

Organization of Immunoglobulin Genes

S. TONEGAWA,* C. BRACK,* N. HOZUMI,* AND V. PIRROTTA†

*Basel Institute for Immunology, Basel 5, Switzerland; †Department of Microbiology, Biocenter, University of Basel, Basel, Switzerland.

Both light and heavy chains of antibody (or immunoglobulin) molecules consist of two regions. About 100 residues at the amino-terminal ends of the polypeptide chains determine the specificity of the molecules and compose *variable* (V) regions, whereas the remaining residues determine the *class* or *subclass* of antibodies to which the molecule belongs and compose *constant* (C) regions. Two types of heterogeneities exist among antibody molecules. The first distinguishes about 20 types of molecules with respect to their C regions. In mice, three (κ , λ_1 , and λ_2) and eight (δ , μ , γ_1 , etc.) different C regions are known for the light and heavy chains, respectively. The second type of heterogeneity distinguishes millions of molecules differing with respect to their V regions. This heterogeneity is often called "antibody diversity," and the question of whether the diversity in the structural genes coding for the V regions (*V genes*) arose in evolution or arises in ontogeny has been one of the most debated subjects in modern immunology (Wigzell 1973; Capra and Kehoe 1974).

Recent hybridization studies with purified immunoglobulin mRNA indicated that the number of germ-line V genes is far too small to account for the observed diversity in the V regions. For instance, there are no more than a few (and probably only one) germ-line V genes for the entire repertoire of the V regions which are associated with the λ -type light chains (Tonegawa 1976; Honjo et al. 1976). These results strongly suggested the existence of some mechanism by which coding information in the V genes is somatically diversified during differentiation of lymphocytes. In mice, the heterogeneity in the V_κ regions (V regions of κ -type chains) is much greater than in the V_λ regions (V regions of λ -type chains) (McKean et al. 1973; Cohn et al. 1974). The hybridization studies, as well as amino acid sequence analysis of myeloma κ chains, suggested that the mouse V_κ genes as a whole constitute a multigene family of as many as a few hundred unique but related sequences (Cohn et al. 1974; Tonegawa et al. 1977c). As with V_λ genes, the coding content of each of these V_κ genes diversifies somatically in the clones of B (bone-marrow-derived) lymphocytes, such that a large V_κ -region repertoire is generated from no more than a few germ-line genes in an adult mouse (Tonegawa et al. 1974; Rabbitts et al. 1975; Tonegawa 1976).

The number of C genes was also estimated by nu-

cleic acid hybridization. These studies showed that both C_κ and C_λ genes are unique or nearly unique (Faust et al. 1974; Stavnezer et al. 1974; Tonegawa et al. 1974; Honjo et al. 1974). Thus, at least in κ chains, the *total* number of germ-line genes is much greater for V regions than for C regions. This led us to investigate the two-"genes"-one-polypeptide-chain hypothesis for antibody chains, which was put forward a decade ago by Dreyer and Bennett (1965). We analyzed DNA from early embryos and from myelomas by digestion with restriction endonucleases and by hybridization with a purified κ -chain mRNA. These experiments strongly suggested that V_κ and C_κ genes are separate in early embryo DNA and that one of the multiple V_κ genes becomes contiguous to the C_κ gene during differentiation of B lymphocytes (Hozumi and Tonegawa 1976). The rearrangement permits the continuous transcription of a full immunoglobulin gene.

To further study the organization and regulation of this gene family, it is useful to clone chromosomal DNA fragments carrying immunoglobulin genes. By combining biochemical enrichment of such a DNA fragment with screening in situ of phage plaques, we were able to clone a λ phage which carries a mouse V_λ gene DNA insert. In this paper we describe isolation of the phage and characterization of the mouse DNA insert. In addition, further analysis with a restriction enzyme of total mouse DNA of various cellular sources is described.

Enrichment of a V_λ Gene

We previously showed that, in a total *EcoRI* digest of early BALB/c embryo DNA, three fragments of 4.5, 2.7, and 2.0 megadaltons contain λ -chain gene sequences (Tonegawa et al. 1977c). The use of a whole (a probe for V_λ and C_λ sequences) as well as a 3'-end-half (a probe for C_λ sequences) λ -chain mRNA in hybridization permitted us to conclude that the λ -chain gene sequences contained in the 4.5-megadalton fragment are exclusively for a C_λ region, whereas those contained in the 2.7- and 2.0-megadalton fragments are for V_λ regions (Tonegawa et al. 1977c). As a first attempt to clone an immunoglobulin gene, we selected the 2.7-megadalton V_λ -gene fragment as the source of mouse DNA.

We incubated the DNA fragments with excess λ -chain mRNA, purified from HOPC-2020 myeloma,

in 70% formamide under the conditions which promote replacement of the coding DNA strand with the RNA, but which do not completely dissociate the two DNA strands (R-loop formation) (White and Hogness 1977). We subjected the nucleic acid mixture to equilibrium centrifugation in CsCl and localized the position of the R-loop structure by hybridization with ^{125}I -labeled HOPC-2020 λ -chain mRNA after removal by alkali of the prehybridized RNA (Fig. 1). When no λ -chain mRNA was added during incubation in formamide (Fig. 1B), the fragment carrying V_λ gene sequences cobanded with the bulk DNA. When λ -chain mRNA was added (Fig. 1A), the major proportion of the V_λ -sequence-carrying DNA banded at a position which was clearly heavier

than the peak of the bulk DNA. The buoyant density at the hybridization peak in Figure 1A was 0.018 g/cm^3 higher than that in Figure 1B.

In a separate experiment we isolated a λgtWES phage which contains as its insert a 6.9-kb mouse DNA fragment (generated by *EcoRI*) carrying ribosomal DNA sequences (S. Tonegawa, in prep.). When the DNA fragment excised from the phage genome was annealed with the purified mouse 18S ribosomal RNA (2.0 kb) and centrifuged to equilibrium in CsCl, the R loop banded at a position 0.023 g/cm^3 denser than the duplex DNA. As a first-order approximation, the density increase of R loops is inversely proportional to the ratio of the lengths of DNA and RNA. The lengths of the V_λ -carrying DNA fragment and the λ -chain mRNA are about 4.8 and 1.2 kb, respectively. We therefore expect that the R loop formed between them is 0.019 g/cm^3 heavier than the DNA duplex. Thus we assume that the major hybridization peak observed in Figure 1A is composed of such an R loop. In Figure 1A, a second hybridization peak is observed in fraction 10. The increment of buoyant density of this peak is approximately twice that of the major hybridization peak, suggesting that this component is a hybrid formed between a single strand of the V_λ DNA fragment and the λ -chain mRNA. This component was not characterized further. The fractions 12–16 shown in Figure 1A were recentrifuged in CsCl. Although the hybridization peak remains at the original position, a large proportion of DNA (i.e., OD_{260} absorbing material) banded at the position where the bulk of the DNA banded in the first centrifugation. The fractions 12–16 shown in Figure 1C were pooled and used for cloning.

