

RESEARCH ARTICLE

Three IgH isotypes, IgM, IgA and IgY are expressed in Gentoo penguin and zebra finch

Binyue Han¹✉, Yan Li²✉, Haitang Han¹, Yaofeng Zhao¹, Qingjie Pan^{2*}, Liming Ren^{1*}

1 State Key Laboratory for Agrobiotechnology, China Agricultural University, Beijing, P. R. China, **2** Key Laboratory of Animal Reproduction and Germplasm Enhancement in Universities of Shandong, Qingdao Agricultural University, Qingdao, P. R. China

✉ These authors contributed equally to this work.

* renlm@cau.edu.cn (LR); qjpan@126.com (QP)



OPEN ACCESS

Citation: Han B, Li Y, Han H, Zhao Y, Pan Q, Ren L (2017) Three IgH isotypes, IgM, IgA and IgY are expressed in Gentoo penguin and zebra finch. PLoS ONE 12(4): e0173334. <https://doi.org/10.1371/journal.pone.0173334>

Editor: Nikolas Nikolaidis, California State University Fullerton, UNITED STATES

Received: October 24, 2016

Accepted: February 20, 2017

Published: April 12, 2017

Copyright: © 2017 Han et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The sequence files are available from the NCBI database (<http://www.ncbi.nlm.nih.gov/>), Genbank accession numbers: Nurse shark IgM, I50731; *Xenopus laevis* IgM, AAH84123; *Xenopus tropicalis* IgM, AAH89670; *Pleurodeles waltl* IgM, CAE02685; Axolotl IgM, A46532; Lizard IgM, ABV66128; Gecko IgM, ABY74509; Chinese soft-shell turtle IgM, ACU45376; Red-eared slider IgM, AFR90255; Chinese alligator IgM1, AFZ39166; Chinese alligator IgM2, AFZ39167; Chinese alligator IgM3, AFZ39168; Siamese crocodile IgM1, AFZ39177; Siamese crocodile IgM2, AFZ39178; Siamese

Abstract

Previous studies on a limited number of birds suggested that the IgD-encoding gene was absent in birds. However, one of our recent studies showed that the gene was definitely expressed in the ostrich and emu. Interestingly, we also identified subclass diversification of IgM and IgY in these two birds. To better understand immunoglobulin genes in birds, in this study, we analyzed the immunoglobulin heavy chain genes in the zebra finch (*Taeniopygia guttata*) and Gentoo penguin (*Pygoscelis papua*), belonging respectively to the order Passeriformes, the most successful bird order in terms of species diversity and numbers, and Sphenisciformes, a relatively primitive avian order. Similar to the results obtained in chickens and ducks, only three genes encoding immunoglobulin heavy chain isotypes, IgM, IgA and IgY, were identified in both species. Besides, we detected a transcript encoding a short membrane-bound IgA lacking the last two CH exons in the Gentoo penguin. We did not find any evidence supporting the presence of IgD gene or subclass diversification of IgM/IgY in penguin or zebra finch. The obtained data in our study provide more insights into the immunoglobulin heavy chain genes in birds and may help to better understand the evolution of immunoglobulin genes in tetrapods.

Introduction

As the characteristic molecules of the adaptive immune system, immunoglobulins (Igs) are only expressed in jawed vertebrates [1, 2]. IgH (immunoglobulin heavy chain) isotype is determined by the heavy chain constant genes [3]. To date, several distinct IgH isotypes, including IgM, IgD/W, IgA, IgY, IgE, and IgG, have been identified in multiple vertebrates. IgM and IgD are the first Igs expressed during B cell development, and coexpression of both Igs on the surface of B cells is a mature B cell marker in mammals. Both IgM and IgD are found in most taxa of vertebrates and are believed to be the most primitive IgH isotypes [4]. IgY, expressed in birds, reptiles and amphibians, is considered a progenitor of the functional mammalian IgG and IgE, and is structurally close to IgE [5].

IgM maintained its primordial four CH-domain structure and has been found in nearly all jawed vertebrates, whereas many internal duplications and deletions have occurred in IgD,

crocodile IgM3, AFZ39216; Python IgM, AFR33764; Snake IgM, AFR33841; Ostrich IgM1, AFA41927; Ostrich IgM2, KR030054; Emu IgM1, KU641024; Emu IgM2, KU641025; Chicken IgM, P01875; Duck IgM, CAC43061; Grey goose IgM, AFM77858; Pigeon IgM, EMC81140; Human IgM, AAS01769; Mouse IgM, CAA24199; Pig IgM, AAC48775; Rabbit IgM, AAA64251; Panda IgM, AAX73309; Platypus IgM, AA037747; Opossum IgM, AF135043; Chicken IgA, AAB22614; Duck IgA, AAK17834; Human IgA, P01876; Mouse IgA, AAB59662; Pig IgA, I47175; Cow IgA, AAC98391; Sheep IgA, AAC64980; Rabbit IgA, S09264; Panda IgA, AAX73304; Dog IgA, AAA56796; Platypus IgA1, AAL17700; Platypus IgA2, AAL17701; X. tropicalis IgX, AAI57651; X. laevis IgX, AAI29629; Ostrich IgA, AFA41929; Emu IgA, KU641023; Axolotl IgX, CAO82107; Opossum IgA, AAC48835; Grey goose IgA, AEF79983; Siamese crocodile IgA1, AFZ39174; Siamese crocodile IgA2, AFZ39175; Siamese crocodile IgA3, AFZ39176; Chinese alligator IgA1, AFZ39164; Chinese alligator IgA2, AFZ39165; Chinese alligator IgA3, AFZ39173; Kea IgA, KFQ46330; Gecko IgA, ABG72684; X. laevis IgY, CAA33212; X. tropicalis IgY, AAH89679; *Pleurodeles waltli* IgY, CAE02686; Axolotl IgY, S31436; Lizard IgY, ABV66132; Gecko IgY, ACF60236; Chinese soft-shell turtle IgY, ACU45374; Snake IgY1, AFR33842; Snake IgY2, AFR33843; Python IgY1, AFR33765; Python IgY2, AFR33766; Chinese alligator IgY1, AFZ39169; Chinese alligator IgY2, AFZ39170; Chinese alligator IgY3, AFZ39171; Siamese crocodile IgY1, AFZ39180; Siamese crocodile IgY2, AFZ39181; Siamese crocodile IgY3, AFZ39182; Chicken IgY, CAA30161; Duck IgY, CAD57004; Grey goose IgY, AFN08753; Ostrich IgY1, AFA41930; Ostrich IgY2, KM510516; Emu IgY1, KU641026; Emu IgY2, KU641027; Human IgG, AH007035; Mouse IgG, CAD32497; Platypus IgG1, AY005781; Platypus IgG2, AY005782; Human IgE, AAB59395; Mouse IgE, AAZ05128; Platypus IgE, AY005780.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

which generated a high degree of structural diversity [6–10]. It is also known that IgD has been lost in some jawed vertebrates such as some mammals and birds [11–14]. The discovery of distinct IgH classes in jawed vertebrates has provided significant insight into the evolution of IgH genes, suggesting that both IgM and IgD (which are orthologous to IgW in lungfish and cartilaginous fish) are the most ancestral IgH classes [15]. In addition to the well-known Igs mentioned above, a number of specific IgH classes have been identified in various species, including IgNAR (in cartilaginous fish), IgZ/T (in teleosts), IgO (in platypus), IgP (in *Pleurodeles waltli*), and IgX and IgF (in amphibians) [8, 16–21].

Birds are highly evolved jawed vertebrates, and the expression and genomic organization of Ig genes in birds exhibit some unusual characteristics. The bursa of Fabricius is a critical immune-related organ involved in avian B cell development that is bird-specific [22, 23]. Previous studies have shown that there are only three IgH isotypes in birds (IgM, IgA and IgY) encoded by the $C\mu$, $C\alpha$ and $C\gamma$ genes, respectively, and clear evidence has shown that the IgD-encoding gene is absent in chickens and ducks [11, 12, 24, 25]. However, in a recent study, one δ gene and two additional μ and ν genes were discovered in certain birds such as the ostrich and emu [4]. This result suggests that the common ancestors of modern birds might express multiple IgH isotypes, including IgD, as well as IgM and IgY subclasses.

