ORGANIZATION, STRUCTURE, AND ASSEMBLY OF IMMUNOGLOBULIN HEAVY CHAIN DIVERSITY DNA SEGMENTS

By YOSHIKAZU KUROSAWA* AND SUSUMU TONEGAWA*

From the Basel Institute for Immunology, 4005 Basel, Switzerland

Complete, active, immunoglobulin variable (V)[1] genes are generated by somatic rearrangement of DNA sequences during the differentiation of lymphocyte precursor cells (1, 2). In the mouse, the conventionally defined V region of the light chains, both λ and κ types, consists of two DNA segments: V_L and J_L, which are separate in the germline genome. The two DNA segments are joined at the 3′ end of the V DNA and the 5′ end of the J DNA to create a complete, light chain V gene (3–8). In contrast, the heavy chain variable gene is encoded by three DNA segments, V_H, diversity (D) (which encodes primarily the third hypervariable region) (9), and J_H. Both V_H-D and D-J_H joinings are necessary to generate a complete, heavy chain V gene (10, 11).

Sequencing studies revealed the presence of two blocks of highly conserved sequences around the recombination sites. In the 5′ flanking region of the J DNAs, a palindromic heptamer, CACTGTG, and a T-rich nonamer, GGTTTTTGT, are conserved, and the sequences complementary to the heptamer and nonamer are ubiquitous in the 3′ flanking region of the germline V DNA segments. It has been proposed that these two blocks of sequences constitute a recognition signal for the putative recombinase (7, 8). The lengths of the spacers between the heptamer and the nonamer are strikingly regular (10, 11). The Vλ and Jλ DNA segments have spacers of 23 and 12 base pairs (bp), respectively (3, 12). In contrast, the Vκ and Jκ spacers are 12 ± 1 and 23 ± 1 bp, respectively (7, 8). In addition, the spacers of all V_H and J_H are 23 ± 1 bp (10, 11, 13–15).

We and others proposed that the putative recombinase for V-J and V-D-J joinings contains two DNA-binding proteins: one recognizing the signal sequences with a shorter (12 bp) spacer and the other with a longer (23 bp) spacer (12/23-bp spacer rule), and all recombinations leading to the generation of a complete V gene, namely, Vκ and Jκ, Vλ and Jλ, V_H and D, and D and J_H joinings, are mediated by the same or similar enzymes following this rule (10, 11). This proposal predicted that the germline D segment would carry the signal sequences with a shorter spacer on each side. This prediction was met by our recent study: two DNA segments identified on the germline genome carried the conserved sequences with a 12-bp spacer on each side (16, 17). We now describe the structure and organization of more germline D segments and present a model explaining how these D segments might become part of the complete V_H gene active in myelomas during differentiation of lymphocytes.

* Present address is the Massachusetts Institute of Technology, Center for Cancer Research, room E17-350, 77 Massachusetts Ave., Cambridge, Massachusetts 02139.

[1] *Abbreviations used in this paper:* D, diversity; kb, kilobase; V, variable; bp, base pairs.

Materials and Methods

Bacteria, Phages, and Cells. Phage Charon 4A and *Escherichia coli* DP50 (Su II$^+$, Su III$^+$) were obtained from F. Blattner (University of Wisconsin, Madison, Wis.) (18). Phage λgtWES-λB and *E. coli* 803 were obtained from P. Leder (NIH, Bethesda, Md.) (19) and K. Murray (Embo Institute, Heidelberg, FRG.), respectively. Lysogens used for preparation of packaging mixtures, BHB 2688 [N205, recA$^-$ (λimm434, cIts b2 red3 Eam4 Sam7)/λ] and BHB 2690 [N205, recA$^-$ (λimm434 cIts b2 red3 Dam15 Sam7)/λ] were obtained from B. Hohn (Friedrisch Miescher Institute, Basel, Switz.).[2] Myelomas J606 and S107 were obtained from M. Cohn (Salk Institute, La Jolla, Calif.), and a myeloma, HOPC8, was obtained from M. Potter (NIH). A cytolytic T cell, D.FL.16, directed against fluorescein was provided by W. Haas (Basel Institute for Immunology) (20). A thymoma, WC2, was provided by A. Coutinho (University of Uriea, Uriea, Switz.).

DNA Blots. Southern-blot hybridization was done essentially as described by Southern (21) and modified by Wahl et al. (22). For cellular DNA, 10 μg of EcoRI-digested DNA was used per gel slot, and for phage DNA, 0.3 μg of DNA. DNA separated on an agarose gel was transferred to a nitrocellulose membrane filter (BA85; Schleicher and Schuell, Inc., Keene, N. H.) and hybridized with 2×10^7 cpm (for cellular DNA) and 5×10^6 cpm (for phage DNA) of the appropriate nick-translated probes.

DNA Cloning. Two recombinant phage libraries previously constructed in this laboratory (23, 24) were screened using the SP2D probe. 2×10^5 plaques from the EcoRI partial library and 6×10^5 plaques from the EcoRI* library were screened by the method of Benton and Davis (25). Clone B1-3 was isolated from a complete EcoRI digest of embryo DNA using the preparative agarose gel method (26). The DNA clones containing the rearranged J_H cluster were also isolated according to the same preparative agarose gel method using the J_H probe (23). Clone RC-1 was isolated by the same method using the FL16 D-J probe. All cloning experiments were carried out under P3-EK2 conditions, in accordance with the National Institutes of Health guidelines for recombinant DNA research, issued in June, 1976.

Heteroduplex Analysis. Electron microscope analysis of heteroduplex DNA was carried out as described by Brack (27).

DNA Sequencing. The 3′ ends of DNA fragments were labeled with four deoxyribonucleoside triphosphate radiolabeled with α-^{32}P (Radiochemical Centre, Amersham, England) and Klenov enzyme (Boehringer, Mannheim Biochemicals, Indianapolis, Ind.), as described previously (26). The labeled fragments were sequenced according to the method of Maxam and Gilbert (28).

Results

At Least Eight SP2-Type D DNA Segments Are Clustered in a 62-kb Region. In our previous study we detected DNA rearrangements within the J_H cluster region of both copies of chromosome 12 in most of the myelomas examined (16). Cloning of the rearranged DNA and characterization of the cloned DNA revealed the nature of these rearrangements. On one chromosome the rearrangement is V_H-D-J_H joining, which is required for the generation of a complete V_H gene active in the myeloma. The rearrangement on the other chromosome is often D-J_H joining, without participation of a germline V_H DNA segment (16). This latter, nonproductive rearrangement was also detected in some cloned T cells. For instance, in a cloned cytolytic T cell directed against a hapten 3-(*p*-sulphenyldiazo)-4-hydroxyphenyl acetic acid (SP) (T cell clone C.SP.2), a germline D segment (referred to as $D_{SP2.1}$) is joined with one of the four J_H segments, J_{H3} (17). When a cloned DNA fragment containing the joined $D_{SP2.1}$-J_{H3} segment and its flanking sequences was used as a hybridization probe, several DNA fragments were detected in an EcoRI digest of embryo DNA (17). Cloning and sequencing analysis of one (5 kilobase [kb]) of these DNA fragments led to identifi-

[2] Collins, J., and B. Hohn. 1978. Cosmid, a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage λ heads. *Proc. Natl. Acad. Sci. U. S. A.* **75**:4242.

cation of a second germline D segment $D_{SP2.2}$ that is homologous to but different from $D_{SP2.1}$ (17). Fig. 1 A shows the location of $D_{SP2.2}$ on the 5-kb fragment. We dissected out the 0.85-kb BglII-BglII fragment from this DNA clone as a hybridization probe (SP2D probe), which contained $D_{SP2.2}$ and its 5′ and 3′ flanking sequences (Fig. 1 A). The SP2D probe detected five intense bands in EcoRI-digested BALB/c embryo DNA (Fig. 2). The intensity of the 5-kb band was stronger than the other four bands that were detected at 18, 6.7, 5.4, and 5.2 kb. A very weak band was also detected at 4.0 kb.

To order the various EcoRI fragments detected by the SP2D probe, we screened an EcoRI partial and an EcoRI* partial library and isolated 8 and 26 clones from the two libraries, respectively. In addition, we isolated another clone (clone B1-3) from the 18-kb region of a preparative agarose gel in which an embryo DNA digested with EcoRI had been electrophoretically fractionated. Fig. 3 shows a Southern gel blot of some of the DNA clones.

