Somatic Recombination 
and Structure of an 
Immunoglobulin Gene

SUSUMU TONEGAWA, CHRISTINE BRACK, 
MINORU HIRAMA, AND 
RITA LENHARD-SCHULLER 
Basel Institute for Immunology 
Basel, Switzerland

INTRODUCTION

Is the organization of DNA sequences altered in cell differentiation? If so, an obvious implication is that such a somatic change might play a key role in cell determination. One of the best systems to which this question can be addressed is that of immunoglobulin genes. Twelve years ago, Dryer and Bennett proposed that a single immunoglobulin chain is encoded in two separate DNA segments, one for the amino terminal half (V region) and the other for the carboxyl terminal half (C region), and that synthesis of a complete chain is preceded by rearrangement of the DNA segments (1). The hypothesis was extended by Gally and Edelman to incorporate a mechanism for somatic generation of the V region diversity (2). Direct investigation into the issue became possible upon discovery of bacterial restriction enzymes. Using a restriction enzyme to digest total cellular DNA of embryo and myeloma cells, Hozumi and Tonegawa produced experimental results that were compatible with the Dryer and Bennett hypothesis (3). Development of the in vitro recombinant technique opened a way for further studies on this problem. This paper deals with isolation and characterization of DNA clones containing mouse λ type light chain genes. The study
not only provided a direct evidence for somatic recombination but also revealed a surprising feature of the structure of an immunoglobulin gene.

SCHEME FOR ISOLATION OF Ig GENES FROM CELLULAR DNA

Various steps involved in the gene cloning experiment are schematically illustrated in Fig. 1. We first digested highly polymerized total cellular DNA with restriction endonuclease EcoRI and fractionated the resulting DNA fragments by electrophoresis in a slab agarose gel. We then identified the DNA fragments carrying specific Ig gene sequences by hybridization. This step was carried out in two different

---

**Fig. 1.** Scheme for isolation of immunoglobulin genes. See text for explanation.
ways. Earlier, we had extracted DNA fragments from the gel slices and carried out hybridization in liquid using °T-labeled, purified Ig mRNA (3). More recently, we synthesized complementary DNA on the purified mRNA template by reverse transcriptase and constructed a plasmid composed of pCRI and Ig DNA sequences. Using chimeric plasmids as the hybridization probes, we applied the blotting technique developed by Southern (4) to a strip of slab gel (5). This procedure obviated the time-consuming extraction step necessary in the liquid hybridization method. Results of the Southern gel blotting of embryo and myeloma DNA’s are illustrated in Fig. 2.

The gel electrophoresis step gave 10- to 100-fold enrichment for various Ig gene-positive fragments. In some cases, we further enriched the relevant DNA fragments by incubating the duplex DNA with excess Ig mRNA in 70% formamide and subjected the mixture to equilibrium centrifugation in a CsCl gradient. Since RNA-DNA hybrids are more stable than DNA duplexes in aqueous formamide under certain conditions, one can construct a structure in which duplex DNA is partially denatured and one of the two strands is replaced by complementary RNA (R-loop structure) (6). Such hybrid molecules are denser than either double- or single-stranded DNA, and therefore can be separated by equilibrium density gradient centrifugation (7). Efficiency of enrichment depends largely on the relative size of the DNA fragments and mRNA molecules. For 4 kb DNA fragments and ~1 kb λ chain mRNA, we obtained enrichment of about fortyfold.

As the cloning vector we used either λgtwes (8) or Charon 4A (9). These phages were attenuated by several genetic tricks to be used specifically in cloning of eukaryotic DNA fragments (EK-2 vectors). When digested with EcoRI, phage DNA’s generate left and right arms as well as one or two DNA fragment(s) that originate from the center section of the genome and are dispensible for phage growth. The center pieces were removed by agarose gel electrophoresis. Vector and mouse DNA were mixed and ligated with T4-ligase at the EcoRI cohesive ends.

We used two different methods in generating phage plaques from recombinant DNA molecules. The first method is to transfect a CaCl₂-treated Escherichia coli with the naked recombinant DNA molecules (10). The second method is to package the DNA molecules in vitro into phage λ coats and to plate viable phage particles on an ordinary E. coli K12 (11).

