Infection, Genetics and Evolution 68 (2019) 265–272

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

# Genome-wide analysis of the ovodefensin gene family: Monophyletic origin, independent gene duplication and presence of different selection patterns

Long Zhang[a,b,*], Dongmei Chen[a,b], Lintian Yu[c], Yi Wei[a,b], Juan Li[d], Caiquan Zhou[a,b,*]

[a] Institute of Ecology, Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong 637009, Sichuan Province, China
[b] Sichuan Wildlife Rehabilitation and Breeding Research Center, China West Normal University, Nanchong 637009, Sichuan Province, China
[c] Guangxi Agricultural Vocational College, Nanning 530007, Guangxi Province, China
[d] Institute of Animal Science, Chengdu Academy of Agriculture and Forestry Sciences, Chengdu 611130, Sichuan Province, China

ARTICLE INFO

*Keywords:*
Host defense peptide
Ovodefensin
Genome-wide identification
Evolutionary analysis

ABSTRACT

Ovodefensins (OvoDs) represent a group of cysteine-rich host defense peptides that are abundant in the egg white. Recent studies have found that ovodefensins are specific to birds and reptiles. However, the entire repertoire and evolutionary relationships of this gene family have not been thoroughly elucidated to date. Following our cross-species and genome-wide computational study, a total of 94 ovodefensin genes with multiple novel cysteine sequence motifs were identified from 22 phylogenetically divergent species. Phylogenetic analysis suggests that a large number of OvoDs evolved by gene duplication after species divergence. Furthermore, the OvoD genes in each species trend to be clustered densely in a syntenic region flanked by the XKR6 and MTMR9 genes, indicating that they are of monophyletic origin and appear to have emerged via independent gene duplication events in snakes, turtles, crocodiles, birds and the green lizard. Furthermore, positive selection sites are located primarily in the mature peptide region of the turtle, lizard and snake OvoD genes. Moreover, the duplicate OvoDAs in birds seem to be maintained in almost identical sequences and functions by strong purifying selection. Genome-wide identification and analyses of the OvoD gene family may greatly improve our understanding of the potential evolutionary relationship scenario of the OvoD gene family. Continued sequence mining and functional studies of OvoDs will be helpful in shedding light on the relationships between OvoDs and other defensin-related gene families.

## 1. Introduction

Defensins comprise a large group of cationic amphipathic peptides with diversified sequences that have been discovered in all three eukaryotic kingdoms (Ganz, 2004; Thomma et al., 2002). As defensins serve a critical role in innate immunity, these peptides are expressed by a variety of cells and are broadly active against bacteria, fungi, viruses and even cancerous cells (Hancock et al., 2012; Zasloff, 2002). In addition to their antimicrobial activities, defensins also play a profound role in immunomodulatory and barrier protective activities, such as activating immune cells, protecting the host from infections and regulating intestinal homeostasis (Hilchie et al., 2013; Mansour et al., 2014). Furthermore, some defensins and defensin-like peptides have been discovered to have roles in pigment (Candille et al., 2007), reproduction (Zhou et al., 2013) and venom toxicity (Rádis-Baptista et al., 2004; Whittington et al., 2008a).

In general, vertebrate defensins are further divided into three subfamilies (α-, β- and θ- defensins) based on their structure and the disulfide bonds of the cysteine residues (Li et al., 2014; Patil et al., 2004; Xiao et al., 2004). Recent research has further identified a subset of molecules that share high similarity with β-defensins, but that are only abundantly expressed in the oviduct; therefore, these peptides have been named as ovodefensins (OvoDs) (Whenham et al., 2015). Accumulating evidence has also indicated that OvoDs can inhibit the activities of microorganisms, although the antimicrobial activities of different OvoD peptides vary (Chattopadhyay et al., 2006; Gong et al., 2010; Hervé et al., 2014; Yu et al., 2018). These observations suggest that OvoDs are likely involved in the immune response and that they also function as other host defense peptides (HDPs). Like other defensins, some OvoDs also have evolved to have multiple functions, including lipase and angiotensin I-converting enzyme inhibitory activities (Naknukool et al., 2011).

* Corresponding authors at: Institute of Ecology, Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong 637009, Sichuan Province, China.
*E-mail addresses:* dr.longzhang@vip.163.com, long.zhang@cwnu.edu.cn (L. Zhang), drcqzhou1@163.com (C. Zhou).

https://doi.org/10.1016/j.meegid.2019.01.001
Received 30 July 2018; Received in revised form 28 November 2018; Accepted 2 January 2019
Available online 03 January 2019
1567-1348/ 

To date, > 35 diversified OvoD genes have been discovered in a large range of birds and reptiles (Whenham et al., 2015). However, the entire repertories and the evolutionary relationships within this gene family in different species remain uncertain. On the other hand, it also remains unknown whether positive selection has been involved in driving the evolution of OvoDs. To address these questions, we cross-species screened the genomic sequences of 25 phylogenetically divergent species based on homology searches and discovered 94 OvoD genes in the present study. Furthermore, systematic analyses were conducted to reveal the phylogenetic relationships, genomic organization and selection pressures of this gene family, as well as to predict the net charges of the various OvoD mature peptides. By uncovering multiple newly identified sequences, our results provide insight into the evolution of the OvoD gene family and may help to reveal the relationships of other defensin and defensin-related genes.

