Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions

(κ-chain mRNA/restriction enzymes/RNA·DNA hybridization)

NOBUMICHI HOZUMI AND SUSUMU TONEGAWA

Basel Institute for Immunology, 487, Grenzacherstrasse, CH-4058 Basel, Switzerland

ABSTRACT   A high-molecular-weight DNA from Balb/c mouse early embryo or from MOPC 321 plasmacytoma (a κ-chain producer) was digested to completion with *Bacillus amyloliquefaciens* strain H restriction enzyme (*Bam*H I). The resulting DNA fragments were fractionated according to size in preparative agarose gel electrophoresis. DNA fragments carrying gene sequences coding for the variable or constant region of κ chains were detected by hybridization with purified, ^{125}I-labeled, whole MOPC 321 κ mRNA and with its 3'-end half. The pattern of hybridization was completely different in the genomes of embryo cells and of the plasmacytoma. The pattern of embryo DNA showed two components, one of which (molecular weight = 6.0 million) hybridized with C-gene sequences and the other (molecular weight = 3.9 million) with V-gene sequences. The pattern of the tumor DNA showed a single component that hybridized with both V-gene and C-gene sequences and that is smaller (molecular weight = 2.4 million) than either of the components in embryo DNA. The results were interpreted to mean that the V_κ and C_κ genes, which are some distance away from each other in the embryo cells, are joined to form a contiguous polynucleotide stretch during differentiation of lymphocytes. Such joining occurs in both of the homologous chromosomes. Relevance of these findings with respect to models for V–C gene joining, activation of a specific V_κ-gene, and allelic exclusion in immunoglobulin gene loci is discussed.

Both light and heavy chains of immunoglobulin molecules consist of two regions: the variable region (V region) and the constant region (C region) (1, 2). Uniqueness (i.e., one copy per haploid genome) of the genetic material coding for C region ("C gene") has been conjectured from normal Mendelian segregation of allotypic markers (3). Nucleic acid hybridization studies have confirmed this notion (4–10). Hybridization studies have also demonstrated that a group of closely related V regions are somatically generated from a few, conceivably even a single, germline gene(s) (V gene) (4–6). They did not, however, give us any reliable estimate of the *total* number of germ line V genes. However, given the enormous diversity of V regions, the existence of multiple germ line V genes seems likely. If this is so, a problem arises: how is the information in the V and C genes integrated to generate a contiguous polypeptide chain? Since V- and C-gene sequences exist in a single mRNA molecule as a contiguous stretch (11), such integration must take place at either the DNA or the RNA level. Integration at the RNA level could result from a "copy-choice" event during transcription or from joining of two RNA molecules after transcription.

We report here experimental evidence for possible joining of V and C sequences at the DNA level.

MATERIALS AND METHODS

Myeloma Tumors. Balb/c mouse plasmacytoma MOPC 321 was kindly provided by Dr. M. O. Potter. Tumors were maintained in mice as described (4).

Abbreviations: V and C genes, genes coding for variable and constant regions, respectively; M_r, molecular weight.

BamH I Restriction Endonuclease. *Bacillus amyloliquefaciens* strain H, originally from Dr. F. Young, was obtained from Dr. T. Bickles at the Biozentrum, Basel. Cells were grown in L-broth. The *Bam*H I endonuclease was prepared according to the method of Wilson and Young (12), except that 10% glycerol was present in all buffers during the phosphocellulose step. Five milligrams of high-molecular-weight embryo or MOPC 321 tumor DNA in a buffer consisting of 6 mM Tris·HCl, pH 7.4, 6 mM MgCl$_2$, and 6 mM 2-mercaptoethanol was incubated with 10^4 units (1 unit is defined as the amount of the enzyme sufficient for digesting 1 μg of phage λ DNA in 60 min at 37° under the above conditions) of the purified enzyme at 37° for 4 hr. In order to investigate completeness of digestion we followed the following scheme. An aliquot of the reaction mixture containing mouse DNA was removed before incubation and mixed with a small amount (ratio of λ DNA to mouse DNA, 1:10) of phage λ DNA. This pilot mixture was incubated in parallel with the main reaction mixture under the conditions described above. After incubation the pilot mixture was electrophoresed in 0.9% agarose. DNA was visualized under ultraviolet light after staining with 1 μg/ml of ethidium bromide (13). Digestion of mouse DNA in the main reaction mixture was considered to be complete when the electrophoresis pattern of the λ DNA in the pilot mixture exhibited no indication of incomplete digestion.

