition sequences involved in V-J joining are underlined. Short, direct repeats which promote deletions are indicated by gray. Insertions are indicated by arrows.

TABLE I  
Comparison of Germ Line  $V_{\kappa}$ -21 Nucleotide Sequences

Compared nucleotide sequences	Leader		Intron	FR1		CDR1		FR2		CDR2		FR3		CDR3	
	s	r		s	r	s	r	s	r	s	r	s	r	s	r
	9.5 × 16	0	0	3 ± 1	0	2 ± 2	0	0	0	0	0	6 ± 6	4 ± 4	3 ± 2	0
9.5 × 18	0	3 ± 3	23 ± 3	20 ± 10	2 ± 2	36 ± 18	18 ± 7	9 ± 9	3 ± 3	0	13 ± 9	4 ± 4	10 ± 4	0	7 ± 7
9.5 × 1.6	0	0	23 ± 3	50 ± 16	8 ± 4	36 ± 18	24 ± 8	9 ± 9	6 ± 4	0	13 ± 9	8 ± 6	13 ± 4	0	27 ± 13
9.5 × 1.5	16 ± 12	0	21 ± 3	45 ± 15	0	36 ± 18	21 ± 8	9 ± 9	0	0	6 ± 6	4 ± 4	11 ± 4	0	20 ± 12
16 × 18	0	3 ± 3	23 ± 3	30 ± 12	3 ± 2	36 ± 18	18 ± 7	9 ± 9	3 ± 3	0	13 ± 9	4 ± 4	10 ± 4	0	13 ± 9
16 × 1.6	0	0	23 ± 3	55 ± 16	2 ± 2	27 ± 15	27 ± 9	9 ± 9	6 ± 4	0	13 ± 9	13 ± 7	8 ± 3	0	47 ± 18
16 × 1.5	16 ± 12	0	20 ± 3	50 ± 16	2 ± 2	36 ± 18	21 ± 8	9 ± 9	0	0	0	8 ± 6	7 ± 3	0	40 ± 16
18 × 1.6	0	3 ± 3	33 ± 4	60 ± 17	3 ± 2	27 ± 15	30 ± 9	0	9 ± 5	0	13 ± 9	8 ± 6	3 ± 2	0	40 ± 16
18 × 1.5	16 ± 12	3 ± 3	28 ± 4	60 ± 17	2 ± 2	36 ± 18	27 ± 9	0	3 ± 3	0	13 ± 9	4 ± 4	1 ± 1	0	33 ± 15
1.5 × 1.6	16 ± 12	0	5 ± 1	5 ± 5	0	0	6 ± 4	0	6 ± 4	0	13 ± 9	4 ± 4	1 ± 1	0	13 ± 9

The nucleotide sequence of all determined  $V_{\kappa}$ -21 germ line sequences were compared with each other. Each change was classified as a replacement substitution (r) if it changed the amino acid, or as a silent substitution (s). The division of the observed number of silent changes by the number of silent sites (25% of total number of sites in the different leader, intron, framework, and hypervariable regions) yields the number of silent changes indicated in %. Replacement substitutions have been calculated as the observed number of replacement sites (75% of the total number of sites in the different parts of the V region). Each group in the intron sequence was scored as one change. The errors shown are the Poisson standard deviations.

TABLE II  
Comparison of  $V_{\kappa}$ -21 Germ Line Nucleotide Sequences with Other Germ Line  $V_{\kappa}$  Sequences

Compared nucleotide sequences	Leader	Intron	FR1	CDR1	FR2	CDR2	FR3	CDR3
9.5 × M41	33 ± 16	—	70 ± 19	97 ± 34	53 ± 22	57 ± 33	42 ± 13	20 ± 20
9.5 × M149	—	—	70 ± 19	73 ± 30	53 ± 22	0	58 ± 16	20 ± 20
9.5 × M167	49 ± 20	—	70 ± 19	80 ± 36	53 ± 22	57 ± 33	50 ± 14	40 ± 28
18 × M41	16 ± 12	—	55 ± 17	97 ± 34	62 ± 24	57 ± 33	42 ± 13	20 ± 20
18 × M149	—	—	75 ± 19	73 ± 30	44 ± 20	0	58 ± 16	20 ± 20
18 × M167	73 ± 24	—	60 ± 17	71 ± 25	62 ± 24	38 ± 27	58 ± 16	40 ± 28
1.5 × M41	33 ± 16	—	55 ± 17	97 ± 34	62 ± 24	57 ± 33	46 ± 14	40 ± 28
1.5 × M149	—	—	45 ± 15	85 ± 32	44 ± 20	0	63 ± 16	40 ± 28
1.5 × M167	82 ± 26	—	60 ± 17	71 ± 25	53 ± 22	57 ± 33	54 ± 15	40 ± 28
M41 × M149	—	—	25 ± 11	61 ± 27	44 ± 20	76 ± 38	33 ± 12	20 ± 20
M41 × M167	33 ± 16	—	40 ± 14	73 ± 30	18 ± 16	95 ± 43	63 ± 16	20 ± 20
M149 × M167	—	—	60 ± 17	61 ± 27	53 ± 22	38 ± 27	54 ± 15	0

Silent substitutions were calculated as described in Table I and are indicated in %. The nucleotide sequence of the germ line  $V_{\kappa}$  sequences of MOPC149 ( $V_{\kappa}$ -13) was determined by Seidman et al. (57), of MOPC41 ( $V_{\kappa}$ -9) by Seidman et al. (4), and MOPC167 ( $V_{\kappa}$ -24) by Selsing and Storb (15).

mutations and one replacement mutation are observed (Fig. 7 and Table I). Secondly, the similarity of the "1.5-kb  $V_{\kappa}$ " to the "1.6-kb  $V_{\kappa}$ " and the "9.5-kb  $V_{\kappa}$ " to the "16-kb  $V_{\kappa}$ " suggests that these genes arose by recent gene duplications. Third, replacement substitutions are generally more frequent in the CDR regions than in framework regions. Finally, silent substitutions are significantly more frequent in the intron, FR1, and CDR1 than in the rest of the V gene subregions. As shown in Table II, this unequal distribution of silent substitutions is not observed when two germ line V genes belonging to two different  $V_{\kappa}$  groups are compared. The unequal distribution of silent substitutions cannot result from selection since they are neutral at the protein level. This phenomenon might be attributed either to correction mechanisms such as gene conversion (51-54) or

to homologous but unequal crossing over (55). Although we cannot exclude the possibility of gene conversion, several facts suggest that the homology among the V $\kappa$ -21 genes within the BALB/c mouse is due to unequal crossing over: First, the flanking regions of these genes exhibit extensive sequence homology. Second, unequal crossing over changes the number of V genes and leads to their rearrangement. Such rearrangement of V $\kappa$ -21 genes can be found in different mouse strains (M. Weigert, personal communication). The distribution of silent substitutions in the V $\kappa$ -21 genes implies that the presumed unequal crossing over occurred between CDR1 and FR2 and was followed by expansion of the V $\kappa$ -21 gene pool. The conclusion can be drawn that unequal crossing over is responsible for the segmental homologies among linked V $\kappa$ -21 genes encoding a group of V regions as it was shown for the  $\alpha$  globulin genes in humans (56).

### Summary

Using a cloned cDNA of a mouse immunoglobulin  $\kappa$  light chain synthesized in a myeloma MOPC321 (V $\kappa$ -21 subgroup C) as a probe we could detect 13 germ line V $\kappa$  gene segments. 11 of these were isolated. Using a set of overlapping cloned segments, we showed that nine of these germ line V $\kappa$  genes are arranged in two linkage clusters and that they all have the same transcriptional orientation (11, 12, 22). These two clusters occupy 90 and 30 kb of chromosomal DNA and contain six and three V $\kappa$ 's, respectively. We determined the complete nucleotide sequences of five germ line V $\kappa$ 's and showed that three of them encode the prototype sequence of V $\kappa$ -21 subgroups B, C, and E. None of these five germ line V $\kappa$ 's encodes the variant amino acid sequences of known V $\kappa$ -21 subgroups. We thus conclude that, as in the  $\lambda_1$  light chains, the variant V regions are encoded by gene segments derived by a few somatic mutations from the corresponding germ line DNA. Such somatic mutations are not restricted to sequences encoding the hypervariable regions; they also occur in sequences encoding framework regions.