## 2. Materials and methods

### 2.1. Sequence identification

A total of 25 phylogenetically divergent species belonging to six different orders (Table 1 and Supplemental file 1) were used to identify the entire repertoire of the OvoD gene family. For each species, a five-step genome search strategy was carried out. (1) All reported OvoD peptide sequences were retrieved from a previous study (Whenham et al., 2015) and aligned by MEGA 6 (Supplemental file 2) (Tamura et al., 2013), and HMMer (V3.1) (Johnson et al., 2010) was used to create a hidden Markov model file from these aligned sequences. (2) The target genome was translated into six open reading frames, and the hidden Markov model file was used as a query to search against the translated amino acid sequences using the hmmsearch with the default settings. (3) All of the positive hits were then checked and verified for the presence of OvoD sequences. (4) For every sequence identified in the previous step, additional iterative TBLASTN (Altschul et al., 1990) searches were conducted until no novel sequence was revealed. (5) The 3000- to 5000-bp genomic sequences up- and downstream of each OvoD motif were retrieved to predict the full-length coding sequences by GeneWise (Birney et al., 2004) and/or GeneScan (Burge and Karlin, 1998). Sequences with a premature stop codon or a frame-shift mutation (insertion or deletion) were considered as pseudogenes. The OvoD genes in the different species were distinguished by a combination of the first letter of the genus and the first three letters of the species. For example, T.gut OvoDA1 refers to gene OvoDA1 of the Zebra finch (*Taeniopygia guttata*). Duplicates of the same subfamily were numbered based on their genomic positions.

### 2.2. Sequence alignment and phylogenetic analyses

All of the deduced OvoD nucleotide sequences were aligned codon-to-codon using the MUSCLE program (Edgar, 2004) with appropriate manual adjustments. To determine the best-fit substitution model for the alignment, jModelTest 2 software (Darriba et al., 2012) was used to calculate the Akaike Information criterion, and the symmetrical (SYM) model + gamma distribution (G) + proportion of invariable sites (I) was chosen for the following Bayesian Evolutionary Analysis. A phylogenetic tree was constructed using the Bayesian Evolutionary Analysis by Sampling Trees (BEAST) software package version 1.8.4 (Drummond and Rambaut, 2007) with the base frequencies of the general time reversible GTR + I + G model set to “All equal” (Zharkikh, 1994). The uncorrelated lognormal relaxed clock model was used to allow for independent substitute rates in each branch, and Yule progress was selected as a tree prior to the Bayesian analysis. Furthermore, a suggested divergence time between birds and reptiles was used as a prior, that is, the priors were set as fitting a normal distribution with the mean and initial value equal to 292 and the standard deviation set to 0.2 (Hedges et al., 2015). Using the above parameters, 100 million generations were performed with the Markov Chain Monte Carlo (MCMC) algorithms, sampling every 5000 steps. The BEAST outputs were analyzed by Tracer version 1.7. With a 10% burn-in, all of the effective sample size (ESS) of the BEAST output were larger than 400. The maximum clade credibility tree was created and visualized using TreeAnnotator and MEGA6 (Tamura et al., 2013), respectively.

### 2.3. Selection pressure analysis

The phylogeny topology was used to estimate the nonsynonymous to synomymous rate ratio (ω = dN/dS) by using the CODEML program from the PAML package version 4.7 (Yang, 2007). In an attempt to test

**Table 1**
Detailed information of the 25 species used in this study.

| Common Name | Scientific name | Abbreviation | Class | Order | Family |
|---|---|---|---|---|---|
| Zebra finch | *Taeniopygia guttata* | T.gut | Aves | Passeriformes | Estrildidae |
| Budgerigar | *Melopsittacus undulatus* | M.und | Aves | Psittaciformes | Psittaculidae |
| Crested ibis | *Nipponia nippon* | N.nip | Aves | Pelecaniformes | Threskiornithidae |
| Chicken | *Gallus gallus* | G.gal | Aves | Galliformes | Phasianidae |
| Ostrich | *Struthio camelus* | S.cam | Aves | Struthioniformes | Struthionidae |
| Painted turtle | *Chrysemys picta* | C.pic | Reptilia | Testudines | Emydidae |
| Chinese soft-shelled turtle | *Pelodiscus sinensis* | P.sin | Reptilia | Testudines | Trionychidae |
| Diamondback terrapin | *Malaclemys terrapin* | M.ter | Reptilia | Testudines | Emydidae |
| Spiny soft-shell turtle | *Apalone spinifera* | A.spi | Reptilia | Testudines | Trionychidae |
| Green sea turtle | *Chelonia mydas* | C.myd | Reptilia | Testudines | Cheloniidae |
| Chinese alligator | *Alligator sinensis* | A.sin | Reptilia | Crocodilia | Alligatoridae |
| Australian saltwater crocodile | *Crocodylus porosus* | C.por | Reptilia | Crocodilia | Crocodylidae |
| American alligator | *Alligator mississippiensis* | A.mis | Reptilia | Crocodilia | Alligatoridae |
| Gharial | *Gavialis gangeticus* | G.gan | Reptilia | Crocodilia | Gavialidae |
| Adder | *Vipera berus* | V.ber | Reptilia | Squamata | Viperidae |
| Mitchell's rattlesnake | *Crotalus pyrrhus* | C.pyr | Reptilia | Squamata | Viperidae |
| Corn snake | *Pantherophis guttatus* | P.gut | Reptilia | Squamata | Colubridae |
| Brown spotted pit viper | *Protobothrops mucrosquamatus* | P.muc | Reptilia | Squamata | Viperidae |
| Burmese python | *Python bivittatus* | P.biv | Reptilia | Squamata | Pythonidae |
| Eastern Garter Snake | *Thamnophis sirtalis* | T.sir | Reptilia | Squamata | Colubridae |
| Timber rattlesnake | *Crotalus horridus* | C.hor | Reptilia | Squamata | Viperidae |
| King cobra | *Ophiophagus hannah* | O.han | Reptilia | Squamata | Elapidae |
| Green Anole | *Anolis carolinensis* | A.car | Reptilia | Squamata | Dactyloidae |
| Central bearded dragon | *Pogona vitticeps* | P.vit | Reptilia | Squamata | Agamidae |
| Japanese gecko | *Gekko japonicus* | G.jap | Reptilia | Squamata | Gekkonidae |

the selective pressure at the amino acid positions, three pairs of site-specific models, including M1a (nearly neutral) against M2a (positive selection), M7 (neutral, β-distribution of $\omega < 1$) against M8 (positive selection, β-distribution of $\omega > 1$) and M8 (positive selection, β-distribution of $\omega > 1$) against M8a (β-distribution of $\omega = 1$), were also implemented in the CODEML program. The likelihood ratio tests (LRTs) were used for comparisons between the three pairs of models, and significant results were determined using $\chi^2$ tests with the corresponding degree of freedom. When the LRT suggested significance, a Bayes empirical Bayes (BEB) procedure was used to estimate the posterior probability of each codon coming from M2a and M8. Sites with a posterior probability > .95 are likely to be under positive selection.