Clone RI-2 contained four EcoRI fragments (5.4, 4.0, 3.8, and 1.2 kb), of which the 5.4-kb band was positive with the SP2D probe. Clone RI-9 contained the 5.4- and 6.7-kb positive fragments, as well as a 2.8-kb negative fragment. Analysis of hetero-

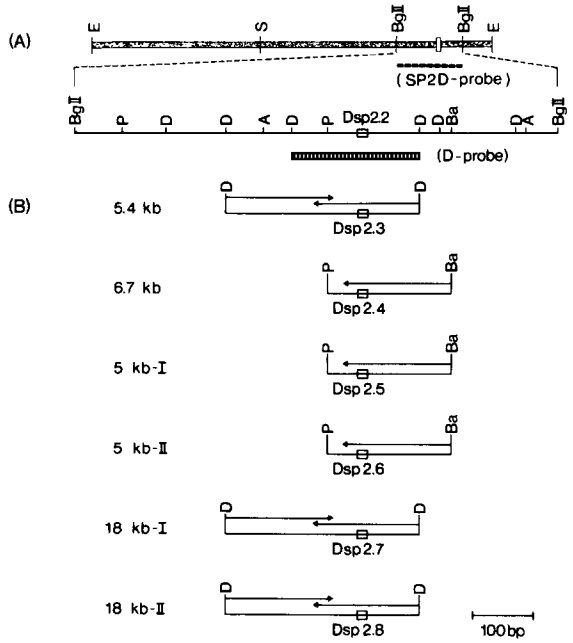

FIG. 1. (A) Restriction-enzyme cleavage map of the EcoRI insert of clone C3H-2 containing $D_{SP2.2}$. The 5-kb EcoRI fragment of clone C3H-2 was described previously (17). The open box indicates the position of the $D_{SP2.2}$ segment. The SP2D probe is a 0.85-kb BglII-BglII fragment. A 200-bp DdeI-DdeI fragment, referred to as D probe, was used in the Southern-blot hybridization to identify the six SP2-type D segments on various DNA clones. (B) Sequencing strategy for the SP2-type D segments. As nucleotide sequences surrounding various SP2-type D segments are very homologous to each other, many of the restriction sites in these regions are the same as those of $D_{SP2.2}$. The DNA fragments containing $D_{SP2.3}$ to $D_{SP2.8}$ were identified using the D probe and sequenced. These fragments are aligned with the 5-kb EcoRI fragment of clone C3H-2 based on common restriction sites. The sequenced regions are indicated by horizontal arrows. E, EcoRI; S, SacI; BgII, BglII; P, PstI; D, DdeI; A, AvaII; Ba, BamHI.

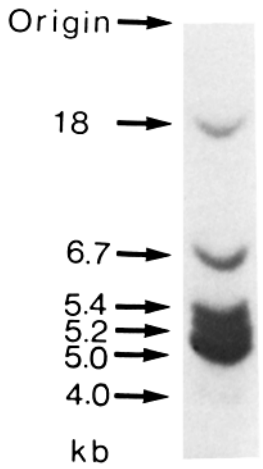

FIG. 2. Southern-blot hybridization of EcoRI-digested BALB/c mouse embryo DNA with the SP2D probe. 10 μg DNA was digested to completion with EcoRI, subjected to electrophoresis on a 0.8% agarose gel, transferred to a nitrocellulose filter, and hybridized with the nick-translated BglII-BglII fragment described in Fig. 1A.

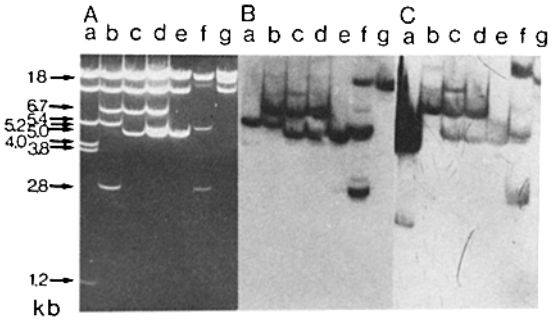

FIG. 3. Southern-blot hybridization analysis of the DNA clones containing SP2D- and FL16D-probe-positive DNA inserts. (A) 0.3 μg of each phage DNA was digested with EcoRI, electrophoresed on a 0.9% agarose gel, and stained with ethidium bromide (0.5 μg/ml). (B) DNA fragments were transferred to a nitrocellulose membrane filter and hybridized with the nick-translated SP2D probe shown in Fig. 1A. (C) The same set of EcoRI digests was subjected to electrophoresis, transferred to a nitrocellulose membrane filter, and hybridized with the nick-translated FL16D probe shown in Fig. 7B. (a) Clone RI-2; (b) clone RI-9; (c) clone RP-1. (d) Clone RP-13; (e) clone RI-3; (f) clone RIS43; (g) clone BI-3.

duplex molecules formed by clone RI-2 and RI-9 confirmed that the 5.4-kb fragments contained by the two clones are identical (data not shown). Clone RP1 contained the 6.7- and 5.0-kb strongly positive fragments. Clone RP13 contained the 6.7- and 5.0-kb positive fragments as well as a 5.2-kb negative fragment. This suggested that the order of the three strongly positive fragments is 5′ 5.4, 6.7, and 5.0 3′. Heteroduplex analysis using the four clones RI-6, RIS42, RIS35, and RIS69, all containing the 6.7-kb fragments, confirmed that the location of the 6.7-kb fragment is between 5.4 kb and 5.0 kb (data not shown). Clone RI-3 contained only 5-kb fragments, but the size of the whole DNA insert was ~15 kb, as determined by electron microscope examination of the heteroduplex molecules formed by this clone and Charon 4A (data not shown). Therefore, we concluded that this clone contains three copies of the 5-kb

fragment. Clone RIS43 contained three positive fragments at 5.2, 2.8, and >8 kb, the last of which was connected to the short phage arm (Fig. 3). Since the strongly positive 2.8-kb fragment was not detected in EcoRI-digested embryo DNA (Fig. 2), one of the two EcoRI sites of this fragment must have been created from the EcoRI* site by joining with the long phage arm. Although we did not isolate a clone containing both 5- and 5.2-kb fragments, we could locate the 5.2-kb fragment between the 5.0- and 18-kb fragments because the >8-kb fragment detected in clone RIS43 could be derived only from the 18-kb fragment. As mentioned above, at least four copies of 5-kb fragments exist, but it is not certain whether some other EcoRI fragments exist between the negative 5.2-kb fragment and the three 5-kb positive fragments.

In summary, as shown in Fig. 4, at least eight EcoRI fragments positive with the SP2D probe are clustered in an ~60-kb region.

Six SP2-Type D Segments Are 17 Nucleotides Long and Carry Two Sets of the Consensus Signal Sequences on Both Sides Separated by a 12-bp Spacer. We determined the nucleotide sequences of six SP2-type D segments. To determine the position of the putative D_{SP2}-type segment in the various cloned EcoRI fragments, we dissected out the 200-bp DdeI-DdeI fragment containing the $D_{SP2.2}$ segment from clone C3H-2 (Fig. 1 A). Using this fragment as a hybridization probe (D probe), we analyzed digests of the various embryo DNA clones. In a first series of experiments, a 0.87-kb BglII-EcoRI fragment derived from the 5.4-kb fragment of clone RI-2, a 0.6-kb BglII-BamHI fragment derived from the 6.7-kb fragment of clone RI-9, a 0.6-kb BglII-BamHI