Enrichment for DNA fragments carrying the V_λ gene was approximately ten-, nine-, and fourfold by agarose gel electrophoresis and by first and second CsCl density gradient centrifugation, respectively. If we assume that other operations carried out between these steps (e.g., ethanol precipitation and dialysis) did not cause loss of specific DNA fragments, the overall enrichment factor was about 360-fold. The final yield was about $2 \mu\text{g}$ from 10 mg *EcoRI*-digested embryo DNA.

Cloning of a V_λ Gene

We inserted the partially purified V_λ DNA fragment in the middle of the λgtWES phage DNA by using T_4 ligase (Mertz and Davis 1973). The phage vector was developed by P. Leder and his coworkers and was approved as an EK-2 vector by the National Institutes of Health Advisory Committee on Recombinant DNA Research (Enquist et al. 1976). Upon transfection of CaCl_2 -treated *Escherichia coli* 803 ($r_1^- m_1^- \text{Su}_{\text{III}}^+$) (Mandel and Higa 1970), we obtained about 6000 plaques from $2 \mu\text{g}$ of the partially purified V_λ DNA. We screened about 4000 plaques by in situ hybridization with ^{125}I -labeled HOPC-2020

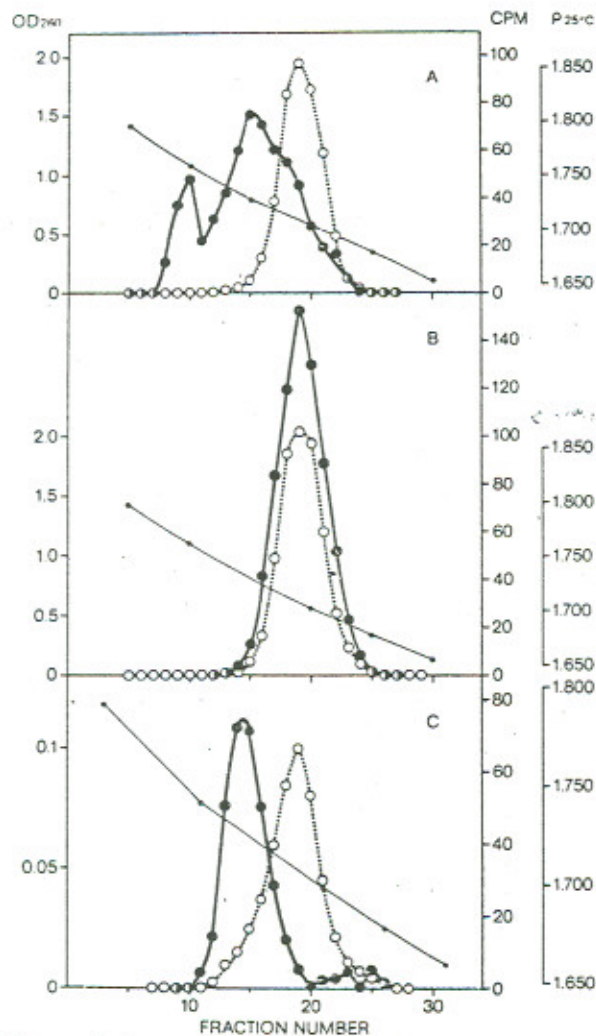


Figure 1. Preparative R-loop formation. *EcoRI*-digested embryonic DNA prefractionated by agarose gel electrophoresis for V_λ gene sequences was incubated in the presence (A) and absence (B) of HOPC-2020 λ -chain mRNA and centrifuged in CsCl as described elsewhere (Tonegawa et al. 1977a). Fractions 12–16 shown in A and another tube which showed profiles nearly identical to those in A were pooled and recentrifuged (C). (O—O) OD_{260} ; (●—●) DNA-RNA hybrid; (—) buoyant density.

mRNA (Kramer et al. 1976). The 38 plaques which produced gray or black autoradiographic spots of varying strength were reexamined by means of a second in situ hybridization with less RNA probe. Since the λ -chain mRNA probe used was about 90% pure and the impurity is distributed among many different RNA species (Tonegawa 1976), each composing a small fraction of the mRNA preparation, the use of smaller amounts of the RNA probe in hybridization favors detection of the DNA clone complementary to the major mRNA component. Out of the 38 plaques reexamined in this way, one plaque, λ gtWES-Ig 13, gave a distinctly stronger autoradiographic spot than the others.

Characterization of the Mouse DNA Insert

Length of the insert. When analyzed by agarose gel electrophoresis, *Eco*RI digestion of the DNA extracted from the clone λ gtWES-Ig 13 yielded, in addition to the left and right arms of the parental λ gtWES genome, one fragment of 3.0 megadaltons. The length of the insert was also determined by electron microscopy using two independent methods. λ gtWES-Ig 13 DNA was digested with *Eco*RI, and the three resulting fragments were measured. Taking PM2 DNA (10 kb) as the internal length standard, the left and right arms of the vector were calculated to be 21.2 and 13.9 kb, respectively; the mouse DNA insert was 4.8 kb.

In heteroduplex preparations made between λ gtWES- λ C and λ gtWES-Ig 13 DNA (Davis et al. 1971), these measurements were confirmed. One large heteroduplex region showed up at the expected position: between 53.1 and 65.2 map units from the left end (Fig. 2). The lengths of the two single-

stranded regions were 4.8 kb (mouse DNA) and 5.5 kb (λ *Eco*RI C fragment).

