

# A Complete Immunoglobulin Gene Is Created by Somatic Recombination

Christine Brack, Minoru Hirama,  
Rita Lenhard-Schuller and Susumu Tonegawa  
Basel Institute for Immunology  
Grenzacherstrasse 487  
Postfach, 4005 Basel 5  
Switzerland

## Summary

Using a pCRI plasmid containing an enzymatically synthesized, full-length DNA transcript of immunoglobulin  $\lambda$  chain mRNA as the hybridization probe in the Southern gel blotting experiments (Southern, 1975), we identified three DNA fragments of 8.6, 4.8 and 3.5 kb in Eco RI-digested total DNA from BALB/c mouse embryos. A fourth fragment of 7.4 kb was found in addition to these three fragments in similarly digested total DNA from a  $\lambda$  chain-secreting myeloma (HOPC 2020). We have cloned the four DNA fragments in an EK-2 phage vector,  $\lambda_{WES}$ , and characterized them with respect to size, type of  $\lambda$  gene sequences contained and position of these sequences in the fragments, using agarose gel electrophoresis, the gel blotting technique and electron microscopic R loop mapping. The embryonic DNA clones Ig 99 $\lambda$ , Ig 25 $\lambda$  and Ig 13 $\lambda$  contain one copy each of  $V_{\lambda I}$ ,  $C_{\lambda I}$  and  $V_{\lambda II}$  sequences, respectively, while the myeloma DNA clone Ig 303 $\lambda$  contains one copy each of  $V_{\lambda I}$  and  $C_{\lambda I}$  sequences that are separated by a 1.2 kb nontranslated DNA segment. Ig 25 $\lambda$  was also shown to contain a DNA segment of approximately 40 base pairs (bp) (J sequence) that lies 1.2 kb away from the  $C_{\lambda I}$  sequence and is homologous to the V-C junction region of a  $\lambda_1$  mRNA. Heteroduplex analysis of the three  $\lambda_1$  DNA clones revealed that Ig 303 $\lambda$  DNA is composed of two parts, one of which is entirely homologous to one end of Ig 99 $\lambda$ , and the other to one end of Ig 25 $\lambda$  DNA. The sequence arrangement observed in the cloned DNA is the same as that in the corresponding cellular DNA. This was shown by identifying certain restriction enzyme cleavage sites on the cloned DNAs and demonstrating the presence of these sites in the total cellular DNA by the gel blotting technique. The site of the homology switch is at the boundary of the V sequence and the 1.2 kb nontranslated DNA segment, and corresponds to the position of the J sequence on the Ig 25 $\lambda$  DNA. We consider the above experimental results the most direct evidence for somatic rearrangement in immunoglobulin genes. We discuss the significance of these findings for the origin of genes in the evolution of higher organisms and in cell differentiation.

## Introduction

Are DNA sequences in the cells of higher organisms rearranged during normal cell differentiation? The restriction enzyme mapping of mouse DNAs strongly suggested that this is the case in lymphocytes for the immunoglobulin genes. Very different patterns of hybridization were obtained when  $\kappa$  light chain mRNA from a  $\kappa$  chain-producing myeloma was hybridized with mouse embryo DNA or with homologous myeloma DNA, both of which had been digested with the restriction endonuclease Bam HI and fractionated by agarose gel electrophoresis (Hozumi and Tonegawa, 1976). Suggestive evidence for a similar DNA rearrangement involving  $\lambda$  chain genes was obtained when DNAs from embryo and from a  $\lambda$  chain-producing myeloma were compared (Tonegawa et al., 1976). Furthermore, a pattern change was detected only for those immunoglobulin genes that are active in a given myeloma cell (Tonegawa et al., 1977a, 1977b). These results indicated that in embryo cells, the DNA sequences coding for the amino terminal half (V region) and for the carboxy terminal half (C region) are separate, and that the two sequences are brought to proximity during the differentiation of B (bone marrow-derived) lymphocytes. An alternative, improbable interpretation was also considered: the pattern difference might result from mutations or base modifications in the enzyme cleavage sites (Hozumi and Tonegawa, 1976).

Two subtypes of mouse  $\lambda$  chains are known,  $\lambda_I$  and  $\lambda_{II}$ , which are characterized by the specific C region sequences  $C_{\lambda I}$  and  $C_{\lambda II}$ . Amino acid sequence studies have so far established seven different  $V_{\lambda I}$  regions and one  $V_{\lambda II}$  region (Weigert et al., 1970; Dugan et al., 1973). Statistical considerations suggest that the mouse is capable of synthesizing many more than seven different V regions of the  $\lambda_I$  subtype (Tonegawa, 1978). The two  $\lambda$  subtype chains each seem to be encoded in a pair of DNA segments, one for the V and the other for the C, that lie in separate sections of the embryo DNA. Restriction enzyme mapping of total cellular DNA combined with hybridization kinetics strongly suggested that there are no more than a few copies (and there is probably only one copy) per haploid genome of the DNA segment encoding each of the four  $\lambda$  chain regions  $V_{\lambda I}$ ,  $C_{\lambda I}$ ,  $V_{\lambda II}$  and  $C_{\lambda II}$  (Tonegawa et al., 1976). Hence the multiple  $V_{\lambda I}$  regions observed in myelomas must have been generated by a somatic process.

To obtain more direct evidence for gene rearrangement, we have isolated, by in vitro recombination with phage  $\lambda$  DNA, DNA fragments from both embryo and myeloma cells that carry part or



all of the  $\lambda$  chain genes. Our previous reports dealt with the cloning and characterization of the embryonic  $V_{\lambda H}$  DNA fragment Ig 13 $\lambda$  (Tonegawa et al., 1977a, 1977c; Tonegawa et al., 1978) and a myeloma DNA fragment containing both  $V_{\lambda}$  and  $C_{\lambda}$  sequences (Brack and Tonegawa, 1977). In both cloned DNA fragments, the protein-encoding sequences are interrupted by nontranslated sequences that have been called introns (Tonegawa et al., 1978) or intervening sequences (Tilghman et al., 1978a). The present report describes the isolation and characterization of two new embryonic DNA clones and cites the results obtained in experiments with the three  $\lambda_1$  DNA clones. The latter provided us with direct evidence for somatic rearrangement of immunoglobulin gene sequences.

## Results

### $V_{\lambda}$ and $C_{\lambda}$ Gene Sequences in Eco RI-Digested Embryo and Myeloma DNA

Our earlier preparative gel electrophoresis and mRNA hybridization experiments identified three  $\lambda$  gene-containing Eco RI DNA fragments in embryo cells. These are 8.6 kb  $C_{\lambda 1}$ , and 4.8 kb  $V_{\lambda}$  and 3.5 kb  $V_{\lambda}$  DNA fragments. The same experiments suggested the presence of a fourth 7.4 kb  $V_{\lambda} + C_{\lambda 1}$  DNA fragment (Tonegawa et al., 1976) in the  $\lambda_1$ -secreting myeloma cells. The results of Southern gel blotting experiments shown in Figure 1 confirmed the presence of the three and four DNA fragments in embryo and myeloma cells, respectively. All of the bands are quite sharp and there is no indication of size heterogeneity within the bands. Figure 1 also shows that the 7.4 kb fragment is absent in the DNA of a  $\kappa$  chain-producing myeloma (MOPC 321).

### Strategy for Cloning $\lambda$ Chain Genes from Cellular DNA

Previous studies on the hybridization kinetics of total DNA have demonstrated that mouse  $\lambda$  chain genes are reiterated in no more than a few copies per haploid genome (Tonegawa, 1976; Honjo et al., 1976; Tonegawa et al., 1976). Since the size of a mouse haploid genome is approximately  $2 \times 10^{12}$  daltons, and the average size of Eco RI fragments is approximately  $3 \times 10^6$  daltons, it would be necessary to screen over a million DNA clones in order to obtain a single  $\lambda$  chain gene if unfractionated DNA is to be used for in vitro recombination. Screening on this scale is not impossible but would be cumbersome. We therefore fractionated the Eco RI-digested DNA by preparative agarose gel electrophoresis and thus enriched it 10–20 fold for  $\lambda$  gene-positive DNA fragments. Some of these DNA fragments were further enriched by R loop formation with a purified  $\lambda_1$  mRNA (Tonegawa et al.,

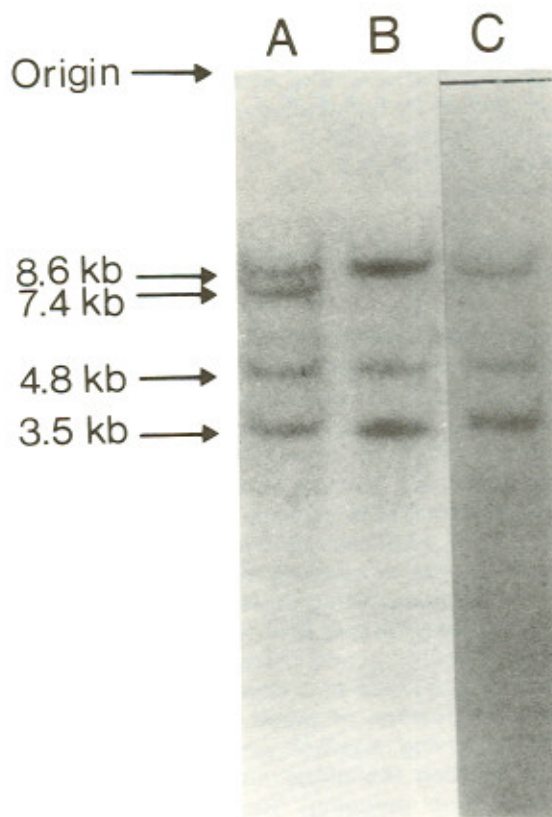


Figure 1.  $\lambda_1$  Gene Sequence-Containing DNA Fragments in Embryo and Myeloma Cells

High molecular weight DNAs extracted from 13 day old BALB/c embryos (B), myelomas H 2020 (a  $\lambda_1$  chain producer) (A) and MOPC 321 (a  $\kappa$  chain producer) (C) were digested to completion with Eco RI, electrophoresed on a 0.9% agarose gel, transferred to nitrocellulose membrane filters and hybridized with a nick-translated Hha I fragment of the plasmid B1 DNA (for details, see Experimental Procedures).

1977c). DNA preparations thus enriched were used for in vitro recombination with phage  $\lambda_{WES}$  DNA. The plaques arising from transfection of the host *E. coli* were screened by the rapid membrane filter method, using an iodinated  $\lambda$ -mRNA or the nick-translated clone B1 DNA as the hybridization probe.

