Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes

Hitoshi Sakano, Yoshikazu Kurosawa, Martin Weigert & Susumu Tonegawa

Basel Institute for Immunology, 487 Grenzacherstrasse, Postfach, 4005 Basel 5, Switzerland

A putative diversity segment of immunoglobulin heavy-chain genes (D segment) has been identified 700 base pairs 5′ to J_{H1} DNA on the germ-line genome of the mouse. This 10-base pair D segment is flanked by two sets of sequences related to

CACTGTG *and* GGTTTTTGT
GTGACAC CCAAAAACA

which are possible recognition sites for a recombinase. The spacer separating the heptamer and the nonamer is 12 base pairs long on both sides of the D segment. As the spacer separating the two signal sequences in V_H DNAs and J_H DNAs is 23 ± 1 base pairs long, the two recombinations required for creation of a complete immunoglobulin V_H gene, a V_H–D joining and a D–J_H joining, follow a 12/23-base pair spacer rule. Allelic exclusion is discussed with respect to D segments.

IT is well established that complete immunoglobulin variable region genes (V genes) are generated by somatic DNA recombination(s) during the development of B-lymphocyte cells[1–15]. In the light-chain genes of the mouse, in both λ- and κ-types, the last 13 COOH-terminal residues of the variable region are encoded by a separate DNA segment called J DNA (J for joining) in the germ-line genome[4–7,9,10]. Coding information for the rest of the V region is contained in an embryonic V DNA segment[3,6,7,9,10]. Accordingly, somatic recombination occurs between the 3′ end of the V DNA and the 5′ end of the J DNA to generate a complete light-chain V gene[4,6,8]. The $J\lambda_1$ (ref. 6) and five Jκ DNAs[9,10] carry two blocks of sequences, one a palindromic heptamer, CACTGTG, and the other a nonamer, GGTTTTTGT, highly conserved in the 5′-noncoding regions. These conserved sequences are invertedly repeated in the 3′-noncoding regions of all embryonic Vλ and Vκ DNAs sequenced to date[3,6,7,9,10]. Another striking feature around the recombination site is that the spacer between the heptamer and the nonamer[12,13] in all V and J DNAs is 12 ± 1 or 23 ± 1 base pairs long. In addition, recombination seems to occur only between a V DNA with the shorter (12-base pair) spacer and a J DNA with the longer (23-base pair) spacer or vice versa (12/23-base pair spacer rule). Based on these observations

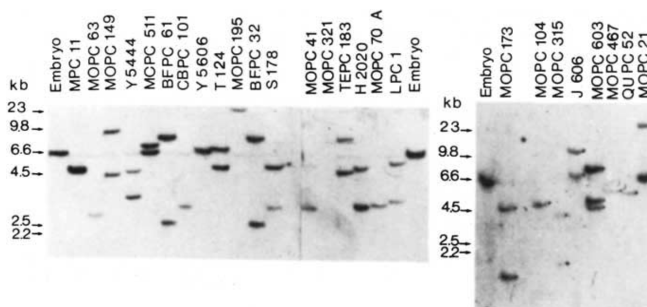

Fig. 1 Southern blot analysis of mouse embryo and myeloma DNAs. Mouse DNAs were digested with *Eco*RI, separated in 0.8% agarose gels and blotted to nitrocellulose filters according to the method of Southern[20]. Filters were then incubated with the nick-translated J_H probe (1.1-kilobase *Sac*I–*Eco*RI fragment, see ref. 11), essentially as described by Wahl *et al.*[39]. *Hin*dIII fragments of phage λ DNA were used as size markers (in kilobases, kb). For the cloning of a 5.2-kilobase Q52 fragment hybridizable with the J_H probe, *Eco*RI digest of the myeloma DNA was fractionated in a 0.8% agarose gel and DNA from the J_H probe-positive fraction was eluted as described previously[40]. Recombinant phages containing the mouse DNA were screened by *in situ* plaque hybridization[41] with the J_H probe.

Early *et al.*[12] and Sakano *et al.*[13] predicted that the same or a similar enzyme mediates the recombinations in both λ- and κ-light-chain genes and that the recombinase consists of two functionally distinct units, one recognizing the heptamer and the nonamer separated by the shorter (12-base pair) spacer and the other recognizing the same signal sequences separated by the longer (23-base pair) spacer.

