

# Immunoglobulin genes of the turtles

Susana Magadán-Mompó · Christian Sánchez-Espinel ·  
Francisco Gambón-Deza

Received: 17 September 2012 / Accepted: 23 November 2012 / Published online: 4 December 2012  
© Springer-Verlag Berlin Heidelberg 2012

**Abstract** The availability of reptile genomes for the use of the scientific community is an exceptional opportunity to study the evolution of immunoglobulin genes. The genome of *Chrysemys picta bellii* and *Pelodiscus sinensis* is the first one that has been reported for turtles. The scanning for immunoglobulin genes resulted in the presence of a complex locus for the immunoglobulin heavy chain (IGH). This IGH locus in both turtles contains genes for 13 isotypes in *C. picta bellii* and 17 in *P. sinensis*. These correspond with one immunoglobulin M, one immunoglobulin D, several immunoglobulins Y (six in *C. picta bellii* and eight in *P. sinensis*), and several immunoglobulins that are similar to immunoglobulin D2 (five in *C. picta bellii* and seven in *P. sinensis*) that was previously described in *Eublepharis macularius*. It is worthy to note that IGHD2 are placed in an

inverted transcriptional orientation and present sequences for two immunoglobulin domains that are similar to bird IgA domains. Furthermore, its phylogenetic analysis allows us to consider about the presence of *IGHA* gene in a primitive reptile, so we would be dealing with the memory of the gene that originated from the bird *IGHA*. In summary, we provide a clear picture of the immunoglobulins present in a turtle, whose analysis supports the idea that turtles emerged from the evolutionary line from the differentiation of birds and the presence of the *IGHA* gene present in a common ancestor.

**Keywords** Immunoglobulin genes · Reptile immunity · Immunoglobulin phylogeny · Evolution

**Electronic supplementary material** The online version of this article (doi:10.1007/s00251-012-0672-7) contains supplementary material, which is available to authorized users.

F. Gambón-Deza (✉)  
Servicio Gallego de Salud (SERGAS) Unidad de Inmunología,  
Hospital do Meixoeiro, Carretera de Madrid s/n,  
Vigo 36210, Pontevedra, Spain  
e-mail: fgambon@gmail.com

C. Sánchez-Espinel  
Shared Unit of Immunology, Vigo University Hospital Complex  
(Hospital Meixoeiro), University of Vigo,  
Edificio de Ciencias Experimentales, Rúa das Abeleiras,  
Campus As Lagoas-Marcosende,  
Vigo 36310, Pontevedra, Spain  
e-mail: cristianespinel@uvigo.es

S. Magadán-Mompó  
Oceanographic Center of Vigo,  
Spanish Institute of Oceanography (IEO), Subida a Radio Faro 50,  
36390 Vigo, Pontevedra, Spain  
e-mail: susana.magadan@gmail.com

## Introduction

Reptiles represent the evolutionary lineage of vertebrates that have adapted to live on land. There are four living orders currently identified in reptiles: Crocodylia that includes crocodiles, gavials, caimans, and alligators; Sphenodontia that only includes the tuatara from New Zealand; Squamata, where we find lizards, snakes, and worm lizards; and the Testudines which includes turtles, terrapins, and tortoises (Janes et al. 2010; Shedlock et al. 2007; Modesto and Anderson 2004).

Immunoglobulins are proteins that are important for defending the body against infections. Many studies have elucidated the immunoglobulin structure and genomic organization in mammals, amphibians, and fish (Paul 2008), yet little is known in reptiles. All reptiles studied, until now, have at least three classes of immunoglobulins: immunoglobulin M (IgM), also found in all vertebrates; immunoglobulin D (IgD) with 11 domains in the constant region (IgD in mammals only have two or three domains); and immunoglobulin Y (IgY) that is already present in amphibians and birds, from which

immunoglobulin G and immunoglobulin E in mammals are derived (Wei et al. 2009; Gambón Deza et al. 2009, 2012).

Recent studies performed in reptilian species of the Squamata order indicate that the lizard *Anolis carolinensis* only possesses three Ig isotypes (Wei et al. 2009; Gambón Deza et al. 2009). Nevertheless, other immunoglobulin types in the gecko *Eublepharis macularius* were described that may have been produced by recombination processes of the IGHM, IGHD, and/or *IGHY* genes. One of these immunoglobulins is similar to immunoglobulin A (named IgA-like) and was generated during evolution by a recombination between IgM and IgY (Deza et al. 2007). The other, denominated IgD2, is a variant of IgD and seems to have been generated by an exon shuffling between IgD and the IgA-like immunoglobulin (Gambón-Deza and Espinel 2008). This IgD2 antibody is highly expressed and may have an evolutionary significance since there is a pseudogene with a similar structure in the genome of the monotreme *Ornithorhynchus anatinus* (Gambón-Deza et al. 2009). Finally, in snakes, which represent a highly evolved group of reptiles, IgM, IgD, and new forms of IgY, called IgYb and IgYc, have been described—the latter with only three domains in its immunoglobulin constant region (Gambón-Deza et al. 2012).

Reptiles are nonmammalian, nonavian amniotes. Birds have evolved from these species (Erickson et al. 2009; Steiger et al. 2009), and their immunoglobulin genes have been studied in ducks and chickens. IgM and IgY (the first animals to be described) were found in both cases. Surprisingly, while no IgD was found, an immunoglobulin gene with opposite orientation and sequence homology to that of IgA of mammals was encountered (Magor et al. 1994; Lundqvist et al. 2001; Zhao et al. 2000).

Studies of immunoglobulins in turtles, or even in the reptilia Testudines order, are scarce. Only a messenger RNA has been described for IgM in red-eared slide turtle (*Trachemys scripta elegans*) (Turchin and Hsu 1996), and sequences of messenger RNA for IgM, IgD, and IgY in the Chinese softshell turtle (*Pelodiscus sinensis*) have recently been published (Xu et al. 2009). Recently, however, the International Painted Turtle Genome Sequencing Consortium has published and made available the genome of painted turtle (*Chrysemys picta bellii*), and the RIKEN institute in Japan has made available the Chinese softshell turtle (*P. sinensis*) genome. This is a great opportunity to provide a clear picture of the immunoglobulins in turtles and to understand their preservation or modification during the reptilian evolutionary history, which is paralleled to that of mammals. In addition, present phylogenetic studies are not yet clear about Testudines (turtles, terrapins, and tortoises) evolutionary location (Lukoschek et al. 2012). The painted turtle immunoglobulin sequences have placed the turtles in the same evolutionary lineage that gave rise to crocodiles and

birds. Therefore, their study can provide answers to the origin of immunoglobulins that are currently found in birds and mammals.