Using the ostrich IgY sequence as a template, we performed a thorough genomic survey using BLAST searches against the 48 recently sequenced bird genomes. The results implied that the penguin and zebra finch might encode two IgY genes. The Gentoo penguin is a flightless aquatic bird that is widely distributed in the Southern Hemisphere and is a representative Sphenisciforme, which is a remarkable order of birds but the structure and expression of Igs have not yet been studied in detail [26]. The zebra finch, which diverged from the chicken (Galliformes) more than 100 million years ago (MYA) [27, 28], is a member of Passeriformes that possesses almost a half of birds. Therefore, to determine whether IgM and IgY subclass diversification is present in the Sphenisciformes or Passeriformes, we chose the Gentoo penguin and zebra finch as model birds. In this study, we cloned and analyzed Ig heavy chain genes in two different species, the Gentoo penguin (*Pygoscelis papua*) and zebra finch (*Taeniopygia guttata*), trying to elucidate the evolution of Ig genes in birds.

Materials and methods

Animals, RNA and DNA isolation, and reverse transcription

Our study was approved by the Animal Care and Use Committee of the China Agricultural University. Gentoo penguin (*Pygoscelis papua*) blood samples were obtained from Haichang Polar Ocean World in Qingdao, Shandong Province, China (36°N, 120°E). Zebra finches (*Poephila guttata*) were purchased from a local market named Guanyuan in Beijing, China (40°N, 116°E). The animals were handled in accordance with the guidelines of China Agricultural University regarding the protection of animals used for experimental and other scientific purposes. Zebra finches were killed after anesthetized excessively when we got them back to the laboratory immediately. All efforts were made to minimize suffering. Total RNA from whole blood and different tissues was extracted using the TRIzol Kit (Tiangen Biotech, Beijing, China). Reverse transcription was performed using M-MLV (Moloney murine leukemia virus) reverse transcriptase following the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). One microgram of RNA was used for cDNA synthesis. Reverse transcription was performed in a 20 μ L volume consisting of 4 μ L of 5 \times RT buffer, 1 μ L 10mM dNTP (deoxy-triphosphatnucleotides solution), 2 μ L 0.1M DTT, 1 μ L RNaseOUT™ (40 U/ μ L), 1 μ L moloney murine leukemia virus (M-MLV) (200 U/ μ L), and so on. All samples were incubated at 25°C for 10 min for primer annealing, then incubated at 37°C for 50 min.

Transcriptome analysis

A total of 10 µg RNA (800 ng/µl) extracted from the whole blood of Gentoo penguins and those derived from the spleen and bursa of Fabricius of zebra finches were subjected to transcriptome sequencing. The RNA libraries were examined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and the ABI StepOnePlus Real-Time PCR System (Life Technologies, Waltham, MA, USA) and then sequenced using an Illumina HiSeq™ 2000 platform (Illumina, San Diego, CA, USA).

3' RACE using J_H-derived primers and cDNA library construction

Total RNA from Gentoo penguin whole blood and zebra finch spleen was reverse-transcribed using the primer NotI-d(T)₁₈ (5' AAC TGG AAG AAT TCG CGG CCG CAG GAA TTT TTT TTT TTT TTT TTT 3'). The 3' RACE PCR parameters were as follows: 94°C for 5 min; followed by 40 cycles at 94°C for 30 s, 60°C for 30 s, 72°C for 90 s; and a final extension at 72°C for 5 min with LA-*Taq* DNA Polymerase (Takara, Dalian, Liaoning Province, China). The primers used were based on the conserved J_H region as follows: Gentoo penguin J_H GSP1 (5' ATC GAC GCG TGG GGC AGC 3') and GSP2 (5' GCA GCG GGA CCT CCG TCA CCG TCT CCT C 3'); zebra finch J_H GSP1 (5' ATT GAC GCC TGG GGC AGC 3') and GSP2 (5' ACC GTC GTC ACC GTC AGC 3'). Two PCR products were cloned into the pMD19-T vector (Takara, Dalian, Liaoning Province, China) to generate Ig cDNA mini-libraries.

Analysis of CDR3 fragments

We used the 5' RACE System for Rapid Amplification of cDNA Ends (Invitrogen, Carlsbad, CA, USA) to study CDR3. The RACE PCR was performed using C_µ-derived primers and total RNA derived from whole blood or the spleen. Positive PCR products were cloned and sequenced. The CDR3 length and amino acid usage were then calculated.

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

Total RNA was isolated from various tissues (liver, spleen, lung, stomach, small intestine, large intestine, and the bursa of Fabricius) using an RNeasy Mini Kit (Qiagen, North Rhine-Westphalia, Hilden, Germany). RNA quantity and quality were assessed with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). First-strand cDNA was synthesized using a QuantiTect Reverse Transcription Kit (Qiagen, North Rhine-Westphalia, Hilden, Germany) and oligo(dT)₂₀ primers. Quantitative RT-PCR (qPCR) was performed using the LightCycler480 and LightCycler480 SYBR Green I Master Mix (Roche, Basel, Basel-Stadt, Switzerland) with the following cycling conditions: 95°C for 5 min; then 40 cycles at 95°C for 10 s, 60°C for 10 s, and 72°C for 10 s. Each sample was run in triplicate. The primers used were as follows: EF1A1 (sense: 5' GCG TGA GCG TGG TAT CAC TA 3', antisense: 5' ACA CCA GCA GCA ACA ATC AG 3'); IGHM (sense: 5' GGA GGT CAC CAT GAG GAA GA 3', antisense: 5' GAG CCA TTG GAG GAG GAT TT 3'); IGHA (sense: 5' GGA CGA AGG GAA CAA CTT CA 3', antisense: 5' GTG GTC GGA GAG GAA AAG GT 3'); IGHY (sense: 5' GGA CTC CAT CAA CGT CCA GT 3', antisense: 5' CAC TCT TCC CGT TCC ACT TC 3').

Sequence alignment and phylogenetic analysis

The accession numbers for the sequences used in the construction of phylogenetic trees in this study are presented in supplemental data (S1 Table). The nurse shark IgM gene was used as an outgroup. The multiple sequence alignments and percent identity calculations were performed

with ClustalW method using the MegAlign program of the LASERGENE bioinformatics computing suite (DNASTAR program). Phylogenetic trees were generated using MrBayes3.2 [29] and viewed in TreeView [30]. Multiple sequence alignments were performed using the ClustalW program [31].

Results

Ig isotypes expressed in the Gentoo penguin

To determine the IgH classes expressed in the Gentoo penguin, transcriptome sequencing was performed on total RNA samples isolated from whole blood. A total of 83,583,032 clean reads ranging in size from 200 bp to 500 bp were generated, accounting for 148,562 annotated unigene sequences. BLAST searches were performed against these unigenes, and three IgH partial fragments (μ , α and ν) were identified. Because distinct IgH isotypes can share identical JH segments, we performed 3' RACE using JH gene segment primers and constructed a blood IgH gene library. From the 200 clones in the library, we identified membrane-bound and secretory forms of IgM and IgA and only the secretory form of IgY. We also detected a transcript encoding a short membrane-bound form of IgA (lacking the last two CH exons); however, no IgD gene was detected. The identities of these genes were verified by phylogenetic analysis.

Ig isotypes expressed in the zebra finch

To better characterize the IgH classes expressed in the zebra finch, transcriptome sequencing was performed on total mixed RNA samples isolated from the spleen and bursa of Fabricius. A total of 55,008,340 clean reads ranging in size from 125 bp to 250 bp were generated, accounting for 43,308 annotated unigene sequences. After BLAST searches against these unigenes, we identified three IgH partial fragments (μ , α and ν). We performed 3' RACE using the JH gene segment primers and RNA samples isolated from the spleen, the tissue believed to have the highest level of IgH gene expression. Consequently, we obtained the secretory form of the IgM, IgA and IgY heavy chain transcripts, similar to those observed in the Gentoo penguin; however, no membrane-bound forms of IgM, IgA or IgY were detected. The identities of these genes were verified by phylogenetic analysis.

To investigate whether the Gentoo penguin and/or the zebra finch express IgD, the degenerate primers based on the conserved C δ regions of other species, including the ostrich, were used. Consistent with the results described above, we did not find any IgD segments.

Characterization of the IgM CH gene

Similar to that in other species, the secretory form of IgM heavy chain constant region amplified from the Gentoo penguin encodes 449 amino acids consisting of four CH domains and shares 64.4% and 48.8% identity with ostrich IgM1 and ostrich IgM2, respectively. Overall, there were three potential N-linked glycosylation sites (N-X-S/T) identified throughout the entire constant region (N-45, N-199 and N-434), all of which are conserved among amphibians, reptiles, birds and humans. However, the N-127 site, which is conserved in other species including the zebra finch, was absent in the Gentoo penguin (Fig 1). The identities of the μ genes were confirmed by phylogenetic analysis (Fig 2).