## 3. Results

### 3.1. Genome-wide identification of OvoDs

Excluding the eastern garter snake, central bearded dragon and Japanese gekko, a total of 94 genes, including 92 functional genes and 2 pseudogenes containing either a frame-shift mutation or a premature stop codon, were identified from the remaining 22 species. Within these 22 species, the number of putative OvoDs genes varies from 1 to 22. A multiple sequence alignment of all 94 amino acid sequences shows that all of the OvoDs share a largely conserved six-cysteine motif (Supplemental file 3). In the previous publication, subfamilies of OvoDs were named based on the spacing between the cysteines (Whenham et al., 2015). For example, a gene with the C-X5-C-X3-C-X11-C-X3-C cysteine spacing is assigned to subfamily A (OvoDA), while molecules presenting the C-X3-C-X3-C-X11-C-X4-C motif belong to OvoDB. In the present study, a large number of sequences with variations in the OvoD motifs, but sharing significant similarity with known genes, were identified for the first time; therefore, to avoid confusion, these genes were named based on their phylogenetic relationships to the earlier termed subfamilies, even if their OvoD motifs did not totally match the aforementioned nomenclature. As in the vertebrate β-defensins (Cheng et al., 2015; Patil et al., 2005), sequence comparison of the OvoDs suggested that the signal peptides are highly conserved among different species, whereas the divergent amino acid positions are mainly concentrated in the remaining parts. With the exception of OvoDDα4s in the Chinese and American alligators, all of the functional OvoD genes exhibit six conserved cysteine residues, with an overall consensus sequence of C-X(2–6)-C-X(3–4)-C-X(10-13)-C-X4-CC. Interestingly, some genes in certain orders contain an extra cysteine in the OvoD motif. All of the OvoDβs in crocodilians contain an extra cysteine reside between the first and the second conserved cysteine residues (Supplemental file 3), while every functional OvoD gene in snakes exhibits an extra cysteine between the positions of the second and third conserved cysteine residues.

### 3.2. Phylogenetic analyses of OvoDs

In order to illuminate the evolutionary relationships among the OvoDs, a phylogenetic tree was constructed by using the nucleotide sequences and a Bayesian approach implemented in the BEAST software package (Drummond and Rambaut, 2007). As shown in Fig. 1, the Bayesian inference tree divided the OvoDs into two distinct groups, with one group including avian, crocodilian and chelonian OvoDs, and the other group including the genes of the green anole and snakes. In birds, the topology of the phylogenetic tree shows that there are two types of bird OvoDBs (OvoDBα and OvoDBβ) with a conserved C-X3-C-X3-C-X11-C-X4-CC cysteine sequence motif. Of these, the OvoDBα genes share more homology with the OvoDA genes, whereas the OvoDBβ genes trend to be grouped as a basal primitive branch of the avian OvoD gene clade. These results suggest that the different types of avian OvoDBs are likely to have emerged before the divergence of birds. Similarly, a total of 21 genes from crocodiles form a distinct clade with a posterior possibility of 0.98. Phylogenetic analysis also shows that a large number of the crocodilian OvoDs are not grouped into intra-species clusters, which indicates that a majority of the crocodilian OvoD genes existed before the split of crocodiles. Importantly, OvoDDβ is conserved as one-to-one orthologous in the four crocodile species, and OvoDDα is present in multiple copies in the genomes of the Chinese alligator, American alligator as well as in the Australian saltwater crocodile, but absent in the gharial. Because the crocodilian OvoDDβs present as the outgroup of the OvoDDαs, we implied that OvoDDβ was the ancestor of OvoDDα and that several OvoDDαs had already duplicated from the OvoDDβs before the four species diverged. The clade of turtle OvoDs demonstrates an interesting structure. Three typical turtle OvoDBs (C-X3-C-X3-C-X11-C-X4-C) are clustered as the outgroup of the other turtle OvoDs, whereas the OvoDBs and OvoDCs of the Chinese soft-shelled turtle and the Spiny soft-shell turtle share more homology. On the other hand, the OvoDCs seem to have been duplicated from the OvoDBs before the five turtle species split and were then amplified to produce multiple copies after species divergence. All of the OvoDs of the green anole lizard are clustered into a large clade and are further divided into the OvoDE and the OvoDF subfamilies. In the OvoDE subfamily, the OvoDE1/OvoDE2 branch was likely to have evolved early, and this branch subsequently gave rise to all of the other OvoDEs. The OvoDF clade contains 11 genes that were probably duplicated from a common ancestor. As supported by a posterior possibility of 0.98, the sequences of the ophidian OvoDs forms a distinct clade (OvoDG clade). Of these sequences, that of the Burmese python OvoDG1 shares more homology with the pseudogene of the Brown spotted pit viper (P.mucOvoDG1ps) and forms as a basal primitive branch of OvoDG clade, suggesting that gene duplications of OvoDG may have emerged before the divergence of the Burmese python and the Brown spotted pit viper and that the pseudogenization occurred after their split. Collectively, the phylogenetic analysis suggests that all of the OvoDs in each subfamily share a common ancestor and that multiple duplications evolved through gene duplication and diversification after species divergence.