## Materials and methods

### Identification of the immunoglobulin heavy chain loci

The complete genome of *C. picta bellii* build 3.0.1 (AHGY00000000.1) and *P. sinensis* PeISin\_1.0 (AGCU00000000.1) deposited in NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) was examined to locate antibody genes. Published immunoglobulin sequences of *A. carolinensis* and *E. macularius* were used to search for painted turtle immunoglobulin sequences. We studied *A. carolinensis* and *E. macularius* sequences and *C. picta bellii* and *P. sinensis* genomes from FASTA files using the Galaxy website (<http://main.g2.bx.psu.edu/>) (Giardine et al. 2005; Blankenberg et al. 2010; Goecks et al. 2010). Different scaffolds and contigs that contain immunoglobulin genes were identified: scaffold group 175 (JH584564.1) and scaffold 2278 (JH586043.1) for immunoglobulin heavy chain (IGH) genes in *C. picta bellii* and scaffolds JH210939 and JH205552 in *P. sinensis*. All sequences were retrieved and analyzed in detail using Vector NTI (Invitrogen).

Identification of the exons coding for the constant heavy (CH) domains was performed using the software FGENESH ([www.softberry.com](http://www.softberry.com)) (Solovyev et al. 2006; Yao et al. 2005) and Augustus (<http://augustus.gobics.de/submission>) (Stanke et al. 2004). Protein products, predicted from previous searches, were compared with reptile sequences and found to be evolutionarily close because there are no RNA studies to date for these genes.

### Phylogenetic studies

Comparative phylogenetic studies were carried out using MEGA5 (Tamura et al. 2011) and sequence alignments using SeaView (Gouy et al. 2010) with the ClustalW and MUSCLE algorithms. Subsequently, the neighbor-joining and maximum likelihood methods were used to produce phylogenetic trees (pairwise deletion, and JTT or WGA matrix). The veracity of our results was tested with 500 replicate bootstrapping execution runs.

### GenBank sequences used

The following IgM accession numbers were used: AAO37747 *O. anatinus*, EU287910 and EU287911 *E. macularius*, FJ605150 *Pelodiscus sinensis*, CAA30613.1 *Heterodontus francisci*, AAB03838 *Trachemys scripta elegans*, ABV66128 *A. carolinensis*, PO1875 *Gallus gallus*,

CAC43061 *Anas platyrhynchos*, MGC69066 *X. laevis*, AAH89670 *Xenopus tropicalis*, A46532 *Axolotl mexicanum*, CAE02685 *Pleurodeles waltl*, AAC48834 *Monodelphis domestica*, CAB37838 *Homo sapiens*.

The following IgD accession numbers were used: EU312156, EU327165, and EU327166 *E. macularius* and ABV66130 *A. carolinensis*. The following IgY accession numbers were used: EF690361 *A. carolinensis*, EU827594 and EU827595 *E. macularius*, CAA46322 *A. platyrhynchos*, S00390 *G. gallus*, and AJ575800 *P. waltl*.

The kappa light chain accession numbers that were used are XP\_003222451.1 *A. carolinensis*, AAA49880.1 *X. laevis*, AAI58339.1 *X. tropicalis*, ABV02006.1 *Macropus eugenii*, AAO84652.1 *Tachyglossus aculeatus*, CAA53284.1 *Equus caballus*, and P01834.1 *H. sapiens*.

## Results

With the recent publication of the *C. picta bellii* and *P. sinensis* genomes, made available by the International Painted Turtle Genome Sequencing Consortium and RIKEN Institute, we searched the immunoglobulin sequences of *A. carolinensis* and *E. macularius* that we had identified previously (Deza et al. 2007; Gambón-Deza and Espinel 2008; Gambón Deza et al. 2009). The turtle genes for the immunoglobulin heavy chain constant regions were found as described in “Materials and methods.” With these data, we were able to perform a detailed study of the immunoglobulins in the evolutionary line that emerged from the first reptiles.

In the IGHC locus, there are 13 sequences coding for different immunoglobulin isotypes in *C. picta belli* and 17 in *P. sinensis* (Fig. 1). The IGHM and IGHD loci are first placed and correspond to the IgM and the reptilian IgD (with 11 exons for CH domains), respectively. Downstream of IgD, there is a region, which includes an *IGHD2* gene and an *IGHY* gene, and that is repeated five times in *C. picta bellii* and seven times in *P. sinensis*. While IGHYs are in the same direction to the IGHM-IGHD, IGHD2 are in the opposite directions. Finally, another sequence coding for IgY was located in each genome without the IGHD2 mate.

### Immunoglobulin M

The locus IGHM contains four exons for CH domains plus TM1 and TM2 exons for transmembrane and cytoplasmic regions, respectively (Fig. 2). Two IgMs have been previously described in turtles, namely, the IgM of *T. scripta elegans* (Turchin and Hsu 1996) and the IgM of *P. sinensis* (Xu et al. 2009). The deduced amino acid sequence in *C. picta bellii* is similar (Fig. 3 and Additional File 1 of Electronic Supplementary Material (ESM)). Unlike those

described for the anole, gecko, and snakes (Wei et al. 2009; Gambón-Deza et al. 2009, 2012), turtles possess a cysteine at the beginning of the CH1 domain, which is required in order to join the light chain. In the CH2 domain, there are two noncanonical cysteines, where one is present in turtles studied (position 118) while the other (position 153) is also present in snakes but absent in the lizard *A. carolinensis* and gecko *E. macularius*. At the beginning of the CH3 domain (position 239), there is a cysteine that is specific for reptiles and is found in lizards, snakes, and turtles. Moreover, the cysteine at position 319 is maintained in most vertebrates. Finally, a cysteine is also present in the secretor tail (position 486), suggesting the possibility of a J-chain bonding to form IgM multimers.

A phylogenetic tree was created with IgM sequences from turtles and other vertebrates (Fig. 3) that supports the claim that turtles shared a common ancestor with birds.