In contrast to the light-chain genes, heavy-chain J DNAs (J_H) do not seem to associate directly with the heavy-chain V DNAs (V_H). In the case of the γ2*b* heavy-chain gene expressed in myeloma MOPC141, the 14-residue peptide comprising the third hypervariable region (HV3) beginning with Val (residue 98) and ending with Thr (residue 111) is encoded neither in the germ-line V DNA nor in the J DNA[13]. A similar observation has been made independently in an α heavy-chain gene[12]. Therefore, it is generally believed that the HV3 region of a heavy chain is encoded by a separate DNA segment, D (D segment for diversity[16]), and that two recombination events, V_H–D and D–J_H joinings, are necessary to form a complete heavy-chain V gene. All germ-line V_H DNAs and J_H DNAs sequenced so far have the conserved heptamer and nonamer separated by the longer spacer. It has therefore been predicted[12,13] that germ-line D segments must contain the same recognition sequences on each side of the coding segment and their spacers must both be short ones, so that the 12/23-base pair spacer rule is adhered to. However, no germ-line D sequence has been reported to date and the predictions outlined above still need to be verified. It was previously observed that the DNA sequences are often rearranged in the J_H cluster in some cloned T-cell lines[17,18], as well as in pre-B-cell lines[18,19]. By the Southern hybridization method using a J_H cluster probe, we observed that DNA rearrangements within the J_H cluster region are not restricted to one allelic chromosome in most of the myelomas. One intriguing possibility was that some of these rearrangements were abortive, representing recombination between a D segment and a J_H segment producing an incompletely rearranged D–J_H structure. We therefore cloned some of the genomic DNA fragments containing the rearranged J_H-cluster sequence from myelomas and characterized the cloned DNA fragments. These studies led to an identification of a potential germ-line D DNA segment, which is reported here. This potential D DNA segment is flanked by the two pairs of conserved sequences in inverted orientations, and the spacers separating the two signal sequences are 12 base pairs long for both.

