Lineage-restricted retention of a primitive immunoglobulin heavy chain isotype within the Dipnoi reveals an evolutionary paradox

Tatsuya Ota*, Jonathan P. Rast†, Gary W. Litman‡, and Chris T. Amemiya§¶

*Department of Biosystems Science, The Graduate University for Advanced Studies (Sokendai), Hayama, Kanagawa 240-0193, Japan; †Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125; ‡Children's Research Institute, University of South Florida, St. Petersburg, FL 33701; and §Virginia Mason Research Center, Seattle, WA 98101

Communicated by Francis H. Ruddle, Yale University, New Haven, CT, December 31, 2002 (received for review December 9, 2002)

The lineage leading to lungfishes is one of the few major jawed vertebrate groups in which Ig heavy chain isotype structure has not been investigated at the genetic level. In this study, we have characterized three different Ig heavy chain isotypes of the African lungfish, *Protopterus aethiopicus*, including an IgM-type heavy chain and short and long forms of non-IgM heavy chains. Northern blot analysis as well as patterns of V_H utilization suggest that the IgM and non-IgM isotypes are likely encoded in separate loci. The two non-IgM isotypes identified in *Protopterus* share structural features with the short and long forms of IgX/W/NARC (referred to hereafter as IgW), which were previously considered to be restricted to the cartilaginous fish. It seems that the IgW isotype has a far broader phylogenetic distribution than considered originally and raises questions with regard to the origin and evolutionary divergence of IgM and IgW. Moreover, its absence in other gnathostome lineages implies paradoxically that the IgW-type genes were lost from teleost and tetrapod lineages.

Tetrapods (land vertebrates) are thought to have shared a common ancestor with a group of fleshy-finned fishes that include the Crossopterygii (coelacanths) and the Dipneusti (lungfishes) (1, 2). Recent systematic studies using both morphological and molecular characters generally have supported a phylogeny in which the lungfishes are the closest relatives of the tetrapods (3–5). Comparative studies on the immunoglobulins at the molecular genetic level have focused on numerous ectothermic vertebrate groups, including chondrichthyans (sharks, rays, chimeras), teleosts (bony fishes), amphibians, and reptilians (reviewed in refs. 6 and 7), but little corresponding information is available for the fleshy-finned fish lineage. Three Ig isotypes have been identified so far at the protein level in two species of lungfishes (8–10) that were generically designated: high molecular weight (IgM), intermediate molecular weight, and low molecular weight (IgN) (9, 10). We report herein the cloning and characterization of several Ig heavy chain (IgH) cDNAs from *Protopterus aethiopicus* and present evidence for a phylogenetic relationship between these different isoforms and genes that have been identified in cartilaginous fishes, thus placing the divergence of IgH isotypes at a far earlier point in vertebrate phylogeny than considered previously.

Materials and Methods

Juvenile *P. aethiopicus* specimens (10–20 cm) were maintained in laboratory aquaria before death, and removal of tissues. RNA, and DNA were isolated by using standard protocols (11). Liver RNA was used as the source for cDNA library construction, whereas erythrocyte DNA was used for genomic library construction.

Generation of V_H Probes. V_H-specific probes were derived by amplifying *Protopterus* genomic DNA with primers that complement sequences in FR1 to FR3 that are common to vertebrate Ig V_H structures (12). Appropriately sized PCR products were subcloned into pBluescript SKII+ and sequenced. A probe specifying a single V_H was used in initial screening.

RNA Isolation, cDNA Library Construction and Screening, and Northern Blot Analysis. Total and poly $(A)^+$ RNA were isolated by using RNAzol (Tel-Test, Friendswood, TX) and an mRNA paramagnetic bead isolation method (Dynal, Oslo), respectively. An oligo(dT)-primed liver cDNA library was generated in λgt11 by using a commercial kit (Amersham Pharmacia). Initial library screening was conducted on quadruplicate nitrocellulose lifts of eight 150-mm plates at a density of ≈25,000–50,000 plaque-forming units per plate. Northern blots were prepared from liver mRNA by using standard methods (11).

Subcloning and DNA Sequencing. Isolated λ cDNAs were subcloned into plasmid (pBluescript SKII) or M13 (mp18) vectors and sequenced bidirectionally by using Applied Biosystems and Li-Cor (Lincoln, NE) DNA automated sequencers. Sequence assembly and analysis was performed by using IG SUITE (Intelligenetics), ALIGNIR (Li-Cor), PROSITE (13), and MEGA2 (14) software. Nucleotide sequences submitted to GenBank are available with the following accession numbers: clone 27, AF437724; clone 41, AF437725; clone 32, AF437726; clone 28, AF437727; clone 22, AF437728; clone 45, AF437729; clone 35, AF437730; clone 26, AF437731; clone 36, AF437732; clone 40, AF437733; and clone 76, AF437734.

Results

Identification of *Protopterus* Ig Genes. Three distinct V_H^+ clones possessing two different constant (C_H) gene segments were initially identified in a juvenile *Protopterus* liver cDNA library. Clone 2 encodes an incomplete IgM heavy chain; clones 4 and 8 encode different IgH chain types. *Protopterus*-specific V_H and C_H probes were derived from the initial clonal isolates and were hybridized individually to replicate filter lifts. Clones 27, 28, and 32 with insert sizes of 2.0, 2.8, and 1.5 kb, respectively, were sequenced.