Using this procedure, we have isolated four different DNA clones, Ig 25 $\lambda$ , Ig 13 $\lambda$  (Tonegawa et al., 1977c), Ig 99 $\lambda$  and Ig 303 $\lambda$  (Brack and Tonegawa, 1977), each carrying one of the four hybridization-positive Eco RI fragments that were detected in embryo DNA and in H 2020 myeloma DNA. Variations in isolation procedures, DNA sources and some of the basic characteristics of the four DNA clones are summarized in Table 1. Ig 25 $\lambda$  and Ig 303 $\lambda$  were cloned from the embryonic 8.6 kb fragment and the myeloma 7.4 kb fragment, respectively. These clones were isolated from DNA of the respective agarose gel fractions without further



Table 1. List of  $\lambda$  Chain Gene Clones

| Clone            | DNA Source       | DNA Pre-enrichment Steps           | Approximate <sup>a</sup> Pre-enrichment Factors | Screening Procedures and Probes  | Approximate Number of Plaques Screened | $\lambda$ Gene <sup>b</sup> Sequences Contained | References                                   |
|------------------|------------------|------------------------------------|---|--|--|---|--|
| Ig 99 $\lambda$  | Embryo<br>3.5 kb | Agarose gel R looping (one cycle)  | 300   |  | 3,000                                  | V <sub>AI</sub>                                 | This paper                                   |
| Ig 25 $\lambda$  | Embryo<br>8.6 kb | Agarose gel                        | 15  | Benton and Davis (1977) with nick-translated cDNA plasmid                | 80,000                                 | C <sub>AI</sub>                                 | This paper                                   |
| Ig 303 $\lambda$ | H2020<br>7.4 kb  | Agarose gel                        | 15  |  | 70,000                                 | V <sub>AI</sub> + C <sub>AI</sub>               | Brack and Tonegawa (1977)                    |
| Ig 13 $\lambda$  | Embryo<br>4.8 kb | Agarose gel R looping (two cycles) | 360   | Kramer, Cameron and Davis (1976) with <sup>32</sup> S]- $\lambda_1$ mRNA | 4,000                                  | V <sub>AII</sub>                                | Tonegawa et al. (1977); Hozumi et al. (1978) |

<sup>a</sup> See Tonegawa et al. (1976) for the definition of the enrichment factor.

<sup>b</sup> This column lists the  $\lambda$  gene sequences assigned to the cloned DNAs. See the text for additional details.

fractionation by R loop formation. The clones Ig 13 $\lambda$  and Ig 99 $\lambda$  were isolated from the embryonic 4.8 and 3.5 kb fragments, respectively. These fragments had been further enriched by R loop formation and equilibrium centrifugation in a CsCl<sub>2</sub> density gradient. Table 1 also gives a very rough estimate of the enrichment factor achieved in each of the four DNA preparations, as well as the number of phage plaques which were screened to isolate each of the four clones. The frequency of the immunoglobulin DNA-positive clones is consistent with our previous conclusion based upon hybridization kinetics: each of the four  $\lambda$  gene sequences is unique or nearly unique (Tonegawa et al., 1976).

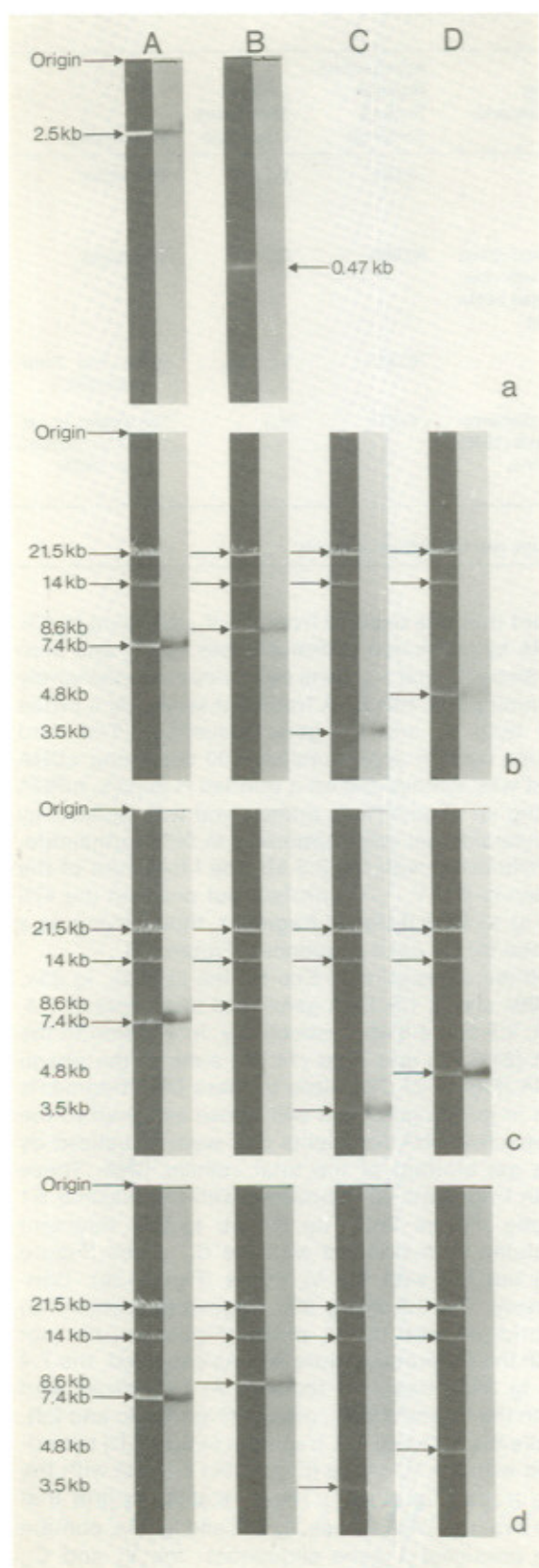
#### Types of $\lambda$ Gene Sequences Contained in the DNA Clones

The hybridization properties of the cellular DNA fragments that were used for cloning (see above) permit a prediction of the kinds of  $\lambda$  gene sequences (V<sub>A</sub>, C<sub>AI</sub> or V<sub>A</sub> plus C<sub>AI</sub>) contained in the isolated clones. They are listed in Table 1. Our previous sequencing studies confirmed that the sequence contained in the Ig 13 $\lambda$  encodes a V<sub>AII</sub> region (Hozumi et al., 1978; Tonegawa et al., 1978), whereas R loop mapping studies demonstrated that Ig 303 $\lambda$  contains both V<sub>A</sub> and C<sub>AI</sub> sequences (Brack and Tonegawa, 1977). To test the validity of the sequence assignments to the other two clones (Ig 99 $\lambda$  and Ig 25 $\lambda$ ), we carried out gel blotting (Southern) experiments with Eco RI-digested, cloned DNA using three different  $\lambda$  gene sequence probes. The first probe was the plasmid clone B1 that contains essentially the whole sequence of a  $\lambda_1$  gene. The second probe was a 470 bp DNA frag-

ment that was excised from the V<sub>AII</sub>-carrying Ig 13 $\lambda$  DNA by restriction endonucleases Hae III and Mbo II. Since V<sub>AI</sub> and V<sub>AII</sub> gene sequences are extensively homologous, this DNA fragment serves as a probe for both V<sub>AI</sub> and V<sub>AII</sub> gene sequences. The third probe was an approximately 400 base long cDNA that was synthesized on a purified H 2020  $\lambda_1$  mRNA using an oligo(dT)<sub>12-18</sub> primer, and was isolated by acrylamide gel electrophoresis in 98% formamide. It hybridized with the 2.5 kb Hha I fragment of the plasmid B1 (V<sub>A</sub> + C<sub>AI</sub> probe), but not with the 470 bp Ig 13 Mbo II-Hae III fragment, thus serving as a probe for C<sub>AI</sub> gene sequences (Figure 2a).

When digested with Eco RI, the Ig 303 $\lambda$ , Ig 25 $\lambda$ , Ig 99 $\lambda$  and Ig 13 $\lambda$  DNA generated fragments of 7.4, 8.6, 3.5 and 4.8 kb, respectively, in addition to the left (21.5 kb) and right (14 kb) arms of the phage DNA (Figure 2). The sizes of these DNA fragments are in good agreement with those assigned to the respective DNA fragments that were visualized by the gel blotting of the total cellular DNA. These DNA fragments all hybridized with the plasmid B1 probe (Figure 2b). The 8.6 kb Ig 25 $\lambda$  fragment (column B) hybridized with the C<sub>AI</sub> probe (Figure 2d) but not with the V<sub>A</sub> probe (Figure 2c). Conversely, the 3.5 kb Ig 99 $\lambda$  fragments (column C) hybridized with the V<sub>A</sub> probe (Figure 2c) but not with the C<sub>AI</sub> probe (Figure 2d). As expected, the 7.4 kb Ig 303 $\lambda$  fragment (column A) hybridized with both the V<sub>A</sub> and the C<sub>AI</sub> probes (Figures 2c and 2d), while the 4.8 kb Ig 13 $\lambda$  fragment (column D) hybridized with the V<sub>A</sub> probe (Figure 2c) but not with the C<sub>AI</sub> probe (Figure 2d). These results confirm that the two new DNA clones, Ig 25 $\lambda$  and Ig 99 $\lambda$ , contain the predicted  $\lambda$  gene sequences: the V<sub>A</sub> and C<sub>AI</sub>





sequences, respectively. Our current nucleotide sequencing studies have demonstrated that the  $V_{\lambda}$  sequence contained in Ig 99 $\lambda$  and Ig 303 $\lambda$  is of the  $\lambda_1$  type (N. Hozumi, O. Bernard and S. Tonegawa, unpublished observations).

#### Location of the $\lambda$ Chain Gene Sequence in the DNA Clones

The position of  $\lambda$  chain gene sequences in the Ig 25 $\lambda$  and Ig 99 $\lambda$  DNA clones was determined by R loop mapping using  $\lambda_1$  mRNA purified from H 2020 myeloma (White and Hogness, 1977).

R loop molecules formed with the 8.6 kb Ig 25 $\lambda$  fragment displayed a double loop structure composed of a 410 nucleotide R loop located 3.9 kb from one end and a 1.2 kb double-stranded DNA loop (Figure 3a and Table 2). This structure closely resembles the triple loop formed by Ig 303 $\lambda$  (Brack and Tonegawa, 1977), except that the Ig 25 $\lambda$  hybrids have a long RNA tail (~260 bases) instead of the second, smaller R loop (Figure 4 and Table 2). Since the Ig 25 $\lambda$  fragment showed homology only with  $C_{\lambda 1}$  and not with  $V_{\lambda}$  sequences (see above), we conclude that the 410 bp R loop contains the  $C_{\lambda 1}$  gene sequence and that the long RNA tail corresponds to the 5' end or V-coding part of the mRNA. A second, short RNA tail (~100 bases) sometimes observed at the other end of the R loop would correspond to the poly(A) sequence at the 3' end of the mRNA. The presence of the 1.2 kb DNA loop indicates that Ig 25 $\lambda$  DNA contains a short homology region that hybridizes to a region near the V-C junction of the mRNA molecule. This second homology region, which we call the J sequence, is separated from the  $C_{\lambda 1}$  sequence by 1200 bp. It is too short to be visualized as a separate R loop, but is strong enough to hold the double loop structure together.