Preparative Agarose Gel Electrophoresis. An apparatus conventionally used for separation of serum proteins was adapted for nucleic acids. Agarose (Sigma, electrophoresis grade) was melted in TA buffer (20 mM Tris-CH$_3$CO$_2$, pH 8.0, 18 mM NaCl, 2 mM Na$_2$EDTA, 20 mM NaCH$_3$CO$_2$) at a final concentration of 0.9% and cast in a Plexiglas tray (40 × 50 × 1 cm) with a central longitudinal partition. Agarose was allowed to solidify at room temperature. With the aid of a scalpel, 5-mm wide slots were made in the agarose at 10 cm from one end. The DNA digested with *Bam*H I was concentrated with 2-butanol by the procedures of Stafford and Bieber (14) and dialyzed against TA buffer. The DNA solution was supplemented with 10 × concentrated TA buffer and bromphenol blue. The mixture was warmed to 47°. Melted agarose was mixed with the warmed DNA solution such that the final mixture contained the original TA buffer, 0.01% dye, and 0.45% agarose. The mixture was transferred to the slot in the agarose sheet and was allowed to solidify at room temperature. In order to avoid overloading, the amount of the DNA was kept at less than 2.1 μg/mm^2 of the cross section of the gel. The gel was covered with plastic wrap and electrophoresis was carried out at 4° for 3 days. For the first 20 hr the electric current was kept at 0.75 mA/cm^2 of gel cross section; it was raised to 3 mA for the remainder of the run.

Extraction of DNA from Agarose Gel. After electrophoresis the gel was cut with a scalpel into 5-mm thick slices. Each slice was transferred to a test tube containing one volume of 5 M

NaClO₄ (15). The gel slices were melted by heating at 60° for 60 min with occasional shaking by hand. In order to remove dissolved agarose as well as $NaClO_4$, DNA was absorbed to a small column of hydroxyapatite (Bio-Gel, DNA grade) and then eluted with 0.4 M potassium phosphate, pH 6.9.

Hybridization. DNA eluted from hydroxyapatite was sonicated in a MSE sonicator to yield fragments 400–500 base pairs long. Purified, sonicated *Escherichia coli* DNA (Sigma type VIII) was added at a final concentration of 100 μg/ml. The mixture was dialyzed against water overnight, and DNA was precipitated with 2 volumes of ethanol. The precipitate was collected by centrifugation, dried under reduced pressure, and dissolved in 25 μl of a mixture consisting of 0.3 M NaCl, 0.03 M sodium citrate, 0.01 M Tris HCl, pH 7.5, 0.001 M Na_2EDTA, and 700–1500 cpm of ^{125}I-labeled mRNA or its 3′-end fragment. The hybridization mixture was covered with a thin layer of mineral oil and heated at 98–100° for 5 min. Annealing was carried out at 70° for 40–48 hr. The processing of the hybridization mixture has been described (4).

Preparation of High-Molecular-Weight DNA. DNA was purified from either 12- to 13-day-old Balb/c mouse embryos (Bomholtgard, Denmark) or MOPC 321 tumors according to procedures described by Gross-Bellard *et al.* (16).

Other Procedures. Methods for purification of κ-chain mRNA of MOPC 321, for isolation of 3′-end half fragment of the mRNA, and for RNA iodination have all been described (4, 6). The κ mRNA preparation used in the present work is the one described in Fig. 1 of ref. 6.

RESULTS

Detection of V- and C-gene sequences in DNA fragments

A straightforward procedure for quantitation of a specific sequence in a particular DNA preparation is to carry out hybridization either in liquid or preferentially on filter paper under conditions where labeled RNA is well in excess over DNA (RNA excess, RNA-driven hybridization) (17). For this procedure it is necessary that the RNA preparation is completely pure. The κ mRNA preparation used in the present experiments, although as pure as the present technology permits, does not meet this condition. From the "fingerprint" analysis of T1 RNase digest and from gel electrophoresis in 99% formamide, the purity of the κ mRNA preparation was estimated to be about 90% (4, 6). For this reason and also because of the difficulty of obtaining a large amount of purified κ mRNA, the RNA saturation procedure is not suitable for quantitation of immunoglobulin genes.

When the purpose is to detect a particular DNA sequence in relative rather than absolute terms, as in the present case, an alternative method is available which circumvents the above problems. In Fig. 1 a varying amount of MOPC 321 DNA was hybridized with ^{125}I-labeled κ mRNA of MOPC 321 or the half containing its 3′-end. The volume of hybridization mixtures, concentration of RNA, and time of incubation were constant. Under these conditions the level of ^{125}I-labeled RNA in the hybrid is a function of the concentration of the complementary DNA sequences. As shown in Fig. 1, the increment in the level of hybridization decreases as the concentration of complementary DNA sequences increases. Thus, when DNA is subjected to a fractionation procedure such as gel electrophoresis, and each fraction is assayed for a specific DNA sequence by this method, the shape of the resulting hybrid peak is flatter than expected from the actual distribution of complementary DNA sequences. (In principle, it is possible to correct each
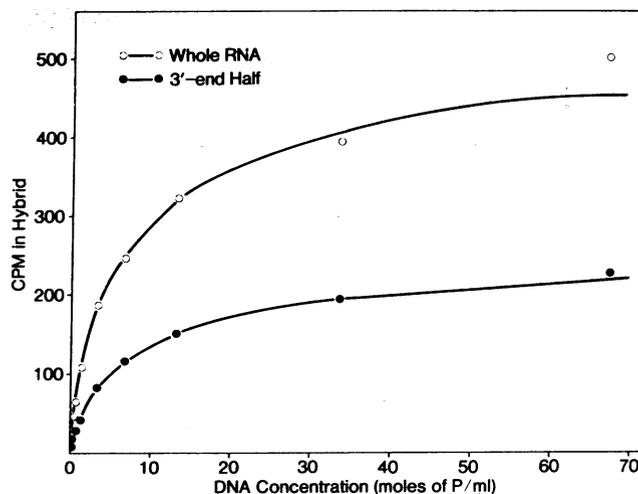

FIG. 1. Dependency of hybridization level on DNA concentration. ^{125}I-Labeled, whole κ mRNA of MOPC 321 (1300 cpm) or its 3′-end half fragment (600 cpm) was annealed with varying indicated amounts of denatured MOPC321 DNA in 0.3 M NaCl, 0.03 M sodium citrate (pH 7) at 70° for 42 hr. The volume of annealing mixture was 0.05 ml. Specific activity of the RNA was 6.5×10^7 cpm/μg. Each point represents an average of duplicate measurement. Intrinsic RNase-resistant counts, 22 cpm for whole RNA and 16 cpm for 3′-end half, are subtracted.

fraction for the nonlinearity of the assay by using Fig. 1 or a "standard curve". We have not done this, however, in Fig. 2.) This disadvantage is more than counterbalanced by the critical advantage; namely, that hybridization of minor RNA impurity does not overshadow that of the major RNA component. The method is particularly suitable where the impurity is distributed among many different RNA species, for in this case, DNA sequences complementary to the contaminating RNAs would be distributed among many different fractions, and hybridization of any particular RNA impurity would not stand out of the general background level. As previously shown, κ mRNA preparations have impurities of exactly this type (4, 6).

Since a RNA probe for V gene sequences is not available (the 3′-end half fragment is a RNA probe for C-gene sequences), V-gene sequences were determined indirectly from the difference in the two hybridization levels obtained by the whole molecule and the half-molecule containing the 3′-end (3′-half). For this purpose two RNA probes with identical specific activity were prepared; hybridization of whole molecules was carried out using twice as much input radioactivity as with 3′-end half-molecules. The final hybridization levels obtained with the whole molecules are just about twice those with the 3′-half in a wide range of DNA concentration (Fig. 1); this justified such an indirect measurement of V-gene sequences.

V- and C-gene sequences in embryo and myeloma DNA fragments generated with BamH I endonuclease

High-molecular-weight embryo and MOPC 321 tumor DNAs were digested with BamH I endonuclease, and the resulting DNA fragments were fractionated according to size by preparative agarose gel electrophoresis. DNA eluted from gel slices was assayed for V- and C-gene sequences by hybridization with ^{125}I-labeled whole or 3′-half κ mRNA of MOPC 321 under the condition described in the last section. The pattern of hybridization is shown in Fig. 2. With embryo DNA, two DNA components of molecular weight (M_r) 6.0 and 3.9 million hybridized with the whole RNA molecules, whereas only the 6.0 million M_r component hybridized with the 3′-half. The fact that only the 6.0 million M_r component exhibits any hybridization
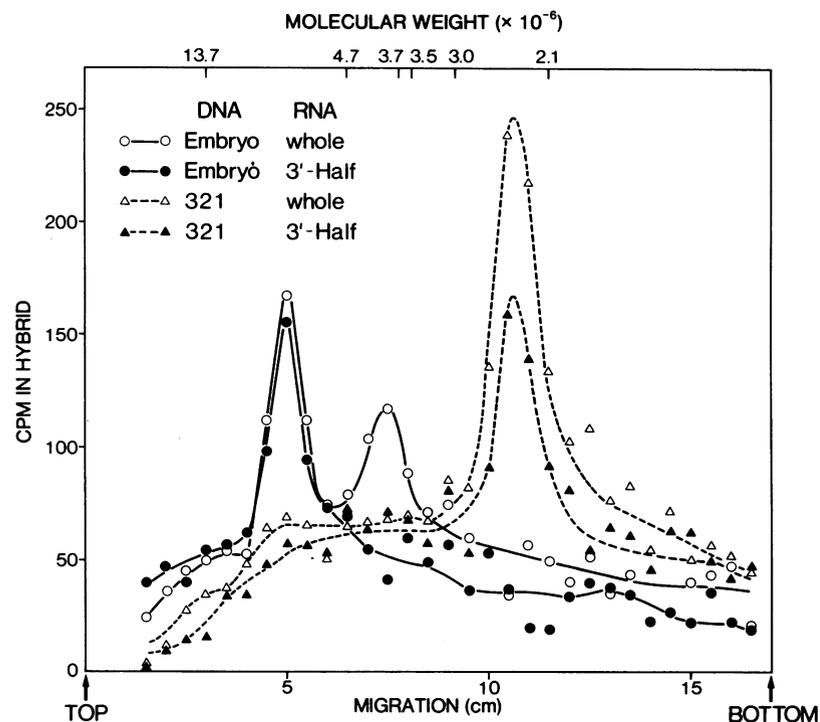
FIG. 2.   Gel electrophoresis patterns of mouse DNA fragments, generated by *Bam*H I, carrying V- or C-gene sequences. ^{125}I-Labeled, whole κ mRNA of MOPC 321 (1220 cpm) or its 3'-end half fragment (600 cpm) was annealed with DNA extracted from gel slices. Intrinsic RNase-resistant counts are subtracted. Conditions of electrophoresis and hybridization are as described in *Materials and Methods*. Hybridization patterns with DNA of two different sources, embryo and MOPC 321 tumor, are superimposed. The molecular weight scale was obtained from phage λ DNA, digested by *Eco*RI (*E. coli*) endonuclease, which was electrophoresed in parallel with mouse DNAs.

above the background level with the 3'-end half indicates that not only the C-region sequences, but also the sequences corresponding to the untranslated region at the 3'-end of mRNA molecule, are in this component. Furthermore, since the 6.0 million M_r component hybridizes almost equally with the two RNA probes, it should not contain appreciable V gene sequences. Thus, the 3.9 million M_r component should contain V-gene sequences as well as sequences corresponding to the untranslated region at the 5'-end of mRNA molecule.