0028–0836/81/160562—04$01.00

© 1981 Macmillan Journals Ltd

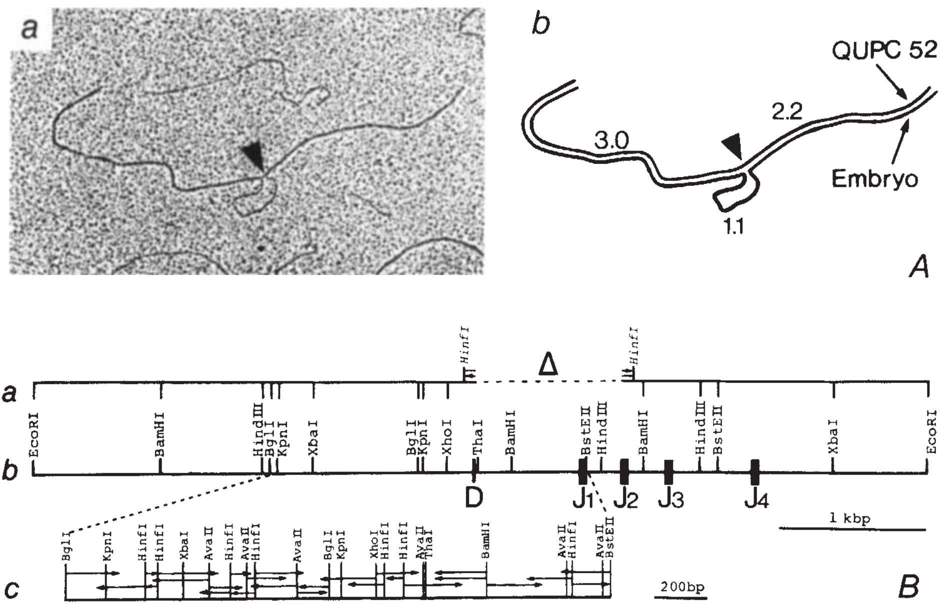

Fig. 2 *A*, Electron micrograph (*a*) and schematic interpretation (*b*) of a heteroduplex molecule formed between two *Eco*RI inserts, one from clone ME184-8 (embryonic J) and the other from clone Q52J (rearranged J from myeloma QUPC52). Procedures used for the heteroduplex formation have been described previously[4]. Clone ME184-8 was prepared by Richard Maki and contains, in a phage λ_{WES} vector, the 6.4-kilobase *Eco*RI embryo DNA fragment carrying the four J_H DNA segments whose sequences were determined previously[13] using another DNA clone referred to as MEP203 (ref. 11). Numbers indicate lengths of various parts of the heteroduplex in kilobases. *B*, Restriction enzyme cleavage maps of the J-containing *Eco*RI fragments. *a*, 5.2-kilobase insert from the myeloma clone Q52J; *b*, 6.4-kilobase insert from the embryonic clone ME184-8. The broken line (Δ) in *Ba* represents a deletion in the myeloma clone; the bars in *Bb* indicate coding regions. The positions of four J_H segments have been determined by R-loop mapping and DNA sequencing[11,13]. The D segment is identified in the present study by DNA sequencing of the 5′ flanking region of J_{H1}. Strategy of the sequencing is shown in *Bc*. kbp, Kilobase pairs; bp, base pairs. The previously published *Bam*HI sites[13] have been corrected.

Frequent rearrangements of the J_H region sequences in myelomas

We analysed *Eco*RI digests of total DNA from 18 different BALB/c myelomas by the Southern gel blotting method[20] using a J_H probe (Fig. 1). The embryo DNA gave a band of 6.4 kilobases as expected from previous studies[11]. All but three (MOPC511, MOPC21 and MOPC321) myelomas gave one, two or three bands that are all different from the embryo band. MOPC511 and MOPC21 each gave two bands, one of which is indistinguishable from the embryo band. MOPC321 gave no band. This myeloma synthesizes no heavy chain and presumably lacks part of the heavy-chain gene sequences in question. In some myelomas a very faint band can be seen at the position of the embryo band, probably due to contamination of non-myeloma cells in the solid tumours from which the DNA was prepared. The variability in the number of bands among various myelomas and in the intensity of bands within a myeloma most probably reflects the fact that myelomas are usually not diploid[21].

We assume that one rearranged band in each myeloma represents the V_H–D–J_H joining required for the formation of a complete heavy-chain gene active in that myeloma. Although the nature of the additional rearrangements observed in many myelomas is unknown, the results show that sequence rearrangements are very frequent in the vicinity of the J_H cluster in all copies of chromosome 12, on which the heavy-chain genes reside[22].

DNA deletion on an abortively-rearranged J_H fragment in myeloma QUPC52

We chose the 5.2-kilobase QUPC52 (Q52) fragment for further analysis and cloned it in the phage vector λ_{WES} (ref. 23). The cloned *Eco*RI fragment was analysed by heteroduplex formation using a germ-line J_H fragment (clone ME184-8) cloned from the *Eco*RI digest of mouse embryo DNA. As shown in Fig. 2*A*, the 5.2-kilobase myeloma fragment carries a 1.1-kilobase deletion, which lies 3.0–4.1 kilobases from the 5′ end of the embryonic J_H fragment. By superimposing the heteroduplex map on the restriction enzyme map of the embryonic clone, we could locate the right end of the deletion near the J_{H2} DNA and the left end about 0.7 kilobases 5′ to the J_{H1} DNA (Fig. 2*B*).

D-like segment is attached to J_H DNA on the abortive clone

To analyse the structure of the DNA segment joined to the J_H region, we purified a 600-base pair *Bgl*I–*Bam*HI fragment (see Fig. 2*Ba*) from the myeloma clone and determined the nucleotide sequence. Figure 3*b* shows the nucleotide sequence of a 110-base pair *Hin*fI–*Hin*fI fragment containing the deletion. As expected, we found the J_{H2} sequence in the sequenced region. By comparing this myeloma sequence with the germ-line J_{H2} sequence (Fig. 3*b*, *c*) we found that the J_{H2} sequence starts with the third nucleotide of the Asp codon at position 101 (number-

