Evidence of IgY Subclass Diversification in Snakes: Evolutionary Implications

Tao Wang, Yi Sun, Wenwei Shao, Gang Cheng, Lingxiao Li, Zubing Cao, Zhi Yang, Huiying Zou, Wei Zhang, Binyue Han, Yang Hu, Liming Ren, Xiaoxiang Hu, Ying Guo, Jing Fei, Lennart Hammarström, Ning Li and Yaofeng Zhao

*J Immunol* 2012; 189:3557-3565; Prepublished online 29 August 2012; doi: 10.4049/jimmunol.1200212

http://www.jimmunol.org/content/189/7/3557

Supplementary Material http://www.jimmunol.org/content/suppl/2012/08/29/jimmunol.1200212.DC1

References This article cites 50 articles, 17 of which you can access for free at: http://www.jimmunol.org/content/189/7/3557.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
Evidence of IgY Subclass Diversification in Snakes: Evolutionary Implications

Tao Wang,* Yi Sun,* Wenwei Shao,* Gang Cheng,* Lingxiao Li,* Zubing Cao,* Zhi Yang,* Huiying Zou,* Wei Zhang,* Binyue Han,* Yang Hu,* Liming Ren,* Xiaoxiang Hu,* Ying Guo,* Jing Fei,* Lennart Hammarström,† Ning Li,* and Yaofeng Zhao*‡*‡

Mammalian IgG and IgE are thought to have evolved from IgY of nonmammalian tetrapods; however, no diversification of IgY subclasses has been reported in reptiles or birds, which are phylogenetically close to mammals. To our knowledge, we report the first evidence of the presence of multiple IgY-encoding (ν) genes in snakes. Two ν genes were identified in the snake Elaphe taeniura, and three ν genes were identified in the Burmese python (Python molurus bivittatus). Although four of the ν genes displayed a conventional four-H chain C region exon structure, one of the ν genes in the Burmese python lacked the H chain C2 region 2 exon, thus exhibiting a structure similar to that of the mammalian γ genes. We developed mouse mAbs specific for the IgY1 and IgY2 of E. taeniura and showed that both were expressed in serum; each had two isoforms: one full-length and one truncated at the C terminus. The truncation was not caused by alternative splicing or transcriptional termination. We also identified the μ and δ genes, but no α gene, in both snakes. This study provides valuable clues for our understanding of Ig gene evolution in tetrapods. The Journal of Immunology, 2012, 189: 3557–3565.

Immunoglobulins are a group of molecules with pivotal roles in adaptive immunity and are present only in jawed vertebrates (1–3). Structurally, a typical Ig molecule consists of two identical H chains and two L chains, where the Ig isotype is based on the C regions. Despite having essentially similar functions, Igs and their encoding genes are extensively diversified in many aspects, such as genomic organization, diversity-generating mechanisms, and the number of classes (as well as subclasses) in different species. Several comparative studies have provided insight into Ig evolution in jawed vertebrates. IgM and IgD are the most primitive H chain (IgH) classes (4), and it is likely that many other IgH isotypes identified in jawed vertebrates evolved from these two classes by gene duplication, recombination, or additional genetic modifications (e.g., the duplication/deletion of a H chain C region (CH) exon or the creation of a hinge exon). Apparently, the eventual functional divergence of Igs was an essential outcome of these genetic events. Ig L chain (IgL) genes have also undergone dramatic genetic changes over hundreds of millions of years of evolution. Four primordial L chain isotypes (λ, κ, σ, and ι-cart) are present in cartilaginous fish, the most primitive jawed vertebrates (5). However, only some of these isotypes are maintained in other groups of jawed vertebrates; for instance, three isotypes (λ, κ, and σ) are conserved in bony fishes and amphibians (5–9), two isotypes (λ and κ) are conserved in reptiles and mammals (1, 10), and λ is the only one found in birds (1, 11, 12).

Ig genes have been relatively less studied in amphibians, reptiles, and birds than in mammals and fishes. Nevertheless, studies of the former three groups have helped to shed light on the evolution of the Ig genes (e.g., in terms of their structural variation and the mechanistic differences in the generation of Ab diversity). A notable example is the finding that chickens use gene conversion instead of V(D)J recombination as a major mechanism of generating Ig diversity (13). Analysis of the Xenopus Ig genes also uncovered a (genetic) hinge-containing Ig class (IgF), which was previously only observed in mammals (14). Along with these findings, many intriguing issues (such as no or very rare diversification of the subclass IgH and a loss of certain IgH and IgL isotypes in selected species) have arisen and have presented challenges to comparative immunology.

The expression of multiple IgH subclasses (mainly IgG and IgA) is common in mammals, with the opossum Monodelphis domestica as the only known exception (15–24). However, no subclass diversification of IgY (a functional equivalent of IgG) or IgA in nonmammalian tetrapods has been reported, although two IgD-encoding genes were recently observed in the reptile leopard gecko (25). IgF in Xenopus tropicalis shows some sequence similarity to IgY, but their genomic structures differ greatly. Although the driving force underlying subclass diversification is known, this process is particularly interesting in the context of Ig.
evolution. It is commonly accepted that both mammalian IgG and IgE are the descendants of IgY (26), despite IgY being structurally similar to IgE but distinct from IgG. Our recent investigation of IgH genes in the most primitive mammal, the duck-billed platypus, uncovered a novel IgH class with four CH domains and a hinge, which sheds light on the evolution of IgY to IgG (22). It seems clear that the divergence of IgY into both IgG and IgE occurred after the emergence of mammals; however, logically, there must have been an initial gene duplication that provided the genetic basis for its ultimate functional divergence. Therefore, it might be expected that some nonmammalian tetrapods express multiple IgY subclasses, although this has not been observed. While this article was under revision, Gambón-Deza et al. (27) reported finding of multiple IgY subclasses in snakes based on analysis of genome and transcriptome data.

