Cell, Vol. S116, S99–S101, January 23, 2004

That Great Time in Basel

Commentary

Susumu Tonegawa*
The Picower Center for Learning and Memory
E17-353
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Brack et al.'s 1978 paper in *Cell* was sort of a culmination of our two years' work in my laboratory at the Basel Institute for Immunology on the somatic rearrangement of immunoglobulin genes that first appeared in Hozumi and Tonegawa's 1976 paper in PNAS. The experiments described in the PNAS paper were carried out in the pre-Southern blot era and depended on a cumbersome, large-scale agarose gel electrophoresis separation of a multimilligram amount of restriction enzyme-digested genomic DNA followed by extraction of fractionated DNA from gel slices and DNA-RNA hybridization in liquid. Neither the nick-translation nor the cDNA cloning technique had been invented, so the hybridization probe had to be Ig mRNA purified from myeloma cells and labeled by iodination. The patterns of hybridization were clearly different between the DNA from the embryonic (i.e., non-Ig-producing) and myeloma (i.e., Ig-producing) cells. The parsimonious interpretation of the data was that the V and C regions of an Ig chain are encoded separately in the embryonic genome but these DNA sequences somatically rearranged to form a contiguous sequence as lymphocytes developed. However, other, albeit less likely, interpretations of the data, such as fortuitous mutations within restriction cleavage sites in the myeloma, could not be excluded. Furthermore, the precise details of the DNA rearrangement could not be determined from the data.

It was clear that in order to acquire further information, the research would have to resort to the emerging new technologies such as cloning of genomic DNA fragments and determining their nucleotide sequences. Today, these tasks are rather trivial, probably requiring no more than one or two weeks of work. That was not the case in the 1970s. First, gene cloning was so novel that concerns were raised about its potential ill effects to health and the environment. As much debate was ongoing in the U.S., Werner Arber and Walter Gehring at the Biozentrum of the University of Basel took the initiative in forming a small local group of scientists from both the University and the Institute who were interested in but concerned with this technology. The consensus of the group was to watch the development in the U.S. and adopt the guidelines being formed at the NIH. Second, none of the DNA manipulation techniques or reagents that are now routinely available for genomic DNA cloning (such as oligonucleotide synthesis, PCR, or genomic DNA libraries) were available, and the cDNA cloning technique (Maniatis et al., 1976) was barely invented. Other than ribosomal RNA genes, which were known to be reiterated by several hundred fold, no so-called unique eukaryotic gene had been cloned.

What we had in our hands were genomic embryonic and myeloma DNA that had been digested with the restriction enzyme *E. coli* R1 (which had to be prepared in my lab from bacteria harvested in a Roche fermenter) and a 90% pure Ig λ chain mRNA preparation from a myeloma line. Following the procedure described in the 1976 PNAS paper, we first fractionated total Eco-R1 digest of emybronic or myeloma DNA by large-scale preparative gel electrophoresis. The DNA extracted from each gel slice was hybridized in liquid with an intact λ light chain mRNA and separately with its 3′-half fragment (i.e., the C gene sequence) to identify the fractions that contained the V_λ gene segment. The V_λ gene-carrying DNA fragments were further enriched by CsCl density gradient centrifugation of R loops formed between double-stranded genomic V_λ DNA fragments and λ mRNA. This Eco R1-digested genomic DNA preparation in which the V_λ gene fragment was enriched by about 350-fold was the source of a partial genomic DNA library that was prepared following the newly established NIH guidelines using a genetically crippled phage λ vector (λgt WES) and *E. coli* host (803) in the P3 facility created at the Institute. Screening of about 4000 plaques with the I^{125}-labeled, 90% pure Ig λ mRNA preparation as the probe led to an identification of a single DNA fragment of 4.8 kb that hybridized with the complete λ mRNA but not its 3′ half and, hence, was deduced to be a V_λ gene clone (Ig 13λ).

