Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 5868–5872, August 1987
Genetics

Extensive families of constant region genes in a phylogenetically primitive vertebrate indicate an additional level of immunoglobulin complexity

F. KOKUBU, K. HINDS, R. LITMAN, M. J. SHAMBLOTT, AND G. W. LITMAN*

Showa University Research Institute, 10900 Roosevelt Boulevard, St. Petersburg, FL 33702

Communicated by Lewis Thomas, May 13, 1987

ABSTRACT **A homologous probe for the constant region of the *Heterodontus francisci* (horned shark) immunoglobulin heavy chain was used to screen a genomic DNA library constructed in bacteriophage λ, and a large number of independent clones were recovered. Their hybridization patterns with segment-specific probes are consistent with the close linkage of heavy-chain constant (C_H), joining (J_H), and variable (V_H) gene segments. Differences in the nucleotide sequences of the first C_H exon of five genes primarily are localized to 5′ positions; extended regions of sequence identity are noted at 3′ positions. The predicted amino acid sequences of each gene are different and are related distantly to the corresponding regions of higher vertebrate immunoglobulins. Gene-specific oligodeoxynucleotide probes were used to establish that at least three of the five genes are transcriptionally active. Quantitative gene titration data are consistent with the large numbers of genes suggested by the library screening analyses. In this representative early vertebrate, it appears that (V_H-diversity-J_H) segments are associated with individual constant region genes that can differ at the predicted protein level.**

The apparent sequence relationships between the various gene families encoding the antigen-binding receptors in mammals have led to considerable speculation concerning their presumed common origin and paths of evolutionary diversification (1, 2). The high degree of relatedness between immunoglobulin genes found in mammals and species representing early periods in the vertebrate radiations has facilitated the isolation and characterization of homologous genes in phylogenetically distant forms (3). Immunoglobulin heavy-chain variable region (V_H) gene families that are as, if not more, complex than those found in higher vertebrates have been encountered (4, 5). Restrictions in antibody diversity observed in lower vertebrates most likely are not related to a paucity of functional genes.

Higher and lower vertebrate V_H, heavy-chain diversity region (D_H), and heavy-chain joining region (J_H) gene segments are remarkably similar. In mammals, segmental elements are linked tandemly over an extended chromosomal region, i.e., $(V_H)_n$-$(D_H)_n$-$(J_H)_n$-$(C_H)_n$, where C_H is the heavy-chain constant region and n varies from 4 to as many as 1000 (6, 7). The genes encoding the constant region are located several million base pairs downstream of the variable gene complex. In *Heterodontus*, a modern representative of the ancient elasmobranchs, individual elements are closely linked with constant regions in multiple ≈10-kilobase "clusters," i.e., $(V_H$-D_H-J_H-$C_H)_n$ (8). This difference in gene organization raises a number of questions concerning the relationship between gene linkage and segmental rearrangement as well as the nature of the extensive constant region gene family that is suggested by the *Heterodontus* model. Studies reported here indicate that different constant region genes indeed are associated with individual (V_H-D_H-J_H)-containing clusters, indicating an additional level of immunoglobulin gene complexity.

MATERIALS AND METHODS

Identification and Isolation of C_H Segments. The procedures for screening a *Heterodontus* genomic DNA library constructed in λ47.1 under conditions of relaxed stringency have been described (5). A 413-base-pair fragment that includes C_H exon 1 as well as 81 base pairs of 5′ and 35 base pairs of 3′ additional sequence was derived by *Alu* I cleavage of a *Pst* I fragment of λ801 (8). This segment was found to be devoid of repeat sequences and was subcloned in pUC13. The insert was purified by polyacrylamide gel electrophoresis and was nick-translated to 2×10^8 cpm/μg. At the initial screening stages, nitrocellulose filters (Schleicher & Schuell) were exposed for 96 hr at −70°C using a Dupont Lightning Plus intensifying screen. Purification of positive λ phage employed minor modifications of established methods.