IgM Heavy Chain. Clone 27 encodes a putative IgM heavy chain consisting of a leader peptide, a single V domain, a DJ region, and four C_H domains. The estimated molecular weight of the deduced amino acid sequence excluding the leader peptide (≈62,900) is consistent with the predicted molecular weight (≈70,000) of the IgM heavy chain (10, 15), if potential N-glycosylation is considered. Clone 27 aligns with various vertebrate IgM heavy chains (Fig. 1), and its assignment as an IgM is unequivocal. Two cysteines and a tryptophan that comprise the core of the Ig fold (16) were observed for the variable Ig domain and three of the four constant Ig domains. Differences in

Abbreviation: IgH, Ig heavy chain.

¶To whom correspondence should be addressed. E-mail: camemiya@vmresearch.org.

EVOLUTION

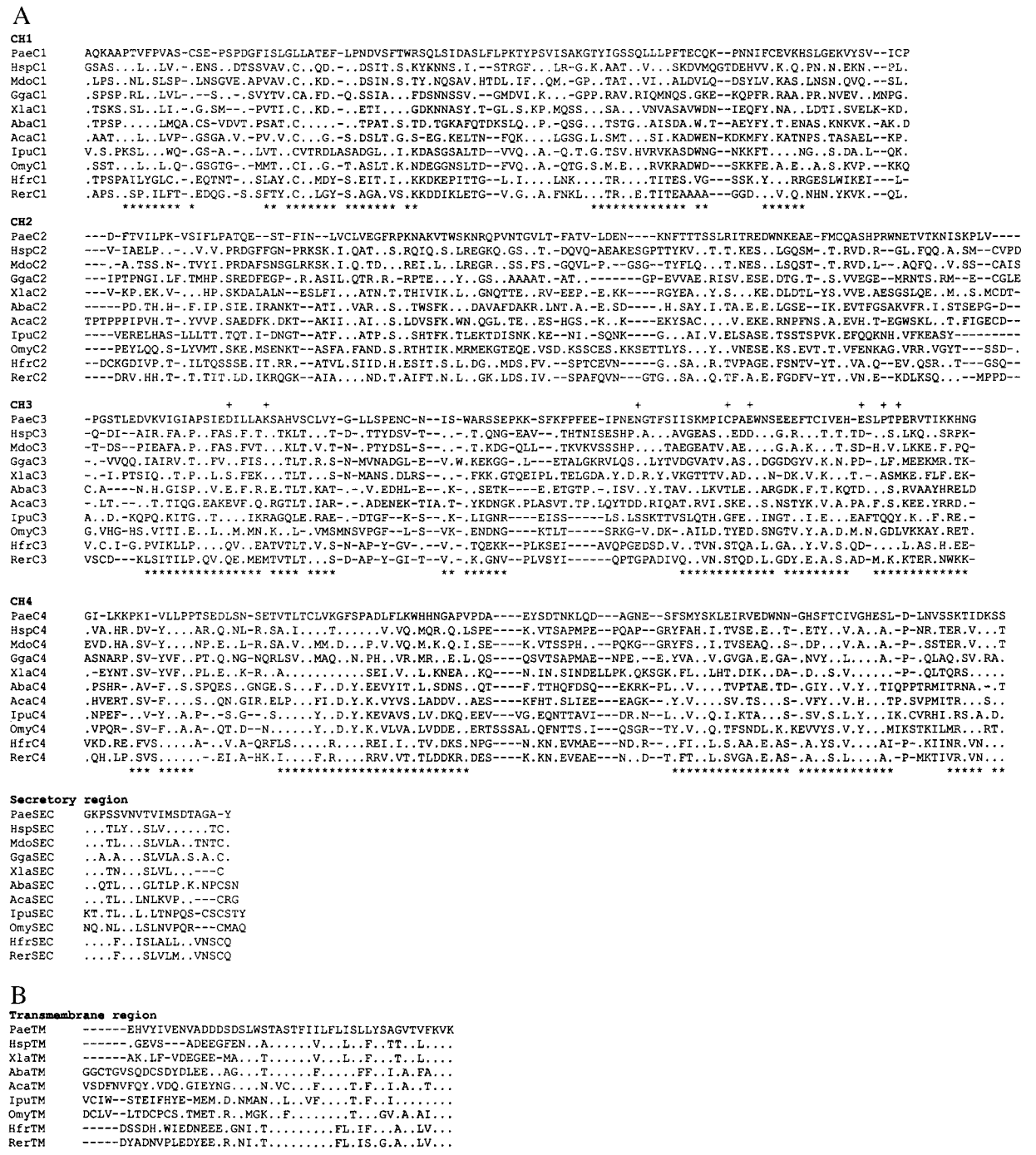

Fig. 1. (*A*) Amino acid alignment of 11 vertebrate Cμ secretory sequences. Sequences were from GenBank: African lungfish (Pae), *P. aethiopicus* (this study, clone 27); human (Hsp), *Homo sapiens* (X57086); short-tailed opossum (Mdo), *Monodelphis domestica* (AF108226); chicken (Gga), *Gallus gallus* (X01613); African clawed frog (Xla), *Xenopus laevis* (M20484); sturgeon (Aba), *Acipenser baeri* (Y13253); bowfin (Aca), *Amia calva* (U12456); channel catfish (Ipu), *Ictalurus punctatus* (M27230); rainbow trout (Omy), *Oncorhynchus mykiss* (S63348); horned shark (Hfr), *Heterodontus francisci* (X07784); little skate (Rer), *Raja erinacea* (M29679). Periods indicate identical residue with African lungfish residue; dashes indicate gaps. +, residues important for complement activation. *, Residues used for subsequent phylogenetic analyses. (*B*) Amino acid sequences of eight vertebrate Cμ transmembrane regions: Pae (this study, clone 24), Hsp (PIR, S14683), Xla (X90768), Aba (AJ245364), Ipu (X52617), Omy (U04616), Hfr (X07785), and Rer (M29678). The sequence for *Amia calva* was taken from Wilson *et al.