With the techniques for visualizing DNA heteroduplexes and R loops emerging, I had anticipated that nucleic acid electron microscopy would be a fast and powerful means for determining the sequence relationship between cloned genomic DNA fragments derived from the pre- and postrearrangement genome as well as for mapping the coding sequences on these DNA fragments. I had heard of the great reputation of a nucleic acid electron microscopist, Christine Brack, in Eduard Kellenberger's lab at the Biozentrum, so I went to see her and asked whether she would join our small team. To my great joy, she accepted the offer and moved across the Rhine to open her shop in the Institute's EM facility in the winter of 1977. This was great timing, because the Ig 13λ had just been cloned and was ready to be analyzed. Within just a few weeks after her arrival at the Institute, Christine came to me with a beaming smile on her face and asked me to follow her to the EM facility. There they were, beautiful R loops with a long tail protruding from one side! Seeing was indeed believing. I was really convinced at that moment that an Ig V gene was indeed separated from the Ig C gene in the embryonic genome and, therefore, had to be somatically rearranged during B cell development as indicated by the Hozumi and Tonegawa experiment (Tonegawa et al., 1977a).

This marked the beginning of a series of truly exciting discoveries that came out of our collaboration during the subsequent two years. It was indeed an amazing period and there is no exaggeration in saying that we

*Correspondence: tonegawa@mit.edu

had a major discovery every few months along the way. For instance, within a few months after the first visualization of the V_λ R loop, my technical assistant Rita Schaller and I succeeded in cloning another Ig λ clone (Ig 303λ), this time from λ chain-positive myeloma cells, which Christine immediately subjected to another set of R loop analysis. The pictures were even more beautiful than those of Ig 13λ; we saw a triple loop structure in which the three loops converged on an almost single point. Christine's experienced eyes immediately concluded that the two smaller loops were composed of R loops while the third longer loop was made up of double-stranded DNA. Some of these triple loop structures also had a tail (poly(A)) protruding from the larger of the two smaller R loops but it was much shorter than the tail (C_λ sequence + poly(A)) attached to the Ig 13λ R loop (Brack and Tonegawa, 1977). These EM pictures were so powerful in directly, instantaneously, and simultaneously revealing multiple new facts regarding the features of not only Ig genes but also eukaryotic genes in general. First, in contrast to the Ig λ-negative embryonic cells, in Ig λ-positive myeloma cells the V and C regions are encoded on a single EcoR1 DNA fragment (7.4 kb) that is distinct from the embryonic Eco-R1 DNA fragment (4.8 kb) containing only the V_λ gene segment. Second, to our astonishment, the V and C region-coding DNA sequences in B lymphocytes (myeloma) were *not* contiguous but were separated by a 1.2 kb DNA sequence that did not appear in the λ chain mRNA! This second observation constituted the first demonstration of an eukaryotic (i.e., nonviral) intron, the viral counterpart of which had been discovered just a few months prior to this (Berget et al., 1977; Chow et al., 1977).

In parallel with these cloning and EM studies, we initiated the nucleotide sequencing study of Ig 13λ with help first from Vincenzo Pirrota of Biozentrum (Hozumi et al., 1978) and subsequently from Allan Maxam and Walter Gilbert at Harvard University (Tonegawa et al., 1978). This was in the spring of 1977 and neither Maxam and Gilbert's nor Sanger's method (Maxam and Gilbert, 1977; Sanger et al., 1977) had been published yet and all operations were carried out manually. These collaborations also generated some surprises. First, the V gene sequence was truncated by about 13 triplets at the 3′ side. Second, there occurred the second intron at the amino-terminal side that separated the signal peptide from the mature amino terminal. Furthermore, the deciphered nucleotide sequence indentified this V gene as that for the $V_{\lambda II}$ type rather than the more common $V_{\lambda I}$ type (Dugan et al., 1973).