The secretory form of the IgM heavy chain constant region from the zebra finch encodes 458 amino acids with a conserved three amino acid motif (TCY) in the carboxy-terminal, which is the site where cysteine is assumed to bind the J chain for polymer formation. The amino acid sequence of the IgM constant region shares 44.2% and 37.8% overall identity with ostrich IgM1 and ostrich IgM2, respectively. These two values are much lower than those for the Gentoo

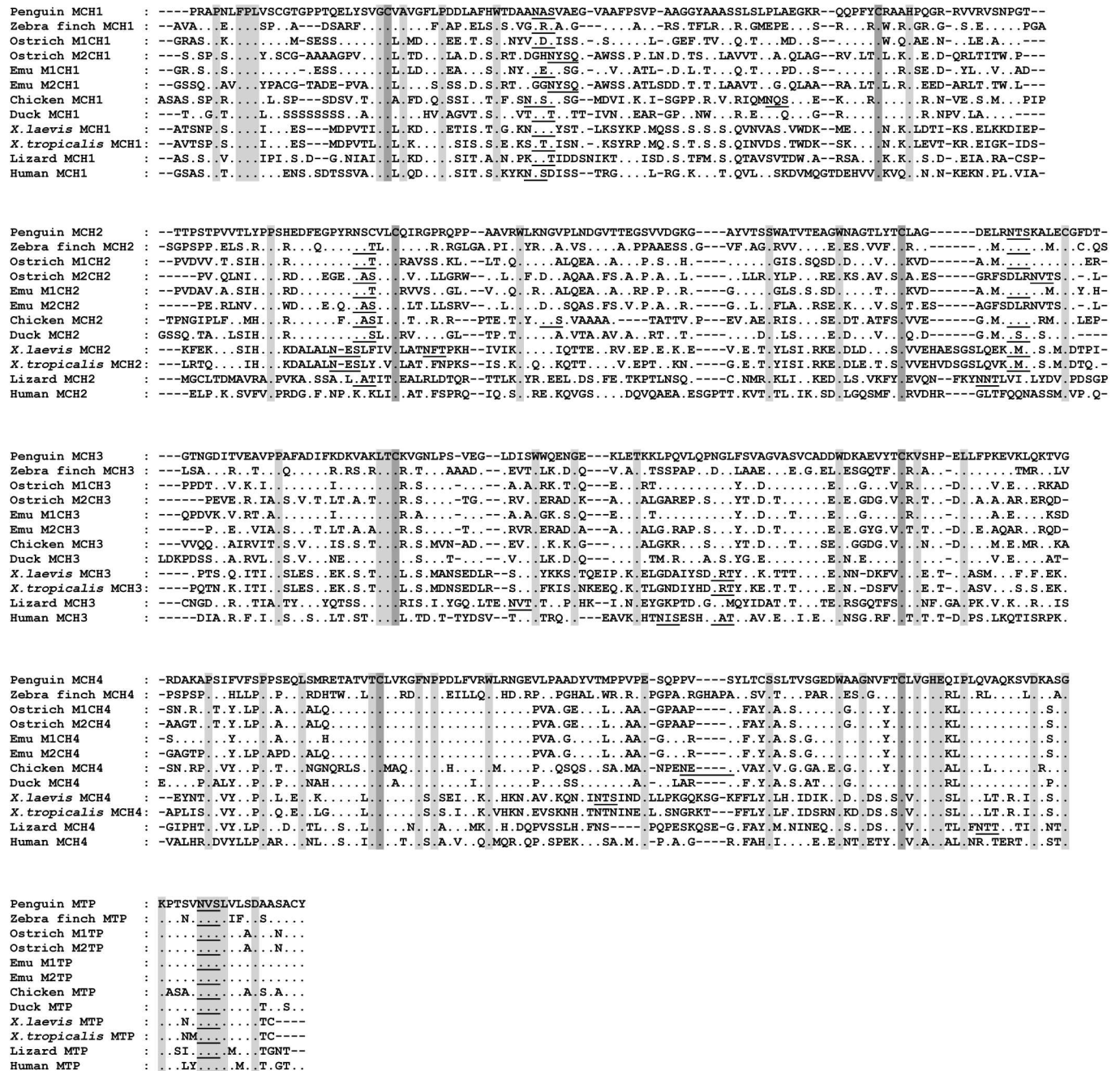


Fig 1. Multiple sequence alignment of the Gentoo penguin, zebra finch, ostrich, emu, chicken, duck, *X. laevis*, *X. tropicalis*, lizard and human Cμ gene. MTP, secretory tailpiece of Cμ gene. Identical amino acid residues were denoted as dots, and conserved amino acid residues were shaded. Canonical cysteines were marked with dark shade. Dashes were used to adjust the sequence alignment. Potential N-linked glycosylation sites were underlined. The alignment was performed using ClustalW with some manual adjustment.

<https://doi.org/10.1371/journal.pone.0173334.g001>

penguin, in accordance with our phylogenetic analysis. There are four conserved N-linked glycosylation sites (N-X-S/T; N-44, N-127, N-202, and N-444) in zebra finch IgM constant region, all of which are conserved among amphibians, reptiles, birds, and mammals (Fig 2).

Sequence alignment of the IgM constant regions of the Gentoo penguin and zebra finch with those of other species by Clustal W demonstrated that the Cμ3 and Cμ4 domains are less

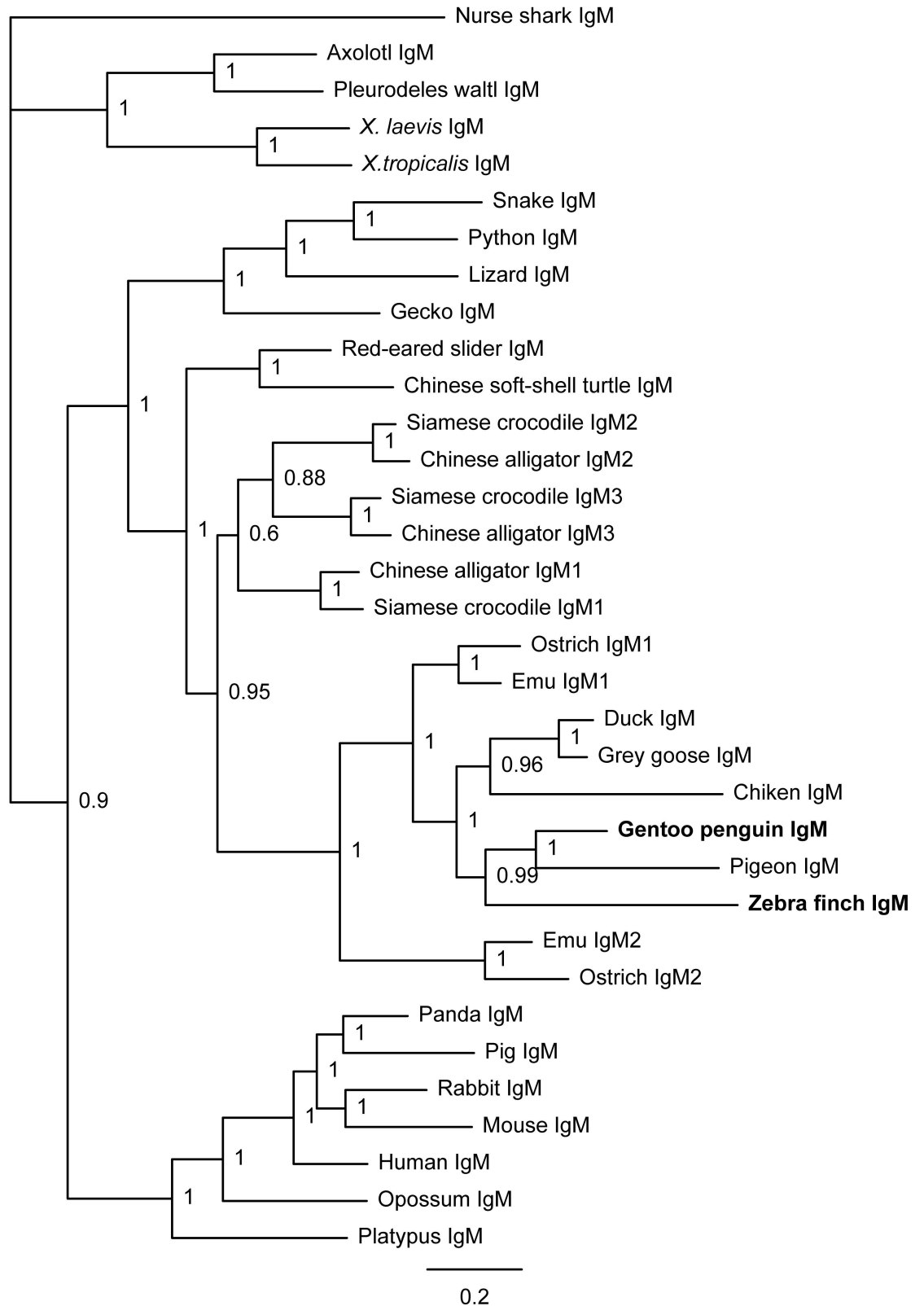


Fig 2. Phylogenetic analysis of IgM-encoding genes in tetrapods. Phylogenetic trees were constructed with MrBayes 3.1.2 using the amino acid sequences of the IgH constant region derived from various vertebrates. The scale bar indicates the genetic distance, and the credibility value is shown for each node. The IgM sequences of the Gentoo

penguin and zebra finch were obtained in this study; all of the other sequences were obtained from NCBI GenBank (for accession numbers see the [Materials and Methods](#) section).

