Expression of IgM, IgD, and IgY in a Reptile, *Anolis carolinensis*

Zhiguo Wei, Qian Wu, Liming Ren, Xiaoxiang Hu, Ying Guo, Gregory W. Warr, Lennart Hammarström, Ning Li and Yaofeng Zhao

*J Immunol* 2009; 183:3858-3864; Prepublished online 28 August 2009;
doi: 10.4049/jimmunol.0803251
http://www.jimmunol.org/content/183/6/3858

**Supplementary Material**
http://www.jimmunol.org/content/suppl/2009/08/28/jimmunol.0803251.DC1

**References**
This article cites 49 articles, 26 of which you can access for free at:
http://www.jimmunol.org/content/183/6/3858.full#ref-list-1

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**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
Expression of IgM, IgD, and IgY in a Reptile, *Anolis carolinensis*1

Zhiguo Wei,2* Qian Wu,2* Liming Ren,* Xiaoxiang Hu,* Ying Guo,* Gregory W. Warr,3† Lennart Hammarström,5 Ning Li,4* and Yaofeng Zhao4*

The reptiles are the last major group of jawed vertebrates in which the organization of the *IGH* locus and its encoded Ig H chain isotypes have not been well characterized. In this study, we show that the green anole lizard (*Anolis carolinensis*) expresses three Ig H chain isotypes (IgM, IgD, and IgY) but no IgA. The presence of the δ gene in the lizard demonstrates an evolutionary continuity of IgD from fishes to mammals. Although the germline δ gene contains 11 CH exons, only the first 4 are used in the expressed IgD membrane-bound form. The μ chain lacks the cysteine in Cμ1 that forms a disulfide bond between H and L chains, suggesting that (as in IgM of some amphibians) the H and L polypeptide chains are not covalently associated. Although conventional IgM transcripts (four CH domains) encoding both secreted and membrane-bound forms were detected, alternatively spliced transcripts encoding a short membrane-bound form were also observed and shown to lack the first two CH domains (VDJ-CH3-CH4-transmembrane region). Similar to duck IgY, lizard IgY H chain (ν) transcripts encoding both full-length and truncated (IgYΔFc) forms (with two CH domains) were observed. The absence of an IgA-encoding gene in the lizard *IGH* locus suggests a complex evolutionary history for IgA in the saurian lineage leading to modern birds, lizards, and their relatives. *The Journal of Immunology*, 2009, 183: 3858–3864.

The Ig classes expressed in mammals are IgM, IgD, IgG, IgE, and IgA. This classification is based on their H chain isotypes, which are encoded by the μ, δ, γ, ε, and α genes, respectively (1, 2). The evolutionary history of these genes is of interest from many perspectives, two of which have received particular attention. The first is the question of how the genes encoding the H chains arose, one from the other, and how their subsequent evolutionary divergence permitted the acquisition of novel Ab functions including transport and secretion into external body fluids such as mucus and milk, or the mediation of allergic reactions. In summary, first, IgM is the only class of Ab that is expressed in all species of jawed vertebrates (1, 2, 5). Second, IgD is a class of Ab that is probably as ancient as IgM, existing in elasmobranchs as a class that was originally named IgW but is now recognized as a homolog of IgD (6). IgD is an enigmatic Ab the structure of which is highly variable (which, in some cases, is due to duplication or loss of Cμ exons; Ref. 7), the function of which is a matter of debate (8), and which has clearly been lost from certain species. In bony fish and pigs, IgD can be expressed as chimeric H chains with inclusion of the IgM Cμ1 domain (9, 10). Although IgD/W has been described in fish (cartilaginous and bony), amphibians, and numerous mammals (6, 9, 11, 12), it is absent in birds and rabbits (13, 14). Third, the major class of low-m.w. Ab in the amphibia, reptiles, and birds is IgY, which is thought to have been derived from IgM by a gene duplication event (5). Fourth, IgY played a central role in Ab evolution in the tetrapod lineage, giving rise to IgA (through an apparent recombination event with IgM; Ref. 15) and to IgG and IgE through a gene duplication event followed by differentiation of structure and function (16). Overlaid on this broad scheme are other important events in the evolutionary history of Abs where certain lineages developed unique classes of Abs, such as the IgNAR in elasmobranchs (17), IgT/Z in some species of bony fish (18, 19), and IgX and IgF in amphibia (12, 20).