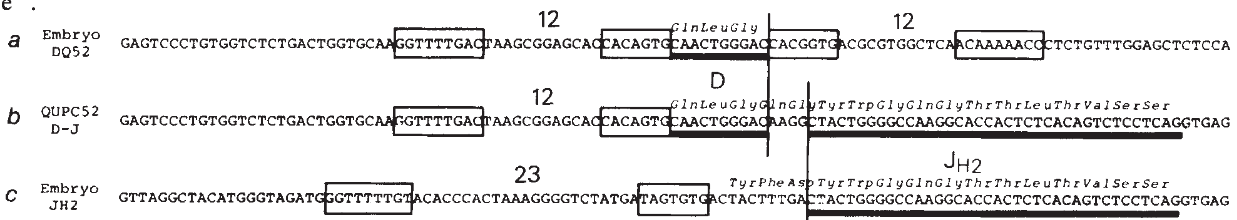

Fig. 3 Nucleotide sequence of the D–J structure in a myeloma clone, Q52J (*b*), and its germ-line sequences, J_{H2} (*c*) and D_{Q52} (*a*). For the sequencing of *b*, a 110-base pair *Hinf*–*Hinf* fragment containing a deletion was purified from a 600-base pair *Bgl*I–*Bam*HI fragment (see Fig. 1*Ba*), both ends labelled with [α-^{32}P]-nucleoside triphosphates, and the strand separated and sequenced in both directions according to the procedures described previously[42]. The embryonic J_{H2} sequence (*c*) was taken from ref. 13. The embryonic D_{Q52} region (*a*) was sequenced according to the strategy shown in Fig. 2*Bc*. Two blocks of conserved sequences,

GGTTTTTGT / CCAAAAACA and CACTGTG / GTGACAC

or their related sequences are boxed. Numbers indicate lengths of spacers separating the two conserved sequences in base pairs. Sequences in *a* and *c* contributing the coding region in *b* are underlined. The corresponding sequences in *b* are also underlined. Note that tetranucleotide AAGG in *b* between the two boundaries of D_{Q52} and J_{H2} is not accounted for by either germ-line sequence. Vertical lines indicate possible recombination sites with D_{Q52} and J_{H2}. Encoded amino acid sequences are in italics.

©1981 Nature Publishing Group

```
KpnI
GGTACCCCCATGGCTTCCCGGAGGCACTGCCCCACCCCCACCATCACAGACCTTTCTCCATAGTTGATAGCTCAGCAGAGTGAATGACAGATGGACCTCCATCTGCTCTTATTTTAAAAA  120

                 XhoI
GAAGACAAACCCCACAGGCTCGAGAACTTTAGCGACTGTTTTGAGAGAAATCATTGGTCCCTGACTCAAGAGATGACTGGCAGTTTGGGGATCAAGAATACCCATACTCTGTGGCTAGTG  240

                                                            12             DQ52              12
TGAGGTTTAAGCCTCAGAGTCCCTGTGGTCTCTGACTGGTGCAAGGTTTTGACTAAGCGGAGCACCACAGTGCAACTGGGACCACGGTGACGCGTGGCTCAACAAAAACCCTCTGTTTGG  360

AGCTCTCCAGGGGCAGCCTGAGCTATGAGGAAGTAGAGAGGCTTGAGAAATCTGAGGAAGAAAAGAGTAGATCTGAGAGGAAAGGTAGCTTTCTGGAGGTCAGGAGACAGTGCAGAGAAG  480

                                                                                                   BamHI
AACGAGTTACTGTGGACAGGTCTTAGATGGGGAAAGAATGAGCAAATGCAAGCATCAGAAGGATAGATGCAATGTCCTGCCAAGGACTTACCAAGAGGATCCCGGACAGAGCAGGCAGGT  600

GGAGTTGACTGAGAGGACAGGGTAGGTGCAGGTTCCTCTCTCGTTTCCTTTCTCCTTCTCCTGTTTCCTTCCTCTCTTGTCACAGGTCTCACTATGCTAGCCAAGGCTAGCCTGAAAGAT  720

TACCATCCTACAGATGGGCCCATCCAGTTGAGTTAAGGTGGAAAGATCTCTCCAAACATCTGAGTTTCTGAGGCTTGGATGCCACTGGGGACGCCAAGGGACTTTGGGCTGGGTTTGGTT  840

GGCCCCACATGTAGGGCTACTTCACTGGGTCTATAATTACTCTGATGTGCTAGGACCAGGGGGCTCAGGTCACTCAGGTCAGGTGAGTCCTGCATCTGGGGACTGTGGGGTTCAGGTGTC  960

                                      22                      JH1                               BstEII
CTAAGGCAGGATGTGGAGAGAGTTTTAGTATAGGAACAGAGGCAGAACAGAGACTGTGCTACTGGTACTTCGATGTCTGGGGCGCAGGGACCACGGTCACCGTCTCCTCAGGTAAGCTGG 1080
```