Identification of the Ig gene sequence-positive plaques was carried out by the replica method developed by Benton and Davis (12). A plaque plate was covered with a dry nitrocellulose membrane filter to
Fig. 2. \(\lambda\) Gene sequence-containing DNA fragments in embryo and myeloma cells. High molecular weight DNA's extracted from 13-day-old BALB/c embryos (B), myelomas HOPC 2020 (a \(\lambda\) chain producer) (A), and MOPC 321 (a \(\kappa\) chain producer) (C) were digested to completion with EcoRI, electrophoresed on a 0.9% agarose gel, transferred to nitrocellulose membrane filters, and hybridized with a nicktranslated \(H\text{ha}I\) fragment of the plasmid B1 DNA. This DNA fragment is 2.5 kb long and contains the full (or near full) \(\lambda\) gene sequence (M. Hiram, G. Matthyssens, and S. Tonegawa, unpublished).

absorb the phage particles and unpackaged phage DNA in the plaques. The filter was briefly dipped in an alkali solution to disrupt the phage particles and also to denature DNA, neutralized in Tris buffer, and baked in a vacuum oven to fix the DNA. The filter was then
incubated under proper conditions with a radiolabeled hybridization probe (as the probe we used $^{125}$I-labeled mRNA or nicktranslated plasmid DNA's containing $\kappa$ or $\lambda$ gene sequences), washed extensively, and subjected to autoradiography. Phages in the plaques that gave positive autoradiographic spots were picked, purified, and propagated in an EK-2 host, E. coli DP50 SupF$^+$ (8).

The efficiency of the in vitro packaging method is at least one, and usually two, orders of magnitude greater than that of transfection (11). We now obtain over $10^5$ plaques by packaging 0.1 $\mu$g of mouse DNA fragments (about 10 kb in length, ligated with 0.3 $\mu$g of Charon 4A or $\lambda$gt$^{wel}$ DNA arms) in a standard 40 $\mu$l packaging mixture prepared from 10 ml each of the two heat-induced lysogen cultures. This scale of packaging is sufficient for the cloning of a unique mammalian gene, if the DNA preparation used is ten- to twentyfold enriched for the particular gene sequence. The simple preparative agarose gel electrophoresis, combined with Southern gel blotting using a nicktranslated cDNA clone probe, easily provides the necessary enrichment. The plaque screening method developed by Benton and Davis is easy, fast, and reliable. With a little care for controlling various factors that affect the plaque size (time of incubation, humidity of the agar plates), one can obtain as many as $10^5$ distinct plaques on a single $20 \times 20$ cm plate, and thereby process as many plaques for in situ hybridization on a single sheet of membrane filter ($19 \times 19$ cm). In summary, the combined use of preparative agarose gel electrophoresis, in vitro packaging, and the Benton and Davis plaque screening technique facilitates cloning of essentially any unique mammalian gene, for which hybridization probes are available, by the handling of a few micrograms of total cellular DNA, less than 100 $\mu$l of the packaging mix, and one or two large agar plate(s). Because the entire experiment is done on a relatively small scale, it is easy to contain and therefore reduce the chances of accidental escape of hypothetically hazardous clones.

DNA CLONES ISOLATED FROM EMBRYO AND MYELOMA CELLS

We have applied the gene cloning procedures described in the last section to the various DNA components from BALB/c embryos (12 days old) and from both $\kappa$-chain secreting (MOPC 321) and $\lambda$-chain secreting (HOPC 2020 and J 558) myelomas. In the present paper we
<table>
<thead>
<tr>
<th>Clones</th>
<th>DNA source</th>
<th>DNA pre-enrichment steps</th>
<th>Approximate pre-enrichment factors</th>
<th>Screening procedures and probes</th>
<th>Approximate number of plaques screened</th>
<th>λ Gene sequences contained</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig 99λ</td>
<td>Embryo</td>
<td>Agarose gel R-looping (one cycle)</td>
<td>300</td>
<td>Benton and Davis (12) with nick-translated cDNA plasmid</td>
<td>3,000</td>
<td>V₉₁</td>
<td>This paper</td>
</tr>
<tr>
<td>Ig 25λ</td>
<td>Embryo</td>
<td>Agarose gel</td>
<td>15</td>
<td></td>
<td>80,000</td>
<td>C₉₁</td>
<td>This paper</td>
</tr>
<tr>
<td>Ig 303λ</td>
<td>HOPC 2020</td>
<td>Agarose gel</td>
<td>15</td>
<td></td>
<td>70,000</td>
<td>V₉₁ + C₉₁</td>
<td>Brack and Tonegawa (16)</td>
</tr>
<tr>
<td>Ig 13λ</td>
<td>Embryo</td>
<td>Agarose gel R-looping (two cycles)</td>
<td>360</td>
<td>Kramer et al. (15) with [¹²⁵I]λ mRNA</td>
<td>4,000</td>
<td>V₉₁</td>
<td>Tonegawa et al. (7)</td>
</tr>
</tbody>
</table>