A second important perspective on the evolution of the Ab genes is the structure of the *IGH* locus. Interest in this question was stimulated by observations that in cartilaginous fish (phylogenetically the most primitive jawed vertebrates), the *IGH* loci show a unique structure. Whereas in the mammals the genes present in the *IGH* locus are arranged in a (VH1-μ-Dμ1-μ-Jμ1-μ-Cμ1) order, the *IGH* locus of cartilaginous fish show a multicluster organization, where multiple clusters, each consisting of (V-Dμ1-Dμ2-Jμ1-Cμ1),...
are arranged in tandem (21–23). These alternative organizations (translocon or multicluster) present different challenges for the orderarily rearrangement and clonal expression by B lymphocytes of the Ab H chains (24).

Although the IGH loci have been extensively investigated in fish, amphibia, birds, and mammals, our current knowledge of the Ig genes in reptiles remains limited. Reptiles are known, on the basis of studies at the protein level, to express two Ig classes, IgM and IgY (25, 26), and an IgM encoding gene has been cloned in turtles (27). Two types of IgY (7S, 5.7S) that differ in the lengths of their H chain have been described at the biochemical level in turtles (28, 29), but it is not known whether the genetic basis of this phenomenon is the same as that underlying the expression of two isoforms of IgY in the duck (30). A gene encoding an IgA-like isotype was recently cloned in the leopard gecko (15), suggesting that the repertoire of Ig classes in reptiles may be similar to that of birds. The IgD-encoding gene, despite its absence in birds, was also recently reported in the gecko (31). A paucity of knowledge of the structure of the reptilian IGH locus and of its encoded Ig H chain isotypes has been a barrier to a more complete understanding of the evolutionary history of Igs in vertebrates. We report here an investigation of the expression of the IGH genes of a reptile, the green anole lizard (Anolis carolinensis). (While our paper was under revision, another group (5) also reported on the expression of IGH in a reptilian species, the green anole lizard, and described the structure of the IGH locus.)

Materials and Methods

Annnotation of the green anole lizard (A. carolinensis) IGH constant genes

Basic local alignment search tool (BLAST) searches were performed against the genome sequence of the lizard (AnoCar1.0 assembly (19); this primer was designed based on the NCBI GenBank under the following accession numbers: EF683585 and EF690357–EF690361 (http://www.ncbi.nlm.nih.gov/sites/entrez).

Animals, RNA isolations, and reverse transcriptions

About 3-year-old green anole lizards were purchased from a local pet market. Total RNA from different tissues was prepared using a TRNzol kit (Tiangen). Recombination signal sequences for Vγ DH, and Jγ gene segments were examined using the program FUZZNUC (http://bioweb.pasteur.fr/sequanal/interfaces/fuzznuc.html).

Cloning of the IγM, IγD, and IγH chain constant region cDNAs

Two IgJ primers, JHs1 (5′-GGGGGAAAAAGAATTTCTG) and JHs2 (ACATTCTCTGTGCTAATCTCAG-3′) were designed based on the retrieved IγM cDNA sequence as a template (27), a BLAST search was performed against the NCBI GenBank under the following accession numbers: EF683585 and EF690357–EF690361. The sequences have been deposited in the NCBI GenBank under the following accession numbers: EF683585 and EF690357–EF690361. (http://www.ncbi.nlm.nih.gov/sites/entrez).

Southern blotting

The above amplified μ (membrane-bound form), δ (membrane-bound form), and ν (secreted form) cDNAs were used as probes, which were all labeled using a PCR dihydrogen synthesize kit (Roche). The hybridization and detection were conducted by following the manufacturer’s dihydrogen hybridization instruction.

Tissue expression of the lizard IGH genes

Tissue expression of the membrane-bound and secreted IgM were detected by PCR using the primers IgMCh3s (5′-CATCATGGGATCCCTCA CA-3′) and IgMtas3 (5′-ATCCCCATCTAGGTAAGCGGGT-3′) or lizard slgMas2. Full-length IgY expression was detected by using primer IgY deletions (5′-TTTCTGTACGCGTGACATGTC-3′) and IgY detections (5′-GGCCTCTTCACCTTCCAGAT-3′), whereas IgY (DFc) expression was detected by using primer IgY detections and IgYDfc deletions (5′-CAACATATTAGTGGCCCATCA-3′). Expression of IgD was detected by using a nested PCR using primers JH1 and IgDTmas1, JH2 and IgDTmas2.