Southern Blotting, Restriction Mapping, RNA Isolation, Subcloning, and DNA Sequencing. These methods were carried out as described (5, 8). The C_H exon 1 probe described above and V_H- and J_H-specific probes described (8) were used in Southern blot analysis (22). RNA gel blots were carried out using 5.1 cm × 7.6 cm × 4 mm gels; transfer time and volume were limited to 2 hr and 12 ml of 20× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), respectively. Dideoxy sequencing on both (+) and (−) strands employed dATP[^{35}S] technology. Standard subcloning as well as successive priming with synthetic oligodeoxynucleotide 17–19 mers were used to extend sequences (5, 8). Primers were synthesized using an Applied Biosystems (Foster City, CA) 380B employing dimethoxytrityl nucleoside phosphoramidite chemistry. Sequence data were organized and analyzed using programs available through the Bionet resource.

Amino Acid Sequencing of a *Heterodontus* C_H Peptide. The isolation and automated sequencing of peptides derived by CNBr cleavage of fully reduced and [^{14}C]carboxamidomethylated heavy chains have been described (5). The NH_2-terminal sequence of peptide CN2 is homologous with the predicted sequence of the first C_H exon (see text). Chymotryptic, tryptic, and V8 staphylococcal protease digests of the peptide were separated by cation-exchange chromatography, followed by reversed-phase HPLC on a C_{18} support. Peptide homogeneity was confirmed by whole

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: V_H, variable region of immunoglobulin heavy chain; D_H, diversity segment of the heavy chain; J_H, joining segment of the heavy chain; C_H, constant region of the heavy chain.

*To whom reprint requests should be addressed.

integer amino acid composition and amino acid sequencing, through the COOH-terminal residue.

RESULTS AND DISCUSSION

Linkage of Segmental Elements. Initially, 6.6×10^5 recombinant phage were screened, and 252 phage, exhibiting considerable variation in hybridization intensity with the homologous probe, were recovered. From this group, 25 phage were partially restriction mapped, and, of these, 22 were found to be independent clones. The observed frequency of clones containing different segments is entirely consistent with the λ insert lengths, intersegment distances, and probe used for selection, i.e., with a C_H probe, $C_H > J_H$-$C_H > V_H$-J_H-C_H, and with a V_H probe, V_H-$J_H > V_H$-J_H-C_H.

Nucleotide Sequence of Different Exon 1 Segments. The nucleotide sequences of the genetic regions corresponding to the first exon of five different *Heterodontus* immunoglobulin C_Hs are shown in Fig. 1. The 5′ boundaries are assigned on the basis of the known sequence of the J_H-C_H junction of *Heterodontus* heavy-chain genes (8). The 3′ assignment is based on the experimentally determined amino acid sequence of a peptide that spans C_H exons 1 and 2 (see below). The acceptor and donor splice junctions of all five sequences are related closely and are consistent with other immunoglobulin splice consensus sequences (9, 10). The occurrence frequency in the five sequences of alternative nucleotides (smaller type) is noted.

Acceptor	Donor
T T C C C A G / C G A C	C A G / G T T A G T
A^2 T^1 A^1	C^1 G^1

All five sequences are closely related, and an efficient alignment of the sequences of 801, 3050, and 3089 can be achieved without introducing insertions or deletions. Insertion of a nucleotide triplet is required to align the sequence of 3083 with the other four sequences. Introduction of a nucleotide triplet at this position as well as the insertion of an additional five contiguous nucleotide triplets is required to align the sequence of 3047. Gene 801, which was used as the initial screening probe, is distinguished from the other genes at guanosine-134, thymidine-156, and cytidine-280 resulting in one neutral and two coding changes (see below). Nucleotide sequence differences occur more frequently in the 5′ than in the 3′ end of C_H exon 1. The extensive 3′-nucleotide identity cannot be explained by preferred codon usage or reduced codon degeneracy. Either the genes have been derived recently from a common ancestor or, more probably, are subject to frequent gene correction.

The Predicted Amino Acid Sequence of the First C_H-Coding Region. The amino acid sequences of C_H exon 1, inferred from the data contained in Fig. 1, are shown in Fig. 2. All five sequences, as well as a limited portion of the constant region of cDNA HC-3, a rearranged *Heterodontus* gene (8), are different. Cysteine residues, corresponding in location to mammalian immunoglobulin intrachain disulfides, are evi-

```
     1                                                                  51
 801 gCG ACA CCG TCT CCC CCA ACC CTT TAC GGC CTG     TGC TCC TGT GAA CAA
3083 --- --- --- --- --- --- --- --C --- --- --- TTC -C- --- --- --- ---
3050 --- --- --- --- --- --- --- --- --- --- ---     --- --- --- --G ---
3089 --A --- --A --- --G --- --- --- --- --T --T     --- --- --- --- ---
3047 --- --- -T- --- --T --C --- --C -T- -T- --- CTC -C- --- -A- --- --G

     52                                                                102
     CCC AAC ACC GAC GGC TCC ATG GCC TAC GGC TGT TTG GCG ATG GAC TAC ATC
     --- --- --- --- A-- --- G-- --- --- --- --- --- --- --- --- --- -C-
     --- --- --- --- --- --- T-- --- --- --- --- --- --A --- --- --- ---
     --- --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- ---
     --- --- --- -G- --- --- --- A-- --- --- --- --- --- --- --- --- TC-

     103                                                               153
     CCC CAA ATC ACC AGC GTT TCC TGG AAG AAA GGT AAA GAA CCG ATC ACC ACT
     --T G-- --- --T -C- --- --- --- --- --- -A- --- --G --- --- --- ---
     --- --- --- --- --- --- --- --- --- --- -A- --T --G --- --- --- ---
     --- --- --- --- --- --- --- --- --- --- -A- --T --G --- --- --- ---
     --T G-- --- --- -CT --- --- --- --- --- -A- --- --- --- --- -G- ---

     154                                                               204
     GGT TTG AAG ACT TAC CCG TCA GTG CTG AAC AAG AAG GGA ACC TAC ACC CTG
     --A --- --- -A- --- T-- --- --- --- --- --- --- --- --- --- --T ---
     --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A-
     --A --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -A-
     --A --- G-- --- --- --- --- --- --- --- --- --- --- --- --- --- -A-

     205                                                               255
     AGC AGC CAG TTA ACC ATC ACC GAG TCA GAG GTG GGA AGC AGC AAA ATC TAC
     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
     --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
     T-- G-- --- --- --- --- --A --- --- --- --- --- --- --- --- --- ---

     256                                                               306
     TGC GAG GTT CGA CGC GGA GAG TCG CTC TGG                     ATC AAG
     --- --- --- --- --- --- --- --A G-- ---                     --- --A
     --- --- --- --- --- --- --- --- G-- ---                     --- ---
     --- --- --- --- --- --- --- --- G-- ---                     --- ---
     --- --- --- --- --- --- --- --A G-- -AT TGC TCT GGG GAC ATG -A- TC-

     307   315
     GAA ATT CCA
     --- --- ---
     --- --- ---
     --- --- ---
     A-- --G ---
```

FIG. 1. Complete nucleotide sequences of five different *Heterodontus* immunoglobulin C_H exon 1 segments. The numbering refers to the triplets that would encode the first and last amino acids of the longest exon 1 (gene 3047). The first nucleotide (g^1) was inferred from the presence of an alanine at this position in the constant region of *Heterodontus* immunoglobulin (8). Codon insertions have been introduced in the sequences of genes 3083 and 3047. Shading identifies the sequence segments that were used to generate gene-specific probes.

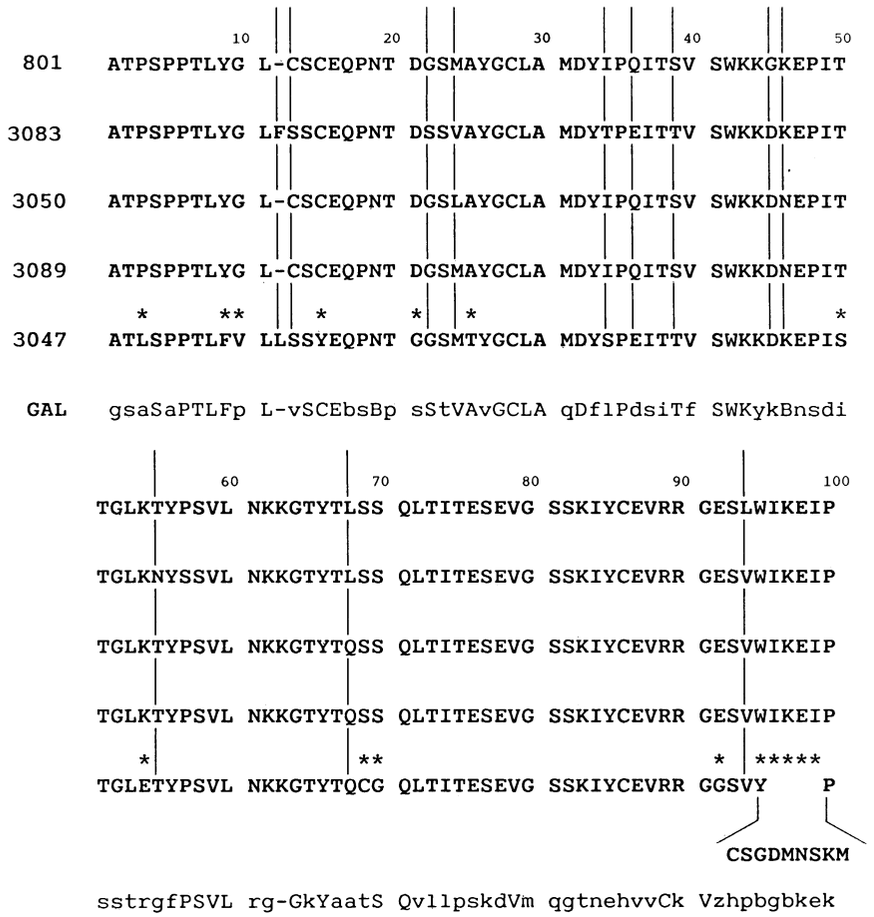

FIG. 2. Maximum predicted amino acid sequence homology of *Heterodontus* immunoglobulin C_H exon 1 regions and comparison to the sequence of human IgM GAL C_H exon 1 (11). The *Heterodontus* sequences are inferred from the nucleotide sequences shown in Fig. 1 and reflect the same codon insertions (−). Positions exhibiting differences are highlighted by vertical lines (|). Additional variation in gene 3047 is shown by * appearing over the respective amino acids. The sequence of GAL is in lower-case letters (e.g., g) that designate nonidentity with the corresponding residues in any of the five *Heterodontus* sequences or in upper-case letters (e.g., G) indicating amino acid identity with at least one of the other sequences. An alignment of *Heterodontus* Cys-86 and human Cys-89 is suggested. Amino acids are designated by the one-letter code.

dent in all five sequences as is the sequence Ser-Trp-Lys from position 41 to position 43. Another tryptophan (at position 95) occurs in four *Heterodontus* sequences but is not found in the mammalian prototype sequence. Cys-15, which is present in four genes and HC-3, corresponds to the heavy–light interchain cysteine found in mammalian immunoglobulins. The structural role of an additional cysteine located further toward the NH_2-terminal end in the three sequences presented here and in HC-3 is not clear. Within this group, 3047 is the most divergent as it encodes five additional amino acids in the hyperconserved COOH terminus and differs at 11 other positions from the consensus amino acid sequence of the four other genes. Two of these positions would encode cysteine residues that cannot be reconciled within the known disulfide bonding patterns of higher vertebrate immunoglobulins. With the mRNA quantitation and direct amino acid sequencing studies referred to below, gene 3047 may not be functional and, unless specified, will not be considered in the remaining comparisons.

Excluding the insertion of Phe-12 in the sequence of 3083, the remaining differences between the genes can be accounted for by single nucleotide changes. Position 24 is the most variable as three different amino acids are encountered in four sequences. The sequence of 801 differs at positions 45 and 94 from the other genes, including gene 3047, which are identical in nucleotide sequence at these positions. Comparison of the genes at other positions where amino acid differences are noted fails to reveal any additional, consistent patterns. Excluding gene 3047, gene 3083 differs most from the consensus sequence of the other genes. Finally, it should be noted that the predicted sequence of HC-3 differs at positions 17 and 18 (Gln-17 and Pro-18) found in all five genes in this study.

Although the lines from which the elasmobranchs and mammals diverged are separated by an extreme period of evolutionary time, the C_H exon 1 amino acid sequences of *Heterodontus* and human IgM GAL (11) are 30–35% homologous. This estimate may be somewhat elevated for the following reasons: (*i*) Amide group assignments are incomplete for the mammalian prototype. (*ii*) Residues vary at single positions in the predicted *Heterodontus* sequences. (*iii*) Single triplet insertions are required for alignment. (*iv*) A 3-amino acid insertion is needed to align intrachain Cys-86. Comparison of the C_H exon 1 sequence to the corresponding region of other immunoglobulin classes reveals an equivalent level of sequence homology with absolute conservation of certain amino acids.

Comparison of Predicted and Determined Amino Acid Sequence. In the course of other studies, considerable primary structure data, corresponding to a significant portion of C_H exons 1 and 2, have been accumulated (Fig. 3). Comparison of the predicted and experimentally determined sequences provides some indication as to the relative contribution these gene families make to the pool of circulating immunoglobulin. The sequences differ only at three positions, and it is unlikely that the Glu-47 difference in C_H exon 1 is due to an error in amino acid sequencing as it was detected in three separate peptides and is present in the corresponding peptides obtained from the immunoglobulin of individual animals. In addition, the codon corresponding to this glutamic acid was incorporated in an immunoglobulin heavy-chain class-specific, mixed oligodeoxynucleotide probe that has been shown to hybridize to an appropriate mRNA size-class in a tissue-specific manner (5). Since the five genes described in this study do not encode a corresponding glutamic acid residue, they most likely are not representative of the particular gene family (subfamily) that dominates in the production of immunoglobulin. The 801 and 3050 genes, however, are members of the more abundant

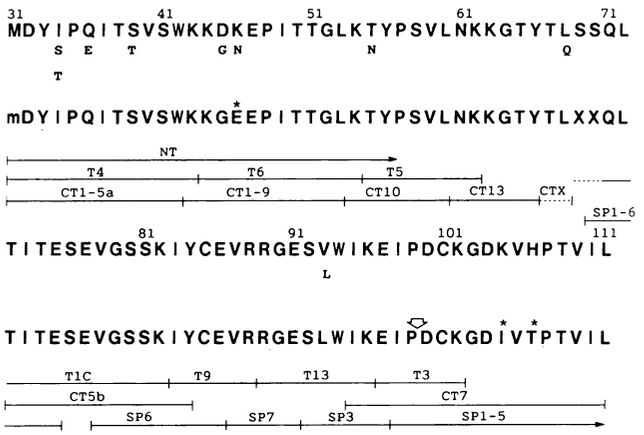

FIG. 3. Comparison of predicted (*Upper*) and determined (*Lower*) amino acid sequences of a region of the *Heterodontus* immunoglobulin heavy chain corresponding to C_H exons 1 and 2. The predicted amino acid sequence is a consensus derived from genes 801, 3083, 3050, and 3089. The most frequently occurring amino acid is shown in larger type and the less frequently occurring residue is shown below in a smaller type. The determined amino acids that show variation from the predicted consensus sequence are noted by *. The NH_2-terminal amino acid sequence of the CNBr fragment (CNH2) was determined from the constituent sequence of its tryptic (T), chymotryptic (CT), and staphylococcal V8 protease (SP) derived peptides, which are numbered accordingly as well as from the direct NH_2-terminal sequence (NT) of the parent fragment. A number occurring after the primary peptide designation, e.g., the "9" in CT1-9, refers to a secondary method used to achieve chemical homogeneity, e.g., reversed-phase HPLC. The first amino acid (M) corresponding to residue 31 of exon 1 (Fig. 2) was inferred from its known interior location within the heavy chain and the absolute specificity of CNBr for methionine; incubation of the *Heterodontus* heavy chain in formic acid alone failed to cleave the polypeptide. The open arrow between amino acids 99 and 100 represents the 3′ boundary of the first exon. This placement is based on the known genomic sequence of exons 1 and 2 and the presence of typical splice donor and acceptor sequences. It was not possible to quantitatively recover amino acids at the NH_2-terminal positions of T1C and SP1-6; these are designated by (X) and (---); the nature of the tryptic cleavage yielding T1C is not clear. CTX was placed by sequence analogy. An arrow indicates that the sequences are not complete. Amino acids are designated by one-letter code.

families of transcriptionally active constant region genes (see below).