* (35). The transmembrane regions of human, lungfish, shark, and skate IgM heavy chain are anchored at the end of the C_H4 domain, whereas those of channel catfish and rainbow trout are anchored at the end of the C_H3 domain. The transmembrane region of bowfin IgH heavy chain can be anchored at the end of C_H4, the end of C_H3, or the middle of the C_H3 domain, whereas the sturgeon transmembrane region seems to be anchored at the end of either the C_H1 (AJ245363) or the C_H4 (AJ245362) domain.

conserved cysteine residues have been noted in IgM heavy chains from both teleost and cartilaginous fishes, specifically Cys^{P409}, which corresponds to mouse Cys^{M414} that is known to be important for complement activation (17), is absent in these species but is present in the lungfish IgM heavy chain (P or M shown for the superscripts of amino acid residues indicates

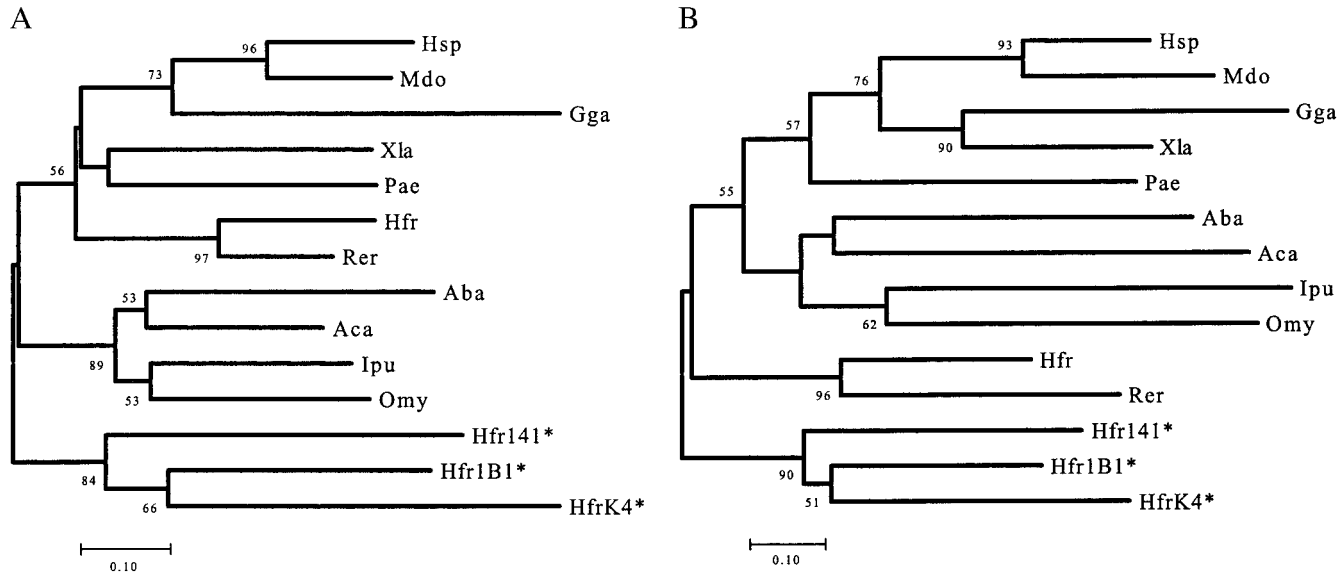

Fig. 2. (*A*) Neighbor-joining tree for C_H4 domains of IgM heavy chain. Pairwise genetic distances were obtained by Poisson correction. Three horn shark light chain sequences (HfrK4, Hfr1B1, Hfr141; ref. 31 and references therein) were used as outgroups and are indicated by asterisks. The number given in each branch is a bootstrap value; bootstrap values <50% are not shown. The branch lengths are measured in terms of the number of amino acid substitutions, with the scale given below the tree. (*B*) Neighbor-joining tree for IgM heavy chain C_H3 domains. Abbreviations for the taxa are given in Fig. 1.

Protopterus or *Mus*, respectively). Other residues important for complement activation that are involved in C1q binding (e.g., homologous residues of mouse, Asp^{M432}, Pro^{M434}, or Pro^{M436}; ref. 18) are conserved in the lungfish with two replacements of biochemically similar amino acids: $Asp^{M417} \rightarrow Glu^{P412}$ and $Asp^{M432} \rightarrow Glu^{P426}$. Furthermore, a site ($Asp^{P397}$) is present in the lungfish IgM heavy chain that corresponds to mouse Asn^{M402}, at which an N-linked glycosaccharide is involved indirectly in C1q binding (19).