### Immunoglobulin D

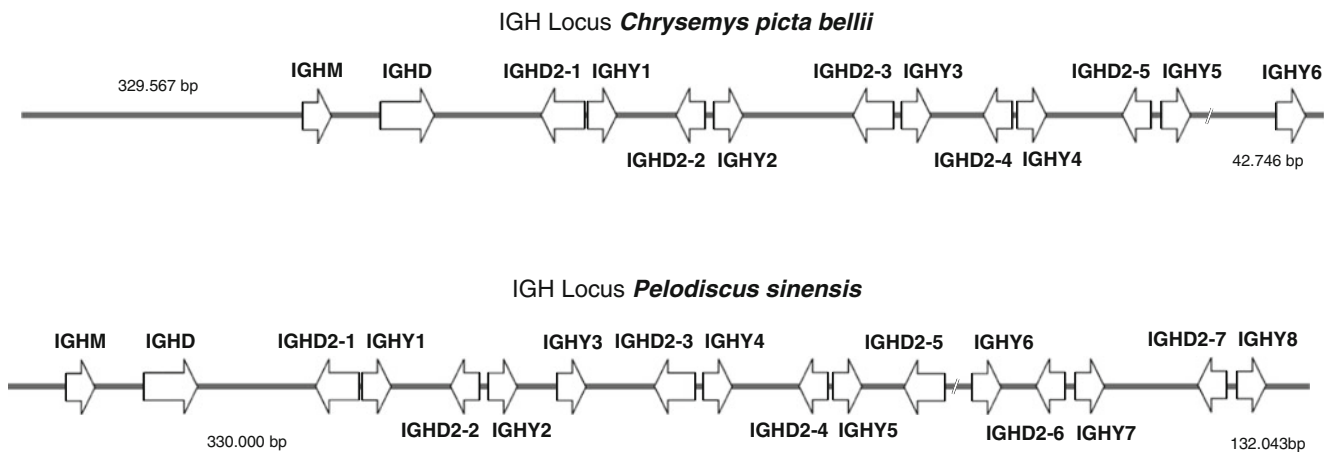
The 11 exons for the IgD immunoglobulin domains are located behind the IGHM (Fig. 1). This confirms the generation of a three-immunoglobulin core (C $\mu$ -C $\delta$ -C $\nu$ ) in the ancestor of all reptiles and its preservation up until the present. The turtle IgD has features that are similar to those described in other reptiles and the platypus IgD. Most non-canonical cysteines and the high glycosylation rate in domains CH7 and CH10 are maintained (Fig. 4 and Additional File 2 of ESM). In *P. sinensis*, the exon that codes for the CH10 domain may not be viable due to the absence in the sequence of the second canonical cysteine and presence of stop codon. The existence of several types of differential splicing in IgD indicates that messenger RNA which is created without the exon could synthesize a functional protein.

The phylogenetic tree supports the domain-to-domain orthology with lizard, gecko, snakes, and platypus. Each group of orthologous domains comes from a corresponding exon. The exonic numbering of this tree will be maintained in subsequent studies in the article (Fig. 4).

### Immunoglobulin Y

We found six IgY gene isotypes in *C. picta belli* and eight in *P. sinensis* (Fig. 1). In *C. picta belli*, five isotypes were found in the contig group 175, while the gene for IgY6 was located in contig 2278. All these genes have four C $\nu$  exons, except in IGHY-5 that has the C $\nu$ 2 duplicated, generating an IgY with five domains (Fig. 2). The second CH2 domain has a sequence that is identical to the CH2 domain of IgY1 (data not shown).

The gene classes for the immunoglobulin Y in *P. sinensis* have a more complex structure. The genes *IGHY1*, *IGHY4*,



**Fig. 1** Genomic organization of the IGH locus in turtles. The organization of the immunoglobulin genes identified in *C. picta bellii* and *P. sinensis* genomes is shown. The transcriptional orientation of genes is

indicated by the direction of the arrows. In both turtles, the genes were found in two scaffolds whose lengths in base pairs are annotated

*IGHY7*, and *IGHY8* have the four epsilon exons for the four immunoglobulin domains. Both the *IGHY2* and *IGHY3* genes have epsilon exons duplicated. In the case of *IGHY2*, the C $\nu$ 3 is duplicated since the two exons are nonviable (due to the presence of a stop codon) and also lack an C $\nu$ 2 exon, suggesting that the gene is a pseudogene. The *IGHY3* gene also lacks the C $\nu$ 2 exon and has duplications of the C $\nu$ 3 and C $\nu$ 4 exons, thereby generating an immunoglobulin with five domains in the constant region. Of great interest is the structure of the *IGHY6* gene, which has three epsilon domains with a structure very similar to the mammalian IGHG (with a missing C $\nu$ 2 exon).

The amino acid sequence alignments of the turtle IgYs with those found in other vertebrates indicate that turtle IgYs do not contain the cysteine that would establish the second intrachain bridge between CH1 and CH2. The non-canonical cysteines for interchain bridging are located in the CH2 domain (Fig. 5 and Additional File 3 of ESM).

All the IgY isotypes are very similar to each other, indicating a recent duplication processes in their genesis (Fig. 5 and Additional File 3 of ESM). As described in the case of the IgM sequences analysis, the phylogenetic tree constructed with IgY indicates that turtles shared a common ancestor with birds (Fig. 5). The high sequence identity between the IgYs within each species and the tree phylogeny can be explained by the diversification of the IgY isotypes that took place in the evolutionary lineage of turtles, or perhaps through a process of concerted gene evolution.

## Immunoglobulin D2

In a previous study, we described the immunoglobulin IgD2 in *E. macularius* (Gambón-Deza and Espinel 2008). This immunoglobulin is a variant of IgD, with the first four CH domains similar to IgD and the last two CH domains similar

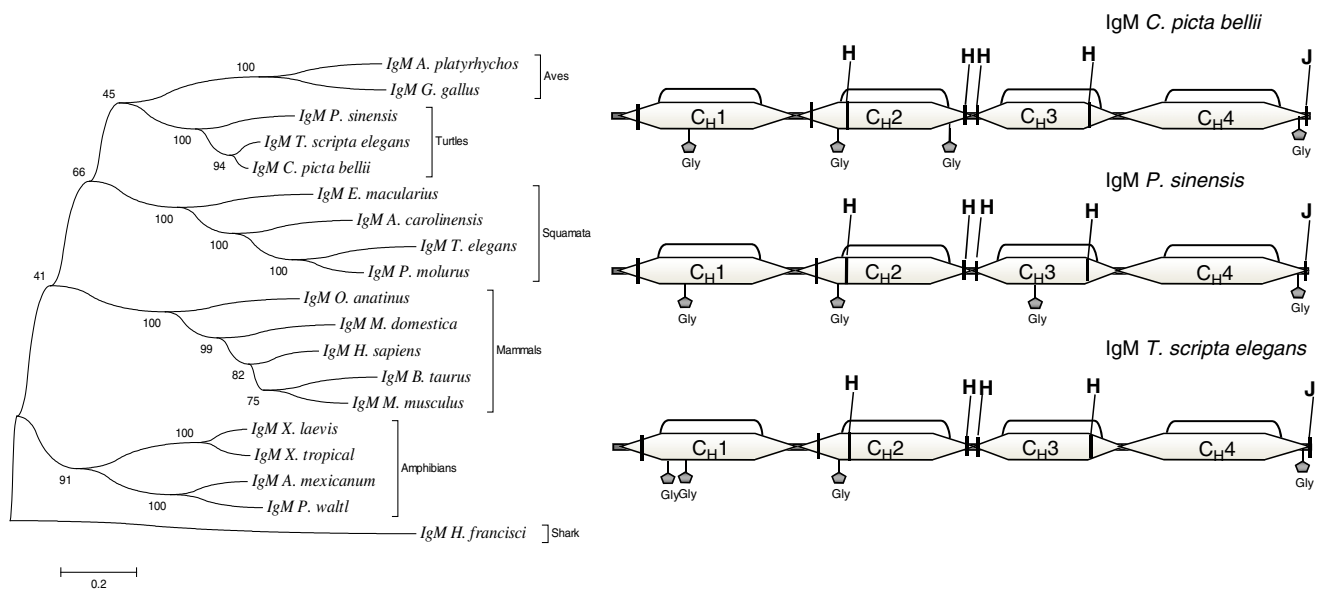
to the CH3 and CH4 domains of IgA-like immunoglobulin (Fig. 6). One year later, an immunoglobulin with the same characteristics was described in the turtle *P. sinensis* (Xu et al. 2009). As shown in Fig. 6, this immunoglobulin corresponds structurally to the IgD2 of *E. macularius*, having the same number and order of CH domains.

Several genes encoding for this kind of immunoglobulin are found in the genome of both turtles. As shown in Fig. 1, we found five in *C. picta bellii* and seven in *P. sinensis* that are oriented in an opposite direction to *IGHM-IGHD* and *IGHY* genes. All of these genes contain C $\delta$  and C $\alpha$  exons. The assignment of domain type that codes for each exon was made from the phylogenetic tree that emerges from the alignment of all domains (Fig. 6 and Additional File 4 of ESM).

The analysis of their sequences indicates that in *C. picta bellii*, only three of the five genes can be functional and two may be pseudogenes (*IGHD2-3* has an altered sequence in the C $\delta$ 2 and C $\delta$ 3 domains, and *IGHD2-5* contains a stop codon in the C $\alpha$ 3 domain). *IGHD2-1* has eight exons (C $\delta$ 1-C $\delta$ 2-C $\delta$ 3-C $\delta$ 2-C $\delta$ 3-C $\delta$ 4-C $\alpha$ 3-C $\alpha$ 4), and the transmembrane exons have not been detected, so it should only be expressed as a secreted form. The other functional immunoglobulin genes, *IGHD2-2* and *IGHD2-4*, have four (C $\delta$ 3-C $\delta$ 4-C $\alpha$ 3-C $\alpha$ 4) and five (C $\delta$ 1-C $\delta$ 2-C $\delta$ 4-C $\alpha$ 3-C $\alpha$ 4) immunoglobulin heavy chain exons and can be expressed as transmembrane forms due to the presence of transmembrane exons (TM1 and TM2) (Figs. 2 and 6).

Of the seven sequences in *P. sinensis*, two could be pseudogenes, only two exons are detected in gene *IGHD2-5*, and, in the gene *IGHD2-2*, a stop codon in the C $\delta$ 2 exon is found. The remaining five *IGHD2* genes have viable structures. Four of them have the same number of exons with the structure C $\delta$ 1-C $\delta$ 2-C $\delta$ 3-C $\delta$ 4-C $\alpha$ 3-C $\alpha$ 4 while *IGHD2-7* has the structure C $\delta$ 1-C $\delta$ 2-C $\delta$ 3-C $\delta$ 4-C $\delta$ 1-C $\delta$ 2-C $\delta$ 3-C $\alpha$ 3-C $\alpha$ 4.





**Fig. 3** Analysis of IgM. On the *left*, a phylogenetic tree was constructed using IgM amino acid sequences obtained from different reptiles, amphibians, birds, mammals, and shark. On the *right*, a schematic representation of the different IgMs identified in the turtles *C. picta bellii*, *P. sinensis*, and *T. scripta elegans* is displayed. The

domains are represented as *hexagons* with the intracatenary bonds. The putative cysteines that establish bonds with heavy chains (*H*) and those that form multimers (*J*) are shown. Potential glycosylation sites are indicated with *pentagons*

domains arose by a recombination between IGHD and IGHD2-1. The IGHD provided the C $\delta$  exons, and the IGHD2-1, the C $\alpha$  exons.