The 3.5 kb Ig 99 $\lambda$  fragment formed a single R loop very similar to the one observed in Ig 13 $\lambda$  (Tonegawa et al., 1977c). It is 380 nucleotides long, lies in the middle of the DNA fragment (that is, 1.65–1.66 kb from either end) and carries a ~340

Figure 2. Type of  $\lambda$  Gene Sequences Contained in the  $\lambda$ -DNA Clones

In (a), the 400–430 base long transcripts of H 2020  $\lambda_1$  mRNA (a  $C_{\lambda 1}$  gene probe) were checked for the absence of possible contamination of  $V_{\lambda}$  gene sequences by hybridization with the 2.5 kb Hha I fragment of plasmid B1 (column A), and with the 0.47 kb Mbo II-Hae III fragment of Ig 13 $\lambda$  DNA (a  $V_{\lambda}$  gene probe) (column B). In (b–d), DNA from the four  $\lambda$  gene clones Ig 303 $\lambda$  (column A), Ig 25 $\lambda$  (column B), Ig 99 $\lambda$  (column C) and Ig 13 $\lambda$  (column D) were digested with Eco RI, electrophoresed in a 0.9% agarose gel, transferred and hybridized with either the  $V_{\lambda} + C_{\lambda 1}$  probe (Figure 2b), the  $V_{\lambda}$  probe (Figure 2c) or the  $C_{\lambda 1}$  probe (Figure 2d). In each pair of pictures, a gel stained with ethidium bromide (1  $\mu$ g/ml) is on the left and an autoradiogram of the blot is on the right. The lengths of the fragments were determined by Eco RI-digested  $\lambda$ -DNA electrophoresed in parallel.



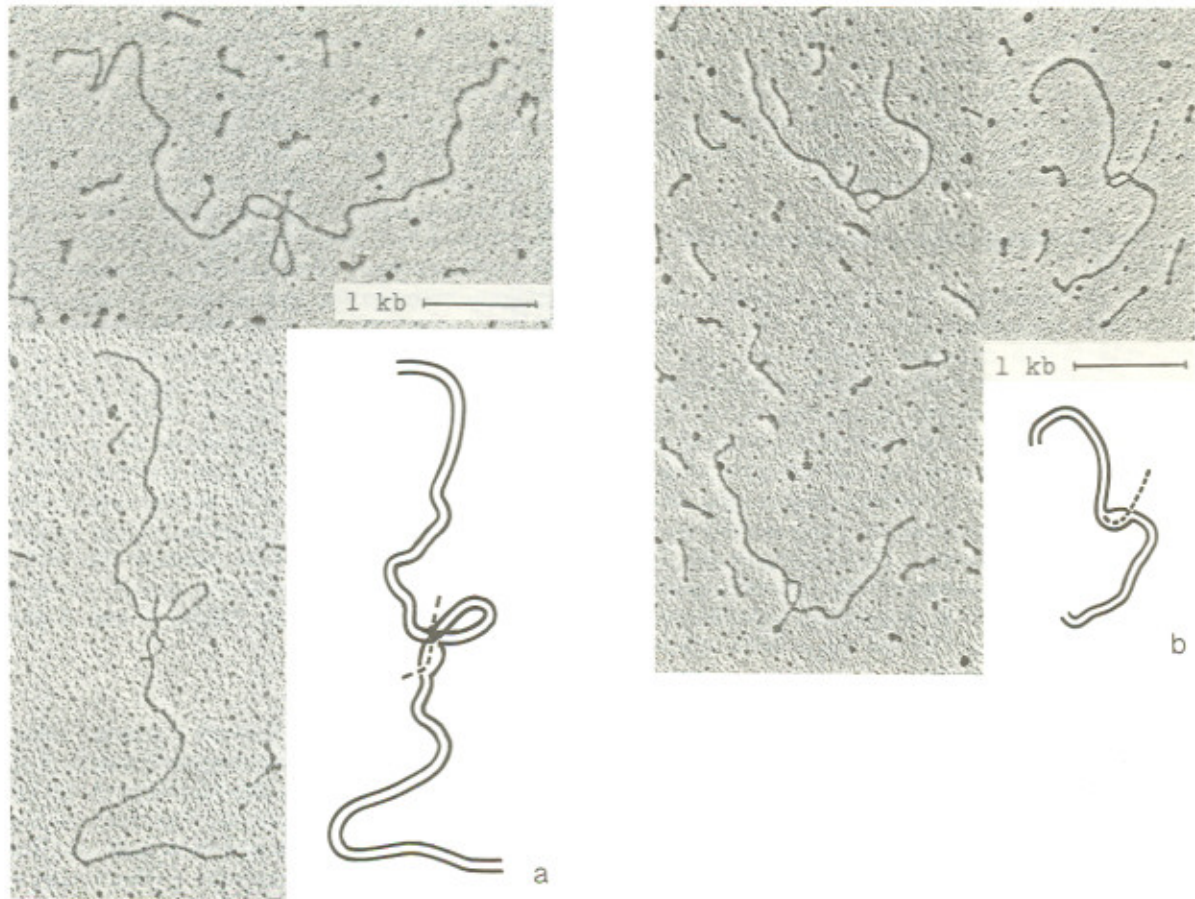


Figure 3. R Loop Molecules Obtained by Hybridizing HOPC 2020  $\lambda_1$  mRNA with the Eco RI Fragments of the DNA Clones  
(a) Ig 25 $\lambda$  DNA displays one R loop corresponding to the C gene, the double-stranded DNA loop and a long RNA tail corresponding to V sequences. The short tail observed in some molecules is the 3' poly(A) tail.  
(b) Ig 99 $\lambda$  DNA has one R loop corresponding to V sequences and a long RNA tail that is composed of the C gene sequences plus poly(A) tail.

Table 2. Length Measurements Made on Electron Micrographs of R Loop Molecules

| DNA Clone        | RNA/DNA* | F†   | N‡ | a               | b               | c               | d                  | e                  | f               |
|------------------|----------|------|----|-----------------|-----------------|-----------------|--------------------|--------------------|-----------------|
| Ig 13 $\lambda$  | 8        | 27.2 | 34 | 3.29 $\pm$ 0.15 | 1.06 $\pm$ 0.22 |                 | 0.388 $\pm$ 0.103  | 0.400 $\pm$ 0.200§ |                 |
| Ig 99 $\lambda$  | 5        | 81.2 | 40 | 1.65 $\pm$ 0.07 | 1.66 $\pm$ 0.11 |                 | 0.375 $\pm$ 0.050  | 0.344 $\pm$ 0.099§ |                 |
| Ig 25 $\lambda$  | 10       | 84.0 | 36 | 3.09 $\pm$ 0.14 | 3.89 $\pm$ 0.15 | 1.22 $\pm$ 0.08 | 0.255 $\pm$ 0.065§ | 0.412 $\pm$ 0.048  | 0.10 $\pm$ 0.05 |
| Ig 303 $\lambda$ | 66       | 50.4 | 53 | 1.66 $\pm$ 0.08 | 3.75 $\pm$ 0.20 | 1.20 $\pm$ 0.09 | 0.380 $\pm$ 0.030  | 0.440 $\pm$ 0.040  | 0.09 $\pm$ 0.04 |

\* Molar ratio of mRNA over DNA.

† (F) frequency of R loop molecules (in %), calculated by screening 250 molecules per sample.

‡ (N) number of molecules measured.

§ RNA tails do not always fully spread out and, therefore, these values cannot be compared directly with those of corresponding R loops. For identification of the various segments, see Figure 4. The values of mean lengths and standard deviations are given in kb.

nucleotide RNA tail at one end (Figure 3b). Since this DNA fragment contains a  $V_{H1}$  sequence (see above), the R loop should contain the  $V_{H1}$  sequence and the RNA tail should correspond to the  $C_{H1}$  sequence plus the poly(A).

Table 2 and Figure 4 summarize the data ob-

tained from the R loop mapping of Ig 25 $\lambda$  and Ig 99 $\lambda$  DNA clones. We have previously reported electron microscopic characterization of R loops formed with Ig 13 $\lambda$  and Ig 303 $\lambda$  DNA clones. Some of the data obtained from the electron microscopic characterization of these two  $\lambda$  gene clones are



also included in Table 2 and Figure 4 because of their significance in the present context of comparative studies. The position of the R loop is unique in each case. In no case have we observed multiple R loops that could indicate the presence of more than one copy of any particular kind of the  $\lambda$  gene sequence on a single DNA fragment.

#### Sequence Homology between the Cloned DNA Fragments

Analysis of sequence homology between the cloned DNA fragments, both within the  $\lambda$  chain genes and in the adjacent regions, may aid in discovering the mechanism by which somatic rearrangement of immunoglobulin genes occurs. The four cloned fragments were therefore hybridized in various combinations and the heteroduplex molecules were analyzed by electron microscopy. We describe below the observations made with various combinations of the three  $\lambda_1$  DNA clones Ig 99 $\lambda$ , Ig 25 $\lambda$  and Ig 303 $\lambda$ . A summary of the results is given in Table 3 and Figure 6. Analogous experiments using the  $V_{\lambda II}$  clone Ig 13 $\lambda$  and the  $\lambda_1$  DNA clones will be described elsewhere.

#### Ig 25 $\lambda$ versus Ig 303 $\lambda$

Ig 25 $\lambda$  hybridized with Ig 303 $\lambda$  DNA formed Y-shaped heteroduplex molecules with two single-stranded and one double-stranded arms (Figure

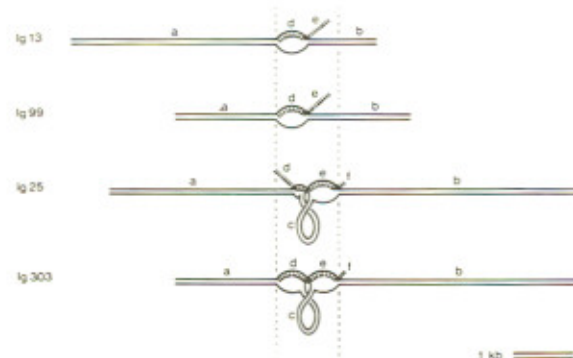


Figure 4. Schematic Representation of R Loops on All Four Cloned Mouse Ig DNA Fragments

The relation between the four clones is shown. See Table 2 for the length measurements of various parts of the R loops.

