# Somatic Variation Precedes Extensive Diversification of Germline Sequences and Combinatorial Joining in the Evolution of Immunoglobulin Heavy Chain Diversity

By Kristin R. Hinds-Frey, Hikaru Nishikata, Ronda T. Litman, and Gary W. Litman

*From the Department of Pediatrics, University of South Florida, All Children's Hospital, St. Petersburg, Florida 33701*

## Summary

In *Heterodontus*, a phylogenetically primitive shark species, the variable ($V_H$), diversity ($D_H$), joining ($J_H$) segments, and constant ($C_H$) exons are organized in individual ~18–20-kb "clusters." A single large $V_H$ family with >90% nucleic acid homology and a monotypic second gene family are identified by extensive screening of a genomic DNA library. Little variation in the nucleotide sequences of $D_H$ segments from different germline gene clusters is evident, suggesting that the early role for $D_H$ was in promoting junctional diversity rather than contributing unique coding specificities. A gene-specific oligodeoxynucleotide screening method was used to relate specific transcription products (cDNAs) to individual gene clusters and showed that gene rearrangements are intra- rather than intercluster. This provides further evidence for restricted diversity in the immunoglobulin heavy chain of *Heterodontus*, from which it is inferred that combinatorial diversity is a more recently acquired means for generating diversity. The observed differences between cDNA sequences selected and the sequences of segmental elements derived from conventional genomic libraries as well as from $V_H$ segment-specific libraries generated by direct PCR amplification of genomic DNA indicate that the $V_H$ repertoire is diversified by both junctional diversity and somatic mutation. Taken together, these findings suggest a heretofore unrecognized contribution of somatic variation that preceded both extensive diversification of the germline repertoire and the combinatorial joining process in the evolution of humoral immunity.

The Ig heavy chain gene family is one of the most complex genetic systems that has been described. The organization of heavy chain genes is well established for the mouse (1–3) and human (1, 4, 5). In both cases, as well as in other mammals, variable ($V_H$), diversity ($D_H$), and joining ($J_H$) segments consisting of several hundred tandemly arrayed $V_H$ and fewer numbers of $D_H$ and $J_H$ elements occupy a chromosomal region of 2–3 × $10^6$ base pairs. Functional Ig $V_H$ genes arise in B lymphocytes by rearrangement of individual segmental elements, so called V-D-J-joining, a process that is mediated by recombination signal sequences (1) and is catalyzed by recombinase enzymes (1, 6, 7). The use of different combinations of $V_H$, $D_H$, and $J_H$ segments is a significant factor in the generation of antibody diversity, as are variations in junctional joining, somatic mutation (1, 2), and, in at least one case, gene conversion (8). The mammalian type of $V_H$ organization is not preserved in all vertebrates. Antibody diversity in *Gallus*, an avian, is generated from the rearrangement of single functional $V_H$ (9) and $V_L$ (10) genes by a process of gene conversion that uses flanking pseudogenes. A similar mechanism is used in another avian species (11).

In elasmobranchs, such as shark and skate, the $V_H$, $D_H$, $J_H$, and $C_H$ (12, 13), as well as $V_L$, $J_L$, and $C_L$ (14), elements of Ig genes are organized in a unique manner. In *Heterodontus francisci* (horned shark), a phylogenetically primitive species, these elements are linked in >100 "clusters," ~18–20 kb, each containing a single $V_H$, two $D_H$, and single $J_H$ and $C_H$ segments (V-$D_1$-$D_2$-J-C) (13). The four recombining elements within the cluster are separated by 300–400 nucleotides, the intervening sequence (IVS)[1] separating $J_H$ and $C_H$ is 8–9 kb (12) (our unpublished observation), and the $C_H$ exons, including the transmembrane segment, encompass ~7 kb (15). The nucleotide and predicted amino acid sequences of $C_H$

[1] *Abbreviations used in this paper:* FR, framework region; IVS, intervening sequence.

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/93/09/0815/10 $2.00
Volume 178 September 1993 815–824

segments of different clusters can vary as much as the $V_H$ region (15, 16). Approximately half of the clusters exhibit evidence of germline "joining" of $V_H$ and $D_H$ (V-D) or $V_H$, $D_H$, and $J_H$ (VDJ) elements (13) in nonlymphoid cells. This type of organization is unique to elasmobranchs and extends to a second, non $\mu$-type $C_H$ isotype (17). The present study focuses on the functional significance of the V-$D_1$-$D_2$-J-C form of gene organization. Using gene-specific oligodeoxynucleotide hybridization, the sequences of cDNA clones are related to their corresponding (parental) germline clusters. The goals of these studies were to determine: if there is combinatorial diversity, the nature of junctional diversity and whether a somatic mutation mechanism is present in this modern representative of an early point in vertebrate phylogeny.

## Materials and Methods

*Animals.* Adult specimens of *H. francisci* (horned shark) were obtained from Pacific Biomarine Supply Co. (Venice, CA). After the animals were killed, all tissues were processed immediately for DNA and RNA isolation. The same animal was used to obtain material for the genomic and the cDNA λ libraries.

*DNA Libraries.* The genomic library was described previously (13). The cDNA library was constructed from spleen mRNA using a kit (Pharmacia LKB Biotechnology, Inc., Piscataway, NJ) and cloned into λgt11.

*Subcloning in M13 mp10/11 and DNA Sequencing.* DNA sequencing was performed by the dideoxy chain termination method (18) as described previously (13).

*DNA Sequences in GenBank.* The GenBank accession numbers for the sequences described herein are: Z11776, 1113; Z11777, 1403; Z11778, 33141; Z11779, 33141CON; Z11780, 12215; Z11781, 12271; Z11782, 12272; Z11783, 12214; Z11784, 12241; Z11785, 12221; Z11786, 12273; Z11787, 12251; Z11788, HN13; Z11790, HN84; Z11791, HN21; Z11792, HN72.

*PCR Analysis of Genomic $V_H$ λ Clones.* All $V_H^+$ plaques were selected from the plating of a genomic library (~7 × $10^5$ recombinants equals approximately two to three genome equivalents), pooled, and eluted into 1× SM. The titer of the resulting pool was 5 × $10^4$ PFU/ml. 1 $\mu$l of this material was used as template for PCR amplification. The phage were heated at 70°C for 5 min to disrupt the phage heads, then amplified using 1403-specific primers for CDR2: CCGAGCTCACCATAGTTCATCATACCAGTACTA with an artificial SstI site (underlined), and $J_H$: CCCCTGCAGTCACCGTCACCAT with a natural PstI site (underlined). 30 cycles of amplification were performed using a 55°C annealing temperature. The ~1.4-kb product band was purified by elution from an agarose gel and cloned into M13. 12 of these clones were confirmed as $V_H^+$ and subjected to sequencing.