The phylogenetic analysis performed with the BEAST software also estimated the divergence time of the OvoD genes. With the most recent common ancestor of birds and reptiles set to 292 ± 0.2 million years ago (Hedges et al., 2015), the BEAST results indicated that the OvoD genes in birds, crocodiles, turtles, snakes and the green anole lizard likely emerged approximately within the same time period. Consistent with the divergence time of the corresponding species, almost all of the species-specific gene duplication events were inferred to have occurred within 10 million years.

### 3.3. Analysis of the genomic structure and localization of the OvoD genes

The structural organization of each OvoD genes was obtained by comparing their putative coding sequences with the genomic sequences. Every intact OvoD gene predicted in the present study consists of two exons separated by a phase-1 intron. Similar to the α-defensins (Lynn and Bradley, 2007; Patil et al., 2004), the first exon is mainly correlated with the signal peptide, while the second exon is approximately equivalent to the mature peptide and other remaining regions (Supplemental files 1 and 3).

To evaluate the genomic organization of the OvoD gene family, we employed a comparative analysis of the OvoD genes against gene markers located in the adjacent regions. Like the avian defensins and cathelicidins (Cheng et al., 2015), OvoD genes ranging from those of snakes to birds are also apparently clustered in a conserved region. The genes are tightly located in a single syntenic chromosomal region between the XKR6 and MTMR9 genes (Fig. 2 and Supplemental file 1), indicating that their locations have been conserved during the evolutionary process. By also considering the topology of the phylogenetic tree, we found that the gene order and transcriptional orientation of the OvoD orthologous are largely conserved in several species. Furthermore, some genes, such as OvoDF2/OvoDF3 and OvoDF9/OvoDF10

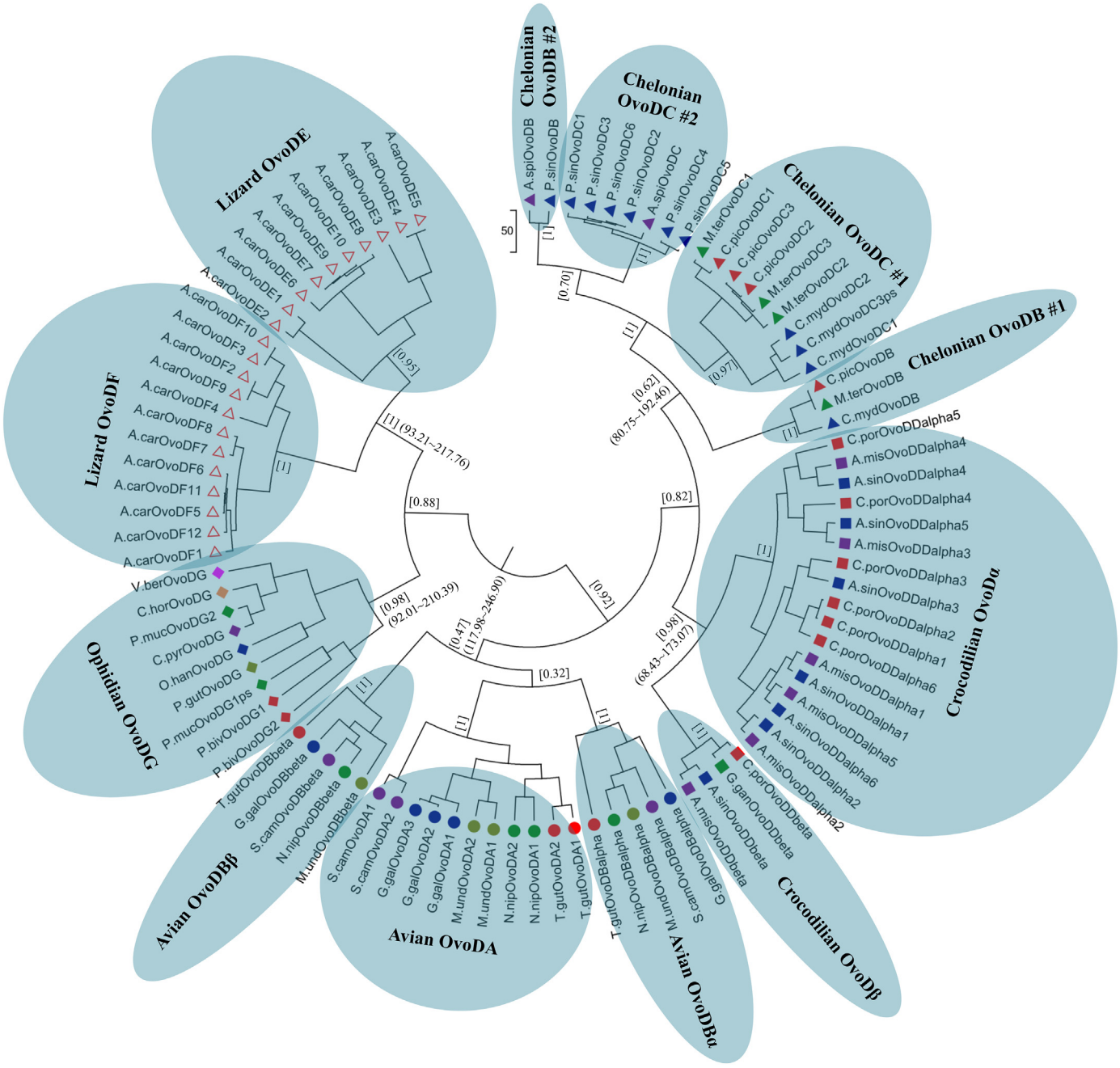


**Fig. 1.** Phylogenetic relationships of the 94 OvoD genes from 23 species. The tree was constructed by using the Bayesian inference method with the nucleotide sequences. The numbers in square brackets indicate the Bayesian posterior probabilities and the numbers in round brackets represent 95% confidence intervals for the divergence times (million years ago).

from the green anole lizard, might have originated from tandem duplication events. The remarkable genomic structure and location conservation of the OvoD gene family strongly suggest that all of the OvoD genes are evolved from a single gene and that rapid gene duplications occurred independently following species divergence.