5a). The lengths of the three arms are 2.87 (long single strand), 1.98 (short single strand) and 5.47 kb (double strand) (Table 3). These measurements indicate that the shorter single-stranded arm and one strand of the double-stranded arm correspond to the Ig 303 $\lambda$  DNA, whereas the longer single-stranded arm and the other strand of the double-stranded arm derive from the Ig 25 $\lambda$  DNA. In the entire 5.5 kb homology region, we observed no local mismatches that would indicate partial non-homology between the two strands. R loop mapping had shown that in both DNA fragments, the  $C_{\lambda I}$  gene lies between 3.8 and 4.2 kb from one end (Table 2). The 5.5 kb homology region therefore contains the  $C_{\lambda I}$  gene. In addition, most if not all of the 1.2 kb DNA segment separating the V and C sequences in the Ig 303 $\lambda$  fragment is highly homologous to the DNA segment of similar length that separates the J and C sequences in the Ig 25 $\lambda$  fragment. The measurements indicate that the J sequence lies very near or at the branch point of the heteroduplex (compare Tables 2 and 3 and Figure 6).

#### Ig 99 $\lambda$ versus Ig 303 $\lambda$

These two fragments also formed Y-shaped heteroduplex molecules with two single-stranded (5.48 and 1.53 kb) and one double-stranded (1.98 kb) arms (Figure 5b and Table 3). The sum of the lengths of the long single-stranded arm and the double-stranded arm corresponds to the length of the Ig 303 $\lambda$  fragment, whereas the short single strand and the second strand of the double-stranded region belong to the Ig 99 $\lambda$  fragment. Again, we did not observe any local nonhomology in the double-stranded part of the heteroduplex. These results indicate that on one end of each fragment, the two DNAs are highly homologous. The  $V_{\lambda}$  sequence lies between 1.66 and 2.0 kb from one end of both Ig 303 $\lambda$  and Ig 99 $\lambda$  (Table 2). Thus the V gene is within the 2.0 kb homology region and lies near the fork of the Y-shaped heteroduplex (Figure 6).

#### Ig 303 $\lambda$ , Ig 25 $\lambda$ and Ig 99 $\lambda$

The results described in the above two sections strongly suggest that a large part of the Ig 303 $\lambda$

Table 3. Length Measurements of Heteroduplex Molecules

| Heteroduplex   | N  | a               | b               | c               | d               |
|--|----|-----------------|-----------------|-----------------|-----------------|
| (1) Ig 303 $\lambda$ $\times$ Ig 25 $\lambda$                          | 51 | 5.47 $\pm$ 0.25 | 1.98 $\pm$ 0.22 | 2.87 $\pm$ 0.20 | —               |
| (2) Ig 303 $\lambda$ $\times$ Ig 99 $\lambda$                          | 48 | 5.48 $\pm$ 0.48 | 1.98 $\pm$ 0.52 | —               | 1.53 $\pm$ 0.13 |
| (3) Ig 303 $\lambda$ $\times$ Ig 99 $\lambda$ $\times$ Ig 25 $\lambda$ | 27 | 5.54 $\pm$ 0.30 | 1.87 $\pm$ 0.19 | 2.85 $\pm$ 0.27 | 1.45 $\pm$ 0.15 |

For identification of the various molecules (1-3) and the different segments a-d, see Figure 6. All lengths (mean length and standard deviation) are given in kb. (N) number of measured molecules.



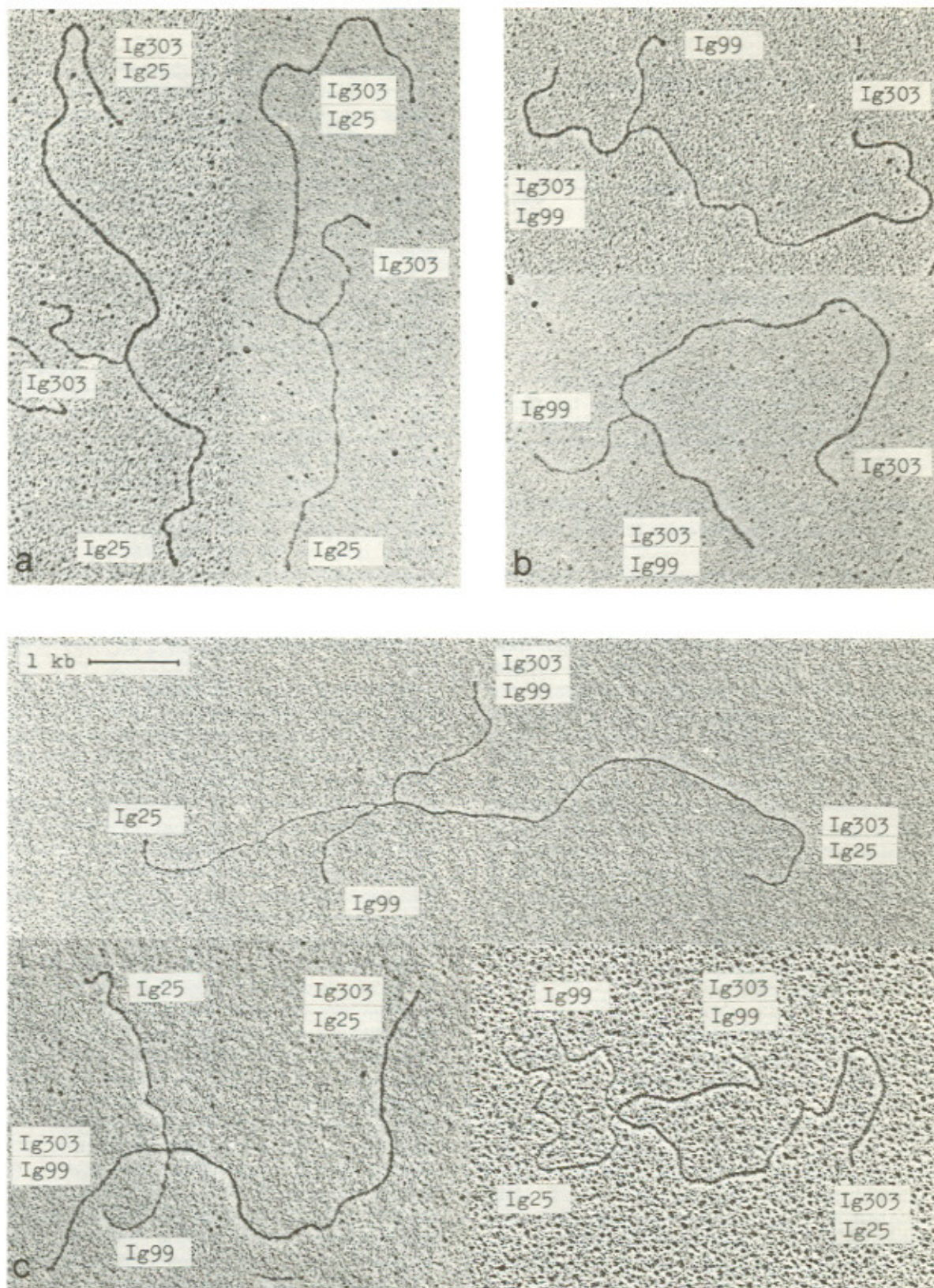


Figure 5. Heteroduplex Molecules Formed by Various Combinations of the Three  $\lambda_1$  Sequence-Containing Cloned DNA Fragments

(a) Myeloma DNA Ig 303 $\lambda$  versus embryo DNA Ig 25 $\lambda$ .

(b) Myeloma DNA Ig 303 $\lambda$  versus embryo DNA Ig.

(c) Combination of all three clones gave double heteroduplex structures.



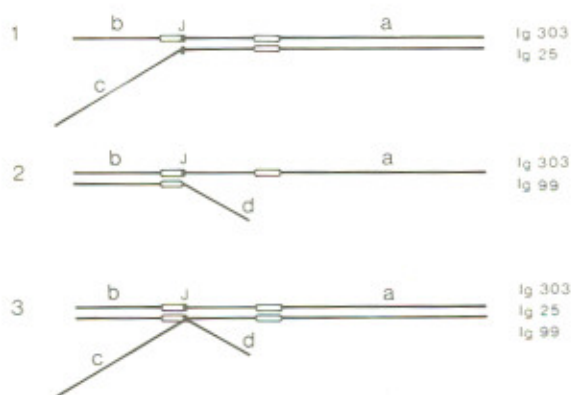


Figure 6. Interpretation of the Heteroduplex Molecules  
See Table 3 for the lengths of the various segments. The positions of V and C gene sequences (white boxes) were deduced from R loop molecules. See the Discussion for the exact position of the J sequence.

fragment (5.5 kb) is homologous to the Ig 25 $\lambda$  DNA, whereas the rest of the molecule (2 kb) is homologous to the Ig 99 $\lambda$  DNA.

The three DNA fragments were mixed, denatured and annealed to confirm this hypothesis. A low proportion of the molecules displayed a double heteroduplex structure or cruciform structure composed of two double-stranded and two single-stranded arms (Figure 5c). Measurements of the arms allowed us to assign each part of the hybrid to the three DNA fragments, as illustrated in Figure 5c. The observed structure can be formed only if each of the two ends of the Ig 303 $\lambda$  fragment is homologous to only one of the two DNA fragments Ig 99 $\lambda$  or Ig 25 $\lambda$ . The two homology regions meet at a point which corresponds to the J sequence in Ig 25 $\lambda$  and lies near the junction between the V sequence and the 1.2 kb intervening sequence of Ig 303 $\lambda$  DNA. The myeloma DNA fragment Ig 303 $\lambda$  thus seems to be entirely composed of DNA segments that are homologous to parts of the two embryonic fragments.

#### Restriction Enzyme Cleavage Sites Present in the Cloned DNA Are Also Present on the Cellular DNA at the Corresponding Positions

It has been questioned whether cloned DNA fragments always retain the sequence organization found in cellular DNA. In the present case, the excellent agreement in the sizes of the cloned DNA fragments and the corresponding cellular DNA fragments (as visualized by the Southern blotting technique) makes it particularly improbable that any gross sequence rearrangement has resulted from the cloning procedures. Nevertheless, it was desirable to obtain additional evidence for retention of the original chromosomal sequence organization in the cloned DNA fragments. To this end,

we compared positions of some of the restriction enzyme cleavage sites on the cloned Eco RI fragments and the corresponding Eco RI fragments of embryo DNA and myeloma DNA. In so doing we chose the enzymes that cleave the cloned DNAs within an intervening sequence (Tonegawa et al., 1978) to enable us to confirm at the same time the presence of the corresponding intervening sequence in the cellular DNA.

The 4.8 kb Ig 13 $\lambda$  DNA carries only one Pst I site, which lies 3.3 kb from one end (Hozumi et al., 1978). This site is within the 93 base intervening sequence that separates the leader-coding and the V region sequences (Tonegawa et al., 1978). The map is illustrated in Figure 7a. When digested with Pst I, the Ig 13 $\lambda$  DNA fragment gave the two expected fragments of 3.3 and 1.5 kb, of which only the latter gave a band of hybridization with the plasmid B1 probe. The sequence homology within the 3.3 kb fragment is short (~120 bases), and a very faint band of the corresponding size was visible only on the original film. An essentially identical band pattern was obtained when the 4.8 kb embryo DNA fraction obtained by Eco RI digestion and electrophoresis was redigested with Pst I and analyzed by the gel blotting technique (Figure 7a).