*See Tonegawa et al. (14) for the definition of the enrichment factor.

*This column lists the λ gene-sequences assigned to the cloned DNA's. See text for more details.
shall restrict ourselves to description and discussion of $\lambda$ gene clones. Four types of $\lambda$ gene clones were isolated. Three clones, Ig 25$\lambda$, Ig 13$\lambda$ and Ig 99$\lambda$, are from the three DNA components visualized by Southern gel blotting of embryo DNA, while the fourth clone is from the DNA component (7.4 kb) that is present only in the myeloma (Fig. 2 and Table I).

**TYPE OF $\lambda$ GENE SEQUENCES CONTAINED IN THE DNA CLONES**

In order to find out what kinds of $\lambda$ gene sequences ($V_\lambda$, $C_{M1}$, or $V_\lambda$ plus $C_{M1}$) are contained in the isolated clones, we carried out gel blotting (Southern) experiments with EcoRI digested, cloned DNA, using three different $\lambda$ gene sequences 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 base pair DNA fragment that was excised from the $V_{M1}$-carrying Ig 13$\lambda$ DNA by restriction endonucleases HaeIII and MboII (13). Since $V_{M1}$ and $V_{MII}$ gene sequences are extensively homologous, this DNA fragment serves as a probe for both $V_{M1}$ and $V_{MII}$ gene sequences. The third probe was an approximately 400 base long cDNA that was synthesized on a purified HOPC 2020 $\lambda_1$ mRNA using the oligo (dT$_{12-18}$) primer and was isolated by acrylamide gel electrophoresis in 98% formamide. Because of the specificity of the priming activity of the oligo(dT), and the size of the cDNA, this probe mostly contains $C_{M1}$ gene sequences.

When digested with EcoRI, the Ig 303$\lambda$, Ig 25$\lambda$, Ig 99$\lambda$, and Ig 13$\lambda$ DNA generated, in addition to the left (21.5 kb) and right (14 kb) arms of the phage DNA, fragments of 7.4, 8.6, 3.5, and 4.8 kb, respectively. 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. The 8.6 kb, Ig 25$\lambda$ fragment hybridized with the $C_{M1}$ probe but not with the $V_\lambda$ probe. Conversely, the 3.5 kb, Ig 99$\lambda$ fragments hybridized with the $V_\lambda$ probe but not with the $C_{M1}$ probe. The 7.4 kb Ig 303$\lambda$ fragment hybridized with both the $V_\lambda$ and the $C_{M1}$ probes, while the 4.8 kb, Ig 13$\lambda$ fragment hybridized with the $V_\lambda$ probe but not with the $C_{M1}$ probe. In addition, our current nucleotide sequencing studies demonstrated that the $V_\lambda$ sequence contained in Ig 99$\lambda$ and Ig 303$\lambda$ are of the $\lambda_1$ type (N. Hozumi, O. Bernard,
and S. Tonegawa, unpublished observations). The assigned λ gene sequences are listed in Table I.

LOCATION OF THE λ CHAIN GENE SEQUENCE IN THE DNA CLONES

The position of λ chain gene sequences in the DNA clones was determined by R-loop mapping (6). The clones Ig 303λ DNA was incubated with HOPC 2020 λ chain mRNA under the conditions for R-loop formation. Upon examination in the electron microscope, more than 50% of the molecules displayed DNA-RNA hybrid regions. However, we did not observe the single R-loop of about 1000-nucleotide length that would have been expected if the mRNA had hybridized with a stretch of DNA corresponding to contiguous V and C gene sequences. Instead, two small R-loops (460 and 380 base long) separated by a double-stranded DNA loop (1.2 kb) were observed (Fig. 3). The interpretation of such a hybrid structure is that one RNA molecule is annealed to two stretches of DNA that are separated by a duplex DNA region. In many cases the hybrid segment or R-loop generated by the longer homology had a short whisker (50–100 nucleotides) at its left end. The whisker is probably the poly(A) at the 3’ end of the mRNA. This suggests that the longer homology is composed of the C gene sequence and that the shorter one is composed of the V gene sequence. The validity of this assumption was confirmed by analysis of the heteroduplexes formed between the Ig 303λ DNA and a Vλ gene-carrying DNA fragment (16).

