

# TWO GENES BECOME ONE: SOMATIC REARRANGEMENT OF IMMUNOGLOBULIN GENES

N. Hozumi and S. Tonegawa, 1976, *Proc. Nat'l. Acad. Sci. USA* 73:3629

For decades, immunologists wondered how the body could generate the multitude of pathogen-fighting immunoglobulins, called antibodies, needed to ward off the vast array of different bacteria and viruses encountered in a lifetime. Clearly, these protective proteins, like all proteins, somehow were encoded in the genome. But the enormous number of different antibodies potentially produced by the immune system made it unlikely that individual immunoglobulin (Ig) genes encoded all the possible antibodies an individual might need. In studies beginning in the early 1970s, Susumu Tonegawa, a molecular biologist, laid the foundation for solving the mystery of how antibody diversity is generated.

## Background

Research on the structure of Ig molecules provided some clues about the generation of antibody diversity. First, it was shown that an Ig molecule is composed of four polypeptide chains: two identical heavy (H) chains and two identical light (L) chains. Some researchers proposed that antibody diversity resulted from different combinations of heavy and light chains. Although somewhat reducing the number of genes needed, this hypothesis still required that a large portion of the genome be devoted to Ig genes. Protein chemists then sequenced several Ig light and heavy chains. They found that the C-terminal regions of different light chains were very similar and thus were termed the constant (C) region, whereas the N-terminal regions were highly variable and thus were termed the variable (V) region. The sequences of different heavy chains exhibited a similar pattern. These findings suggested that the genome contains

a small number of C genes and a much larger group of V genes.

In 1965, W. Dryer and J. Bennett proposed that two separate genes, one V gene and one C gene, encode each heavy chain and each light chain. Although this proposal seemed logical, it violated the well-documented principle that each gene encodes a single polypeptide. To avoid this objection, Dryer and Bennett suggested that a V and C gene somehow were rearranged in the genome to form a single gene, which then was transcribed and translated into a single polypeptide, either a heavy or light chain. Indirect support for this model came from DNA hybridization studies showing that only a small number of genes encoded Ig constant regions. However, until more powerful techniques for analyzing genes came on the scene, a definitive test of the novel two-gene model was not possible.

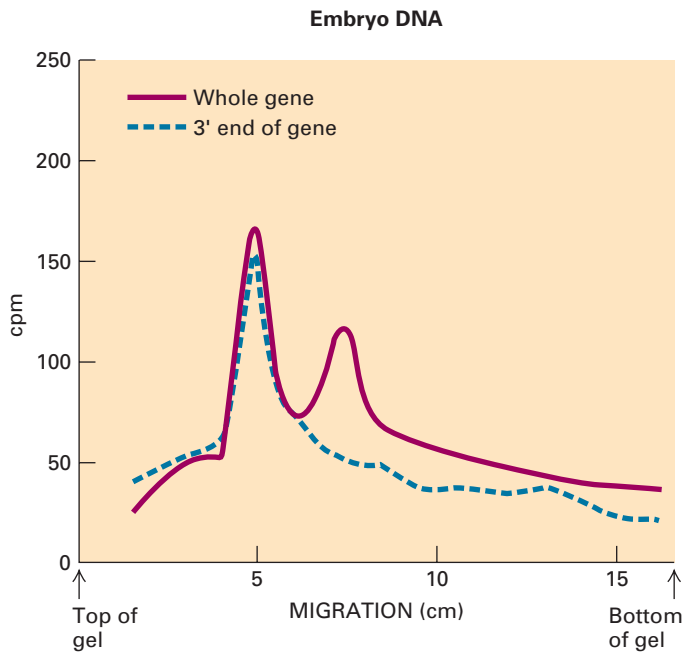
## The Experiment

Tonegawa realized that if immunoglobulin genes underwent rearrangement, then the V and C genes were most likely located at different points in the genome. The discovery of restriction endonucleases, enzymes that cleave DNA at specific sites, had allowed some bacterial genes to be mapped. However, because mammalian genomes are much more complex, similar mapping of the genes encoding V and C regions was not technically feasible. Instead, drawing on newly developed molecular biology techniques, Tonegawa devised another approach for determining whether the V and C regions were encoded by two separate genes. He reasoned that if rearrangement of the V and C genes occurs, it must happen during differentiation of Ig-secreting B cells from embryonic cells. Furthermore, if

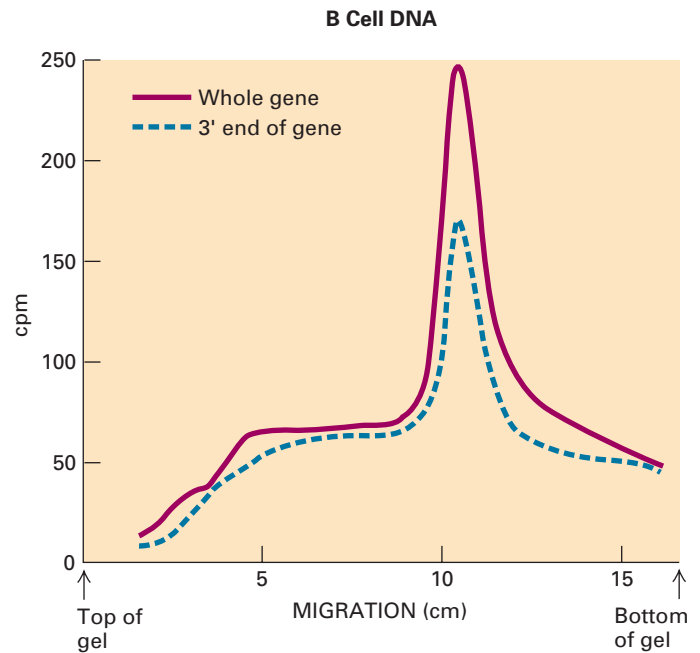
rearrangement occurs, there should be detectable differences between un-rearranged germ-line DNA from embryonic cells and the DNA from Ig-secreting B cells. Thus, he set out to see if such differences existed using a combination of restriction-enzyme digestion and RNA–DNA hybridization to detect the DNA fragments.