Analysis of the expressed Ig isotypes in the small intestine

Small intestine total RNA was reverse-transcribed using the primer Norl-d(T)14 (5′-AAGCGATACAGAGAAGATTTTTTTTTTTTTTTTTTTTTTTTTTTT-3′) and Moloney murine leukemia virus reverse transcriptase. One round of 3′-RACE PCR was conducted using a mixture of IgM (5′-GAGGCTGTGATGCTGGATAGA-3′) and IgY (5′-GGAGGGGGATAGGACCAAC-3′) as primer and R2 (5′-AGAAATCGCCGCGAGAAG-3′) as sense primer and R2 (5′-AAATTGCCATGCTTATAGGAGA-3′) as antisense primer. The 3′-RACE PCR parameters were 94°C for 5 min for 1 cycle; followed by 40 cycles of 94°C for 30 s, 66.5°C for 30 s, and 72°C for 90 s and a final extension at 72°C for 7 min. The polymerase used was LA-TaqDNA Polymerase (Takara). The resulted amplified cDNA clones were sent for sequencing to identify the nature of the inserts.

DNA and protein sequence computations

DNA and protein sequence editing, alignments, and comparisons were performed using the DNAStar program. Phylogenetic trees were made using ClustalW (22) or ClustalX (32) (33) and viewed in TreeView (34). Multiple sequence alignments were performed using ClustalW. Accession numbers for the sequences (Table 3) used in phylogenetic analysis are as follows. α or γ gene: chicken, S40610; cow, AF109167; duck, A5134754; human, X14940; horned shark, X07781; lungfish, BAF1347; sheep, AF515673; catfish, U67437; mouse, J00475; platypus, AY055780; Xenopus laevis, BC072981; gene: catfish, U67437; sheep, AF109167; mouse, J00475; platypus, AY055780; gene: catfish, X52617; chicken, X01613; cow, J011096; dog, J011547; human, X14940; horned shark, X07781; lungfish, AF437724; mouse, Y00818; nurse shark, M90285; platypus, AY168639; skate, M29679; trout, X65261; X. laevis, BC084123; zebrational, AFX21840. μ gene: chicken, X01715; duck, X78273; X. laevis, X15114. γ gene: zebrational, X634732; trout, AY872256; gene: sandbar shark, U40560; lungfish, AF437727; nurse shark, U51450.

Results

The lizard IGH locus

Using the turtle μ gene sequence as a template (27), a BLAST search was performed against the green anole lizard whole-genome shotgun sequences deposited in the Ensemble database. A 369-kb
The lizard \( \mu \) gene

Four exons encoding sequences homologous (43.8% similarity at the overall protein level) to the turtle \( \mu \) C\( \mu\)1–4 sequences were identified downstream of the \( J_\mu \) locus. Phylogenetic analysis indicates that the identified gene is the lizard \( \mu \) gene (Fig. 3). Alignment of the lizard \( \mu \) sequence (inferred amino acids) with those from other species revealed a distinct pattern with regard to the distribution of noncanonical cysteines; i.e., those cysteines not involved in the intradomain disulfide bond (supplemental Fig. 3). Surprisingly, the cysteine residue in the N-terminal region of C\( \mu \)1, typically involved in disulfide bonding of heavy and L chains, is absent (supplemental Fig. 3). This was confirmed in both the retrieved genomic sequence and by sequencing of RT-PCR products generated using forward primers in the JH region and reverse primers in the 3’-region of the \( \mu \) sequence. In contrast, three noncanonical cysteines were observed in the C\( \mu \)2 domain (supplemental Fig. 3). The cysteine residues that covalently polymerize pentameric IgM in mammals (35) are also found in the C\( \mu \)3 and secreted tail of the lizard \( \mu \) gene (supplemental Fig. 3). Three potential N-linked glycosylation sites, one in C\( \mu \)1, one in C\( \mu \)2, and one in the secreted tail, are conserved in reptiles, birds, and Xenopus (supplemental Fig. 3). The first and third (invariant NVS) glycosylation sites are also conserved between reptiles and mammals (36, 37), suggesting a functional importance (37, 38).