*PCR Analysis of Genomic DNA.* Direct PCR amplification of genomic DNA was carried out using primers specific for the hyperconserved area of framework region[1] (FR1): 5'-CCGAGCTCTTGACTCAGCCAGAAGCA-3', and contain a 2-nt spacer plus SstI site (underlined), and the λ1403/λ33141-type CDR2: 5'-CCCTGCAGAGGGGCGTAGTACTGGTAT-3', containing a 2-nt spacer plus PstI site (underlined). *Heterodontus* genomic DNA (~500 ng) was denatured for 5 min at 94°C and amplified for 30 cycles consisting of a 1-min denaturation cycle at 94°C, 2-min annealing cycle at 65°C, and 3-min extension cycle at 72°C in a 100-$\mu$l reaction volume. A second round of PCR amplification was carried out using either 1 or 10 $\mu$l of this product. The final product was digested with SstI and PstI, and cloned in M13 mp18/19. Positive clones were identified, single-strand templates prepared, and sequenced.

*Probe Construction and Library Screening.* It was determined empirically using λ $V_H$ clones of known sequence that 33mer probes could both discriminate between closely related sequences and provide a usable signal when screening a library. Probes <33 nucleotides tended to give a high number of false positives as well as high background. Probes >33 nucleotides did not give adequate discrimination between closely related DNA sequences. The 33mer probes used were synthesized using a synthesizer (380B; Applied Biosystems, Inc., Foster City, CA) and were complementary to the most variable positions within CDR2. Hybridization was at 60°C for 6 h in 6× SSC (1× = 0.15 M NaCl, 0.015 M NaCitrate), followed by four washes in 2× SSC, also at 60°C. Hybridization was to duplicate lifts, with a third lift hybridized to a *Heterodontus* $V_H$ probe. Exposure times were 18–36 h. 20 150-mm plates with ~3× $10^4$ plaques on each were used in the genomic library screening. Eight plates with ~2 × $10^4$ plaques/plate were used for cDNA library screenings. Positive plaques were those with signals on all three lifts. These were plaque purified and the specificity of the probe confirmed. All positive clones selected with the λ1403 and λ1113 CDR "specific" probes were closely related when sequenced.

## Results

*$V_H$ Families.* Using a homologous (*Heterodontus* $V_H$) screening strategy employing a single probe, HXIA (19), which is equivalent to the $V_H$ sequence of λ1403 (13), several hundred $V_H^+$ clones were detected in a representative *Heterodontus* genomic λ library prepared from the DNA of a single animal. The complete nucleotide sequence of the $V_H$ through the $J_H$ elements has been determined for 24 clones. With a single exception, the $V_H$ segments of all of the 24 clones are >90% identical at the nucleotide level with HXIA (λ1403) and thus are classified as members (defined as sequences sharing >70% nucleotide identity) of the same Ig gene family, $V_H$I. The exception, λ1113 (13) (and below), hybridizes weakly with the selection probe HXIA and is only ~61% identical at the nucleotide sequence level to the most closely related member of the *Heterodontus* $V_H$I gene family, although it possesses extensive sequence identity in FR3. No other weakly hybridizing clones were noted. The nucleotide sequence of λ1113 from 3' of $V_H$ through $J_H$ has been reported previously (13) and consists of a recombination signal sequence arrangement in which either the $D_H$1 and $D_H$2 segments or the $D_H$1 segment alone can be joined productively, in contrast to $V_H$1 genes in which $D_H$1 and $D_H$2 or $D_H$2 independently can be joined, according to the 12/23 rule of recombination (1). Despite repeated efforts to identify related genes using different genomic (and cDNA) libraries, it appears that the λ1113 cluster represents a monotypic $V_H$ family ($V_H$II). Attempts to detect additional $V_H$ families in several different genomic and cDNA libraries, using both $J_H$ and $C_H$ selection, have been unsuccessful, although it could be argued that unique $V_H$ families might be associated with appreciably different $J_H$ or $C_H$ segments.

*Germline Diversity of $D_H$ and $J_H$ Segments: $C_H$ Variation.* The $D_H$ coding regions of the V-$D_1$-$D_2$-J genes, with

the exception of λ1113, are related very closely at the nucleotide level (Fig. 1). λ1403, the gene cluster that has been examined in greatest detail, differs at one nucleotide in $D_H1$. Two of the λ1403-type genes, λ1315 and λ32004, share one nucleotide difference from λ1403 in $D_2$. These restrictions in sequence variation in the λ1403-like $D_H$'s not only limit germline diversity but also complicate unequivocal identification of the "parental" origin of cDNAs (see below). Of the 24 $J_H$ segments that have been sequenced, essentially all variation is in the 5′ 11 nucleotides (13; see below). This portion typically is eliminated in the process of productive gene rearrangement. As reported previously, Ig gene clusters in *Heterodontus* are associated with unique C regions (16) that potentially can facilitate identification of the "parental" origin of cDNAs.

*Identification of λ1403-related $V_H$-$D_H$-$J_H$ Clusters.* Sequence comparisons indicate that regardless of the high degree of nucleotide (and amino acid sequence) sharing, CDR2 remains the most informative region in terms of distinguishing different $V_H$ gene clusters, with the greatest amount of nucleotide variation being centered at nucleotides 166–174 (see Fig. 2 *a*, below). Under high stringency conditions, 33mer probes complementing this substituted region exhibit relatively little nonspecific hybridization and allow discrimination between known phage clones that differ in the sequences of CDR2s. Genomic clone λ1403, which contains $V_H$, $D_H1$, $D_H2$, $J_H$ (V-$D_1$-$D_2$-J-C configuration), as well as the first two $C_H$ exons, was selected for the initial analyses of somatic variation. To determine the specificity of the oligonucleotide probe strategy as well as to estimate the extent of the λ1403-related gene family, a genomic library (∼$7.2 \times 10^5$ recombinants, corresponding to approximately two to three genomes) was screened. Only three genomic clones (λ33141, λ33181, and λ33132) hybridized with the 1403 CDR2 probe, and the complete nucleotide sequence ($V_H$→$J_H$) was determined for each clone. λ33132 is identical to λ1403 in both coding and noncoding regions; λ33141 and λ33181 are identical to each other and differ from λ1403 by one nucleotide in FR1, one nucleotide in FR3, two nucleotides in CDR2 (Fig. 2, *a* and *b*), and by one nucleotide in $D_H1$ (Fig. 1), but have identical IVSs. The restriction maps of λ1403 and λ33141 differ significantly from each other (data not shown).

```
              DH1                            DH2

 1403     GGT ACA GCg GTG GGT           A TAT ACT GGA TGG
33141     GGT ACA GCA GTG GGT           A TAT ACT GGA TGG
 1315     GGT ACA GCA GTG GGT           A TAT ACT GGg TGG
 2807     GGT ACA GCA GTG GGT           A TAT ACT GGA TGG
32004     GGT ACA GCA GTG GGT           A TAT ACT GGg TGG

 1113     GGT ACt aCA GTG GGTat         t TAT ACT GGg TGG
```

**Figure 1.** Nucleotide sequence differences in $D_H$ segments of V-$D_1$-$D_2$-J-type genes. In V-$D_1$-$D_2$-J-type genes, $D_H1$ and $D_H2$ sequences are shown in arbitrary triplets that are not intended to correspond to codons and are arranged to emphasize sequence identity with λ1403. Differences (including the two additional nucleotides in λ1113 $D_H1$) are in lowercase bold, as is the **g** at position 10 in several $D_H2$ sequences. λ1113 exhibits the highest degree of divergence from the λ1403 prototype.