#### Relationship of the IgD2 with bird IgA

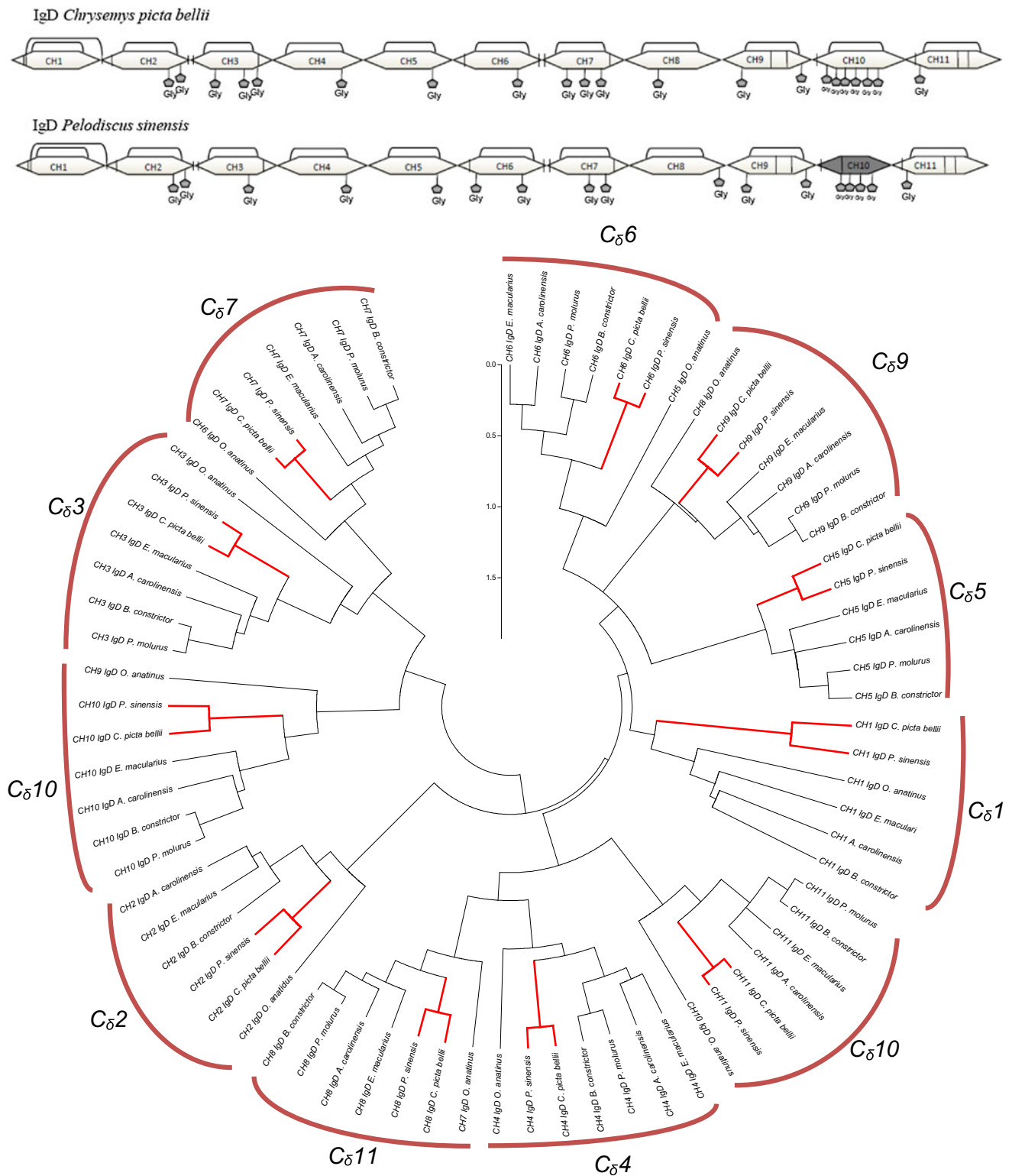
It is worth noting that the IGHD2 location (between the *IGHM-IGHD* and *IGHY* genes) and the direction (laying in the opposite direction to the *IGHM-IGHD* and the *IGHY* genes) are very similar to that found in avian IGHA, suggesting an evolutionary relationship. To determine whether the IGHA found in birds and IGHD2 evolved from a common ancestor, we aligned the domain amino acid sequences obtained from *A. platyrhynchos* and *G. gallus* IgA, *P. sinensis* IgD2 domains, and *C. picta bellii* IgD2-1 and IgD domains. As shown in Fig. 7, the resulting phylogenetic tree constructed from these sequences suggests a clear evolutionary relationship between the last two domains of IgD2 and the CH3 and CH4 domains from the bird IgA. This analysis also indicates that CH1 and CH2 of the IgD2 arise from the duplication of C $\delta$  exons. This duplication happened after the bird C $\alpha$ 1 and C $\alpha$ 2 divergence, suggesting that these IgA domains do not derive from the *IGHD2* gene.

Therefore, we can conclude that the two terminal domains of the turtle IgD2 and the two terminal domains of the bird IgA are orthologues; however, any orthologue exon to the bird C $\alpha$ 1 and C $\alpha$ 2 has been detected in the turtles' IGH locus.

#### Discussion

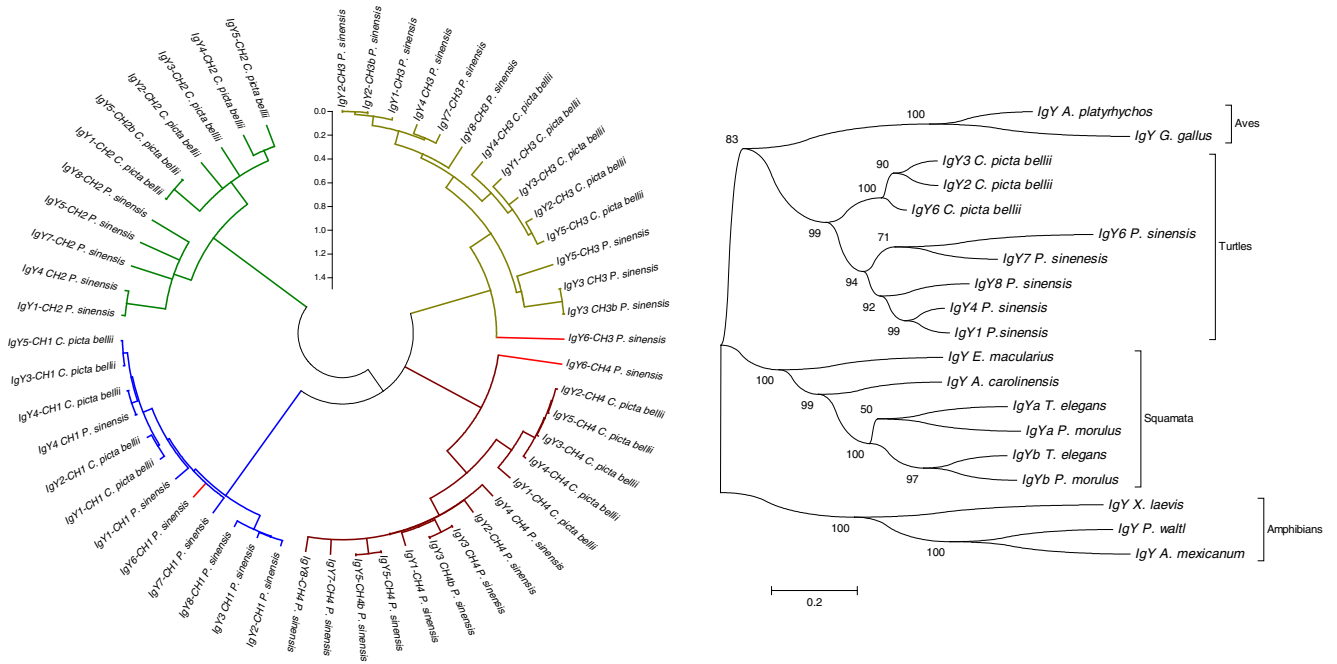
Reptile evolution may be divided into two lineages: one line gave rise to lizards and snakes (Squamata) while the other to turtles, crocodiles, and birds. With respect to immunoglobulins, there are significant studies in birds (Lundqvist et al. 2006; 2001; Higgins and Warr 1993), and recently, these genes have also been described in several Squamata (Wei et al. 2009; Gambón-Deza et al. 2012; 2009; Gambón-Deza and Espinel 2008). From this information, we can affirm that the common ancestor of Squamata and birds must have had *IGHM* and *IGHY* genes. In this work, we describe the immunoglobulin genes in turtles, providing new insights into the origin of the *IGHD2* and *IGHA* genes.