Another interesting issue arising from studies of bird and reptile Ig genes is the loss of certain Ig isotypes; for instance, no IgD or κ L chains have been found in birds. Accompanying the absence of an IgD-encoding gene (δ) in chickens and ducks is the inversion of the α gene (28, 29), which seems to suggest that the bird IgH C gene region locus has experienced a major genetic rearrangement. Our recent finding of the absence of an IgA-encoding gene in the reptile anole green lizard provides additional support for this hypothesis (30). Recently, an IgA-like class was identified in the reptile Eublepharis macularius (31). This class is evidently not orthologous to bird and mammalian IgA but seems to have been derived from a recent recombination of IgM and IgY: the first two CH domains of this reptile IgA-like class are highly similar to IgY, and the last two CH domains are highly similar to IgM (30, 31).

IgA is specifically involved in protection against infection at mucosal surfaces, and its absence in certain reptiles is unexpected because nearly all tetrapods express IgA or its functional equivalent (e.g., IgX in amphibians) in their mucosal tissues (1, 2, 32). Even in bony fishes, IgT acts as a mucosal Ig (33), suggesting that a specialized Ig class for mucosal immune protection is especially important. In humans, some patients with IgA deficiency suffer from recurrent infections of the respiratory and gastrointestinal tracts and/or comorbid autoimmune diseases (34).

To further address the above-mentioned issues, we analyzed the IgH genes and their expression in two species of snakes.

Materials and Methods

Animals, RNA and DNA isolation, and reverse transcription

Snakes (Elaphe taeniura) were purchased from a snake farm located in Qingzhou, Shandong Province, China. An approximately 1-m-long captive-bred Burmese python (Python molurus bivittatus) was purchased in Tianjin, China. Genomic DNA was isolated using a standard phenol-chloroform extraction method. Total RNA from different tissues was prepared using a TRizol kit (Tiangen Biotech, Beijing, China). Reverse transcription was conducted using Moloney murine leukemia virus reverse transcriptase, following the manufacturer’s instructions (Invitrogen, Carlsbad, CA).

The amplification of IgM C region cDNA fragments using degenerate primers

Two degenerate primers (IgMUp: 5′-TGC SCK KKY WCT TTT CCC MCT C-3′; IgMDown: 5′-CTC GTG TCC CAC CAC GCA AGT-3′) were designed according to previously published IgM CH cDNAs of different species. RT-PCR with these two primers and snake spleen cDNA was carried out under the following conditions: 95°C for 5 min; 35 cycles of 95°C for 15 s, 60°C for 15 s, and 72°C for 1 min; and a final extension at 72°C for 7 min. The polymerase used was LA-Taq DNA polymerase (Takara, Dalian, China). The resultant 1.25-kb PCR product was cloned into the pMD19-T vector (Takara) and sequenced.

5′ RACE

Three IgC region-specific primers (RTigM: 5′-GCA GCT GTG TAT GTC-3′; IgMGsp1: 5′-GTT CCG CAAT TAC TGT GTC TGG G-3′; IgMGsp2: 5′-TGA TTT CAG GAA ATT CTT GGC GA-3′) were designed using the obtained snake IgM CH cDNA sequence. We used the 5′ RACE System for Rapid Amplification of cDNA Ends (Invitrogen). RTigM was used to synthesize first-strand cDNA. The RACE PCRs were performed by following the manufacturer’s instructions. The resultant PCR products were cloned into pMD19-T and sequenced.

3′ RACE and construction of the snake mini-IgH cDNA libraries

We used the 3′ RACE System for Rapid Amplification of cDNA Ends (Invitrogen). The RACE PCRs were performed by following the manufacturer’s instructions. The snake IgH gene-specific primer (IgH: 5′-GGG CAA AGG AAC AAT GTG CAC CTT CA-3′) was derived from the snake IgH joining-segment coding sequence. The PCRs were performed using both the spleen and intestine total cDNAs under the following conditions: 95°C for 5 min; 35 cycles of 95°C for 30 s, 65°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-Taq DNA polymerase. The resultant PCR products were cloned into pMD19-T to construct both the spleen and intestine mini-IgH cDNA libraries separately. These libraries were subject to further sequencing and PCR analysis for identification of the inserts.

Southern blotting

Nine pairs of primers were used to amplify the cDNA fragments (all from single CH exons) of IgH C regions in two types of snakes. These primers were as follows: snake IgM-CH2U: 5′-CGC TCT GTA AAC GTC ATC TCT-3′; snake IgM-CH1L: 5′-TTC ACA GTC TAT GAC ATT GAC GA-3′; snake IgD-CH1U: 5′-GAG CCA ACA AAA GAT CCT TCC ATC-3′; snake IgD-CH1L: 5′-GTC AAG ATA GCT GGG AGG GTA CCT TCT-3′; snake IgY1-CH1U: 5′-CCA ATC ACA TGC AGT CCT CTT CC-3′; snake IgY1-CH4L: 5′-AGG TTT CCA GGA ATT TGT TCC TAT G-3′; snake IgY2-CH4U: 5′-TCT TTG CCT CTC CTC ACC GAG A-3′; snake IgY2-CH4L: 5′-TTA CCC TGG GTT GAG ATG TTC C-3′; python IgY1-CHU: 5′-CCT CAC ATC CCT TCC GTA GAA C-3′; python IgY1-CHL: 5′-GAG GAT GGT GGC GTA TCT GGT TTA G-3′; python IgD-CH1U: 5′-AAG ATT CTG ACA GTG TAA CTG TCT GG-3′; python IgD-CH1L: 5′-CAI CTT GCT TCA GAG GTG GCC AT-3′; snake IgY1-CH1U: 5′-TTC CCA GAT GAG ACA ACA CAG CAA G-3′; python IgY1-CHL: 5′-ACC CTG AGA TTT TCT GAT CCT C-3′; python IgY2-CH1U: 5′-AGG AGA ACT CAA CCC AGG AAC C-3′; python IgY2-CH1L: 5′-CTG GAT GGC GTA TCT CCT TCC TAC G-3′; python IgY3-CH1U: 5′-GTC CCG TAT GAG TTT ACG ATC TCT C-3′; python IgY3-CH4L: 5′-TCC TCC CAT TGA AGT GAC CCA CAC-3′; python IgY3-CH4L: 5′-ATG TTT CCA GGA AGA CAG TAT TTT C-3′; and python IgY3-CH4L: 5′-ATG TTT CCA GGA AGA CAG TAT TTT C-3′. These cDNA fragments were labeled using a PCR DIG Probe Synthesis kit, and hybridization and detection were performed using the DIG High-Prime DNA Labeling and Detection Starter Kit II, following the manufacturer’s instructions.