As the map in Figure 7b shows, the 3.5 kb Ig 99 $\lambda$  DNA fragment also carries only one Pst I site that lies within the intervening sequence analogous to the intervening sequence of Ig 13 $\lambda$  (N. Hozumi, O. Bernard and S. Tonegawa, unpublished results). Digestion of Ig 99 $\lambda$  DNA with this enzyme generated two fragments of 1.6 and 1.9 kb, of which only the latter exhibited easily detectable hybridization. Pst I digestion of the 3.5 kb embryo DNA fragment that had been obtained by Eco RI digestion gave a fragment that co-migrated with the 1.9 kb Pst I-Eco RI fragment of Ig 99 $\lambda$  DNA.

Xba I maps of Ig 25 $\lambda$  and Ig 303 $\lambda$  DNA are given in Figures 7c and 7d. On each of the two clones, one Xba I site is present within the 1.2 kb intervening sequence that separates the V- and C-coding sequences on Ig 303 $\lambda$  DNA and the J- and C-coding sequences on Ig 25 $\lambda$  DNA (N. Hozumi, O. Bernard, C. Brack and S. Tonegawa, unpublished observation). In agreement with the maps, gel blots of Xba I-digested Ig 25 $\lambda$  and Ig 303 $\lambda$  DNAs gave 3.2 kb, and 3.2 and 0.9 kb hybridization-positive fragments, respectively. These fragments were also present in the digests of the corresponding cellular DNA fractions—that is, the 8.6 kb embryo (Figure 7c) and 7.4 kb H 2020 myeloma DNAs (Figure 7d).

In summary, restriction enzyme cleavage sites uniquely identified within the intervening sequences of the cloned DNAs are also present at the corresponding positions of the cellular DNA. Over-



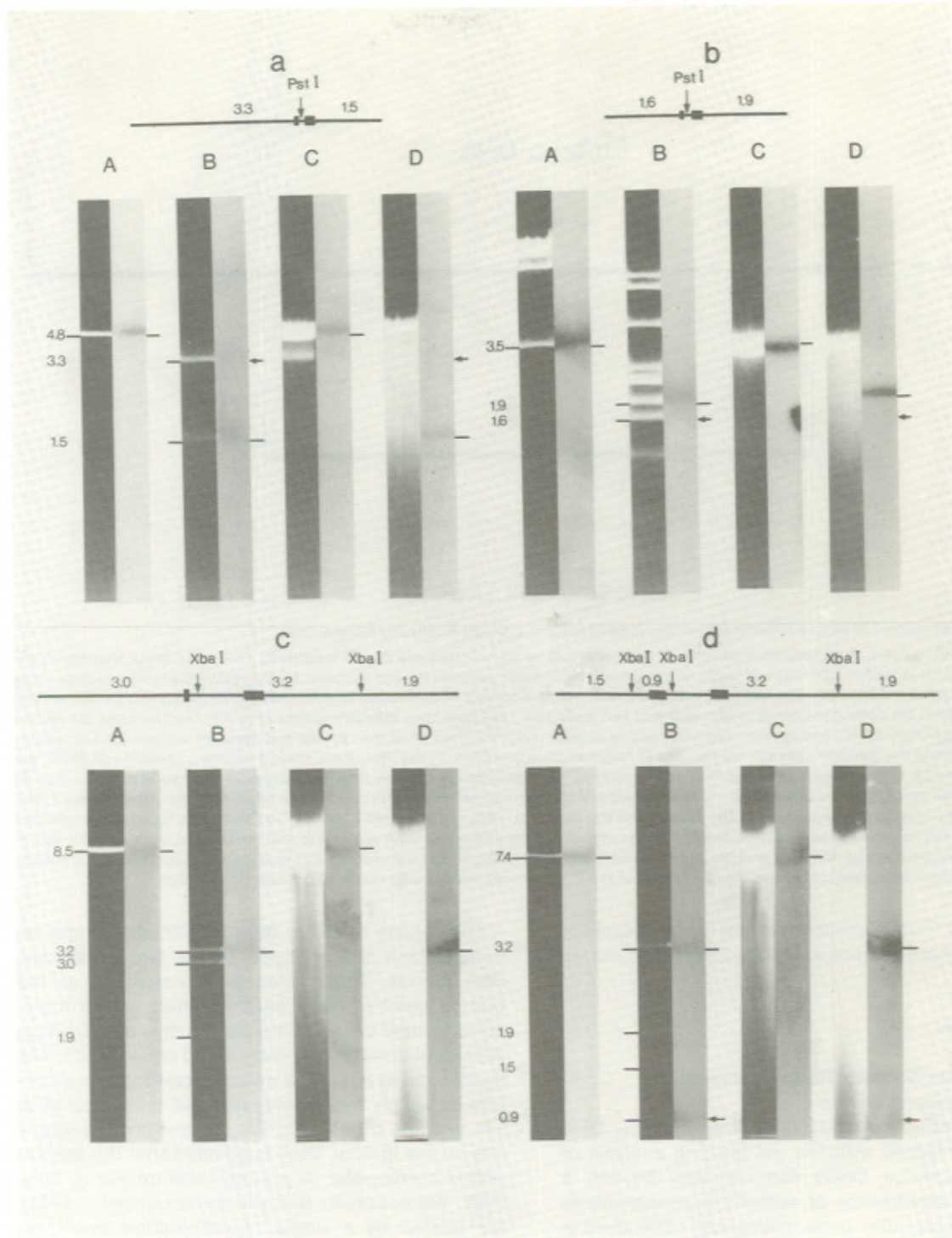


Figure 7. Comparison of Pst I and Xba I Cleavage Sites on Cloned DNAs and Corresponding Cellular DNA Fragments Generated by Eco RI Digestion

Whole embryo or H 2020 myeloma DNA was digested with Eco RI and fractionated by preparative agarose gel electrophoresis. The 4.8 kb  $V_{\lambda 11}$ -containing (a), 3.5 kb  $V_{\lambda 1}$ -containing (b) or 8.6 kb  $C_{\lambda 1}$ -containing (c) embryonic DNA fragments, and the 7.4 kb  $V_{\lambda 1} + C_{\lambda 1}$ -containing (d) myeloma DNA fragment were identified by hybridization with  $^{125}\text{I}$ -labeled  $\lambda_1$  mRNA, isolated, digested with Pst I (for the 4.8 and 3.5 kb fragments) or with Xba I (for the 8.6 and 7.4 kb fragments), electrophoresed on a 0.8% agarose gel, blotted and hybridized with the nick-translated plasmid BI DNA. The Eco RI DNA fragments of the corresponding  $\lambda$  chain gene clones Ig 13 $\lambda$  (a), Ig 99 $\lambda$  (b), Ig 25 $\lambda$  (c) and Ig 303 $\lambda$  (d) were also digested with Pst I (for Ig 13 $\lambda$  and Ig 99 $\lambda$ ) or with Xba I (for Ig 25 $\lambda$  and Ig 303 $\lambda$ ) and subjected to the same procedures. In each double column designated by A, B, C and D, gels stained with ethidium bromide are on the left and autoradiograms of the blots are on the right. Columns A and B are undigested and digested cloned DNAs, respectively, while columns C and D are undigested and digested cellular DNA fragments, respectively. At the top of each of the panels are Pst I or Xba I maps of the cloned DNAs. In (a and b), the two filled boxes indicate positions of leader- (small box) and V- (large box) coding sequences. In (c), the boxes indicate positions of J- (small box) and C- (large box) coding sequences, while in (d), they indicate those of V- (left) and C- (right) coding sequences. In (b), columns A and B, whole Ig 99 $\lambda$  phage DNA (instead of the 3.5 kb Eco RI fragment from the phage) was used after digestion with Eco RI. Bands visible only on the original autoradiograms are indicated by arrows. Numbers indicate fragment sizes in kilobases.



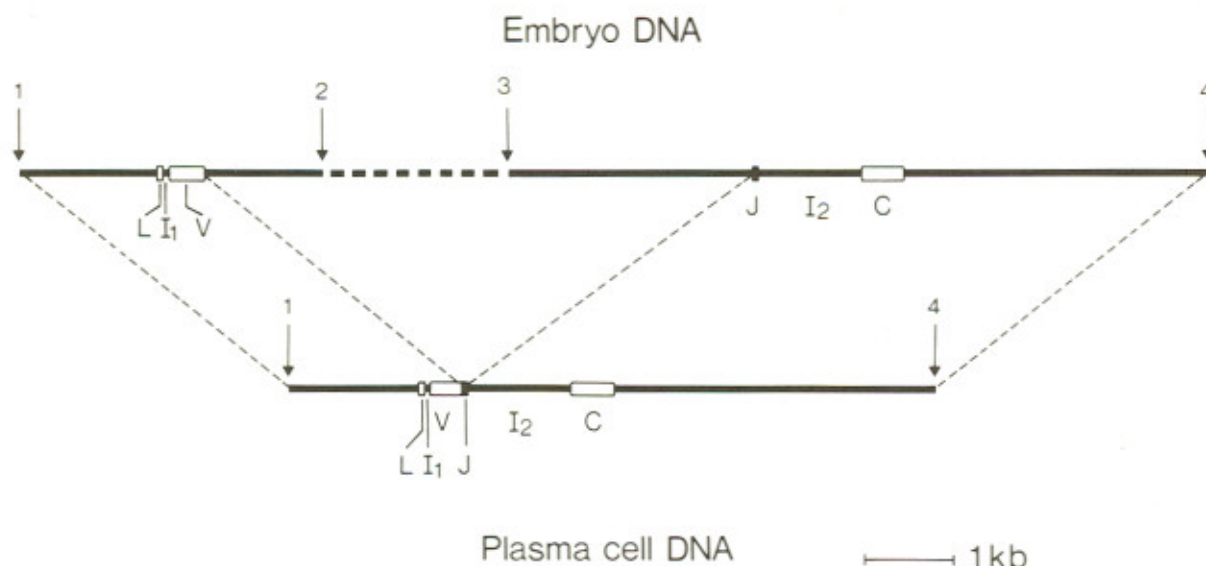


Figure 8. Arrangement of Mouse  $\lambda_1$  Gene Sequences in Embryos and  $\lambda_1$  Chain-Producing Plasma Cells

In embryo DNA, a full  $\lambda_1$  gene sequence consists of two parts that lie on two separate Eco RI fragments. On one of these fragments, the coding sequence is further split into two parts, one for most of the leader peptides (L) and the other for the rest of the leader peptides plus the variable region peptides (V). The two coding sequences are separated by a 93 nucleotide long intervening sequence ( $I_1$ ). On the second Eco RI fragment, the coding sequence is also split into two parts by a 1250 base long intervening sequence ( $I_2$ ). The two parts are for the constant region peptides (C) and approximately 13 residue peptides near the junction of the variable and constant regions (J). The relative orientation of and the distance between the two Eco RI fragments are unknown. In the DNA of a  $\lambda_1$  chain-producing myeloma (H 2020), the  $\lambda_1$  gene sequence is rearranged as a result of one (or more) recombination(s) that involves sequences in the two embryonic Eco RI fragments. One recombination takes place at the ends of the V and the J sequences and brings the two sequences into direct contact. The limits of the corresponding sequences in the embryo and the myeloma DNAs are indicated by thin dotted lines. The figure is not intended to imply that the recombination results in deletion or looping-out of the embryonic DNA sequences that lie between the V and the Eco RI site 2, or between the Eco RI site 3 and the J. Neither is it intended to imply that the embryo and myeloma V sequences are identical. Additional short intervening sequences may be present in the C sequences. Arrows with numbers indicate Eco RI sites.

all results thus confirm that no gross sequence rearrangements can have resulted from the cloning operation.