Fig. 4 Nucleotide sequence of the 5′-flanking region of J_{H1} DNA. Strategy for the sequencing is shown in Fig. 2*Bc*. The J_{H1} sequence and some of the 5′-flanking sequence have been published elsewhere[12,13,43]. The D_{Q52} and J_{H1} segments are underlined. Two blocks of conserved sequences,

GGTTTTTGT / CCAAAAACA and CACTGTG / GTGACAC

or their related sequences are boxed. Numbers indicate the lengths of the spacer separating the two conserved sequences in nucleotide bases.

ing is after Kabat *et al.*[24]). The myeloma sequence of the 3′-flanking region of J_{H2} is identical to the embryonic sequence (data not shown). Thus, the right end of deletion was identified between the second and third nucleotides in the Asp codon (Fig. 3*b*). We then examined the sequence preceding the J_{H2} on the myeloma clone. As shown in Fig. 3*b*, a palindromic heptamer, CACAGTG, and a T-rich sequence, GGTTTTGAC, closely related to the nonamer GGTTTTTGT, are present 14 and 26 base pairs, respectively, 5′ to the recombination site. The spacer separating the heptamer and the nonamer is 12 base pairs long. These structural features of the DNA attached to the Q52 J_{H2} sequence fit the predictions made for the D segments. We could not find any V-like sequence on the 5′ side of the recombination site, and this would have been easy to detect, as the sequence at the end of V DNA coding for residues 90–94 is highly conserved[21].

Identification of a germ-line D segment

If the rearrangement on the Q52 clone represents the deletion between a D segment and a J_H segment, the germ-line D sequence must be on the 6.4-kilobase embryonic J_H fragment. As shown in Fig. 2, the left end of the deletion on the Q52 clone has been mapped 0.7 kilobases 5′ to the J_{H1} DNA on the germ-line clone. We therefore extensively sequenced the 5′-flanking region of the J_H DNA cluster on the embryonic clone. Figure 4 shows the complete nucleotide sequence of the 1-kilobase region upstream from J_{H1} DNA. As predicted, about 700 base pairs 5′ to the J_{H1} we identified the same sequence as found on the 5′ side of Q52 J_{H2} DNA. In this germ-line sequence another set of the palindromic heptamer and the T-rich nonamer are invertedly repeated on the 3′ side (Figs 4 and 3*a*). Again the spacer separating the heptamer and the nonamer is 12 base pairs. Accordingly, this germ-line sequence has the structure predicted for the germ-line D segment. In the D-like segment identified here, a 10-base pair DNA stretch is flanked by two palindromic heptamers,

CACAGTG / GTGTCAC and CACCGTG / GTGGCAC

and further flanked by A+T-rich nonamers,

GGTTTTGAC / CCAAAACTG and ACAAAAACC / TGTTTTTGG

12 base pairs outside. This 10-base pair segment can code for three amino acid residues, Gln-Leu-Gly, Asn-Trp-Asp or Thr-Gly-Thr. Although no published mouse HV3 sequence corresponds to these hypothetical peptide sequences, we tentatively refer to this potential D segment as D_{Q52}.

The tetranucleotides, AAGG, adjacent to the recombination site on the Q52 J_{H2} DNA cannot be explained by either the germ-line J_{H2} or the D_{Q52} sequence (Fig. 3), and at present the origin of this sequence is unknown. We looked for another germ-line D segment encoding the tetramer between the D_{Q52} and J_{H1} sequences and found none (Fig. 4). As shown in Fig. 3*a*, the coding sequence in the embryonic D_{Q52} segment is followed by pentanucleotides CACGG. If we assume one nucleotide deletion and one nucleotide substitution in this pentamer, the tetramer AAGG could be generated. These events might have occurred somatically on recombination (see the accompanying article[17] for more discussion on this point).

In the hope of finding another D segment, we extensively sequenced the 1.3-kilobase segment 5′ to D_{Q52} according to the strategy shown in Fig. 2*Bc* (sequence not shown). However, detection of another D or D-like sequence was unsuccessful: therefore, germ-line D segments do not seem to be as tightly clustered as are J_H DNAs.

