Two distinct immunoglobulin heavy chain isotypes in a primitive, cartilaginous fish, *Raja erinacea*

Fiona A.Harding[1,2], Chris T.Amemiya[1,4], Ronda T.Litman[1,3], Nicholas Cohen[2] and Gary W.Litman[1,3,*]

[1]Tampa Bay Research Institute, 10900 Roosevelt Boulevard, St Petersburg, FL 33716, [2]University of Rochester, Department of Microbiology and Immunology, Box 672, Rochester, NY 14642, [3]Department of Pediatrics, University of South Florida, All Children's Hospital, 801 Sixth Street South, St Petersburg, FL 33701 and [4]Lawrence Livermore National Laboratory, Biomedical Sciences Division, L-452 Livermore, CA 94588, USA

Received June 25, 1990; Revised and Accepted September 28, 1990

Genbank accession nos M35185, M29672 – M29679+

ABSTRACT

Immunoglobulin heavy chain genes in *Raja erinacea* (little skate) are organized in clusters consisting of V_H, D_H, J_H segments and C_H exons (1). An immunoglobulin heavy chain μ-like isotype that exhibits 61 – 91% nucleotide sequence identity in coding segments to the *Heterodontus francisci* (horned shark) μ-type immunoglobulin is described. The overall length of the μ-type clusters is ~16 kb; transmembrane exons (TM1 and TM2) are located 3 to C_H exon 4 (C_H4). In three of four TM-containing genomic clones, a significant deletion is present in TM1. A second isotype of *Raja* immunoglobulin heavy chain genes has been detected by screening a spleen cDNA library with homologous *Raja* V_H- and C_H1-specific probes complementing the respective regions of the μ-like isotype. Weak hybridization with V_H-specific probes and no discernable hybridization with C_μ-specific probes were considered presumptive evidence for a second immunoglobulin isotype that nominally is designated as X-type. The V_X region of the X-type cDNA is ~60% identical at the nucleotide (nt) level to other *Raja* V_H segments and thus represents a second V_H family. Putative D_X and J_X sequences also have been identified. The constant region of the X-type immunoglobulin heavy chain gene consists of two characteristic immunoglobulin domains and a cysteine-rich carboxy terminal segment that are only partially homologous with the μ-like isotype. Genomic Southern blotting indicates that the V and C segments of both immunoglobulin heavy chain isotypes are encoded by complex multigene families. V_X- and different C_X-specific probes hybridize to different length transcripts in northern blot analyses of *Raja* spleen RNA suggesting that the regulation of expression of the X-type genes may involve differential RNA processing.

INTRODUCTION

Multiple classes of immunoglobulin are present in all sarcopterygian species, including the lungfishes, amphibians, reptiles, avians and mammals (2). During the differentiation of B lymphocytes, switch recombination allows the same antigen-binding specificity that is determined by V_H and V_L sequences, to be expressed in the context of different secondary biological properties specified by different heavy chain constant region isotypes (3,4). Immunoglobulin class switching depends on the tandem linear order of the genes encoding various constant region isotypes (5) and is influenced by environmental stimuli such as LPS and IL-4 (6,7). Switch recombination occurs via a pathway distinct from that described for immunoglobulin and T cell receptor (TCR) V-(D)-J recombination, although both can lead to the deletion of intervening sequences (8). Unique 2 – 10 kb class switch signal sequences located 5′ to each constant region gene (except δ) form a basic recognition motif for the recombination of a specific constant region with the rearranged variable region. The biological effects of class switching are apparent; however, the triggering mechanisms of this complex process are not well understood.

The majority of the elasmobranch species studied to date possess only a single immunoglobulin class; prolonged immunization fails to result in the production of additional immunoglobulin classes (9) or in an increase in antibody affinity. In one elasmobranch, *Heterodontus francisci* (horned shark), multiple immunoglobulin heavy chain gene clusters each consisting of a single variable (V_H), two diversity (D_1 and D_2), a joining (J_H) segments and a μ-like constant region (C_H) have been described (10). Each cluster spans approximately 16 kb and in at least two different individuals, approximately one-half of the clusters exhibit V_HD_H- or $V_HD_HJ_H$-joining in the germline of non-lymphoid cells (11). Considerable microheterogeneity has been observed in the μ-like C_H exons of *Heterodontus* immunoglobulin genes (12) and rearrangement of V-D_1-D_2-J may occur only within individual clusters (10) (unpublished

* To whom correspondence should be addressed

+ and EMBL accession nos X15124, X15125

observations). Due to the high degree of nucleotide similarity between V_H segments of *Heterodontus*, switch recombination between clusters is difficult to discern and cannot be ruled out.

Raja erinacea, a related elasmobranch that is representative of a separate phylogenetic order, possesses μ-like heavy chain constant region (C_H) genes that exhibit a similar cluster organization (1). Certain members of the Rajiformes (skates and guitarfish) possess two antigenically distinct molecular weight classes of immunoglobulin (13,14). These immunoglobulins share antigenically cross-reactive light chains and possess 70,000 and 45–50,000 M_r heavy chains, respectively; there is no information available regarding the primary structures of these molecules. Both the high molecular weight (HMW) and low molecular weight (LMW) forms are present in significant quantities in non-immune serum. The HMW and LMW forms are co-expressed by a significant percentage of *Raja* spleen cells (15) only during early ontogenetic development. Adult spleen cells, as well as the majority of embryonic cells, are positive for one or the other but never both classes of immunoglobulin, suggesting isotypic exclusion (15). The genetic regulation of differential isotype expression in *Raja* may be of considerable interest in terms of understanding heavy chain isotype expression in higher vertebrates. In this study, we report the isolation and determination of the complete nucleotide sequences of the heavy chains of two distinct immunoglobulin classes found in *Raja*.

MATERIALS AND METHODS

Animals

Adult specimens of *Raja erinacea* were obtained from the Marine Biological Laboratory, Woods Hole, MA. Tissues were processed immediately after the animals were sacrificed.

DNA Libraries

A genomic library was constructed in λDASH™ (Stratagene) from testes high molecular weight DNA that had been partially digested with *Sau*3A, as described (16). The library was amplified on bacterial host strain P2392, a P2 lysogen of LE392. Approximately 0.9 genome was recovered, based on a genome size of 7 pg/haploid genome (17), an average insert size of 17–19 kb and proportional representation.