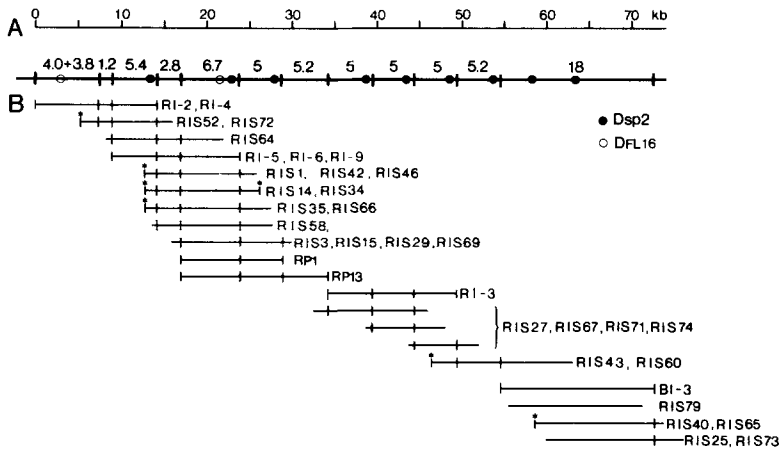

FIG. 4. (A) Linkage map of SP2D probe-positive EcoRI fragments. The complete EcoRI cleavage map is shown. The size of each EcoRI fragment is indicated in kb. ●, positions of SP2-type segments; ○, FL16-type D segments. The positions in the 5.2-kb and 18-kb EcoRI fragments have not been precisely determined and therefore should be considered as tentative. The copy number of 5-kb fragments is at least four. The boundary between the 5.2-kb fragment, which does not contain SP2-type D, and the 5-kb fragment is uncertain. The order of 4-kb and 3.8-kb EcoRI fragments is unknown. (B) DNA inserts contained in various EcoRI partial clones (RI-2–RI-9, RP1, RP13), EcoRI* clones (RIS1–RIS79) and clone B1-3. The inserts are arranged to show their overlapping regions. Identification of overlapping regions has been done by Southern-blot hybridization analysis as shown in Fig. 3 and by heteroduplex analysis (see text). The EcoRI sites are indicated by vertical lines. An asterisk indicates an EcoRI site created from an EcoRI* restriction site by ligation to a vector phage arm. Each of four EcoRI* clones (RIS27, RIS67, RIS71, and RIS74) contains at least one copy of the 5-kb SP2D-probe-positive fragment, but the exact boundaries of the insert in these clones have not been determined.

fragment derived from the 5-kb fragment of clone RI-3, a 3.5-kb EcoRI-BglII fragment (18-kb-I) derived from the 18-kb fragment of clone Bl-3, and a 2.4-kb BglII-BglII fragment (18-kb-II) derived from the same 18-kb EcoRI fragment were identified as D-probe-positive fragments and isolated by acrylamide gel electrophoresis.

In a second series of experiments, these fragments were further digested with appropriate enzymes and 330-bp DdeI-DdeI fragments (Fig. 1 B), derived from 5.4-kb, 18-kb-I, and 18-kb-II, were identified as D-probe-positive fragments and sequenced. The D segment sequence on the 6.7-kb fragment was determined by sequencing a PstI-BamHI fragment (Fig. 1 B). The 0.6-kb BglII-BamHI fragments derived from clone RI-3 (see above) should contain three different copies, each arising from one of the three different EcoRI fragments (Fig. 4). After labeling the ends with a mixture of four α-^{32}P-deoxynucleotide triphosphates and digesting the fragments with PstI, we separated the various DNA components by acrylamide gel electrophoresis, once in 8% acrylamide gel under native conditions, and once in 8% acrylamide gel under denaturing (i.e., strand-separation) conditions. These separation procedures allowed fractionation of the DNA fragments that have the same size but slightly different sequences. The sequences of two different D segments belonging to the 5-kb band could be determined by analysis of fragments fractionated in this way.

The sequences of the six SP2-type germline D segments ($D_{SP2.3}$ to $D_{SP2.8}$) determined in the present study, as well as one D segment ($D_{SP2.2}$) determined previously, are shown in Fig. 5. All seven D segments are 17 nucleotides long and very homologous in sequence. They carry the signal sequences on each side separated by a 12-bp spacer. The relative orientation of the various D segments has been deduced by partial mapping of the EcoRI fragments with selected restriction enzymes (BglII, PstI, BamHI, DdeI), combined with the sequence data. All six D segments have the same 5′-3′ orientation.

Identification of a New Type of D Segment. The nucleotide sequences of the seven SP2-type D segments are considerably homologous to the majority of the D regions of the assembled V genes present in the myelomas and hybridomas studied to date (see Fig. 11 and Discussion). However, there is a notable single-base substitution between the SP2-type germline D regions and the D regions of many myelomas or hybridoma heavy chains: namely, the SP2-type D regions have a heptamer, TA*T*GGTA, in the center, whereas the corresponding sequences of the majority of the myeloma or hybridoma D regions are TA*C*GGTA (Fig. 11), suggesting that these D regions are encoded by a family of D segments different from D_{SP2}.

We previously showed that the J_H sequence is often rearranged in a nonproductive way, both in myelomas (16) and cloned T cells (17). In the case of one myeloma (QUPC52) and two T cell clones (C.SP.2 and B6.1) the nonproductive rearrangements were shown to be D-J joinings. Cloning and DNA sequencing analysis led to the identification of two families (D_{SP2} and D_{Q52}) of germline D segments. To identify a third family of germline D segments we cloned DNA fragments from myelomas and cloned T cells that had been judged to contain a rearranged J_H sequence based on Southern-gel-blotting analyses. Among these cloned DNA fragments, those myeloma DNA clones containing the rearranged, complete V_H gene were identified by hybridization with the respective heavy chain messenger RNA. The rest of the myeloma DNA clones and all T cell DNA clones were analyzed by the Southern-gel-blotting

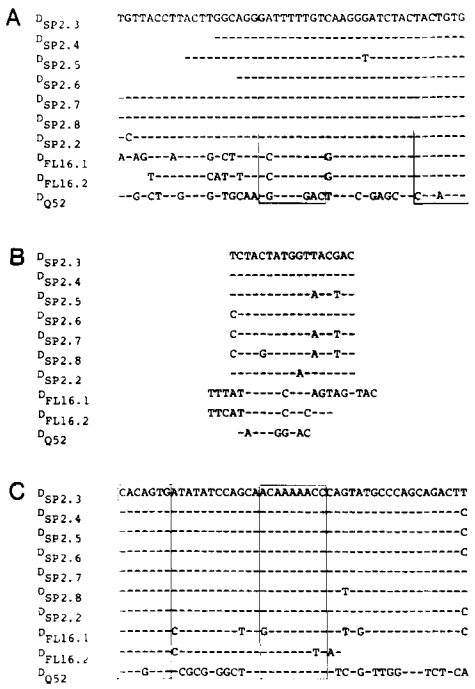

FIG. 5. Comparison of the nucleotide sequences of 10 germline D segments. (A) 5′-flanking region. (B) Coding region. (C) 3′-flanking region. The sequences of $D_{SP2.2}$ and D_{Q52} were described previously (16, 17). The sequences of $D_{SP2.3}$ to $D_{SP2.8}$ were determined using the strategy shown in Fig. 1B. The sequencing strategy of $D_{FL16.1}$ is shown in Fig. 7. The sequence of $D_{FL16.2}$ was determined as follows: The 1.8-kb BglII-BglII fragment shown in Fig. 10Ab was isolated from the 6.7-kb EcoRI fragment of RP13 and digested with HinfI. A 110-bp fragment was identified as the D-containing fragment using the 200 bp AvaII-HinfI fragment as a probe, which contains $D_{FL16.1}$ and is shown in Fig. 7Bb, and sequenced. $D_{SP2.3}$ was used as a reference sequence for comparison with the others. A dash (—) indicates the same nucleotide as that of $D_{SP2.3}$. The signal nonamers and heptamers for the putative recombinase are enclosed by boxes.

procedure using the SP2D probe, with the hope that it will cross-hybridize with a D segment of an expected new family.

Fig. 6 shows the results of the gel-blotting experiment of five DNA clones. The 12-kb insert of clone J606A and the 10-kb insert of clone HOPC8A gave strong hybridization signals with both the J_H and SP2D probes, suggesting that each of these DNA inserts contain a J_H segment joined with an SP2-type D segment. In contrast, the 5.5-kb inserts of clones S107B and WC2-J1 hybridized strongly with the J_H probe and very weakly with the SP2D probe. The same was true for the 3.5-kb insert of clone D16-J2. These results suggested to us that each of the last three DNA clones contains an expected new D segment joined with a J_H segment. We chose the 3.5-kb fragment of clone D16-J2 and the 5.5-kb fragment of clone S107B for further analysis.