### 3.4. Net charge estimation of the mature OvoD peptides

To determine the cationicity of the OvoDs, we predicted the net charges of their deduced mature peptides. The predicted net charges for the mature OvoD peptides in each subfamily are shown in Fig. 3. The net charges of the mature peptides ranged from −1.2 to 9.8, with an average net charge within each subfamily higher than two, indicating that the cationicity of OvoDs may be essential for maintaining their biological functions. The highest average net charge was observed in the bird OvoDBβ cluster, whereas the lowest average value was observed in the lizard OvoDE lineage. Furthermore, we noticed that the net charge was highly divergent in certain subfamilies, but relatively stable in some specific groups, such as the bird OvoDA and crocodile OvoDDβ subsets. Because positively charged amino acid residues are important for establishing tight interactions between HDPs and target microorganisms (Brogden, 2005), these cationic OvoDs likely possess strong antimicrobial activity.

### 3.5. Positive selection in the different OvoD clusters

Because our evolutionary analyses suggested that all of the OvoD genes obtained in the present study evolved after the divergence of the

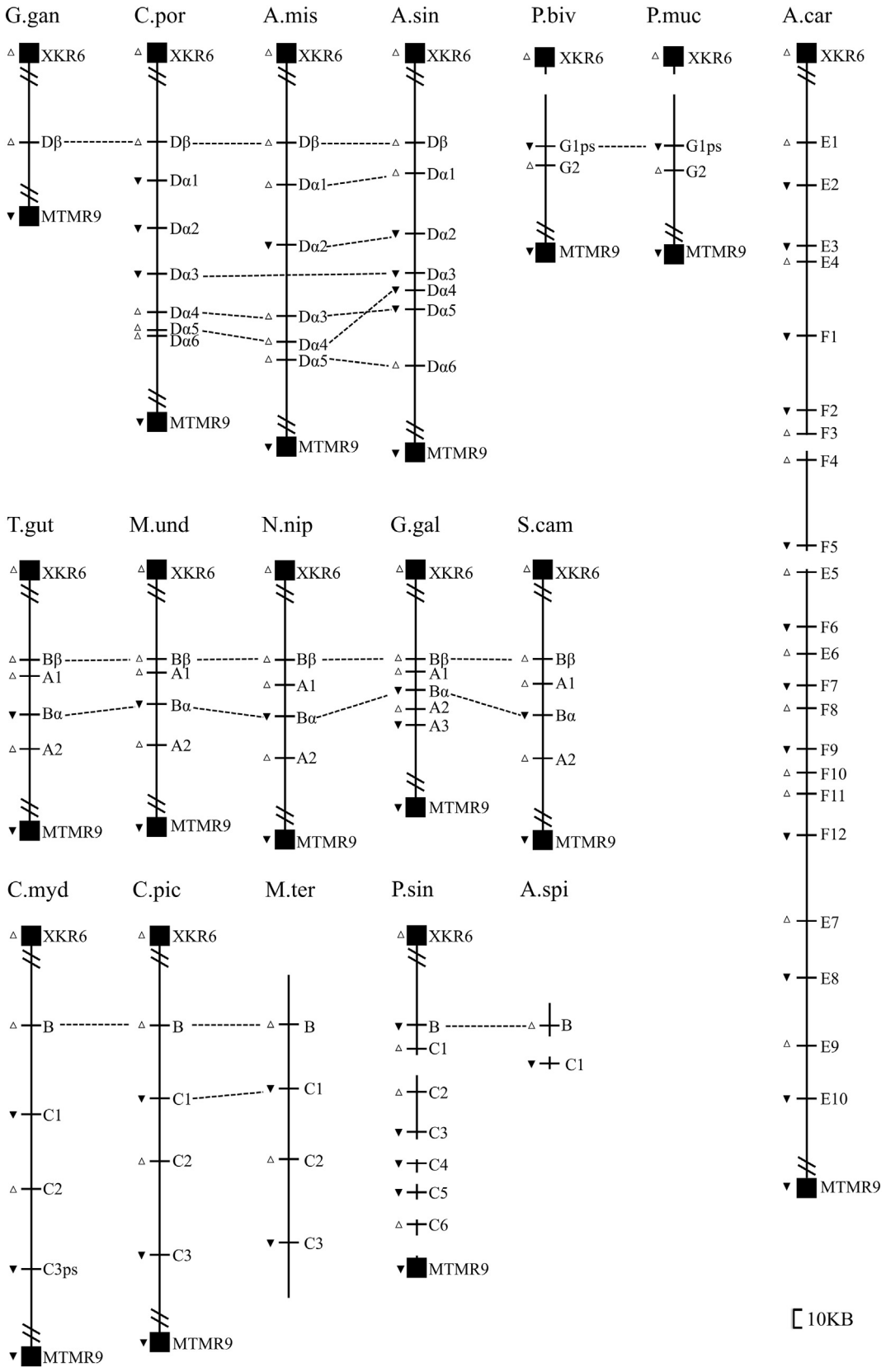


**Fig. 2.** Genomic organization of the OvoD gene clusters in the different species. The location of each gene is represented by the position of the six-cysteine motif, and the orientation of transcription is indicated by a triangle. The genes with solid triangles are transcribed opposite to the ones with open triangles. A broken line in a gene cluster is an indication of a gap in the genomic DNA sequence. Sequences outside the OvoD gene cluster in each species are omitted by slanted lines. The orthologous OvoDs are connected with dashed lines. Only species that contains at least two OvoD genes are shown.