RNA Isolation and cDNA Library Construction

Total RNA was isolated from the spleen of an individual specimen of *Raja* using the guanidium isothiocyanate/CsTFA technique according to the supplier's specifications (Pharmacia, Piscataway, NJ) and enriched for poly A^+ RNA by elution from an oligo dT-cellulose column. Poly A^+-selected RNA was converted to double-stranded, *Eco*RI adaptor-linkered cDNA using a commercial synthesis kit (Pharmacia) that employs a modification of a previously described procedure (18). A cDNA library consisting of 1.3×10^6 recombinant PFUs was constructed in λgt11. All cDNA library screenings were performed on the unamplified portion of the library in order to facilitate isolation of unique clones.

Filter Hybridization

Nitrocellulose lifts were hybridized in 0.6M NaCl, 0.02M EDTA, 0.2M Tris, 0.5% SDS, 0.1% sodium pyrophosphate (SET) for 12 h at 65°C, washed at 52°C in 1× SSC, 0.1% SDS, 0.05% sodium pyrophosphate and subjected to autoradiography at −70°C for 12 h.

Probes

Heterodontus derived probes utilized in the present study were: TM/5301, containing 84 nts of exon 4, the joined transmembrane (TM) exons and the 3′ untranslated segment of *Heterodontus* cDNA 5301 (19); SEC, an oligonucleotide 24mer complementing the *Heterodontus* secretory (SEC) segment (19); C_H801, a germline exon 1-specific probe derived from genomic clone 801 (12); 6121, a probe derived from a cDNA containing approximately half of exon 1 plus exons 2,3,4 including the entire SEC portion of C_H4 (19); and 5301/3,4 a 510 nt subinsert containing exons 3 and 4 from cDNA 5301 (19).

All other probes were derived from *Raja*: $110V_H$, a ~900 nt *Xba*I genomic fragment from skate clone Re110 (this is V_H-J_H-specific) (1); $107V_H$, a 1.5 kb *Xba*I genomic fragment from Re107 that contains a complete (unrearranged) V_H region (1); SkC1, a 1.8 kb *Eco*RI-*Sst*I fragment from genomic clone Re107 (see below; Fig. 2) containing C_H1; SkC2, a 1.6 kb *Sst*I-*Eco*RI fragment from Re107 that adjoins SkC1 and contains C_H2. 113ex3, a 1.3 kb *Eco*RI fragment derived from genomic clone Re113 that contains C_H3. SkTM was isolated from genomic clone (Re27054) that was identified by hybridization with TM/5301 and consists of a 1.5 *Eco*RI fragment possessing a deletion within the TM coding region (see below; Fig. 3A, 3B). $4bC_H$ consists of the entire C_H coding region from *Raja* cDNA clone Re4b and was derived using the polymerase chain reaction (PCR), employing primers that complement the 5′ end of the J region and the 3′ end of SEC. The complete sequence of the constant region of Re4b is illustrated below (Fig. 1).

V_X and C_X are DNA probes of ~450 and ~750 nts complementing the V and C regions of the X-type isotype, respectively and were derived by PCR from cDNA Re20 using sequence-specific primers. C_X1 and C_X2-SEC are DNA probes (235 and 347 nts, respectively) that were generated by PCR using primers specific for the predicted first exon and 3′ predicted coding segment of Re20. Insert DNAs were used as templates and all PCR products were gel-purified prior to use.

DNA Labelling

DNA probes were labelled to 10^8-10^9 cpm/μg using a modified version of the random hexanucleotide priming method as described (1). Oligonucleotide probes were end-labelled to a specific activity of ~10^8 cpm/μg by T4 polynucleotide kinase (Bethesda Research Laboratory) in the presence of γ^{32}P-ATP.

Genomic Southern Analyses

Total high molecular weight DNA was isolated from *Raja* testes (20). Following standard digestion with various restriction enzymes and electrophoretic separation, the DNA was partially nicked by UV irradiation, denatured and transferred to Zeta-Probe™ (Bio-Rad) in 1M NH_4OAc/20× SSC (1× SSC = 0.15M NaCl, 0.015M NaCitrate) for 20 h. The filters were baked at 80°C under vacuum for 2 h. Hybridization to DNA probes and washing were performed as described (16).

Northern Blot Analyses

Poly A^+ RNA was separated in a 1.5% agarose/2.2M formaldehyde gel, then transferred to Zeta-Probe™ in 20× SSC. The blots were dried under vacuum for 2 h at 80°C. Hybridization with DNA probes was carried out in SET as described for library lifts.