An alternative strategy for identifying λ1403-like genes also was used. PCR products were generated, using a λ1403/λ33141 CDR2-specific and a $J_H$ primer, from a pool (∼250 clones) of $V_H^+$ λ clones that were identified by cross-hybridization with the HXIA probe under conditions of moderate stringency (19). 12 of these clones were sequenced and four were found to be identical to λ33141, three were identical to λ1403, four were identical to λ33141 in the coding region, but had either 1- or 2-bp differences in the IVSs, and one had a 1-bp deletion in the IVS. The IVS differences most likely represent PCR-induced substitutions, or cloning artifacts, as similar IVS changes have been noted when individual clones are amplified by PCR (our unpublished observations).

To investigate the possibility that other λ1403/λ33141-type genes had not been integrated in the library, genomic DNA was directly amplified using a PCR procedure that used a 5′ FR1 primer complementing a highly conserved region in all sequenced shark $V_H$ genes and a 3′ CDR2 primer. The 3′ end of this latter, CDR2 primer, was complementary to the region unique to these two genes (5′-ATACCAGTACTA[165]-CGCCCCT[183],-3′; see Fig. 2, *a* and *b*), which should prevent the amplification of other $V_H$ segments. After two rounds of amplification, the products were cloned in M13. Over 70 clones were sequenced and found to correspond only to the λ1403 and λ33141 germline genes. Thus, through three separate forms of analysis, it was not possible to detect other genes of the λ1403/λ33141 types that could have given rise to closely related but distinct cDNAs.

*Expression of the λ1403 and λ33141-type Genes.* The spleen is the primary immune organ in elasmobranchs. Approximately $1.6 \times 10^4$ spleen cDNA recombinant plaques were screened with the λ1403–33mer CDR2 probe under the same hybridization wash conditions that were used in screening the genomic library, and 15 strongly hybridizing plaques were detected. 12 of these were sequenced from the 5′ of $V_H$ FR1 through $C_H1$, and eight unique sequences were obtained. In addition a closely related clone, HN13, was recovered from the same *Heterodontus* cDNA library in the course of an unrelated experiment. The $V_H$ segments of the nine sequences are highly homologous to λ1403 (Fig. 2 *a*) and λ33141 (Fig. 2 *b*). Seven cDNA clones have two FR and two CDR nucleotides that are unique to λ1403 and two cDNA clones are λ33141 like.

The $V_H$ segments of cDNA clone 12215 and genomic clone λ1403 are identical (Fig. 2 *a*). The other clones typically exhibit one or two differences in the sequence segment complemented by the CDR2-specific 33mer probe and up to 11 nucleotide differences from λ1403 throughout the $V_H$ region. In the λ1403-related cDNAs, CDR1 and CDR2, which collectively constitute <20% of the $V_H$ region, exhibit 44% of the sequence differences. Overall, there are almost twice as many replacement as there are neutral substitutions. Only replacement substitutions occur in FR3; however, it should be noted that many of these occur near the junctional boundary and may represent changes that are