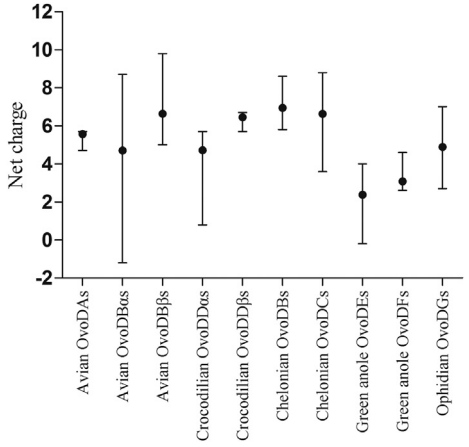


**Fig. 3.** Net charge comparison of the putative OvoD mature peptides. The mature peptides of the OvoD genes were predicted based on the amino acid alignment in a previous study (Whenham et al., 2015). The net charge of each peptide was estimated by using the on-line software at https://pepcalc.com/, and the average net charge and net charge ranges are shown by lineage.

different orders or classes, we attempted to identify positively selected sites in birds, crocodiles, turtles, the green lizard and snakes, as well as for the entire OvoD family. In the avian and crocodilian OvoD clades, three pairs of LRTs yielded consistent results suggesting that no significant positive selection had acted on the OvoDs within these linages (Supplemental file 4). In contrast, multiple sites seemed to be under strong positive pressure in turtles, green lizard and snakes.

By mapping the deduced positively selected sites to the sequence logo, we found that the positively selected sites were primarily located in the mature peptide (Fig. 4). This pattern is similar to findings in many other HDPs, and the high frequency amino acid substitutions that occurred in the mature peptides might help the immune system to cope with rapid pathogen evolution (Hughes, 1999; Semple et al., 2003; Tu et al., 2015). The positive selection analysis of the entire OvoD gene family further showed that the dN/dS radio in each model was lower than one, indicating that the whole gene family is under weak purifying selection (natural selection that prevents the fixation of deleterious alleles). However, comparing the M7 and M8 site models showed that two amino acids located in the mature peptide are under positive selection in the whole OvoD gene family.

## 4. Discussion

Following our genome-wide screening performed using a combination of hidden Markov models and the BLAST program, the entire OvoD gene family repertoire, including multiple novel genes, has been identified in 22 species. On account of low genome quality in certain regions, we failed to obtain the signal peptide sequences of some genes. Because some defensin genes, such as AvBD11 (Xiao et al., 2004) and AcBD14 (Dalla Valle et al., 2012), contain two tandem repeats of the six-cysteine motif, we cannot exclude the possibility that some OvoD genes might form a similar pattern. Unlike the typical vertebrate defensins, which are characterized by the presence of a conserved six-cysteine motif in the mature peptide, two groups of peptides, namely the crocodile OvoDDβs and snake OvoDGs, were found to encode mature peptides with seven cysteine residues. Based on the observation that chicken OvoDA and marine turtle OvoDB exhibit different disulfide bridging patterns (Chattopadhyay et al., 2006; Hervé et al., 2014), we suggest that OvoDs can have multiple disulfide bonding patterns, albeit the six cysteine residues are highly conserved. Although lack of a cysteine residue may disrupt the disulfide bonding patterns of the OvoDDα4s in the Chinese and American alligators, these peptides with different disulfide bonding patterns may also have similar antimicrobial activities and biological functions like rattusin (Patil et al., 2013) and Defr1(Taylor et al., 2009). With a large number of newly identified sequences, the detailed phylogenetic topology presented in this study indicates that all of the OvoDs have evolved separately within each order via gene duplication followed by divergence. Consistent with a previous report, bird and turtle OvoDBs are classified into different cluster even though they contain the same OvoD motif (Whenham et al., 2015; Yu et al., 2018). Furthermore, our genomic location analysis shows that the OvoDs are closely linked in a syntenic region flanked by XKR6 and MTMR9. These observations collectively suggest that the OvoD genes in each species share a common ancestor and arose via independent duplications events after speciation. Despite the lack of tissue expression patterns in many other species, such tightly linked structures in the genome are very likely to lead to their synergistic expression in the oviduct, at least in the case of the avian OvoDs (Gong et al., 2010; Whenham et al., 2015). In addition to the OvoD gene family, multiple defensin-related genes are also located in the region franked by the XKR6 and MTMR9 genes. The current gene annotation in the NCBI database shows that at least two AvBD13-like genes, LOC102929355 and LOC102450934, are also located in a region

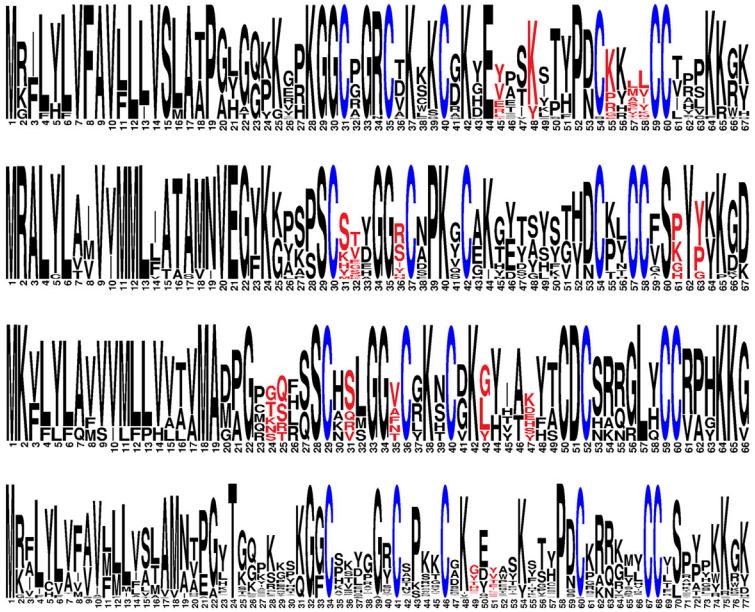


**Fig. 4.** Sequence logo and positively selected sites in the different OvoD clusters. Sequence logos were generated via the WEBLOGO (http://weblogo.berkeley.edu/logo.cgi), and the six conversed cysteines and the positive selection sites (M7 vs. M8 model) are highlighted in blue and red, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