## Discussion

### Evidence for Somatic Rearrangement of Immunoglobulin Genes

The heteroduplex analysis of the three  $\lambda_1$  DNA clones combined with the gel blotting analysis of the total cellular DNAs demonstrated beyond a doubt the occurrence of somatic rearrangements of immunoglobulin gene sequences. The double heteroduplex structure generated by co-annealing the three cloned  $\lambda_1$  DNAs is incompatible with the alternative interpretation (see Hozumi and Tonegawa, 1976; also Introduction) of the results obtained by restriction enzyme mapping of total cellular DNA. The observed heteroduplex structures cannot be artifacts of DNA cloning. The length of each of the cloned DNA fragments coincides well with that of the corresponding cellular DNA fragments visualized by the gel blotting technique. Furthermore, certain restriction enzyme sites identified in the cloned DNAs are also present at corresponding positions in cellular DNA (Figure 7).

The results reported here identified a single recombination site on each of the two embryonic DNA clones. These sites were visualized as the branch point of the Y-shaped heteroduplex molecules formed between the cloned myeloma DNA (Ig 303 $\lambda$ ) and either of the two cloned embryonic DNAs (Ig 99 $\lambda$  and Ig 25 $\lambda$ ). The double heteroduplex structure in which two single-stranded tails, one of Ig 99 $\lambda$  and the other of Ig 25 $\lambda$ , extend from a single site on the Ig 303 $\lambda$  DNA suggested that the branch points correspond to a single site on the Ig 303 $\lambda$  DNA. We conclude that the three cloned  $\lambda_1$  DNAs are related by a single recombination event, as illustrated in Figure 8. Embryonic DNA recombines at the right end of the V sequence and the left end of the J sequence to generate the sequence arrangement present in myeloma DNA. The entire 1.2 kb intervening sequence in the Ig 303 $\lambda$  DNA originates from the Ig 25 $\lambda$  DNA. Measurements of various parts of the R loops and heteroduplex structures described in the present work are entirely consistent with this model. Our previous nucleotide sequencing studies have shown that the  $V_{H1}$  region encoded by Ig 13 $\lambda$  is approximately 15 residues shorter than the V region defined by amino acid sequencing (Tonegawa et al., 1978).



Since the  $V_{\lambda 1}$  and  $V_{\lambda 2}$  genes are closely related in evolution, we expect that the  $V_{\lambda 1}$  gene carried in Ig 99 $\lambda$  has a similar structure. The missing nucleotide sequences are most probably provided by the J region of Ig 25 $\lambda$  DNA. It should be added that the present experimental results do not allow us to distinguish among the several alternative modes of sequence rearrangement discussed previously (Hozumi and Tonegawa, 1976), because the DNA segments involved in the rearrangement events may be much longer than the DNA fragments studied so far.

The presence in H 2020 DNA of embryonic DNA fragments carrying  $\lambda_1$  sequences (Figure 1) is consistent with the hypothesis that the joining of V and C sequences takes place in only one of two homologous chromosomes present in a diploid plasma cell. We have previously reported an analogous observation concerning  $\kappa$  chain gene sequences (Tonegawa et al., 1977a). If this hypothesis is correct, it would conveniently explain the allelic exclusion of immunoglobulin gene expression (Pernis et al., 1965). The results, however, are subject to other interpretations. Although the hybridization kinetics indicated that there are no more than a few copies of germ line  $V_{\lambda 1}$  DNA sequences and that the number is too few to account for the observed diversity in  $V_{\lambda 1}$  regions, they could not definitely prove that there is only one copy of the  $V_{\lambda 1}$  DNA sequence per haploid genome (Tonegawa, 1976; Honjo et al., 1976; Tonegawa et al., 1976). The same degree of uncertainty also applies to the copy number of  $C_{\lambda}$  DNA sequences. There might be, for instance, two copies each of  $V_{\lambda}$  and  $C_{\lambda}$  DNA sequences per haploid genome, and joining of V and C might take place on both of the homologous chromosomes, although only in one of the two gene pairs on a single chromosome. Another possibility is that both  $V_{\lambda}$  and  $C_{\lambda}$  DNA sequences duplicate during lymphocyte differentiation and that the joining involves the duplicated copies. Yet more trivial possibilities arise because myeloma cells are polyploid (Yoshida, Imai and Potter, 1968).

#### Gene in Pieces

R loop mapping demonstrated the presence of a 1.2 kb intervening sequence on both the Ig 303 $\lambda$  and the Ig 25 $\lambda$  DNA. The resolution of cytochrome spreadings will not allow detection of intervening sequences shorter than approximately 100 nucleotides. Indeed, the 93 base long intervening sequence near the region corresponding to the amino terminal of the Ig 13 $\lambda$  DNA was revealed only by nucleotide sequence determination (Tonegawa et al., 1978). Our recent nucleotide sequencing studies (O. Bernard, N. Hozumi and S. Tonegawa, manuscript in preparation) established that both Ig 303 $\lambda$  and Ig 99 $\lambda$  DNA contain an intervening se-

quence equivalent to that of Ig 13 $\lambda$  both in length and position. These findings are incorporated into Figure 8. It should be added that additional short intervening sequences may be revealed in the C DNA sequence by the nucleotide sequencing now in progress in this laboratory. In any case, in each of the three  $\lambda_1$  DNA clones, the protein-coding sequences are arranged in discrete pieces. For instance, the somatically rearranged, complete  $\lambda_1$  gene in the myeloma (Ig 303 $\lambda$ ) consists of at least three DNA segments, one coding for the leader, one for the V region and one for the C region. The intervening sequences separating the three coding segments are present in the original mouse DNA and are not introduced during the cloning procedure. The unique restriction enzyme cleavage sites that have been identified within the intervening sequences of the cloned DNA were also demonstrated in the uncloned cellular DNA (Figure 7).

Recent studies on other genes of eucaryotes (Glover and Hogness, 1977; Jeffreys and Flavell, 1977; Wellauer and Dawid, 1977; Breathnach, Mandel and Chambon, 1977; Goodman, Olson and Hall, 1978; Tilghman et al., 1978a) and their viruses (Klessig, 1977; Berget, Moore and Sharp, 1977; Aloni et al., 1977; Chow et al., 1977) have revealed several cases of this unexpected gene structure: protein-encoding DNA interspersed with silent sequences. Intervening sequences are probably transcribed together with the protein-encoding DNA and subsequently excised during the maturation of pre-mRNA. One recent experiment concerning the mouse globin  $\beta$  chain gene seems to support this hypothesis (Tilghman et al., 1978b). These observations have led us to propose an additional evolutionary pathway for the creation of genes in higher organisms (Tonegawa et al., 1978). By this pathway, a new gene can be created from two or more separate DNA segments upon the emergence, at the boundary of the DNA segments, of mutations that generate signals for RNA splicing. If the new polypeptide chain coded by the spliced RNA has survival value, such mutations may become fixed in evolution. Since the splicing does not have to be 100% efficient, creation of the new gene need not destroy the old—usually a disadvantageous event.

#### Somatic Rearrangement as a Mechanism for Gene Control in Cell Differentiation

We assume that RNA splicing is an intramolecular reaction. This will restrict the operation of the gene creation mechanism described above to the space of a single transcription unit. One way to enlarge the effectiveness of this gene creation mechanism is to shuffle DNA segments by introducing them into transcription units. DNA sequences thus newly introduced are then available for splicing with



preexisting sequences. Actual evolutionary use of such DNA sequences depends upon the emergence of mutations leading to new splicing signals at proper positions in the transcription unit. We hypothesize that many genes in higher organisms have arisen through such evolutionary processes.

Does a higher organism utilize such a gene creation mechanism in the normal process of cell differentiation? We believe that the immunoglobulin genes are the perfect example. The J sequence seems to play a key role here by providing a "bridge" between DNA recombination and RNA splicing. Its left half most probably contains a nucleotide sequence for a site-specific recombination with the V DNA segment, while the sequence in the right half would almost certainly be involved in the RNA splicing event that connects the V and C sequences.

When a gene is created in this manner during ontogeny only in a particular subpopulation of cells composing an organism, the recombination itself can provide a novel mechanism for gene control in cell differentiation. The somatic rearrangement involving particular immunoglobulin gene sequences seems to be restricted to a small subpopulation of cells. Arrangement of a  $\kappa$  chain sequence in DNAs of several nonlymphatic adult tissues was identical to that of embryo DNA when analyzed by the electrophoresis-hybridization assay (Tonogawa et al., 1977a). Analogous experiments carried out using DNA from a  $\lambda_1$  chain-producing myeloma and a  $\kappa$  sequence probe and vice versa indicated that there is a mutual exclusion in the rearrangement of  $\kappa$  and  $\lambda$  chain DNA sequences (Tonogawa et al., 1977a). The results shown in Figure 1 provided additional evidence—that is, that the 7.4 kb Eco RI fragment carrying the rearranged full  $\lambda_1$  gene sequence is absent in the DNA of  $\kappa$ -producing MOPC 321.

One can postulate many variations of a gene control mechanism operated by somatic DNA rearrangement at the molecular level. For instance, in the case of immunoglobulin genes, a V sequence-carrying DNA segment that has been transcriptionally silent may be excised and inserted into a constitutively active transcription unit containing the C sequence. Alternatively, the act of V DNA insertion itself may create a new promoter at the very site of insertion. As we discussed previously (Brack and Tonogawa, 1977), a gene control mechanism directly dependent upon somatic sequence rearrangement in DNA seems to fulfill most easily the "one lymphocyte clone—one light chain" rule of the immune system. The questions concerning how generally somatic DNA rearrangement occurs and how the rearrangement-dependent gene control mechanism operates in developmental pathways remain to be answered.