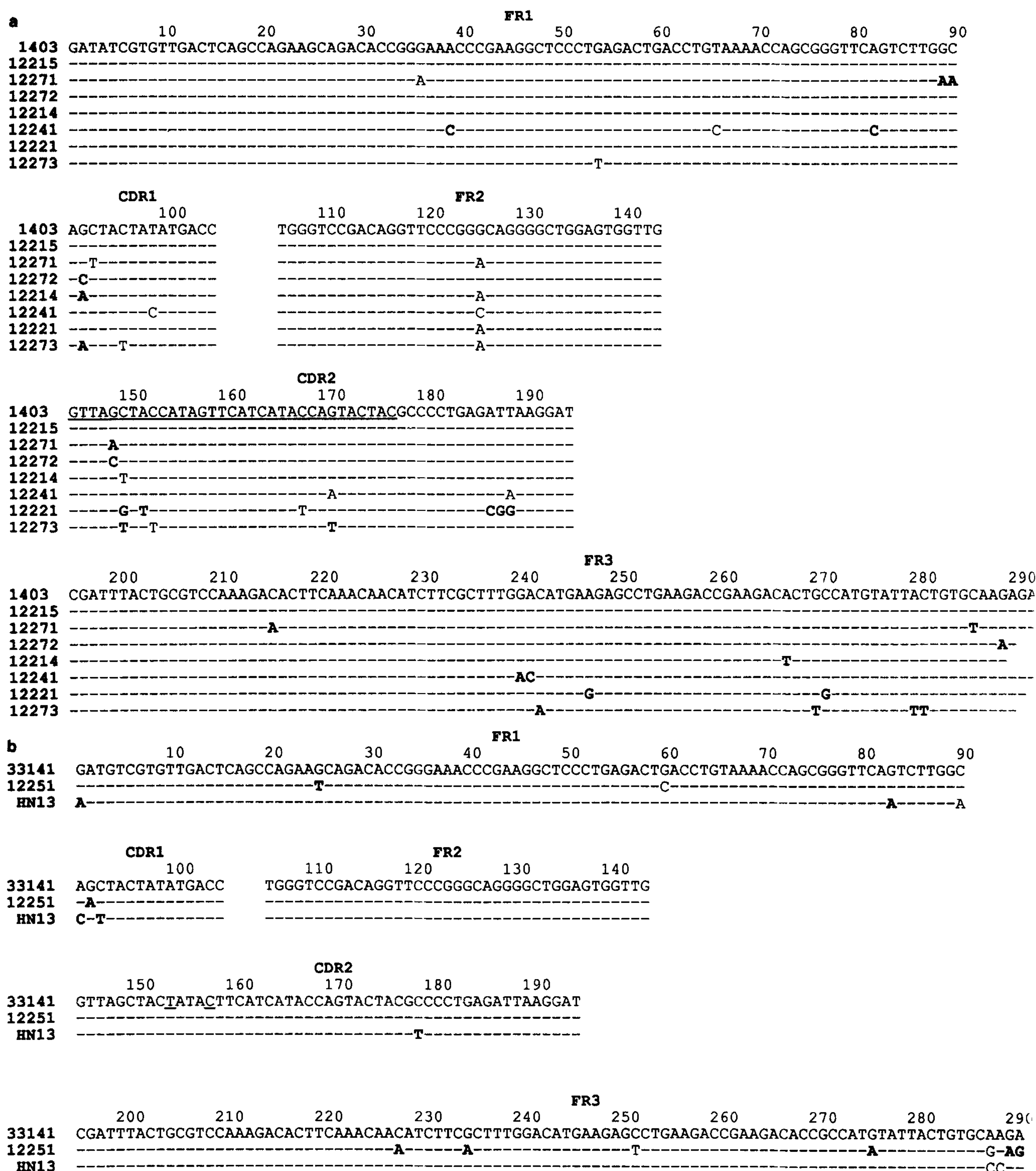

**Figure 2.** (*a*) Nucleotide alignment of the $V_H$ segment of λ1403 with the sequences of seven different cDNAs that represent the putative rearrangement products of this gene cluster. Nucleotide identities are indicated (–); uppercase designates neutral substitutions, and uppercase bold indicates replacement substitutions. For two nucleotide changes within a single codon, both nucleotides are indicated in bold even if one of the changes would not result in a replacement substitution. The 33mer CDR2-"specific" probe used to select the cDNAs is underlined. (*b*) Nucleotide alignment of λ33141 with the sequences of two cDNAs that represent the putative rearrangement products of this germline gene. The two CDR2 nucleotides that distinguish λ33141 from λ1403 are underlined; all other designations are as in *a*.

introduced during functional rearrangements. In three instances, there are two contiguous nucleotide changes and, in one case, there are three nucleotide changes, when compared with the prototype sequence. Nucleotide 125, which is the last base in a SmaI restriction site, is highly polymorphic in the germline of other non-λ1403/λ33141-type clones (K. Hinds-Frey, unpublished observations) as well as in the cDNA sequences, and may represent a mutational hot spot although it does not impart a coding change.

Furthermore, λ33141 is distinguished from λ1403 by two nucleotide substitutions in CDR2 and two FR substitutions. The CDR2 regions of cDNA clone 12251 and genomic clone λ33141 are identical (Fig. 2 *b*), while HN13 shares the distinct nucleotides but differs at a single base from the CDR2 of λ33141. The cDNA clones exhibit several differences at the 3′ of $V_H$, which could represent somatically introduced junctional variation (Fig. 2 *b*; see below). The number of replacement substitutions exceeds neutral substitutions. The relatively few λ33141-type cDNAs that were recovered limit further interpretation of these data.

*Inter- or Intracluster Rearrangement of D.* The nucleotide sequences of the putative $D_H1$ and $D_H2$ regions of the λ1403 33mer-selected cDNAs are shown in Fig. 3. As illustrated in Fig. 1, $D_H1$ and $D_H2$ regions of the V-$D_1$-$D_2$-J genes characterized thus far are nearly identical; the $D_H1$ segment of λ1403 differs from the other $D_H1$ sequences by a single base substitution. This base difference is present in all seven of the λ1403-type genes but is not found in either of the λ33141-type genes, suggesting that the rearrangement of $V_H$ and $D_H1$ occurred within the λ1403 cluster (and that an equivalent intracluster rearrangement took place in λ33141). The $D_H2$ segments of λ1403 and λ33141 are identical to each other, and to other genes, precluding distinction of cluster-restricted $D_H2$ rearrangements.

The putative $D_H1$ regions of all of the cDNAs possess a core of bases that are identical to the germline genes. The core segments of clones 12241 and 12271 exhibit single base differences from germline $D_H1$ that can be accounted for by point mutations. Although unlikely, these cDNAs may have been transcribed from germline clusters that are refractory to genomic cloning, but in all other regards are identical to λ1403 (see discussion below). Although somatic reorganization of *Heterodontus* Ig gene segments of the λ1403 type ($V_{23}$-$12D_123$-$12D_212$-23J) can exclude $D_H1$, according to the 12/23 rule, from 4 to 10 of the 15 germline-encoded $D_H1$ nucleotides are detected in the various cDNAs. It is notable that cDNA 12215, which exhibits the least change from the parental cluster in the $V_H$ segment, also has the most extensive contribution from germline $D_H1$ and $D_H2$ segments.

*D Region Diversity Is Generated by Extensive Junctional Modification.* The overall lengths of the $D_H$ regions in these cDNAs range from 14 to 36 nucleotides, compared with the 28 nucleotides encoded in the genomic $D_H$ segments, assuming that appropriate junctional boundaries have been assigned. As these cDNAs presumably represent the rearrangement products of the λ1403 and λ33141 genomic clones, it can be estimated that from 1 to 10 nucleotide additions can occur at each of the various junctions. In this series, ∼65% of the junctional additions are G/Cs and at least one nucleotide addition occurs in at least one of the joining junctions, $V_H/D_H1$, $D_H1/D_H2$, or $D_H2/J_H$, in all of the cDNAs. In three cDNAs, nucleotide additions occur at all three junctions. Nucleotide deletion followed by N region additions (2) are associated with the rearrangements of $V_H$ genes in *Heterodontus*. It appears that germline $D_H$ segments are used in all three reading frames, although the extensive diversity generated by the deletion and addition of bases at the junctions during segmental rearrangement compromises unequivocal alignment of germline $D_H$ segments and the specific cDNAs (data not shown).

*$J_H$ and $C_H$ Show Evidence for Intracluster Rearrangement.* The complete (including 5′) sequences of the $J_H$ segments of λ1403 and λ33141 are identical to each other but are unique at several positions from the other $J_H$ segments characterized in the course of these (Fig. 4) and earlier studies (13). The sequences of the 5′ $J_H$ segments of λ1403- and λ33141-selected cDNAs, with the exception of a single difference in cDNA 12214 and a three nucleotide substitution in 12272,

| CLONE | $V_H/D_H1$ | $D_H1$ | $D_H1/D_H2$ | $D_H2$ | $D_H2/J_H$ |
|---|---|---|---|---|---|
| 1403 | | **GGTACAGCgGTGGGT** | | **ATATACTGGATGG** | |
| 12215 | CGTA | **ACAGCgGTGG** | CGA | **ATACTGGAT** | ATTCGGCGCG |
| 12271 | CGGGCGC | **GTACAcCg** | | **TACTGG** | G |
| 12272 | AA | **GTACAGCgG** | CG | **GGAT** | ACAACAA |
| 12214 | TGGT | **GCgG** | ACC | **GAT** | |
| 12241 | TGGG | **GCgaTGG** | ACA | **ATACTGGA** | |
| 12221 | | **GCgGTGGGT** | | **CTGG** | GA |
| 12273 | A | **GTACAGCgGT** | CC | **ATAT** | CGG |
| | | | | | |
| 12251 | CGGGCA | **ACAGCAG** | GG | **CTGGATGG** | |
| HN13 | T | **GGTACAGCAG** | | **AT** | T |
| 33141 | | **GGTACAGCAGTGGGT** | | **ATATACTGGATGG** | |