R-loop molecules formed with the 8.6 kb, Ig 25λ fragment displayed a double loop structure composed of a 410 nucleotide R-loop at 3.9 kb from one and, and a 1.2 kb double-stranded DNA loop (Fig. 3). This structure closely resembles the triple loop of Ig 303λ, except that the Ig 25λ hybrids have a long RNA tail (~260 bases) instead of the second, smaller R-loop. Because the Ig 25λ fragment showed homology only with Cλ, but not with Vλ sequences (see above), we conclude that the 410 base pair R-loop contains the Cλ 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) observed sometimes 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 25λ DNA contains a short homology region that hybridizes to a region near the V-C junction of the mRNA molecule. This second homology, which we call the J sequence, is separated from the Cλ
Fig. 3. R-loop molecules obtained by hybridizing HOPC 2020 λ mRNA with the EcoRI fragments of the DNA clones. (a) Ig 303 λ DNA showed two R-loops corresponding to the V and C genes that are separated by the double-stranded DNA loop of about 1.2 kb. The short tail at the end of one loop is the 3' poly(A). (b) Ig 25 λ 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. (c) Ig 99 λ 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.

sequence by 1200 base pairs. 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 λ fragment formed a single R-loop very similar to the one observed in Ig 13 λ (7). It is 380 nucleotides long, lies in the middle of the DNA fragment, i.e., 1.65–1.66 kb from either end, and carries a ~340 nucleotide RNA tail at one end. Since this DNA fragment contains a V_M sequence (see above), the R-loop should contain the V_M sequence and the RNA tail should correspond to the C_M sequence plus the poly(A).
SEQUENCE HOMOLOGY BETWEEN THE CLONED DNA FRAGMENTS

Analysis of sequence homology between the cloned DNA fragments, both within the λ chain genes and in the adjacent regions, may be helpful 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 analyzed by electron microscopy. We shall describe below the observations made with various combinations of the three λ, DNA clones Ig 99λ, Ig 25λ, and Ig 303λ. A summary of all the results is given in Fig. 4.

Ig 25λ VERSUS Ig 303λ

Ig 25λ hybridized with Ig 303λ DNA to form Y-shaped heteroduplex molecules with two single-stranded and one double-stranded arms (Fig. 5). The lengths of the three arms are: 2.87 kb (long single strand), 1.98 kb (short single strand), and 5.47 kb (double strand). The measurements indicate that the shorter single-stranded arm and one strand of the double-stranded arm correspond to the Ig 303λ DNA, whereas the longer single-stranded arm and the other strand of the double-stranded arm derive from the Ig 25λ DNA. In the entire 5.5 kb homology region we observed no local mismatches that would indicate partial nonhomology between the two strands. R-loop mapping had

![Diagram](Fig. 4. Interpretation of the heteroduplex molecules. The position of V and C gene sequences (white boxes) was deduced from R-loop molecules.)
Fig. 5. Heteroduplex molecules formed by various combinations of the three $\alpha$ sequence-containing cloned DNA fragments. (a) Myeloma DNA Ig 303$\alpha$ versus embryo DNA Ig 99$\alpha$; (b) myeloma DNA Ig 303$\alpha$ versus embryo DNA Ig 25$\alpha$; (c) combination of all three clones gave double heteroduplex structures.
shown that in both DNA fragments the C\textsubscript{\Lambda} gene lies between 3.8 and 4.2 kb from one end. The 5.5 kb homology region therefore contains the C\textsubscript{\Lambda} gene. In addition, most if not all of the 1.2 kb DNA segment separating the V and C sequences in the Ig 303\textsubscript{\Lambda} fragment is highly homologous to the DNA segment of similar length that separates the J and C sequences in the Ig 25\textsubscript{\Lambda} fragment. The measurements indicated that the J sequence lies very near or at the branch point of the heteroduplex.