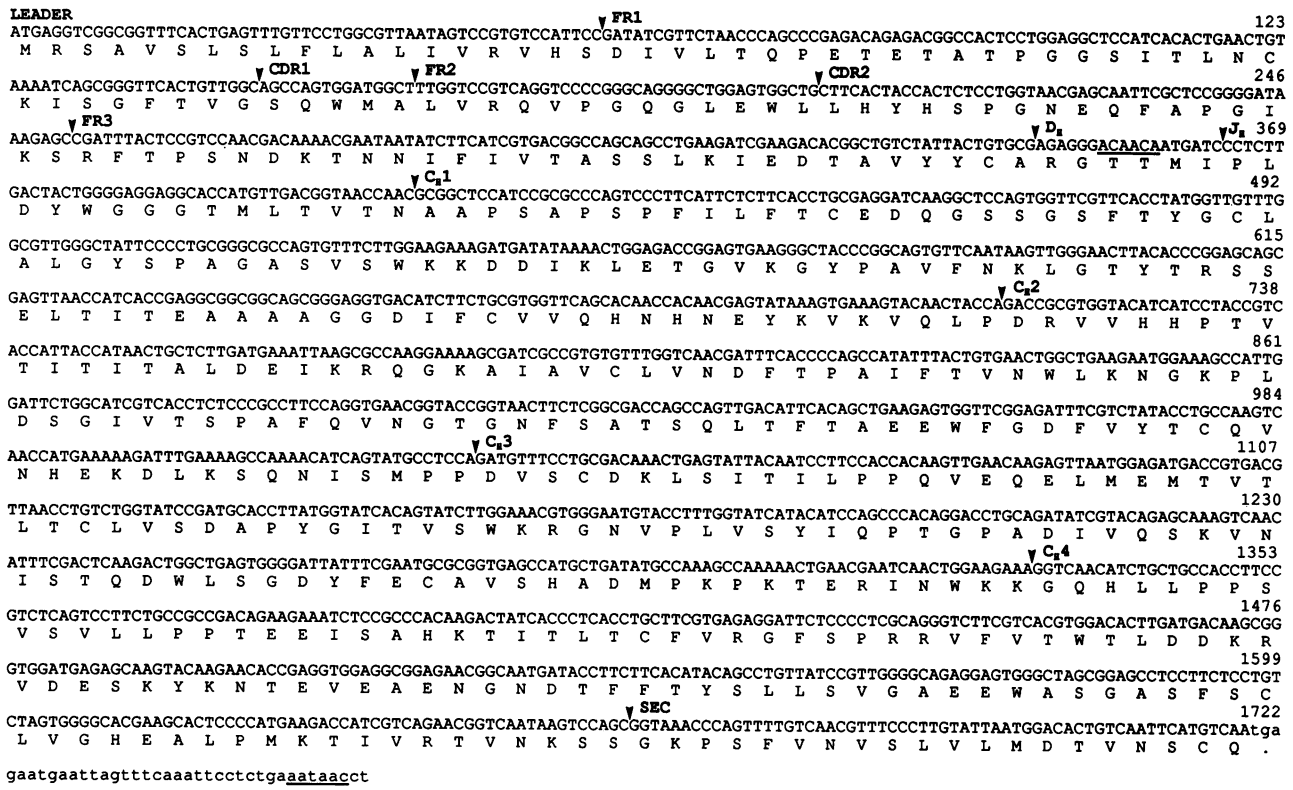

Figure 1. Nucleotide and predicted amino acid (one letter code) sequence of *Raja* cDNA clone Re4b. Nucleotides are numbered along the right hand margin. Leader, framework region (FR), complementarity determining region (CDR), diversity (D) region, joining (J) region and secretory (SEC) segment boundaries are assigned by comparison to previously published information (16). L: 1 to 57; FR1: 58 to 147; CDR1: 148 to 162; FR2: 163 to 201; CDR2: 202 to 252; FR3: 253 to 345; D: 346 to 363; J_H: 364 to 408; C_H1: 409 to 711; C_H2; 712 to 1029; C_H3: 1030 to 1329; C_H4: 1330 to 1659; SEC: 1660 to 1719. C_H exon boundaries are assigned by comparison to *Heterodontus* (19) and *Raja* (see below) genomic C_H exons. The six nt D_1-like sequence is underlined. The termination codon 1720–1722 is in lower case; the putative polyadenylation signal sequence also is underlined.

Sequence Analysis

All subcloning was done in commercially available M13 Rfs. DNA sequences were determined by the dideoxynucleotide chain termination method (21) using α-^{35}SdATP and modified T7 DNA polymerase, Sequenase™ (United States Biochemical Co.). The primary strategy for extending sequences utilized specific 18mer primers. Routine analyses of DNA sequences were made using IFIND and hydropathicity indices were estimated using algorithms available in the PEP option of the Intelligenetics Suite, copyrighted programs of Intelligenetics, Inc., Palo Alto, CA.

RESULTS

Constant Region μ Genomic Organization

An unamplified *Raja* spleen cDNA library was screened with a homologous V_H-J_H-specific probe ($110V_H$ (1)). One $V_HJ_H^+$ isolate with an insert length of ~1.8 kb was sequenced in its entirety. The complete nucleotide and predicted amino acid sequence of Re4b is shown in Fig. 1. Nucleotide sequence identity between the V_H segment of Re4b and $110V_H$ is 83%. Amino acid identity across the V_H region (including complementarity determining regions (CDRs)) is 73%. Framework (FR) and CDR boundaries have been assigned by comparison to $110V_H$ (1). A typical J_H region is found 18 nts 3′ to the V_H region. A six nt motif found in the prototypic genomic clone Re107 D_1 segment (1) is present in the Re4b D region. The relative contributions of D_1, D_2 and potential N-region additions to the Re4b D region cannot be determined in the absence of the germline 'parent' cluster. By comparison to *Heterodontus* C_H sequences, Re4b would represent a full copy length immunoglobulin heavy chain cDNA encompassing all four C_H exons, including the SEC segment of exon 4 and some 3′ untranslated (3′UT) sequence.

Table I. Percent identity of *Raja* and *Heterodontus* C_H exons. Alignments were produced by the IFIND program. Gaps in the alignments were scored as a mismatch. a) Re4b (EMBL accession #M29679) C_H1 is six nts longer than genomic C_H1. b) Re4b C_H3 is six nts shorter than genomic C_H3. c) *Heterodontus* genomic C_H3 is three nts longer than *Raja* genomic C_H3. d) *Raja* genomic exons 1 and 2 were derived from clone Re107 (EMBL accession #M29676 and M29674, respectively), exon 3 from Re113 (EMBL accession #M29675) and exon 4 and the SEC segment from Re27041 (EMBL accession #M29677).

	CH1	CH2	CH3	CH4	SEC
	Re4b Constant Region Segments				
% Nucleotide Identity					
Raja Genomic Exons	87[a]	89	84[b]	97	95
3050 CH Exons	64	61	62	76	91
% Amino Acid Identity					
Raja Genomic Exons[d]	81	79	70	97	100
3050 CH Exons	48	46	50	63	90

	Raja Genomic Exons				
% Nucleotide Identity					
Heterodontus Genomic Exons	65	61	62[c]	77	89
% Amino Acid Identity					
Heterodontus Genomic Exons	47	44	49	65	85