**Figure 3.** Analysis of germline $D_H$ regions and corresponding cDNA sequences. Alignments of λ1403 and λ33141 germline $D_H1$ and $D_H2$, and corresponding cDNA sequences, are based on absolute sequence identity. $D_H1$, $D_H2$, and identical cDNA nucleotides are in upper case bold. The nucleotide difference in λ1403 $D_H1$ that is shared with the seven cDNAs is in lower case bold. Additions (not accounted for in germline sequences) at $V_H/D_H1$; $D_H1/D_H2$; and $D_H2/J_H$ junctions are indicated in upper case. The two differences between the germline $D_H1$ segments and their corresponding cDNAs are in lower case. The most extensive ($D_H1$ and $D_H2$) contiguous contribution is in cDNA 12215. The **AT** dinucleotide in the $D_H2$ homology alignment of HN13 has alternative placements or may represent a nucleotide addition that arose in junctional joining. In 12271, an alignment is suggested in which the $D_H1$ contribution is maximized; however, it is noted that the sequence GCG at the 3′ of the $V_H/D_H1$ boundary is identical with the GCG $D_H1$ core that has been preserved by all other cDNAs.

```
 1403  ACTATCTCGACTACTGGGGACAAGGGACCATGGTGACGGTGACTGCA
12215      -------------------------------------------
12271    ---------------------------------------------
12272       ----CTC-----------------------------------
12214  ----C------------------------------------------
12241       ------------------------------------------
12221     ------------------------------------------A-
12273   --------------------------------------------T-

12251  ----------------------------------------A-G----
 HN13   ----------------------------------------------
33141  ACTATCTCGACTACTGGGGACAAGGGACCATGGTGACGGTGACTGCA
```

**Figure 4.** Nucleotide sequence identities in the $J_H$ region of cDNAs that derive from germline clusters λ1403 and λ33141. The nucleotides that appear to be specific for the λ1403 and λ33141 (λ1403-type) sequences are underlined. Sequence identities are shown (–). Neutral substitutions are indicated in upper case and replacement substitutions are in bold.

are identical to the sequences of λ1403 and λ33141 genes. This finding is consistent with intracluster joining of $J_H$ segmental elements, as has been suggested above for $V_H$ and $D_H1$ joining. The single difference in clone 12214 and the differences at the 3′ of several $J_H$s of the λ1403 and λ33141 types most likely represent somatic substitutions.

The $C_H1$ exons of *Heterodontus* Ig genes differ significantly from each other at the 5′ (16). Lesser degrees of variation occur in the other $C_H$ exons (15). The sequences of $C_H1$ of λ1403 and λ33141 differ at 10 positions from the $C_H1$ consensus sequence of the ∼25 germline $C_H$ genes that have been characterized to date (15, 16; and our unpublished observations) but are identical to each other. The $C_H1$ sequences of all nine cDNAs are identical to the corresponding germline λ1403/33141 sequence (not shown), indicating that the $J_H/C_H1$ splice is of an intra- vs. intercluster nature. In addition to establishing that rearrangements are intra- vs. intercluster, the complete sequence identity between germline and cDNA $C_H1$ sequences indicates that variation primarily is restricted to the V, D, and J regions of the gene, a finding entirely consistent with the V region–restricted nature of somatic mutation.

*λ1113-restricted Clusters and Expression of the λ1113-type Gene.* As indicated above, λ1113 is a unique (presumably monotypic) gene family in which an inverted $D_H1$-$D_H2$ IVS changes the orientation of the functional recombination signal sequence spacers, i.e., V23-12$D_1$12-23$D_2$12-23J (λ1113-type) vs. V23-12$D_1$23-12$D_2$12-23J (λ1403-type gene). In addition, λ1113 differs from all other V-$D_1$-$D_2$-J-type genes characterized to date by a three-nucleotide insertion in CDR2 (Fig.

```
a    FR1                                                                                       CDR1
             10        20        30        40        50        60        70        80        90       100
1113 GATATTGTGCTGACCCAGTCAGCTTCGGTGACGAAAAGCCCGGAGAATCTCACAAGCTTTCCTGCTCCGTGAGCGGATTCAGTCTGGACAGCTACTATGT
HN84 ----------------------------------------------------------------------------------------------------
HN21 ------------------------------------------------------------------------------------------A---------
HN72 ---------------------A-----------------------------------A--------------------------------A---------

          FR2                                    CDR2
             110       120       130       140        150       160       170       180       190
1113 GCAC TGGGTTAAGCAAGCTCCAGGGAAGGGGCTGGAGTGGCTC CTTGCTTACAGAAAACCATCGGATACCATCTATTACGCGCCTGGAGTCCAAGGT
HN84 ---- --------------------------------------- ------------------------------------------------------
HN21 ---- --------------------------------------- ------------------C-----------T-----------------------
HN72 ---- --------------------------------------- ------------------------------------------------------

     FR3
      200       210       220       230       240       250       260       270       280      288
1113 CGAATCATTCCTTCCAGAGCCTCTTCCACAACCTATATTGAGATCAAAAGTCTGAGATCCGAAGACACCGCCATGTATTACTGTGCAAGA
HN84 ------------------------------------------------------------------------------------------
HN21 ---------------------------------------------------------------------------------------
HN72 --------------------------------------------------------------------------------T--------G

b    V/D1           D1                          D1/D2                      D2               D2/J

1113                GGTACtaCAGTGGGTat                                     tTATACTGGgTGG
HN84 AAAGC                    TGG                                               TGGgT            TTA
HN21 GAAATAC        GGTACta                     TGATGGATA                      TGGg             GC
HN72                GGTAC   AGTGGGTat             ATCTCC                         GGgTGG           CG

     J
1113 ACATCATGGACTACTGGGGACAAGGGACCATGGTGACGGTGACTGCA
HN84       -----------------------------------------
HN21  --C-----------------------------------------T--
HN72       ---------------------------------------T-
```

**Figure 5.** (*a*) Nucleotide alignment of λ1113 with the sequences of three different cDNAs that represent the putative rearrangement products of this gene cluster. For the $V_H$ and $J_H$ segments, nucleotide identities are indicated (–); upper case designates neutral substitutions and upper case bold indicates replacement substitutions. The probe sequence used to select the cDNAs is underlined. (*b*) $D_H1$, $D_H2$, and corresponding cDNA nucleotides are in upper case bold, except for nucleotides that appear to be unique to the λ1113 gene (see Fig. 1), which are shown in lower case bold, as is the **g** at position 10 of $D_H2$ that is found in half of the $D_H2$ segments compared in Fig. 1, including λ1113; additional nucleotides at the $V_H/D_H1$, $D_H1/D_H2$, and $D_H2/J_H$ boundaries are indicated in upper case bold. Potential homologies in these regions are indicated by inspection of vertical alignments.