<https://doi.org/10.1371/journal.pone.0173334.g002>

divergent than the $C\mu 1$ and $C\mu 2$ domains. We did not find any evidence indicating IgM subclass diversification in the Gentoo penguin or the zebra finch.

Characterization of the IgA CH gene

As an important first line of defense, IgA is the principal antibody class in mucosal secretions [32]. The IgA sequences obtained from both the Gentoo penguin and zebra finch encode four CH domains. In addition, we detected a transcript encoding a short membrane-bound form of IgA (lacking the last two CH exons) in the Gentoo penguin. The short IgA consists of 252 amino acids and shares a 50.6% identity with the ostrich IgA. The four CH domains of the secretory IgA and the short membrane-bound IgA were comparatively analyzed among several birds (Gentoo penguin, zebra finch, ostrich, emu, chicken, and duck), reptile (lizard), amphibians (*X. tropicalis* and *X. laevis*) and mammal (Human) (Fig 3). The amino acid sequences of full length $C\alpha$ gene in the Gentoo penguin and zebra finch share 56.1% and 36.6% identity with the $C\alpha$ gene in ostrich, respectively. The degree of sequence identity between the Gentoo penguin and the ostrich is higher than between the zebra finch and ostrich, which agrees with the phylogenetic analysis of the IgA gene (Fig 4). Protein sequence alignment of the $C\alpha$ genes revealed that ten conserved cysteine residues are distributed in an identical pattern between the different species. There are two N-linked glycosylation sites in CH1, CH2, and one in the IgA secretory tail, all of which are conserved in birds (Fig 3).

Characterization of the IgY CH gene

Initial BLAST searches against the 48 recently sequenced bird genomes showed that the penguin and zebra finch may have two IgY genes. However, only a single IgY heavy chain constant region was identified from both the Gentoo penguin and zebra finch in our study. The secretory form of IgY includes four CH domains. Compared with ν in the Gentoo penguin, the ν gene in zebra finch shows a more conserved cysteine distribution relative to IgY-encoding genes in other birds, reptiles and amphibians. Sequence alignment with other species revealed two cysteines in the $C\nu 1$, which suggests that these molecules associate with the light chain. Moreover, eight additional conserved cysteines were observed in $C\nu 2$ - $C\nu 4$ across all of the species examined. The secretory form of the IgY constant region in the Gentoo penguin contains three potential N-linked glycosylation sites, whereas the zebra finch IgY constant region contains two potential N-linked glycosylation sites. The N-413 site is exclusive to the Gentoo penguin. Compared with IgY2 in ostrich and emu, IgY1 shows a less conserved cysteine distribution. For example, IgY2 features a conserved double cysteine in the N-terminal of CH1 and another cysteine in the N-terminal of CH2. The differences were marked with black boxes (Fig 5).

Using the amino acid sequences of IgY-encoding genes from different jawed vertebrates, we constructed an IgY phylogenetic tree. The results suggest that subclass gene diversification likely occurred during the early stages of bird evolution but after their divergence from non-avian reptiles (Fig 6).

Analysis of rearranged VDJ fragments

To analyze the expressed VDJ fragments, we designed primers based on the μ gene constant region and amplified the VH regions associated with IgM heavy chains using 5' RACE. For both the Gentoo penguin and zebra finch, all of the obtained sequences began with the same

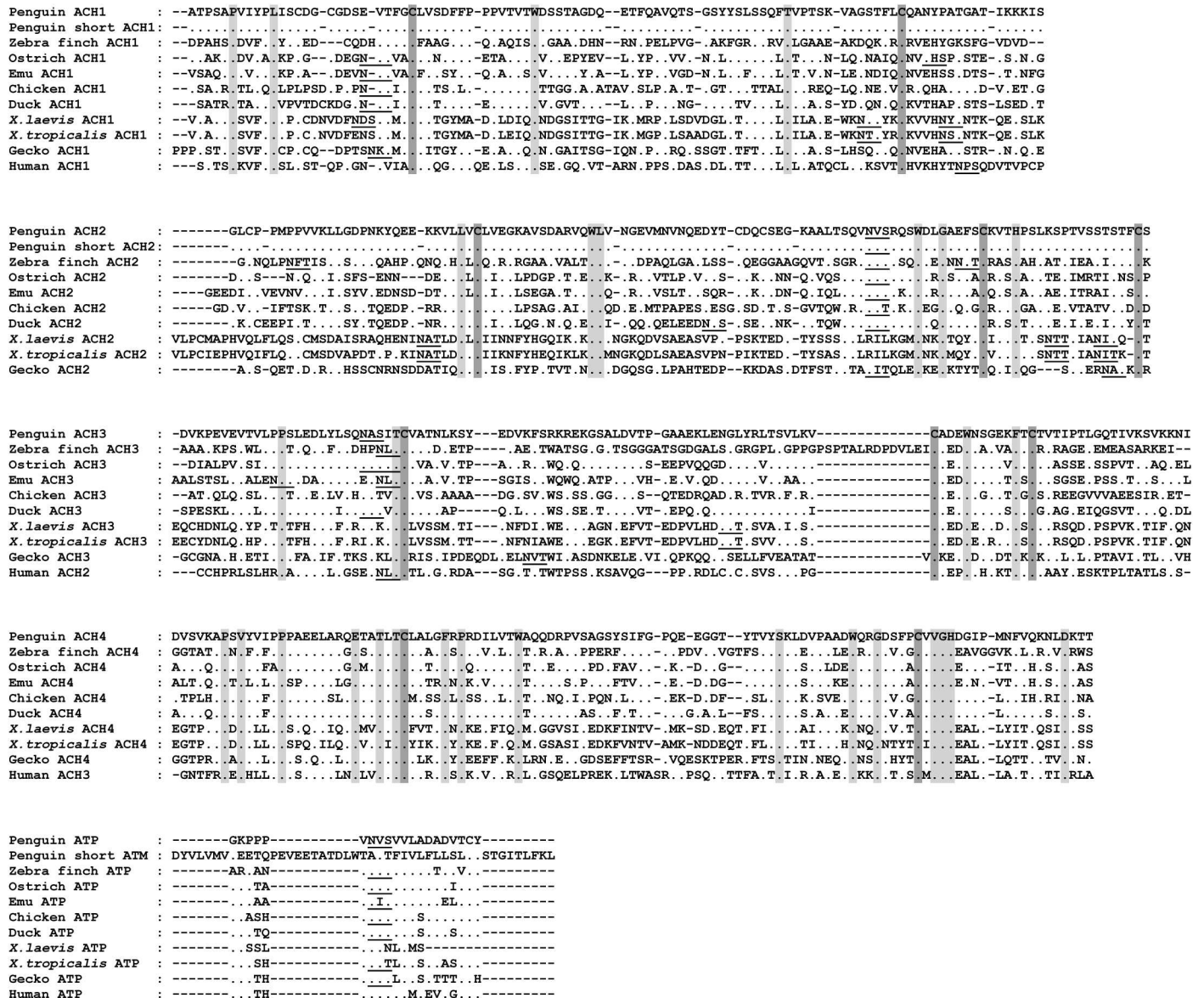


Fig 3. Multiple sequence alignment of the Gentoo penguin, zebra finch, ostrich, emu, chicken, duck, *X. laevis*, *X. tropicalis*, gecko, human *Ca* gene and Gentoo penguin short *Ca* gene. ATP, secretory tailpiece of *Ca* gene. ATM, transmembrane domain of *Ca* gene. Identical amino acid residues were denoted as dots, and conserved amino acid residues were shaded. Canonical cysteines were marked with dark shade. Dashes were used to adjust the sequence alignment. Potential N-linked glycosylation sites were underlined. The alignment was performed using ClustalW with some manual adjustment.