Fig. 7A shows the restriction map of the 3.5-kb fragment of clone D16-J2. We isolated a 0.68-kb BamHI-BamHI fragment containing the putative D-J joining region (Fig. 7A) and determined its nucleotide sequence (Fig. 8). The D segment, referred to as $D_{FL16.1}$ carries the heptamer and nonamer recognition sequences separated by a 12-bp spacer. As expected, the $D_{FL16.1}$ sequence is homologous to the D_{SP2} sequences (Fig. 11) but has a heptamer, TA*C*GGTA instead of TA*T*GGTA, in

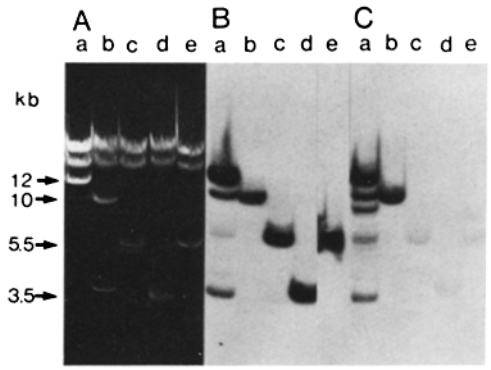

FIG. 6. Southern-blot hybridization analysis of various myelomas and T cell DNA clones containing rearranged J_H sequences. The 12-, 10-, 5.5-, 3.5-, and 5.5-kb fragments detected by the J_H probe (1.1-kb SacI-EcoRI fragment, see Maki et al. [23]) in the EcoRI digests of myelomas J606, HOPC8 and S107; a cytolytic T cell, D.FL.16; and a thymoma, WC2, DNA, respectively, were cloned in the λWES vector and these DNA clones were analyzed. (a) clone J606A; (b) clone HOPC8A; (c) clone S107B; (d) clone D16-J2; (e) clone WC2-J1; (A) 0.3 μg of each phage DNA was digested with EcoRI, subjected to electrophoresis on a 0.9% agarose gel and stained with ethidium bromide (0.5 μg/ml). Clones HOPC8A and S107B contain an additional EcoRI fragment co-cloned with the relevant DNA fragment. In addition to the 12-kb insert, clone J606A gave several weakly stained bands. These bands are considered to have arisen from the 12-kb fragment by deletion of part of its sequence, which occurred during the propagation of the DNA clone in *E. coli*. (B) DNA fragments were transferred to a nitrocellulose membrane filter and hybridized with the nick-translated J_H probe. (C) The same set of EcoRI digests was subjected to electrophoresis, transferred to a nitrocellulose membrane filter, and hybridized with the nick-translated SP2D probe.

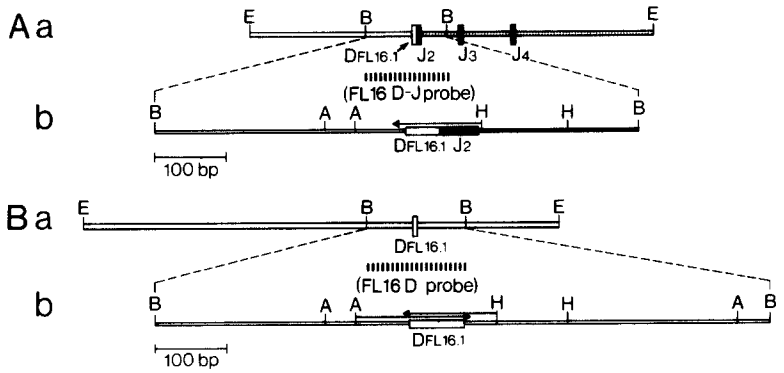

FIG. 7. Restriction-enzyme cleavage maps of 3.5-kb EcoRI insert of clone D16-J2 (A) and 4-kb EcoRI insert of clone RC-1 (B). Heteroduplex analysis, using clone D16-J2 and clone ME184-8 (23) containing J_H segments in the germline configuration, mapped the recombination site at or near the J_{H2} segment. The analysis of the heteroduplex molecule between clone D16-J2 and clone RC-1 indicates the 5'-flanking regions of $D_{FL16.1}$ are the same between these two clones, but the positions of EcoRI sites are different by polymorphism because the T cell clone D.FL.16 is derived from DBA mice. ■, J_H DNA segments; □, $D_{FL16.1}$ segment identified by DNA sequencing. ▬, FL16DJ probe (a 0.68-kb BamHI-BamHI fragment) and the FL16D probe (a 0.85-kb BamHI-BamHI fragment). Nucleotide sequences of the recombination region and the germline $D_{FL16.1}$ segment were determined using the AvaII-HinfI fragments, indicated by the horizontal arrows, respectively. E, EcoRI; B, BamHI; H, HinfI; A, AvaII.

the coding region. The nucleotide sequence of the D-J joining region of clone S107B was similarly determined. The sequence shown in Fig. 12 suggests that this D-J joining occurred also between $D_{FL16.1}$ and J_2.

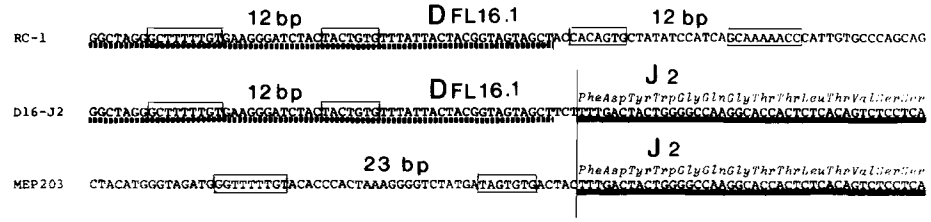

FIG. 8. Nucleotide sequence of the recombination regions of clones RC1, D16-J2 and MEP203. In clone D16-J2, rearrangement was observed between $D_{FL16.1}$ and J_{H2}. The sequence of the J_{H2} region of the embryonic clone MEP203 was determined previously (11). Homologous regions between RC-1 and D16-J2 as well as D16-J2 and MEP203, respectively, are underlined. The vertical line designates the recombination sites. Note that trinucleotide TCT around the DJ joint in clone D16-J2 is not accounted for by either the germline $D_{FL16.1}$ or J_{H2} sequences. The signal nonamer and heptamer for the putative recombinase are in boxes. The lengths of the spacers separating these two sequences are indicated in base pairs.

When EcoRI-digested BALB/c embryo DNA was analyzed by the Southern-gel-blotting procedure using the BamHI-BamHI fragment (FL16DJ probe) (Fig. 7A) containing $D_{FL16.1}$ and its 5′ flanking region as well as J_2 and its 3′ flanking region, we detected an intense band at 4.0 kb in addition to the previously characterized 6.4-kb band (23) containing the embryonic J_H cluster (data not shown). We cloned this 4.0-kb EcoRI fragment in λWES (clone RC-1) and mapped the putative germline D segment within the 0.8-kb BamH1-BamH1 fragment using the FL16DJ probe (Fig. 7B). We then determined the nucleotide sequence of the D segment according to the strategy shown in Fig. 7Bb. As shown in Fig. 8, this embryo clone contains a 23-nucleotide-long D segment flanked by 12-bp-spacer signal sequences. In the 35-bp 5′ flanking region and in the first 21-bp-long coding region, the sequences of clones RC-1 and D16-J2 are identical. This indicates that the D segment of clone RC-1 is indeed the germline counterpart of D16-J2. However, three base pairs occur around the D-J joint that cannot be accounted for by either the germline $D_{FL16.1}$ or J_2 sequence. We previously observed similar cases (16, 17). The implications of these nucleotides are discussed below.

We mapped the $D_{FL16.1}$ segment to the 5′ end of the SP2D cluster (Fig. 4); as shown in Fig. 3C, the 4.0-kb EcoRI fragment contained in the EcoRI-partial clone RI-2 hybridized strongly with the FL16D probe.