Detection of IgH gene expression in different tissues by quantitative real-time PCR

cDNA samples of eight tissues (heart, liver, spleen, lung, kidney, intestine, stomach, and testis) were used to detect IgM, IgY1, and IgY2 expression by quantitative real-time PCR. PCRs were performed using the LightCycler 48 and LightCycler 480 SYBR Green I Master. Each sample was run in triplicate. The snake EF1α gene was chosen as the internal control. The PCRs were performed using both the spleen and intestine total cDNAs under the following conditions: 95°C for 10 min; 35 cycles of 95°C for 10 s, 60°C for 20 s, and 72°C for 15 s; and a final extension at 72°C for 7 min. The polymerase used was LA-Taq DNA polymerase. The resultant PCR products were cloned into pMD19-T to construct both the spleen and intestine mini-IgH cDNA libraries separately. These libraries were subject to further sequencing and PCR analysis for identification of the inserts.

Preparation of snake IgMs, IgY1-, and IgY2-specific mAbs

Mouse mAbs against snake IgM, IgY1, and IgY2 were prepared by Abmart (Shanghai, China). Briefly, cDNA fragments of snake IgM (CH1–CH3), IgY1 (CH2–CH4), and IgY2 (CH2–CH4) genes were cloned into pEFl-28a (+) expression vector (Novagen, Darmstadt, Germany) and fused with a His tag. These fused cDNA fragments were expressed in an Escherichia coli-expression system and purified using Ni-column chromatography. The purified proteins and Freund’s adjuvant were used to immunize BALB/c
mice. Hybridoma cells were prepared by fusing the splenocytes derived from the immunized mice with Sp2/0 mouse myeloma cells. The specificity of each secreting mAb was examined using ELISA. Selected hybridomas were injected into the mouse peritoneal cavity to induce ascites.

Western blotting

Serum was prepared by centrifugation of the whole blood at 3000 × g after incubation for 3 h at 4°C to exclude the erythrocytes. Denatured serum treated with DTT (5.0 μM) and/or peptide-N4-(N-acetyl-[β]-glucosamino) asparagine amidase (PNGase F) and endo-α-N-acetylgalactosaminidase (NEP, Ipswich, MA) was separated on 12% SDS denaturing polyacrylamide gels and transferred onto Amersham Hybond N+ membranes (GE Healthcare, Little Chalfont, U.K.). Nonspecific binding was blocked with TBST containing 5% nonfat milk. The membranes were then washed three times with TBST. HRP-conjugated goat anti-mouse secondary Abs (Santa Cruz Biotechnology, Santa Cruz, CA) were added and incubated for 1 h at room temperature. Following five washes with TBST, the protein bands were visualized with ECL (Thermo Fisher Scientific, Waltham, MA) by exposing them to film (Kodak, New York, NY).

Immunoprecipitation

Immunoprecipitation was performed using protein G magnetic beads (NEB), following the manufacturer’s instructions. The immunoprecipitated targeted proteins were excised from the gels and subjected to liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis.

DNA and protein sequence computations

DNA and protein sequence editing, alignments, and comparisons were performed using the DNASTAR program. Phylogenetic trees were generated using MrBayes3.1.2 (35) and viewed in TreeView (36). Multiple-sequence alignments were performed using ClustalW.

Results

Analysis of the expressed V and J segments in the snake

We initially set out to analyze the expressed V and J segments in the snake E. taeniura by cloning the 5′ RACE products of IgH chain cDNAs. We first amplified a 1.25-kb IgM H chain C region fragment using two degenerate primers (IgMUp and IgMDown) that were designed according to previously published IgM H chain sequences of other species. The identity of the amplified fragment was confirmed by sequencing and basic local alignment search tool analysis. RACE primers were subsequently designed and used to amplify the V and J segments that were associated with IgH chains. We sequenced 335 clones, of which 271 clones with intact H chain variable segments (VH) and JH segments (coding sequences) were used in further analyses. In total, 245 clones contained unique rearranged VDJ sequences (CDR3). Six distinct JH segments could be identified in these 245 clones, but they were present at different frequencies (JH1, 33/245; JH2, 7/245; JH3, 19/245; JH4, 77/245; JH5, 28/245; and JH6, 81/245) (Supplemental Fig. 1A). The JH6 sequences were present in almost one third of these clones. We did not analyze the usage of DH gene segments because of a lack of germ-line DH sequences. Further analysis of the CDR3 length in these 245 IgH chain clones showed that it varied from 4 to 18 aa, with CDR3 lengths of 8 aa (37/245), 9 aa (44/245), and 10 aa (41/245) being the most common. The average length was 9.55 ± 2.34 aa, which is slightly longer than that in lizards (8.9 aa) (30), Xenopus (8.6 aa), and mice (8.7 aa) (37).

All expressed VH in the 245 clones could be classified into seven families (VH I–VH VII), based on the criterion that the members in a single family share ≥70% sequence identity (38, 39) (Supplemental Fig. 1B). More than half of the expressed VH sequences belonged to one family, VH I (154/245), which was followed in frequency by VH II (28/245) and VH III (25/245). Phylogenetic analysis of the snake VH families together with the VH genes of other species showed that five families (VH I, VH II, VH IV, VH V, and VH VI) fell into clans II and III, which are previous classifications defined according to the mammalian VH genes (40), and the remaining two families were separately clustered with fish sequences, suggesting that the snake shares VH genes of both fish and mammals (Fig. 1).

Expression of IgM, IgD, and two IgY subclasses in the snake E. taeniura

The distinct IgH isotypes can share identical JH segments. To analyze the expressed IgH isotypes in the snake, we performed 3′ RACE using JH-derived primers and total cDNAs from both the spleen and intestine. The 3′ RACE products derived from the two tissues were cloned into a T vector separately to produce two IgH-specific mini-libraries, which were subject to further PCR analysis and sequencing to identify the inserts. A total of 1019 clones in the spleen library was analyzed; 61 IgM, 784 IgY1, 173 IgY2, and 1 IgD H chain clones were identified. Most clones in the small intestine library (914 clones) were IgM (890/914); only 19 IgY1 and 5 IgY2 clones were identified. As in the green anole lizard and Chinese soft-shelled turtle (30, 41), no IgA H chain clone was identified.

Characterization of the snake IgM CH gene

The identity of the snake μ was confirmed by phylogenetic analysis (Fig. 2). The complete μ H chain C region cDNA (coding for a secreted form) was obtained by sequencing the cloned 3′ RACE products. This cDNA encoded four CH domains. The deduced amino acid sequence showed a conserved structure compared with IgMs of other tetrapods. Phylogenetic analysis suggested that the snake μ is most closely related to its counterpart in the green anole lizard (Fig. 2). As in other species, the snake μ contained two conserved N-linked glycosylation sites: one in the CH1 domain, and one in the CH4 domain (Supplemental Fig. 2A). Similar to the lizard μ, the snake μ also lacked the cysteine residue in the N terminus of the CH1 domain, which is usually used to covalently associate with L chains (Supplemental Fig. 2A).

Further sequencing of the genomic sequence of the IgM CH showed that the four CH domains were respectively encoded by four CH exons in the snake (data not shown). The presence of this gene in the snake genome was confirmed by Southern blotting (Fig. 3A), and quantitative real-time PCR showed that it was highly expressed in the spleen, lung, and intestine (Fig. 4A).

Characterization of the snake IgD CH gene

In our screening of the spleen mini-IgH–specific library, we identified a 1.2-kb IgH transcript. Further analysis indicated that this cDNA encoded three CH domains and appeared to represent a secreted IgD H chain transcript. This secreted IgD H chain was mainly expressed in the spleen and was also detectable in the lung, stomach, and intestine (Fig. 4B). The genomic δ genes usually comprise multiple (up to 11) CH-encoding exons in non-mammalian tetrapods and share a conserved transmembrane (TM) region (42). To further characterize the snake δ gene, we used a JH primer and a degenerate primer spanning the 6TM to amplify the IgD H chain transcripts. This attempt generated four major PCR bands with sizes of 3.6, 2.4, 2.0, and 1.4 kb (Fig. 5). Sequencing of these bands revealed that the longest band encoded 11 CH domains, whereas the remaining three encoded 7, 6, and 4 CH domains. The δ gene in both the green anole lizard and the leopard gecko contains 11 CH exons at the genomic level (25, 30). A thorough comparison of the snake 3.6-kb product (11 CH domains) with the homologous sequences of lizard and gecko suggested that they are closely related; their CH domains exhibit a nearly perfect
1:1 correspondence in their order based on the sequence homology (data not shown).

We elucidated the genomic organization of the δ gene by determining the sizes of the introns using long-distance PCR. From this result, it was clear that the above-described amplified 2.4-kb transcript (seven CH) was generated by splicing the first four CH exons (CH1–4) onto the last three exons (CH9–11), whereas the 2.0-kb transcript contained the first six CH exons (CH1–6). The 1.4-kb transcript was a splicing product of the first four CH exons (CH1–4) and the TM exons (Fig. 5). These findings reveal a more complicated splicing pattern in the snake than in the green anole lizard, because the lizard exclusively expresses the shortest 4-CH transcript, despite the presence of 11 CH exons in the germ-line δ locus.

Two IgY subclasses are expressed in the snake

Two distinct IgY H chain transcripts (termed IgY1 and IgY2, or γ1 and γ2 for the genes, respectively) were identified in the snake. These transcripts exhibited 68% sequence identity at the cDNA level and 58.5% identity at the protein level. Although only two IgY H chains were identified at the transcriptional level, probes using CH1 exons of IgY1 and IgY2 both could detect multiple bands in Southern blotting on the snake genomic DNA (Fig. 3A). This strongly suggested that there are additional γ genes present in the snake genome that might be pseudogenized.

The apparent differences between the deduced amino acid sequences of these genes were that IgY1 lacked a cysteine in the N terminus of the CH1 domain, which is necessary for association with L chains (Supplemental Fig. 2B), and the two genes showed different N-linked glycosylation sites in their CH1 and CH4 domains (Supplemental Fig. 2B). Quantitative real-time PCR showed that both IgY1 and IgY2 were highly expressed in the spleen, as well as at relatively high levels in the lung and intestine (Fig. 4C, 4D). Although the total IgY (IgY1 and IgY2) expression level was approximately the same as that of IgM in the spleen, it was expressed at much lower levels than IgM in the lung, stomach, and intestine (Fig. 4E), suggesting that IgM plays a role in mucosal defense in the absence of IgA.

The detection of IgM, IgY1, and IgY2 in the serum

To detect the expression of IgM, IgY1, and IgY2 at the protein level, we developed mouse mAbs specific for the H chain C regions of these isotypes. The specificity of these mAbs was first tested by Western blotting using the corresponding GFP-fused CH expressed in CHO cells. Because IgY1 and IgY2 share a relatively high sequence identity, all initially developed mAbs for these two isotypes were tested on both GFP-IgY1 and GFP-IgY2 (Supplemental Fig. 3). After the initial screening process, three mAbs—4E7 (specific to IgM), 1A11 (specific to IgY1), and 4G6 (specific to IgY2)—were chosen for Western blotting of the serum samples.
The IgM H chain was predicted to be ∼65 kDa in size. Under reducing conditions, the actual size of the IgM H chain in serum was ∼85–95 kDa, with variable molecular weights (Fig. 6A), suggesting that the serum IgM was heterogeneously glycosylated. The protein being glycosylated was confirmed when the serum sample was treated with PNGase F to remove N-linked glycosylations, which produced a sharp ∼65-kDa band (Fig. 6A). Under nonreducing conditions, a weaker band of ∼450 kDa and a stronger band, far larger than 500 kDa, were observed (Fig. 6B), indicating that the IgM is present in serum as both dimers and polymers.

Western blotting analysis for serum IgY1 and IgY2 generated somewhat unexpected results, because two bands (∼60 and ∼70 kDa) were detected for each subclass (Fig. 6C). The sizes of these bands were approximately equal to or slightly greater than predicted (∼62 kDa). The size difference between the two detected bands for each IgY subclass was not caused by glycosylation, because the sizes of the bands were equally reduced when treated with PNGase F (removing N-linked saccharides) (Fig. 6C) and were not changed by endo-α-N-acetylgalactosaminidase treatment (removing O-linked saccharides) (data not shown). These data suggest that both IgY subclasses were glycosylated by N-linked, but not by O-linked, saccharides. In addition, they suggest that both IgY subclasses were expressed as two isoforms with different sizes and that glycosylation was not associated with the size difference. It is notable that the sizes of the lower bands for both IgY1 and IgY2 (∼55 kDa) were even smaller than predicted (62 kDa), indicating that they might be derived from an alternative splicing of the two IgY genes that both had a portion of the IgY coding sequence deleted. However, further investigation indicated that this was not the case. No alternative splicing of either IgY gene was observed using RT-PCR with several pairs of primers. To exclude the possibility that the two bands detected were not IgY H chain proteins, we performed an immunoprecipitation of serum IgY2 with the Ab generated earlier (Fig. 6D). Both bands were pulled down and subjected to LC-MS/MS (Supplemental Fig. 4), which revealed that both bands were indeed IgY2 and that the lower band represented a truncated form at the C terminus, although exactly where it was truncated could not be determined. Under nonreducing and denaturing conditions, both IgY subclasses were detected as two major bands of ∼240 and 280 kDa, in addition to a weak band of ∼170 kDa (Fig. 6E). The 170-kDa band was roughly equal to the molecular mass of an IgY monomer (two H chains plus two L chains). However, the two major bands (∼240 and 280 kDa) seemed to be aggregates of four H chain subunits, because under reducing conditions, the molecular mass of two variants of each IgY subclass were ∼60 and ∼70 kDa.
Association of the snake IgY H chains with L chains

Compared with IgY2, IgY1 lacks the cysteine in the N terminus of the CH1 domain, which is necessary for association with L chains (Supplemental Fig. 2B). In Western blotting under nonreducing conditions, both IgY1 and IgY2 were similarly detected as three bands (170, 240, and 280 kDa), of which the weak 170-kDa band is roughly equal to the molecular mass of an IgY monomer with two H chains and two L chains. To address whether the snake IgY H chains are associated with L chains, we analyzed whether L chains could be pulled down together with IgY H chains by immunoprecipitation. We first determined the sequence of the snake Ig L chain C region cDNA, which was amplified with degenerated primers. LC-MS/MS analysis of an ~25-kDa band pulled down by the IgY2-specific mAb showed a good match to the snake Ig L chain sequence (Fig. 6F, Supplemental Fig. 4B). Because the same pattern was detected in nonreducing Western blotting for IgY1 and IgY2, we speculate that IgY1 H chains are able to associate with L chains. The association is through either covalent bonds, which are formed by cysteines in unusual positions, or noncovalent bonds, like in bullfrogs (43).

The identification of three IgY subclass-encoding genes in the Burmese python (P. molurus bivittatus)

The genome of the Burmese python was recently sequenced using Illumina technology and deposited in the National Center for Biotechnology Information GenBank under accession number AEQU010000000 (44). By performing basic local alignment search tool against the genome sequence, we identified μ, δ, and three ν genes (ν1, ν2, and ν3) but no α. The presence of these genes in the python genome was confirmed by Southern blotting (Fig. 3B), and all five of these genes encoded actively transcribed RNAs, as demonstrated by RT-PCR. The μ gene encoded 4 CH domains, and the δ gene encoded 11; two ν genes (ν1 and ν2) encoded 4 CH domains. However, the ν3 gene encoded three CH domains at the genomic and RNA levels, but it lacked the CH2 domain.

**FIGURE 3.** Southern blotting of IgH C region genes in the snake *E. taeniura* and Burmese python. (A) Detection of IgH C region genes in the snake *E. taeniura*. (B) Detection of IgH C region genes in the Burmese python. E, EcoRI; H, HindIII; P, PstI.

**FIGURE 4.** The relative expression levels of μ, δ, ν1, and ν2 in different tissues. The data are representative of three independent experiments. The y-axis indicates fold normalized expression. Vertical lines are the standard deviations from the mean. (A) Expression level of μ in different tissues. (B) Expression level of the secreted δ in different tissues. (C) Expression level of ν1 in different tissues. (D) Expression level of ν2 in different tissues. (E) Relative expression level of μ to both ν1 and ν2 in different tissues.
corresponding to those of \(v1\) and \(v2\). Therefore, the Burmese python \(v3\) gene shows a similar genomic structure as the mammalian \(\gamma\) genes except that a dedicated genetic hinge exon was not found.

When a phylogenetic tree including all five \(v\) genes from both examined snake species was constructed, these genes were not clustered in a species-specific manner but instead were intermixed (Fig. 2). This finding clearly suggests that the subclass diversification of at least the \(v1\) and \(v2\) genes occurred before the two species diverged. Because \(E. taeniura\) and \(P. molurus bivittatus\) belong to Caenophidia and Henophidia, respectively, two superfamilies of the suborder Serpentes (snakes), it may be expected that most snakes in these two superfamilies express more than one IgY subclass.

**Discussion**

Investigations of Ig genes of two closely related groups of animals, birds and reptiles, have led to unusual findings in the context of Ig gene evolution. For example, although its counterparts or functional equivalents have been observed in most jawed vertebrates (2, 33), the IgA H chain encoding gene, which is important for mucosal defense, has been lost in some reptiles and shows an inverted orientation in birds (28–30). It is also notable that birds have lost the genes encoding both IgD and the \(k\) L chain (28, 29). Compared with mammalian species, it is surprising that both reptiles and birds rarely show diversification of IgH subclasses, a common phenomenon, even in the most primitive mammals, the monotremes (22). In contrast, in both birds and reptiles, the only known species that expresses IgH subclasses is the leopard gecko, which possesses two IgD genes within its IgH locus (25).

We were not able to clone the \(\alpha\) gene from the snake \(E. taeniura\) or to identify the \(\alpha\) gene from the genome data of the Burmese python, indicating that the gene might be absent in these species. This result is consistent with our study of the lizard IgH genes (30). If this result is correct, it will be interesting to examine which Ab class is responsible for mucosal immunity in these species.

**Figure 5.** Alternative splicing of the snake \(d\) gene. CH domain encoding exons are indicated by Arabic numbers. 1, Amplified PCR products; M, DNA marker; Sec, secreted exon; TM, transmembrane exon.

**Figure 6.** Western blotting detection of IgM, IgY1, and IgY2 in the serum and the immunoprecipitation of IgY2. (A) The detection of IgM. Lane 1: serum treated only with SDS and DTT; lane 2: serum treated with SDS, DTT, and PNGase F. (B) The detection of IgM under nonreducing conditions (without DTT treatment). (C) The detection of IgY1 (lanes 1 and 2) and IgY2 (lanes 3 and 4). Lanes 1 and 3: serum treated with SDS and DTT; lanes 2 and 4: serum treated with SDS, DTT, and PNGase F. (D) Immunoprecipitation of IgY2. Lane 1: mouse mAb used for immunoprecipitation; lane 2: immunoprecipitated IgY2 together with the mouse mAb. (E) The determination of the molecular mass of IgY1 (lane 1) and IgY2 (lane 2) under nonreducing conditions. (F) Silver staining of immunoprecipitated IgY2 under reducing conditions. Lane 1: mouse mAb used for immunoprecipitation; lane 2: immunoprecipitated IgY2. The arrow indicates the putative L chain band that was subject to LC-MS/MS.
species because, even in bony fishes, a specialized Ab class, IgT, has been confirmed to function in mucosal immunity (33).

In the current study, we found that snakes can express multiple IgY subclasses at both the mRNA and protein levels. This is a significant finding when considered in an evolutionary context. Ig subclass diversification in mammals is largely restricted to IgG and IgA, which are the key players in humoral and mucosal immunity. The corresponding functional divergence in complement activation and opsonization of these subclasses enables a more subtle immune response to infection with different pathogens. The divergence of IgY may enable the snakes to respond more specifically to varied infections by pathogens through the effector functions of the IgY Fc regions. More importantly, the identification of IgY subclasses provides additional support for the commonly accepted hypothesis that mammalian IgG and IgE were evolutionarily derived from IgY (26).

The functional and structural divergence of IgG and IgE must have been preceded by gene duplication in the mammalian lineage. Evidence for this model was recently obtained from a study of the IgH locus of the duck-billed platypus, a primitive mammal, in which an IgO-encoding gene was identified (22). The gene encodes four CH domains that are homologous to IgY, IgE, and IgG, as well as a hinge region attached to the CH2 domain. Because IgO is phylogenetically related to IgY, IgE, and IgG, and it possesses four CH domains, Gambão-Deza et al. (45) even suggested that it was an IgY equivalent in platypus. Despite having four CH domains, IgO also possesses a hinge attached to its CH2 and, thus, looks like an evolutionary intermediate between IgY (or IgE) and IgG. The coexistence of IgO and IgE in the platypus IgH gene locus strongly suggests that duplication of four-CH Ig genes (IgY or IgY-like genes) should have preceded the divergence of IgG and IgE in mammals. The finding of multiple IgY classes in snakes shows that the structural divergence of this Ig class occurred early in nonmammalian tetrapods. During the evolution of IgY to IgG, an evident outcome is the loss of the CH2 domain. In this regard, it is interesting that the v3 gene in the python already lost the CH2-encoding exon, showing a similar organization as the mammalian γ genes, although no genetic hinge exon was observed. It would be interesting to examine whether these IgY subclasses are dedicated to different effector functions, similar to IgG and IgE in mammals. The IgY subclasses in the two snakes investigated share ~60–80% sequence identity at the protein level and 70–90% sequence identity at the DNA level. It is evident that variations of some key residues, such as cysteine, and N-linked glycosylation sites are present between these subclasses, which may potentially lead to structural or functional divergence. In mammals, functional sequence divergence of IgG subclasses has been found in the sites that interact with FcRs (15). Because no homologs of the corresponding mammalian IgG FcRs have been analyzed in nonmammalian tetrapods, and the FcR binding sites of mammalian IgG are not conserved in the corresponding positions of IgY, it is difficult to predict whether the sequence variations between the snake IgY subclasses influence their receptor-binding capacity. Nonetheless, some sequence sites in mammalian IgG that are known to be essential for complement or protein G binding can be observed in IgY, and sequence variations are present in these sites among different IgY subclasses. It can be speculated that at least some of the sequence variations among these IgY subclasses should be associated with functional divergence.

Compared with other IgY subclasses, IgY1 in both E. taeniura and the Burmese python lack the cysteine in the N terminus of the CH1 domain, which is necessary for association with L chains (Supplemental Fig. 2B). However, unlike H chain-only Abs in camels and sharks (46, 47), our data suggest that IgY1 most likely is still associated with L chains. The association is through either covalent bonds, which are formed by cysteines in unusual positions, or noncovalent bonds, like in bullfrogs (43).

Another interesting finding in this study is that a single IgY gene can be expressed as two variants with different molecular masses in the serum and that one variant is truncated at the C terminus, although it is not known where the truncation occurs. The truncation is not a result of alternative splicing, or, as in ducks, γ gene transcription can be terminated at different sites, which generates both a full-length and a truncated IgY(ΔFc) (48, 49). In both cases, the IgY variants are encoded by different RNA transcripts. This does not apply to snake IgY; we could identify only a single transcript encoding a secreted IgY (by both 3′ RACE and conventional PCR). It is likely that both IgY variants are translated from a single mRNA transcript. It is not known whether translational recoding, a mechanism that can cause frame shifting and, thus, the production of two or more peptides from a single mRNA (50), is involved in the expression of the snake IgY variants.

In summary, this study finds that nonmammalian tetrapods can express multiple IgY subclasses. The present findings provide valuable insight into the functional divergence of IgY into IgG and IgE in mammals.

Disclosures

The authors have no financial conflicts of interest.

References

Supplemental figure legends

Supplemental Figure 1. Sequence alignment of the six JH segments and phylogenetic analysis of the snake VH families. (a): Sequence alignment of the six JH segments. (b): Phylogenetic analysis of the snake VH families.

Supplemental Figure 2. Alignment of amino acid sequence of IgM and IgY of different species. Dots denote identical amino acids, whereas dashes indicate gaps. Canonical cysteines and tryptophans are shaded, and the conserved N-linked glycosylation sites are in red and underlined. (a) Alignment of amino acid sequence of IgM of different species. (b) Alignment of amino acid sequence of IgY of different species. Potential sites for protein A, G and C1q complement are indicated.

Supplemental Figure 3. Western blotting test of the specificity and cross-reactivity of mouse mAbs for IgY1 and IgY2. (a) Test of mAbs for IgY1. Lane A-G were loaded with GFP-fused IgY1 protein which was over-expressed in CHO 293T cells while lane a-g were loaded with GFP-fused IgY2. Anti-GFP antibody was used as positive control in lane A&a to estimate the specificity, 6 mAbs for IgY1 were applied to the remaining lanes. The mAb was chosen only if it reacts strongly with GFP-fused IgY1 but not with GFP-fused IgY2. (b) Lane A-F were loaded with GFP-fused IgY1 proteins and lane a-f were loaded with GFP-fused IgY2. Anti-GFP antibody was used in lane A and F, a and f, 4 mAbs for IgY2 were applied to the remaining lanes.
Supplemental Figure 4. LC-MS/MS analysis of immunoprecipitated snake I gY2. (a). Analysis of heavy chains. (b). Analysis of light chains. Matched peptides were shown in bold red.
Supplemental Figure 1.

a.

JH1: TTTGATGTATGGGGACAAGGAACTGCCGTGACAGTAACTAGT
  FD V W G Q G T A V T V T S
JH2: TTCGATGTCTGGGGACCCGTTACTCCAGTCACGTCACTTCA
  F D V W G P G T P V T V T S
JH3: TTCGGTTATTGGGGCAAAGGAACCATGGTACGGTCTCTTCA
  F G Y W G Q G T M V T V S S
JH4: TTGGATTACTGGGGCAAGGAACATGGTACCGTACGCTCA
  L D Y W G K G T M V T V S S
JH5: TTTGACTATTGGGAAGGCAACTCTGGTACCGTACCTCTTCA
  F D Y W G K G T S V T V S S
JH6: TTCGAGTACTGGGGTAAAGGAACATCGGTTAGTCACCTCA
  F E Y W G K G T S V V V T S

b.
**Supplemental Figure 2.**

a.

```
Snake M: ESPKAPSLFLIPS--GDNSTGADIALQIAKNFLPNSI5DFPS---WNLPLQNYQNVYKFKISL---GGTHTAAPQAKVSLDS--WNSLSKPYGAKHPSQGKVTVSVERQGFGPSREPMSL--
Python M: .P.S........-----EK...ST..V..SL........DAVT.-------E.LSAESP.........RNPA...F......E.....EI--.--.TY.....F...TD.S...NG.VA.AKP-------QQ...-
X.laevis M: ATS.S.---------CGESMDP--VT..S.L....ET..S.TGKDN.A...YSTGGLSYK.VMQ...-S...S.S...VN.ASA...-VQDNIEQF...QN...LDTI.SVELKDPVW.VKPVVSI-
Duck M: AT.RG.T....L-.CSSSS.SSSLY.V..V.DGHV.AGV..-..TDVTNATVATTI.N..EARG-P.GNW...RLEL.P.QEKGQ--Q.R.Q.F.P.R.R.NP.VLA.SNP.SSQ.TA.V---
Platypus M: AT.G..L....V-.CGDSNQESSL.V..L...TL..DFKNNTV.TKEFINY.TV...QA.ST..LMI.ATEVMKQDDYVF...AV.N.R.-.FLRVAQPS..PVUS.H-
Horse M: .T.T.D....V-.CG-PSDLXSLV.V..L..RD...DT..GT..DNA.S.VRDIK.MNQGTVV..VQ..N.EKN..-PLPVIAELPKVSVF
```
b.

Snake Y1: GTPKAP-TVFPLSPODQVSTSDNVSVGGKSYFPETATQVNSGAGTSGIHHPPFVNHSS-GHYTHSSLTITPCQSSETPCQDVHEATHTRINKNIER--CDHEPV
Snake Y2: AS.T.........--CNQDILT.L.Q........N.........MLG.--N.........S.........VSK.Q....S.E...N.A...R.D.P.--EVRS.A
Turtle Y: --.R--S....TS--CFGEA.T.Q.TF...O.........V.........STP--N--VKY.V.S.LQT.S.L.SLS.QV.VSDDSS.DN.YHT.K.F..SSS.T.E.FK----SST.Q
Chicken Y: A.S.TS.PRLY...A-CCS--DSAVPPAV...LSPSSAG--GQSGQR---G.ATAVAG---RTSV.FPVKL.PVR3--PGKRRK.VEGAPGD--ALLK.EVQCVRDPVFV.