franked by XKR6 and MTMR9 in the green sea turtle and the Chinese soft-shelled turtle, respectively. Importantly, the sequence analysis applied in Pfam (Finn et al., 2013) suggests that this type of AvBD13-like gene is more similar to an invertebrate big defensin than a vertebrate β-defensin (data not shown). Current understandings suggest that vertebrate α-, β-, θ- and invertebrate big defensins have evolved from a single ancestor (Shafee et al., 2016), while vertebrate β-defensins emerged from invertebrate big-defensins through exon shuffling or intronization (Zhu and Gao, 2013). Due to the similarities in tertiary structure, genomic organization as well as intron phase between OvoDs and big defensins (Hervé et al., 2014; Zhu, 2008), OvoDs also likely originated from big defensins. Alternatively, OvoDs might have evolved from a specific β-defensin gene, as β-defensins are pervasive in vertebrates (Hollox and Abujaber, 2017) and OvoDs have only been found in birds and reptiles. Additionally, our homolog screening also obtained some sequences containing either a myotoxin or reptilian defensin motif from the genome of snakes and the Japanese gekko (data not shown), indicating that OvoDs and some reptilian defensin-related genes likely share a common ancestor. To draw more conclusive evolutionary scenario for the OvoDs, more defensin-related genes are needed to be studied in the future.

Consistent with observations in avian β-defensins (Cheng et al., 2015), positive selection was detected in the mature peptides of multiple, but not all OvoD clades. Given that amino acid substitutions at positively selected sites can result in significant modification of antimicrobial activities, the high frequency of positive selection sites occurred in the mature peptides may be drove by pathogen mediated selection (Hughes, 1999). Gene duplication followed by positive selection plays a major role in the evolution of new biological functions (Zhang, 2003). As a consequence, some genes identified in the present study might not be solely expressed in the albumen-producing region of the oviduct or may have evolved to perform functions other than antimicrobial defense, like the platypus defensin-like peptides (Whittington et al., 2008b) and the snake crotamines (Hargreaves et al., 2014). Conversely, the low amino acid divergence seen in the bird and crocodile OvoDs is probably due to the possibility that those genes were selected and maintained as the most efficacious for each species (Chapman et al., 2016). Interestingly, we observed that all of the bird OvoDAs in each species are likely to have evolved in concert as a unit. The genomic organizations of the gene order are highly conserved as OvoDBβ, OvoDA1, OvoDBα and OvoDA2 from ostrich to zebra finch, strongly suggesting that these genes already existed in the ancestral bird genome. However, the species-specific OvoDAs clusters generated by the Bayesian inference phylogenetic analysis implied that the paralogous OvoDAs likely evolved after the divergence of birds. Based on the extremely high conversation in the amino acid sequences of the OvoDA lineage, we speculated that the OvoDA genes might be under strong purifying selection to prevent functional divergence. Gene duplication sometimes generates two identical genes with the same function, and these duplicates can only be stably maintained in the genome if the increase in the abundance of the gene product is beneficial (Zhang, 2003). As one of the most abundant protein groups present in the egg white (Gong et al., 2010; Sun et al., 2017), higher levels of the protein or RNA products of the avian OvoDs are probably in high demand. In this regard, the extra copies of the OvoDA genes are stably maintained in the genome to generate higher levels of their gene products. Furthermore, in light of the fact that only chicken OvoDA2 and OvoDA3 are expressed in the isthmus (Gong et al., 2010), we cannot exclude the possibility that subfunctionaliztion has occurred at the gene expression level of the OvoDA genes. However, the reason why the number of OvoD genes is so variable among the different species remains obscure. It would be particularly interesting to unravel the possible mechanism that drives the duplication and divergence of the OvoD gene family.

## 5. Conclusions

A total of 94 OvoD genes from snakes to birds were obtained by cross-genome analysis in our study. Phylogenetic and genome organization analyses together indicate that the OvoDs in different species share a common ancestor and that a series of independent duplication events likely occurred in multiple species. Importantly, our evolutionary analyses further suggest that different selective pressures may have preferentially acted on specific lineages. Additional evolutionary analysis and functional characterization of the OvoD genes, may lead to much better understanding of this gene family.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.01.001.

## Competing interests

The authors declare that there are no competing interests.

## Authors' contributions

LZ and CZ conceived and designed the experiments. LZ, DC, LY, YW, JL and CZ analyzed the data. LZ wrote the paper. All authors reviewed and approved the manuscript.

## Acknowledgements

This research was supported by the Science and Technology Department of Sichuan Province (Grant NO. 2015JY0249), the Doctoral Scientific Research Funds (Grant NO. 17E077) and the Fundamental Research Funds of China West Normal University (Grant NO. CXTD2017-4 and 18B030).