Size of FL16D Family. To determine the repertoire of the D_{FL16} family, we analyzed EcoRI-digested embryo DNA using a hybridization probe containing $D_{FL16.1}$ (dissected from the embryo clone RC-1 and referred to as FL16D probe in Fig. 7Ba). In addition to the intense 4.0-kb band that contains $D_{FL16.1}$ the probe detected bands of medium intensity at 6.7 and 5.0 kb and light bands at 18, 10, and 5.2 kb. Since SP2D probe also detected bands at 18-, 6.7-, 5.2-, and 5.0-kb (Fig. 2), we speculated that these bands were observed in the radioautograph shown in Fig. 9 because of cross-hybridization of the FL16D probe with SP2-type D segments. Indeed, FL16D probe hybridized weakly with the corresponding EcoRI fragments of various SP2D clones (Fig. 3C).

To confirm this point, we digested the SP2D-containing clones with BglII and analyzed the digests using SP2D and FL16D probes (Fig. 10). Clone *a* gave a major hybridization-positive band at 0.87 kb when analyzed using the SP2D probe. (In this case, clone *a* means the 5.4-kb EcoRI insert isolated from clone RI-6. See the legend

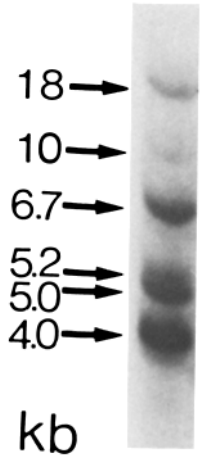

FIG. 9. Southern-blot hybridization of EcoRI-digested BALB/c mouse embryo DNA with the FL16D probe. 10 μg DNA was digested to completion with EcoRI, subjected to electrophoresis on a 0.8% agarose gel, transferred to a nitrocellulose filter, and hybridized with the nick-translated BamHI-BamHI fragment described in Fig. 7B.

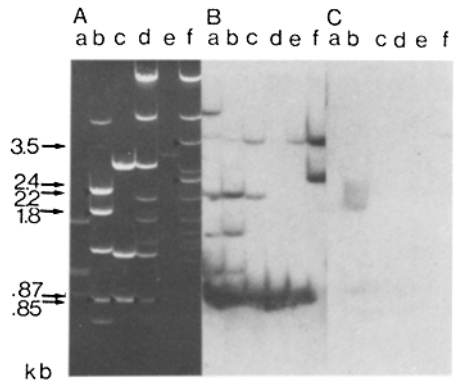

FIG. 10. Southern-blot-hybridization analysis of BglII digests of various EcoRI fragments containing SP2D segments with SP2D probe and FL16D probe. (A) 5.4-kb, 6.7-kb, 5-kb, and 5.2-kb EcoRI inserts were isolated from clones RI-6, RIS69, and RIS43, respectively. 0.03 to 0.1 μg DNA was digested with BglII. 0.3 μg of phage DNA of clone RI-3 and BI-3 were digested with EcoRI and BglII. All the digests were subjected to electrophoresis on a 1% agarose gel and stained with ethidium bromide (0.5 μg/ml). (a) 5.4 kb of clone RI-3, (b) 6.7 kb of clone RI-6, (c) 5 kb of clone RIS69, (d) clone RI-3, (e) 5.2 kb of clone RIS43, (f) clone BI-3. (B) DNA fragments were transferred to a nitrocellulose membrane filter and hybridized with the nick-translated SP2D probe. (C) The same set of BglII digests were subjected to electrophoresis, transferred to a nitrocellulose membrane filter, and hybridized with the nick-translated FL16D probe. ▲, fragments positive to SP2D and FL16D probe.

to Fig. 10 for explanation of the others.) Likewise, clones *b*, *c*, *d*, and *e* each gave a major band at 0.85 kb. In the case of clones *a*, *b*, and *d*, the DNA sequencing study confirmed the presence of an SP2-type D segment in the corresponding hybridization-positive BglII fragment. We expect that the 0.85-kb band in clones *c* and *e* also contains an SP2-type D fragment. The SP2D probe detected two major bands at 3.5 and 2.4 kb in the BglII digest of clone *f*, which were shown to contain $D_{SP2.8}$ and $D_{SP2.7}$, respectively (see above). In addition to these major bands, a number of minor

bands can be observed in Fig. 10 B. In all but two cases (2.2 kb of clone *b* and 1 kb of clone *a*), no corresponding DNA fragment is present in an amount expected from the complete enzyme digestion (Fig. 10 A). We therefore conclude that most of these hybridization-positive bands occur because of incomplete digestion.

The FL16D probe hybridized weakly with most of the BglII fragments containing SP2-type D segments, but not with those BglII fragments lacking them, except for the 2.2- and 1.8-kb fragments arising from the 6.7-kb EcoRI fragment of clone RI-6. It is thus unlikely that the 18-, 5.2- and 5.0-kb EcoRI fragments contain D segments belonging to the D_{FL16} family.

The 2.2- and 1.8-kb fragments of the 6.7-kb EcoRI fragment were sequenced, and a second member of the D_{FL16} family was identified (referred to as $D_{FL16.2}$). Note that BglII splits the region homologous to the FL16D probe into two parts. The nucleotide sequence of $D_{FL16.2}$ shown in Fig. 5 demonstrates that this D segment is 17 bp long and, like all other D segments studied to date, is flanked by 12-bp signal sequences.

The 10-kb fragment that hybridized weakly with the FL16D probe may contain another D segment similar to $D_{FL16.1}$ or $D_{FL16.2}$. Because two neighboring D segments are at least ~5 kb, it is unlikely that the 10-kb EcoRI fragments contain more than two copies of FL16-type D segments. Thus, the maximum number of D_{FL16} families seems to be four.

Discussion

Structure and Diversity of Germline D Segments. To date, we have determined the nucleotide sequences of 10 different germline D segments. All of these D segments are homologous in sequence, suggesting that they share a common primordial sequence. The 10 D segments can be classified into three families based on the extent of sequence similarity in the coding as well as immediately adjacent regions (Fig. 5). The seven D_{SP2} ($D_{SP2.2}$ to $D_{SP2.8}$) together with two more unsequenced D segments on the 5- and 5.2-kb EcoRI fragments (see Fig. 5) constitute one family. They are clustered, the distance between the two adjacent segments being either ~5 or ~10 kb (Fig. 4). The two D_{FL16} segments, $D_{FL16.1}$ and $D_{FL16.2}$, compose a second family. They are located in the 5′ end region of the D_{SP2} cluster. The 10-kb EcoRI fragment detected by the FL16D probe (Fig. 9) may contain the third member of this family, but this must be confirmed by DNA sequencing. D_{Q52} seems to constitute the third family by itself. This D segment has been mapped ~700 bp 5′ to the J_H cluster (16).

Among the three D families, D_{Q52} is much less related to either D_{SP2} or D_{FL16} than the latter two are related to each other. Thus, the conservation of the signal heptamer and nonamer sequences stands out when D_{Q52} is compared with the other D segments (Fig. 5). Also well conserved is the length of the spacer between the heptamer and the nonamer: all 10 D segments carry the signal sequences with exactly 12-bp spacers. In contrast, the distance between the two signal heptamers—namely, the length of the coding region—is variable from one family to another, and also within the family, in the case of D_{FL16}. This length variation seems to have arisen by insertion or deletion of a short block of sequence. For instance, deletion of a heptamer, TAGTAG, from $D_{FL16.1}$ (and one base substitution elsewhere) can generate $D_{FL16.2}$. Base substitution also seems to play a role in diversification of the D-coding regions, which is most evidently seen among the D_{SP2} members. Substitutions are greater in the coding regions than in the spacer regions. Thus, the D-coding regions seem to diversify