1) and a six-nucleotide deletion in FR3 (Fig. 5). The $D_H1$ segment of λ1113 (Fig. 1) differs at two positions and also is two nucleotides longer than the $D_H1$ segments of the other V-$D_1$-$D_2$-J genes characterized thus far. The $D_H2$ differs by a single 5′ nucleotide from the V-$D_1$-$D_2$-J-type gene (Fig. 1); however, this potentially informative base difference is not present in the three cDNAs. The $J_H$s cannot be distinguished readily from other V-$D_1$-$D_2$-J-type genes. The $C_H1$ exon of λ1113 contains a three-nucleotide addition and 18 additional nucleotide changes that distinguish it from the $C_H1$ exons of all other V-$D_1$-$D_2$-J genes characterized to date.

A 33mer probe specific for the CDR2 of λ1113 was used to screen a genomic λ library as described above for λ1403. Only three positive clones were identified in this screening and these had identical or overlapping restriction maps as well as identical V-$D_1$-$D_2$-J sequences, including IVSs.

The 33mer probe then was used to select λ1113-like cDNAs. Only three positive clones (HN21, HN72, and HN84) were identified (Fig. 5). All three cDNAs possess the three additional λ1113-type nucleotides in CDR2 and the six nucleotide deletion in FR3 identified in λ1113, compelling evidence that these cDNAs derived from λ1113 or a closely related (indistinguishable) gene (Fig. 5 *a*). HN84 is identical to λ1113 in all coding regions; the other two cDNAs exhibit limited differences. With all three λ1113-type cDNAs, junctional diversity resembles that seen with the λ1403- and 33141-type genes (Fig. 5 *b*). The contribution of $D_H1$ in HN72 is considerable, although, the nature of the TAC deletion (by homology to λ1113 $D_H1$) in the D region of HN72 is uncertain (similar effects have been noted previously [20]). As indicated in Fig. 1, the $D_H1$ region of λ1113 has distinct nucleotides ($T^6A^7$), as well as two additional nucleotides ($A^{16}T^{17}$) that are not seen in other $D_H1$ segments. The two unique nucleotides, $T^6A^7$, are present in HN21. The absence of these in HN72 and HN84 is uninformative owing to extensive $D_H1$ deletion (HN21) and the TAC deletion (HN72) indicated above. HN72 contains two nucleotides that correspond to $A^{16}T^{17}$ of λ 1113. All three cDNAs share the $G^{10}$ identified in the $D_H2$ of λ1113 (Fig. 1), further supporting the conclusion that these elements originate in the λ1113 cluster. As in the case with the λ1403- and λ33141-type gene clusters, germline $D_H$ sequences in λ1113 transcripts are used in all three reading frames (not shown).

The nucleotide sequence of the $J_H$ region of HN84 is identical to the corresponding sequence of λ1113. Both HN21 and HN72 exhibit replacement substitutions at the 3′ of $J_H$, similar to findings for $J_H$ regions in several λ1403- and λ33141-type genes (Fig. 5 *b*). HN21 differs from the putative parental prototype sequence at an additional 5′ position. Because the sequence of the $J_H$ segment of λ1113 does not vary significantly from the $J_H$ consensus sequence, it is not possible, in this case, to provide unequivocal evidence for intra- vs. intercluster $J_H$ use. The sequences of the $C_H1$ exons of cDNA HN21, HN72, and HN84 are identical to the unique (see above) corresponding sequence of λ1113 (not shown). Intracluster rearrangement of elasmobranch $V_H$ genes is entirely consistent with these observations.

## Discussion

*Heterodontus* represents the most phylogenetically distant form in which information is available about antibody gene structure and organization. A remarkable degree of sequence homology exists between the coding segments and recombination signal sequence elements of Ig genes in *Heterodontus* and all vertebrate species characterized to date (13, 15, 19). Although strong $V_H$ homology and characteristic μ-type heavy chain organization and regulation resemble mammalian Ig, close linkage of segmental elements (cluster organization), presence of both $D_H1$ and $D_H2$ elements, and absence of an invariant Ig gene promoter are characteristic of TCRs. The dramatic changes in gene organization that have occurred during vertebrate phylogeny raise important questions about the functional regulation of Ig genes that possess alternative patterns of segmental arrangement as well as the somatic mechanisms that expand germline Ig gene repertoires.

In addition to the unique gene organization, the Ig heavy chain gene system of *Heterodontus* is distinguished from that of higher vertebrates by a high degree of identity in the individual members of the primary $V_H$ family ($V_H1$) as well as by the extremely high degree of sequence identity in both the $D_H$ and $J_H$ segmental elements, as indicated above. At the nucleotide level, the individual members of the $V_H1$ family are >90% related, in marked contrast to the extensively diversified $V_H$ families in mammals (3, 4, 21) and the amphibian *Xenopus laevis* (22, 23). As there probably are ∼200 gene clusters and ∼50% of these are joined in the germline, the entire repertoire of *Heterodontus* heavy chain genes apparently arises from ∼100 different clusters containing closely related $V_H$ genes. Presumably, this restricts the potential of the antibody repertoire. The only non-$V_H1$ gene cluster, λ1113, would contribute only minimally to $V_H$ diversity. As indicated above, this gene most likely arose through inversion of the $D_1$-$D_2$ IVS (13), an effect that has been observed in vitro (24) and has been proposed to represent a significant component in the evolution of Ig genes (13, 24). The inversion may relieve this particular cluster of the gene correction effect that we suggest could account for the high degree of sequence identity in both coding and noncoding regions, and indeed, the IVS regions of this gene are related only distantly to other $V_H$ genes (13).

To address the role of somatic mechanisms in the generation of Ig diversity, a strategy has been devised that permits the identification of the rearrangement products of specific gene clusters. In these experiments, CDR2-specific probes have been used both to select cDNAs related to a specific germline gene cluster and to determine how many other candidate (parental) gene clusters are present in the germline. Regarding the latter, identification of all the individual members of a multigene family is confounded by the requirement that the relevant gene sequences be clonable and that the specific replication of the clones containing these genes (sequences) is not compromised relative to the replication of clones containing other genes. Because of these uncertainties, which are not easily amenable to experimental resolution, it is difficult to establish that all members of a germline