Expression of C_H Genes. Termination codons are not present in any of the C_H exon 1 sequences. To further examine the functional status of the different genes, sequence segments specific to each were identified (Fig. 1). Regions differing significantly from the other four genes are present in genes 3047, 3083, and 3089, whereas gene 3050 varies only at one central position from several other genes. Oligodeoxynucleotides that would complement the mRNAs corresponding to these positions were used as probes in RNA blot-hybridization analyses of purified *Heterodontus* total spleen RNA (Fig. 4). Genes 801, 3050, and 3083 hybridize to various degrees. The relative size of the mRNA transcripts resembles that observed in similar analyses of spleen mRNA that employed V_H-specific probes (5), suggesting that functional transcripts are being detected. Quantitative interpretation of the results, however, is complex as a probe for the 3050 gene will hybridize, under conditions of moderate stringency, with closely related members of this extensive gene family (see below). Furthermore, the failure to detect a specific transcript with probes for gene 3047 or gene 3089 does not necessarily imply functional inactivity, since total mRNA is being characterized and a potentially functional gene may not have undergone a productive rearrangement or may be

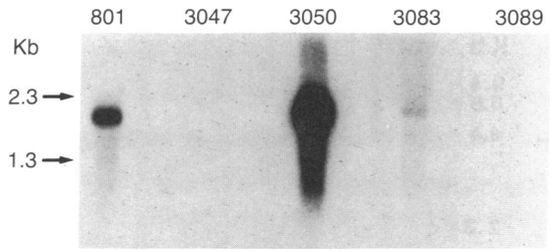

FIG. 4. RNA transfer blot analysis of spleen RNA using gene-specific oligodeoxynucleotide probes. RNA (10 μg per track) was fractionated using a 1.5% agarose/2.2 M formaldehyde gel and transferred to nitrocellulose. Synthetic oligodeoxynucleotide (19-mer) probes, inversely complementing the portions of each C_H sequence shown in Fig. 1, were labeled, hybridized to the blots, and washed as described (5). The specific activity of oligodeoxynucleotide probes was 1.2×10^8 cpm/μg, and hybridization mixes contained 2.5×10^5 cpm/ml. Autoradiographic exposure shown was 48 hr; a 96-hr exposure failed to indicate hybridization with the 3047 and 3089 probes.

transcribed at a reduced level. Even at relaxed stringency, the probes do not hybridize with mRNA derived from tissues other than spleen (data not illustrated). It is likely that at least three of the five gene families are functional.

Quantitative Gene Titration. The data presented thus far are consistent with a relatively large family of C_H genes. To more accurately assess the extent of multigene complexity, quantitative gene titrations were carried out with the presumably functional genes 801, 3050, and 3083. *Pst* I digestion was employed in these analyses as this restriction endonuclease failed to reveal major restriction fragment length polymorphism when DNAs isolated from 10 individual animals were compared (data not illustrated). Using the gene 801 C_H probe, under conditions of relaxed stringency, the majority of hybridization is localized to an ≈3.7-kilobase region that contains fragments of at least two different lengths, most evident in the 801 (homologous) titration (Fig. 5). Although assignment of absolute numbers for genes 801 and 3050 is confounded by the similar restriction fragment lengths and by the presence of flanking sequences in the probe, it is likely that the two gene families consist of >50 members, whereas gene 3083 may represent a single-copy gene. Additional C_H genes are represented by other components that hybridize at various intensities. When 17 different λ isolates were digested with *Pst* I and hybridized with the gene 801 C_H probe, six genes were found to be of the 3050 type, two were of the 801 type, four (including gene 3083) were of other types, and five could not be assigned since one of the informative *Pst* I sites is distal to the λ cloning insertion site (data not illustrated). The gene titration, restriction fragment length distribution, and expression level data are consistent with the relative occurrence frequencies being gene 3050 > gene 801 > gene 3083 and with the total number of C_H genes exceeding that found in mammals by at least one order of magnitude.