Hybridization with λ -Chain mRNA and cDNA

In Table 1, hybridization properties of the whole λ gtWES-Ig 13 DNA with various RNA probes are shown. The DNA hybridized well with HOPC-2020 whole λ mRNA, whereas the levels of hybridization obtained with the 3'-end-half of the same mRNA were no higher than when the same mRNA was hybridized to a clone carrying mouse ribosomal DNA. The results suggest that λ gtWES-Ig 13 DNA contains V_{λ} sequences and lacks C_{λ} gene sequences of HOPC-2020 mRNA. The same DNA showed no hybridization with a MOPC-321 κ -chain mRNA. Because these experiments were conducted with DNA baked on a Millipore filter, the efficiency of hybridization was not high. To circumvent this problem, full-length cDNA was synthesized from λ -chain mRNA with the avian myeloblastosis virus reverse transcriptase, and this DNA was hybridized with excess λ gtWES-Ig 13 DNA in liquid. The results are shown in Table 2. When assayed by hydroxylapatite, at least 60% of the cDNA, prepared either from MOPC-104E λ -chain mRNA or HOPC-2020 λ -chain mRNA, hybridized with the cloned DNA. When nuclease S_1 was used to remove the tail (and possibly some mismatched bases), the hybridization levels were reduced by about one-half. The mean melting point of the hybrid thus formed was 84°C in 0.12 M NaPO₄, as assayed by the hydroxylapatite method. It is not surprising that the hybridization of cDNA, as assayed by hydroxylapatite, is incomplete (up to 64%), since the λ mRNA used in preparing the cDNA was about 80% pure and the efficiency of hy-

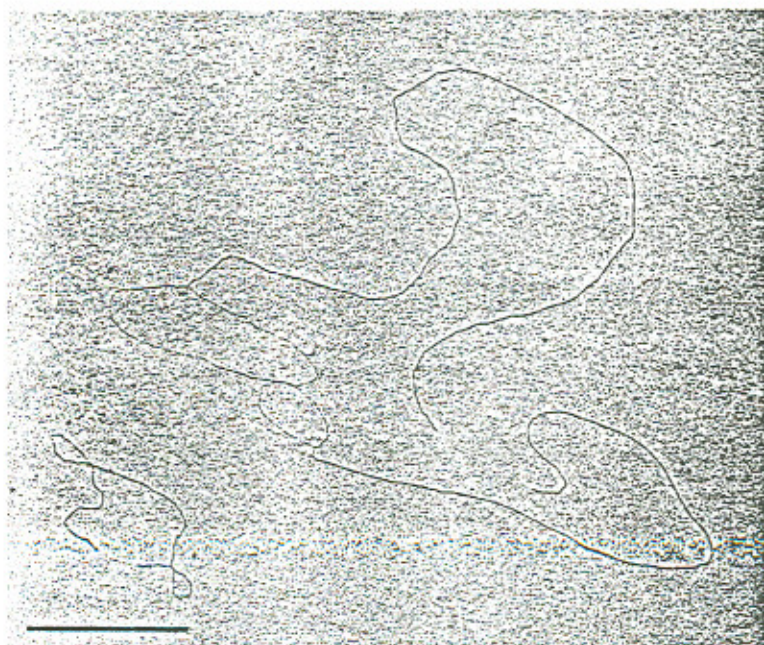


Figure 2. Electron micrograph of a hybrid molecule formed between λ gtWES- λ C and λ gtWES-Ig 13 DNA showing one large heteroduplex region. The bar represents 1 μ m.

Table 1. Hybridization of ^{125}I -labeled Light-chain mRNA and Its Fragments with $\lambda\text{gtWES-Ig 13 DNA}$

DNA on filter	μg	^{125}I mRNA	Input (cpm)	Hybrid (cpm)	Percentage of input
$\lambda\text{gtWES-Ig 13}$	25	λ chain (whole)	16,000	2,943	15.6
$\lambda\text{gtWES-Ig 13}$	50	λ chain (whole)	16,000	3,397	21.2
$\lambda\text{gt-rD}$	25	λ chain (whole)	16,000	136	0.85
$\lambda\text{gt-rD}$	50	λ chain (whole)	16,000	172	1.08
$\lambda\text{gtWES-Ig 13}$	25	λ chain (3'-end-half)	9,500	136	1.39
$\lambda\text{gtWES-Ig 13}$	50	λ chain (3'-end-half)	9,500	163	1.72
$\lambda\text{gt-rD}$	25	λ chain (3'-end-half)	9,500	125	1.32
$\lambda\text{gt-rD}$	50	λ chain (3'-end-half)	9,500	132	1.39
$\lambda\text{gtWES-Ig 13}$	25	κ chain (whole)	15,000	211	1.35
$\lambda\text{gtWES-Ig 13}$	50	κ chain (whole)	15,000	144	0.93
$\lambda\text{gt-rD}$	25	κ chain (whole)	15,000	158	1.00
$\lambda\text{gt-rD}$	50	κ chain (whole)	15,000	186	1.20

DNA was heat-denatured and fixed on a Millipore filter by the method of Gillespie and Spiegelman (1965). Hybridization was carried out in $2\times\text{SSC}$ buffered with 0.05 M PIPES (pH 7.1) at 69°C for 14 hr. Specific activity of RNA was $3\times 10^7\text{ cpm}/\mu\text{g}$ for whole and 3'-end-half λ -chain mRNA and $3.2\times 10^7\text{ cpm}/\mu\text{g}$ for κ mRNA. The hybrid was assayed by RNase treatment (RNase A, $20\ \mu\text{g}/\text{ml}$; RNase T₁, 2 units/ml, in $2\times\text{SSC}$). $\lambda\text{gt-rD}$ designates a λWES phage in which the center *EcoRI* fragment was replaced with a 6.9-kb mouse DNA fragment carrying ribosomal DNA sequences. The ^{125}I -labeled 3'-end-half of the λ -chain mRNA and the whole κ -chain mRNA used in this experiment hybridized well with corresponding cDNA (data not shown).

bridization under these conditions is about 90%. These results are consistent with the hypothesis that the $\lambda\text{gtWES-Ig 13}$ clone carries V_λ gene sequences.

Length and Location of Homology Region

Further evidence for the existence of V_λ gene sequences was obtained by electron microscopic examination of R loops. HOPC-2020 λ -chain mRNA and $\lambda\text{gtWES-Ig 13 DNA}$ were incubated under conditions of R-loop formation (Thomas et al. 1976; Tonegawa

et al. 1977a). We observed only a relatively low proportion of hybrid molecules. About 30% of the molecules displayed a small R loop, and some of the R loops had a small RNA tail at one end (Fig. 3). Measurement of 89 such hybrid molecules showed that the R loop is at a unique position: 63 ± 1 map units from the left end (Fig. 4). Since the right *EcoRI* site on this chimeric molecule is at 65.2 map units, we conclude that the region to which HOPC-2020 λ -chain mRNA hybridizes lies within the mouse DNA.

The lengths of the R loop and the RNA tails were 400 ± 100 nucleotides and 200–600 nucleotides, respectively. We think that the large variations in the length measurements of the RNA tail result from that fact that the mRNA molecules are not fully denatured and do not completely extend under the spreading condition used.