## References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Birney, E., Clamp, M., Durbin, R., 2004. GeneWise and Genomewise. Genome Res. 14, 988–995.
Brogden, K.A., 2005. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat. Rev. Microbiol. 3, 238.
Burge, C.B., Karlin, S., 1998. Finding the genes in genomic DNA. Curr. Opin. Struct. Biol. 8, 346–354.
Candille, S.I., Kaelin, C.B., Cattanach, B.M., Yu, B., Thompson, D.A., Nix, M.A., Kerns, J.A., Schmutz, S.M., Millhauser, G.L., Barsh, G.S., 2007. A β-defensin mutation causes black coat color in domestic dogs. Science 318, 1418–1423.
Chapman, J.R., Hellgren, O., Helin, A.S., Kraus, R.H., Cromie, R.L., Waldenström, J., 2016. The evolution of innate immune genes: purifying and balancing selection on β-defensins in waterfowl. Mol. Biol. Evol. 33, 3075–3087.
Chattopadhyay, S., Sinha, N.K., Banerjee, S., Roy, D., Chattopadhyay, D., Roy, S., 2006. Small cationic protein from a marine turtle has β-defensin-like fold and antibacterial and antiviral activity. Proteins: Structure, Function, and Bioinformatics 64, 524–531.
Cheng, Y., Prickett, M.D., Gutowska, W., Kuo, R., Belov, K., Burt, D.W., 2015. Evolution of the avian β-defensin and cathelicidin genes. BMC Evol. Biol. 15, 188.
Dalla Valle, L., Benato, F., Maistro, S., Quinzani, S., Alibardi, L., 2012. Bioinformatic and molecular characterization of beta-defensins-like peptides isolated from the green lizard Anolis carolinensis. Developmental & Comparative Immunology 36, 222–229.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., 2014. Pfam: the protein families database. Nucleic Acids Res. 42, 222–230.
Ganz, T., 2004. Defensins: antimicrobial peptides of vertebrates. C. R. Biol. 327, 539–549.
Gong, D., Wilson, P.W., Bain, M.M., McDade, K., Kalina, J., Hervé-Grépinet, V., Nys, Y., Dunn, I.C., 2010. Gallin; an antimicrobial peptide member of a new avian defensin family, the ovodefensins, has been subject to recent gene duplication. BMC Immunol. 11, 12.
Hancock, R.E., Nijnik, A., Philpott, D.J., 2012. Modulating immunity as a therapy for bacterial infections. Nat. Rev. Microbiol. 10, 243.
Hargreaves, A.D., Swain, M.T., Hegarty, M.J., Logan, D.W., Mulley, J.F., 2014. Restriction and recruitment—gene duplication and the origin and evolution of snake venom toxins. Genome Biol. Evol. 6, 2088–2095.

Hedges, S.B., Marin, J., Suleski, M., Paymer, M., Kumar, S., 2015. Tree of life reveals clock-like speciation and diversification. Mol. Biol. Evol. 32, 835–845.
Hervé, V., Meudal, H., Labas, V., Réhault-Godbert, S., Gautron, J., Berges, M., Guyot, N., Delmas, A.F., Nys, Y., Landon, C., 2014. Three-dimensional NMR structure of hen egg gallin (chicken ovodefensin) reveals a new variation of the β-defensin fold. J. Biol. Chem. 289, 7211–7220.
Hilchie, A.L., Wuerth, K., Hancock, R.E., 2013. Immune modulation by multifaceted cationic host defense (antimicrobial) peptides. Nat. Chem. Biol. 9, 761.
Hollox, E.J., Abujaber, R., 2017. Evolution and Diversity of Defensins in Vertebrates, Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts. Springer, pp. 27–50.
Hughes, A., 1999. Evolutionary diversification of the mammalian defensins. Cell. Mol. Life Sci. CMLS 56, 94–103.
Johnson, L.S., Eddy, S.R., Portugaly, E., 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinforma. 11, 431.
Li, D., Zhang, L., Yin, H., Xu, H., Trask, J.S., Smith, D.G., Li, Y., Yang, M., Zhu, Q., 2014. Evolution of primate α and θ defensins revealed by analysis of genomes. Mol. Biol. Rep. 41, 3859–3866.
Lynn, D.J., Bradley, D.G., 2007. Discovery of α-defensins in basal mammals. Dev. Comp. Immunol. 31, 963–967.
Mansour, S.C., Pena, O.M., Hancock, R.E., 2014. Host defense peptides: front-line immunomodulators. Trends Immunol. 35, 443–450.
Naknukool, S., Hayakawa, S., Ogawa, M., 2011. Multiple biological functions of novel basic proteins isolated from duck egg white: duck basic protein small 1 (dBPS1) and 2 (dBPS2). J. Agric. Food Chem. 59, 5081–5086.
Patil, A., Hughes, A.L., Zhang, G., 2004. Rapid evolution and diversification of mammalian α-defensins as revealed by comparative analysis of rodent and primate genes. Physiol. Genomics 20, 1–11.
Patil, A.A., Cai, Y., Sang, Y., Blecha, F., Zhang, G., 2005. Cross-species analysis of the mammalian β-defensin gene family: presence of syntenic gene clusters and preferential expression in the male reproductive tract. Physiol. Genomics 23, 5–17.
Patil, A.A., Ouellette, A.J., Lu, W., Zhang, G., 2013. Rattusin, an intestinal α-defensin-related peptide in rats with a unique cysteine spacing pattern and salt-insensitive antibacterial activities. Antimicrob. Agents Chemother., 54, 1823–1831.
Rádis-Baptista, G., Kubo, T., Oguiura, N., da Silva, A.P., Hayashi, M., Oliveira, E., Yamane, T., 2004. Identification of crotasin, a crotamine-related gene of Crotalus durissus terrificus. Toxicon 43, 751–759.
Semple, C.A., Rolfe, M., Dorin, J.R., 2003. Duplication and selection in the evolution of primate β-defensin genes. Genome Biol. 4, R31.
Shafee, T.M., Lay, F.T., Hulett, M.D., Anderson, M.A., 2016. The defensins consist of two independent, convergent protein superfamilies. Mol. Biol. Evol. 33, 2345–2356.
Sun, C., Liu, J., