

This information is current as of August 8, 2022.

The *IgH* Locus of the Channel Catfish, *Ictalurus punctatus*, Contains Multiple Constant Region Gene Sequences: Different Genes Encode Heavy Chains of Membrane and Secreted IgD

Eva Bengtén, Sylvie M.-A. Quiniou, Tor B. Stuge, Takayuki Katagiri, Norman W. Miller, L. William Clem, Gregory W. Warr and Melanie Wilson

J Immunol 2002; 169:2488-2497; ; doi: 10.4049/jimmunol.169.5.2488 http://www.jimmunol.org/content/169/5/2488

**References** This article **cites 36 articles**, 14 of which you can access for free at: http://www.jimmunol.org/content/169/5/2488.full#ref-list-1

Why The JI? Submit online.
Rapid Reviews! 30 days\* from submission to initial decision
No Triage! Every submission reviewed by practicing scientists
Fast Publication! 4 weeks from acceptance to publication
\*average

Subscription Information about subscribing to The Journal of Immunology is online at:

- Permissions
   Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html
- **Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts



# The *IgH* Locus of the Channel Catfish, *Ictalurus punctatus*, Contains Multiple Constant Region Gene Sequences: Different Genes Encode Heavy Chains of Membrane and Secreted IgD<sup>1</sup>

Eva Bengtén,<sup>2</sup>\* Sylvie M.-A. Quiniou,<sup>3</sup>\* Tor B. Stuge,<sup>4</sup>\* Takayuki Katagiri,<sup>5</sup>\* Norman W. Miller,\* L. William Clem,\* Gregory W. Warr,<sup>†</sup> and Melanie Wilson\*

The  $\delta$ -chain of catfish IgD was initially characterized as a unique chimeric molecule containing a rearranged VDJ spliced to C $\mu$ 1, seven C domain-encoding exons ( $\delta$ 1– $\delta$ 7), and a transmembrane tail. The presence of cDNA forms showing splicing of  $\delta$ 7 to an exon encoding a secretory tail was interpreted to indicate that membrane ( $\delta$ m) and secreted ( $\delta$ s) forms were likely expressed from a single gene by alternative RNA processing. Subsequent cloning and sequence analyses have unexpectedly revealed the presence of three  $\delta$  C region genes, each linked to a  $\mu$  gene or pseudogene. The first (*IGHD1*) is located 1.6 kb 3' of the functional C $\mu$  (*IGHM1*). The second (*IGHD3*) is positioned immediately downstream of a pseudo C $\mu$  (*IGHM3P*), ~725 kb 5' of *IGHM1*. These two  $\delta$  genes are highly similar in sequence and each contains a tandem duplication of  $\delta$ 2- $\delta$ 3- $\delta$ 4. However, *IGHD1* has a terminal exon encoding the transmembrane region, whereas *IGHD3* has a single terminal exon encoding a secreted tail. The occurrence of *IGHD3* immediately downstream of a  $\mu$  pseudogene indicates that the putative  $\delta$ s product may not be expressed as a chimeric  $\mu\delta$  molecule. Western blots and protein sequencing data indicate that an *IGHD3*-encoded protein is expressed in catfish serum. Thus, catfish  $\delta$ m transcripts appear to originate from *IGHD1*, whereas  $\delta$ s transcripts originate from *IGHD3* rather than, as previously inferred, from a single expressed  $\delta$  gene. The third  $\delta$  (*IGHD2*) is associated with a pseudo C $\mu$  (*IGHM2P*); its presence is inferred by Southern blot analyses. *The Journal of Immunology*, 2002, 169: 2488–2497.

he predominant teleost serum Ig has been well characterized at both the protein and molecular levels and is considered to be a tetrameric homolog of IgM (reviewed in Refs. 1 and 2). In contrast, much less is known about the teleost homolog(s) of IgD. Before the description of a  $\delta$  homolog in channel catfish, Ictalurus punctatus (3), IgD had been found only in primates and rodents (4–7). The catfish  $\delta$  homolog, initially identified at the cDNA level, was unique in that it was expressed as a chimeric molecule consisting of a rearranged VDJ, a Cµ1 exon, and seven novel C region domains, some of which showed sequence homology to mammalian  $\delta$  by phylogenetic analyses. Fulllength cDNAs for the membrane ( $\delta m$ ) form and partial cDNAs for the secreted ( $\delta$ s) form of catfish  $\delta$  were identified (3). In addition, these earlier studies demonstrated that the catfish  $\delta 1$  exon was located 1.6 kb 3' of the  $\mu$  gene and that identifiable class-switch sequences were absent from this short intron. These prior findings suggested that

<sup>2</sup> Address correspondence and reprint requests to Dr. Eva Bengtén, Department of Microbiology, University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216-4505. E-mail address: ebengten@microbio.umsmed.edu catfish  $\delta$  message, like that of mammalian  $\delta$ , is produced by alternative mRNA splicing rather than class switching involving chromosomal recombination. Further evidence for alternative mRNA splicing was also provided by the results of studies with a catfish clonal B cell line, 3B11, which was observed to express both  $\mu$  and  $\delta$  messages containing identical VDJ rearrangements (3).

cDNA homologs of  $\delta$  have subsequently been identified in Atlantic salmon, *Salmo salar* (8), and Atlantic cod, *Gadus morhua* (9). As in catfish, the  $\delta$  transcripts in these species included Cµ1 as the first C region exon and the  $\delta$  gene was found directly downstream of the µ gene. In addition to the  $\delta$ m and  $\delta$ s cDNAs identified in catfish, evidence for the expression of a catfish  $\delta$ s protein has been obtained; i.e., an anti-peptide Ab specific for the carboxyl terminus of  $\delta$ s identified (by Western analysis) a 180-kDa serum protein, the predicted size of  $\delta$ s (10).

It is now clear that the catfish *IgH* locus is complex and contains several  $\mu$  and  $\delta$  genes. The first described and most intensively studied catfish  $\mu$  gene (3, 11–13) is termed *IGHM1*, and the  $\delta$  gene immediately 3' of it is *IGHD1*. A second  $\mu$  gene, determined to be a pseudogene and therefore termed *IGHM2P*, has been described ~725 kb upstream of *IGHM1* (14, 15). It was initially believed that secreted and membrane-bound forms of catfish IgM and IgD originated from the *IGHM1* and *IGHD1* genes. This conclusion was supported by the pseudogene status of *IGHM2P*, which contains only remnants of exons C $\mu$ 1 and C $\mu$ 2 (14). Reported here is the evidence that the catfish *IgH* locus contains three  $\delta$  constant region genes or pseudogenes, each linked to a  $\mu$  gene or pseudogene. The complete characterization of two of these catfish  $\delta$  genes has led to considerable revision in our understanding of the expression of IgD in this species.

# **Materials and Methods**

#### Experimental animals

Outbred channel catfish (1-2 kg) were obtained from ConAgra (Isola, MS) and second-generation homozygous MHC-matched gynogenetic channel

<sup>\*</sup>Department of Microbiology, University of Mississippi Medical Center, Jackson, MS 39216; and <sup>†</sup>Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC 29425

Received for publication February 19, 2002. Accepted for publication June 26, 2002.

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

<sup>&</sup>lt;sup>1</sup> This work was supported by grants from the National Institutes of Health (R01AI-19530 and R01 GM62317), the National Science Foundation (MCB 9807531), and the U.S. Department of Agriculture (99-35204-7844).

<sup>&</sup>lt;sup>3</sup> Current address: Catfish Genetics Research Unit, U.S. Department of Agriculture-Agricultural Research Service, Stoneville, MS 38776.

<sup>&</sup>lt;sup>4</sup> Current address: Division of Hematology, Stanford University School of Medicine, Stanford, CA 94305-5156.

<sup>&</sup>lt;sup>5</sup> Current address: Hubbard Center for Genome Studies, University of New Hampshire, Durham, NH 03824.

catfish were obtained from the Catfish Genetics Research Unit, U.S. Department of Agriculture-Agricultural Research Service (Stoneville, MS) (16, 17) and maintained in individual tanks (18).

### Genomic library construction and screening

Genomic DNA was prepared from erythrocytes of a single outbred catfish and a recombinant catfish genomic library was created in  $\lambda$  phage. Briefly, partially *Sau*IIIA-digested DNA was size-separated by centrifugation on a sucrose density gradient and fragments of ~15–20 kb were ligated and packaged in Lambda EMBL-4 (Stratagene, La Jolla, CA). The library was screened with catfish  $\delta 2/\delta 3$  and  $\delta 6/\delta 7/\delta$  transmembrane (TM)<sup>6</sup> probes. Seven hybridizing clones were identified and phage clones,  $\lambda IgD2$ ,  $\lambda IgD3$ , and  $\lambda IgD6$ , were chosen for further analyses. Overlapping restriction fragments from each of the three recombinant genomic clones were subcloned into either pUC19 or pBluescript (Stratagene) and completely sequenced on both strands (Biotechnology Resource Laboratory, Medical University of South Carolina, Charleston, SC). Sequences were analyzed using DNAstar (DNAstar, Madison, WI) and Vector NTI Deluxe v4.0.1 (InforMax, North Bethesda, MD).

A recombinant genomic bacterial artificial chromosome (BAC) library was prepared with DNA isolated from brain tissue of a single gynogenetic catfish (S. M.-A. Quiniou, unpublished data). Briefly, high-molecular weight DNA was partially digested with *Hin*dIII enzyme and ligated into pBeloBac11 vector (19) (Invitrogen, Carlsbad, CA). Approximately 50,000 bacterial colonies were plated on Luria-Bertani plates containing 12.5  $\mu g/ml$  chloramphenicol and screened with a catfish  $\delta7$  probe. Six hybridizing clones, BAC 1–6, were identified; each contained an insert of ~130 kb.

#### Southern blotting and mapping of BAC clones

Genomic DNA was prepared from erythrocytes of individual gynogenetic or outbred fish by lysis in TES buffer (10 mM Tris (pH 8), 10 mM EDTA, 400 mM NaCl, 0.2% SDS) containing 100  $\mu$ g/ml proteinase K. After overnight digestion at 37°C the DNA was extracted with 6 M NaCl (20). BAC DNA was isolated using standard alkaline lysis. Genomic DNA (10  $\mu$ g) or recombinant BAC DNA (0.2  $\mu$ g) was digested with *Eco*RI separated on a 1% agarose gel and transferred by capillary action onto Hybond-N+ charged nylon membranes (Amersham Pharmacia Biotech, Arlington Heights, IL) using standard techniques. Hybridizations were performed in UltraHyb hybridization solution (Ambion, Austin, TX) at 42°C as per the manufacturer's recommended instructions. The membranes were washed four times at high stringency (62°C with 0.2× SSC plus 0.1% SDS) and subjected to autoradiography.

Recombinant BAC clones (clone 2 and 4) were mapped by digesting 1  $\mu$ g of DNA with *PvuI*, *MluI*, *ClaI*, *Bss*HII, *Sgr*AI, and *NotI* (New England Biolabs, Beverly, MA) to completion, either separately or in combinations, and separated by pulsed field gel electrophoresis (CHEF-MAPPER; Bio-Rad, Hercules, CA) on 1% agarose gels (Seakem LE; BioWhittaker, Rockland, ME) in 0.5× TBE along with Lambda Ladder PFG marker (New England Biolabs) using the following parameters: current, 6 V/cm; angle, 120 degrees; pulse interval ramping, 5–15 s; run time, 15 h; temperature, 14°C. The sizes of the DNA fragments were calculated using the GelExpert software (NucleoTech, Hayward, CA). The DNA was transferred and hybridized as above.

#### Probes

Probes for Southern blot and screening were amplified by PCR using AmpliTaq (PE Applied Biosystems, Foster City, CA) according to the manufacturer's recommended protocol. Typical parameters were as follows: 1 min at 95°C, 1 min at 61°C, and 2 min at 72°C for 30 cycles. The specific  $\mu$  and  $\delta$  probes (Table I) were random prime labeled with [<sup>32</sup>P]dCTP using the Megaprime labeling kit (Amersham Pharmacia Biotech).

#### Plasmids, transient transfections, and reporter assays

The enhancer-less plasmid pFVH-CAT, containing a goldfish V<sub>H</sub> promoter and the chloramphenicol acetyl transferase (CAT) reporter gene, and the positive control plasmid pFVH-CAT-ELF11, containing the catfish IgH enhancer  $E\mu3'$ , have been described by Magor et al. (11). The catfish test construct,  $E\delta6.1$ , was prepared by ligating into pFVH-CAT a 2824-bp *Hind*III/*Bam*HI fragment that contains the putative enhancer region of the *IGHM3P/IGHD3* gene; i.e., nt 5813–8637 of genomic phage AlgD6. This fragment contains seven octamer motifs and two  $\mu$ E5 motifs that match

Probes	Nucleotide Location	Accession No.	Gene
Сµ1	2133-2425	X79482	IGHM1
Cµ2	4208-4500	X79482	IGHM1
Cµ3	4960-5230	X79482	IGHM1
Intron $\psi C\mu 3 - C\mu 4$	4954-5330	AF363449	IGHM3P
IpTc1	3004-3809	AF363449	IGHM3P
Ĉδ1	734-1025	U67437	M5 (cDNA)
Сб2	1048-1288	U67437	M5 (cDNA)
Сб3	1336-1612	U67437	M5 (cDNA)
Сб7	2652-2794	U67437	M5 (cDNA)
85/86/87	2221-2885	U67437	M5 (cDNA)
$\delta 2/\delta 3$	1048-1496	U67437	M5 (cDNA)

those in the previously described catfish  $E\mu 3'$  enhancer (11). Plasmid pCMV-LUC served as an internal control for transfection efficiencies (21). Plasmid DNAs were purified using QiaFilter Plasmid Maxi kit (Qiagen, Valencia, CA).

The catfish clonal B cell line 1G8 was cultured as described (22, 23). For transfection, 3.5 pmol (10–16  $\mu$ g) of the test plasmids were introduced by electroporation into 1G8 together with 2 µg of pCMV-LUC as an efficiency control. In all samples, pUC19 was used as carrier DNA to adjust the total amount of DNA to 20  $\mu$ g. Electroporation was as described by Bengtén et al. (21) and experiments were conducted in triplicate. Expression of reporter constructs was measured 30 h posttransfection. Briefly, cells were harvested and washed in PBS and extracts were prepared in  $1 \times$ cell culture lysis reagent (Promega, Madison WI) for luciferase and CAT assays. Luminescence from pCMV-LUC was used to normalize the activity levels for transfection efficiency with luciferase activity measured in triplicate using the Promega luciferase assay system and a LumiCount (Packard Instrument, Meriden, CT). CAT activity was measured using the liquid phase assay as described by Askovic and Baumann (24). Briefly, cell extracts were incubated with 0.25  $\mu$ Ci [<sup>3</sup>H]acetyl-coenzyme A (Amersham Pharmacia Biotech) for 2 h. After incubation the <sup>3</sup>H-acetylated chloramphenicol was extracted into toluene and quantified by liquid scintillation.

#### Abs and Western blots

Anti- $\delta$ s polyclonal Ab (pAb) was produced by immunizing BALB/c mice with a multiantigenic peptide (MAP) (25) comprised of seven copies of the  $\delta$ s carboxyl terminus (FTEETIYFDENKYEQLLTAPSRP). Mice were injected i.p. with 100  $\mu$ g of peptide in CFA and three subsequent immunizations were given in IFA. Blood was collected from the tail vein. Polyclonal ascitic fluid was generated by injecting immunized mice with 0.5 ml pristane followed 2 wk later with an injection of 2 × 10<sup>6</sup> nonsecreting myeloma cells SP2/0 Ag-14 (26). Anti- $\delta$ s mAbs were generated according to standard methods (27). The  $\delta$ S MAP was synthesized by the Biotechnology Resource Laboratory (Medical University of South Carolina).

Serum from individual catfish was separated on 8% SDS-PAGE gels under reducing conditions. The proteins were electrophoretically transferred to nitrocellulose Hybond ECL (Amersham Pharmacia Biotech). The membranes were blocked with 15% fat-free milk in TBS-T (TBS plus 0.05% Tween 20) and incubated with either anti-δs pAb or anti-δs mAb 2E5 (IgG2a<sub>κ</sub>). Following four washes in TBS-T the membranes were incubated with HRP-conjugated goat anti-mouse IgG and developed using the ECL Western blot detection kit (Amersham Pharmacia Biotech) according to the manufacturer's protocol.

#### Purification of serum IgD

The IgD protein was purified from catfish serum using gel filtration and preparative SDS-PAGE. Briefly, 7–9 ml of serum was applied to an XK50/60 column of Sephacryl S-300 (Amersham Pharmacia Biotech) in 0.5 M Tris (pH 8), 0.15 M NaCl, 0.5 mM EDTA, and the IgD-containing fractions were determined by Western blot analyses. These were pooled and concentrated ~5-fold using a Centriprep YM-10 (Amicon, Beverly, MA). Twenty-five microliters (maximum volume of well) of the concentrated IgD pool were then electrophoresed under reducing conditions on a 5% SDS polyacrylamide gel. The gel was stained with Coomassie blue using standard conditions and the  $\delta$  protein band was excised. Gel slices were washed twice in 50% acetonitrile in HPLC-grade water and snap-frozen in liquid nitrogen. Samples were sent to the Harvard Microchemistry Facility (Harvard University, Boston, MA) and were subjected to proteolytic digestion and sequencing. Sequence analysis was performed by

<sup>&</sup>lt;sup>6</sup> Abbreviations used in this paper: TM, transmembrane; BAC, bacterial artificial chromosome; MAP, multiantigenic peptide; CAT, chloramphenicol acetyl transferase; pAb, polyclonal Ab.

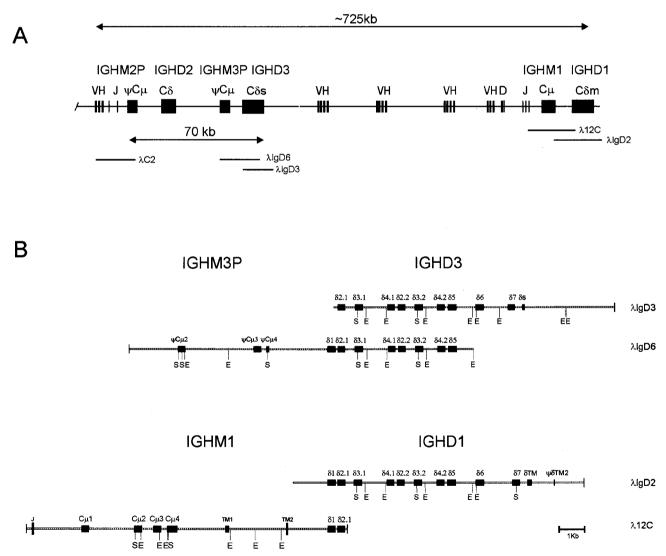
microcapillary reversed-phase HPLC nanoelectrospray tandem mass spectrometry on a Finnigan LCQ quadrupole ion trap mass spectrometer (28).

## Results

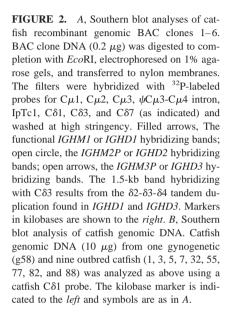
# Cloning and mapping of $\delta$ genes in the catfish IgH locus

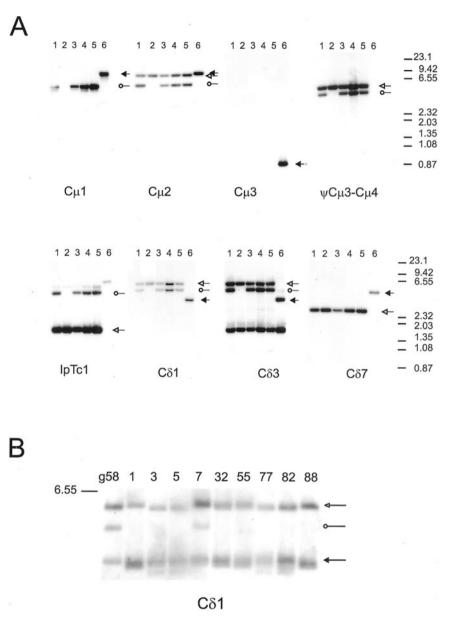
Previous cloning and sequence analysis of the catfish *IgH* locus revealed that exons 1 and 2 of an expressed  $\delta$  gene resided ~1.6 kb 3' of the *IGHM1* TM2 ( $\mu$ TM2) exon (3). To extend the analysis of this catfish  $\delta$  gene, a recombinant genomic  $\lambda$  phage library from a single outbred fish was screened with probes corresponding to the 5' and 3' regions of catfish  $\delta$  cDNA. Seven recombinant clones hybridized with both probes, and the three with the longest inserts (termed  $\lambda$ IgD2,  $\lambda$ IgD3, and  $\lambda$ IgD6; GenBank accession nos. AF363448, AF363450, and AF363449, respectively) were selected for further analysis. Complete mapping and sequencing of the inserts within these phage revealed that they represented two separate contigs of 11245 and 18486 bp, respectively, which included two distinct, albeit very similar,  $\delta$  genes (Fig. 1). Both  $\delta$  genes contained exons for  $\delta 1$ ,  $\delta 2$ ,  $\delta 3$ , and  $\delta 4$ , followed by a tandem duplication of  $\delta 2$ - $\delta 3$ - $\delta 4$ , followed by  $\delta 5$ ,  $\delta 6$ ,  $\delta 7$  and an exon encoding for either the TM region or the  $\delta s$ .

The sequence of the  $\lambda$ IgD2 insert overlapped with that of the previously sequenced clone  $\lambda$ 12C (GenBank accession no. X79482) (11); therefore, these sequences could be assembled into a contig (Fig. 1*B*) spanning the C $\mu$  (*IGHM1*) and C $\delta$  (*IGHD1*) genes, from the proximal J<sub>H</sub> region to 2 kb downstream of the *IGHD1* TM exon. The second  $\delta$  gene was characterized from a contig of two phage:  $\lambda$ IgD6 and  $\lambda$ IgD3 (Fig. 1*B*). This contig extends upstream of the  $\delta$  gene and includes three remnant exons ( $\psi$ C $\mu$ 2,  $\psi$ C $\mu$ 3, and  $\psi$ C $\mu$ 4) of a pseudo  $\mu$  gene. Interestingly, the  $\lambda$ IgD6/3 contig did not contain or overlap with the previously described catfish pseudo  $\mu$  gene (14) now termed *IGHM2P*. Consequently, there are two  $\mu$  pseudogenes in the catfish *IgH* locus, *IGHM2P* and *IGHM3P* (described in this paper). To map the organization of the multiple  $\mu$  and  $\delta$  genes present in the catfish *IgH* locus, catfish BAC clones hybridizing with probes for  $\delta$ 7 were



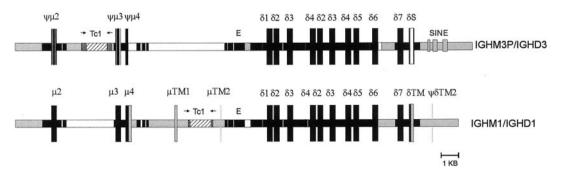
**FIGURE 1.** The catfish  $I_{gH}$  locus. A, Map showing the six identified C region genes of the catfish. Linkage between two catfish  $C_{\mu}$  genes (*IGHM1* and *IGHM2P*) was previously established (15). Linkage among *IGHM2P*, *IGHM3P*, and the two catfish C $\delta$  genes (*IGHD2* and *IGHD3*) was established by restriction mapping, pulsed field gel electrophoresis, and Southern blot analyses of recombinant BAC4. Phage clones  $\lambda$ IgD2,  $\lambda$ IgD3, and  $\lambda$ IgD6 were isolated in this study and phage  $\lambda$ 12C (3, 12) and  $\lambda$ C2 (14) have been reported previously. *B*, Physical maps of catfish genomic  $\lambda$  phage clones  $\lambda$ IgD2,  $\lambda$ IgD3, and  $\lambda$ IgD6 showing contigs linking *IGHM1* with *IGHD1* and *IGHM3P* with *IGHD3*. Restriction sites (S, *Sst*I; E, *Eco*RI) are marked. The overlapping, previously described phage clone  $\lambda$ 12C is shown for comparison. Schematics are drawn to scale.





isolated. Southern blots of six *Eco*RI-restricted BAC clones were sequentially hybridized with different probes for  $\mu$ ,  $\delta$ , and a Tc1 transposable element (Fig. 2*A*). Because  $\lambda$ IgD3,  $\lambda$ IgD6,  $\lambda$ IgD2, and  $\lambda$ 12C had been completely sequenced the identity of the hy-

bridizing bands was readily determined. The restriction digests and blot hybridizations demonstrated that one clone (BAC6) contained *IGHM1* and *IGHD1*, i.e., restriction fragments hybridizing with  $C\mu 1$ ,  $C\mu 2$ , and  $C\mu 3$  probes exhibiting sizes corresponding to the



**FIGURE 3.** Comparisons of two  $\mu\delta$  gene regions of the catfish *IgH* locus. Inverted repeats of the IpTc1/Mariner elements are marked by arrows and E marks the IgH enhancer regions. Black indicates sequence similarity (>79% identity) between the two regions, gray indicates that the sequences are different (<55% identity), and white represents areas not present in one gene but present in the other. Short interspersed repetitive elements (SINE) are marked. The *IGHM1/IGHD1* map is based on the previously described  $\lambda$ 12C (12) and  $\lambda$ IgD2. Schematics are to scale.

expected fragment sizes calculated from the  $\lambda 12C$  sequence. Likewise, the sizes of fragments hybridizing with C $\delta 1$ , C $\delta 3$ , and C $\delta 7$  probes correspond to the expected sizes as determined from the  $\lambda IgD2$  sequence. The other five clones (BAC1–5) each contained the *IGHM3P* gene (Fig. 2A, open arrows) and BACS 1, 3, 4, and 5 also contained the previously described *IGHM2P* gene (Fig. 2A, open circle). For example, the 5.4-kb band hybridizing with  $\psi C\mu 3$ -

Cμ4 was distinctive for *IGHM3P* as predicted by the  $\lambda$ IgD6 sequence, while the 4.8-kb hybridizing Cμ2 band was distinctive for *IGHM2P* as predicted from the previously described  $\lambda$ 2C sequence (14). Thus, the two  $\psi\mu$  genes are linked within 130 kb of each other. Furthermore, both  $\psi\mu$  genes are linked to  $\delta$  sequences; the linkage of *IGHM3P* and *IGHD3* was established by sequencing of  $\lambda$ IgD6 (Fig. 1*B*) and the linkage of *IGHM2P* and *IGHD2* was

CTGAAACTGGAGACTTCTTCCATAAATGTTAAATAAATGTCTCCTTATAGAAAGCTTTACAATGTCAATGATTGTAAACATTTTTTGA	3903	λ12C
AGCT. T. A	1558	λIgD6
AGCTT.A		0
TAAATAACACATTTTTTTAAAGTGTTTATTATTAGGCTTATAGAATGTCAACTACACCAGTACCTGCA-TAATTGTTACTATAG	3986	λ12C
	1645	λIgD6
		-
AAACAATAACAATATCGAGTGCATTAATACAAAATATAGATAATTTAATTAA	4086	λ12C
	1741	λIgD6
	9289	λ2C
${\tt GTA-TAAAAATA-GCATTTTTGTGACCTT-ATATACTGTGGCAACACAACCTATTTCTAAACTGCAATACCTGCTGTGCCCATTGCTGCTG-ATATGGAT$	4182	λ12C
<b>TT</b> TC	1841	$\lambda$ IgD6
TTTC	9389	λ2C
${\tt CTGTGTGTTTGTTTGCAGTTGAGAGAGAGAGAGAGCTCCATGCATCTCTGCTTCTAACAACTCCAACCCAAACAGAAATAGAAATAGAACAGCTACCTTCGT$	4280	λ12C
CGCTCCC	1937	λIgD6
T	9483	λ2C
VERELHASLLLTTPTQTEIDNGTATFV		λ12C
VETELPASLLLTTPTQTEIPNGTAT		λIgD6
VETELPASLLLTTPTQTEIPNGTAT		λ2C
CTGCTTAGCTACACCATTTTCACCTAAATCACACACATTTAAGTGGACTCTTGAAAA-GACAG-ACATCAGTAATAAGGTCAAAGAAGAAACACACATCATGAAACAACAACAACAACAACAACAACAACAACAACAACAA	4368	λ12C
	2018	λIgD6
	9561	λ2C
CLATPFSPKSHTFKWTLE KTDIS NKVKEN		λ12C
SA*LHNFHLNHTHLSGLVKAW SSK IRS		λIgD6
SA*LHNFHLNHTHLSGLVKACHYFRSLFNRS		λ2C
- ATAGTAAGCCAGAATAAAGGTAACTT-CACTGCCATAAGTGTTTTGGAACTCAGCGCCAGTGAATGGACAAGCTCAACTTCTCCCAGTTAAGTGCGAATTGAACTCAGCGCAATTGAACTGCAAATTGAACTGCAAATTGAACTGCAAATTGAACTGCAAATTGAACTGCAAATTGAACTGCAAATTGAACTGCAAATTGAACTGAACTGAATGGACAAGTGCAAATTGAACTGCAAATTGAACTGCAAATTGAACTGAACTGAATGGACAAGTGCAAATTGAAATGGAACTGAATGGAATTGAAATGGAACTGAATGGAATGGACAAGTGCAAATTGAAATGGAAATTGAAATGGAAATGGAAATTGAAATGGAAATTGAAATGGAAATTGAAATGGAAATGGAAATGGAAATTGAAATGGAAATTGAAGTGCAAATGGAAATGGAAATGGAAATGGAAATGAAGTGCAAATGGAATTGAAGTGCAAATTGAAATGGAATGAACTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATTGAAGTGCAATGGAATTGAAGTGCAATGGAATGGAATGGAAATGGAAATGGAATGGAATGGAATGGAATGGAATGAATGGAATGGAATTGAAGTGCAAATTGAAGTGGAATGGAATGGAAATGGAAG	4466	λ12C
-ATAGTAAGCCAGAATAAAGGTAACTT-CACTGCCATAAGTGTTTTGGAACTCAGCGCCAGTGAATGGACAAGCTCAACTTCTCCAGTTAAGTGCGAATT T. TT, TG, G, C, C, T, A, A, A, GG, T, T		
$T \dots TT \dots T \dots G \dots G \dots G \dots G \dots G \dots \dots G \dots \dots \dots \dots$	2113	λIgD6
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671	λIgD6 λ2C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671	λIgD6 λ2C λ12C
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671	λIgD6 λ2C λ12C λIgD6
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671	λIgD6 λ2C λ12C
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671	λIgD6 λ2C λ12C λIgD6
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671	λIgD6 λ2C λ12C λIgD6 λ2C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559	λIgD6 λ2C λ12C λ12C λIgD6 λ2C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187	λIgD6 λ2C λ12C λIgD6 λ2C λIgD6 λ12C λ12C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744	λIgD6 λ2C λ12C λIgD6 λ2C λIgD6 λ12C λ12C λ12C λ12C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744	λIgD6 λ2C λ12C λIgD6 λ2C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744	λIgD6 λ2C λ12C λIgD6 λ2C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744	λIgD6 λ2C λ12C λIgD6 λ2C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744	λιgD6 λ2C λ12C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12D λ2C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744 4658	λιgD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ12C λ12C λ12C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C λ13D6 λ2C λ12C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744 4658 2286	λιgD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744 4658 2286	λιgD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744 4658 2286 9843	λιgD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744 4658 2286 9843 4710	λιgD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ12C λ13D λ12C λ12C λ12C λ13D λ2C λ12C
$\begin{array}{c} TTT. \dots TG. \dots .G^{-} \dots C. \dots \dotsA. T. A. A. A. A. AGG. \dots T. \dots .T. \dots T. \dots T TTT TG. \dots .G^{-} \dots .C. \dotsA. T. A AA AGG. \dots TT T T T T T T T T A A$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383	λιgD6 λ2C λ1gC λ1gD6 λ2C
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383	λιgD6 λ2C λ1gC λ1gD6 λ2C
$\begin{array}{c} TTT. \dots TG. \dots .G- \dots .C. \dots \dotsA.T. A. A. A. \dots AGG. \dots T. \dotsT. \dots T. \dots T T T T T T T T T T T T T T T T A A A T A A A A A T$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878	λιgD6 λ2C λ12C λ12C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C
T. TT T. G G CA.T. A A A GG T T T T T T G G CA.T. A A A GG T T T T T T G G CA.T. A A A A TGG T T T T T T T T T A A TGG T T T T T T T G G G G G G A G T T T T A A A A A A A A TGG T T T T T T T T G G A C T T T T T T T C. A G A G A C T A A A A A A A A A T T T	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800	λιgD6 λ2C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12D λ12C λ12D λ12C
$\begin{array}{c} TTT. \dots TG. \dots .G- \dots .C. \dots \dotsA.T. A. A. A. \dots AGG. \dots T. \dotsT. \dots T. \dots TTT. $	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800	λιgD6 λ2C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12D λ12C λ12D λ12C
$\begin{array}{c} T. TT \dots T. G. \dots G^{-} \dots C. \dots \dots P^{A}, T. A \dots A \dots A \dots GG \dots T \dots $	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483	λιgD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C
$ \begin{array}{c} T. TT \dots T. G. \dots G- \dots C. \dots \dotsA. T. A. \dots A. \dots A. \dots GG. \dots T. \dots T. \dots T. \dots T. \dots T. \dots T. \dots $	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870	λιgD6 λ2C λ12C λ12C λ1gD6 λ2C
$\begin{array}{c} T. TT \dots T. G. \dots G^{-} \dots C. \dots \dots P^{A}, T. A \dots A \dots A \dots GG \dots T \dots $	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870	λιgD6 λ2C λ12C λ12C λ1gD6 λ2C
$\begin{array}{c} TTT TGG CA. T. A. A. A AGG TT T T T TT A AGG TT T T TTTA. T. G TT$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870 2578	λιgD6 λ2C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12C λ12D λ12C
$ \begin{array}{c} TTT TGG CA. T. A. A. A A GG T T T T T T T T A A$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4800 2483 2678	λιgD6 λ2C λ12C λ1gD6 λ1gD6 λ1gD6
$\begin{array}{c} T. TT, \dots T. G, \dots G- \dots C, \dots \dotsA, T. A, \dots A, \dots A, TGG, \dots T, \dots T, \dots T, \dots T, \dots T, \dots T, \dots T$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870 2578 2678 2778	λιgD6 λ2C λ12C λ1gD6 λ1gD6 λ1gD6
$\begin{array}{c} T. TT. \dots T. G. \dots G^{-} \dots C. \dots G^{-} \dots A. T. A. \dots A. \dots A. \dots GG. \dots T. \dots T. \dots T. \dots T. \dots T. \dots T. \dots $	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4800 2578 2678 2578 2678 2778	λιgD6 λ2C λ12C λ12C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ2C λ1gD6 λ12C λ1gD6 λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ1gD6 λ2C λ12C λ1gD6 λ1gD6 λ1gD6
$ \begin{array}{c} T. TT TGG^{-} CA. T. A A A GG T T T T T T A. T. G T A. T A A. T. G T T T T T T A. A A. T A. T A. A A. T A. A A. T A. A A. T A. A A A. T A. A A. A A. A A. A$	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870 2578 2678 2678 2878 2978	λιgD6 λ2C λ12C λ12C λ1gD6 λ1gD6 λ1gD6
$ \begin{array}{c} T. TT. \dots T. G. \dots G- \dots C. \dotsA. T. A. \dots A. \dots A. \dots A. \dots GG. \dots T. \dots T. \dots T. \dots T. \dots T. \dots T. \dots $	2113 9671 4559 2187 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870 2578 2678 2778 2678 2778 2878 2978 3078	λιgD6 λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ1gD6 λ1gD6
$ \begin{array}{c} T. TT, \dots T. G, \dots G^{-} \dots C, \dots Q^{-} A, T. A, \dots A, \dots A, \dots A, \dots GG, \dots T, \dots T, \dots T, \dots T, \dots T, \dots T, \dots $	2113 9671 4559 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870 2578 2678 2678 2778 2878 2978 3078 3178	λιgD6 λ2C λ12C λ12D λ2C λ12D6 λ2C λ12D6 λ2C λ12D6 λ2C λ12C λ12D6 λ2C λ12C λ12D6 λ2C λ12C λ12D6 λ2C λ12C λ12D6 λ2C λ12C λ12D6 λ2C λ12C λ12D6 λ12C λ12D6 λ12C λ12D6 λ12C λ12D6 λ12C λ12D6 λ12C λ12D6 λ12D
$ \begin{array}{c} T. TT. \dots T. G. \dots G- \dots C. \dotsA. T. A. \dots A. \dots A$	2113 9671 4559 9744 4658 2286 9843 4710 2383 9878 4800 2483 4870 2578 2678 2678 2778 2878 2978 3078 3178	λιgD6 λ1gD6 λ2C λ12C λ1gD6 λ2C λ12C λ1gD6 λ1gD6 λ26 λ26 λ26 λ26 λ26 λ26 λ26 λ2

**FIGURE 4.** Nucleotide and amino acid sequence alignments of the catfish IgH genes and Tc1 elements. *A*, Comparison of the  $\psi C\mu^2$  and functional  $C\mu^2$  exons. The *IGHM3P* sequence initiates 5' of  $\psi C\mu^2$  and nt 1459–3218 of  $\lambda$ IgD6 are compared with nt 3881–4870 of  $\lambda$ 12C (*IGHM1*) (12) and nt 8996–9878 of  $\lambda$ 2C (*IGHM2P*) (14). Exon boundaries are marked by  $\downarrow$  and the start of the exons are labeled C2. Periods indicate identity and dashes show gaps. Stop codons are indicated by asterisks. The inferred amino acid sequences of  $C\mu^2$  and the  $\psi C\mu^2$  exons are shown *below* the second base of the codon; those that are identical to the functional  $C\mu^2$  are in bold. The nucleotide numbers and  $\lambda$  phage clone names are to the *right*.

suggested by the presence of the "second"  $C\mu 2$ ,  $C\delta 1$ , and  $C\delta 3$ hybridizing bands (Fig. 2A, open circle) found in BAC1, 3, 4, and 5. In addition, restriction mapping with *PvuI*, *MluI*, *ClaI*, *Bss*HII, *Sgr*AI, and *Not*I enzymes and Southern blot analyses of BAC4 demonstrated that *IGHM2P* and *IGHD2* are linked on an 11-kb *Bss*H II fragment and *IGHD2* and *IGHD3* are linked within 70 kb of each other (Fig. 1A). Comparison of the recombinant BAC and  $\lambda$  phage hybridization results combined with Southern blot analysis of genomic DNA from 10 individual catfish hybridized with a  $\delta 1$  probe demonstrated that *IGHD1* and *IGHD3* are present in all catfish studied. Furthermore, they indicate that all Ig C region genes are readily detectable by Southern blot and are included in the map shown in Fig. 1A, albeit with some allelic variation in the hybridizing *IGHD2* band (Fig. 2B). Thus, this evidence clearly indicates that the catfish *IgH* locus contains three  $\mu$  and  $\delta$  genes.

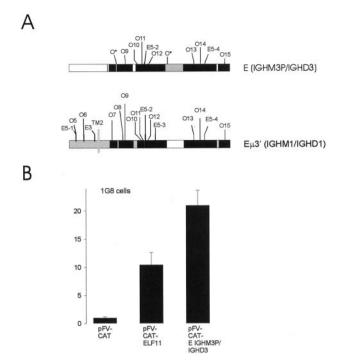
## Catfish $\mu$ and $\delta$ genes

The complete sequence of  $\lambda$ IgD2,  $\lambda$ IgD3, and  $\lambda$ IgD6 permits direct comparisons of the *IGHM1* vs *IGHM3P* and the *IGHD1* vs *IGHD3* genes (Fig. 3). The three  $\mu$  genes are very different, i.e., the *IGHM1* gene is clearly functional (12), whereas *IGHM3P* and *IGHM2P* (14) are highly degenerate pseudogenes. *IGHM3P* contains three remnant  $\mu$  exons,  $\psi$ C $\mu$ 2,  $\psi$ C $\mu$ 3, and  $\psi$ C $\mu$ 4, with a

remnant of a Tc1/mariner transposable element located between  $\psi C\mu 2$  and  $\psi C\mu 3$ . The  $\psi C\mu 2$  exon of *IGHM3P* is very similar but not identical to the  $\psi C\mu 2$  of *IGHM2P* (Fig. 4). Both contain five stop codons and, except for the first conserved 25 amino acids, are out of frame as compared with the functional  $C\mu 2$  of *IGHM1*. In contrast, the introns preceding the  $C\mu 2$  exons are highly conserved. The location of the Tc1 element of IGHM3P is identical to the location of the Tc1 element termed IpTc2 described in IGHM2P (14). However, the Tc1 element of IGHM3P is more complete because it possesses coding blocks A through F and is flanked by terminal repeats (29). IpTc2 lacks terminal repeats and contains only coding block E and part of block D. The  $\psi C\mu 3$  and  $\psi C \mu 4$  remnant exons of IGHM3P are located directly 3' of the Tc1 element and Southern blot analysis suggests that they are also present in IGHM2P. The  $\psi C\mu 3$  remnant is in frame and 55% identical at the nucleotide level to functional  $C\mu 3$ , with conservation of encoded amino acid residues throughout the domain. However, an 18-aa stretch in the second half of the C $\mu$ 3 domain, including the Cys necessary for the interchain disulfide bond, is missing. The  $\psi C \mu 4$  exon consists of only 108 nt corresponding to the first third of the functional C $\mu$ 4 (79% identical at the nucleotide level) but does not encode any Cys residues (Fig. 5). No identifiable  $\mu$ TM exons are found 3' of the C $\mu$ 4 exon remnant, but a region showing

ACAGCCACAGGTGAAAATAACTGGACCTTCCACGAGGAGATTCTGATCAAAAGAGCC-GGCCAGCT-C-GAGTGCAGGGCCGAGGGAGACACGGG 		
AAGAGCATTAAATGGCTTATTGGAAATAGAGAGATCTCTTCTCTATCA-AATCTATCTTCGAAAACGACGGTTTCACTCCAAACCCACAGCGG CGTTT.G.TTGT.ATTC.GGAG.CA.CTACATA.TGGA.TC.T.A.C K S I K W L I G N R E I S S L S N L S S K T T V S L Q T H I G N V L I V K R F L L L R N S H I Y K H S G H T H H D G	т. 5002	
↓ ANGAGTGGATCAATGGCACCGAATTCATCTGTGAGGTGGAACATGAAGCATTCACTCAACAGTATGAAAAAGTAACCTTCAAAAGAGAAAATGGTA .TGAA E E W I N G T E F I C E V E H E A F T Q Q Y E K V T F K R E N M S G A M A P K S P K R K		
АТССТТТТАСТАСТАТАТАСАТАТТССТССТСТТААТТТСТАТСАААААА	А 5150 Жааса 5458	$\lambda IgD$ $\lambda 12C$
$ \downarrow_{C4} \\ TATTTTGA-CTTAATTGTGATTTCATGTTTATAGGCAATCCGGAATTCCCCAAGGTTTACTTGCTCGCTC$	ACCC 5557	λ12C
TGACTTGCTATGTTAAAGACTTCTACCCTAAGGAGGTGGCTGGTCTTGGCTTGTTAACGATAAACAAGTGGAAGAAGTGGTCGGCTATGAGCAGA GGTATACAC	5389	
CACTGCAGTTATCGACAGAAACAACCTCTTTTCAGTGTACAGCCAGC		λ12C λ12C
- TATCATGAGTCCATCAAGGACTGTGTGCGCCACATATCCAGATCCATCGCTAAAGATTCAAAAACGCCCACCTTAGTGAATCTCACCCTGACCAA1 Y H E S I K D C V R H I S R S I A K D S K T P T L V N L T L T N		λ12c λ12c

**FIGURE 5.** Alignment of the nucleotide sequence of catfish *IGHM3P* and *IGHM1* genes comparing the  $\psi$ Cu3,  $\psi$ Cu4, and functional C $\mu$ 3 and C $\mu$ 4 exons. Nucleotides 4719–5389 of  $\lambda$ IgD6 and nt 4871–5883 of  $\lambda$ 12C are shown. Exon boundaries are marked by  $\downarrow$  and the start of the exons is labeled C3 or C4. Periods indicate identity and dashes show gaps. The inferred amino acid sequences of the  $\psi\mu$  exons and the C $\mu$ 3 and C $\mu$ 4 exons are shown *below* the second base of the codon; those that are identical are in bold. The nucleotide numbers and  $\lambda$  phage clone names are to the *right*.



**FIGURE 6.** *A*, Map of the IgH chain enhancer regions of *IGHM3P/ IGHD3* and *IGHM1/IGHD1*. Black indicates sequence similarity (>79% identity) between the two regions, gray indicates sequence differences (<55% identity), and white represents areas not present in one gene but present in the other. O, Octamer; O\*, octamers only present in *IGHM3P/ IGHD3*; E3,  $\mu$ E3 motif; E5,  $\mu$ E5 motif. Octamers are numbered by their order of occurrence in  $\lambda$ 12C (according to Ref. 11). *B*, Relative CAT activity resulting from transient transfection of catfish IgH enhancer regions into catfish 1G8 B cells. Constructs contained the previously described catfish E $\mu$ 3' enhancer (pFV-CAT-ELF11) (11) or the catfish IGHM3P/ IGHD3 enhancer region (pFV-CAT-E $\delta$ 6.1). Results are the average of two transfections conducted in triplicate. CAT activities are normalized in correlation to the activity of the cotransfected CMV-LUC plasmid and given as mean  $\pm$  SD, relative to the transcriptional activity of the empty vector control (pFV+CAT), which is set to 1.

strong sequence conservation with the  $E\mu3'$  transcriptional enhancer of *IGHM1* (11) is present (Figs. 3 and 6A). This *IGHM3P*associated enhancer region contains two octamers and a  $\mu$ E5 motif (O10, O11, and E5-2) that are known to comprise the minimal functional unit of  $E\mu3'$  critical for enhancer function (30). In addition, this putative enhancer contains five other motifs conserved in the  $E\mu3'$  enhancer: four octamers and one  $E\mu5$  motif, plus two additional unique octamer motifs. Furthermore, this putative enhancer was tested for activity by transient transfection. To this end a *Bam*HI/*Hin*dIII fragment encompassing nt 5813–8637 of  $\lambda$ IgD6 was cloned into the pFV-CAT reporter plasmid (11). This catfish IGHM3P/IGHD3 enhancer construct was transfected by electroporation into the catfish clonal autonomous B cell line 1G8 together with the control plasmid pCMV-LUC and assessed for enhancer activity. This fragment exhibited strong enhancer activity, which is greater than that of the catfish E $\mu$ 3' (pFV-CAT-ELF11; 11) from the *IGHM1* gene (Fig. 6*B*).

In marked contrast to the major sequence differences found within the three  $\mu$  genes, the two  $\delta$  genes compared here are highly conserved. They exhibit a high degree of identity at both the exon and intron levels (Fig. 3 and Table II) and both appear to be functional, because all exons have consensus splice donor and acceptor signals and are in frame. The duplicated exons within each gene also exhibit strong homologies with nucleotide sequence identities ranging from 91 to 100% (Table II). The only exception involves comparing  $\delta 4.1$  to  $\delta 4.2$  in either locus:  $\delta 4.1$  is 5 aa shorter than  $\delta 4.2$  and they are 76 and 62% identical at the nucleotide and amino acid levels, respectively. The splice acceptor/donor sequences of the exons in IGHD3 are identical to those found in the IGHD1 gene, except for the acceptor splice site of  $\delta 6$ . A different  $\delta 6$  acceptor splice site in IGHD3 appears to be formed as a result of 7 nt differences (underlined), TTTGTATTTATTCAG/AC as compared with TATTTGTAG/GCTCAGAC. The shift in splice site usage (confirmed by cDNA sequencing) adds two additional amino acids to the exon. This IGHD3 &6 splice site reflects the introduction of an AG dinucleotide and the preferential use of the more 5' AG as a splice acceptor site (31, 32). In toto there are 24 insertions/deletions between these two  $\delta$  genes, with most being short, i.e., 14 are <5 bp. The largest insertion was found in the *IGHD1* gene where the  $\delta 6/\delta 7$  intron contains an 809-bp insertion flanked by TATA repeats (Fig. 3).

The most significant differences between the catfish *IGHD1* and *IGHD3* genes reside in their terminal exons (Fig. 7A). The *IGHD1* gene ends with an exon encoding a typical Ig-like TM region, whereas the *IGHD3* gene terminates with the shorter  $\delta$ s exon. Thus, the truncated  $\delta$ s cDNA sequences previously obtained were not derived by alternative RNA processing from a primary transcript of the *IGHD1* gene (3). Interestingly, the nucleotide sequences encoding the first 15 aa of the  $\delta$ s and  $\delta$ TM exons are almost identical and, when  $\delta$ s and  $\delta$ TM are aligned, the nucleotide series encoding the last 9 aa of  $\delta$ s match with the 3' untranslated region of  $\delta$ TM, albeit in the second reading frame (Fig. 7A). Thus,  $\delta$ s differs from  $\delta$ TM primarily by a 121-bp deletion. The presence

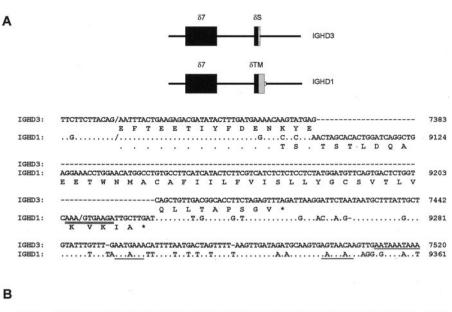
Table II. Comparisons of the nucleotide and amino acid sequences of catfish  $\delta$  exons and the nucleotide sequences and lengths of  $\delta$  introns

IGHD3 vs IGHD1		Tandem Duplication Within           3 vs IGHD1         IGHD3		Tandem Duplication Within <i>IGHD1</i>			IGHD3 vs IGHD1		Length (nt)			
Exon	nt <sup>a</sup>	aa <sup>a</sup>	Exon	nt <sup>a</sup>	aa <sup>a</sup>	Exon	nt <sup>a</sup>	aa <sup>a</sup>	Intron	nta	IGHD3	IGHD1
δ1	97.4	92.2							δ1/δ2	100	82	82
δ2.1	98.3	95.9	δ2	92.5	86.7	δ2	97.3	95.9	δ2.1/δ3.1	98.3	384	344
δ3.1	98.0	94.1	δ3	99.0	95.9	δ3	100	100	δ3.1/δ4.1	97.0	971	966
δ4.1	100	100	δ4	73.4	58.5	δ4	76.0	62.1	δ4.1/δ2.2	100	130	130
δ2.2	94.2	89.2							δ2.2/δ3.2	97.6	372	368
δ3.2	98.4	96.1							δ3.2/δ4.2	83.6	990	566
δ4.2	100	100							δ4.2/δ5	98.3	119	119
δ5	96.3	90.9							δ5/δ6	95.9	751	778
δ6	94.4	89.6							δ6/δ7	50.2	923	1086
δ7	95.8	92.7							$\delta 7/\delta ter^b$	91.6	312	300

Downloaded from http://www.jimmunol.org/ by guest on August 8, 2022

<sup>b</sup> Terminal exon =  $\delta S$  in *IGHD3* and  $\delta TM$  in *IGHD1*, respectively.

FIGURE 7. A. Schematic and alignment of the nucleotide and inferred amino acid sequences of catfish &s and &TM. Black indicates sequence similarity; hatching indicates the sequence that encodes only for  $\delta s$ . The  $\delta TM$  sequence, which is absent in  $\delta s$ , is shown in gray. The splice between  $\delta 7$  and the terminal exons is marked by "/." Identities are indicated by periods, gaps by dashes, and stop codons by asterisks. Inferred amino acids are below the second base of each codon. The potential splice site in  $\delta TM$  is double-underlined and polyadenylation sites are underlined. B, Sequence of the  $\psi\delta TM2$ exon found in IGHD1. The potential splice acceptor site is marked by "/." Nucleotide numbers are shown to the right.

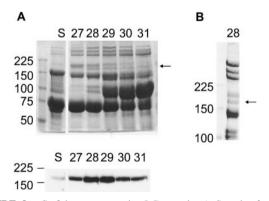


IGHD1: 10042//CCCCCTTCTTGATCTTGAATTAACTTTAAGAACAACAG/GTCAAATTAAGCAATTTAGTCAAGAAATTAGAGGG 10113 V K \*

of only a single  $\delta$ TM exon was unexpected because the TM regions of most other Igs, except IgA, are encoded by two exons (reviewed in Refs. 1 and 2). A potential 2-aa  $\delta$ TM2 (VK) exon is found 857 nt downstream of *IGHD1* (Fig. 7*B*) and a potential splice donor site is located in the  $\delta$ TM exon (marked in Fig. 7*A*). However, RT-PCR protocols performed using forward primers for Cµ1 and various  $\delta$  exons ( $\delta$ 1,  $\delta$ 2,  $\delta$ 3,  $\delta$ 5,  $\delta$ 6, and  $\delta$ 7) in combination with reverse primers for regions 3' of  $\delta$ TM2 yielded no products; therefore, it was concluded that  $\delta$ TM2 is a pseudoexon, designated  $\psi\delta$ TM2.

To determine whether the *IGHD3* gene is expressed, polyclonal mouse antisera (pAb anti- $\delta$ s) and a mAb (2E5) were generated against a MAP of the catfish  $\delta$ s carboxyl terminus. Initial Western blot analyses revealed that both Abs reacted with a serum protein of ~180 kDa. Moreover, this reactivity could be completely in-

hibited by the addition of excess synthetic peptide (10). Identification of the 180-kDa protein as  $\delta$  was confirmed by protein purification and peptide sequencing. Fig. 8*A* shows a SDS-PAGE and a corresponding Western blot of fractions obtained by gel filtration of catfish serum on Sephacryl S-300. Aliquots of concentrated fraction 28 were electrophoresed on 5% SDS-PAGE (Fig. 8*B*), and the presumed  $\delta$  protein band was excised from each lane and subsequently sequenced by microcapillary reversed-phase HPLC nanoelectrospray tandem mass spectrometry. Table III lists the  $\delta$ peptide sequences obtained by mass spectrometric analysis. Three of the peptides, one each from  $\delta 1$ ,  $\delta 5$ , and  $\delta 6$ , could only have originated from *IGHD3*; i.e., each contains specific amino acid residues encoded by the exons of that gene. Amino acids that are specific and conserved for the  $\delta 2.1$ ,  $\delta 2.2$ ,  $\delta 3.2$ , and  $\delta 4.2$  exons found in both  $\delta$  genes were also identified (Table III), suggesting



**FIGURE 8.** Catfish serum contains IgD protein. *A*, Samples from unseparated serum (S) and fractions 27-31, obtained by gel filtration of catfish serum on Sephacryl S-300, were separated under reducing conditions by 8% SDS-PAGE and stained with Coomassie blue (*top*). Western blot analysis was performed with a duplicate gel using the anti- $\delta$ s mAb 2E5 (*bottom*). *B*, Aliquots from fraction 28 were separated under reducing conditions by 5% SDS-PAGE. The  $\delta$ s-containing band (as indicated by the arrow) was excised and sequenced by tandem mass spectrometry.

Table III. Tandem mass spectrometry peptide sequences corresponding to the inferred amino acid sequences of catfish  $\delta$  exons

Sequences	Corresponding δ Exon
TFGALSQLSINAEQWNEGTAFTCK	$\delta 1^a$
IFSQTWSK	δ1
AEPTSKPLIR	$\delta 2.1^b$
LEKPGLMSVLTDSEVTASCVVETVHNTK	δ2
NWQT <u>I</u> K	$\delta 2.2^b$
TINIL/EPVQK	δ2/δ3
TPTVVIR	δ3
GDS <u>A</u> VLECAAR	$\delta 3.2^{b}$
SKQWTSNSIENLF/GDPSVEL <u>LVI</u> SS <u>V</u> D <u>K</u>	$\delta 3/\delta 4.2^{b}$
VSDQD <u>PLKP</u> V <u>E</u> K	$\delta 4.2^b$
STSICA/VTPDFAQK	$\delta 4.2^b/\delta 5$
AQVYLLAPSISDMR	δ5
LNDF <u>S</u> IVWK	$\delta 5^a$
ILSPSDDDLSGVR	δ6
E <u>DG</u> TFSCVVSH	$\delta 6^a$

<sup>a</sup> Underlined amino acids are found only in exons of the IGHD3 gene.

<sup>b</sup> Underlined amino acids are unique to that specific duplicated exon of both *IGHD1* and *IGHD3* genes.

that the duplicated  $\delta$  exons are expressed. Based on densitometry of Coomassie blue-stained gels the concentration of  $\delta$  protein in catfish serum is ~0.04 mg/ml.

## Discussion

The studies reported in this paper, in conjunction with previous reports (3, 12, 14, 15), reveal a previously unexpected complexity in the catfish IgH locus; i.e., three  $\mu$  genes and three  $\delta$  genes are present. These almost certainly arose as the result of large gene duplication events. The 3'-most  $\mu$  and  $\delta$  genes in the locus (IGHM1 and IGHMD1) are fully functional and encode H chains of catfish IgM and IgD (3, 11-13). In the 5' region of the locus (~725 kb upstream of IGHM1), the previously reported  $\psi\mu$  gene (IGHM2P), associated with a germline-fused VDJ and additional  $V_{\rm H}$  and  $J_{\rm H}$  elements (14, 15), was found to be linked to a  $\delta$  gene (IGHD2) which has not yet been sequenced. Downstream of IGHD2 a third IgH chain gene cluster was identified, which contains a second  $\psi\mu$  gene (IGHM3P), a strong enhancer, and a functional  $\delta$  gene (*IGHD3*). The presence of transposable elements, particularly Tc1/mariner-like elements (which are highly represented in catfish) (33), within these IgH genes is quite likely related to the duplication events that produced the complex structure of the catfish IgH locus. It has been suggested that remnant Tc1 elements, mobilized by transposases encoded elsewhere in the genome, played a role in the evolution of the vertebrate IgH locus (34) and may also have been involved in the large scale duplications of the catfish IgH locus. It is also likely that Tc1 elements were responsible for the disruptions of structure that have occurred in the two pseudo  $\mu$  genes. For example, the deletion of  $\mu$ TM1,  $\mu$ TM2, and part of the E $\mu$ 3' enhancer region in *IGHM3P* is presumably the result of excision of the IpTc1 originally present in this region, followed by DNA repair.

In stark contrast to the great disparity between the catfish duplicated  $\mu$  genes, the two duplicated  $\delta$  genes (*IGHD1* and *IGHD3*) are remarkably conserved. The major difference between the two genes involves the terminal exons: the IGHD1 gene encodes a polypeptide with a typical Ig cytoplasmic tail, whereas the terminal exon of the IGHD3 gene encodes a terminal secretory tail. The finding of cDNAs containing these two alternative terminal exons clearly indicates that both genes are transcribed (3). In a previous report it was suggested that catfish IgD, as now considered to be encoded by the IGHD1 gene, may function as a B cell Ag receptor. Its sequence relatedness, location immediately downstream of the functional  $\mu$  gene, and coexpression by alternative mRNA processing pathways with  $\mu$  in some B cells are consistent with catfish  $\delta$  being a true homolog of mammalian  $\delta$  (3). The only evidence to date for a functional IgD TM protein is the existence of full-length δm transcripts encoded by the IGHD1 gene. All catfish full-length  $\delta$  cDNAs identified to date consist of a rearranged VDJ spliced to  $C\mu 1$ , the  $\delta 1-\delta 7$  domains, and  $\delta TM$ . The catfish  $\delta TM$  encodes for a typical Ig-like TM region and there is no indication from the sequence data that it could not be expressed or function as an Ag receptor on the B cell surface. The finding of a second functional  $\delta$  gene encoding the previously reported putative secreted form of  $\delta$  (3) was unexpected. Coomassie blue staining and Western blot analyses using pAb and mAb specific for the C-terminal  $\delta s$  revealed that  $\delta$  polypeptides are found in normal catfish serum at levels similar to IgD in humans (Fig. 8 and Refs. 10 and 23). The origin of serum IgD from the IGHD3 gene was subsequently confirmed by protein purification and tandem mass spectrometry peptide sequencing, although the N-terminal sequence of serum IgD was not determined. Catfish full-length  $\delta s$  transcripts have also not yet been identified. 5'-RACE approaches have identified the longest  $\delta s$  clones as beginning with 110 bp of an as-yet-unidentified

exon spliced to  $\delta 1$  (E. Bengtén, unpublished observations). Because IGHD3 is located 5' of the IGHM1 and IGHD1 genes, \deltas may not use the same VDJ rearrangements used by IGHM1 and IGHD1 to form  $\mu$ m,  $\mu$ s, and  $\delta$ m. Therefore, if  $\delta$ s contains a VDJ rearrangement it could use the V<sub>H</sub> segments 5' of the IGHM2P gene (14). If this is the case then  $\delta m$  and  $\delta s$  may not be produced by the same cell. However, it cannot be ruled out that  $\delta m$  and  $\delta s$ share the same VDJ rearrangements by a mechanism of intrachromosomal homologous recombination as demonstrated for Ig genes in mouse hybridomas (35, 36). It is also possible that separate short sterile transcripts of VDJ and CH genes could be joined by a transplicing mechanism as documented for Ig isotypes in transgenic mouse models (37, 38). At this time it is not known which of these possibilities are operative in channel catfish. Similarly, it is not known whether the IGHD2 gene, identified only by Southern blot hybridization (Fig. 2), encodes a functional polypeptide.

From the results of previous studies and those presented here, it can be concluded that the downstream C genes of the catfish *IgH* locus (*IGHM1* and *IGHD1*) are most likely fully functional, encoding  $\mu$ m,  $\mu$ s, and  $\delta$ m. The upstream genes of the locus (*IGHM2P*, *IGHM3P*, *IGHD2*, *IGHD3*) most probably resulted from a large duplication event that included most (if not all) of the original locus and that was followed by a second duplication of a smaller region. The upstream genes derived from hypothetical ancient duplication events include multiple Ig elements that are apparently expressible: these include V<sub>H</sub> genes, a J<sub>H</sub> element, a fused V<sub>H</sub>DJ<sub>H</sub>, and, as shown here, an intact  $\delta$ s gene and a functional enhancer. Whether or not the *IGHD3*-encoded  $\delta$  polypeptides function in the immune response is currently unresolved.

# References

- DuPasquier, L., and M. Flajnik. 1999. Origin and evolution of the vertebrate immune system. In *Fundamental Immunology*. W. E. Paul, ed. Lippincott-Raven, Philadelphia, p. 605.
- Bengtén, E., M. Wilson, N. Miller, L. W. Clem, L. Pilström, and G. W. Warr. 2000. Immunoglobulin isotypes: structure, function, and genetics. *Curr. Top. Microbiol. Immunol.* 248:189.
- Wilson, M., E. Bengten, N. W. Miller, L. W. Clem, L. Du Pasquier, and G. W. Warr. 1997. A novel chimeric Ig heavy chain from a teleost fish shares similarities to IgD. *Proc. Natl. Acad. Sci. USA* 94:4593.
- Rowe, D. S., and J. L. Fahey. 1965. A new class of human immunoglobulins. I. A unique myeloma protein. J. Exp. Med. 121:171.
- Abney, E. R., and R. M. Parkhouse. 1974. Candidate for immunoglobulin D present on murine B lymphocytes. *Nature 252:600.*
- Martin, L. N., G. A. Leslie, and R. Hindes. 1976. Lymphocyte surface IgD and IgM in non-human primates. Int. Arch. Allergy Appl. Immunol. 51:320.
- Ruddick, J. H., and G. A. Leslie. 1977. Structure and biologic functions of human IgD. XI. Identification and ontogeny of a rat lymphocyte immunoglobulin having antigenic cross-reactivity with human IgD. J. Immunol. 118:1025.
- Hordvik, I., J. Thevarajan, I. Samdal, N. Bastani, and B. Krossoy. 1999. Molecular cloning and phylogenetic analysis of the Atlantic salmon immunoglobulin D gene. Scand. J. Immunol. 50:202.
- Stenvik, J., and T. O. Jorgensen. 2000. Immunoglobulin D (IgD) of Atlantic cod has a unique structure. *Immunogenetics* 51:452.
- Miller, N., M. Wilson, E. Bengten, T. Stuge, G. Warr, and W. Clem. 1998. Functional and molecular characterization of teleost leukocytes. *Immunol. Rev.* 166:187.
- Magor, B. G., M. R. Wilson, N. W. Miller, L. W. Clem, D. L. Middleton, and G. W. Warr. 1994. An Ig heavy chain enhancer of the channel catfish *Ictalurus punctatus*: evolutionary conservation of function but not structure. J. Immunol. 153:5556.
- Wilson, M. R., A. Marcus, F. van Ginkel, N. W. Miller, L. W. Clem, D. Middleton, and G. W. Warr. 1990. The immunoglobulin M heavy chain constant region gene of the channel catfish, *Ictalurus punctatus*: an unusual mRNA splice pattern produces the membrane form of the molecule. *Nucleic Acids Res.* 18:5227.
- Hayman, J. R., S. H. Ghaffari, and C. J. Lobb. 1993. Heavy chain joining region segments of the channel catfish. J. Immunol. 151:3587.
- Ghaffari, S. H., and C. J. Lobb. 1999. Structure and genomic organization of a second cluster of Ig heavy chain gene segments in the channel catfish. J. Immunol. 162:1519.
- Ventura-Holman, T., and C. J. Lobb. 2002. Structural organization of the immunoglobulin heavy chain locus in the channel catfish: the *IgH* locus represents a composite of two gene clusters. *Mol. Immunol.* 38:557.

- Hogan, R. J., G. C. Waldbieser, C. A. Goudie, A. Antao, U. B. Godwin, M. R. Wilson, N. W. Miller, L. W. Clem, T. J. McConnell, W. R. Wolters, and V. G. Chinchar. 1999. Molecular and immunologic characterization of gynogenetic channel catfish (*Ictalurus punctatus*). *Marine Biotechnol.* 1:317.
- van Ginkel, F. W., N. W. Miller, C. J. Lobb, and L. W. Clem. 1992. Characterization of anti-hapten antibodies generated in vitro by channel catfish peripheral blood lymphocytes. *Dev. Comp. Immunol.* 16:139.
- Kim, U. J., B. W. Birren, T. Slepak, V. Mancino, C. Boysen, H. L. Kang, M. I. Simon, and H. Shizuya. 1996. Construction and characterization of a human bacterial artificial chromosome library. *Genomics* 34:213.
- Miller, S. A., D. D. Dykes, and H. F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16: 1215.
- Bengtén, E., S. Strömberg, A. Daggfeldt, B. G. Magor, and L. Pilström. 2000. Transcriptional enhancers of immunoglobulin light chain genes in Atlantic cod (*Gadus morhua*). *Immunogenetics* 51:647.
- Miller, N. W., V. G. Chinchar, and L. W. Clem. 1994. Development of leukocyte cell lines from the channel catfish (*Ictalurus punctatus*). J. Tissue Cult. Methods 16:117.
- Miller, N. W., M. A. Rycyzyn, M. R. Wilson, G. W. Warr, J. P. Naftel, and L. W. Clem. 1994. Development and characterization of channel catfish long term B cell lines. J. Immunol. 152:2180.
- Askovic, S., and R. Baumann. 1997. Alternative system for detection and mapping of activation domains. *BioTechniques* 22:948.
- Posnett, D. N., H. McGrath, and J. P. Tam. 1988. A novel method for producing anti-peptide antibodies: production of site-specific antibodies to the T cell antigen receptor β-chain. J. Biol. Chem. 263:1719.
- Szalai, A. J., M. T. Norcum, J. E. Bly, and L. W. Clem. 1992. Isolation of an acute-phase phosphorylcholine-reactive pentraxin from channel catfish (*Ictalurus punctatus*). Comp. Biochem. Physiol. B 102:535.

- Harlow, E. D., and D. Lane. 1988. Monoclonal antibodies. In Antibodies: A Laboratory Manual. Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY, p. 139.
- Eng, J. K., A. L. McCormick, and J. R. Yates III. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom. 5:976.
- Henikoff, S. 1992. Detection of *Caenorhabditis* transposon homologs in diverse organisms. *New Biol.* 4:382.
- Cioffi, C. C., D. L. Middleton, M. R. Wilson, N. W. Miller, L. W. Clem, and G. W. Warr. 2001. An IgH enhancer that drives transcription through basic helixloop-helix and Oct transcription factor binding motifs: functional analysis of the Eμ3' enhancer of the catfish. J. Biol. Chem. 276:27825.
- Chen, S., K. Anderson, and M. J. Moore. 2000. Evidence for a linear search in bimolecular 3' splice site AG selection. Proc. Natl. Acad. Sci. USA 97:593.
- Chua, K., and R. Reed. 2001. An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing. *Mol. Cell. Biol.* 21:1509.
- Liu, Z. J., P. Li, H. Kucuktas, and R. A. Dunham. 1999. Characterization of nonautonomous Tc1-like transposable elements of channel catfish (*Ictalurus punctatus*). *Fish Physiol. Biochem.* 21:65.
- Magor, B. G., D. A. Ross, L. Pilström, and G. W. Warr. 1999. Transcriptional enhancers and the evolution of the *IgH* locus. *Immunol. Today* 20:13.
- Baker, M. D., N. Pennell, L. Bosnoyan, and M. J. Shulman. 1988. Homologous recombination can restore normal immunoglobulin production in a mutant hybridoma cell line. *Proc. Natl. Acad. Sci. USA 85:6432.*
- Baker, M. D., L. R. Read, P. Ng, and B. G. Beatty. 1999. Intrachromosomal recombination between well-separated, homologous sequences in mammalian cells. *Genetics* 152:685.
- Shimizu, A., M. C. Nussenzweig, T. R. Mizuta, P. Leder, and T. Honjo. 1989. Immunoglobulin double-isotype expression by trans-mRNA in a human immunoglobulin transgenic mouse. *Proc. Natl. Acad. Sci. USA* 86:8020.