We thank Drs. M. Gefter, C. Steinberg, and T. Imanishi-Kari for critically reading the manuscript. In particular we are grateful to Dr. M. Weigert for unpublished protein sequence data. We also thank Dr. H. G. Zachau for MPC11 L+LF and MPC11 LF DNA.

Thanks go to Laura Moser and Lynn Gisler for the superb job of typing this manuscript.

*Received for publication 15 September 1983.*

### References

1. Tonegawa, S., C. Brack, N. Hozumi, and R. Schuller. 1977. Cloning of an immunoglobulin variable region gene from mouse embryo. *Proc. Natl. Acad. Sci. USA.* 74:3518.
2. Brack, C., M. Hirama, R. Lenhard-Schuller, and S. Tonegawa. 1978. A complete immunoglobulin gene is created by somatic recombination. *Cell.* 15:1.
3. Lenhard-Schuller, R., B. Hohn, C. Brack, M. Hirama, and S. Tonegawa. 1978. DNA clones containing mouse immunoglobulin  $\kappa$  chain genes isolated by in vitro packaging into phage  $\lambda$  coats. *Proc. Natl. Acad. Sci. USA.* 75:4709.
4. Seidman, J. G., E. E. Max, and P. Leder. 1979. A  $\kappa$ -immunoglobulin gene is formed by site-specific recombination without further somatic mutation. *Nature (Lond.)* 280:370.
5. Wu, T. T., and E. A. Kabat. 1970. An analysis of the sequences of the variable



- regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* 132:211.
6. Poljak, R. J., L. M. Amzel, H. P. Avey, B. L. Chen, R. P. Phizackerley, and F. Saul. 1973. Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2.8-Å resolution. *Proc. Natl. Acad. Sci. USA.* 70:3305.
  7. Padlan, E. A., D. M. Segal, T. F. Spande, D. R. Davies, S. Rudikoff, and M. Potter. 1973. Structure at 4.5 Å resolution of a phosphorylcholine-binding Fab. *Nature New Biol.* 245:165.
  8. Weigert, M. G., T. M. Cesari, S. J. Yonkovich, and M. Cohn. 1970. Variability in the lambda light chain sequences of mouse antibody. *Nature (Lond.)* 228:1045.
  9. Cohn, M., R. Langman, and W. Geckeler. 1980. Diversity 1980. *In Immunology* 80. M. Fougereau and J. Dausset, editors. Academic Press, New York. 153.
  10. Bernard, O., N. Hozumi, and S. Tonegawa. 1978. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell.* 15:1133.
  11. Bothwell, A. L. M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D. Baltimore. 1981. Heavy chain variable region contribution to the NP<sup>b</sup> family of antibodies: somatic mutation evident in a  $\gamma$ 2a variable region. *Cell.* 24:625.
  12. Crews, S., J. Griffin, H. Huang, K. Calame, and L. Hood. 1981. A single V<sub>H</sub> gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. *Cell.* 25:59.
  13. Tonegawa, S., C. Steinberg, S. Dube, and A. Bernardini. 1974. Evidence for somatic generation of antibody diversity. *Proc. Natl. Acad. Sci. USA.* 71:4027.
  14. Tonegawa, S. 1976. Reiteration frequency of immunoglobulin light chain genes: further evidence for somatic generation of antibody diversity. *Proc. Natl. Acad. Sci. USA.* 73:203.
  15. Selsing, E., and U. Storb. 1981. Somatic mutation of immunoglobulin light-chain variable-region genes. *Cell.* 25:47.
  16. Gershenfeld, H. K., A. Tsukamoto, I. L. Weissman, and R. Joho. 1981. Somatic diversification is required to generate the V<sub>κ</sub> genes of MOPC 511 and MOPC 167 myeloma proteins. *Proc. Natl. Acad. Sci. USA.* 78:7674.
  17. Loh, E., B. Black, R. Riblet, M. Weigert, J. M. Hood, and L. Hood. 1979. Myeloma proteins from NZB and BALB/c mice: structural and functional differences. *Proc. Natl. Acad. Sci. USA.* 76:1395.
  18. Morse, H. C. III, R. Riblet, R. Asofsky, and M. Weigert. 1978. Plasmacytomas of the NZB mouse. *J. Immunol.* 121:1969.
  19. Potter, M. 1977. Antigen-binding myeloma proteins of mice. *Adv. Immunol.* 25:141.
  20. McKean, D. J., M. Bell, and M. Potter. 1978. Mechanisms of antibody diversity: multiple genes encode structurally related mouse  $\kappa$  variable regions. *Proc. Natl. Acad. Sci. USA.* 75:3913.
  21. Weigert, M., L. Gatmaitan, E. Loh, J. Schilling, and L. Hood. 1978. Rearrangement of genetic information may produce immunoglobulin diversity. *Nature (Lond.)* 276:785.
  22. Kemp, D. J., B. Tyler, O. Bernard, N. Gough, S. Gerondakis, J. M. Adams, and S. Cory. 1981. Organization of genes and spacers within the mouse immunoglobulin V<sub>H</sub> locus. *J. Mol. Appl. Genet.* 1:245.
  23. Blattner, F. R., B. G. Williams, A. E. Blechl, K. Denniston-Thompson, H. E. Faber, L.-A. Furlong, D. J. Grunwald, D. O. Kiefer, D. D. Moore, J. W. Schumm, E. L. Sheldon, and O. Smithies. 1977. Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. *Science (Wash. DC)* 196:161.
  24. Leder, P., D. Tiemeier, and L. Enquist. 1977. EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the  $\lambda$ gtWES system.