C_H4 is the most conserved IgM heavy chain constant region domain (20). In phylogenetic analyses, the lungfish IgM and tetrapod IgM heavy chain C_H4 domains cluster with the corresponding regions of IgM that have been identified in cartilaginous fishes but not with the C_H4 domains from representative actinopterygians (bony fishes) (Fig. 2*A*). This finding was entirely unanticipated because actinopterygians are more closely related phylogenetically to tetrapods and lungfishes than to the cartilaginous fishes, although it was also observed in earlier phylogenetic analysis (e.g., ref. 21). Conversely, phylogenetic analyses of the other $C\mu$ domains demonstrate closer relationships between domains of actinopterygian, tetrapod, and lungfish rather than with those of cartilaginous fishes (e.g., Fig. 2*B*). Because the splicing pattern of the transmembrane form of IgM heavy chain of teleost fishes differs from higher vertebrates as discussed below, it seems that the teleost IgM heavy chain constant regions have been subject to some unique evolutionary event, such as exon shuffling, gene conversion, or genomic recombination, around the C_H3–C_H4 region.

Non-IgM Heavy Chain. Several structural/functional inferences can be made from the predicted peptide sequence of a second IgH-like molecule (clone 32): (*i*) strong homology is evident for the V_H gene segment of this clone and other V_H gene segments; (*ii*) a β-bulge motif, which is significant for V_H–V_L interactions at the V and DJ junction regions, is present (see below); and (*iii*) a cysteine residue in the C_H1 domain is consistent with a light chain disulfide bridge (Fig. 3). The predicted protein product of clone 32 consists of a leader peptide, one V domain, a DJ region, and two C domains. This structure may correspond to that of the Ig (low molecular weight)/IgN heavy chain (10, 15). The estimated molecular weight of the gene product of clone 32, excluding the leader peptide, is ≈36,500, equivalent to the estimated molecular weight of Ig (low molecular weight) heavy chain. The two consecutive cysteine residues present in the C_H1 domain raise the possibility of an extra intradomain disulfide bridge, as described in other species (22). Four additional cysteine residues are located at various parts of the C_H region, some of which are potentially involved in the formation of inter-heavy chain disulfide bridges.

A third cDNA (clone 28) is predicted to encode an IgH consisting of a single V domain, a DJ region, and seven C domains (Fig. 3). Whereas the predicted leader peptides and V gene segments found in clones 32 and 28 differ to some degree, the nucleotide sequences of the first two C domains as well as J gene segments are identical. This observation is consistent with alternative RNA processing at a single Ig locus, and comparisons of the predicted sequences of clones 28 and 32 are reminiscent of observations with skate IgX (IgW/IgNARC) in which the short and long forms of heavy chain encode two and six C domains, respectively (23–26). Owing to the extraordinarily large genome size of *Protopterus* (>20 times larger than that of human; ref. 27), it will be difficult to definitively resolve whether the different gene products are produced by alternative splicing or from two discrete loci.

A phylogenetic analysis of the lungfish IgH isotype has been carried out (Fig. 4; a bootstrap consensus tree was obtained because many basal branching points were poorly resolved). Because the numbers of C_H domains range from two to seven in different vertebrate IgH isotypes, it was necessary to identify and compare homologous domains. The C_H1 domain was selected as a reference because it is present in all IgH and shares, with limited exceptions, one conserved function, i.e., covalent association with IgL. The phylogenetic analysis indicates that the C_H1 domain of this lungfish IgH is related closely to that of IgW found in cartilaginous fish. Although it would be of considerable significance, the relationship of the IgH in lungfish to teleost fish IgD cannot be resolved, because the first IgD domain corresponds to C_H1 of IgM (28).

The C-terminal of clone 28 contains a cysteine residue as well as a potential N-glycosylation site, which are characteristics of

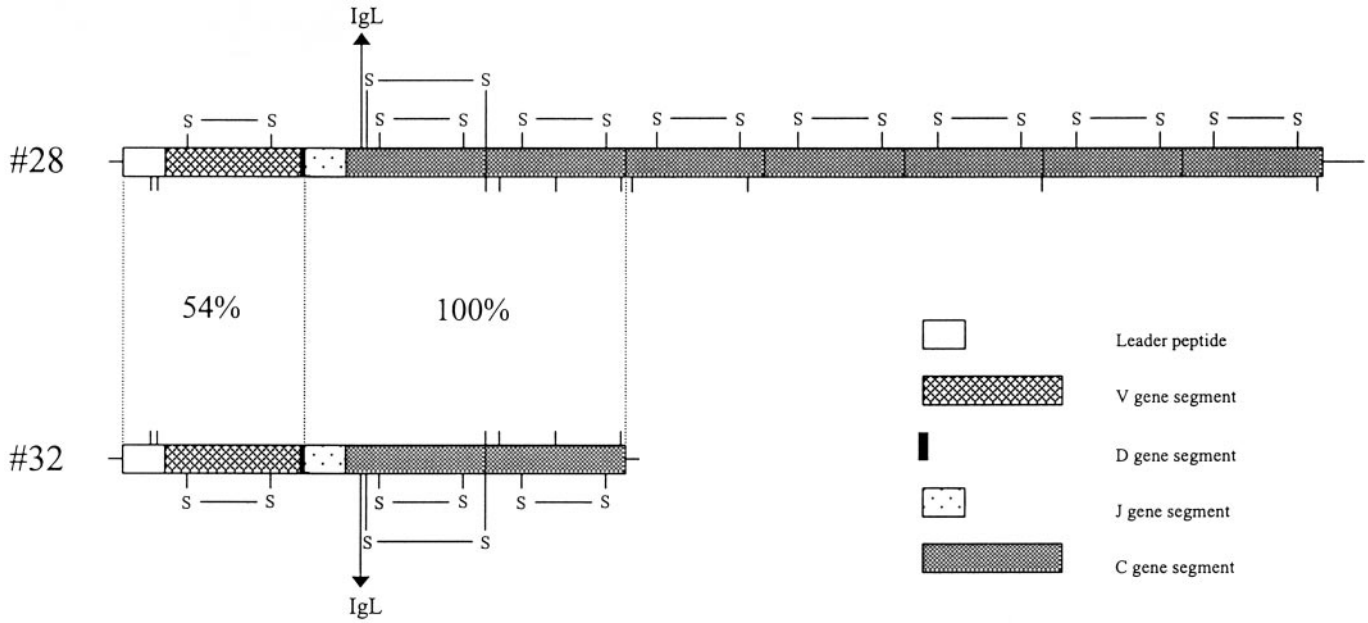

Fig. 3. Long form (clone 28) and short form (clone 32) of IgW-like heavy chains. Vertical lines indicate sulfhydryl groups of cysteine residues, which can form interchain disulfide bridges. Potential intrachain disulfide bridges are indicated by S–S bridges. The numbers shown are the proportions of nucleotide identity at the V–D region and the J–C region.

the secretory tail of vertebrate IgM, mammalian and avian IgA, and cartilaginous fish IgW heavy chain, but not of other IgH, including the teleost IgD heavy chain. Amino acid comparisons of the secretory tails show relatively high degrees of similarity (50% identity) between the lungfish and cartilaginous fish IgW heavy chain sequences, even though the lungfish sequence contains an extra C-terminal residue. Collectively, these observations further support a direct relationship of the lungfish IgH to cartilaginous fish IgW heavy chain.

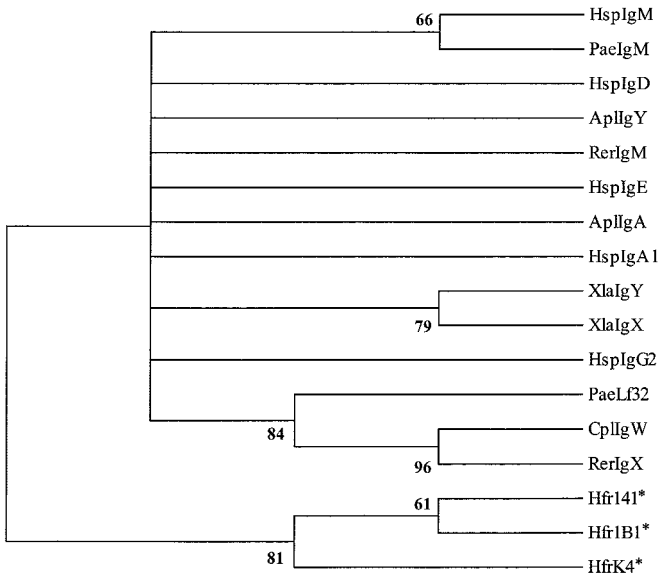

Fig. 4. Bootstrap consensus tree for C_H1 domains of IgH. Pairwise genetic distances were estimated by Poisson correction, and the phylogenetic tree was constructed by the neighbor-joining method. Sequences were from GenBank or the Protein Information Resource: human (Hsp), *H. sapiens* IgA1(J00220), IgD (PIR:AAB21246), IgE (J00222), IgG2 (J00230), and IgM (X57086), duck (Apl), *Anas platyrhyncos* IgY (X65219) and IgA (U27222); African clawed frog (Xla), *X. laevis* IgX (PIR:S03186) and IgY (X15114); African lungfish (Pae), *Protopterus aethiopicus* IgM (this study, clone 27) and Lf32 (this study, clone 32); carcharhine shark (Cpl), *Carcharhinus plumbeus* (U40560); and little skate (Rer), *R. erinacea* (PIR:S12839). Three horn shark light chain sequences (HfrK4, Hfr1B1, and Hfr141) (ref. 31 and references therein) were used as outgroups and are denoted by asterisks.

Variable Region Diversity. V_H regions from 13 IgH^+ clones have been sequenced, and eight different V_H gene families have been identified thus far (Fig. 5). There were, in addition, four other V_H gene segments known from a lungfish (E. Hsu, personal communication), each of which may represent a different V_H family. The amino acid difference between the V_H regions of different families varies from ≈35% to ≈64%. The variability, including insertions/deletions, is particularly extensive in the complementarity-determining regions; framework regions (FR) vary to a lesser degree. A conserved cysteine in the V_H region of clone 76 has been replaced with a serine and may be indicative of a pseudogene. All V_H regions, with the exception of atypical clone 45, encode the Gly-Leu-Glu-Trp motif (or variants thereof) found in FR2, which is essential for V_H/V_L dimerization through β-bulge formation (29, 30) along with the (Trp/Phe)-Gly-X-Gly motif in the DJ region.

Although the number of genes being compared is limited, the same V_H gene families are associated with the same isotypes (Fig. 5). Association of V_H gene families with a particular isotype is unusual and observed only rarely in higher vertebrates. However, a similar restricted utilization of V gene families has been described in cartilaginous fish IgM, IgW, and Ig light chains (31–33), which are encoded at separate clusters throughout the genome (ref. 24; T.O. and C.T.A., unpublished work).

RNA Expression. Northern blots were hybridized with probes complementing two distinct V_H and C_H genes (from clones 2 and 8; Fig. 6). The V_H probe derived from clone 2 was shown to hybridize strongly with a 2-kb fraction and relatively weakly with a 1.6-kb fraction. The V_H probe (V_H V family) derived from clone 8 hybridized strongly with 1.6- and 2.8-kb fractions but failed to hybridize with a 2-kb fraction. The 2-kb transcripts hybridized with V_H probes derived from clone 2, which encodes

```
                FR1                              CDR1     FR2              CDR2                    FR3

VH family I
Lf2 (IgM)
Lf22 (Ig?)    QVVLTETGGDAMKPGETLRLTCKTSGFDFS --SSWMG WVRQQPER-GLDWIG E-I-NS-GGGT--IHYHPSFKG RFTISRDNSKNLLYLDLGSLKSEDTAVYFCAR

VH family II
Lf27 (IgM)    AVEIKETGGGIWKPGDSITLTCRCSGFTFT --DYYMH WARQQPGS-GLEWVG M-VRNPANGGT--TEYSPTLRD RVTISRDNSRSEVYLQIKTLKAEDSALYYCAR

VH family III
Lf41 (IgM)    DIVLTQSGSEVKQPGQSVNLRCTGSGFDMN --SYYIE WVRQFPGG-GFEWLV QYL-KP-GAGSN--YVNPTVQG RFSGSVD-SSS-FYLQMNNLQVADTAMYYCAR

VH family IV
Lf45 (IgM)    SIVLTQKELEQGSPGSTIVLQCPFSGFSPS --DHHVS WVRQTPGKTELEWLG G-S-RT-GQTT---YISPAVKG KMTPSTI-GST-AQLQINQLRPEDTAVYYCAR

VH family V
Lf4/8 (IgW)
Lf32 (IgW)    QVKLEESGGRAVTPGASLQLNCKATGFDFG PGGYAMN WVRQQPGK-GLEWVA YHY-TG-SGGTNYIYYSPSVQG RFTISRDVPQRLLHLDMKNVKTEDTGTYYCAR
Lf35 (IgW)    QVKLEESGGRAVTPGASLQLNCKATGFDFG PGGYAMN WVRQQPGK-GLEWVA YHY-TS-SGGTNYIYYSPSVQG RFTISRDVPQRLLHLDMKNVKTEDTGTYYCAR

VH family VI
Lf26 (IgW)    AITLKETGGGVKKPGDTTRLTCQVSGFSLS --NNDVS WVRQSPGQ-GLQWLG T-M-WD-DGST---DLSPSIKD RTSISRDVPKGEVYLDISKMKTEDTGTYYCAR

VH family VII
Lf28 (IgW)    DIVLSQSAAEVKPPGQSLKLKCAVTGFDVS --SYYMS WVKQAPGK-GFEWIV SYY-DS-NPSNH--NYNPTFQS RFSASKD-SSN-FYLQMNNLQVADTSVYYCAR
Lf36 (IgW)    DIVLSQSAAEVKPPGQSLKLKCAVTGFDVS --SYYMS WVKQAPGK-GFEWIV SYY-DS-NPSNH--NYNPTFQS RFSASKD-SSN-FYLQMNNLQVADTSVYYCAT
Lf40 (IgW)    DIVLSQSAAEVKPPGQSLKLKCAVTGFDVS --SYYMS WVKQAPGK-GFEWIV SYY-DS-NPSNH--NYNPTFQS RFSASKD-SSN-FYLQMNNLQVADTSVYYCAR

VH family VIII
Lf76 (IgW)    QVSLTETGGGVKKPGESLRLSCKGSGYNFG --SYGLS WVRQQPGK-GLEWVA Y-I-S--SGSTN-IYYSPSLKD RTTISRDNSQSLLYLDMTNLRLEDTGTYYSAR
```

Fig. 5. Inferred amino acid sequences of the V_H regions of *P. aethiopicus* clones. FR, framework region; CDR, complementarity-determining region. V_H regions of clone 2 were designated V_HI; clones 4 and 8 were designated V_HV. IgH-type was confirmed by success/failure of PCR with using C_H1 primers unique to IgM or IgW isotypes. The isotype designation of clone 22 was not of the IgM or IgW type (unpublished data).

an IgM heavy chain, whereas the 1.6- and 2.8-kb transcripts hybridized with the probes derived from clone 8, which would complement the short and long forms, respectively, of the IgW-like heavy chain. This observation was confirmed by hybridization analyses using C_H-specific probes (Fig. 6). Taken together, the hybridization patterns observed are consistent with the association of different V_H gene families with different heavy chain isotypes.