During the autumn of 1977, Rita and I continued our cloning effort. By this time, a postdoctoral fellow, Minoru Hirama, had succeeded in producing the λ cDNA clone and in setting up the Southern blotting method (Southern, 1975) in the lab. As shown in Figure 1 of Brack et al.'s paper, the λ probe gave three hybridizing bands in embryonic DNA, while the λ chain-positive myeloma DNA gave, in addition to the three bands, the fourth band of 7.4 kb which corresponded to the V- and C-positive genomic DNA fragment cloned in Ig 303λ. One of the remaining three bands (4.8 kb) corresponded to the V fragment cloned in Ig 13λ. Thus, our task was to clone the remaining 8.5 kb and 3.5 kb fragments. Our earlier preparative gel electrophoresis experiment suggested that the longer of the two fragments harbors the C gene (Tonegawa et al., 1977b), while the shorter included the V gene, presumably $V_{\lambda I}$ gene because the $V_{\lambda II}$ gene had already been found in the 4.8 kb fragment of Ig 13λ (Tonegawa et al., 1978) (the probe was λ_I cDNA, which would not cross hybridize with the putative $C_{\lambda II}$ gene). By the winter of 1978, all of the genomic DNA fragments that appeared in Figure 1 were cloned using a method more akin to the one used today, namely filter-based screening of a genomic DNA library with a nick-translated cDNA probe. These materials constituted the basis for Brack et al.'s paper in which the relationship among the entire set of Ig λ_I chain genomic DNA fragments from both embryonic and myeloma DNA was determined by electron microscopy. As the title of the paper states, it established beyond doubt that a complete immunoglobuline gene (in this case the mouse λ_I light chain gene) is created by somatic recombination between V- and C-encoding DNA segments that lie separately and in distance in the embryonic genome. Furthermore, the R loop analysis of the embryonic C gene-containing DNA fragment identified the short genomic DNA segment that encoded the missing V region piece (~13 amino acids). This DNA segment was discovered as an independent entity at about 1.2 kb upstream of the C gene segment and was named a J segment for joining because the somatic joining seemed to occur directly at the 5′ end of this segment and the 3′ end of the embryonic V gene segment, a conclusion confirmed quickly by nucleotide sequencing (Bernard et al., 1978).

The discovery of the J segment and the demonstration that somatic recombination occurred *within* and not at the end of a classically defined V region were a complete surprise and had a profound implication in the role and mechanism of the somatic origin of antibody diversity, that is, diversification of V region-coding sequences and, hence, antibody specificity could occur somatically by a random (or quasirandom) selection of a pair of V and J segments in a given B cell lineage from a pool of multiple and distinct V segments and multiple and distinct J segments. The occurrence and importance of this combinatorial diversification mechanisms were quickly confirmed by the discoveries of multiple V and J segments in the Ig κ light chain locus and their modular combinatorial uses in the formation of complete κ genes (Lenhard-Schuller et al., 1978; Max et al., 1979; Sakano et al., 1979; Weigert et al., 1978). This somatic diversification mechanism was subsequently shown to be even more pronounced in the Ig heavy chain locus because of the involvement of the D gene segments, the third type of V region-encoding DNA segments (Early et al., 1980; Kurasawa and Tonegawa, 1982; Kurasawa et al., 1981; Sakano et al., 1981). Another important message was harbored in Brack et al.'s paper regarding somatic diversification mechanisms for antibody diversity. The frequency of the $V_{\lambda I}$ clones was about one in one million among the Eco-R1 digest. This frequency was very similar to the frequency with which $C_{\lambda I}$ clones were identified and indicated that the $V_{\lambda I}$ gene segment is unique in the embryonic genome. On the other hand, *at least* eight different V amino acid sequences had already been known among the collection of myeloma λ_I chains (Weigert et al., 1970). Thus, the Brack et al. paper provided evidence for the earlier predictions (Cohn et al., 1974;

Tonegawa, 1976; Tonegawa and Steinberg, 1976) that somatic hypermutation is another mechanism for the generation of antibody diversity.

References

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. USA *74*, 3171–3175.

Bernard, O., Hozumi, N., and Tonegawa, S. (1978). Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell *15*, 1133–1144.

Brack, C., and Tonegawa, S. (1977). Variable and constant parts of the immunoglobulin light chain gene of a mouse myeloma cell are 1250 nontranslated bases apart. Proc. Natl. Acad. Sci. USA *74*, 5652–5656.

Brack, C., Hirama, M., Lenhard-Schuller, R., and Tonegawa, S.A. (1978). A complete immunoglobulin gene is created by somatic recombination. Cell *15*, 1–14.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. Cell *12*, 1–8.