Generality of the 12/23-base pair spacer rule

The present study extended the 12/23-base pair spacer rule[12,13] to germ-line D segments. It thus seems that all recombinations leading to formation of a complete V gene are mediated by the same or similar enzymes. These enzymes are thought to contain two DNA-binding proteins, one for a DNA segment carrying the heptamer and nonamer separated by a short (12-base pair) spacer and the other for a DNA segment having essentially the same short sequences separated by a long (23-base pair) spacer[12,13]. As the two pairs of the putative recognition sequences are in an inverted orientation relative to the corresponding coding DNA segments, they could form a transient intermediary duplex if it is stabilized by the recombinase[9,13].

D–J_H joining accompanies a deletion

It seems to be a general rule that sequence rearrangement involving immunoglobulin genes accompanies a deletion of the DNA sequence which separates a pair of joining DNA segments on a germ-line genome. Experimental evidence for this rule was first obtained for V–J joining of the λ_1-type light-chain gene[9], and was confirmed in a subsequent experiment[25] in the κ-type gene. There is also evidence that deletions accompany heavy-chain joining[18,26]. Furthermore, a number of groups reported results which strongly suggest that the same rule applies to the 'switch recombination' of heavy-chain genes[27–30,44]. All the evidence is based on hybridization analysis of mouse DNA using

©1981 Nature Publishing Group

an appropriate probe. The heteroduplex molecule, formed between the embryonic J_H clone and the Q52 J_H clone, provides the most direct evidence for the involvement of deletion in D–J_H joining. As the size of the cloned DNA fragments corresponds well to those of the fragments detected by Southern gel blotting analysis, the deletion cannot be attributed to cloning artefacts (see Fig. 1 for clone Q52 J_H).

D–J_H joining and allelic exclusion

Unlike all previous autosomal gene studies, only the gene on one of the two homologous chromosomes is active in a given cell for immunoglobulin chains[31–33]. This 'allelic exclusion' is manifested in several ways for light chains. A common form is that one copy of the chromosome contains a rearranged, complete gene, and the other retains the germ-line configuration[4]. Alternatively, both alleles are rearranged, but one abortively in several different forms (refs 5, 18, 33–36, 45–48 and P. Kennedy and S.T., unpublished). At the heavy-chain locus, rearrangements on both copies of the chromosome seem to be very common (Fig. 1), and lack of rearrangement as observed in light chains does not seem to be a common form of allelic exclusion. The demonstration that the heavy-chain locus includes D DNA segments suggests a unique way in which this locus can be abortively rearranged. As shown here for myeloma QUPC52, the unexpressed chromosome contains a D–J segment. Although allelic exclusion at the heavy-chain locus could be manifested by other forms of abortive rearrangement observed in the light-chain loci, the existence of D segments may provide additional forms. One might imagine that in addition to D–J joining V–D joining can occur.

The analysis of myeloma DNA indicates that rearrangement in the vicinity of J DNA segments is much more frequent in heavy-chain genes than in light-chain genes. This distinction also applies to normal B cells: little, if any J_H was observed in the germ-line configuration[37], whereas unrearranged $J\kappa$ is observed in significant amounts[38]. The reason for this difference is not clear. However, it is reasonable to assume that the frequency of joining of two separate DNA segments increases as the number of segments increases and/or as the distance between them decreases. If there are many D segments (see the accompanying article[17]) and if they are close to J segments as D_{Q52} is, D–J joining could account for more frequent abortive rearrangement in the heavy-chain locus.

We thank Lisa Sultzman, Lena Angman, Patricia Riegert and André Traunecker for technical assistance, Richard Maki and William Roeder for reading the manuscript, and Michael Potter for myelomas. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche & Co.

Received 7 January; accepted 11 March 1981.

1. Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3628–3632 (1976).
2. Brack, C. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5652–5656 (1977).
3. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 3171–3175 (1978).
4. Brack, C., Hirama, M., Schuller, R. & Tonegawa, S. *Cell* **15,** 1–14 (1978).
5. Lenhard-Schuller, R., Hohn, B., Brack, C., Hirama, M., & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4709–4713 (1978).
6. Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15,** 1133–1144 (1978).
7. Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 3881–3885 (1978).
8. Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280,** 370–375 (1979).
9. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280,** 288–294 (1979).
10. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3450–3454 (1979).
11. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2138–2142 (1980).
12. Early, P. *et al. Cell* **19,** 981–992 (1980).
13. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286,** 676–683 (1980).
14. Bernard, O. & Gough, N. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3630–3634 (1980).
15. Honjo, T. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
16. Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. *Nature* **283,** 35–40 (1980).
17. Kurosawa, Y. *et al. Nature* **290,** 565–570 (1981).
18. Cory, S., Adams, J. M. & Kemp, D. J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4943–4947 (1980).
19. Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209,** 1366–1369 (1980).
20. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
21. Yoshida, T. H., Imai, H. T. & Potter, M. *J. natn. Cancer Inst.* **41,** 1083–1089 (1968).
22. Hengartner, H., Meo, T. & Müller, E. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4494–4498 (1978).
23. Tiemeier, D. C., Enquist, L. & Leder, P. *Nature* **263,** 526–528 (1976).
24. Kabat, E. A., Wu, T. T. & Bilofsky, H. in *Sequences of Immunoglobulin Chains* (NIH, Bethesda, 1979).
25. Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 6022–6026 (1980).
26. Cory, S. & Adams, J. M. *Cell* **19,** 37–51 (1980).
27. Honjo, T. & Kataoka, T. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2140–2144 (1978).
28. Yaoita, Y. & Honjo, T. *Nature* **286,** 850–853 (1980).
29. Rabbitts, T. H. *et al. Nature* **283,** 351–356 (1980).
30. Cory, S. *et al. Nature* **285,** 450 (1980).
31. Pernis, B. *et al. J. exp. Med.* **122,** 853–875 (1965).
32. Cebra, J. S., Colberg, J. E. & Dray, S. *J. exp. Med.* **123,** 547–558 (1966).
33. Max, E. E. *et al. Cell* **21,** 793–799 (1980).
34. Seidman, J. G. & Leder, P. *Nature* **276,** 790–795 (1978).
35. Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 1937–1941 (1980).
36. Seidman, J. G. & Leder, P. *Nature* **286,** 779–783 (1980).
37. Nottenberg, C. & Weissman, I. L. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
38. Joho, R. & Weissman, I. L. *Nature* **284,** 179–181 (1980).
39. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3683–3687 (1979).
40. Sakano, H. *et al. Nature* **277,** 627–633 (1979).
41. Benton, W. D. & Davis, R. W. *Science* **196,** 180–182 (1977).
42. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65,** 498–502 (1980).
43. Newell, N., Richards, J. E., Tucker, P. W. & Blattner, F. *Science* **209,** 1128–1132 (1980).
44. Coleclough, C., Cooper, D. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1422–1426 (1980).
45. Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287,** 603–607 (1980).
46. Choi, E., Kuehl, M. & Wall, R. *Nature* **286,** 776–779 (1980).
47. Schnell, H. *et al. Nature* **286,** 170–173 (1980).
48. Steinmettz, M., Zachau, H. G. & Mach, B. *Nucleic Acids Res.* **6,** 3213–3229 (1979).

Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes

Yoshikazu Kurosawa, Harald von Boehmer, Werner Haas, Hitoshi Sakano, André Trauneker & Susumu Tonegawa

Basel Institute for Immunology, 487 Grenzacherstrasse, Postfach, 4005 Basel 5, Switzerland

The finding that the diversity (D) and joining (J_H) but not the variable (V_H) DNA segments of mouse immunoglobulin heavy-chain genes are joined in the DNA of some cloned cytolytic T cells, led to identification and sequencing of three different D DNA segments. Two segments identified on the embryo DNA carry on both the 5′ and 3′ sides two sets of characteristic sequences separated by a 12-base pair spacer, which have been implicated as recognition signals for a recombinase. The third segment, identified in a form joined with a J_H DNA segment in a T cell, carries the recognition signal on the 5′ side. These results support the 12/23-base pair model for somatic generation of immunoglobulin V genes, and rule out the possibility that the cytolytic T cells use assembled V_H, D and J_H sequences to encode their antigen receptors.

THE variable region of an immunoglobulin chain is encoded in multiple DNA segments scattered along a chromosome of a germ-line genome[1–10]. These DNA segments are assembled into a continuous stretch with concomitant deletion of the spacer

0028–0836/81/160565—06$01.00

© 1981 Macmillan Journals Ltd