- Science (Wash. DC)*. 196:175.
25. Collins, J., and B. Hohn. 1978. Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage  $\lambda$  heads. *Proc. Natl. Acad. Sci. USA*. 75:4242.
  26. Schnell, H., M. Steinmetz, and H. G. Zachau. 1980. An unusual translocation of immunoglobulin gene segments in variants of the mouse myeloma MPC11. *Nature (Lond.)*. 286:170.
  27. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503.
  28. Wahl, G. M., M. Stern, and G. R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. *Proc. Natl. Acad. Sci. USA*. 76:3683.
  29. Matthyssens, G., and T. H. Rabbitts. 1980. Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region. *Proc. Natl. Acad. Sci. USA*. 77:6561.
  30. Maki, R., A. Traunecker, H. Sakano, W. Roeder, and S. Tonegawa. 1980. Exon shuffling generates an immunoglobulin heavy chain gene. *Proc. Natl. Acad. Sci. USA*. 77:2138.
  31. Roeder, W., R. Maki, A. Traunecker, and S. Tonegawa. 1981. Linkage of the four  $\gamma$  subclass heavy chain genes. *Proc. Natl. Acad. Sci. USA*. 78:474.
  32. 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.
  33. Sakano, H., K. Hüppi, G. Heinrich, and S. Tonegawa. 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature (Lond.)*. 280:288.
  34. Brack, C. 1981. DNA electron microscopy. *Crit. Rev. Biochem.* 10:113.
  35. Sakano, H., J. H. Rogers, K. Hüppi, C. Brack, A. Traunecker, R. Maki, R. Wall, and S. Tonegawa. 1979. Domains and the hinge region of an immunoglobulin heavy chain are encoded in separate DNA segments. *Nature (Lond.)*. 277:627.
  36. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499.
  37. Bentley, D. L., and T. H. Rabbitts. 1981. Human V $\kappa$  immunoglobulin gene number: implications for the origin of antibody diversity. *Cell*. 14:613.
  38. Heinrich, G. 1982. Structure and organization of immunoglobulin light-chain genes. Ph.D. thesis, University of Basel, Switzerland.
  39. Rose, S. M., W. M. Kuehl, and G. P. Smith. 1977. Cloned MPC 11 myeloma cells express two kappa genes: a gene for a complete light chain and a gene for a constant region polypeptide. *Cell*. 12:453.
  40. Seidman, J. G., and P. Leder. 1980. A mutant immunoglobulin light chain is formed by aberrant DNA- and RNA-splicing events. *Nature (Lond.)*. 286:779.
  41. Julius, M. A., D. J. McKean, M. Potter, and M. Weigert. 1981. Expression of kappa chains of the V $\kappa$ -21 group in *Mus musculus* and related species. *Mol. Immunol.* 18:11.
  42. Valbuena, O., K. B. Marcu, M. Weigert, and R. P. Perry. 1978. Multiplicity of germ line genes specifying a group of related mouse  $\kappa$  chains with implications for the generation of immunoglobulin diversity. *Nature (Lond.)*. 276:780.
  43. Givol, D., R. Zakut, K. Effron, G. Rechavi, D. Ram, and J. B. Cohen. 1981. Diversity of germ-line immunoglobulin V<sub>H</sub> genes. *Nature (Lond.)*. 292:426.
  44. Proudfoot, N. 1980. Pseudogenes. *Nature (Lond.)*. 286:840.
  45. Rabbitts, T. H. 1979. The organization of immunoglobulin genes. In *Defense and Recognition IIB-Structural Aspects*. *Int. Rev. Biochem.* 23:49.
  46. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature (Lond.)*. 302:575.
  47. Shimizu, A., N. Takahashi, Y. Yamawaki-Kataoka, Y. Nishida, T. Kataoka, and T.