<https://doi.org/10.1371/journal.pone.0173334.g003>

leading peptide-encoding sequence. This strongly suggests that only one VH gene can participate in VDJ recombination, and that gene conversion is the primary mechanism for antibody diversity in these species. CDR3 plays an important role in determining antibody specificity and affinity, and the length of the CDR3 region is important for the capacity to bind antigen [33]. Among the 251 clones obtained from the Gentoo penguin, 48 unique sequences were further analyzed to determine CDR3 length. The results showed that CDR3 varies from 5 to 19 amino acids, with the most common length being 12 amino acids. For the zebra finch, 200 cDNA fragments that contained 119 unique CDR3 regions were sequenced. The length of

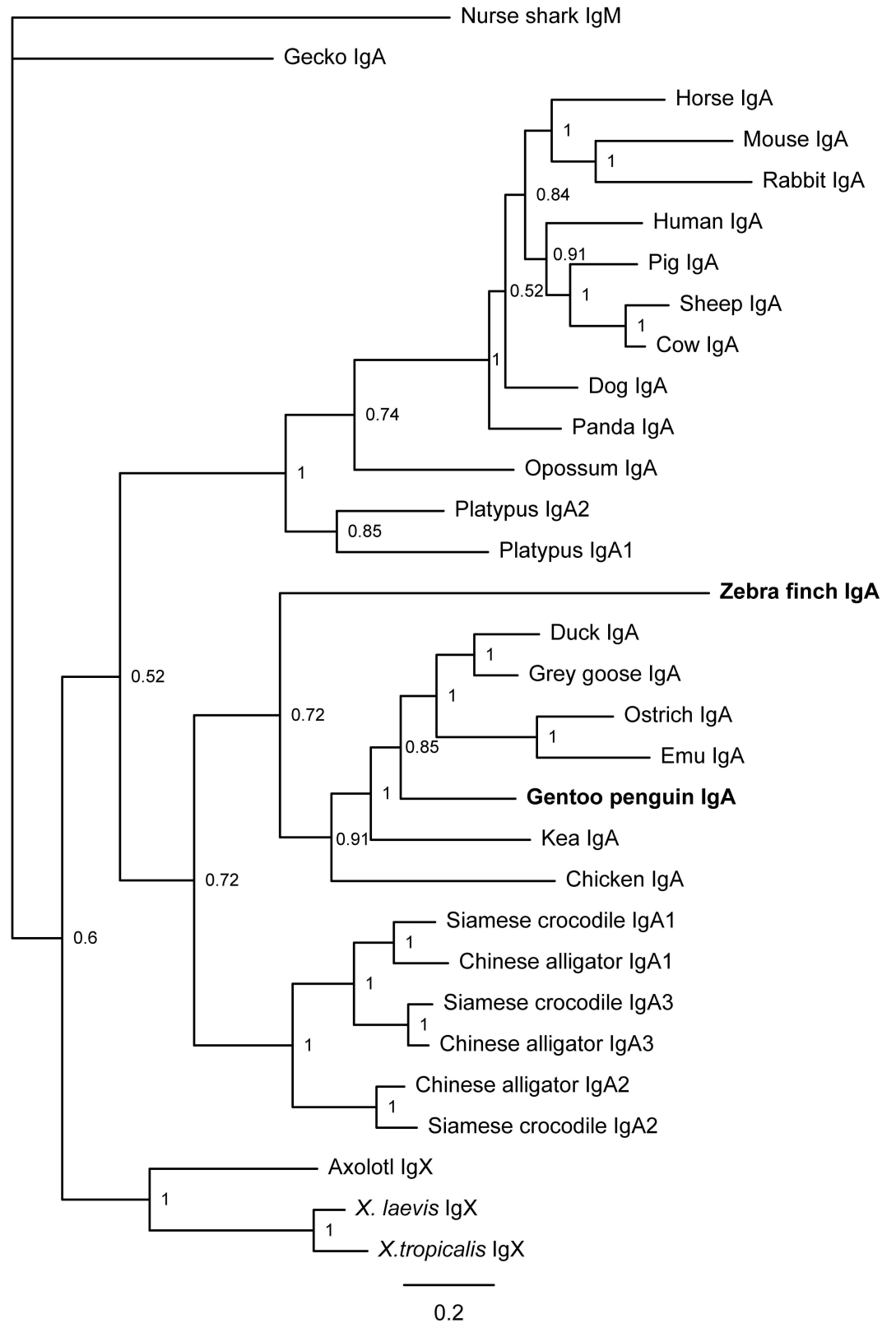


Fig 4. Phylogenetic analysis of IgA-encoding genes in tetrapods. Phylogenetic trees were constructed with MrBayes 3.1.2 using the amino acid sequences of the IgH constant region derived from various vertebrates. The scale bar indicates genetic distance, and the credibility value is shown for each node. The IgA sequences from the

Gentoo penguin and zebra finch were obtained in this study; all of the other sequences were obtained from NCBI GenBank (for accession numbers see the [Materials and Methods](#) section).

<https://doi.org/10.1371/journal.pone.0173334.g004>

CDR3 varies from 14 to 19 residues, which is much smaller range than that of the Gentoo penguin. The length of the span is much shorter in the Gentoo penguin and zebra finch than in the ostrich and emu ([Fig 7](#)). There are more than two J_H gene segments in the Gentoo penguin and zebra finch (17 and 18, respectively).

Quantitative real-time PCR (q-PCR) in the zebra finch

The expression pattern of immunoglobulin transcripts were examined in various tissues (liver, spleen, lung, stomach, small intestine, large intestine, and bursa of Fabricius) in the zebra finch using q-PCR. IgM is highly expressed in immune-related organs such as the spleen and the bursa of Fabricius, as expected. IgY is highly expressed in the spleen and at relatively high

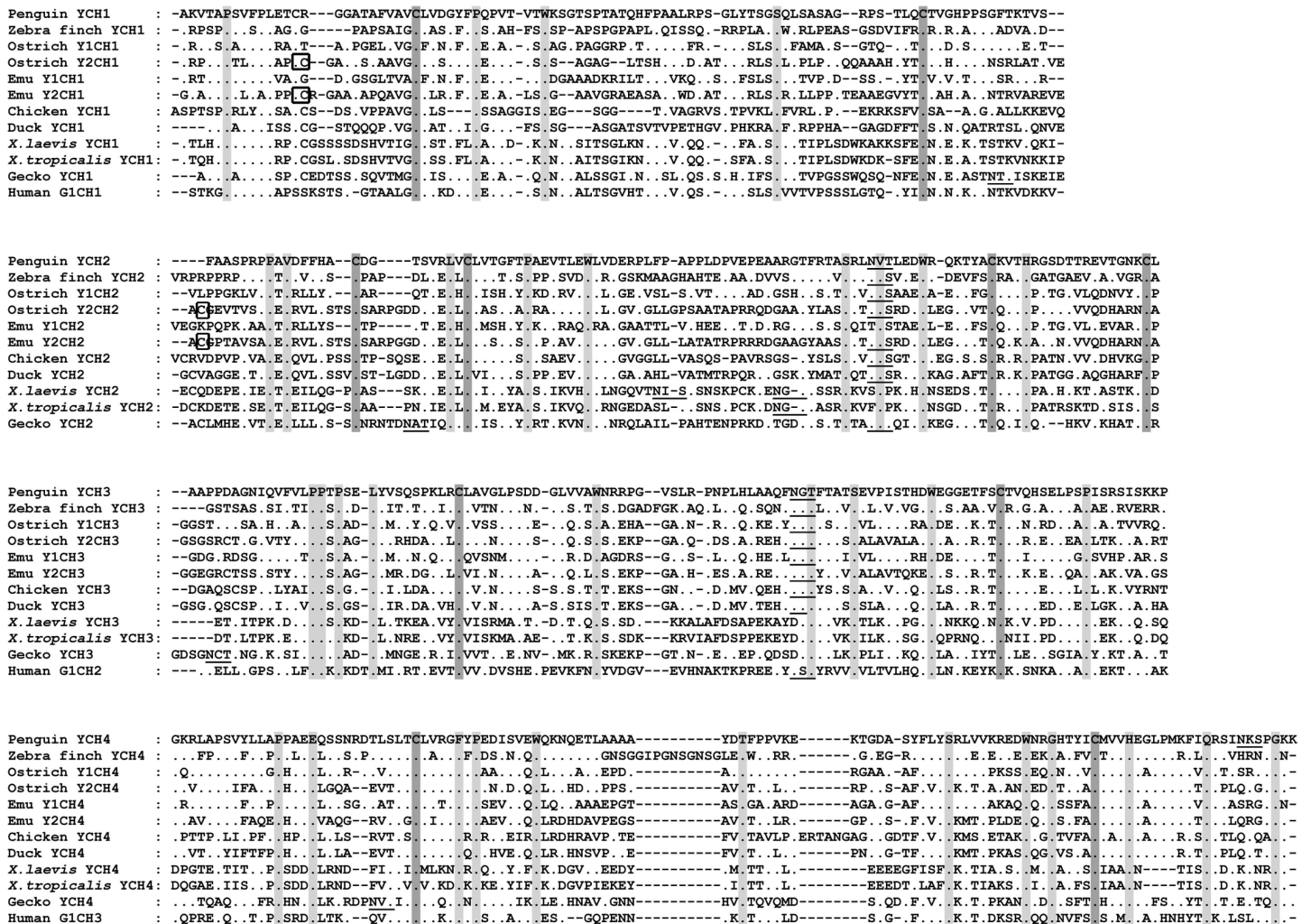


Fig 5. Multiple sequence alignment of the Gentoo penguin, zebra finch, ostrich, emu, chicken, duck, *X. laevis*, *X. tropicalis*, gecko *Cu* gene and human *Cy1* gene. Identical amino acid residues were denoted as dots, and conserved amino acid residues were shaded. Canonical cysteines were marked with dark shade. Dashes were used to adjust the sequence alignment. Potential N-linked glycosylation sites were underlined. The alignment was performed using ClustalW with some manual adjustment. The black boxes were used to mark the different cysteines between IgY1 and IgY2.

<https://doi.org/10.1371/journal.pone.0173334.g005>

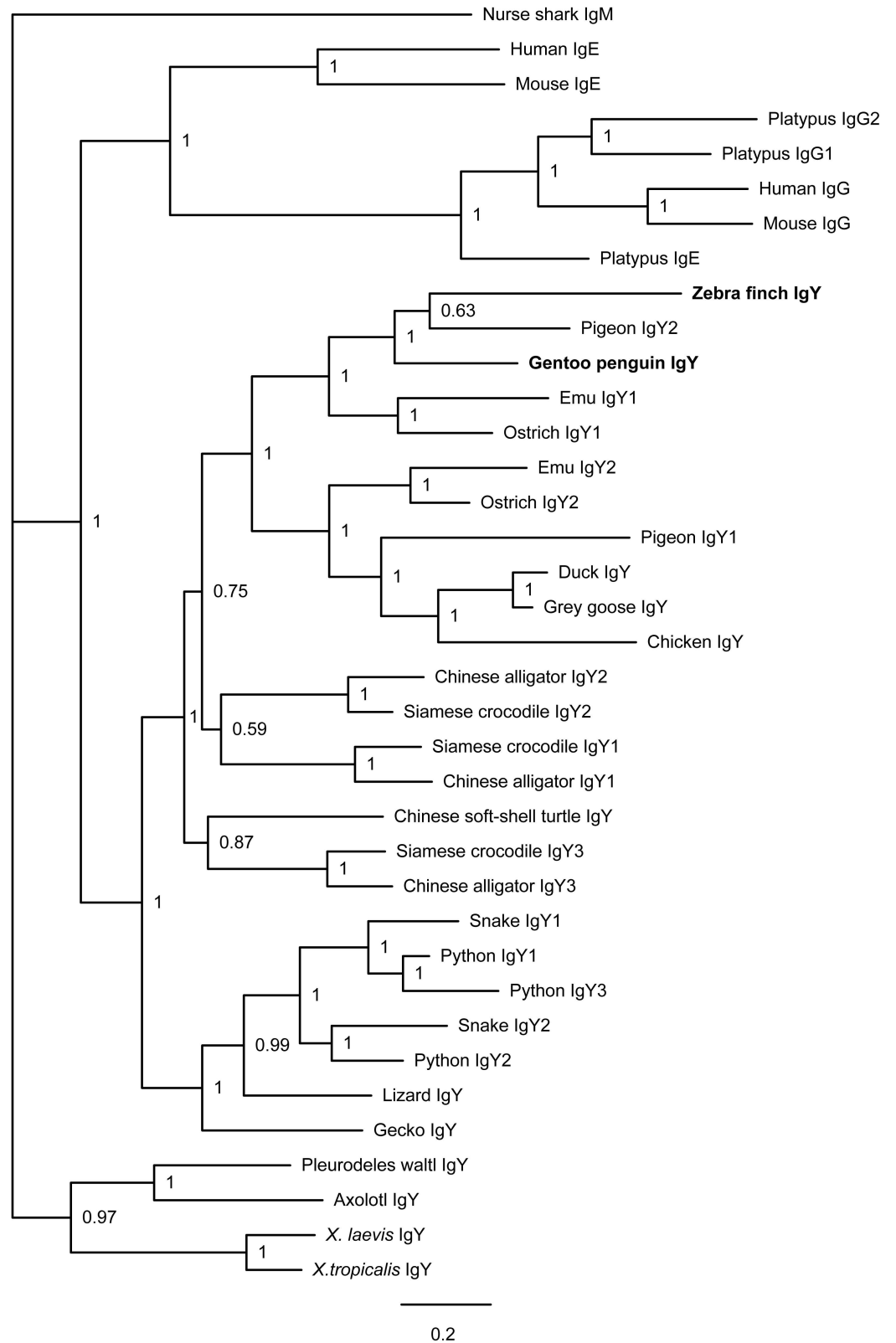


Fig 6. Phylogenetic analysis of IgY-encoding genes in tetrapods. Phylogenetic trees were constructed with MrBayes 3.1.2 using the amino acid sequences of the IgH constant region derived from various vertebrates. The scale bar indicates genetic distance, and the credibility value is shown for each node. The IgY sequences of the

Gentoo penguin and zebra finch were obtained in this study; all of the other sequences were obtained from NCBI GenBank (for accession numbers see the [Materials and Methods](#) section).

<https://doi.org/10.1371/journal.pone.0173334.g006>

levels in the liver, the bursa of Fabricius and the intestine. IgA is usually highly expressed in mucosal tissue and is the main antibody class in mucosal secretions, acting as an important first line of defense [32]. In accordance with the expected results, our q-PCR data showed that zebra finch IgA is primarily expressed in the small intestine (Fig 8).

Discussion

The ostrich (*Struthio camelus*) is a primitive avian species of the order Struthioniformes, having diverged from other avian lineages approximately 140 MYA [34]. Based on the recently sequenced genomes of 48 bird species and the high-throughput transcriptome sequencing of immune-related tissues, we demonstrated that the ostrich (*S. camelus*) possesses a functional δ gene similar to that of reptiles, such as lizards and turtles, as well as IgM and IgY subclass diversification, as in crocodylians [4,28,35]. Therefore, this study was performed for the following reasons: to determine whether IgM and IgY subclass diversification is present in the Sphenisciformes or Passeriformes and, ultimately, how widespread this phenomenon is in birds. However, in this study, only three Ig isotypes (IgM, IgA and IgY) were identified, and there was no evidence of a functional IgD gene or of subclass diversification in the Gentoo penguin or zebra finch, which is the case for most birds.