To examine the splicing pattern of the lizard \( \mu \) gene transcript, we used primers in J\( \mu \)1, TM, or secreted tail sequences to perform nested RT-PCR amplifications. This approach revealed the presence of 4-C\( \mu \)1 secreted and membrane-bound forms (VDJ-C\( \mu \)1-C\( \mu \)2-C\( \mu \)3-C\( \mu \)4-s; VDJ-C\( \mu \)1-C\( \mu \)2-C\( \mu \)3-C\( \mu \)4-TM). Surprisingly, a short, membrane-bound encoding transcript, lacking the first two \( \mu \) domains (i.e., VDJ-C\( \mu \)1-C\( \mu \)2-C\( \mu \)3-C\( \mu \)4-s) was also observed (Fig. 4; supplemental Figs. 4–6). However, Northern blotting suggested that this short IgM transcript was a very minor form as it was not detectable using this approach (data not shown). Transcripts encoding the secreted form were detected only with four \( \mu \) exons (VDJ-C\( \mu \)1-C\( \mu \)2-C\( \mu \)3-C\( \mu \)4-s) in both PCR and Northern experiments (Figs. 4 and 5, supplemental Fig. 4, and data not shown). Transcripts encoding both the membrane-bound and secreted forms of \( \mu \) were detectable in multiple tissues (Fig. 5).

Characterization of the rearranged VDJ-C\( \mu \) fragments

To analyze the expressed VDJ-C\( \mu \) sequences, we used four V\( \mu \) sense primers that matched most of the V\( \mu \) genes identified in the current genome assembly (including those not assembled to scaffold 369), and an antisense primer derived from the IgM C\( \mu \)1 to amplify the rearranged VDJ-C\( \mu \) fragments by RT-PCR. We sequenced 92 clones, which provided 34 uniquely rearranged VDJ sequences (supplemental Fig. 7). Among these 34 clones, only 1 (LDA24) was expressed from the 8 V\( \mu \) gene segments (V\( \mu \)7) identified in scaffold 369. The frequency of J\( \mu \) utilization was J\( \mu \)1 (0), J\( \mu \)2 (7), J\( \mu \)3 (1), J\( \mu \)4 (3), J\( \mu \)5 (4), J\( \mu \)6 (5), J\( \mu \)7 (5), J\( \mu \)8 (7), and J\( \mu \)9 (2); see supplemental Fig. 7. The most frequently used J\( \mu \) segment were D\( \mu \)11 (or DH20, \( n = 10 \)) and D\( \mu \)2 (or DH4, \( n = 6 \)). The average length of the CDR3 was 8.9 codons, which is nearly the same as that in Xenopus (8.6 codons) and mice (8.7 codons); see Ref. 39. On average, the D gene plus N and P nucleotides contributed 43.1% of the CDR3 length, a ratio lower than that observed in X. laevis (57%) and adult mice (57%) (39).

Identification of the lizard \( \delta \) gene

An array of 11 C domain-encoding exons spanning ~12 kb was identified further downstream of the \( \mu \) gene. Two typical Ig transmembrane region-encoding exons were also observed 1 kb downstream of the C\( \delta \)11 exon. A BLAST search using the translated protein sequence from the 11 exons showed homology to the Xenopus and fugu IgD H chains, supporting the notion that it is the \( \delta \) ortholog in lizards. This conclusion was also supported by a phylogenetic analysis (Fig. 3 and supplemental Figs. 8 and 9).

All the \( \delta \) C\( \delta \)1 domains show C1-set Ig superfamily structure with seven strands, and canonical cysteines at invariant positions (Ref.
In the CH1, the second canonical tryptophan is replaced by a phenylalanine, in CH4 by a glycine, and in CH8 by a glutamine. As in Xenopus IgD (and in lungfish IgW; Refs. 6, 12, and 41), the N-terminal region of CH1 contains two consecutive cysteines, either of which could potentially provide the disulfide bond to the L chain.

**FIGURE 3.** Phylogenetic analysis of the lizard Ig genes. The tree was constructed using protein sequences of the last CH domains (CH8 for the lizard IgD, CH7 for the X. tropicalis IgD, and CH6 for the catfish IgD that show maximal homology to the mammalian δCH3 were used). Except the sequences of the lizard IgM, IgD, and IgY (obtained in this study), and X. tropicalis IgF, IgM, and IgX (12), the remaining Ig sequences were obtained from the NCBI GenBank.