Discussion

The Evolution of IgM. In higher vertebrates, IgM is produced first during development and is the most significant in the early phase of the immune response. In its membrane-bound form, IgM functions as a B cell-specific ligand binding receptor, and its expression is critical to progressive B cell differentiation. IgM-type genes comprise the only class of IgH that has been identified in all jawed vertebrates (20). All known secretory forms of IgM heavy chain consist of one V domain, a DJ region, and four C_H domains, as in the case of lungfish. However, the transmembrane region of lungfish IgM is contiguous with the C_H4 domain (clone 24; Fig. 1), a situation also found in mammals, avians, amphibians, and cartilaginous fish. In contrast, membrane-bound forms of teleost IgM consist of two or three C_H domains owing to a unique form of RNA processing (34), which likely was established at an early point in the evolution of bony fishes (35, 36). Phylogenetic analysis of C_H4 domains of vertebrate IgM molecules indicates that the Neopterygii, a group that includes the teleosts as well as primitive bony fishes such as bowfins and sturgeons, forms a distinct cluster of their own (Fig. 2). Taken together, these findings suggest that the lungfish IgM heavy chain is more similar to those of tetrapods than to those found in neopterygian fishes. This is consistent with the generally held hypothesis that the lobe-finned fishes rather than actinopterygian fishes are the closest relatives of tetrapods.

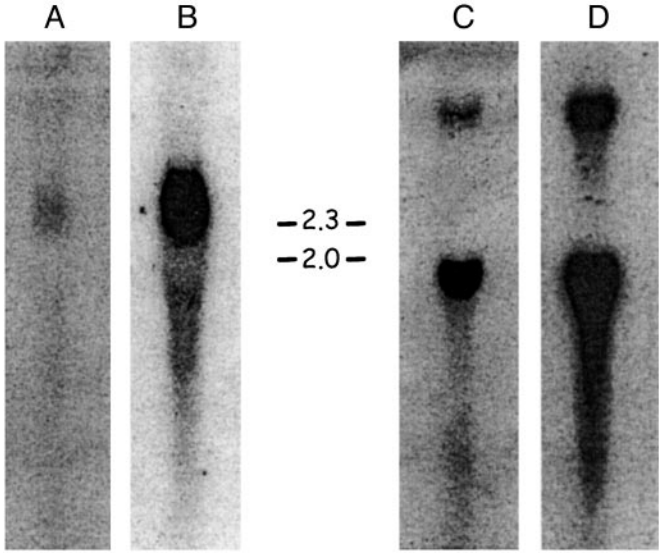

Fig. 6. Northern blot analysis of two different isotypes suggests differential usage of V_H and C_H genes. Probes complementing the V_H and C_H regions of clones 2 (IgM-type) and 8 (IgW-type) were hybridized as described. (*A*) Hybridization of V_H of clone 2. (*B*) Hybridization of C_H of clone 2. (*C*) Hybridization of V_H of clone 8. (*D*) Hybridization of C_H of clone 8. Sizes are in kb.

Lungfish IgW-Like Heavy Chain. In addition to IgM, a second isotype of Ig (IgW) has been identified in the earliest jawed vertebrates, raising questions as to the primordial IgH (23, 25, 37). Based on phylogenetic analysis of the C_H of the lungfish clones presented here as well as their heavy chain structure and patterns of exon utilization, we propose that this isotype system is orthologous to cartilaginous fish IgW heavy chains. In cartilaginous fish, the IgM and IgW heavy chain gene loci are encoded by >100 separate, paralogous loci (24–26). Although the number of loci encoding the IgH in lungfish is not yet known, it is unlikely that it is as great as that seen in the cartilaginous fishes, as the nucleotide sequences of the C_H regions examined thus far exhibit little divergence, with the exception of the few amino acid differences in the C_H1 domain of IgM heavy chain that may represent allelic polymorphisms (clone 41). The distinct patterns of usage of V_H gene segments between the IgM and the

IgW heavy chains are consistent with their being encoded at independent loci. Furthermore, nucleotide sequence analysis suggests that these short and long forms derive by alternative splicing, although the possibility that they are products of recently duplicated genes cannot be ruled out. As indicated above, resolution of this issue requires extensive genomic characterization, which is technically confounded by the unusually large size of the lungfish genome (27).

The presence of both the IgM and IgW isotypes in cartilaginous fishes as well as lungfish implies that the duplication of these genes predates the origin of bony fishes analogous to the duplication of Ig light chain to produce κ and non-κ (λ) types. Importantly, this report demonstrates that the IgW system is not restricted to the Chondrichthyes and has a broader phylogenetic distribution than considered originally. No clear homologs of IgW have been identified in the human genome, or for that matter in amphibians, reptiles, and avians. Although perhaps premature, it is notable that in terms of subunit composition, the evolution of Ig genes has, on one hand, been highly conserved throughout the phylogeny of jawed vertebrates; however, the organization and complexity of genes is highly variable in those taxa examined to date.

We thank Mary West and Barbara Pryor for editorial assistance and Ronda Litman, Asaf Halevi, and Wendy Lippmann for technical assistance. This work was funded, in part, by National Science Foundation Grants IBN-9614940 and IBN-9905408 (to C.T.A.), National Institutes of Health Grants R37-AI23338 (to G.W.L.) and R24-RR14085 (to C.T.A.), the Center for Human Genetics (Boston University School of Medicine), and the Benaroya Research Institute at Virginia Mason.