## Experimental Procedures

### Bacteria and Phages

*Escherichia coli* LE 392/ThyA, DP50 (Su II<sup>+</sup>, Su III<sup>+</sup>) and phage  $\lambda$ gt<sub>10</sub>- $\lambda$ B were gifts from P. Leder and his colleagues (NIH, Bethesda, Maryland). (Leder, Tiemeier and Enquist, 1977). *E. coli* 803 (r<sub>h</sub><sup>-</sup>, m<sub>k</sub><sup>-</sup>, Su III<sup>+</sup>) originally from K. and N. E. Murray (University of Edinburgh) was obtained through W. Arber (Biocenter, Basel, Switzerland); *E. coli*  $\chi$ 1776 was obtained from R. Curtis and his colleagues (University of Alabama); plasmid pCRI was from J. Carbon (University of California, Santa Barbara). The cloning experiments were carried out in a P3 facility in accordance with the NIH Guidelines issued in June 1976.

### Enrichment of $\lambda$ Chain Sequence-Positive Eco RI DNA Fragments

Extraction of high molecular weight DNA, preparative agarose gel electrophoresis, use of iodinated  $\lambda_1$  mRNA (from H 2020 myeloma) in the detection of DNA fragments carrying immunoglobulin V and C sequences, and extraction of DNA fragments from agarose gel have all been previously described (Hozumi and Tonegawa, 1976). Procedures for preparative R loop formation based upon the method of Thomas, White and Davis (1976) have also been described previously (Tonogawa et al., 1977c). For the 4.0 kb embryonic DNA fragments, only one cycle of centrifugation in a CsCl density gradient was carried out.

### Ligation and Transfection

Ligation and transfection were carried out as described by Tonogawa et al. (1977c), except that the left and right arms of  $\lambda$ gt<sub>10</sub>- $\lambda$ B DNA were prepared from  $\lambda$ gt<sub>10</sub>- $\lambda$ B by preparative agarose gel electrophoresis. More specifically, for the 7.4 kb embryonic fragments, 17.5  $\mu$ g of gel-purified DNA and 67  $\mu$ g of the  $\lambda$ gt<sub>10</sub>- $\lambda$ B DNA arms were ligated in a volume of 0.7 ml. For the 3.5 kb embryonic fragments, 14  $\mu$ g of DNA purified by R loop formation and 29  $\mu$ g of the  $\lambda$ gt<sub>10</sub>- $\lambda$ B DNA arms were ligated in a volume of 0.75 ml.

### Plaque Screening by in Situ Hybridization

Procedures described by Benton and Davis (1977) were followed, except that nick-translated, cloned  $\lambda$  chain cDNA (in pCRI) was used as the hybridization probe. Phage particles and naked DNA in the plaques were transferred to membrane filter disks (Schleicher-Schuell BA85, diameter 8.0 cm). The filter disks were dipped for ~1 min into a denaturing solution (0.1 N NaOH and 1.5 M NaCl) and then also for ~1 min into a neutralizing solution [0.2 M Tris-HCl (pH 7.5), 0.3 M NaCl, 0.03 M Na-citrate (pH 7.5)], dried for a few hours at room temperature and baked in vacuo at 80°C for 2 hr. Up to 100 baked filters were incubated in a single batch at 65°C for 6 to 12 hr in a precoat solution consisting of 1 x Denhardt's solution [0.02% each of Ficoll, bovine serum albumin and polyvinylpyrrolidone (Denhardt, 1966)], 6 x SSC, 0.05 M PIPES-NaOH (pH 7.8) and 0.1 M KPO<sub>4</sub> (pH 7), using 6 ml per filter, and then transferred without drying to a hybridizing solution of the following composition: nick-translated plasmid B1 DNA, 3-6 x 10<sup>5</sup> cpm per filter; 100  $\mu$ g/ml of sonicated calf thymus DNA (Sigma); 50  $\mu$ g/ml of sonicated *E. coli* DNA (Sigma); 5  $\mu$ g/ml poly(A) (Miles); 1 mM Tris-HCl (pH 7.5); 0.2 mM Na<sub>2</sub>-EDTA; 0.1 M KPO<sub>4</sub> (pH 7); 0.5% SDS; 1 x Denhardt's solution; and 4 x SSC. In preparing this hybridization mixture, the first six components were mixed in one tenth of the total final volume, and the DNA was denatured by heating at 98-100°C for 5 min, followed by quick cooling in ice water; the last four components, mixed in the remaining volume of H<sub>2</sub>O, were then added at 65°C. A final volume of 3 ml per filter was used, and hybridization was carried out in a 65°C waterbath for 12-18 hr. The nonspecifically bound P<sup>32</sup> probe was then removed by washing at 65°C, first in a solution of 4 x SSC, 0.5% SDS, 1 x Denhardt's (20 ml per filter, 30 min immersion in each of three consecutive batches of solution), and then in a solution of 2 x SSC, 0.5% SDS (30 ml per filter, 60 min immersion in two consecutive batches of solution). During each step of the washing procedure, the filters were transferred individually from



one solution to the next, with excess liquid blotted off, and the solutions were agitated occasionally to increase the efficiency of washing. After drying at room temperature for approximately 1 hr, the filters were mounted on aluminum sheets covered with plastic-backed filter paper, marked with radioactive ink and covered with plastic wrap. Autoradiography was carried out at  $-70^{\circ}\text{C}$  using Kodak X-Omat R film and Fuji Mach 2 tungsten screens, with exposure times of 24–36 hr (Laskey and Mills, 1977).

#### Construction of a $\lambda$ Chain cDNA Clone (B1) and Preparation of Hybridization Probe

A plasmid pCRI clone that carries an enzymatically synthesized DNA complementary to HOPC 2020  $\lambda_1$  chain mRNA (970 bases long) in the Eco RI site was constructed according to the technique of Maniatis et al. (1976) by the poly(dA)·poly(dT) method (Lobban and Kaiser, 1973). The  $\lambda_1$  DNA is 970 bases long as determined by the S1 nuclease method of Hofstetter et al. (1976). For screening of phage plaques and filter hybridization of cloned DNAs, whole plasmid DNA was nick-translated (see below) and used as the hybridization probe. For filter hybridization of cellular DNA, the plasmid B1 DNA was digested with Hha I, and the 2.5 kb fragment containing the full  $\lambda_1$  DNA sequence was isolated by acrylamide gel (5%) electrophoresis. The isolated Hha I fragment was nick-translated in a 50  $\mu\text{l}$  reaction mixture with the following composition: 0.05 M Tris-HCl (pH 7.8), 0.005 M  $\text{MgCl}_2$ , 0.01 M  $\beta$ -mercaptoethanol, 1  $\mu\text{g}$  DNA, 3 units of *E. coli* DNA polymerase I (Boehringer) and 100 pmoles each of  $\alpha$ - $^{32}\text{P}$ -labeled dATP, dCTP, dGTP and TTP (Amersham-Searle; spec. act. 300 Ci/mmmole). No additional DNAase was used. The mixture was incubated at  $14^{\circ}\text{C}$  for 2 hr, extracted with phenol and passed through a small column of Sephadex G-150.

#### Electrophoretic Visualization of $\lambda$ Chain Sequence-Positive DNA Fragments (Southern Gel Blotting Technique)

Highly polymerized cellular or phage DNA was digested with Eco RI under standard conditions (Greene, Betlach and Boyer, 1974). For the cellular DNA, the extent of digestion was monitored by adding phage  $\lambda$  DNA (one tenth of the cellular DNA) in the digestion mixture and visualizing the Eco RI- $\lambda$  DNA bands in a pilot electrophoresis run. Digestion of the cellular DNA was considered to be complete when the DNA mixture was incubated with an amount of Eco RI that is 3 times more than necessary for the complete digestion of the admixed  $\lambda$  DNA. Eco RI-digested DNAs (10  $\mu\text{g}$  of cellular DNA; 0.1  $\mu\text{g}$  of phage DNA; slot cross-section, 8 x 4 mm) were electrophoresed on a 0.9% slab agarose gel in TA buffer (0.04 M Tris-acetate (pH 7.9), 0.005 M Na-acetate, 0.001 M  $\text{Na}_2\text{EDTA}$ ) and transferred overnight onto nitrocellulose membrane filters (Schleicher-Schuell, BA85) according to the procedures described by Southern (1975). The filters were washed for 10 min in 2 x SSC, air-dried and baked for 3 hr at  $80^{\circ}\text{C}$  in a vacuum oven. The baked filters were wetted with 2 x SSC and incubated at  $68^{\circ}\text{C}$  in a precoating mixture consisting of 5 x SSC, 0.1 M  $\text{KPO}_4$  (pH 7.0) (Breathnach et al., 1977) and the Denhardt's mixture. The precoated filters (5 x 9 cm or 10 x 9 cm) were transferred while wet to a plexiglass slit (1.5 mm thick) containing a hybridization mixture (4 or 10 ml) with the following composition: 0.1 M  $\text{KPO}_4$  (pH 7.0), 1 mM  $\text{Na}_2\text{EDTA}$ , 5 x SSC, 1 x Denhardt's solution, 10  $\mu\text{g}/\text{ml}$  of poly(A), 100  $\mu\text{g}/\text{ml}$  of poly(A), 100  $\mu\text{g}/\text{ml}$  of sonicated, heat-denatured salmon sperm DNA (Sigma), 0.5% SDS and 5 x  $10^6$  cpm/ml (for cellular DNA) or 5 x  $10^6$  cpm/ml (for phage DNA) heat-denatured, nick-translated plasmid B1 DNA. Hybridization was at  $68^{\circ}\text{C}$  for 16 hr. The filters were washed at  $68^{\circ}\text{C}$  first in the precoating solution supplemented with 0.5% SDS and 100  $\mu\text{g}/\text{ml}$  of sonicated, heat-denatured salmon sperm DNA (200 ml per filter, 60 min immersion in each of six consecutive batches of solution), then in a solution of 0.1 M  $\text{KPO}_4$  (pH 8.4), 1 x SSC and 0.5% SDS (200 ml per filter, 60 min immersion in each of three consecutive batches of solution followed by one overnight wash), and finally in a solution of 0.1 x SSC and 0.5% SDS (200 ml per filter, one 30 min immersion). Filters were rinsed in 2 x SSC at room temperature, air-dried,

mounted on plastic-backed paper and covered with plastic wrap. Autoradiography was as described above, except that exposure times were 7 days in the case of the cellular DNA.

#### Enzymes

T4 ligase was purchased from Miles. *E. coli* DNA polymerase I was obtained from Boehringer; pancreatic DNAase was from Worthington. Avian myeloblastosis virus (AMV) reverse transcriptase was a gift from J. Beard. Endonucleases Eco RI and Bam HI were gifts from N. Hozumi, and Hae III from J. Summers. Calf thymus terminal transferase was from C. P. Kung. Xba I, Mbo II and Pst I were purchased from New England Bio Labs.