More accurate measurement of the position and length of the homology region was made by electron microscopic examination of R loops formed with the λ -chain mRNA and the purified 4.8-kb *EcoRI* fragment, as well as hybrid molecules formed with the same DNA fragment and full-length cDNA synthesized on MOPC-104E λ -chain mRNA (in prep.). The results obtained from these experiments confirmed the conclusion drawn above and indicated that the homology region is approximately 400 nucleotides long and that it is localized between 3290 and 3690 (± 200) nucleotide pairs from the left end of the 4.8-kb mouse DNA insert. Figure 4 shows the maps of $\lambda\text{gtWES-Ig 13}$, $\lambda\text{gtWES-}\lambda\text{C DNA}$, and the mouse DNA insert as obtained from electron microscopy.

Restriction Enzyme Cleavage Sites

To characterize the mouse DNA insert further, we determined the cleavage sites of several restric-

Table 2. Hybridization of ^{32}P λ -Chain cDNA with Ig 13 DNA

^{32}P -labeled cDNA prepared on	C_0t	Fraction of ^{32}P counts in hybrid	
		hydroxylapatite assay (%)	nuclease S_1 assay (%)
HOPC-2020 mRNA	0	4.4	2.7
HOPC-2020 mRNA	5×10^{-2}	61	27
HOPC-2020 mRNA	6.7×10^{-2}	61	27
MOPC-104E mRNA	2.2×10^{-1}	64	30

Isolation of Ig 13 DNA, synthesis of full transcript cDNA, and annealing were carried out as described in the legend to Fig. 2. For the hydroxylapatite assay, the annealing mixtures were made to 0.12 M NaPO_4 by the addition of water, and were loaded on a small hydroxylapatite column pre-equilibrated with 0.12 M NaPO_4 , pH 6.9, at 60°C . The column was washed with 9 ml of the same buffer and the hybrid was eluted with 6 ml of 0.4 M NaPO_4 , pH 6.9. The hybrid fraction was determined by dividing the radioactivity in the 0.4 M NaPO_4 fraction by the total radioactivity recovered from the column. For the nuclease S_1 assay, the annealing mixture was divided into two equal parts. DNA in one part was directly precipitated with trichloroacetic acid, and DNA in the other part was precipitated after treatment with nuclease S_1 (see legend to Fig. 2). The hybrid fraction was determined by subtracting the intrinsic S_1 -resistant counts (1.2%) from the ratio of the S_1 -resistant counts and the total TCA precipitable counts. $C_0t = 0$ sample was prepared by placing the heat-denatured sample directly in a dry-ice ethanol bath. C_0t refers to the concentration of DNA in moles of nucleotide per liter \times incubation time in sec.

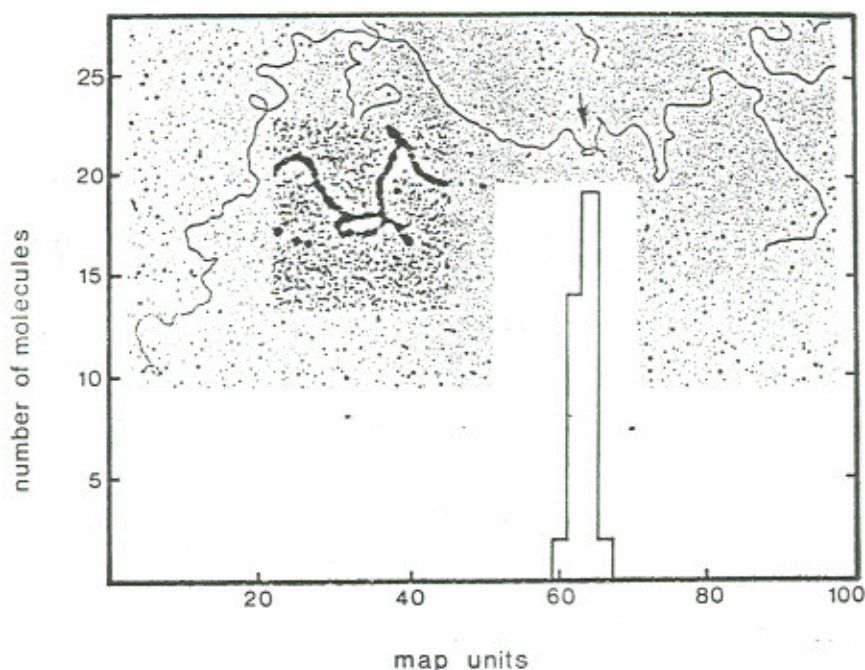


Figure 3. Electron micrograph of an R-loop molecule formed by hybridizing myeloma HOPC-2020 λ -chain mRNA with λ gtWES-Ig 13 DNA. The R loop with an RNA tail is shown at higher magnification in the insert. The histogram was made from one representative set of measurements from the R loops made by incubation at 57°C and high salt, and spread from 70% formamide. The R loop is at a unique position 63 ± 1 map units from the left end of the λ gtWES-Ig 13 molecule.

tion enzymes. Two kinds of experiments were carried out. In one series of experiments, the 4.8-kb *EcoRI* fragment was digested by several restriction enzymes which cleave DNA relatively infrequently. The digests were electrophoresed in 1.2% agarose gel, and the DNA fragments containing sequences homologous to HOPC-2020 λ -chain mRNA were identified by the Southern transfer technique (Southern 1975). In some cases, mere comparison of these results with the homology map obtained by electron microscopy (Fig. 4) allowed us to order the fragments. In other cases, fragment order was deduced from results obtained by single- and multiple-enzyme digestion. Figure 5A summarizes these results.

In the other series of experiments, a more extensive cleavage-site map of the 1.5-kb *HaeIII* fragment was obtained by the partial-digestion method of Smith and Birnstiel (1976). The 4.8-kb *EcoRI* fragment was terminally labeled with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ by T_4 polynucleotide kinase. The 1.5-kb *HaeIII* fragment was isolated by agarose gel electrophoresis and was partially digested with *HinfI*, *MboII*, and *AluI*. The positions of the cleavage sites determined by this method are summarized in Figure 5B. These results

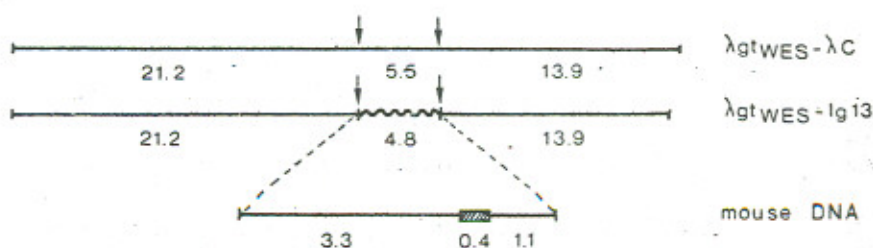
were confirmed by total digestion of the *HaeIII* fragment (data not shown).

Nucleotide Sequences

Ultimate evidence that the mouse DNA insert contains λ -chain gene sequences comes from direct DNA sequencing. We have determined a partial sequence of the *HinfI* C fragment. This fragment is about 290 nucleotides long and includes part of the homology region and the sequences immediately adjacent to it (Fig. 5B). According to the map in Figure 5B, there should be an *MboII* restriction site near the center of this fragment. We labeled the 5' ends of the *HinfI* fragment with polynucleotide kinase, cut it with *MboII*, and determined the nucleotide sequences of the two halves by the method of Maxam and Gilbert (1977).

The results of the left and right halves are shown in Figures 6 and 7, respectively. To determine the fit of the observed sequences to that predicted from the amino acid sequence of the λ chains, all possible *HinfI* sites in the predicted sequence were aligned with the DNA sequences. Figure 6 shows that the possible *HinfI* site at amino acid positions 63 and

Figure 4. Maps of λ gtWES- λ C and λ gtWES-Ig 13 DNA. The wavy line indicates the mouse DNA insert. The hatched box represents the region homologous to HOPC-2020 λ -chain mRNA.



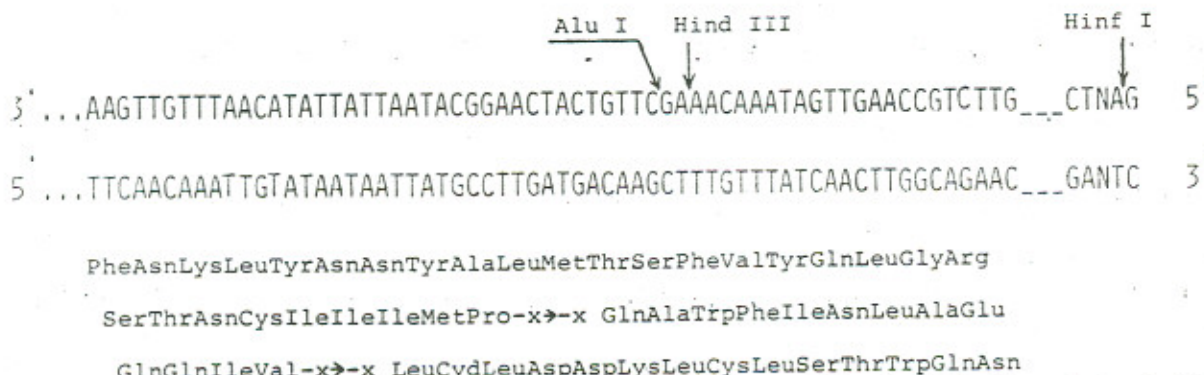


Figure 7. Partial nucleotide sequences of the right half of the *Hinf* C fragment. Experimentally determined nucleotide sequence is shown at the top. In the second row, the complementary sequence is shown. The latter sequence was transformed into amino acid sequences by reading it in all three frames. These amino acid sequences are shown under the complementary nucleotide sequence. Positions of *Alu*I and *Hinf*I cleavage sites are indicated by arrows. (x) Indicates a termination triplet.

acid sequences obtained by reading the complementary sequence in each of the three frames. Amino acid sequences read by two possible frames run into double termination codons. The third frames yielded amino acid sequences which bear no similarity to either the C_I or C_{II} region. Alternatively, we have tried to align the observed and predicted nucleotide sequences using as reference points the *Hinf* and the *Alu*I recognition sequences. No alignment was possible which gave less than 40% mismatching, even using the most favorable codon assignments. Whether or not some residual sequence homology exists between this region and C_λ genes remains to be determined. We conclude that a C_λ gene is not contiguous to the V_λ gene.

In collaboration with A. Maxam, W. Gilbert, R. Tizard, and O. Bernard, we recently determined more extensively the nucleotide sequences in the homology region and the regions immediately adjacent to it (in prep.). The result confirms the sequences shown in Figures 6 and 7.

Rearrangement of Immunoglobulin Genes

The DNA cloning technique is a powerful tool for analyzing the organization and function of eukaryotic genes. There is no doubt that the analysis of various immunoglobulin gene clones from different cellular origins will reveal many microscopic details of the arrangement and rearrangement of DNA sequences. For a better understanding of the various problems associated with this gene family, it would be useful to combine these studies with analysis of total cellular DNA at a more macroscopic level.

When limit *Bam*HI digests of total DNA from BALB/c embryos or MOPC-321 myeloma (a κ -chain producer) were fractionated by agarose gel electrophoresis and the DNA in each fraction was hybridized with a whole κ -chain mRNA (a probe for V_κ and C_κ gene sequences), profiles of hybridization

were drastically different with the two DNAs (Hozumi and Tonegawa 1976). We interpreted the profile difference as a result of a rearrangement of immunoglobulin genes which takes place during differentiation of lymphocytes. The rearrangement is thought to bring one of the multiple V genes in contiguity to a C gene, thereby permitting the continuous transcription of a full immunoglobulin gene. An alternative explanation of the results, namely, that accumulation of multiple mutations or base modifications leading to either loss or gain of *Bam*HI sites generated the observed pattern difference, was considered to be unlikely (Hozumi and Tonegawa 1976). Another possible, also unlikely, interpretation was to ascribe the pattern difference to massive scrambling of DNA sequences which might accompany generation or propagation of the myeloma cells. To exclude these trivial interpretations, we extended the analysis to DNAs of other sources.

In Figure 8B a *Bam*HI digest of kidney DNA from adult BALB/c mice was analyzed with a whole κ -chain mRNA purified from MOPC-321 myeloma, as well as with its 3'-end-half. Two DNA components of 6.0 and 3.9 megadaltons hybridized with the whole RNA molecule, whereas only the 6.0-megadalton component hybridized with the 3'-end-half. Overall, hybridization patterns were indistinguishable from those obtained with embryonic DNA (Fig. 8A) (Hozumi and Tonegawa 1976). Essentially identical patterns were obtained with DNAs of other adult tissues such as liver and brain. The hybridization patterns of the whole κ -chain mRNA with DNA from J558 myeloma were also indistinguishable from those obtained with embryonic DNA (Fig. 8A). In this myeloma, a λ chain is produced and no κ -chain gene is expressed. DNA from two other λ -chain-producing myelomas (MOPC-104E and HOPC-2020) gave essentially the same result (not shown). Conversely, when DNA from a κ -chain-producing myeloma (MOPC-321) was analyzed with a purified λ -chain mRNA as hybridization probe (HOPC-2020), the hybridization pattern was indistinguishable from that