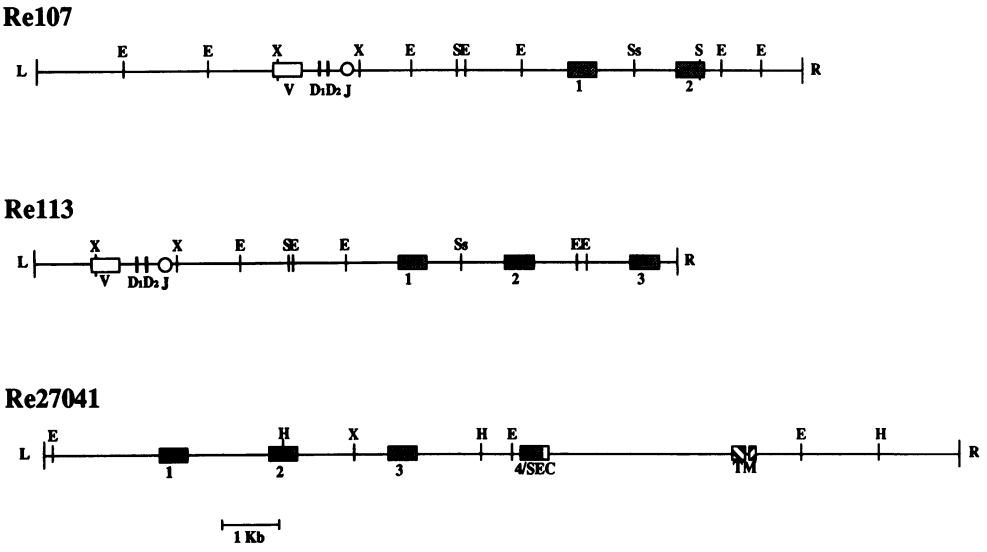

Figure 2. Partial restriction maps of representative *Raja* immunoglobulin heavy chain genomic clones. V_H segments are indicated by (V) and open rectangles, D_H segments by (D_1 or D_2) and bold black vertical lines, J_H segments by (J) and open circles, C_H exons by corresponding number and shaded boxes, secretory (SEC) segment by an open box and transmembrane (TM) segments by cross-hatched boxes. C_H1-4 assignments are based on identity with *Heterodontus* gene sequences (19) and by comparisons to a *Raja* cDNA (see text). Restriction endonucleases are E=*Eco*RI, H=*Hind*III, S=*Sal*I, Ss=*Sst*I, X=*Xba*I; L and R refer to the left and right arms of the β phage vectors; scale in kb pairs is indicated.

The location of predicted cysteine and tryptophan residues also is consistent with four typical immunoglobulin C_H domains. Nucleotide identities between the putative constant region domains of Re4b and *Heterodontus* genomic clone 3050 exons (19) range from 61% for C_H2 to 91% for the SEC segment of exon 4 (Table 1). The sequence AATAAC is found 28 nts 3′ to the termination codon of the SEC segment; this sequence differs from the consensus polyadenylation signal by one nt (underlined) (22).

The C_H exon boundaries of cDNA Re4b were assigned by comparison to *Raja* C_H exon sequences derived from different genomic clones. Nucleotide identities between the cDNA and genomic C_H1 and C_H2 of Re107 (Fig. 2) are 87 and 89%, respectively (Table 1). This high degree of nucleotide sequence identity between Re4b and the Re107 gene sequences indicates that they represent similar C_H exons. At the level of predicted amino acid sequence, C_H1 and C_H2 domains of Re4b and Re107 are 81 and 79% related, respectively. Nucleotide sequence identity between C_H3 of Re4b and Re113 (Fig. 2) is ~84% and the predicted amino acid sequence identity is 70% (Table 1).

In order to facilitate identification of a genomic clone containing all four C_H exons, a skate genomic library was screened with 5301/TM (19). TM^+ clones were isolated and screened with probes for the other C_H exons. Re27041 hybridizes with SkC1, SkC2, 113ex3, 5301/TM (Fig. 2), but does not hybridize with a homologous V_H probe. The sequence of C_H4 in Re4b is 97% identical at both the nucleotide and predicted amino acid sequence levels to that of Re27041 C_H4 (Table 1). The predicted amino acid sequence of the SEC segments of Re4b and Re27041 are identical. Two consensus polyadenylation sites are located 27 nts 3′ to SEC and 4 nts downstream of the first AATAAA, as in *Heterodontus* (19).

Nucleotide identities between comparable genomic C_H exons of *Raja* and *Heterodontus* range from 61% for C_H2 to 77% for C_H4 and are even higher if the SEC portion of C_H4 is included (Table 1). Amino acid identities between genomic C_H exons of *Raja* and *Heterodontus* are considerably lower, varying from 44% for C_H2 to 65% for C_H4 and 85% for the SEC segment.

Genomic clone Re27041 contains an intact TM sequence ~1.8 kb 3′ to C_H4/SEC. The coding segments of this sequence are shown in Fig. 3A. TM1 and TM2 are separated by a 130 nt intervening sequence (IVS) (not shown). The presence of separate TM exons is consistent with the assignment of this immunoglobulin locus as μ-like. Because the length of the IVS separating C_H1 and C_H2 is the same in Re107, Re113 and Re27041, and the IVS separating V_H and C_H1 are similar in Re107 and Re113 (Fig. 2), a reasonable estimate of the overall genomic linkage distance (V_H→TM) of a *Raja* Ig gene cluster is ~16 kb, similar to that described in *Heterodontus*.

The nucleotide sequence of a $5301/TM^+$ *Eco*RI fragment from genomic clone Re27054 was determined and is compared to *Heterodontus* (5301/TM) and *Raja* TM-containing sequences (Fig. 3A). Homology of Re27054TM to all other TM^+ isolates terminates abruptly in TM1 (Fig. 3A arrow). A splice consensus sequence is not associated with the homology breakpoint. An in-frame stop codon was detected 15 (predicted) amino acids 3′ to the homology breakpoint in clone Re27054 (Fig. 3A). TM2-like sequences, typically exhibiting strong phylogenetic conservation, were not detected within 683 nts 3′ to this stop codon (data not shown). A consensus polyadenylation site (AATAAA) is located 565 nts 3′ to the homology breakpoint.

The *Eco*RI fragment of Re27054 containing this putative TM1-like fragment was used to screen a portion of the unamplified spleen cDNA library. A C_H4^+/TM^+ cDNA (ReTM6) was identified and sequenced. ReTM6 has an intact, joined TM region that is 96% identical at the nucleotide level and 100% identical at the amino acid level to the Re27041TM exons. By comparison to ReTM6 the deletion in Re27054 includes 11 predicted amino acids representing 25% of TM1, all of TM2 and 35 nts of 3′ untranslated sequence (not shown). Therefore, the entire deletion in Re27054 encompasses 33 nts of TM1 coding sequence, the 134 nts IVS (including splice consensus sequences), TM2 and 35 nts of 3′ untranslated sequence or 202 nts total (Fig. 3B). This atypical TM-like sequence is exhibited by two genes in addition to 27054; thus