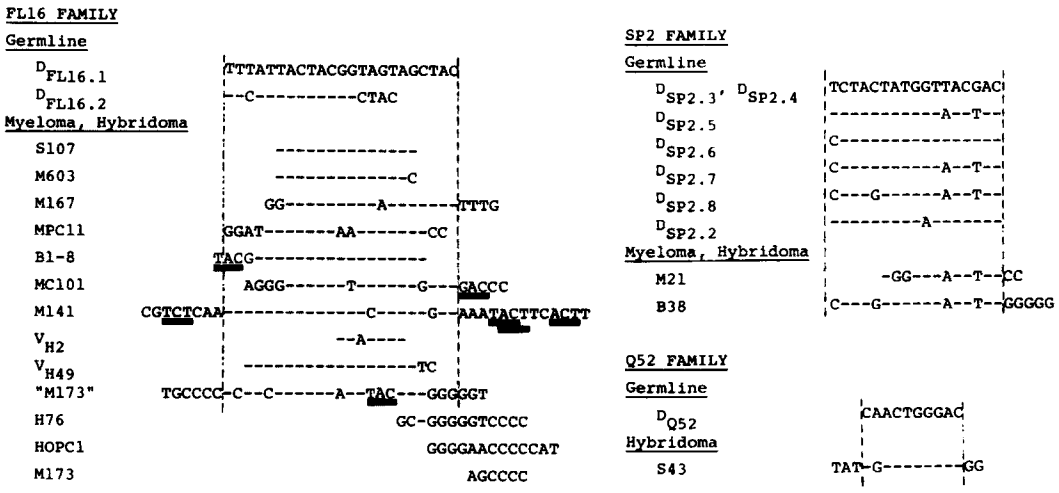

FIG. 11. Nucleotide sequences of germline D segments and D regions of complete heavy-chain genes. The nucleotide sequences of 10 germline D segments and the D regions of 16 complete V_H genes are listed. To determine the boundaries between V_H and D as well as between D and J_H in the assembled V gene, it was necessary to compare the nucleotide sequences of three regions, namely the 3′ end of the germline V gene, the assembled V_H-D-J_H and the 5′ end of the germline J_H segment. For M141 (11), anti-phosphorylcholine (PC) myelomas (S107, M603, M167) (10), and anti-(4-hydroxy-3-nitrophenyl)acetyl (NP) hybridomas (B1-8, S43) (15) three sets of nucleotide sequences are available. The nucleotide sequences in the 3′ ends of these three different germline V genes are highly conserved. Assuming the other germline V genes also have the same degree of sequence homology, the V_H-D boundaries of complete genes were deduced for those V genes whose germline V genes have not been sequenced, namely MPC11 (29), H76 (14), V_{H2} (30), V_{H49} (30), MC101 (31), and "M173" (17). The D sequence of B38 is unpublished (P. Early, personal communication). These sequences are classified into three families based on the extent of sequence homology. D sequences of H76, HOPC1, and M173 cannot be classified into these families. The sequences of $D_{FL16.1}$, $D_{SP2.3}$, and D_{Q52} are used as a reference sequence of each family for comparison with the others. A dash (—) indicates the same nucleotide as the reference. Vertical broken lines indicate the 5′ and 3′ boundaries of the 23-bp $D_{FL16.1}$, 17-bp $D_{SP2.3}$, and 10-bp D_{Q52}, respectively, and are shown for comparison of the lengths of various germline and myeloma (or hybridoma) D sequences. The heavy underlines indicate the short sequences homologous to the L or R sequences of SP2D (see text).

rapidly, both by base substitution and insertion or deletion of an oligonucleotide. This makes sense because diversification of D sequences is an effective way for increasing antibody repertoire.

Comparison of Germline D Segments with D Regions of Complete Heavy Chain Genes. Fig. 11 shows a comparison of the nucleotide sequences of the 10 germline D segments with those of the D regions of 16 complete V_H genes studied to date. Based on sequence homology, the majority of the 16 somatic D sequences can be assigned to one of the three germline D families. The exceptions are H76, HOPC1, and M173, which are rich in G-C pairs and correspond to an unidentified germline D family. Each of the three germline D families has at least one somatic D sequence assigned, but in no case do the germline and somatic sequences match exactly. Does this mean that we have not yet identified a germline D segment corresponding to any of the known somatic D sequences? We believe that this is not the case because of the following reasons. We have determined the nucleotide sequences of all or most members of each of the three germline D families. It is highly unlikely that only the remaining few D segments yet to be sequenced will match the somatic sequences. On

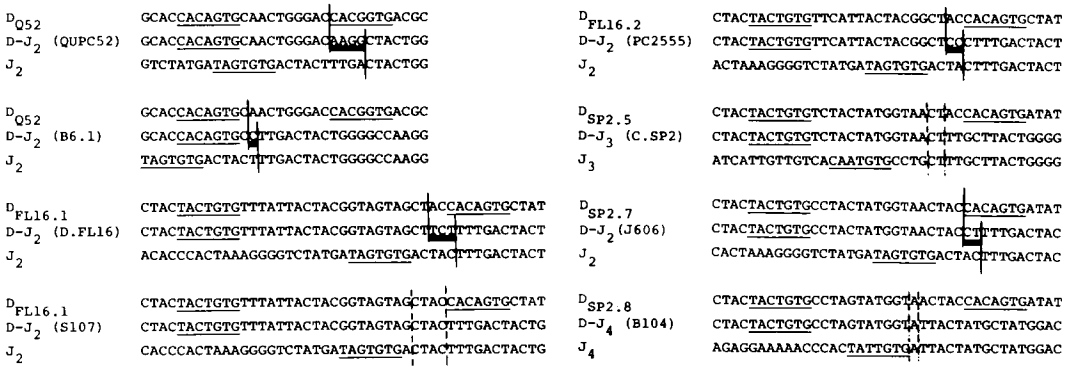

FIG. 12. Nucleotide sequences of eight joined D-J segments isolated from myelomas and cloned T cells. The sequences of the D-J joining regions were compared with the sequence of corresponding germline D and J_H. The sequences of D_{Q52}, D-J_2 derived from myeloma QUPC52, D-J_2 from T cell B6.1, and D-J_3 from T cell C.SP.2 were also described previously (16, 17). The germline J_H sequences were also published (11). The sequences of D-J_2 derived from myeloma PC2555, D-J_2 from myeloma J606, and D-J_4 from B cell B104 are unpublished (H. Sakano, C. Wood, and P. Early, personal communications). The other sequences are described in this paper. The signal heptamer for the putative recombinase is underlined. Vertical lines designate the recombination sites. Heavy underlining indicates the extra nucleotide that can be accounted for by neither the germline D nor J sequences. In the case of S107, C.SP.2, and B104, the exact recombination sites cannot be assigned because the same nucleotides are overlapped around the D-J joining sites among the germline D, joined D-J, and germline J clones. Vertical broken lines designate the overlapped region.

the other hand, it is also unlikely that the hybridization experiments failed to detect germline D segments corresponding to the somatic D sequences, at least in some cases listed in Fig. 11. For instance, the 14-bp S107 D sequence is identical to the center portion of $D_{FL16.1}$ and the 5′ 17-bp portion of the B38 D sequence is the same as that of $D_{SP2.8}$. Note that under conditions of annealing employed, hybridization was shown to occur to a detectable level between less-homologous pairs, namely the FL16D probe and most members of the D_{SP2} family, or the SP2D probe and members of the D_{FL16} family (Figs. 3 and 10). Further support for our belief is as follows: We and others determined the nucleotide sequences of eight joined D-J segments isolated from myelomas and cloned T cells. As shown in Fig. 12, in all cases, the sequence of the D segment and its 5′ flanking region is the same as one of the 10 germline sequences that we identified. Unless there are special sets of D segments that are never involved in abortive D-J joining, this result indicates that only a few more D segments may remain to be found.

How then do we explain the sequence discrepancy between the germline and somatic D sequences? Three types of sequence discrepancy can be observed in Fig. 11. The first type is single- or occasionally double-base substitutions. The second type is the absence of several nucleotides at one or both ends of the somatic sequence. The third type of discrepancy is the presence of several extra nucleotides at one or both ends of the somatic sequence. The first and second types of sequence discrepancy can be explained by known mechanisms, namely, by somatic mutation and by modulation of the V-D and/or D-J joining sites, respectively. It has been shown first in mouse λ light chains (12, 32, 33) and recently in heavy chains (15, 34, 35) that somatic mutations occur in immunoglobulin genes primarily in the three complementarity-determining regions, of which the D region is one. It has also been shown that the joining end of an immunoglobulin gene segment is not fixed to a particular phospho-

diester bond, but can vary from one joining event to another within the range of several base pairs (7, 8, 17). Combinations of these two known mechanisms allow generation of several somatic D sequences listed in Fig. 11. Examples are S107, M603, and V_{H2} D regions that can be accounted for by $D_{FL16.1}$.