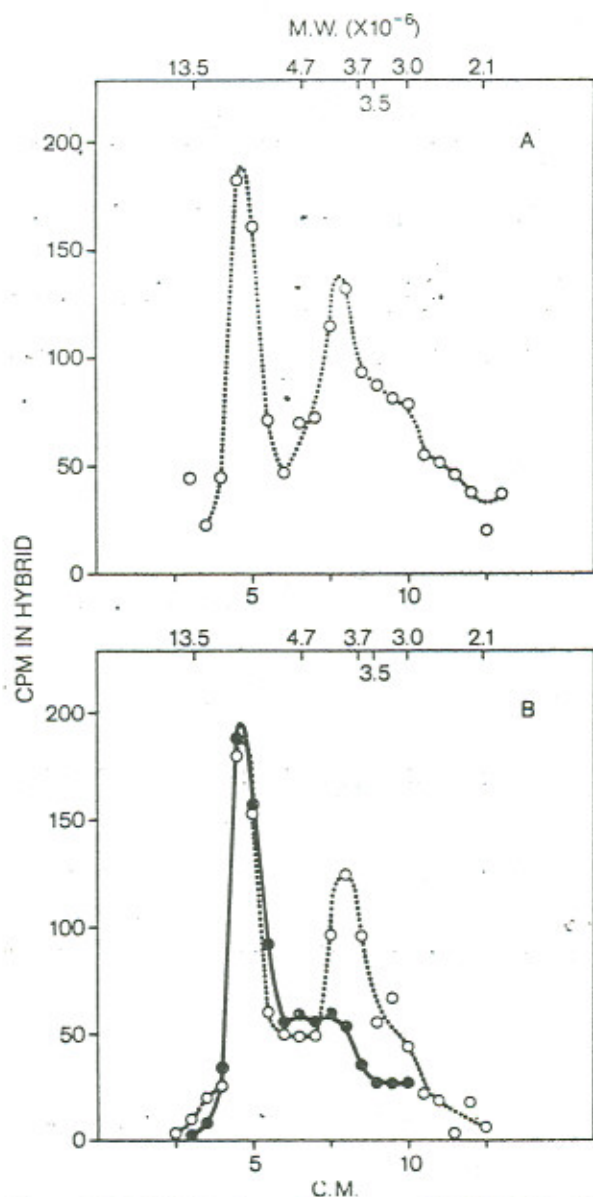


Figure 8. Gel electrophoresis patterns of λ J558 DNA (A) and kidney DNA (B) digested with *Bam*HI. Whole MOPC-321 RNA (\circ - \cdots - \circ) (1250 cpm, 7×10^7 cpm/ μ g) or its 3'-end-half fragment (\bullet - \cdots - \bullet) (600 cpm, 7×10^7 cpm/ μ g) was hybridized to fractionated DNA. Digestion of DNA, preparation of RNA probes, electrophoresis, and hybridization procedures were as described elsewhere (Tonegawa 1976; Hozumi and Tonegawa 1976). The number at the top of each panel in this figure and in Figs. 9 and 10 indicates the molecular weight (in megadaltons) of *Eco*RI-digested phage λ DNA used as migration markers.

obtained with embryo DNA by the same mRNA (Tonegawa et al. 1977b). These results render the trivial interpretations described above unlikely.

An Inactive V Gene Remains Unjoined with the Corresponding C Gene in Plasma Cells

Our earlier hybridization studies, as well as those of others, indicated that a group of closely related

V regions are somatically generated from a few, probably single, germ line gene(s) (Tonegawa et al. 1974; Rabbitts et al. 1975; Tonegawa 1976; Honjo et al. 1976; Tonegawa et al. 1977c). Such a V-region group is best approximated to the subgroup as defined by Cohn et al. (1974). We previously presented direct demonstration of separate germ-line V genes for two V regions of different subgroups (Tonegawa et al. 1977c). The results are reproduced in Figure 9A for the present context. Embryonic DNA digested with *Bam*HI enzyme was analyzed with two mRNAs coding for two κ chains of different subgroups, MOPC-321 and MOPC-21. These two κ chains show little homology in their V regions, whereas they have identical sequences in the C region (Milstein and Svasti 1971; McKean et al. 1973). As expected, both RNAs hybridized with the 6.0-megadalton com-

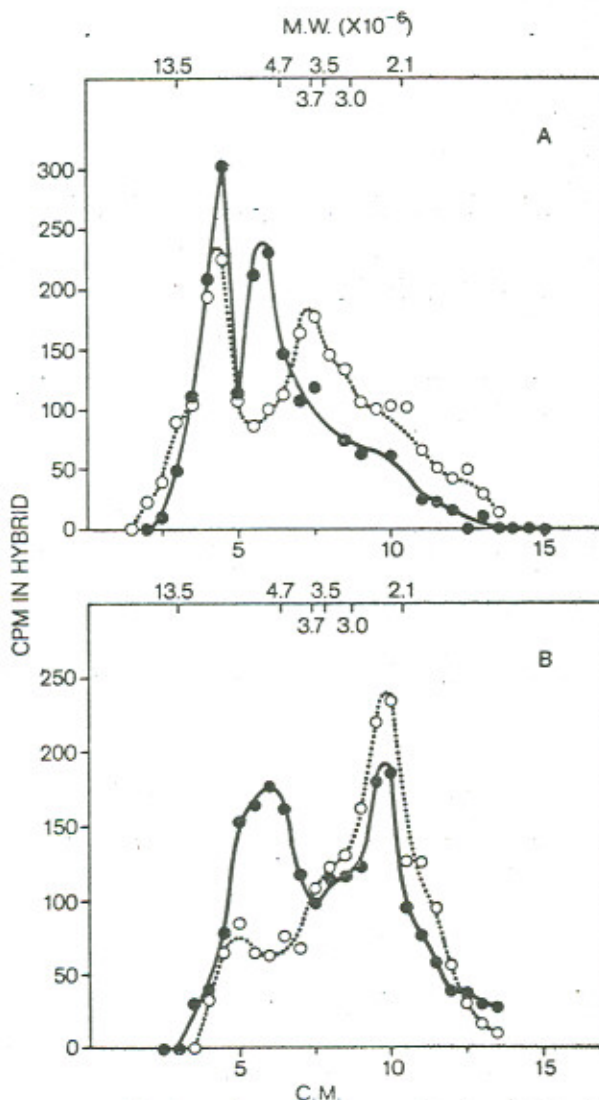


Figure 9. Gel electrophoresis patterns of embryo DNA (A) and MOPC-321 DNA (B) digested with *Bam*HI. Whole MOPC-321 mRNA (\circ - \cdots - \circ) (1250 cpm, 7×10^7 cpm/ μ g) or MOPC-21 mRNA (\bullet - \cdots - \bullet) (1220 cpm, 8×10^7 cpm/ μ g) was annealed with extracted DNA.

ponent which carries the C_{κ} gene (see above). In addition, each of the two RNAs hybridized with a second, but mutually different, DNA component. These DNA components of 5.0 and 3.9 megadaltons should carry MOPC-21 and MOPC-321 V-gene sequences, respectively.