IgM subclass diversification is not common in avian reptiles (birds), although it is common in non-avian reptiles (Crocodylia, Testudines and Squamata) [36, 37]. To date, it has been reported that there are three functional μ genes in the crocodylian IgH locus and two in the ostrich and emu, and at least four Squamata species have more than one μ gene [4, 35, 38]. The isotypes mentioned above originated from duplications of the exons that encode for the $C\mu$ located at the 5' end. Phylogenetic analyses suggested that the duplication in birds occurred

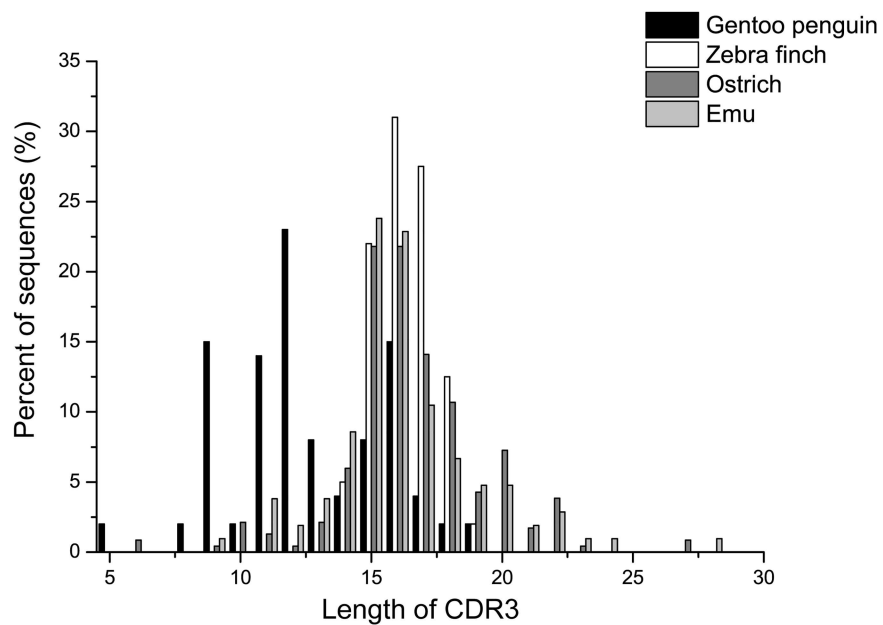


Fig 7. CDR3 length distribution in different species. The x-axis represents CDR3 length. The y-axis represents the relative frequency of each length.

<https://doi.org/10.1371/journal.pone.0173334.g007>

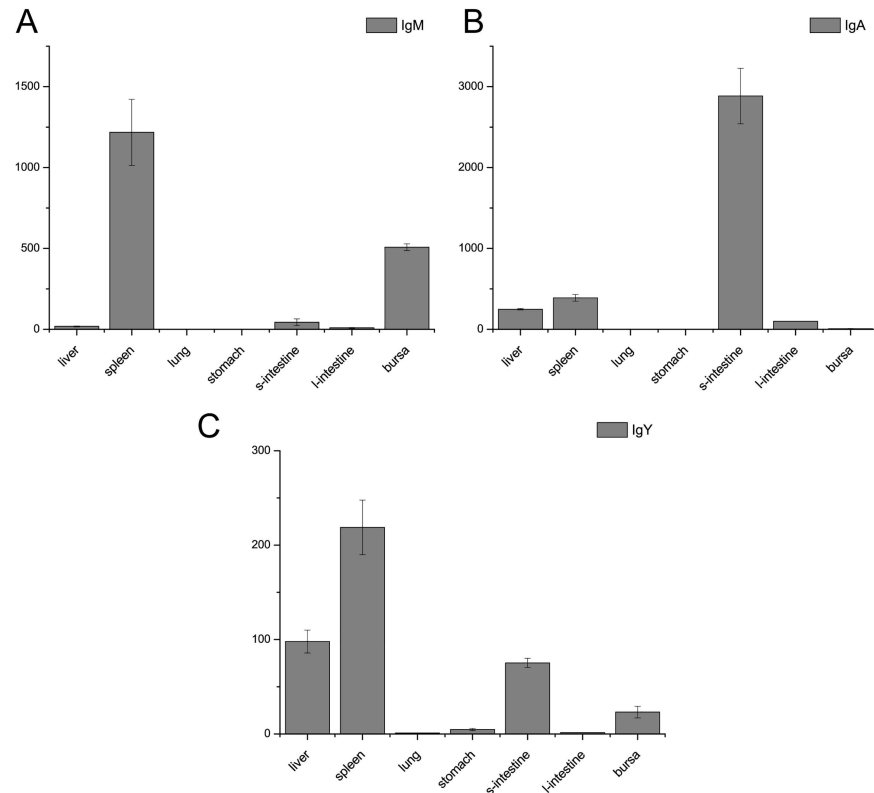


Fig 8. Expression of three immunoglobulin genes in the zebra finch. The data are representative of three independent experiments. The y-axis indicates fold changes in normalized expression. The vertical lines indicate standard deviations from the mean. The *EF1A1* gene was used as an internal control, and the values shown in the figure were calculated using the $\Delta\Delta Ct$ method. s-intestine: small intestine; l-intestine: large intestine. (A) Expression of IgM in the zebra finch. (B) Expression of IgA in the zebra finch. (C) Expression of IgY in the zebra finch.

<https://doi.org/10.1371/journal.pone.0173334.g008>

after the divergence of Aves and reptiles and before the divergence of modern birds. Although the ostrich and emu both have two μ genes, many other birds, including the Gentoo penguin and zebra finch, evolutionarily retained only the $\mu 1$ gene and lost the $\mu 2$ gene (Table 1).

As shown in the phylogenetic trees, the gene subclass diversification of ν genes was similar to IgM, also occurring in the early stages of bird evolution. Although the ostrich and emu, as primitive birds, encode two ν genes, many other birds retained only one during evolution. Unlike that almost all birds retain IgM1, our data show that some birds retain IgY1, while others retain IgY2 (Table 1).

Table 1. The IgH gene isotypes in different birds.

Order	Species (references)	IgM1	IgM2	IgD	IgA	IgY1	IgY2
Struthioniformes	Ostrich [4]	✓	✓	✓	✓	✓	✓
Casuariiformes	Emu [4]	✓	✓	✓	✓	✓	✓
Sphenisciformes	Penguin	✓			✓	✓	
Passeriformes	Zebra finch	✓			✓	✓	
Galliformes	Chicken [12]	✓			✓		✓
Anseriformes	Duck [25]	✓			✓		✓
Anseriformes	Goose [39]	✓			✓		✓

<https://doi.org/10.1371/journal.pone.0173334.t001>

In summary, based on transcriptome sequencing and additional experiments, we successfully identified three immunoglobulin heavy chain genes (μ , α and ν) expressed in the Gentoo penguin and zebra finch. Furthermore, a single leader peptide in the expressed heavy chain V regions indicates that gene conversion also plays a major role in the generation of antibody diversity in these two birds. Our study enriches the current knowledge of IgH gene evolution in birds and provides more data to elucidate the evolutionary history of the adaptive immune system.

Supporting information

S1 Table. The accession numbers used in the construction of phylogenetic trees.
(DOCX)

Acknowledgments

The authors wish to thank Drs. Shoubin Zhao, Ming Zhang, Di Yu, Peng Wang, and Cuncun Ke for their help and suggestions.

Author Contributions

Conceptualization: YZ LR HH.

Data curation: BH YL.

Formal analysis: BH YL.

Funding acquisition: YZ QP.

Investigation: BH YL.

Methodology: BH YL.

Project administration: LR HH.

Resources: YZ LR HH QP.

Software: BH.

Supervision: YZ LR HH QP.

Validation: BH YL.

Visualization: BH LR.

Writing – original draft: BH.

Writing – review & editing: BH LR YZ.

References

1. Flajnik MF. Comparative analyses of immunoglobulin genes: surprises and portents. *Nat Rev Immunol*. 2002; 2(9):688–98. <https://doi.org/10.1038/nri889> PMID: 12209137
2. Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature reviews Genetics*. 2010; 11(1):47–59. PubMed Central PMCID: PMC3805090. <https://doi.org/10.1038/nrg2703> PMID: 19997068
3. Bengten E, Wilson M, Miller N, Clem LW, Pilstrom L, Warr GW. Immunoglobulin isotypes: structure, function, and genetics. *Current topics in microbiology and immunology*. 2000; 248:189–219. PMID: 10793479