Since V and C genes are in separate DNA fragments whose size is much larger than the size of either gene, they are probably some distance away from each other in the embryo genome. However, the possibility that the enzyme cleaved contiguously arranged V and C genes near the boundary is not entirely eliminated. Cleavage sites are possible in the nucleotides coding for the amino acids at positions 93–95 and 99–100 (18). However, the probability that either of these amino acid sequences provides the exact nucleotide sequence required is low.

The pattern of hybridization is completely different in the DNA from the homologous tumor (Fig. 2). Both RNA probes hybridized with a new DNA component of M_r 2.4 million. There are no indications that either of these RNA probes hybridizes with other DNA components above the general background level. These results indicate that both V and C genes, or the entire sequences represented in the mRNA molecule [except for the unlabeled poly(A)sequence], are contained in the 2.4 million M_r component in the tumor genome. The whole RNA hybridizes with this component nearly twice as well as does the 3'-end half, thereby supporting this notion. Hence, the $V_κ$ and $C_κ$ genes, which are most likely some distance away from each other in the embryo genome, are brought together in the plasma cells expressing this particular $V_κ$ gene, presumably to form a contiguous nucleotide stretch. The fact that

neither the 6.0 nor the 3.9 million M_r component exists in the plasma cell genome indicates that such rearrangement of immunoglobulin genes takes place in all of the homologous chromosomes.

DISCUSSION

We have shown that the pattern of *Bam*H I DNA fragments that carry immunoglobulin V- or C-gene sequences is completely different in the genomes of mouse embryo cells and of a murine plasmacytoma. The pattern of embryo DNA shows two components, one of which (M_r = 6.0 million) hybridizes with C-gene sequences and the other (M_r = 3.9 million) with V-gene sequences. The pattern of tumor DNA shows a single component that hybridizes with both V-gene and C-gene sequences and that is smaller (M_r = 2.4 million) than either of the components in embryo DNA. The straightforward interpretation of these results is that $V_κ$ and $C_κ$ genes, which are some distance away from each other in the embryo cells, are joined to form a contiguous polynucleotide stretch during differentiation of lymphocytes. Such joining occurs in both of the homologous chromosomes.

An alternative explanation of the results, namely, that accumulation of mutations or base modifications leading to either loss or gain of *Bam*H I sites generated the observed pattern difference, is not impossible. On this view, there would have to be a *Bam*H I site close to the V–C junction in embryo DNA. This *Bam*H I site would have to be lost by mutation or by base modification in the MOPC 321 tumor. By itself, such an alteration would cause the appearance of a single 9.9 million M_r component in the tumor. To achieve the M_r of the single component actually observed in the tumor (2.4 million), there would have to be new *Bam*H I sites created by mutation between the $V_κ$ gene and the nearest site on either side. Since there is no reason why there should be any selective pressures in-
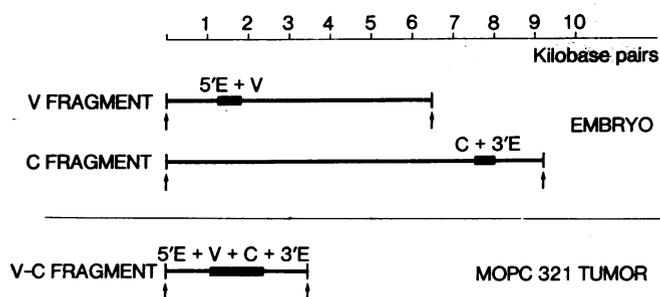
FIG. 3.   Mouse DNA fragments carrying V_κ and C_κ genes. DNA fragments were generated by *Bam*H I restriction endonucleases. Arrows indicate *Bam*H I sites. 5'E and 3'E designate base sequences corresponding to untranslated regions of a κ-chain mRNA molecule at the 5'- and 3'-end, respectively. V and C designate base sequences corresponding to variable and constant regions, respectively. The relative position of these sequences within each fragment is deduced from the present results within the framework of either of the latter three models depicted in Fig. 4.

volving *Bam*H I sites, the occurrence of three base alterations would seem to be quite unlikely. Furthermore, such pattern changes are not unique to this particular combination of enzyme and DNA. With *Eco* RI enzyme and embryo and HOPC 2020 DNA (a λ-chain producer) we have obtained results that lead us to a similar conclusion (Hozumi and Tonegawa, unpublished).