- Honjo. 1981. Ordering of mouse immunoglobulin heavy chain genes by molecular cloning. *Nature (Lond.)*. 289:149.
48. Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot. 1980. The structure and evolution of the human  $\beta$ -globin gene family. *Cell*. 21:653.
  49. Farabaugh, P. J., U. Schmeissner, M. Hofer, and J. H. Miller. 1978. Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli*. *J. Mol. Biol.* 126:847.
  50. Lomedico, P., N. Rosenthal, A. Efstratiadis, W. Gilbert, R. Kolodner, and R. Tizard. 1979. The structure and evolution of the two nonallelic rat preproinsulin genes. *Cell*. 18:545.
  51. Engel, R. 1981. Intergenic conversion and reiterated genes. *Nature (Lond.)*. 290:191.
  52. Baltimore, D. 1981. Gene conversion: some implications for immunoglobulin genes. *Cell*. 24:592.
  53. Dildrop, R., M. Brüggemann, A. Radbruch, K. Rajewsky, and K. Beyreuther. 1982. Immunoglobulin V region variants in hybridoma cells. II. Recombination between V genes. *EMBO (Eur. Mol. Biol. Org.) J.* 1:635.
  54. Schreier, H. P., A. L. M. Bothwell, B. Mueller-Hill, and D. Baltimore. 1981. Multiple differences between the nucleic acid sequences of the IgG2a<sup>a</sup> and IgG2a<sup>b</sup> alleles of the mouse. *Proc. Natl. Acad. Sci. USA*. 78:4495.
  55. Smith, G. P. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science (Wash. DC)*. 191:528.
  56. Lauer, J., C.-K. J. Schen, and T. Maniatis. 1980. The chromosomal arrangement of human  $\alpha$ -like globin genes: sequence homology and  $\alpha$ -globin gene deletions. *Cell*. 20:119.
  57. Seidman, J. G., A. Leder, M. H. Edgell, F. Polsky, S. M. Tilghman, D. C. Tiemeier, and P. Leder. 1978. Multiple related immunoglobulin variable-region genes identified by cloning and sequence analysis. *Proc. Natl. Acad. Sci. USA*. 75:3881.
  58. Kabat, E. A., T. T. Wu, and H. Bilofsky. 1979. Sequences of Immunoglobulin Chains. NIH Publication No. 80-2008, Bethesda, MD.