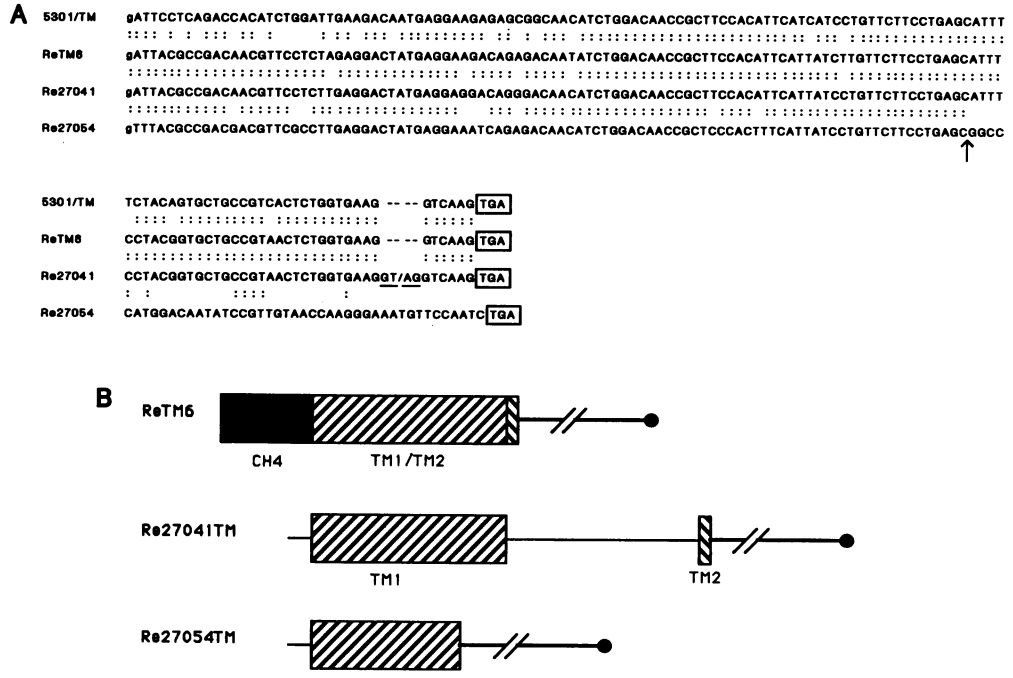

Figure 3. A.Sequence alignments of *Heterodontus* (5301/TM, selection probe) and *Raja* TM segments. ReTM6 (EMBL accession #M29673) and 5301/TM are TM^+ cDNA clones. Genomic clones Re27041 (EMBL accession #M29678) and Re27054 (EMBL accession #M35185) possess an intact and truncated TM regions, respectively. Splice acceptor nucleotides for the C_H4 splice donor site are in lower case letters. TM1−2 IVS splice consensus sequence (GT/AG) for Re27041 is underlined and the TM1−2 IVS in Re27041 is 130 nts (not shown). The homology breakpoint between Re27054 and all other isolates is indicated by an arrow. Termination codons are enclosed in boxes. B.Schematic organization of *Raja* TM exons in cDNA ReTM6 and in Re27041 and Re27054 (genomic) configuration. C_H4 is represented by a shaded box and TM1 and TM2 are shown by cross hatched boxes. IVSs are thin lines and 3′ untranslated (3′UT) segments are bold lines. The distance from TM2 to the putative polyadenylation site (represented by closed circles) is 599 nts for Re TM6. The presence of ~150 nts of sequence 3′ to the putative polyadenylation site in ReTM6 (not shown) implies that this site was not used to form the end of this mRNA. The predicted second polyadenylation site and poly A^+ tract was not found in this clone. The polyadenylation site in 27041 is 586 nts 3′ to TM2 and differs from consensus by one nt. The distance from the end of the 'coding segment of Re27054 to a potential polyadenylation site is 565 nts.

three out of four otherwise unique 5301/TM^+ genomic isolates (data not shown) are truncated.

A Second Constant Region Isotype is Present in *Raja Erinacea*

Approximately 1.2×10^5 unamplified recombinant cDNA pfu were screened in duplicate using V_H- (110V_H (1)) and C_H1-specific (SkC1) probes. Six weakly hybridizing V_H^+, C_H1^- plaques were detected. Clone Re20 was found to contain a 1.2 kb insert; inserts of less than 200 bp were identified in the other five clones. The complete nucleotide sequence of Re20 is illustrated (Fig. 4). The 5′ (V_H) sequence of Re20 is 60% identical at the nucleotide level to 110V_H; by definition, this represents a unique V_H family. The predicted amino acid sequences of Re110 and Re107 V regions are 46% and 47% related to Re20, respectively (Fig. 5A). This is in contrast to the 78% predicted amino acid homology between the V regions of Re110 and Re107. The V segment (V_X) is separated by 24 nts from a unique but identifiable J region ('J_X') (Fig. 5B). The relative contributions of putative D_X segments to this region cannot be determined without identification and characterization of genomic D_X sequences; however, identity is apparent between this region and known *Raja* μ-type immunoglobulin D segments (1). It is apparent that the J_H segment of Re107, a μ-type gene, is closely related to the J_X sequence of Re20; however, an extended segment of eight nts at the putative D/J_X boundary also is identical to the sequence of the D_2 segment of Re107 reported previously. The most likely explanation is that the X-type gene utilizes a similar D_2 and that in this particular case of D_2 joining, there has been junctional (deletion of nucleotides) and no N-type diversity at the D_2/J_H junction (assuming the general scheme of elasmobranch immunoglobulin gene organization). The more 5′ identities between the cDNA and the μ-type genomic sequences are consistent with the utilization of a partially homologous D_1 segment and possible V-D_1/D_1-D_2 junctional and N diversity. The distributions of cysteine, tryptophan and other critical amino acids in the sequence region 3′ to J_X suggests that there at least two immunoglobulin domains (C_X1, C_X2/SEC) (Fig. 4). Re20 constant region domains C_X1 and C_X2 are defined arbitrarily by 100 predicted amino acids. The N terminal 13 amino acids of the putative C_X2 contains five proline residues. The 3′ end of the C_X2 segment (44 amino acids) encodes five cysteine residues and terminates in TAA. Based on hydropathicity indices, it is unlikely that this 44 amino acid stretch is involved in membrane anchoring. Multiple hybridizing components were apparent in Southern blots of *Raja* genomic DNA hybridized with V_X- and C_X-specific probes (Fig. 6), consistent with the designation of the X-type components as members of a multigene family.