Evolutionary Significance. In mammals, the variable region gene segments are separated by extreme chromosomal distances from fewer than 10 closely linked constant region genes (6, 13). During somatic development, sequence-mediated recombination, as well as differential mRNA splicing, can result in the association of a single, joined variable region (V_H-D_H-J_H) with different C_H genes that encode various immunoglobulin classes (14). Secondary biological properties associated with these heavy-chain classes arise from limited primary structure substitutions, overall polymeric differences, as well as through the association of a heavy-chain class with other polypeptides, e.g., joining chain or secretory component. The chromosomal order of the C_H genes corresponds, in part, to the sequence of their expression during B-lymphocyte development (15, 16).

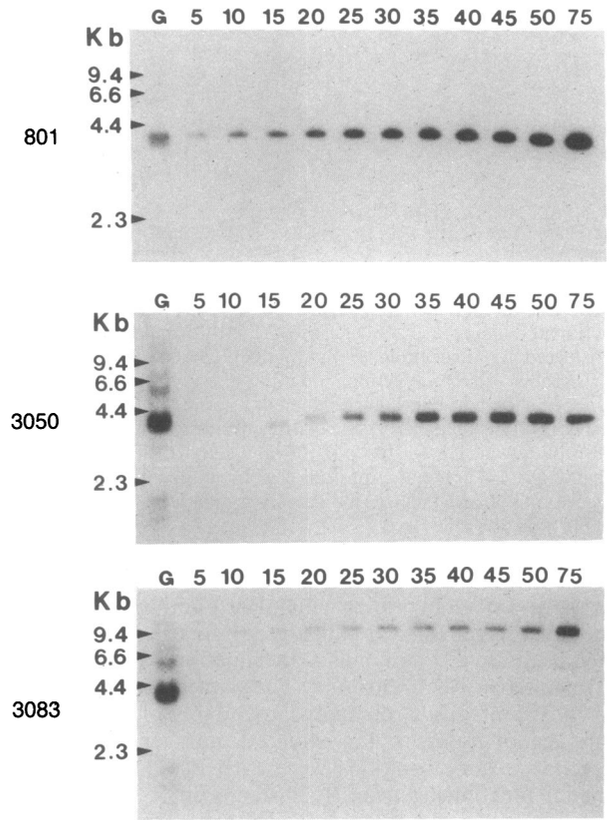

FIG. 5. Relative gene content of *Heterodontus*. High molecular weight genomic DNA (lane G) (10 μg) was digested to completion with *Pst* I and subjected to agarose gel (0.8%) electrophoresis in parallel with various quantities of *Pst* I-digested λ clones 801, 3050, and 3083 corresponding to 5–75 genomic equivalents [based on the value of 17.5 pg of DNA per 2*N* cells (12)] as indicated. After partial depurination, the DNAs were transferred to nitrocellulose, and moderate stringency blot hybridization and wash conditions were as described (5). The 801 C_H exon 1 probe was labeled to 2×10^8 cpm/μg and used at 10^6 cpm/ml. Autoradiographic exposures were 10 hr for clone 801, 24 hr for clone 3050, and 96 hr for clone 3083. Not shown are the results of an experiment in which 10 μg of human genomic DNA was added to the same amounts of the λ clones without effect on signal strength. It was concluded that the nitrocellulose binding capacity was not exceeded by a 10-μg genomic digest, thus supporting quantitative interpretation of these analyses.