The rest of the somatic D sequences listed in Fig. 11, which constitute the overwhelming majority, show the third type of sequence discrepancy in addition to base substitutions. The presence of several extra base pairs at the end of the somatic D sequence unaccounted for by the corresponding germline D sequence had independently been observed during the nucleotide sequence analysis of joined D-J segments isolated from myelomas and cloned T cells (16, 17). Fig. 12 compiles the nucleotide sequences of eight such segments studied to date. In five out of the eight cases, one to four extra nucleotides occur around the D-J joining region that can be accounted for neither by the germline D nor J sequences. We believe that this phenomenon and the aforementioned third type of sequence discrepancy between germline and somatic sequences are equivalent: they reflect the feature of the mechanism by which a germline D segment undergoes joining. One possible mechanism is that the extra nucleotides are inserted during the joining event. It should however be mentioned that equivalent phenomena have not been observed in the V-J joining of light chains. An alternative, but not necessarily exclusive, possibility is that the extra nucleotides arise from independent germline D segments. According to this hypothesis, two or more germline D segments join to code for a somatic D sequence in some cases. Below we will present a specific model for D-D joining.

Possible Dual Signal in the SP2-Type D Segments. We previously speculated on the possible occurrence of D-D joining as part of the V_H gene assembly process (17).

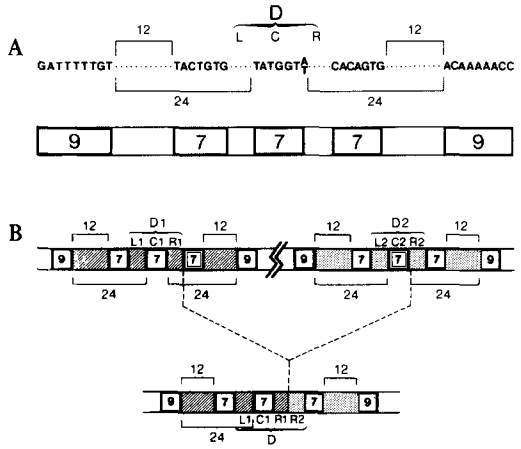

FIG. 13. Model for joining multiple germline D segments. (A) The anatomy of a D_{SP2} segment is shown in two forms. A 17-bp D_{SP2} segment is hypothesized to be composed of three regions, the 5′ pentamer L, the core pentamer C, and the 3′ pentamer R. The core pentamer is thought to have an equivalent function as the signal heptamers flanking the D segment. On both the 5′ and 3′ sides the length of the spacer between the signal heptamer and signal nanomer is 12 bp, whereas that of the core heptamer and signal haptamer is 24 bp. (B) A hypothetical joining of two SP2-type D segments, D1 and D2, is illustrated. The joining occurred using the 12-bp spacer signal at the 3′ side of the D1 segment and the 24-bp spacer signal at the 5′ side of the D2 segment and generated the composite D segment consisting of the D1 and R2. This recombinant D segment carries both a 12-bp and a 24-bp spacer signal at the 5′ side, but only a 12-bp spacer signal at the 3′ side.

According to the 12/23-bp spacer rule, this hypothesis requires a germline D segment having the signal sequence with a 23-bp spacer on at least one side. As described above, none of the 10 D segments studied to date seemed to fulfill this requirement. However, close examination of the nucleotide sequences of the SP2-type D segments reveals a possibility that they carry dual recognition signals. As shown in Fig. 13 A, we could imagine that the 17-bp SP2-type-D segments are composed of three regions, the central or core (C) heptamer, TATGGT$\begin{pmatrix}T\\A\end{pmatrix}$, being flanked by the 5′ or left (L) pentamer, $\begin{pmatrix}T\\C\end{pmatrix}$CTAC, and the 3′ or right (R) pentamer, AC$\begin{pmatrix}G\\T\end{pmatrix}$AC. The core heptamers show considerable similarity to the signal heptamers, TACTGTG and CACTGTG, which flank the D segments. Particularly noteworthy is the second A and the fifth T that are conserved without exception among the signal heptamers of all germline V gene segments, V_L, J_L, V_H, D, and J_H (11; and Fig. 5). In addition, the core heptamers, particularly those with A as the last base, share with the signal heptamers the palindromic or pseudo-palindromic structure around the central base pair. If we assign the core heptamers the function equivalent to that assigned to the signal heptamers, the spacer between the newly adopted signal heptamer (i.e., core heptamer) and the signal nonamer is 24 bp on each side of a SP2-type D segment (Fig. 13 A). This 24-bp spacer is apparently acceptable as the 23-bp-type spacer (or a longer spacer) because we have a precedent in $J_{\kappa 4}$ (7, 8).

Fig. 13 B illustrates a hypothetical joining of two SP2-type D segments based on the aforementioned dual-signal model. Here, the joining is thought to occur between D_1 and D_2 using the right (or 3′) 12-bp spacer signal of D_1 and the left (or 5′) 24-bp spacer signal of D_2, generating a composite D segment having the R_2 segment attached to the 3′ side of the D_1 segment. The newly formed D segment has lost the 24-bp signal at the 3′ side but still carries a 12-bp signal brought in by the R_2 segment. It could therefore join at the 3′ side with another D segment using the 24-bp signal of the latter, or join with a J_H segment, which can provide only a 23-bp spacer. In an analogous manner, a D segment can grow to the left (5′ side) by addition of a single or multiple L segment(s) at a time, or it can join with a V_H segment. R and L segments of SP2D segments are pentamers. However, since the joining end is apparently flexible in the range of several base pairs (17), the growth of a D-coding region by the aforementioned mechanism does not necessarily occur by five or multiples of five base pairs. Unlike SP2-type D segments, the $D_{FL16.1}$ segment carries only a 12-bp spacer signal on each side (note that L is 8 bp long and R is at least 9 bp long in $D_{FL16.1}$). It can participate in D-D joining if the partner provides a 24-bp spacer signal as the D_{SP2}-type D segments would do.

In the proposed model, the core sequence is eliminated from the recombinant whenever a D segment provides a 24-bp spacer signal. Thus, the core sequence of a D_{SP2}-type D segment may be lost during successive D-D joining events. In contrast, the corresponding region of a D_{FL16}-type D segment is protected from loss because it cannot provide a 24-bp spacer signal. This may explain, at least in part, why more D regions of the complete V_H genes studied to date contain the sequence corresponding to the core segment of the FL16-type (or TACGGTA-type) D segment rather than that of SP2-type (or TA*T*GGTA-type) D segment. This is not to suggest that the core heptamers of SP2-type D segments never appear in the D regions of complete heavy-

chain V_H genes. Examples are seen in M21 and B38 (Fig. 11). The D_{SP2} core will also be protected once it acquires at least one round of extension on both the 5′ and 3′ sides. In addition, an SP2-type D segment can directly join with germline V_H and J_H segments. That the D_{SP2} core appears less frequently than the D_{FL16} core in the somatic D sequences suggests that joining of multiple D segments is frequent.

One additional indirect piece of evidence supports the D-D joining model. A short homology seems to occur between the L or R regions of D_{SP2} segments and the 5′ or 3′ extensions of some of the somatic D regions, where homology to the $D_{FL16.1}$ sequence is absent. For instance, the trimers, TAC and ACT, in the 3′ extension of M141 may originate from the same sequence found in the R regions of D_{SP2}. Other possible examples are indicated in Fig. 11.

If validated by more direct evidence, the proposed mechanism provides extreme versatility in the use of germline D segments for generation of a diverse set of complete V_H genes. It will be particularly so if multiple D segments can assemble independently of their order in the germline genome. This may be accomplished by an exchange of DNA sequences between two homologous chromosomes or two sister chromatids, or by excision and insertion of a DNA sequence. As mentioned above (Fig. 12), one and four nucleotides are inserted between D_{Q52} and J_{H2} in the recombinants isolated from myeloma QUPC52 and killer T cell clone B6.1, respectively. One likely explanation is that these nucleotides derive from independent germline D segments, and yet no D segment seems to be present between D_{Q52} and J_{H2} in the germline genome (16). Thus, these recombinants may be examples of D-D joinings that occurred by one of the mechanisms described above.