**Ig 99\textsubscript{\Lambda} versus Ig 303\textsubscript{\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 (Fig. 5). The sum of the lengths of the long single-stranded arm and the double-stranded arm corresponds to the length of the Ig 303\textsubscript{\Lambda} fragment, whereas the short single strand and the second strand of the double-stranded region belong to the Ig 99\textsubscript{\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 DNA's are highly homologous. The V\textsubscript{\lambda} sequence lies between 1.66 and 2.0 kb from one end of both Ig 303\textsubscript{\Lambda} and Ig 99\textsubscript{\Lambda} DNA's. Therefore the V gene is within the 2.0 kb homology region and lies near the fork of the Y-shaped heteroduplex (Fig. 5).

**Ig 303, Ig 25, and Ig 99**

The results described in the last two sections strongly suggest that a large part of the Ig 303 fragment (5.5 kb) is homologous to the Ig 25 DNA whereas the rest of the molecule (2 kb) is homologous to the Ig 99 DNA.

In order to confirm this, the three DNA fragments were mixed, denatured, and annealed. A low proportion of the molecules displayed a double heteroduplex structure or cruciform structure composed of the double-stranded and two single-stranded arms (Fig. 5). Measurements of the arms allowed us to assign each part of the hybrid to the three DNA fragments as illustrated in Fig. 4. The observed structure can be formed only if each of the two ends of the Ig 303 fragment is homologous to only one of the two DNA fragments Ig 99 or Ig 25. The two homology regions meet at a point which corresponds to the J sequence in Ig 25 and lies near the junction between the V sequence and the 1.2 kb intron of Ig 303 DNA. The myeloma DNA fragment, Ig 303, thus
seems to be entirely composed of DNA segments that are homologous to parts of the two embryonic fragments.

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 DNA's demonstrated beyond doubt the occurrence of somatic rearrangements of immunoglobulin gene sequences. The double heteroduplex structure generated by co-annealing the three cloned $\lambda_1$ DNA's is incompatible with the trivial alternative interpretation (3) of the results obtained by restriction enzyme mapping of total cellular DNA. The observed heteroduplex structures can not 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 DNA's are also present at corresponding positions in cellular DNA (5).

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 DNA's (Ig 99$\lambda$ and Ig 25$\lambda$). That the branch points correspond to a single site on the Ig 303$\lambda$ DNA was suggested by 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. We conclude that the three cloned $\lambda_1$ DNA's are related by a single recombination event as illustrated in Fig. 6. 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 intron 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. Further support has come from our more recent nucleotide sequencing studies (O. Bernard, N. Hozumi, and S. Tonegawa, unpublished results). These studies revealed that the Ig 303 V DNA segment codes for the polypeptide chain consecutively for a length corresponding to a complete V-region as defined by amino acid sequence studies, whereas the Ig 99$\lambda$ V DNA segment ceases to
Fig. 6. Arrangement of mouse $\lambda$ gene sequences in embryos and $\lambda$ chain-producing plasma cells. In embryo DNA a full $\kappa$ gene sequence is split into two parts that lie separately on two Eco RI fragments. On one, 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 intron ($I_1$) (N. Hozumi, O. Bernard, and S. Tonegawa, unpublished observation). On the second Eco RI fragment the coding sequence is also split into two parts by a 1250 base long intron ($I_2$). The two parts are for the constant region peptides (C) and about 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 $\lambda$ chain-producing myeloma (HOFC 2020) the $\lambda$ gene sequence is rearranged as a result of one or more recombination(s) that involve 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 directly in contact. The limits of the corresponding sequences in the embryo and the myeloma DNA's 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. It is not intended to imply that the embryo and myeloma V sequences are identical. Additional short introns may be present in the C sequences. Arrows with numbers indicate Eco RI sites.

code at residue 97. This suggests that the nucleotides necessary to code for the extra amino acids at the end of the Ig 303$\lambda$ V region are contributed by the J sequence on the Ig 25$\lambda$ DNA. Our correct sequencing study of the J region confirmed this contention.