In *Heterodontus*, each (V_H-D_H-J_H) segment is linked closely to an individual C_H segment. Assuming that recombination does not take place between ([V_H-D_H-J_H]-C_H) clusters, selection may act on the gene complex rather than on its individual components. The different C_H exon 1 segments may be associated with specialized, secondary biological properties as well as with primary antibody function, through the adjacency of this domain to the antigen combining site. Close linkage of segmental elements has been shown with other antigen-binding receptors, although no case is as extreme as that reported here in terms of the numbers of segments involved. A single, functional λ light-chain variable-joining-constant cluster, occupying ≈4 kilobases, has been described in birds (17). Different, closely linked λ light-chain joining and constant region genes are linked to λ light-chain variable region genes in mammals, although such clusters occupy considerably longer distances (18–20). Different joining segments are closely linked to constant segment regions in some T-cell receptor families, but only in the case of the murine T-cell γ receptors (21) are the linkage distances as close as in *Heterodontus*. If the close linkage and large numbers of different C_H regions seen in *Heterodontus* are representative of an ancestral variable region gene family, this pattern was abandoned in favor of fewer C_H regions at some stage of evolutionary development predating the mammalian radiations. Exclusive association of one constant region with a specific (V_H-D_H-J_H) segment was replaced as combinatorial diversity, presumably facilitated by extended chromosomal distances, became increasingly more important in the generation of antibody diversity. The genetic mechanisms that serve to stabilize the extremely complex multigene family found in this contemporary representative of a phylogenetically ancient species are of considerable interest and could provide important insight into the diversification of the antigen binding receptors.

The editorial assistance of Ms. Kari D. Johnson is appreciated. This work was supported by a grant from the National Institutes of Health (AI-23338). Computer analyses utilized the Bionet Resource that is supported by Grant U41-RR-01685-03.

1. Hood, L., Kronenberg, M. & Hunkapiller, T. (1985) *Cell* **40,** 225–229.
2. Davis, M. M. (1985) *Annu. Rev. Immunol.* **3,** 537–560.
3. Litman, G. W. & Hinds, K. R. (1987) in *Evolution and Vertebrate Immunity: The Antigen-Receptor and MHC Gene Families*, eds. Kelsoe, G. & Schulze, D. (Univ. Texas Press, Austin), pp. 35–51.
4. Litman, G. W., Murphy, K., Berger, L., Litman, R., Hinds, K. & Erickson, B. W. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 844–848.
5. Litman, G. W., Berger, L., Murphy, K., Litman, R., Hinds, K. & Erickson, B. W. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 2082–2086.
6. Tonegawa, S. (1983) *Nature (London)* **302,** 575–581.
7. Livant, D., Blatt, C. & Hood, L. (1986) *Cell* **47,** 461–470.
8. Hinds, K. R. & Litman, G. W. (1986) *Nature (London)* **320,** 546–549.
9. Kawakami, T., Takahashi, N. & Honjo, T. (1980) *Nucleic Acids Res.* **8,** 3933–3945.
10. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20,** 313–319.
11. Watanabe, S., Barnikol, H. U., Horn, J., Bertram, J. & Hilschmann, N. (1973) *Z. Physiol. Chem.* **354,** 1505–1509.
12. Schwartz, F. J. & Maddock, M. B. (1986) in *Indo-Pacific Fish Biology: Proceedings of the Second International Conference on Indo-Pacific Fisheries*, eds. Uyeno, T., Arai, R., Taniuchi, T. & Matsuura, K. (The Ichthyological Soc. Japan, Tokyo), pp. 148–157.
13. Shimizu, A., Takahashi, N., Yaoita, Y. & Honjo, T. (1982) *Cell* **28,** 499–506.
14. Honjo, T. & Habu, S. (1985) *Annu. Rev. Biochem.* **54,** 803–830.
15. Yancopoulos, G. D., Desiderio, S. V., Paskind, M., Kearney, J. F., Baltimore, D. & Alt, F. W. (1984) *Nature (London)* **311,** 727–733.
16. Yancopoulos, G. D. & Alt, F. W. (1986) *Annu. Rev. Immunol.* **4,** 339–368.
17. Reynaud, C. A., Anquez, V., Dahan, A. & Weill, J. C. (1985) *Cell* **40,** 283–291.
18. Blomberg, B., Traunecker, A., Eisen, H. & Tonegawa, S. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3765–3769.
19. Miller, J., Bothwell, A. & Storb, U. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3829–3833.
20. Reilly, E. B., Blomberg, B., Imanishi-Kari, T., Tonegawa, S. & Eisen, H. N. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2484–2488.
21. Traunecker, A., Oliveri, F., Nicholas, A. & Karjalainen, K. (1986) *EMBO J.* **5,** 1589–1593.
22. Litman, G. W., Berger, L. & Jahn, C. L. (1982) *Nucleic Acids Res.* **10,** 3371–3380.