4. Han B, Yuan H, Wang T, Li B, Ma L, Yu S, et al. Multiple IgH Isotypes Including IgD, Subclasses of IgM, and IgY Are Expressed in the Common Ancestors of Modern Birds. *Journal of immunology*. 2016; 196(12):5138–47.
5. Warr GW, Magor KE, Higgins DA. IgY: clues to the origins of modern antibodies. *Immunology today*. 1995; 16(8):392–8. [https://doi.org/10.1016/0167-5699\(95\)80008-5](https://doi.org/10.1016/0167-5699(95)80008-5) PMID: 7546196
6. Sun Y, Wei Z, Hammarstrom L, Zhao Y. The immunoglobulin delta gene in jawed vertebrates: a comparative overview. *Developmental and comparative immunology*. 2011; 35(9):975–81. <https://doi.org/10.1016/j.dci.2010.12.010> PMID: 21182859
7. Ramirez-Gomez F, Greene W, Rego K, Hansen JD, Costa G, Kataria P, et al. Discovery and characterization of secretory IgD in rainbow trout: secretory IgD is produced through a novel splicing mechanism. *Journal of immunology*. 2012; 188(3):1341–9.
8. Zhao Y, Cui H, Whittington CM, Wei Z, Zhang X, Zhang Z, et al. Ornithorhynchus anatinus (platypus) links the evolution of immunoglobulin genes in eutherian mammals and nonmammalian tetrapods. *Journal of immunology*. 2009; 183(5):3285–93.
9. Li L, Wang T, Sun Y, Cheng G, Yang H, Wei Z, et al. Extensive diversification of IgD-, IgY-, and truncated IgY(deltaFc)-encoding genes in the red-eared turtle (*Trachemys scripta elegans*). *Journal of immunology*. 2012; 189(8):3995–4004.
10. Wang T, Sun Y, Shao W, Cheng G, Li L, Cao Z, et al. Evidence of IgY subclass diversification in snakes: evolutionary implications. *Journal of immunology*. 2012; 189(7):3557–65.
11. Lundqvist ML, Middleton DL, Hazard S, Warr GW. The immunoglobulin heavy chain locus of the duck. Genomic organization and expression of D, J, and C region genes. *The Journal of biological chemistry*. 2001; 276(50):46729–36. <https://doi.org/10.1074/jbc.M106221200> PMID: 11592961
12. Zhao Y, Rabbani H, Shimizu A, Hammarstrom L. Mapping of the chicken immunoglobulin heavy-chain constant region gene locus reveals an inverted alpha gene upstream of a condensed epsilon gene. *Immunology*. 2000; 101(3):348–53. PubMed Central PMCID: PMC2327085. <https://doi.org/10.1046/j.1365-2567.2000.00106.x> PMID: 11106938
13. Lanning DK, Zhai SK, Knight KL. Analysis of the 3' Cmu region of the rabbit Ig heavy chain locus. *Gene*. 2003; 309(2):135–44. PMID: 12758129
14. Wang X, Olp JJ, Miller RD. On the genomics of immunoglobulins in the gray, short-tailed opossum *Monodelphis domestica*. *Immunogenetics*. 2009; 61(8):581–96. PubMed Central PMCID: PMC2880577. <https://doi.org/10.1007/s00251-009-0385-8> PMID: 19609519
15. Ohta Y, Flajnik M. IgD, like IgM, is a primordial immunoglobulin class perpetuated in most jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(28):10723–8. PubMed Central PMCID: PMC1636022. <https://doi.org/10.1073/pnas.0601407103> PMID: 16818885
16. Amemiya CT, Haire RN, Litman GW. Nucleotide sequence of a cDNA encoding a third distinct *Xenopus* immunoglobulin heavy chain isotype. *Nucleic acids research*. 1989; 17(13):5388. PubMed Central PMCID: PMC318127. PMID: 2503814
17. Zhao Y, Pan-Hammarstrom Q, Yu S, Wertz N, Zhang X, Li N, et al. Identification of IgF, a hinge-region-containing Ig class, and IgD in *Xenopus tropicalis*. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(32):12087–92. PubMed Central PMCID: PMC1567701. <https://doi.org/10.1073/pnas.0600291103> PMID: 16877547
18. Danilova N, Bussmann J, Jekosch K, Steiner LA. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. *Nature immunology*. 2005; 6(3):295–302.
19. Hansen JD, Landis ED, Phillips RB. Discovery of a unique Ig heavy-chain isotype (IgT) in rainbow trout: Implications for a distinctive B cell developmental pathway in teleost fish. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(19):6919–24. PubMed Central PMCID: PMC1100771. <https://doi.org/10.1073/pnas.0500027102> PMID: 15863615
20. Greenberg AS, Avila D, Hughes M, Hughes A, McKinney EC, Flajnik MF. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature*. 1995; 374(6518):168–73. <https://doi.org/10.1038/374168a0> PMID: 7877689
21. Schaerlinger B, Bascove M, Fripiat JP. A new isotype of immunoglobulin heavy chain in the urodele amphibian *Pleurodeles waltl* predominantly expressed in larvae. *Molecular immunology*. 2008; 45(3):776–86. <https://doi.org/10.1016/j.molimm.2007.06.356> PMID: 17681605
22. Parvari R, Ziv E, Lentner F, Tel-Or S, Burstein Y, Schechter I. Analyses of chicken immunoglobulin light chain cDNA clones indicate a few germline V lambda genes and allotypes of the C lambda locus. *The EMBO journal*. 1987; 6(1):97–102. PubMed Central PMCID: PMC553362. PMID: 3107981

23. Reynaud CA, Anquez V, Grimal H, Weill JC. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell*. 1987; 48(3):379–88. PMID: [3100050](#)
24. Choi JW, Kim JK, Seo HW, Cho BW, Song G, Han JY. Molecular cloning and comparative analysis of immunoglobulin heavy chain genes from *Phasianus colchicus*, *Meleagris gallopavo*, and *Coturnix japonica*. *Veterinary immunology and immunopathology*. 2010; 136(3–4):248–56. <https://doi.org/10.1016/j.vetimm.2010.03.014> PMID: [20398946](#)
25. Lundqvist ML, Middleton DL, Radford C, Warr GW, Magor KE. Immunoglobulins of the non-galliform birds: antibody expression and repertoire in the duck. *Developmental and comparative immunology*. 2006; 30(1–2):93–100. PubMed Central PMCID: PMCPMC1317265. <https://doi.org/10.1016/j.dci.2005.06.019> PMID: [16150486](#)
26. Subramanian S, Beans-Picon G, Swaminathan SK, Millar CD, Lambert DM. Evidence for a recent origin of penguins. *Biology letters*. 2013; 9(6):20130748. PubMed Central PMCID: PMC3871361. <https://doi.org/10.1098/rsbl.2013.0748> PMID: [24227045](#)
27. Brown JW, Rest JS, Garcia-Moreno J, Sorenson MD, Mindell DP. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. *BMC biology*. 2008; 6:6. PubMed Central PMCID: PMC2267772. <https://doi.org/10.1186/1741-7007-6-6> PMID: [18226223](#)
28. Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, et al. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*. 2015; 526(7574):569–73. <https://doi.org/10.1038/nature15697> PMID: [26444237](#)
29. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012; 61(3):539–42. PubMed Central PMCID: PMCPMC3329765. <https://doi.org/10.1093/sysbio/sys029> PMID: [22357727](#)
30. Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*. 1996; 12(4):357–8. PMID: [8902363](#)
31. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, et al. Multiple sequence alignment with the Clustal series of programs. *Nucleic acids research*. 2003; 31(13):3497–500. PubMed Central PMCID: PMC168907. PMID: [12824352](#)
32. Woof JM, Kerr MA. The function of immunoglobulin A in immunity. *The Journal of pathology*. 2006; 208(2):270–82. <https://doi.org/10.1002/path.1877> PMID: [16362985](#)
33. Roman T, De Guerra A, Charlemagne J. Evolution of specific antigen recognition: size reduction and restricted length distribution of the CDRH3 regions in the rainbow trout. *Eur J Immunol*. 1995; 25(1):269–73. <https://doi.org/10.1002/eji.1830250144> PMID: [7843242](#)
34. Pereira SL, Baker AJ. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. *Molecular biology and evolution*. 2006; 23(9):1731–40. <https://doi.org/10.1093/molbev/msl038> PMID: [16774978](#)
35. Cheng G, Gao Y, Wang T, Sun Y, Wei Z, Li L, et al. Extensive diversification of IgH subclass-encoding genes and IgM subclass switching in crocodylians. *Nature communications*. 2013; 4:1337. <https://doi.org/10.1038/ncomms2317> PMID: [23299887](#)
36. Modesto SP, Anderson JS. The phylogenetic definition of reptilia. *Syst Biol*. 2004; 53(5):815–21. <https://doi.org/10.1080/10635150490503026> PMID: [15545258](#)
37. Shedlock AM, Botka CW, Zhao S, Shetty J, Zhang T, Liu JS, et al. Phylogenomics of nonavian reptiles and the structure of the ancestral amniote genome. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(8):2767–72. PubMed Central PMCID: PMC1815256. <https://doi.org/10.1073/pnas.0606204104> PMID: [17307883](#)
38. Olivieri DN, Garet E, Estevez O, Sanchez-Espinel C, Gambon-Deza F. Genomic structure and expression of immunoglobulins in Squamata. *Molecular immunology*. 2016; 72:81–91. <https://doi.org/10.1016/j.molimm.2016.03.003> PMID: [26998711](#)
39. Huang T, Wu K, Yuan X, Shao S, Wang W, Wei S, et al. Molecular analysis of the immunoglobulin genes in goose. *Developmental and comparative immunology*. 2016; 60:160–6. <https://doi.org/10.1016/j.dci.2016.02.024> PMID: [26921669](#)