**FIGURE 4.** Splicing patterns of the lizard Ig H chain genes. a, RNA splicing patterns of the IgM H chains; b, RNA splicing patterns of the IgY H chains; c, RNA splicing pattern of the IgD H chain. The CH exons are indicated in Arabic numbers.

**FIGURE 5.** Tissue expression of the lizard Ig genes. Although all other Ig genes were detected by one round of RT-PCR, IgD (transmembrane form) was detected only by two rounds (nested) of RT-PCR. TM, membrane-bound form; \( \gamma \) truncated form, IgY(ΔFc).
domain to domain comparisons of the lizard and human δ sequences revealed a relatively high identity between the lizard C_H7 and the human C_H2 (30.8%), and between the lizard C_H8 and human C_H3 (38.3%; supplemental Table I). Two potential N-linked glycosylation sites in the C_H7 and C_H8 are conserved across species when compared with the mammalian δ sequences (supplemental Fig. 10). Strikingly, C_H10 contains seven potential N-linked glycosylation sites.

The presence of 11 δC exons suggested that the lizard δ gene may be expressed as a long, multiple domain H chain as in Xenopus and bony fish (6, 9, 12). However, a nested RT-PCR using primers derived from the conserved JH sequence and the predicted δTM sequence generated a single product of ~1.4 kb. Sequencing of this amplicon showed that the first 4 δC exons were directly spliced to the δTM exon, with the remaining exons (C_H5–C_H11) being spliced out (Fig. 4 and supplemental Fig. 11). We were unable to identify the 3′-end of a secreted form using 3′-RACE PCR, due, perhaps, to the overall low level of expression of δ transcripts. The encoded lizard IgD H chain is thus more similar in size to mammalian IgD than to that of bony fish and Xenopus (Fig. 6). The membrane form of the lizard δ gene, however, appears to be weakly expressed given that even a two-round, nested RT-PCR generated only weak ampiclons in the spleen and lung (Fig. 5).

Identification of the lizard v gene

Approximately 17 kb downstream of the δTM2 exon, a third Ig gene consisting of four C_H-domain-encoding exons, plus two transmembrane region-encoding exons, was identified. A comparison of sequences, as well as phylogenetic analysis, indicated that this was the lizard v gene (Fig. 3). The entire gene, including four C_H- and two TM region-encoding exons, spans ~8 kb of DNA.

Alignment of the lizard IgY H chain (v) sequence with those of other species revealed several conserved features (supplemental Fig. 12). These included the three cysteines (potentially linking H chains and L chains) in the N-terminal region of the C_H1 and three noncanonical cysteines located in the N- and C-terminal regions of the C_H2 domain, all of which are conserved across the species examined. The two potential N-linked glycosylation sites in the lizard C_H2 and C_H3 are also found in birds, but not in Xenopus.

RT-PCR using primers for the Jv and secreted tail (or TM region) followed by sequencing of the amplicons revealed that the lizard IgY is expressed in secreted and membrane-bound forms that contain all four C-region exons (Fig. 4 and supplemental Figs. 13 and 14).

IgY(ΔFc) is expressed in the lizard

Based on serological data, a truncated IgY form has previously been reported in turtles (28). This is reminiscent of the situation in ducks, in which both full-length IgY and IgY(ΔFc) lacking the Fc region (C_H3) and C_H4, have been reported (16, 30). To confirm whether a truncated form is also expressed in the green anole lizard, we searched the C_H2-C_H3 intron for a potential transcriptional termination signal (or polyadenylation addition signal, AATAAA). Such signal sequences were found 464 bp downstream of the C_H2 exon. To identify the putative expression of the truncated IgY H chains, RT-PCR was conducted using primers for the predicted 3′-untranslated region and conserved Jv sequences. The sequences of the amplicon generated showed that a truncated IgY transcript was indeed present (Fig. 4 and supplemental Figs. 15–17) but expressed much weaker than the four-C_H forms in many tissues as shown by Northern blotting (data not shown).