The scanning of the painted turtle genomic sequences revealed the presence of *IGHM*, *IGHD*, and *IGHY* genes, which are similar to those previously found in Squamata (lizards and snakes). So, we may deduce that these immunoglobulins were present in early reptiles 300 million years ago. These three immunoglobulins were already present in amphibians and, although the structure of IgM and IgY has remained similar to those found today, the amphibian IgD has been modified through evolution. The IGHD sequence itself described in the frog *X. tropicalis* (Qin et al. 2008; Zhao et al. 2006), coding for an IgD with eight CH domains, proves the presence of duplications. On the other hand, the *IGHD* gene described in reptiles gives rise to an immunoglobulin with 11 CH domains that has survived for at least 300 million years without modifying its basic structure. It suggests essential



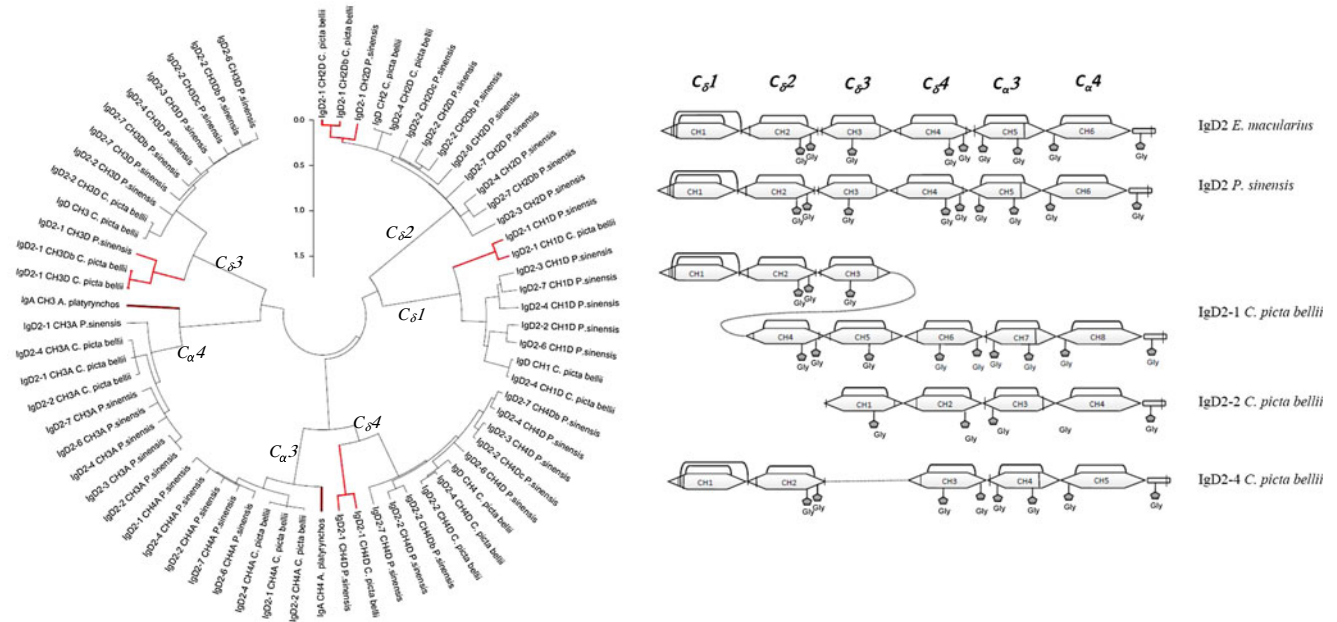
**Fig. 4** Shown at the top of the figure is a schematic representation of the IgD identified in the turtles *C. picta bellii* and *P. sinensis*. The domains are represented as hexagons with the intracatenary bonds. The putative cysteines that establish bonds with heavy chains (*H*) are shown. Potential glycosylation sites are indicated with pentagons. The CH10 *P. sinensis* is shaded to emphasize important alterations in the sequence due to the absence of the second canonical cysteine. In

the lower part of the figure, a phylogenetic tree constructed using the amino acid IgD domain sequences identified in turtles, other reptiles (snakes, lizard, and gecko), and in the platypus is shown. It indicates the IgD domain-to-domain orthology between these species. The tree was constructed using the evolution algorithm, JTT matrix, and differences by sites activated with gamma parameter 2.5. The GenBank accession numbers are indicated in “Materials and methods”



**Fig. 5** Analysis of IgYs. The *left figure* shows the phylogenetic tree constructed from amino acid sequences coding for all Cv exons. The tree identifies the relationship between the exon sequences and assigns the correct numbering. The *right figure* shows the phylogenetic tree

constructed using the IgY amino acid sequences obtained from different reptiles, amphibians, and birds. The tree branches represent all turtle IgY sequences that emerged from the evolutionary line of birds

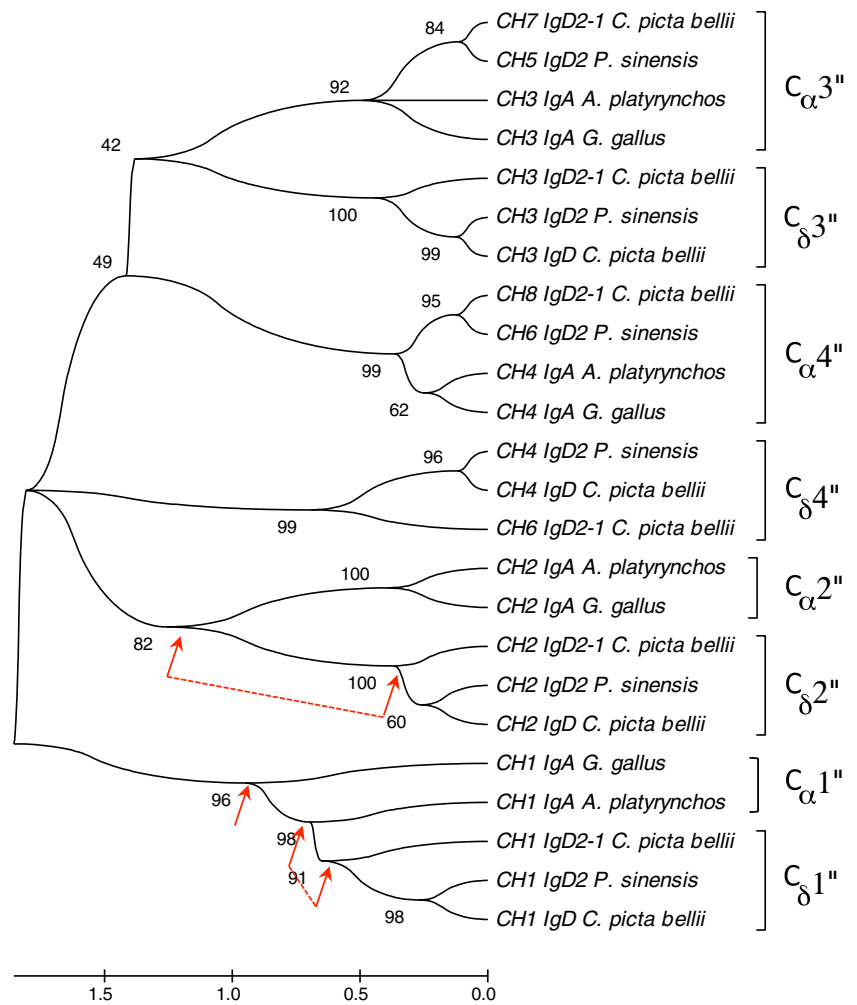


**Fig. 6** The *left of the figure* represents the phylogenetic tree which is used to study the relationship between IgD2 and IgD. The tree was constructed after the alignment of the amino acid immunoglobulin domain sequences of IgD from *C. picta bellii* and IgD2 from *C. picta bellii* and *P. sinensis*. The exons coding for each domain group are also indicated. The maximum likelihood, WGA matrix, was used for obtaining this tree. The IgD domains are marked in *red* while the IgD2-1 domains are drawn in *brown* to demonstrate an earlier origin

to other IgD2 domains. To the *right of the figure*, the structure of IgD2 is identified in turtles (*P. sinensis* and *C. picta bellii*). The domains are displayed as *hexagons* with the intracatenary bonds. The noncanonical cysteines are drawn with a *line* and the *points* of N-glycosylation are represented with *pentagons*. The exons that code each domain are indicated at the *top of the figure*. The graphics of IgD2-1 and IgD2-4 from *C. picta bellii* have been adapted so that all domains are placed with the corresponding exon



**Fig. 7** The evolutionary relationship between IgD2 and bird IgA. The amino acid sequences obtained from the *A. platyrhynchos* and *G. gallus* IgA domains, *P. sinensis* IgD2 domains, and *C. picta bellii* IgD2-1 and IgD domains were aligned, and then, an unrooted phylogenetic tree was constructed. The constant heavy chain domains (CH) are grouped according to the exon which originated from them. Red arrows indicate the origin of the CH1 and CH2 domains of the IgD2-1. They arose from the duplication of IGHD. For phylogenetic tree, the algorithm used was maximum likelihood, WGA matrix, and bootstrapping of 200 replicates



immune functions that remain to be discovered. The obtained immunoglobulin sequences support the conclusions provided by others about the origin of turtles as a sister group to birds and crocodylians, contradicting the long-held view that they were basal amniotes (Chiari et al. 2012).

The *IGHY* genes found in turtle indicate the presence of a process of duplication of these genes in the evolutionary line. Probably, once they have made the duplications, a process of concerted evolution can explain the phylogenetic tree that emerges from the study (Fig. 5). The turtle ancestor must have already had several *IGHY* and nonhomologous recombination processes between different genes should have given the homogenization of the sequences and the presence of genes with duplicate epsilon domains. Considering that turtles diverged from other reptiles in a recent time to the divergence of mammals, this process of duplication and concerted evolution may be a remnant of the process that must have occurred to generate the *IGHE* and *IGHG* genes of mammals. The presence of the *IGHY-6* gene with three domains is similar to the *IGHG* mammalian gene.

We recently described a new isotype in *E. macularius* (Gambón-Deza and Espinel 2008), the IgD2, which is

defined as an immunoglobulin that has arisen by duplication of the IgD and the IgA domains. In turtles, immunoglobulins with these characteristics also exist, as well as in different evolutionary lines as compared with Squamata. Different IgD2 with a variable number of domains exist in turtles; however, all have their first domains from the IgD and two terminal domains that are similar to the CH3 and CH4 of IgA. Between *E. macularius* and the turtles' (*P. sinensis* and *C. picta bellii*) IgD2, there is no full-domain orthology. All have the same structural characteristics yet do not share a common ancestor, since the evolutionary relationship between the IgD and the IgD2 from the same species demonstrates that these IgD2 were created independently in their own evolutionary line.

The results obtained regarding the IgD2 in turtles indicate that the original gene was *IGHD2-1*. The other *IGHD2* genes have recently been generated by combination between *IGHD* gene (providing the  $C_{\delta}$  exons) and *IGHD2-1* (providing the  $C_{\alpha}$  exons). This may explain why some of these new *IGHD2* are pseudogenes. Similar to the *IGHY* genes, a concerted evolution process must have occurred in order to explain the identity between the sequences.

The presence of IgA domains in the turtle IGH locus and the current avian IgA indicate that the first reptiles must have had other immunoglobulin apart from IgM, IgD, and IgY. Our results suggest that it could be the IgA that probably arose from the evolution of the amphibian immunoglobulin X (Zhao et al. 2006). This IgA was most likely lost after recombining (in an inverted transcriptional direction) the *IGHA* and the *IGHD* genes to generate the turtle *IGHD2-1* gene. Within the evolutionary line from which birds originated, other changes took place that also prompted the loss of *IGHD*. Further genomic studies in other reptiles, belonging to the bird evolutionary line, are needed to verify this explanation of the data. As a result, we can firmly establish that the remnant IgA in reptiles corresponds to the IgA found in birds. The analysis of the American alligator (*Alligator mississippiensis*) and saltwater crocodile (*Crocodylus porosus*) genomes, recently published by the International Crocodylian Genomes Working Group ([www.crocgenomes.org](http://www.crocgenomes.org)), shows the presence of *IGHA* (summit to publication) and supports this hypothesis.

All these results, maintenance of a common reptile core (C $\mu$ -C $\delta$ -C $\nu$ ) and emergence of new immunoglobulin isotypes (different IgY and IgD2) or loss (IgA), suggest that immunoglobulin loci are shaped by a complex model of birth-and-death evolution (Nei and Rooney 2005). In this process, new genes are created by duplication, some of them are maintained, and others are deleted or diversified by accumulation of mutations or recombination process. All these processes seem to have acted to create the current turtle immunoglobulin heavy chains loci, and probably, they must have been the cause of the emergence of the mammalian isotypes.

In conclusion, the increasing availability of animal genome information is very useful for understanding the genetic basis of immunoglobulin diversity and the evolutionary divergence of the immunoglobulin loci in vertebrates.

**Acknowledgments** The authors would like to thank Dr. David Olivieri (Escuela Superior de Ingeniería Informática University of Vigo, Spain) for the critical appreciation of the manuscript.

## References

- Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M et al (2010) A web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol* 19(10):1–21
- Chiari Y, Cahais V, Galtier N, Delsuc P (2012) Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). *BMC Biol* 10:65
- Deza FG, Espinel CS, Beneitez JV (2007) A novel IgA-like immunoglobulin in the reptile *Eublepharis macularius*. *Dev Comp Immunol* 31(6):596–605

- Erickson GM, Rauhut OWM, Zhou Z, Turner AH, Inouye BD, Hu D et al (2009) Was dinosaurian physiology inherited by birds? Reconciling slow growth in archaopteryx. *PLoS One* 4(10):e7390
- Gambón Deza F, Sánchez Espinel C, Magadán Mompó S (2009) The immunoglobulin heavy chain locus in the reptile *Anolis carolinensis*. *Mol Immunol* 46(8–9):1679–1687
- Gambón-Deza F, Espinel CS (2008) IgD in the reptile leopard gecko. *Mol Immunol* 45(12):3470–3476
- Gambón-Deza F, Sánchez-Espinel C, Magadán-Mompó S (2009) The immunoglobulin heavy chain locus in the platypus (*Ornithorhynchus anatinus*). *Mol Immunol* 46(13):2515–2523
- Gambón-Deza F, Sánchez-Espinel C, Mirete-Bachiller S, Magadán-Mompó S (2012) Snakes antibodies. *Dev Comp Immunol* 38(1):1–9
- Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P et al (2005) Galaxy: A platform for interactive large-scale genome analysis. *Genome Research* 15:1451–1455
- Goecks J, Nekrutenko A, Taylor J (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11(8):R86
- Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27(2):221–224
- Higgins DA, Warr GW (1993) Duck immunoglobulins: structure, functions and molecular genetics. *Avian Pathol* 22(2):211–236
- Janes DE, Organ CL, Fujita MK, Shedlock AM, Edwards SV (2010) Genome evolution in Reptilia, the sister group of mammals. *Annu Rev Genom Hum Genet* 11:239–264
- Lukoschek V, Scott Keogh J, Avise JC (2012) Evaluating fossil calibrations for dating phylogenies in light of rates of molecular evolution: a comparison of three approaches. *Syst Biol* 61(1):22–43
- Lundqvist ML, Middleton DL, Hazard S, Warr GW (2001) The immunoglobulin heavy chain locus of the duck. Genomic organization and expression of D, J, and C region genes. *J Biol Chem* 276(50):46729–46736
- Lundqvist ML, Middleton DL, Radford C, Warr GW, Magor KE (2006) Immunoglobulins of the non-galliform birds: antibody expression and repertoire in the duck. *Dev Comp Immunol* 30(1–2):93–100
- Magor KE, Higgins DA, Middleton DL, Warr GW (1994) One gene encodes the heavy chains for three different forms of IgY in the duck. *J Immunol* 153(12):5549–5555
- Modesto S, Anderson J (2004) The phylogenetic definition of reptilia. *Syst Biol* 53(5):815–821
- Nei M, Rooney A (2005) Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* 39:121–152
- Paul WE (2008) *Fundamental immunology*. Lippincott, Philadelphia
- Qin T, Ren L, Hu X, Guo Y, Fei J, Zhu Q et al (2008) Genomic organization of the immunoglobulin light chain gene loci in *Xenopus tropicalis*: evolutionary implications. *Dev Comp Immunol* 32(2):156–165
- Shedlock A, Botka C, Zhao S, Shetty J, Zhang T, Liu J et al (2007) Phylogenomics of nonavian reptiles and the structure of the ancestral amniote genome. *Proc Natl Acad Sci U S A* 104(8):2767–2772
- Solovyev V, Kosarev P, Seledsov I, Vorobyev D (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol* 7(Suppl 1):S10.1–S10.12
- Stanke M, Steinkamp R, Waack S, Morgenstern B (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* 32(Web Server issue):W309–W312
- Steiger SS, Kuryshev VY, Stensmyr MC, Kempnaers B, Mueller JC (2009) A comparison of reptilian and avian olfactory receptor gene repertoires: species-specific expansion of group gamma genes in birds. *BMC Genom* 10:446

- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S et al (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739
- Turchin A, Hsu E (1996) The generation of antibody diversity in the turtle. *J Immunol* 156(10):3797–3805
- Wei Z, Wu Q, Ren L, Hu X, Guo Y, Warr GW et al (2009) Expression of IgM, IgD, and IgY in a reptile, *Anolis carolinensis*. *J Immunol* 183(6):3858–3864
- Xu Z, Wang GL, Nie P (2009) IgM, IgD and IgY and their expression pattern in the Chinese soft-shelled turtle *Pelodiscus sinensis*. *Mol Immunol* 46(10):2124–2132
- Yao H, Guo L, Fu Y, Borsuk LA, Wen TJ, Skibbe DS et al (2005) Evaluation of five ab initio gene prediction programs for the discovery of maize genes. *Plant Mol Biol* 57(3):445–460
- Zhao Y, Rabbani H, Shimizu A, Hammarström L (2000) Mapping of the chicken immunoglobulin heavy-chain constant region gene locus reveals an inverted alpha gene upstream of a condensed epsilon gene. *Immunology* 101(3):348–353
- Zhao Y, Pan-Hammarström Q, Yu S, Wertz N, Zhang X, Li N et al (2006) Identification of IgF, a hinge-region-containing Ig class, and IgD in *Xenopus tropicalis*. *Proc Natl Acad Sci U S A* 103(32):12087–12092