GENE IN PIECES

The R-loop mapping demonstrated the presence of a 1.2 kb intron both on the Ig 303$\lambda$ and the Ig 25$\lambda$ DNA. The resolution of cytochrome spreadsings will not allow detection of introns shorter than about 100
nucleotides. Indeed, the 93 base long intron near the region corresponding to the amino terminal of the Ig 13λ DNA was revealed only by nucleotide sequence determination (13). Our recent nucleotide sequencing studies (O. Bernard, N. Hozumi, and S. Tonegawa) established that both Ig 303λ and Ig 99λ DNA contain an intron equivalent to that of Ig 13λ both in length and position. These findings are incorporated into Fig. 6. It should be added that additional short introns 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 λ DNA clones the protein-coding sequences are arranged in discrete pieces. For instance, the somatically rearranged, complete λ gene in the myeloma (Ig 303λ) consists of at least three DNA segments, one coding for the leader, one for the V region, and one for the C region. The introns 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 introns of the cloned DNA were also demonstrated in the uncloned cellular DNA (5).

Recent studies on other genes of eukaryotes (17–22), as well as their viruses (23–26) have revealed several cases of this unexpected gene structure: informational DNA interspersed with introns (silent sequences). Introns are probably transcribed together with the informational DNA and subsequently excised during maturation of pre-mRNA. One recent experiment concerning the mouse globin β chain gene seems to support this hypothesis (27). These observations have led us to propose an additional evolutionary pathway for creation of genes in higher organisms (13). By this pathway, a new gene can be created from two or more separate DNA segments upon 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 be fixed in evolution. As the splicing does not have to be 100% efficient, creation of the new gene need not destroy the old: usually a disadvantageous event.

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 into a transcription unit are then available for splicing with pre-existing sequences. Actual evolutionary use of such DNA sequences depends on the emergence of mutations lead-
ing to new splicing signals at the proper positions in the transcription unit. We hypothesize that many genes in higher organisms have arisen through such evolutionary processes.

SOMATIC REARRANGEMENT AS A MECHANISM FOR GENE CONTROL IN CELL DIFFERENTIATION

Does a higher organism utilize such a gene creation mechanism in the normal process of cell differentiation? We consider that the immunoglobulin genes are the perfect example. Here, the J sequence seems to play a key role 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 seem to be restricted to a small subpopulation of cells. Arrangement of a κ chain sequence in DNA's of several nonlymphatic adult tissues was identical to that of embryo DNA when analyzed by the electrophoresis–hybridization assay (28; also M. Hirama, unpublished observation). Analogous experiments carried out using DNA from a λ chain-producing myeloma and a κ sequence probe and vice versa indicated that there is a mutual exclusion in rearrangement of κ and λ chain DNA sequences (28). The results shown in Fig. 2 gave an additional example, namely that the 7.4 kb EcoRI fragment carrying the rearranged full λ gene sequence is absent in the DNA of κ-producing MOPC 321. It should be added that presence of the 7.4 kb fragment is not peculiar to HOPC 2020 myeloma DNA. DNA’s from other λ1-producing myelomas studied (MOPC 104E and J 558) also contained this fragment. [The \( V_{\lambda 1} \) regions synthesized by these myelomas are identical and differ from HOPC 2020 \( V_{\lambda 1} \) region by two residues (29), but all \( V_{\lambda 1} \) regions are believed to share an identical germ line V gene (30).] Conversely, absence of the 7.4 kb fragment is not peculiar to MOPC 321 DNA. DNA’s from other κ chain-producing myelomas studied (TEPC 124 and MOPC 21) gave no such fragment (M. Hirama, unpublished observations).

On the molecular level one can conceive many variations of a gene control mechanism operated by somatic DNA rearrangement. For in-
stance, 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 previously discussed (16), a gene control mechanism directly dependent on somatic sequence rearrangement in DNA seems to fulfill most easily the "one lymphocyte clone—one light chain" rule of the immune system.

ACKNOWLEDGMENTS

We are very grateful to Dr. B. Hohn for invaluable help in the packaging experiments. We thank Dr. W. Arber, Dr. J. Carbon, Dr. R. Curtis, Dr. P. Leder, and Drs. K. and N. E. Murray for bacteria and phage strains. We also thank Dr. J. Beard, Dr. N. Hozumi, Dr. C. P. Kung, and Dr. J. Summers for enzymes, and Dr. G. Matthysens who participated in the construction of the Bl plasmid. We are grateful to Mr. G. Dostoornikoo, Mr. A. Traunecker, and Mrs. R. Hiestand for their expert technical assistance.

REFERENCES