Putative switch regions are present upstream of the μ and v but not δ genes

Expression of H chain isotypes other than IgM (IgD) in tetrapods is accomplished through class switch recombination, in which a functional VDJ region is translocated from upstream of an expressed μ gene to a position immediately 5′ of the H chain gene to be expressed (42). The recombination is mediated by switch (S) regions, which are located upstream of the μ gene and of the other H chain genes (except for δ, which is expressed by RNA processing rather than by class-switch recombination). The switch regions are typically repeat rich (in mammals; Ref. 42) and contain short repetitive sequences rich in AGCT motifs (43). Analysis by dot-plot under relatively relaxed conditions (window, 30 bp; percent match, 70%) of the JH-μ intron identified an ~3.2-kb DNA repetitive sequence block (supplemental Fig. 18a). In this 3.2-kb DNA region, 12 CTGGG (both strands) and 18 CTGAG (both strands) motifs that characterize the mammalian Sμ were found. The region is also rich in the AGCT motifs (supplemental Fig. 18a) that can promote isotype switching (43), suggesting that this region is most likely the Sμ. Using the same method, an ~5-kb Sv was also identified (supplemental Fig. 18c). When dot-plot analysis was applied to the μ-δ intron, some repetitive sequences (mainly due to the presence of two (ATT)_n microsatellite sites) could be identified. This intron is, however, very much less rich in AGCT motifs than are the putative Sμ and Sv regions (supplemental Fig. 18b), suggesting that there is no Sδ region in the lizard. The δ gene is thus likely to be expressed through cotranscription together with the μ gene, followed by alternative processing of the primary transcript, as occurs in other species that express IgD.

Discussion

Making use of the recently released whole-genome sequence combined with experimental investigation of gene expression, we report that the anole green lizard expresses three IgH isotypes (IgM, IgD, and IgY) but no IgA. These findings have implications for our
understanding of the evolutionary history of the vertebrate \textit{IGH} locus and the \textit{Ab} classes it encodes (Fig. 7). These results, as discussed below, also allow a more complete reconstruction of the evolutionary history of the vertebrate \textit{IGH} locus and suggest conclusions (some unexpected) regarding the structural and functional plasticity of this locus and of the \textit{Ab} that it encodes.

First, the demonstration of IgD in a lizard makes clear that IgD, an \textit{Ab} once thought to be present only in primates, shows an evolutionary continuity from fish, through the amphibia and reptiles to the mammals (Fig. 7). In those taxa such as the gecko; 2, IgD was lost during the emergence of the birds; 3, IgG and IgE evolved from IgY, and hinge regions for IgD, IgG and IgA were developed.

FIGURE 7. Phylogeny of the IGH isotypes in jawed vertebrates. Approximate times (unit, million years) of vertebrate divergence are shown in ovals. The filled circles indicate the first appearance of an Ig isotype. Genetic events: 1, IgA was lost during speciation of reptiles, being retained in some species such as the gecko; 2, IgD was lost during the emergence of the birds; 3, IgG and IgE evolved from IgY, and hinge regions for IgD, IgG and IgA were developed.

the \(\mu\) and \(\nu\) genes, with the first two C\textsubscript{H} domains being derived from the \(\nu C_4 H-1-2\) and the last two derived from \(\mu C_3-4\), a scenario that also applies to bird \(\alpha\) and \textit{Xenopus} IgX \textit{H} chains (15). The hypothesis that \(\alpha\) originated from a recombination between the \(\mu\) and \(\nu\) genes explains why only IgM and IgA can form polymeric Abs and also why IgM and IgA share the same polymeric Ig receptor (46, 47).

Indeed, sequence comparisons of the gecko \(\alpha\) with the lizard \(\mu\) and \(\nu\) show strikingly high similarities at the inferred protein level between the gecko \(C_4-1\) and the lizard \(\nu C_4-1\) (59.0% identity), and between the gecko \(C_3-4\) and the lizard \(\nu C_3-4\) (59.6% identity; supplemental Table II and supplemental Fig. 19). Dot plot analysis of the lizard \(\nu\) and \(\nu\) with the gecko \(\alpha\) gene supports the same conclusion (supplemental Fig. 20). Although IgA contain 4 C\textsubscript{H} domains in birds and reptiles, mammalian IgA possess only three and a hinge region. Phylogenetic analysis suggested that the C\textsubscript{H}2 domain (of 4-C\textsubscript{H} IgA) was lost in mammals (supplemental Fig. 21).