gene family have been detected. While approaches such as PCR amplification of *Heterodontus* genomic DNA (complete clusters) and direct sequencing of PCR-generated templates are attractive, it is not possible to efficiently derive V-$D_1$-$D_2$-J sequences due to the near identity of joined VDJ and VD-J genes, which possesses shorter IVS sequences and are amplified to the exclusion of the V-$D_1$-$D_2$-J forms (K. Hinds-Frey, unpublished observations). Nevertheless: (*a*) the detection in this study of identical λ1403 genes in the libraries of two different individuals; (*b*) the isolation of the closely related λ33141 gene, which may represent an allelic form; (*c*) the inability to detect additional, related gene clusters by screening of conventional and PCR-generated (FR1-CDR2) libraries; (*d*) the inability to detect additional λ1403/λ33141-type genes by direct amplication of FR1-CDR2 sequences; and (*e*) the high degree of concordance between germline (V, $D_1$, $D_2$, J, and $C_H$) and cDNA sequences with λ1403, λ33141, the monotypic λ1113, and a fourth, unique cluster (λ2807; results not shown) are consistent with the identification of the rearrangement products of an individual gene cluster. Furthermore, the findings that the $C_H1$ sequences of the λ1403/λ33141 and the λ1113 cDNAs are identical with their respective germline counterparts would be inconsistent with the origin of these transcripts from other clusters. While hypotheses could be invoked that would explain all somatic changes as having arisen from as yet undetected gene clusters that also cannot be amplified by direct PCR, such claims are difficult to substantiate based on these observations. Finally, such hypotheses would be entirely inconsistent with the various estimates of the number of gene clusters in this species, some of which we based on direct, quantitative Southern blot analysis of C region exons (13, 16).

The unique sequence characteristics of $D_H1$'s, the 5' $J_H$ segments, as well as the $C_H1$ exons of genes λ1403 and λ33141, relative to the other genes that have been described, provide a means whereby intra- vs. intercluster gene rearrangements potentially can be distinguished. It is most likely that rearrangements occur within clusters, as is suggested from the analyses of $V_H/D_H1$, $V_H/J_H$, and $V_H/C_H$ association patterns. This restriction may relate to the remarkably close linkage of V-$D_1$-$D_2$-J (300–350 nucleotides separate the segmental elements), and is consistent with the presence of $V_H$ clusters (of uncharacterized organization) on multiple chromosomes in *Heterodontus* (Amemiya et al., unpublished observations). Whereas spatial proximity is not an obligatory requirement of segmental rearrangements, certain components of the joining process are influenced by the positions of the recombining elements, e.g., close proximity may be the primary factor in the selective joining of $D_H$ and $J_H$ segments (2). The intersegmental distances of the *Heterodontus* elements are several orders of magnitude less than the minimal linkage distances (most proximal $V_H$, $D_H$, and $J_H$ elements) in the mammalian Ig gene loci and may drive recombination of immediately adjacent elements. In terms of the evolution of this highly diversified multigene family, selective pressures may operate on the entire cluster vs. individual elements.

Despite the near homogeneity in the sequences of germline $D_H$ segments, the high degree of diversity in the D regions of cDNAs that derive from λ1403 (and λ33141) is remarkable. The D region, which includes both $D_H$ and N additions, ranges from 5 to 13 amino acids; however, no D region sequence in these comparison groups shares more than three amino acids with any other cDNA and typically show only two amino acids in common. An average of six unique amino acids (range, 4–9) are introduced in the D regions of the seven cDNAs in the λ1403 comparison group, whereas the V segments of these same genes, which are 10 times longer, average only four replacements, with a range of 0–7 changes. In λ1403-type genes, it is neither germline diversity nor somatic mutation but D region variation that contributes extensive sequence diversity. This diversity arises from junctional N diversity and use of $D_H$ sequences in all three reading frames. These findings in a lower vertebrate are in marked contrast to the human $D_H$ region, in which ~30 germline $D_H$ gene segments have been characterized (25–27). There is no $D_H$ consensus sequence, and these D's range in length from 11 to 44 nucleotides (23). The murine D locus is similarly diversified (23); however, two groups of three identical D segments are present among the 15 avian $D_H$ elements that have been described previously (28).

Without exception, all relevant germline cDNA comparisons suggest that somatic mutation or a highly restricted form of gene conversion is giving rise to the observed sequence variation. Mutation also is observed in $D_H$ and $J_H$ segments; however, no mutations (changes) are noted in the $C_H$ segments of any of the cDNAs in this study. Thus, the $C_H1$ exon, which typically is localized 8–9 kb from $J_H$, does not appear to be targeted for mutation, consistent with the previously described restriction of the Ig mutational (hypermutation) mechanism to the V region, and immediately flanking sequences (29, 30).

The existence of a somatic mutation mechanism seems contrary to reports that the hapten-specific immune response of elasmobranchs is not associated with affinity maturation or interindividual variation (31), even after prolonged periods of antigenic stimulation (32). As indicated above, the cDNAs were not selected as components of an antigen-driven selection process. Rather, cDNAs in these unimmunized animals are being related to parental clusters, without any knowledge of their specificity. The only real requirement for cDNA detection was imposed at the screening stage and would eliminate from study only those clones in which the CDR2 had been so mutated as to render the clones unreactive with the CDR2-specific oligonucleotide probe. While presumed mutation effects are apparent and occur preferentially in CDRs, caution is warranted against any conclusions that such a process either alters primary antibody recognition or occurs with sufficient frequency to effectively alter the specificity and/or affinity of the antibody response. The absence of germinal centers in this (and other lower vertebrate) species may result in a failure to select somatic mutations. Alternatively, this (and other lower vertebrate) species may use other mechanisms of selection.

The sequence of joining and mutational events in the *Het-*

*erodontus* system is unknown. No sterile transcripts or cDNAs without an open reading frame were detected in the course of these studies. In addition, because it is not possible to follow somatic changes in a successive, chronological manner, it is impossible to say when the putative mutations may have occurred, and, indeed, mutation may have occurred before joining. The existence of two and three contiguous mutations, observed in these studies, has been reported in antigen-driven selection systems (33, 34). None of the cDNAs exhibited >4% variation from the prototype gene, a figure reported as the upper range of somatic mutation (29, 33). However, this type of comparison is complicated by the >90% identity between *Heterodontus* $V_H$ genes.

*Heterodontus* may represent the most phylogenetically primitive system in which a true counterpart of the rearranging Ig gene system is present. The central issue in this investigation was to determine what role, if any, various somatic processes have in the adaptive immune response in a species representative of this primitive level of evolutionary development. The presence of only minimally diversified D coding segments suggests strongly that the D-mediated recombination process was an early event in the evolution of adaptive immunity. Facilitated joining rather than the actual genetic diversification of coding sequences most likely represents an earlier stage in the evolution of Ig gene diversity. Furthermore, the detection of extensive somatic variation and junctional diversity, with an apparent restriction to intracluster gene rearrangements, indicates that somatic variation preceded combinatorial diversity and extensive diversification of V gene families in the evolutionary acquisition of Ig diversity.

We are grateful to Dr. L. Du Pasquier for critical reading of this manuscript and valuable suggestions and thank Barbara Pryor for editorial assistance.

This work was supported by National Institutes of Health grant AI-23338. Dr. G. W. Litman is the Hines Professor at the University of South Florida College of Medicine, All Children's Hospital.

Address correspondence to G. W. Litman, Department of Pediatrics, University of South Florida, All Children's Hospital, 801 Sixth Street South, St. Petersburg, FL 33701.