There is also an alternative explanation for the absence of the embryonic DNA components in the tumor, namely, that the V–C gene joining took place in only one of the homologous chromosomes and that the other chromosome(s) has been lost during propagation of the tumor. In view of the known chromosome abnormalities of murine plasmacytomas, we cannot eliminate this trivial possibility.

Fig. 3 summarizes our interpretation of the experimental results. What is the mechanism by which the integration of V- and C-gene sequences is brought about? In the past, several models have been proposed, and some of these models are schematically illustrated in Fig. 4. In the "copy-insertion" model a specific V gene is duplicated and the copy is inserted at a site adjacent to a C gene (19). In this model, the embryonic sequence context of a V gene should be retained in the lymphocyte expressing that particular V gene. Since the embryonic DNA fragment carrying the MOPC 321 V gene does not seem to exist in the genome of this tumor, our results are clearly incompatible with this model. In the "excision-insertion" model, a specific V gene is excised into an episome-like structure, and this in turn is integrated adjacent to a C gene (20). In the "deletion" model, DNA in the interval between a particular V gene and the corresponding C gene loops out, is excised, and diluted out upon subsequent cell multiplication (21). In the "inversion" model, V genes and a C gene are arranged in a chromosome in opposite directions and a segment of chromosome between a particular V gene (inclusive) and a C gene (exclusive) is inverted. The latter three models are both consistent with the experimental results presented here.

It is easy to think of variants of these models. In one variant of the copy-insertion model, the C gene is copied and the copy is inserted next to the V gene. In another variant, both V and C genes are copied, and the copies inserted at a third location. These variants are just as incompatible with our results as is the copy-insertion model illustrated in Fig. 4. There are analogous variants of the "excision-insertion" model and they are just as compatible with the experimental results as the excision-insertion model illustrated in Fig. 4.

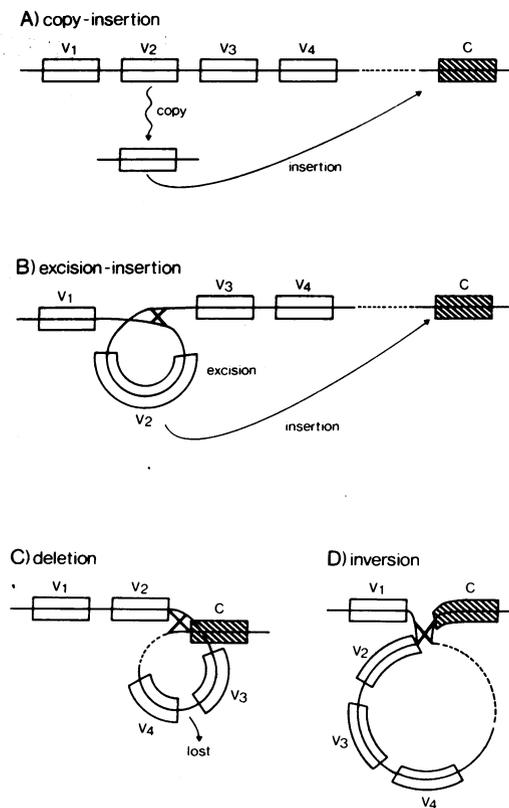A "committed" bone marrow-derived (B) lymphocyte and

FIG. 4.   Models for V–C gene joining at DNA level. See *text* for explanation.

a plasma cell produce antibody of only one specificity (22–24). In particular, it expresses only one light chain V gene. If there are multiple V_κ genes, there must exist a mechanism for the activation of one particular V_κ gene. In the light of the present findings, one intriguing possibility is that activation of a V gene is intimately coupled with its joining to a C gene. For instance, in the "excision-insertion" model, a promotor site may be created by the insertion of the excised V-DNA fragment. This would activate that particular V gene for transcription. In fact, on this view the current terminology of "V genes" and "C genes" is inappropriate. Rather, there are two segments of DNA, one specifying the V region and the other specifying the C region. The gene is *created* by joining.