The position of IgA-encoding genes in the \textit{IGH} locus differs greatly between birds and mammals. In mammals it is found at the 3’-end of the locus, but in birds it is found (in an inverted transcriptional orientation) between the \(\mu\) and \(\nu\) genes. It is possible that, following its creation, the \(\alpha\) gene was in an unstable chromosomal location that favored deletion and/or recombination events, thus explaining the presence of the \(\alpha\) gene in different sites in birds and mammals and its absence in the lizard.

A broad overview of the Abs present in the jawed vertebrates, including the lizard, leads to two conclusions. First, certain classes of Ab (such as IgD or IgA) can be either present or absent from otherwise related species without any apparent effects on fitness. Second, structural variants of Abs (such as IgY) that are predicted to impair important functions are also well tolerated. An example of the latter is the presence in the lizard of two forms of IgY, one of which is of normal size (four-C-domain \textit{H} chain), but the other is truncated. The observation in the lizard of a message that encodes a truncated form of IgY, termed IgY(\(\Delta\text{Fc}\)) (16), mirrors findings made in ducks (30) and in turtles (28, 29). The IgY(\(\Delta\text{Fc}\)) of the lizard appears to be a structural equivalent of a F(ab’\textsubscript{2}) fragment, with a V\textsubscript{\textit{H}}-C\textsubscript{\textit{H}2}-C\textsubscript{\textit{H}3} domain structure of the \textit{H} chain. Such a fragment would be predicted to maintain virus-neutralizing functions but to be inactive in opsonization and complement fixation (16, 48), although some potential selective advantages of IgY(\(\Delta\text{Fc}\)) have been proposed, such as an inability to initiate anaphylactic reactions (48). In this study, it was clearly shown that the truncated \(\nu\) chain of the lizard has a genetic basis different from that described in ducks. In ducks, the C-terminal sequence of the \(\nu(\Delta\text{Fc})\) chain results from the use of a small additional exon that is present in the C\textsubscript{H}2-C\textsubscript{H}3 intron (supplemental Fig. 17) (30). However, the C-terminal encoding sequence and 3’-untranslated region of the lizard IgY(\(\Delta\text{Fc}\)) transcripts are located immediately downstream of the C\textsubscript{H}2 exon, encoding a slightly longer tail (GKTSCGLH; supplemental Figs. 16 and 17). Thus, processing of the primary \(\nu\) transcript in the lizard involves a choice between splicing from the donor site at the end of the CH2 exon to the acceptor site of the C\textsubscript{H}3 exon or cleavage/polyadenylation of the transcript at a site downstream of the C\textsubscript{H}2 splice donor site. This is reminiscent of the alternative RNA processing pathways associated with the expression of the secreted vs membrane bound forms of Ig \textit{H} chains (49). The expression of the membrane form of an Ig \textit{H} chain typically involves splicing of the TM1 exon into a cryptic site (consensus GGGGTAAAA) within the terminal spliced exon. The C\textsubscript{H}4 exons of the lizard \(\mu\) and \(\nu\) genes both use cryptic splice sites of this sequence for generating messages that encode the membrane-anchored form of IgM or IgY. In contrast, the 3’ splice site of the C\textsubscript{H}2 exon of the lizard \(\nu\) gene (AG/ GTAAG) conforms more closely to the canonical splice donor sequence (AG/GTGAG).
Thus, the IgY(ΔFc) of the lizard and of the duck arise from different genetic events and are examples of convergent evolution. This would argue for a selective advantage for the presence of IgY(ΔFc) in the Ab responses of lizards, sea turtles, and ducks, but the biological significance of IgY(ΔFc) expression in any of these species has yet to be established.

Overall, these observations on the Ig genes of the lizard reinforce several conclusions concerning the vertebrate IGH locus and the Abs it encodes. First, this locus shows remarkable evolutionary flexibility in the number of classes of Ab that are expressed in different species of jawed vertebrate. Furthermore, the patterns of expression (Fig. 7) show little consistency, outside the universal presence of IgM. Second, any single class of vertebrate Ab (such as IgD, IgA, or IgY) can show substantial structural variation, in ways that do not clearly correlate with obvious differentiated functions. Third, it is clear that Abs show both structural plasticity and functional redundancy. Although the evolutionary history of the Ab family is now better described, we are still far from understanding the functional significance of the remarkable diversity that has been uncovered.

Disclosures

The authors have no financial conflict of interest.

References