*Received for publication 3 March 1993 and in revised form 28 May 1993.*

## References

1. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature (Lond.).* 302:575.
2. Blackwell, T.K., and F.W. Alt. 1989. Mechanism and developmental program of immunoglobulin gene rearrangement in mammals. *Annu. Rev. Genet.* 23:605.
3. Brodeur, P.H., G.F. Osman, J.J. Mackle, and T.M. Lalor. 1988. The organization of the mouse IgH-V locus: dispersion, interspersion, and the evolution of $V_H$ gene family clusters. *J. Exp. Med.* 168:2261.
4. Berman, J.E., S.J. Mellis, R. Pollock, C.L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C.R. Cantor, and F.W. Alt. 1988. Content and organization of the human Ig $V_H$ locus: definition of three new $V_H$ families and linkage to the Ig $C_H$ locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:727.
5. Matsuda, F., E.K. Shin, H. Nagoaka, R. Matsumura, M. Haino, Y. Fukita, S. Taka-ishi, T. Imai, J.H. Riley, R. Anand, E. Soeda, and T. Honjo. 1993. Structure and physical mapping of 64 variable segments in the 3' 0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nature Genetics.* 3:88.
6. Oettinger, M.A., D.G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science (Wash. DC).* 248:1517.
7. Lieber, M.R. 1991. Site-specific recombination in the immune system. *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 5:2934.
8. Becker, R.S., and K.L. Knight. 1990. Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. *Cell.* 63:987.
9. Reynaud, C.-A., A. Dahan, V. Anquez, and J.-C. Weill. 1989. Somatic hyperconversion diversifies the single $V_H$ gene of the chicken with a high incidence in the D region. *Cell.* 59:171.
10. Reynaud, C.-A., V. Anquez, H. Grimal, and J.-C. Weill. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell.* 48:379.
11. McCormack, W.T., L.M. Carlson, L.W. Tjoelker, and C.B. Thompson. 1989. Evolutionary comparison of the avian IgL locus: combinatorial diversity plays a role in the generation of the antibody repertoire in some avian species. *Int. Immunol.* 1:332.
12. Hinds, K.R., and G.W. Litman. 1986. Major reorganization of immunoglobulin $V_H$ segmental elements during vertebrate evolution. *Nature (Lond.).* 320:546.
13. Kokubu, F., R. Litman, M.J. Shamblott, K. Hinds, and G.W. Litman. 1988. Diverse organization of immunoglobulin $V_H$ gene loci in a primitive vertebrate. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:3413.
14. Shamblott, M.J., and G.W. Litman. 1989. Genomic organization and sequences of immunoglobulin light chain genes in a primitive vertebrate suggest coevolution of immunoglobulin gene organization. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3733.
15. Kokubu, F., K. Hinds, R. Litman, M.J. Shamblott, and G.W. Litman. 1988. Complete structure and organization of immunoglobulin heavy chain constant region genes in a phylogenetically primitive vertebrate. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1979.
16. Kokubu, F., K. Hinds, R. Litman, M.J. Shamblott, and G.W.

Litman. 1987. Extensive families of constant region genes in a phylogenetically primitive vertebrate indicate an additional level of immunoglobulin complexity. *Proc. Natl. Acad. Sci. USA.* 84:5868.
17. Harding, F.A., C.T. Amemiya, R.T. Litman, N. Cohen, and G.W. Litman. 1990. Two distinct immunoglobulin heavy chain isotypes in a primitive, cartilaginous fish, Raja erinacea. *Nucleic Acids Res.* 18:6369.
18. Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
19. Litman, G.W., L. Berger, K. Murphy, R. Litman, K.R. Hinds, and B.W. Erickson. 1985. Immunoglobulin $V_H$ gene structure and diversity in *Heterodontus*, a phylogenetically primitive shark. *Proc. Natl. Acad. Sci. USA.* 82:2082.
20. Sanz, I. 1991. Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions. *J. Immunol.* 147:1720.
21. Kodaira, M., T. Kinashi, I. Umemura, F. Matsuda, T. Noma, Y. Ono, and T. Honjo. 1986. Organization and evolution of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* 190:529.
22. Hsu, E., J. Schwager, and F.W. Alt. 1989. Evolution of immunoglobulin genes: $V_H$ families in the amphibian *Xenopus*. *Proc. Natl. Acad. Sci. USA.* 86:8010.
23. Haire, R.N., C.T. Amemiya, D. Suzuki, and G.W. Litman. 1990. A high degree of immunoglobulin $V_H$ gene complexity in a lower vertebrate: *Xenopus laevis*. *J. Exp. Med.* 171:1721.
24. Lewis, S.M., J.E. Hesse, K. Mizuuchi, and M. Gellert. 1988. Novel strand exchanges in V(D)J recombination. *Cell.* 55:1099.
25. Siebenlist, U., J.V. Ravetch, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Human immunoglobulin D segments encode in tandem multigenic families. *Nature (Lond.).* 294:631.
26. Ichihara, Y., M. Abe, H. Yasuli, H. Matsuoka, and Y. Kurosawa. 1988. At least five $D_H$ genes of human immunoglobulin heavy chains are encoded in 9-kilobase DNA fragments. *Eur. J. Immunol.* 18:649.
27. Ichihara, Y., H. Matsuoka, and Y. Kurosawa. 1988. Organization of human immunoglobulin heavy chain diversity gene loci. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:4141.
28. Reynaud, C.-A., V. Anquez, and J.-C. Weill. 1991. The chicken D locus and its contribution to the immunoglobulin heavy chain repertoire. *Eur. J. Immunol.* 21:2661.
29. Kim, S., M. Davis, E. Sinn, P. Patten, and L. Hood. 1981. Antibody diversity: Somatic hypermutation of rearranged $V_H$ genes. *Cell.* 27:573.
30. Lebecque, S.G., and P.J. Gearhart. 1990. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5′ boundary is near the promoter, and 3′ boundary is ~1 kb from V(D)J gene. *J. Exp. Med.* 172:1717.
31. Mäkelä, O., and G.W. Litman. 1980. Lack of heterogeneity in anti-hapten antibodies of a phylogenetically primitive shark. *Nature (Lond.).* 287:639.
32. Litman, G.W., J. Stolen, H.O. Sarvas, and O. Mäkelä. 1982. The range and fine specificity of the anti-hapten immune response: phylogenetic studies. *J. Immunogenet.* 9:465.
33. Levy, N.S., U.V. Mallpiero, S.G. Lebecque, and P.J. Gearhart. 1989. Early onset of somatic mutation in immunoglobulin $V_H$ genes during the primary immune response. *J. Exp. Med.* 169:2007.
34. Jacob, J., G. Kelsoe, K. Rajewsky, and U. Weiss. 1991. Intraclonal generation of antibody mutants in germinal centres. *Nature (Lond.).* 354:389.