Cohn, M. Blomberg, B., Geckeler, W., Raschke, W., Riblet, R., and Weigert, M. (1974). First order considerations in analyzing the generator of diversity. In The Immune System, Genes, Receptors, Signals, E.E. Sercarz and A.R.F. Williamson, eds. (New York: Academic Press), pp. 89–118.

Dugan, E.S., Bradshaw, R.A., Simms, E.S., and Eisen, H.N. (1973). Amino acid sequence of the light chain of a mouse myeloma protein (MOPC-315). Biochemistry *12*, 5400–5416.

Early, P., Huang, H., Davis, M., Calame, K., and Hood, L. (1980). An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: V_H, D and J_H. Cell *19*, 981–992.

Hozumi, N., and Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proc. Natl. Acad. Sci. USA *73*, 3628–3632.

Hozumi, N., Brack, C., Pirrotta, V., Lenhard-Schuller, R., and Tonegawa, S. (1978). Characaterization of a mouse DNA lone containing an immunoglobulin variable region gene. Nucleic Acids Res. *5*, 1779–1799.

Kurasawa, Y., and Tonegawa, S. (1982). Organization, structure, and assembly of immunoglobulin hevy chain diversity DNA segments. J. Exp. Med. *155*, 201–218.

Kurasawa, Y., von Boehmer, H., Haas, W., Sakano, H., Traunecker, A., and Tonegawa, S. (1981). Identification of D segments of immunolglobulin heavy-chain genes and their rearrangement in T lymphocytes. Nature *290*, 565–570.

Lenhard-Schuller, R., Hohn, B., Brack, C., Hirama, M., and Tonegawa, S. (1978). DNA clones containing mouse immunoglobulin kappa chain genes isolated by in vitro packaging into phage lambda coats. Proc. Natl. Acad. Sci. USA *75*, 4709–4713.

Maniatis, T., Kee, S., Efstratiadis, A., and Kafatos, F. (1976). Amplification and characterization of a beta-globin gene synthesized in vitro. Cell *8*, 163–182.

Max, E.E., Seidman, J.G., and Leder, P. (1979). Sequences of five potential recombination sites encoded close to an immunoglobulin kappa constant region gene. Proc. Natl. Acad. Sci. USA *76*, 3450–3454.

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA *74*, 560–564.

Sakano, H., Hiuppi, K., Heinrich, G., and Tonegawa, S. (1979). Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature *280*, 288–294.

Sakano, H., Kurasawa, Y., Weigert, M., and Tonegawa, S. (1981). Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. Nature *290*, 562–565.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503–517.

Tonegawa, S. (1976). Reiteration frequency of immunoglobulin light chain genes: further evidence for somatic generation of antibody diversity. Proc. Natl. Acad. Sci. USA *73*, 203–207.

Tonegawa, S., and Steinberg, C. (1976). Too many chains—too few genes. In The Generation of Antibody Diversity: A New Look, A. Cunningham, ed. (New York: Academic Press), pp. 175–182.

Tonegawa, S., Brack, C., Hozumi, N., and Schuller, R. (1977a). Cloning of an immunoglobulin variable region gene from mouse embryo. Proc. Natl. Acad. Sci. USA *74*, 3518–3522.

Tonegawa, S., Brack, C., Hozumi, N., Matthyssens, G., and Schuller, R. (1977b). Dynamics of immunoglobulin genes. Immunol. Rev. *36*, 73–94.

Tonegawa, S., Maxam, A.M., Tizard, R., Bernard, O., and Gilbert, W. (1978). Sequence of a mouse germline gene for a variable region of an immunoglobulin light chain. Proc. Natl. Acad. Sci. USA *75*, 1485–1489.

Weigert, M., Cesari, I.M., Yondovich, S.J., and Cohn, M. (1970). Variability in the lambda light chain sequences of mouse antibody. Nature *128*, 1045–1047.

Weigert, M., Gatmaitan, L., Loh, E., Schilling, J., and Hood, L. (1978). Rearrangement of genetic information may produce immunoglobulin diversity. Nature *276*, 785–790.