Snake Y1: APRVHLHSSCNFKSGDSGTIALV.FIKSFYPKEINVENCVVGSHQGGLTFTEPPFRRDSNTYTFSTNSVYNNIFTQEDNLEGNTYEVTHAASQTRKSRARKCEDAI-SHFSTP
Python Y1: ...............E.R....P..L.....I.Q.............H............A.Y.....H.VSG....TNT...A............E....KV.........DS-..I.SG
Python Y3: ..............................................................-N.ICCG
Snake Y2: G.E..VF....GSR--....Q.W.Q.SG....PLEIQ.K...RS....R.YNY..WQNPG......A.........O.........V.....E...E...KM..K.K..GGS-DCP.GD
Lizard Y: .D.R........S.RST.AS.E....ILSN.....VS...I.GKS...PSY...RQ.AVG.....T SA......A.....TI.T.K...G...TMRA...K...DS-TGQCDG
Gecko Y: T.B.L..RNDTIA.A.Q...L.SG...RTVKN....N.RQ1L1.PAH.N.RK.ATGD...T.TA...V.......K.K...TI...I.Q--HKVEHK.T..RGDSQNTCA
Turtle Y: .PQ.VY--.KLT.P--.D.L.HVSG...TPVT...D.GES.V.HGE.A.AT.AGGR...H.NAVS--FMQDKM.TGR.S.POTG.TEDDY--K.VP--TTTSS
<table>
<thead>
<tr>
<th>LM1S</th>
<th>D</th>
<th>Clq</th>
<th>OD</th>
<th>SpG</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>350</td>
<td>360</td>
<td>370</td>
<td>380</td>
<td>390</td>
<td>400</td>
<td>410</td>
</tr>
<tr>
<td>Snake</td>
<td>Y1 :</td>
<td>ITTYLLFPPDKLYSRRKAIHVTVQNLQDEGLKISITREKQVAVHEVIATENQGFTAVSRPLQAGDSHEFTTEVSVEYPLGPIEKTIYTKSREKSKAPIIHYVF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Python Y1 :</td>
<td>.N....A.R....NK....T....S.E....E....E.M....V....R....E....D.G.L....T....V....S.I....R....E....I....</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Python Y3 :</td>
<td>.N....R....O.G....T....PK....I....E.M....V....K....T....H....T.A....G.P....T.R....T....C.V....S.I....R....E....I....</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snake Y2 :</td>
<td>.N.V....T.RA....D....S....S....N....N....Q....T....NEHLN.DPIERTD.P....Y....V....G.YS....G.V....V.H....SFTV....V....S.T.G....I....KVF.P.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Python Y2 :</td>
<td>.N.V....T.RA....T.D....S.LS....G....ENDQ....FT....G....GNN.N.D....AET....Q....Y....I....S....FHN....D.G.I....A....AV....T.I.S....LAG....S....KVSIF.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lizard Y :</td>
<td>.V....K.S....T.N....L....D.P....V.S....G....E....V....D....KASLD.KFPS....S....L....Y....V....A....ILTRQGD.E....T....V....V.K....LKG....VTT....SVYLFK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gecko Y :</td>
<td>.KVSIT....T.A.R....MGENFRLEIG....TS....ENY....M....V.R.K....P....GTLN....P....EPAQDEBD....L.K....I.K....TLAG....Y....T.L....HSSGIA....Y.S....A....KKG....PTAQQVYLYR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turtle Y :</td>
<td>.LVFTV.S.EE....VQOSP....VTL....PNDAD.RNV....TGLPS....P....QH....K....S....I....ITR.E....DG....KYS....HSE....T....I....S....KQG....MSP....SVYLR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duck Y :</td>
<td>.QIFVV....S.GS....QBD....VHE....PSDAS....S....S.G....R....DPMVL....HF....S....S....A....T....TLAG....R....T.D....QHED....E....S....A....HAG....V.T....Y....PTF.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chicken Y :</td>
<td>.QL.AI....S.GE....LD....LAA....FSDS....S.VT....S....GNNLR....DPMVLQ....HF....S.Y....S....A....V....T....G....R....T....QHEE....T....S....L....S....VYRNTGPTTP....Y...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**LM1S**

- **SpA, SpG**

**SpG**

- **C1q**
Supplemental Figure 3.
**Supplemental Figure 4.**

**a.**

60kd : ASPTAPTVFPLSNCQDITTLSSVIGCLVKGYFPQFAPTVQWNSGAITSIGHNFPPVMLGNSHYTHSSQLTIWVSKWQSESFECNVNAATSTR1KRRIDRPEVRSAP70kd : ASPTAPTVFPLSNCQDITTLSSVIGCLVKGYFPQFAPTVQWNSGAITSIGHNFPPVMLGNSHYTHSSQLTIWVSKWQSESFECNVNAATSTR1KRRIDRPEVRSAP

------- 10 ------- 20 ------- 30 ------- 40 ------- 50 ------- 60 ------- 70 ------- 80 ------- 90 ------- 100 ------- 1

60kd : GPEVHFHSSCGSRDGTIQLWCQISGFYPKPLEIQWKVGSRLLRPRYNYPWQNPQGTYFSTASVQTVNQDMLWGEVNYCEVTQHAASQTMKSKAKKCEGGSDCP570kd : GPEVHFHSSCGSRDGTIQLWCQISGFYPKPLEIQWKVGSRLLRPRYNYPWQNPQGTYFSTASVQTVNQDMLWGEVNYCEVTQHAASQTMKSKAKKCEGGSDCP

------- 10 ------- 120 ------- 130 ------- 140 ------- 150 ------- 160 ------- 170 ------- 180 ------- 190 ------- 200 ------- 210 ------- 220

60kd : GDINVYILPPTFRAALYIDRNKISCVNNLQNEQGLKITWSREKNEHLNPDPIETDEPNGTYVESRLGVYSQDWGEVFTCVCVEHPSFVTPTKITSKTRGKSIA70kd : GDINVYILPPTFRAALYIDRNKISCVNNLQNEQGLKITWSREKNEHLNPDPIETDEPNGTYVESRLGVYSQDWGEVFTCVCVEHPSFVTPTKITSKTRGKSIA

------- 220 ------- 230 ------- 240 ------- 250 ------- 260 ------- 270 ------- 280 ------- 290 ------- 300 ------- 310 ------- 320 ------- 330

60kd : PVVFVFPPHRDELNPETTVSLCTCMITGFYPKEADVQWLSNYPLPENYVSTSMKSSHDSFFRYSKLNIPRADWEREASFTCVMVHEALPMKFTQRNISKTQGKK70kd : PVVFVFPPHRDELNPETTVSLCTCMITGFYPKEADVQWLSNYPLPENYVSTSMKSSHDSFFRYSKLNIPRADWEREASFTCVMVHEALPMKFTQRNISKTQGKK

**b.**

lambda : GQPPVSPSVDSTQFPEAQEIKTKKATLVCLLSGFHPPAFLK5KVDG<5ETTSGVETTKASKEDKYLASSYLT<DSNYGDHTYICEV<NGKTTIKALTGSGS