Poly A^+-selected RNAs from liver and spleen were examined by northern blotting using V_X- and C_X-specific probes. The presence of three different sized messages in spleen (Fig. 7) was noted with both V_X- and C_X-specific probes; no hybridization was evident with liver RNA. In northern blot analyses of *Raja* spleen poly A^+-selected RNA, V_X- and C_X1-specific probes hybridize at equivalent intensity to RNA species migrating at gel positions corresponding to 3.3, 1.3 and 1.0 kb (Fig. 7, lanes 1

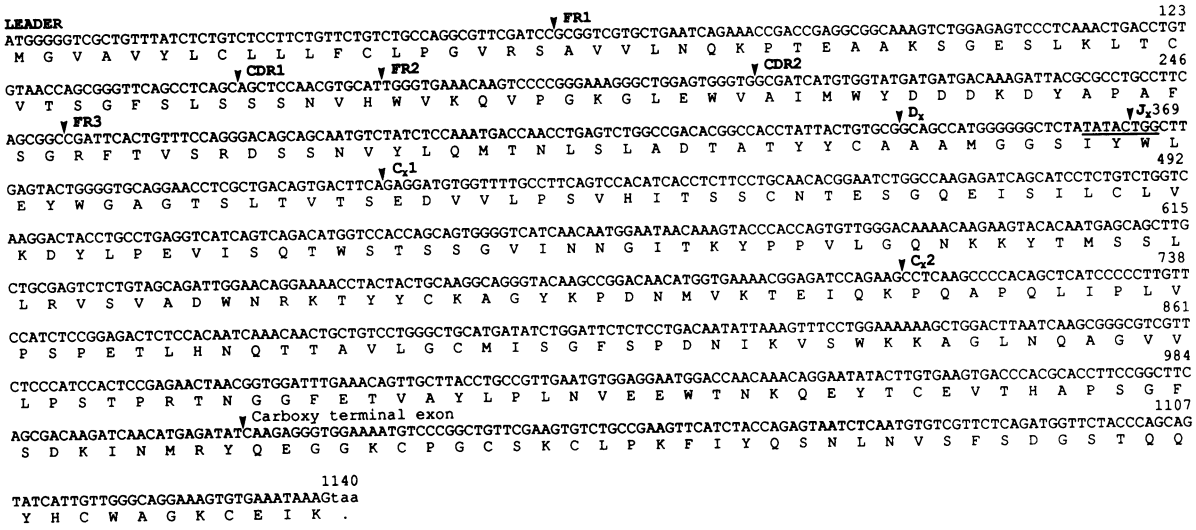

Figure 4. Nucleotide and predicted amino acid (one letter code) sequence of *Raja* cDNA clone Re20 (EMBL accession # M29672). Nucleotides are numbered in right margin. Leader, framework (FR), complementarity determining region (CDR), diversity (D) region and joining (J) boundaries were determined by comparison to previously published information (1). D_2-like sequences are underlined. L: 1 to 57. FR1: 58 to 147. CDR1: 148 to 162. FR2: 163 to 201. CDR2: 202 to 252. FR3: 253 to 339. D region: 340 to 363. J: 364 to 408. Constant region: 409 to 1140. The assignments of constant region boundaries are based only on considerations of immunoglobulin domain length and are tentative. The predicted location of the carboxy terminal exon is indicated directly. Termination codon is in lower case; only eight additional nts were present 3′ to taa in this clone, thus it is not possible to locate a polyadenylation signal sequence.

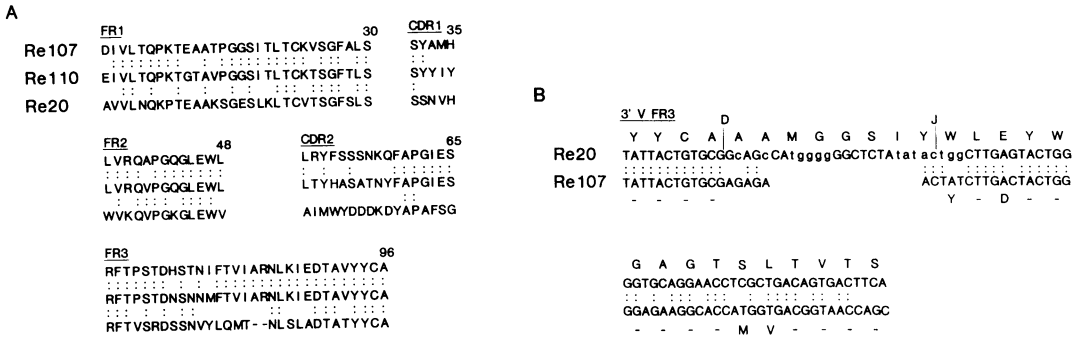

Figure 5. A. Alignments of the predicted amino acid sequences of the V_H regions of *Raja* genomic clones Re107 (EMBL accession # X15124), Re110 (EMBL accession # X15125) and Re20 (V_X). Framework (FR) and complementarity determining region (CDR) assignments are based on previously published information (16). Alignment gaps have been introduced to maximize sequence identity indicated by (:). FR1: 1 to 30. CDR1: 31 to 35. FR2: 36–48. CDR2: 49–65. FR3: 66 to 96. B. Nucleotide and predicted amino acid sequence of the 3′ of V (V_X see 7A.), putative D (D_X) and J (J_X) regions of Re20. D and J assignments are tentative and correspond to Fig. 6. Nucleotide identity between Re20 and Re107 is indicated by (:). Predicted amino acid sequence identities in the corresponding V_H and J_H segments of genomic clone Re107 are shown immediately below as (−) and differences in J are indicated in one letter code. Nucleotide identity of Re20 to Re107D_H1 and D_H2 segments are indicated by lowercase letters in the sequence of Re20; the extended sequence identity at the D/J junction of Re20 also is indicated in Fig. 6. Predicted amino acid identities in D_1 and D_2 are not shown.

and 2), whereas a probe specific for C_X2 and the putative C-terminal segment hybridizes only with the RNA species corresponding to 3.3 and 1.3 kb (Fig. 7, lane 3). These results suggest the significant presence of an RNA species that lacks all or a considerable length of C_X2.

DISCUSSION

Constant Region μ Genomic Organization

The skate *Raja erinacea* exhibits a clustered immunoglobulin μ genomic organization similar to that described for *Heterodontus*. The entire non-rearranged coding unit from the V_H segments through the TM spans ~16 kb in both species. Multiple copies of the V_H-C_H linked clusters are present in the genome. The inter-exon distances between the individual C_H exons, however, are much longer in *Raja* than in *Heterodontus*. The appreciably longer J_H-C_H IVS in *Heterodontus* accounts for the overall similar cluster sizes in these two species. It is possible that there is selective pressure to maintain a ~16 kb overall cluster length irrespective of inter-exon distances, although this relationship may be inconsequential.

Another notable difference between *Raja* and *Heterodontus* involves the large percentage of apparently non-functional TM segments. In this study, three out of four genomic clones from *Raja* possess deletions in the 3′ end of TM1 and an absence of TM2 within the TM1-TM2 linkage distance described for a presumably functional *Raja* immunoglobulin gene clone (Re27041). The genes exhibiting the deletions are presumed to be pseudogenes; however, elements of the TM-deleted clusters may recombine with other genes possessing functional TM

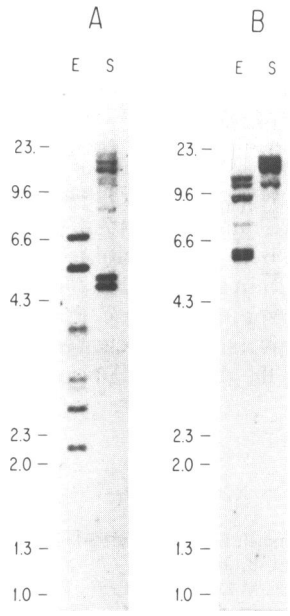

Figure 6. Southern blot of *Raja* genomic DNA. Testes DNA (10μg/lane) was digested with *Eco*RI (E) or *Sst*I (S) and transferred to Zeta-Probe™ as described in Materials and Methods. Duplicate strips from the same blot were probed with: A) V_X, the V region of cDNA Re20 or B) C_X, the complete constant region of cDNA Re20. Probes were labelled to $\sim 10^9$ cpm/ml. Hybridization was carried out at $\sim 3 \times 10^5$ cpm/ml; hybridization and washing conditions were as described (16). Standards are in kb.

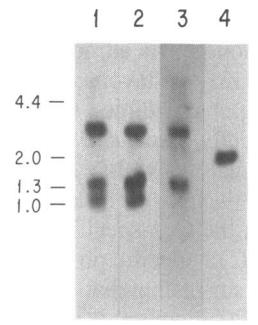

Figure 7. *Raja* Northern blot/hybridization analysis. *Raja* spleen poly A^+ RNA (2μg/track) was fractionated in a 1.5% agarose/2.2M formaldehyde gel, transferred to Zeta-Probe™ (Bio-Rad) and hybridized with: 1) V_X, the V region of cDNA Re20; 2) C_X1, a probe complementing the first exon of cDNA Re20; 3) C_X2, a probe complementing the entire second exon through 3′ coding segment of cDNA Re20; and 4) $4bC_H$, the entire C_H coding region of cDNA Re4b. Sizes (in kb) are given on left. C_X1 corresponds to nucleotides 464–699 of Re20; C_X2 corresponds to nucleotides 829–1176 of Re20 (Fig. 6).

sequences. Alternatively, a non-functional TM region may essentially regulate the exclusive expression of a secreted form of immunoglobulin, although this would be contrary to our understanding of membrane bound antibodies as primary antigen receptors. Enhanced expression of secretory immunoglobulin through any mechanism would account for the low number of TM^+ cDNA clones that have been detected.

A Second Constant Region Isotype

A cDNA representing a second heavy chain isotype (C_X) has been detected in *Raja*. The open reading frame in Re20 is consistent with three typical immunoglobulin domains ('V_X-C_X1-C_X2') as well as an additional, cysteine-rich region that may serve as the equivalent of a secretory segment. The predicted amino acid sequence of Re20 would encode a 40,000 M_r heavy chain molecule, consistent with the molecular mass of the 'second-class' heavy chain detected in any other rajiform species (14). It is possible that Re20 corresponds to this distinct heavy chain constant region isotype.

Comparisons of Re20 exons to *Heterodontus* and *Raja* μ constant region exons at the nucleotide and predicted amino acid level indicates relatedness between Re20 C_X1 and *Heterodontus* C_H1 at 21/96 positions. *Heterodontus* C_H exons lack significant nucleotide or predicted amino acid identity with Re20 C_X2. Re20 C_X2 exhibits a similar level of predicted amino acid sequence identity to both exons 2 and 4 of Re4b (26/106 and 23/106, respectively). As the only significant amino acid and nucleotide identity was between *Heterodontus* C_H exons, and *Raja* μ- and X-type exons, the μ and × clusters in *Raja* may have evolved from a primordial μ-type cluster.

Nucleotide and predicted amino acid identities between *Raja* C_H exons and most mammalian, amphibian and avian constant region exons exhibit considerably less identity than between the two *Raja* isotypes, and *Raja* μ-type and *Heterodontus* μ-type exons. An exception is the human $C_\mu1$, which demonstrates 18/106 predicted amino acid identities to Re20 C_X2. The significance of this finding is unclear. As the X-type constant region is a member of a multigene family, the X-type cluster represented by Re20 may exhibit a fortuitous level of predicted amino acid identities to the human $C_\mu1$ sequence. Similarly, due to the fact that the X-type constant region exons are members of a multigene family, it is difficult to perform 'molecular clock' analyses on the one nucleotide sequence represented by Re20 and other vertebrate constant region exons.

Due to the lack of predicted amino acid and nucleotide identity, it is unlikely that the X-type cluster is an evolutionary forerunner to the typical immunoglobulin isotypes of higher vertebrates. The X-type gene may be a 'lost' isotype in the evolutionary history of constant regions. An immunoglobulin with an apparent molecular weight consistent with a variable and two constant region domains also is present in at least one avian, lungfish and a reptilian. A two domain constant region (with an additional unrelated sequence segment) may have evolved independently at two points during the radiation of the vertebrate lineage, represented in the skate and these other species.

The V_X sequence is ~60% related to the previously described family of V_H segments in *Raja*. Although J_H μ-type segments in *Raja* vary considerably in predicted amino acid sequence, the J region of Re20 exhibits amino acid motifs that are found in other J segments (1). The D region of Re20 possesses nucleotide sequences similar to those found in germline μ-type D_H segments. An eight nt sequence which overlaps the first predicted codon of the J_X region is identical to Re107 D_2. Whether this is a coincidence or represents highly conserved D elements between the X- and μ-type clusters cannot be resolved without identifying the germline segments that gave rise to Re20.

The predicted amino acid sequence of Re20 suggests that two typical constant region domains (segments) are present as well as a shorter carboxy terminal segment. Since a 1.3 kb band is visualized with both V_H- and C_H-specific probes on northern blots of spleen RNA, it is unlikely that this particular cDNA derives from a potentially degraded mRNA. Without sequence data for genomic clones, it is difficult to predict exon boundaries. However, a proline-rich region of ~13 amino acids is located between the putative ends of C_X1 and the start of C_X2 and may

correspond to the hinge-region found in higher vertebrate immunoglobulin. Although it is attractive to speculate that the cysteine rich terminus may be involved in dimer formation, the 320,000 M_r immunoglobulin dimer found in related rajiform species are not covalently associated (13).

RNA blot transfer analyses with V_X- and C_X-specific probes show the presence of ~3.3, ~1.3 and ~1.0 kb hybridizing components. Whereas the C_X1-specific probe hybridizes to all three RNA classes, a C_X2-specific probe hybridizes only to the 3.3 and 1.3 kb bands, suggesting that the 1.0 kb RNA may lack C_X2 and the unique SEC segment. It is unclear whether the 1.3 and 1.0 kb products arise from alternative splicing of a 3.3 kb (or larger) precursor RNA or whether these are synthesized by different clusters. The 3.3 kb component is distinct from the 2.2 kb μ-type mRNAs that hybridize with $4bC_H$ (Fig. 7) as well as with V_H (data not shown) probes.

At this time, based on the results described above, it is reasonable to conclude that differential RNA processing of a common μ-X-type RNA and isotype switching (μ-type/X-type) are not likely, since a non-homologous V region (V_X) is found in clone Re20 and there is a lack of tandem repetitive DNA structures in the J_H to C_H1 IVS of *Raja* genomic clones (data not shown) such as those mediating immunoglobulin class switching in higher vertebrates (23). Furthermore, no V_X or C_X sequences are detected by hybridization within 10 kb 3′ of the C_H TM sequences (unpublished observation). Therefore, although *Raja erinacea* possess two distinct V region families and two immunoglobulin heavy chain isotypes, these appear to be regulated differently than different isotypes found in the higher vertebrates. The genomic organization and the regulation of expression of the gene encoding the second isotype currently is under study.

CONCLUSIONS

It has been suggested that with respect to immunoglobulin heavy chain gene organization, the elasmobranchs are representative of a primitive form of rearranging gene (24). The multiple clustered organization of elasmobranch immunoglobulin loci presumably reflects the evolutionary decision to duplicate entire V-C linked clusters, whereas the organization of the higher vertebrate immunoglobulin genes is associated with duplication of individual segmental elements (10). The immunoglobulin heavy chain μ-like loci in *Raja* are similar to those found in *Heterodontus*; however, the second heavy chain isotype that is associated with a unique V region family suggests an appreciable divergence from the *Heterodontus* system. Along with previous observations regarding the diversity of J_H sequences (1), studies reported here indicate that *Raja* expands diversity through variation in the sequence of both V_H and C_H coding regions.

The respective phylogenetic lineages represented by *Raja* and *Heterodontus* probably did not diverge until after the putative initial duplication event of V-C linked μ-like clusters. The second immunoglobulin isotype presumably diverged from one (or more) of these clusters. It is possible that the second immunoglobulin isotype present only in *Raja* represents the evolutionary forerunner of the μ-type cluster (present in both *Raja* and *Heterodontus*). *Raja* is the most phylogenetically primitive species in which the presence of two distinct isotypes has been unequivocally documented. The developmental regulation of these two isotypes and their functional significance in terms of antigen recognition at the cell surface are of considerable interest. Based on these results, an in-depth analysis of the genetic regulation of this system appears to be approachable.

ACKNOWLEDGEMENTS

The authors would like to thank B. Pryor for editorial assistance. This work is supported by NIH grant AI23338 (GWL); FAH is supported by NIH Training Grant AI07285 from the USPHS. CTA was supported by a postdoctoral fellowship for molecular studies of evolution from the Alfred P. Sloan Foundation. GWL is the Hines Professor of Pediatric Research. This work is submitted to the University of Rochester in partial fulfillment of requirements for the degree of Doctor of Philosophy.

REFERENCES

1. Harding, F.A., Cohen, N. and Litman, G.W. (1990) Nuc. Acids Res. 18, 1015–1020.
2. Litman, G.W., Frommel, D., Chartrand, S.L., Finstad, J. and Good, R.A. (1971) Immunochem. 8, 345–349.
3. Abney, E.R., Cooper, M.D., Kearney, J.F., Lawton, A.R. and Parkhouse, R.M.E. (1978) J. Immunol. 120, 2041–2049.
4. Pernis, B., Forni, L. and Luzzati, A.L. (1976) Cold Spring Harbor Symposia 41, 175–183.
5. Marcu, K.B. (1982) Cell 29, 719–721.
6. Lutzker, S., Rothman, P., Pollock, R., Coffman, R. and Alt, F.W. (1988) Cell 53, 177–184.
7. Esser, C. and Radbruch, A. (1989) EMBO J. 8, 483–488.
8. Wabl, M., Meyer, J., Beck-Engeser, G., Teukhoff, M. and Burrows, P.D. (1985) Nature 313, 687–689.
9. Shankey, T.V. and Clem, L.W. (1980) J. Immunol. 125, 2690–2698.
10. Hinds, K.R. and Litman, G.W. (1986) Nature 320, 546–549.
11. Kokubu, F., Litman, R., Shamblott, M.J., Hinds, K. and Litman, G.W. (1988) EMBO J. 7, 3413–3422.
12. Kokubu, F., Hinds, K., Litman, R., Shamblott, M.J. and Litman, G.W. (1987) Proc. Natl. Acad. Sci. USA 84, 5868–5872.
13. Kobayashi, K., Tomonaga, S. and Kajii, T. (1984) Mol. Immunol. 21, 397–404.
14. Kobayashi, K. and Tomonaga, S. (1988) Mol. Immunol. 25, 1115–1120.
15. Kobayashi, K., Tomonaga, S., Teshima, K. and Kajii, T. (1985) Eur. J. Immunol. 15, 952–956.
16. Litman, G.W., Berger, L., Murphy, K., Litman, R., Hinds, K.R. and Erickson, B.W. (1985) Proc. Natl. Acad. Sci. USA 82, 2082–2086.
17. Stingo, V. (1979) Genetica 50, 227–239.
18. Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263–269.
19. Kokubu, F., Hinds, K., Litman, R., Shamblott, M.J. and Litman, G.W. (1988) EMBO J. 7, 1979–1988.
20. Blin, N. and Stafford, D.W. (1976) Nucleic Acids Res. 3, 2303–2308.
21. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
22. Levitt, N., Briggs, D., Gil, A. and Proudfoot, N.J. (1989) Genes & Development 3, 1019–1025.
23. Kataoka, T., Miyata, T. and Honjo, T. (1981) Cell 23, 357–368.
24. Litman, G.W., Shamblott, M.J., Haire, R., Amemiya, C., Nishikata, H., Hinds, K., Harding, F., Litman, R. and Varner, J. (1989) in Progress in Immunology VII, Melchers, F. Ed. pp. 361–368, Springer-Verlag, Berlin.