Is a V_{κ} gene for a given subgroup joined with a C_{κ} gene in the myeloma which synthesizes a κ chain carrying a V region of another subgroup? That this is probably not the case is shown in Figure 9B, where MOPC-321 DNA was analyzed with the homologous (MOPC-321) and heterologous (MOPC-21) κ mRNA. As we reported previously, the homologous κ mRNA hybridized with a major DNA component of 2.4 megadaltons (Hozumi and Tonegawa 1976). This component carries both V_{κ} MOPC-321 and C_{κ} gene sequences. While the 6.0-megadalton, embryonic C_{κ} gene component disappears, the 5.0-megadalton V_{κ} MOPC-21 gene fragment remains at the embryonic position (this band was atypically broad in this particular experiment). These results, as well as those described in the last section (i.e., analysis of κ -gene sequences in DNA from adult tissues and from λ -chain-producing myelomas), suggest that there is a strict correlation between the V-C-gene joining and activation of the joined immunoglobulin genes.

Does the V-C-gene Joining Take Place in Both of the Homologous Chromosomes?

The apparent disappearance of the embryonic V- and C-gene components (Hozumi and Tonegawa 1976) in MOPC-321 DNA suggested that the joining event leads to homozygosity. As we pointed out previously (Tonegawa et al. 1977c), homozygosity could result from the loss of one homolog followed by reduplication of the other, or it could result from somatic recombination between the centromere and the immunoglobulin locus. The alternative view—that joining takes place in both chromosomes—presents a problem since there is more than one V gene (see above). There is no intrinsic reason why the same V gene should be joined on both homologs. The above results, however, could be explained by the known abnormality of the karyotype of mouse myelomas. Although MOPC-321 is subtetraploid (S. Tonegawa, unpubl.), we have no information as to the number of chromosomes on which the κ genes lie. Thus it is possible that this myeloma has lost the homolog(s) on which the V_{κ} MOPC-321 and C_{κ} genes are located separately.

To investigate this possibility, we analyzed DNA from another myeloma, TEPC-124 (Fig. 10). MOPC-321 and TEPC-124 κ chains are different only by three amino acids in the V regions (McKean et al. 1973) and belong to a single subgroup. The two V regions, therefore, presumably share the same germ-line V gene. Since nucleotide sequences in the two mRNAs are extensively homologous (Tonegawa et al. 1974; S. Tonegawa, unpubl.), MOPC-321 κ mRNA and its 3'-end-half were used as the hybridization

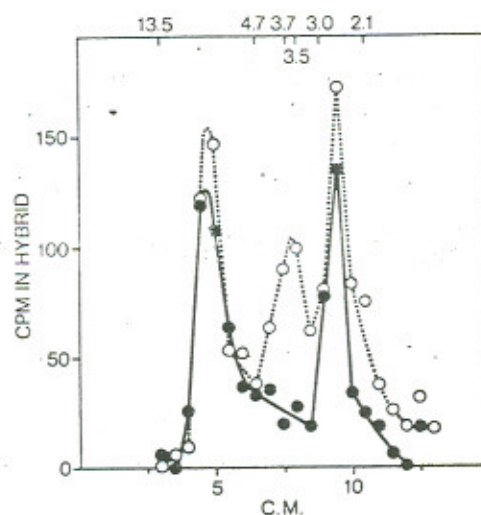


Figure 10. Gel electrophoresis patterns of TEPC-124 DNA digested with *Bam*HI. Whole MOPC-321 mRNA (○---○) or its 3'-end-half fragment (●—●) was annealed with extracted DNA.

probes instead of the homologous mRNA. Three major DNA components hybridized with the whole κ mRNA, two of which hybridized also with the 3'-end-half fragments (Fig. 10). The size and hybridization properties of these DNA components suggest that the overall pattern is a composite of the two patterns obtained when embryonic and MOPC-321 DNA were analyzed by the same RNA probes. Thus the principal difference in the two hybridization patterns, one of embryonic DNA and the other of TEPC-124 tumor, is the addition of the 2.4-megadalton component in the latter. This component hybridized with both whole and 3'-end half RNA probes. Each of the two components that hybridized with the 3'-end-half should contain a complete C gene sequence, for our recent study with a κ -chain cDNA clone indicates that there is no *Bam*HI cleavage site in the C gene (G. Matthyssens and S. Tonegawa, unpubl.). If the two V regions indeed share the same germ-line V gene, the straightforward interpretation of the above results is that the V gene is joined with the C gene only in one of the two homologs in plasma cells.

For immunoglobulin loci, only one allele is expressed in any given lymphocyte (allelic exclusion) (Pernis et al. 1965). This is not the case for any other autosomal gene studied until now. Considering the strict correlation between joining of a pair of V and C genes and the expression of the joined V gene, the above results with TEPC-124 DNA may be relevant to the mechanism of allelic exclusion. On the other hand, the apparent homozygosity observed in MOPC-321 DNA can also conveniently explain allelic exclusion, if this is a naturally occurring event. In this case, heterozygosity observed in TEPC-124 DNA could result from accidental acquisition of a homolog from a nonlymphatic cell during generation or propagation of the tumor. We feel that the

matter will be clarified only by analysis of DNA from natural lymphocyte clones.

CONCLUSIONS

The nucleotide-sequence determination of a cloned, embryonic V_{λ} gene directly demonstrated that V genes are separate from a corresponding C gene in embryonic cells. Analysis by restriction enzymes of total cellular DNA from various sources strongly suggested that the two separate immunoglobulin genes become continuous during differentiation of B lymphocytes. There seems to be a strict correlation between the joining event and activation of the joined genes. Cloning of more immunoglobulin genes from embryo and plasma cells will not only provide direct demonstration of such a gene-joining event but also help in the elucidation of a possible relationship of the event to gene activation mechanisms.

The gene-cloning experiments were carried out in a P-3 laboratory in accordance with the National Institutes of Health guidelines issued in June 1976. The phage vector used was approved, in January 1976, as an EK-2 vector by the National Institutes of Health Advisory Committee on Recombinant DNA Research. After completion of the work, in April 1977, we were informed that this approval was withdrawn. In compliance with the new rule, we are now growing the λ gtWES-Ig 13 clone on *E. coli* DP50. Retention of the original amber mutations in this phage clone was kindly confirmed by Dr. B. Hohn at the Biocenter, Basel.