1. Carroll, R. L. (1988) *Vertebrate Paleontology and Evolution* (Freeman, New York).
2. Helfman, G. S., Collette, B. B. & Facey, D. E. (1997) *The Diversity of Fishes* (Blackwell, Oxford).
3. Meyer, A. & Wilson, A. C. (1990) *J. Mol. Evol.* **31,** 359–364.
4. Meyer, A. & Dolven, S. I. (1992) *J. Mol. Evol.* **35,** 102–113.
5. Marshall, C. & Schultze, H. P. (1992) *J. Mol. Evol.* **35,** 93–101.
6. Du Pasquier, L. & Flajnik, M. (1998) in *Fundamental Immunology*, ed. Paul, W. E. (Lippincott–Raven, Philadelphia), pp. 605–650.
7. Litman, G. W., Anderson, M. K. & Rast, J. P. (1999) *Annu. Rev. Immunol.* **17,** 109–147.
8. Marchalonis, J. J. (1969) *Aust. J. Exp. Biol. Med. Sci.* **47,** 405–419.
9. Litman, G. W., Wang, A. C., Fudenberg, H. H. & Good, R. A. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 2321–2324.
10. Litman, G. W. (1976) in *Relationships Between Structure and Function of Lower Vertebrate Immunoglobulins*, eds. Hildemann, W. H. & Benedict, A. A. (Plenum, New York), pp. 217–228.
11. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
12. Amemiya, C. T., Ohta, Y., Litman, R. T., Rast, J. P., Haire, R. N. & Litman, G. W. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6661–6665.
13. Bairoch, A., Bucher, P. & Hofmann, K. (1997) *Nucleic Acids Res.* **25,** 217–221.
14. Kumar, S., Tamura, K., Jakobsen, I. & Nei, M. (2001) *Bioinformatics* **17,** 1244–1245.
15. Marchalonis, J. J. (1977) *Immunity in Evolution* (Harvard Univ. Press, Cambridge, MA).
16. Lesk, A. M. & Chothia, C. (1982) *J. Mol. Biol.* **160,** 325–342.
17. Davis, A. C., Roux, K. H., Pursey, J. & Shulman, M. J. (1989) *EMBO J.* **8,** 2519–2526.
18. Arya, S., Chen, F., Spycher, S., Iseman, D. E., Shulman, M. J. & Painter, R. H. (1994) *J. Immunol.* **152,** 1206–1212.
19. Wright, J. F., Shulman, M. J., Isenman, D. E. & Painter, R. H. (1990) *J. Biol. Chem.* **265,** 10506–10513.
20. Bengtén, E., Wilson, M., Miller, N., Clem, L. W., Pilstrom, L. & Warr, G. W. (2000) *Curr. Top. Microbiol. Immunol.* **248,** 189–220.
21. Greenberg, A. S., Avila, D., Hughes, M., Hughes, A., McKinney, E. C. & Flajnik, M. F. (1995) *Nature* **374,** 168–173.
22. Fellah, J. S., Kerfourn, F., Wiles, M. V., Schwager, J. & Charlemagne, J. (1993) *Immunogenetics* **38,** 311–317.
23. Harding, F. A., Amemiya, C. T., Litman, R. T., Cohen, N. & Litman, G. W. (1990) *Nucleic Acids Res.* **18,** 6369–6376.
24. Anderson, M., Amemiya, C., Luer, C., Litman, R., Rast, J., Niimura, Y. & Litman, G. (1994) *Int. Immunol.* **6,** 1661–1670.
25. Greenberg, A. S., Hughes, A. L., Guo, J., Avila, D., McKinney, E. C. & Flajnik, M. F. (1996) *Eur. J. Immunol.* **26,** 1123–1129.
26. Anderson, M. K., Strong, S. J., Litman, R. T., Luer, C. A., Amemiya, C. T., Rast, J. P. & Litman, G. W. (1999) *Immunogenetics* **49,** 56–67.
27. Pedersen, R. A. (1971) *J. Exp. Zool.* **177,** 65–79.
28. Wilson, M., Bengtën, E., Miller, N. W., Clem, L. W., Du Pasquier, L. & Warr, G. W. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 4593–4597.
29. Chothia, C., Novotny, J., Bruccoleri, R. & Karplus, M. (1985) *J. Mol. Biol.* **186,** 651–663.
30. Lascombe, M. B., Alzari, P. M., Poliak, R. J. & Nisonoff, A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9429–9433.
31. Rast, J. P., Anderson, M. K., Ota, T., Litman, R. T., Margittai, M., Shamblott, M. J. & Litman, G. W. (1994) *Immunogenetics* **40,** 83–99.
32. Shen, S. X., Berstein, R. M., Schluter, S. F. & Marchalonis, J. J. (1996) *Immunol. Cell Biol.* **74,** 357–364.
33. Ota, T., Sitnikova, T. & Nei, M. (2000) *Curr. Top. Micorbiol. Immunol.* **248,** 221–245.
34. Ross, D. A., Wilson, M. R., Miller, N. W., Clem, L. W. & Warr, G. W. (1998) *Immunol. Rev.* **166,** 143–151.
35. Wilson, M. R., Ross, D. A., Miller, N. W., Clem, L. W., Middleton, D. L. & Warr, G. W. (1995) *Dev. Comp. Immunol.* **19,** 165–177.
36. Ota, T., Nguyen, T.-A., Huang, E., Detrich, H. W. & Amemiya, C. T. (2003) *J. Exp. Zool. B* **295,** 45–58.
37. Bernstein, R. M., Schluter, S., Shen, S. & Marchalonis, J. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 3289–3293.