He began by isolating genomic DNA from mouse embryos and from mouse B cells. To simplify the analysis, he used a line of B-cell tumor cells, all of which produce the same type of antibody. The genomic DNA was then digested with the restriction enzyme *Bam*HI, which recognizes a sequence that occurs relatively rarely in mammalian genomes. Thus, the DNA was broken into many large fragments. He then separated these DNA fragments by agarose gel electrophoresis, which separates biomolecules on the basis of charge and size. Since all DNA carries an overall negative charge, the fragments were separated based on their size. Next, he cut the gel into small slices and isolated the DNA from each slice. Now, Tonegawa had many fractions of DNA pieces of various sizes. He then could analyze these DNA fractions to determine if the V and C genes resided on the same fragment in both B cells and embryonic cells.

To perform this analysis, Tonegawa first isolated from the B-cell tumor cells the mRNA encoding the major type of Ig light chain, called  $\kappa$ . Since a RNA is complementary to one strand of the DNA from which it is transcribed, it can hybridize with this strand forming a RNA–DNA hybrid. By radioactively labeling the entire  $\kappa$  mRNA, Tonegawa produced a probe for detecting which of the separated DNA fragments contained the  $\kappa$  chain gene. He then isolated the 3' end of the



▲ **FIGURE 1** Experimental results showing that the genes encoding the variable (V) and constant (C) regions of  $\kappa$  light chains are rearranged during development of B cells. These curves depict the hybridization of labeled RNA probes, specific for the



entire  $\kappa$  gene (V + C) and for the 3' end that encodes the C region, to fractions of digested embryonic or B-cell DNA separated by agarose gel electrophoresis. [Adapted from N. Hozumi and S. Tonegawa, 1976, *Proc. Nat'l. Acad. Sci. USA* 73:3629.]

$\kappa$  mRNA and labeled it, yielding a second probe that would detect only the DNA sequences encoding the constant region of the  $\kappa$  chain. With these probes in hand—one specific for the combined V + C gene and one specific for C alone—Tonegawa was ready to compare the DNA fragments obtained from B cells and embryonic cells.

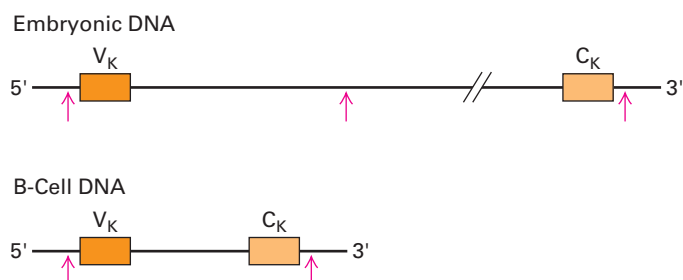
He first denatured the DNA in each of the fractions into single strands and then added one or the other labeled probe. He found that the C-specific probe hybridized to different fractions derived from embryonic and B-cell DNA (Figure 1). Even more telling, the full-length RNA probe hybridized to two *different* fractions of the embry-

onic DNA, suggesting that the V and C genes are not connected and that a cleavage site for *Bam*HI lies between them. Tonegawa concluded that during the formation of B cells, separate genes encoding the V and C regions are rearranged into a single DNA sequence encoding the entire  $\kappa$  light chain (Figure 2).

### Discussion

The generation of antibody diversity was a problem awaiting development of powerful molecular techniques to answer it. Tonegawa went on to clone V-region genes and prove that the rearrangement must occur somatically.

These findings impacted genetics as well as immunology. Where once it was believed that every cell in the body contained the same genetic information, it became clear that some cells take that information and alter it to suit other purposes. In addition to somatic rearrangement, Ig genes undergo a variety of other alterations that allow the immune system to create the diverse repertoire of antibodies necessary to react to any invading organism. Our current understanding of these mechanisms rests on the foundation of Tonegawa's fundamental discovery. For this work, he received the Nobel Prize for Physiology and Medicine in 1987.



◀ **FIGURE 2** Schematic diagram of  $\kappa$  light-chain DNA in embryonic cells and B cells that is consistent with Tonegawa's results. In embryonic cells, cleavage with the *Bam*HI restriction enzymes (red arrows) produces two different sized fragments, one containing the V gene and one containing the C gene. In B cells, the DNA is rearranged so that the V and C genes are adjacent, with no intervening cleavage site. *Bam*HI digestion thus yields one fragment that contains both the V and C genes.