Acknowledgments

We thank Drs. M. Potter, M. Cohn, M. Weigert, J. W. Beard, T. Bickle, P. Leder, M. Sugiura, and M. Takanami for providing strains and enzymes. We thank Dr. B. Hohn for analyzing the genetic markers carried in λ phages. We also thank Dr. G. Matthysens for a gift of λ -chain cDNA. The technical assistance of Mr. G. Dastoornikoo, Ms. R. Schuller, Ms. P. Riegert, and Mr. A. Traunecker is highly appreciated. The excellent secretarial help of Ms. K. Perret-Thurston is also appreciated.

REFERENCES

- CAPRA, J. D. and J. M. KEHOE. 1974. Antibody diversity: Is it all coded for by the germ line genes? *Scand. J. Immunol.* 3: 1.
- COHN, M., B. BLOMBERG, W. GECKLER, W. RASCHKE, R. RIBLET, and M. WEIGERT. 1974. First order considerations in analyzing the generator of diversity. In *The immune system, genes, receptors, signals* (ed. E. E. Sercarz et al.), p. 89. Academic Press, New York.
- DAVIS, R. W., M. SIMON, and N. DAVIDSON. 1971. Electron microscope heteroduplex methods for mapping regions of base sequence homology in nucleic acids. *Methods Enzymol.* 21D: 413.
- DREYER, W. J. and J. C. BENNETT. 1965. The molecular basis of antibody formation: A paradox. *Proc. Natl. Acad. Sci.* 54: 864.
- DUGAN, E. S., R. A. BRADSHAW, E. S. SIMMS, and H. N. EISEN. 1973. Amino acid sequence of the light chain of a mouse myeloma protein (MOPC-315). *Biochemistry* 12: 5400.
- ENQUIST, L., D. TIEMEIER, P. LEDER, R. WEISBERG, and N. STERNBERG. 1976. Safer derivatives of bacteriophage λ gt λ c for use in cloning of recombinant DNA molecules. *Nature* 259: 596.
- FAUST, C. H., H. DIGGELMANN, and B. MACH. 1974. Estimation of the number of genes coding for the constant part of the mouse immunoglobulin kappa light chain. *Proc. Natl. Acad. Sci.* 71: 2491.
- GILLESPIE, D. and S. SPIEGELMAN. 1965. A quantitation assay for DNA-RNA hybrids with DNA immobilized on a membrane. *J. Mol. Biol.* 12: 829.
- HONJO, T., S. PACKMAN, D. SWAN, and P. LEDER. 1976. Quantitation of constant and variable region genes for mouse immunoglobulin λ chains. *Biochemistry* 15: 2780.
- HONJO, T., S. PACKMAN, D. SWAN, M. NAU, and P. LEDER. 1974. Organization of immunoglobulin genes: Reiteration frequency of the mouse κ chain constant region gene. *Proc. Natl. Acad. Sci.* 71: 3659.
- HOZUMI, N. and S. TONEGAWA. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci.* 73: 3628.
- KRAMER, R. A., J. R. CAMERON, and R. W. DAVIS. 1976. Isolation of bacteriophage λ containing yeast ribosomal RNA genes: Screening by in situ RNA hybridization to plaques. *Cell* 8: 227.
- MANDEL, M. and A. HIGA. 1970. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* 53: 159.
- MAXAM, A. M. and W. GILBERT. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* 74: 560.
- McKEAN, D., M. POTTER, and L. HOOD. 1973. Mouse immunoglobulin chains. Pattern of sequence variation among kappa chains with limited sequence differences. *Biochemistry* 12: 760.
- MERTZ, J. E. and R. W. DAVIS. 1972. Cleavage of DNA by R_1 restriction endonuclease generates cohesive ends. *Proc. Natl. Acad. Sci.* 69: 3370.
- MILSTEIN, C. and J. SVAJTI. 1971. Expansion and contraction in the evolution of immunoglobulin gene pools. *Prog. Immunol.* 1: 33.
- PERNIS, B., G. CHIAPPINO, A. S. KELUS, and P. G. H. GELL. 1965. Cellular localization of immunoglobulins with different allotypic specificities in rabbit lymphoid tissues. *J. Exp. Med.* 122: 853.
- RABBITS, T. H., J. M. JARVIS, and C. MILSTEIN. 1975. Demonstration that a mouse immunoglobulin light chain messenger RNA hybridizes exclusively with unique DNA. *Cell* 6: 5.
- SMITH, H. O. and M. L. BIRNSTIEL. 1976. A simple method for DNA restriction site mapping. *Nucleic Acids Res.* 3: 2387.
- SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel-electrophoresis. *J. Mol. Biol.* 98: 503.
- STAVNEZER, J., R. C. C. HUANG, E. STAVNEZER, and J. M. BISHOP. 1974. Isolation of messenger RNA for an immunoglobulin kappa chain and enumeration of the genes for the constant region of kappa chain in the mouse. *J. Mol. Biol.* 88: 43.
- THOMAS, M., J. R. CAMERON, and R. W. DAVIS. 1974. Viable molecular hybrids of bacteriophage lambda in eukaryotic DNA. *Proc. Natl. Acad. Sci.* 71: 4579.
- TONEGAWA, S. 1976. Reiteration frequency of immunoglobulin light chain genes: Further evidence for somatic generation of antibody diversity. *Proc. Natl. Acad. Sci.* 73: 203.

- TONEGAWA, S., C. BRACK, N. HOZUMI, and R. SCHULLER. 1977a. Cloning of an immunoglobulin variable region gene from mouse embryo. *Proc. Natl. Acad. Sci.* 74: 3518.
- TONEGAWA, S., N. HOZUMI, C. BRACK, and R. SCHULLER. 1977b. Arrangement and rearrangement of immunoglobulin genes. In *Regulation of the immune system: Genes and the cells in which they function. ICN-UCLA Symp.* (in press).
- TONEGAWA, S., N. HOZUMI, G. MATTHYSSENS, and R. SCHULLER. 1977c. Somatic changes in the content and context of immunoglobulin genes. *Cold Spring Harbor Symp. Quant. Biol.* 41: 877.
- TONEGAWA, S., C. STEINBERG, S. DUBE, and A. BERNARDINI. 1974. Evidence for somatic generation of antibody diversity. *Proc. Natl. Acad. Sci.* 71: 4027.
- WHITE, R. L. and D. S. HOGNESS. 1977. R-loop mapping of the 18S and 28S sequences in the long and short repeating units of *Drosophila melanogaster* rDNA. *Cell* 10: 177.
- WIGZELL, H. 1973. Antibody diversity: Is it all coded for by the germ line genes? *Scand. J. Immunol.* 2: 199.