Summary

We have identified, cloned, and sequenced eight different DNA segments encoding the diversity (D) regions of mouse immunoglobulin heavy-chain genes. Like the two D segments previously characterized (16, 17), all eight D segments are flanked by characteristic heptamers and nonamers separated by 12-bp spacers. These 10 D segments, and several more D segments identified but not yet sequenced, can be classified into three families based on the extent of sequence homology. The SP2 family consists of nine highly homologous D segments that are all 17-bp long and clustered in a chromosomal region of ~60 kb. The FL16 family consists of up to four D segments, two of which were mapped in the 5′ end region of the SP2-D cluster. The two FL16D segments are 23 and 17 bp long. The third, the Q52 family, is a single-member family of the 10-bp-long D_{Q52}, located 700 bp 5′ to the J_H cluster. We argue that the D-region sequences of the majority of heavy chain genes arise from these germline D segments by various somatic mechanisms, including joining of multiple D segments. We present a specific model of D-D joining that does not violate the 12/23-bp spacer rule.

We thank Lena Ångman, Patricia Riegert and André Traunecker for competent technical assistance. We also thank Charles Wood, Hitoshi Sakano, and Philip Early for communicating unpublished results, Michael Potter and Melvin Cohn for providing myelomas, and Werner Haas for providing a T cell clone. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche and Co., Ltd., Basel, Switzerland.

Received for publication 14 September 1981.

References

1. Dreyer, W. J., and J. C. Bennett. 1965. The molecular basis of antibody formation: a paradox. *Proc. Natl. Acad. Sci. U. S. A.* **54:**864.
2. Hozumi, N., and S. Tonegawa. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. U. S. A.* **73:**3628.
3. Tonegawa, S., A. M. Maxam, R. Tizard, O. Bernard, and W. Gilbert. 1978. Sequence of a mouse germline gene for a variable region of an immunoglobulin light chain. *Proc. Natl. Acad. Sci. U. S. A.* **74:**3171.
4. Brack, C., M. Hirama, R. Lenhard-Schuller, and S. Tonegawa. 1978. A complete immunoglobulin gene is created by somatic recombination. *Cell.* **15:**1.
5. Lenhard-Schuller, R., B. Hohn, C. Brack, M. Hirama, and S. Tonegawa. 1978. DNA clones containing mouse immunoglobulin κ chain genes isolated by in vitro packaging into phage λ coats. *Proc. Natl. Acad. Sci. U. S. A.* **74:**4709.
6. Seidman, J. G., and P. Leder. 1978. The arrangement and rearrangement of antibody genes. *Nature (Lond.).* **276:**790.
7. Sakano, H., K. Hüppi, G. Heinrich, and S. Tonegawa. 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature (Lond.).* **280:**288.
8. Max, E. E., J. G. Seidman, and P. Leder. 1979. Sequences of five recombination sites encoded close to an immunoglobulin κ constant region gene. *Proc. Natl. Acad. Sci. U. S. A.* **76:**3450.
9. Schilling, J., B. Clevinger, J. M. Davie, and L. Hood. 1980. Amino acid sequence of homogeneous antibodies to dextran and DNA rearrangements in heavy-chain V-region gene segments. *Nature (Lond.).* **283:**35.
10. Early, P., H. Huang, M. Davis, K. Calame, and L. Hood. 1980. An immunoglobulin heavy-chain variable region gene is generated from three segments of DNA: V_H, D and J_H. *Cell.* **19:**981.
11. Sakano, H., R. Maki, Y. Kurosawa, W. Roeder, and S. Tonegawa. 1980. Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. *Nature (Lond.).* **286:**676.
12. Bernard, O., N. Hozumi, and S. Tonegawa. 1978. Sequences of mouse immunoglobulin light-chain genes before and after somatic changes. *Cell.* **15:**1133.
13. Rabbitts, T. H., G. Matthyssens, and P. H. Hamlyn. 1980. Contribution of immunoglobulin heavy-chain variable-region genes to antibody diversity. *Nature (Lond.).* **284:**238.
14. Bernard, O., and N. M. Gough. 1980. Nucleotide sequence of immunoglobulin heavy-chain joining segments between translocated V_H and μ constant region genes. *Proc. Natl. Acad. Sci. U. S. A.* **54:**864.
15. Bothwell, A. L. M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D. Baltimore. 1981. Heavy-chain variable region contribution to the NP^b family of antibodies: somatic mutation is separable from germline combinatorial events. *Cell.* **24:**625.
16. Sakano, H., Y. Kurosawa, M. Weigert, and S. Tonegawa. 1981. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. *Nature (Lond.).* **290:**562.
17. Kurosawa, Y., H. von Boehmer, W. Haas, H. Sakano, A. Traunecker, and S. Tonegawa. 1981. Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes. *Nature (Lond.).* **290:**565.
18. Blattner, F. R., B. G. Williams, A. E. Blechl, K. Denniston-Thompson, H. E. Faber, L. Furlong, D. J. Grunwald, D. O. Kiefer, D. D. Moore, J. W. Schumm, E. L. Sheldon, and O. Smithies. 1977. Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. *Science (Wash. D. C.).* **196:**161.
19. Leder, P., D. Tiemeier, and L. Enquist. 1977. EK2 derivatives of bacteriophage lambda

useful in the cloning of DNA from higher organisms: the λgtWES system. *Science (Wash. D. C.)*. **196:**175–177.

20. Haas, W., J. Mathur-Rochat, H. Pohlit, M. Nabholz, and H. von Boehmer. 1980. Cytotoxic T cell responses to haptenated cells. III. Isolation and specificity analysis of continuously growing cells. *Eur. J. Immunol.* **10:**828.
21. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:**503.
22. Wahl, G. M., M. Stern, and G. R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenoyloxymethyl paper and rapid hybridization by using dextran sulfate. *Proc. Natl. Acad. Sci. U. S. A.* **76:**3683.
23. Maki, R., A. Traunecker, H. Sakano, W. Roeder, and S. Tonegawa. 1980. Exon shuffling generates an immunoglobulin heavy-chain gene. *Proc. Natl. Acad. Sci. U. S. A.* **77:**2138.
24. Roeder, W., R. Maki, A. Traunecker, and S. Tonegawa. 1981. Linkage of the four γ subclass heavy-chain genes. *Proc. Natl. Acad. Sci. U. S. A.* **78:**474.
25. Benton, W. D., and R. W. Davis. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. *Science (Wash. D. C.)*. **196:**180.
26. Sakano, H., J. H. Rogers, K. Hüppi, C. Brack, A. Traunecker, R. Maki, R. Wall, and S. Tonegawa. 1979. Domains and the hinge region of an immunoglobulin heavy chain are encoded in separate DNA segments. *Nature (Lond.)*. **277:**627.
27. Brack, C. 1981. DNA electron microscopy. *Crit. Rev. Biochem.* **10:**113–169.
28. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65:**499.
29. Zakut, R., J. Cohen, and D. Givol. 1980. Cloning and sequence of the cDNA corresponding to the variable region of immunoglobulin heavy-chain MPC 11. *Nucleic Acids Res.* **8:**3591.
30. Hood, L., M. Davis, P. Early, K. Calame, S. Kim, S. Crews, and H. Huang. 1980. Two types of DNA rearrangements in immunoglobulin genes. *Cold Spring Harbor Symp. Quant. Biol.* **45:**887.
31. Honjo, T., T. Kataoka, Y. Yaoita, A. Shimizu, N. Takahashi, Y. Yamawaki-Kataoka, T. Nikaido, S. Nikai, M. Obata, T. Kawakami, and Y. Nishida. 1980. Organization and reorganization of immunoglobulin heavy-chain genes. *Cold Spring Harbor Symp. Quant. Biol.* **45:**913.
32. Weigert, M. G., I. M. Cesari, S. J. Yonkovich, and M. Cohn. 1970. Variability in the lambda light-chain sequences of mouse antibody. *Nature (Lond.)*. **228:**1045.
33. Tonegawa, S. 1976. Reiteration frequency of immunoglobulin light-chain genes: further evidence for somatic generation of antibody diversity. *Proc. Natl. Acad. Sci. U. S. A.* **73:**203.
34. Crews, S., J. Griffin, H. Huang, K. Calame, and L. Hood. 1981. A single V_H gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. *Cell.* **25:**59.
35. Gerhard, P. J., N. D. Johnson, R. Douglas, and L. Hood. 1981. IgG antibodies to phosphocholine exhibit more diversity than their IgM counterparts. *Nature (Lond.)*. **291:**29.