#### Preparation of $V_H$ and $C_H1$ Sequence Probes

##### $V_H$ Sequence Probe

A 470 nucleotide long DNA fragment that consisted of the nearly complete  $V_{H11}$  gene sequence (Hozumi et al., 1978; Tonegawa et al., 1978) was isolated from Ig 13A DNA by sequential digestion of the whole phage DNA with Hae III and Eco RI, followed by electrophoresis in a 5% (w/v) acrylamide gel and by digestion of the isolated 1.5 kb Eco RI-Hae III fragment with Mbo II, followed by electrophoresis in a 6% (w/v) acrylamide gel. The DNA fragments were eluted from the gel as previously described (Maxam and Gilbert, 1977) and nick-translated.

##### $C_H1$ Sequence Probe

$^{32}\text{P}$ -labeled DNA complementary to purified  $\lambda_1$  chain mRNA (from H 2020) was synthesized by the AMV reverse transcriptase. The incubation mixture consisted of the following components in 60  $\mu\text{l}$ : 50 mM Tris-HCl (pH 8.3), 10 mM  $\text{MgCl}_2$ , 40 mM KCl, 10 mM dithiothreitol, 0.1 mM each of dATP, dCTP, dGTP and dTTP, 250  $\mu\text{Ci}$  each of  $^{32}\text{P}$ -dATP and  $^{32}\text{P}$ -dCTP (both from New England Nuclear), 2  $\mu\text{g}$  of  $\lambda_1$  mRNA, 0.6  $\mu\text{g}$  of oligo (dT)<sub>12-18</sub> (P.L. Biochemicals) and 3 units of AMV reverse transcriptase. The incubation was at  $37^{\circ}\text{C}$  for 1 hr. To free the DNA transcript from the RNA and the proteins, the reaction mixture was incubated for 12 hr with 0.3 N NaOH at  $37^{\circ}\text{C}$ , neutralized, extracted with phenol and passed through a 3 ml Sephadex G-150 column in  $\text{H}_2\text{O}$ . The excluded material was precipitated from 2.5 vol of ethanol in the presence of 5  $\mu\text{g}$  of *E. coli* tRNA, dissolved in 90% formamide and electrophoresed in a 5% (w/v) acrylamide gel polymerized in 98% formamide, according to the procedures described previously by Maniatis et al. (1976). After electrophoresis, the gel was wrapped in a thin plastic sheet and exposed for 5 min to Kodak X-Omat R film. The lengths of the transcripts were estimated from  $^{32}\text{P}$ -labeled, Hind III-digested SV40 DNA that was electrophoresed in parallel. A 2 mm gel slice containing 400–430 base long transcripts was excised and the DNA was extracted as described above (Maxam and Gilbert, 1977).

#### R Loops

R loops were prepared essentially as described by Thomas et al. (1976). Hybridization was carried out in a volume of 72  $\mu\text{l}$  of the following solution: 70% (v/v) of formamide (Merck pro-analysis, 3 x crystallized, conductivity 28  $\mu\text{MHO}$ ), 100 mM PIPES, 20 mM Tris, 5 mM EDTA (pH 7.8), 0.56 M NaCl, 12  $\mu\text{g}/\text{ml}$  DNA, 16  $\mu\text{g}/\text{ml}$  mRNA. Samples were incubated at  $56^{\circ}\text{C}$  for 6–14 hr.

#### Heteroduplex

For heteroduplex formation (Westmoreland, Szybalski and Ris, 1969; Davis, Simon and Davidson, 1971), either the purified mouse DNA fragments or total Eco RI digests of phage DNA clones were used. Equimolar amounts (2–5  $\mu\text{g}/\text{ml}$ ) of each fragment were denatured in a total volume of 15  $\mu\text{l}$  of 0.1 M NaOH, 15 mM EDTA at  $30^{\circ}\text{C}$ . Formamide was then added to a final concentration of 50% (v/v) and Tris (pH 8.5) to a final concentration of 100 mM. Renaturation was allowed to occur at room temperature for periods of 2–4 hr.

#### Electron Microscopy

The formamide method (Westmoreland et al., 1969; Davis et al.,



1971) was used for spreading R loops and heteroduplexes. The spreading solution (50  $\mu$ l) was made up of 65–70% (v/v) of formamide (3x crystallized), 100 mM Tris, 10 mM EDTA (pH 8.5), 100  $\mu$ g/ml of cytochrome c (Sigma grade IV, CNBr-treated). DNA was added immediately before spreading to a final concentration of 0.5–1  $\mu$ g/ml. Aliquots of 10  $\mu$ l were spread on a hypophase of 20% formamide, 20 mM Tris, 2 mM EDTA, adjusted to pH 8.5–8.6 with NaOH. Samples were picked up on collodion-coated grids, stained in  $10^{-4}$  M uranyl acetate in 90% ethanol, dehydrated in isopropanol and rotary-shadowed with platinum at an angle of 6°. In the first experiments, DNAs of phage fd (6.3 kb) and phage PM<sub>2</sub> (10 kb) were added as internal length standards. The fragment of lg 13 $\lambda$  was thus determined to be 4.8 kb, and was used in later experiments as a length standard. Photographs were taken with a Philips EM 300 at a magnification of 10,000 $\times$ . Molecules were measured on 10x enlarged negatives with a Numonics digitizer. Data were stored and analyzed with a Hewlett Packard 9825 table computer.

#### Acknowledgments

We thank Drs. W. Arber, J. Carbon, R. Curtis, P. Leder, K. Murray and N. E. Murray for bacteria and phage strains. We are also grateful to Drs. J. Beard, N. Hozumi, C. P. Kung and J. Summers for enzymes, and Dr. G. Matthysens, who participated in the construction of the B1 plasmid. We are grateful to Mr. G. Dastoor-nikoo, Mr. A. Traunecker and Mrs. R. Hiestand for their expert technical assistance.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 24, 1978; revised July 5, 1978

#### References

- Aloni, Y., Dhar, R., Laub, O., Horowitz, M. and Khoury, G. (1977). *Proc. Nat. Acad. Sci. USA* 74, 3686.
- Benton, W. D. and Davis, R. W. (1977). *Science* 196, 180.
- Berget, S. M., Moore, C. and Sharp, P. A. (1977). *Proc. Nat. Acad. Sci. USA* 74, 3171.
- Brack, C. and Tonegawa, S. (1977). *Proc. Nat. Acad. Sci. USA* 74, 5453.
- Breathnach, R., Mandel, J. L. and Chambon, P. (1977). *Nature* 270, 314.
- Chow, L. T., Gelin, R. E., Broker, T. R. and Roberts, R. J. (1977). *Cell* 12, 1.
- Davis, R. W., Simon, M. and Davidson, N. (1971). In *Methods in Enzymology*, 21D, L. Grossman and K. Moldave, eds. (New York: Academic Press), p. 413.
- Denhardt, D. (1966). *Biochem. Biophys. Res. Commun.* 23, 641.
- Dugan, E. S., Bradshaw, R. A., Simms, E. S. and Eisen, H. N. (1973). *Biochemistry* 12, 5400.
- Glover, D. M. and Hogness, D. S. (1977). *Cell* 10, 167.
- Goodman, H. M., Olson, M. V. and Hall, B. D. (1978). *Proc. Nat. Acad. Sci. USA* 75, 5453.
- Greene, P. J., Betlach, M. C. and Boyer, H. W. (1974). In *Methods in Molecular Biology*, 7, R. B. Wickner, ed. (New York: Marcel Dekker), p. 88.
- Hofstetter, H., Schamböck, A., Van den Berg, J. and Weissmann, C. (1976). *Biochim. Biophys. Acta* 454, 587.
- Honjo, T., Packman, S., Swan, D. and Leder, P. (1976). *Biochemistry* 15, 2780.
- Hozumi, N. and Tonegawa, S. (1976). *Proc. Nat. Acad. Sci. USA* 73, 3628.
- Hozumi, N., Brack, C., Pirrotta, V., Lenhard-Schuller, R. and Tonegawa, S. (1978). *Nucl. Acids. Res.* 5, 1779.

- Jeffreys, A. J. and Flavell, R. A. (1977). *Cell* 12, 1097.
- Klessig, D. F. (1977). *Cell* 12, 9.
- Kramer, R. A., Cameron, J. R. and Davis, R. W. (1976). *Cell* 8, 227.
- Laskey, R. A. and Mills, A. D. (1977). *FEBS Letters* 82, 314.
- Leder, P., Tiemeier, D. and Enquist, L. (1977). *Science* 196, 175.
- Lobban, P. E. and Kaiser, A. D. (1973). *J. Mol. Biol.* 78, 453.
- Maniatis, T., Sim, G. K., Efstratiadis, A. and Kafatos, F. C. (1976). *Cell* 8, 163.
- Maxam, A. M. and Gilbert, W. (1977). *Proc. Nat. Acad. Sci. USA* 74, 560.
- Pernis, B., Chiappino, G., Kelus, A. S. and Gell, P. G. H. (1965). *J. Exp. Med.* 122, 853.
- Southern, E. M. (1975). *J. Mol. Biol.* 98, 503.
- Tilghman, S. M., Tiemeier, D. C., Seidman, J. G., Peterlin, B. M., Sullivan, M., Maizel, J. V. and Leder, P. (1978a). *Proc. Nat. Acad. Sci. USA* 75, 725–729.
- Tilghman, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P. and Weissmann, C. (1978b). *Proc. Nat. Acad. Sci. USA* 75, 1309–1313.
- Thomas, M., White, R. L. and Davis, R. W. (1976). *Proc. Nat. Acad. Sci.* 73, 2294.
- Tonegawa, S. (1976). *Proc. Nat. Acad. Sci. USA* 73, 203.
- Tonegawa, S., Hozumi, N., Matthysens, G. and Schuller, R. (1976). *Cold Spring Harbor Symp. Quant. Biol.* 41, 877.
- Tonegawa, S., Brack, C., Hozumi, N. and Pirrotta, V. (1977a). *Cold Spring Harbor Symp. Quant. Biol.* 42, in press.
- Tonegawa, S., Brack, C., Hozumi, N., Matthysens, G. and Schuller, R. (1977b). *Immunol. Rev.* 36, 73.
- Tonegawa, S., Brack, C., Hozumi, N. and Schuller, R. (1977c). *Proc. Nat. Acad. Sci. USA* 74, 3518.
- Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. and Gilbert, M. (1978). *Proc. Nat. Acad. Sci. USA* 75, 1485.
- Weigert, M. G., Cesari, I. M., Yonkovich, S. J. and Cohn, M. (1970). *Nature* 228, 1045.
- Wellauer, P. K. and Dawid, I. B. (1977). *Cell* 10, 193.
- Westmoreland, B. C., Szybalski, W. and Ris, W. (1969). *Science* 163, 1343.
- White, R. L. and Hogness, D. S. (1977). *Cell* 10, 177.
- Yoshida, T. H., Imai, H. T. and Potter, M. (1968). *J. Nat. Cancer Inst.* 41, 1083.