For immunoglobulin loci, only one allele is expressed in any given lymphocyte (22). This is not the case with most autosomal genes. Our results suggest an interesting explanation for allelic exclusion—that is, that the two homologous chromosomes are, in any given plasma cell, homozygous. Homozygosis could result from the loss of one homologue followed by reduplication of the other or from somatic recombination between the centromere and the immunoglobulin locus (25, 26). The alternative view, that joining takes place in both chromosomes, presents a problem if there is more than one V segment. There is no intrinsic reason why the same V segment should be joined on both homologues.

The present results, when combined with those of our previous reports, demonstrate that both content and context of immunoglobulin genes are altered during differentiation of lymphocytes. Whether such a genetic event is common in other eukaryotic gene systems remains to be seen.

1. Hilschmann, N. & Craig, L. C. (1965) *Proc. Natl. Acad. Sci. USA* **53,** 1403–1409.
2. Titani, K., Whitley, E., Avogardo, L. & Putnam, F. W. (1965) *Science* **149,** 1090–1092.
3. Milstein, C. & Munro, A. J. (1970) *Annu. Rev. Microbiol.* **24,** 335–358.
4. Tonegawa, S., Steinberg, C., Dube, S. & Bernardini, A. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 4027–4031.
5. Tonegawa, S., Bernardini, A., Weimann, B. J. & Steinberg, C. (1974) *FEBS Lett.* **40,** 92–96.
6. Tonegawa, S. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 203–207.
7. Faust, C. H., Diggelmann, H. & Mach, B. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 2491–2495.
8. Rabbitts, T. H. (1974) *FEBS Lett.* **42,** 323–326.
9. Stavnezer, J., Huang, R. C. C., Stavnezer, E. & Bishop, J. M. (1974) *J. Mol. Biol.* **88,** 43–63.
10. Honjo, T., Packman, S., Swan, D., Nau, M. & Leder, P. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 3659–3663.
11. Milstein, C., Brownlee, G. G., Cartwright, E. M., Javis, J. M. & Proudfoot, N. J. (1974) *Nature* **252,** 354–359.
12. Wilson, G. A. & Young, F. E. (1975) *J. Mol. Biol.* **97,** 123–125.
13. Sharp, P. A., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **12,** 3055–3063.
14. Stafford, D. W. & Bieber, D. (1975) *Biochim. Biophys. Acta* **378,** 18–21.
15. Fukes, M. & Thomas, C. A., Jr. (1970) *J. Mol. Biol.* **52,** 395–397.
16. Gross-Bellard, M., Dudet, P. & Chambon, P. (1973) *Eur. J. Biochem.* **36,** 32–38.
17. Birnstiel, M. L., Sells, B. H. & Purdon, I. F. (1972) *J. Mol. Biol.* **63,** 21–39.
18. McKean, D., Potter, M. & Hood, L. (1973) *Biochemistry* **12,** 760–771.
19. Dreyer, W. J., Gray, W. R. & Hood, L. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32,** 353–367.
20. Gally, J. A. & Edelman, G. M. (1970) *Nature* **227,** 341–348.
21. Kabat, D. (1972) *Science* **175,** 134–140.
22. Pernis, B., Chiappino, G., Kelus, A. S. & Gell, P. G. H. (1965) *J. Exp. Med.* **122,** 853–876.
23. Mäkelä, O. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32,** 423–430.
24. Raff, M. C., Feldman, M. & De Petris, S. (1973) *J. Exp. Med.* **137,** 1024–1030.
25. Ohno, S. (1974) in *Chromosomes and Cancer*, ed. German, J. (John Wiley & Sons, New York) pp. 77–94.
26. Gally, J. A. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32,** 166–168.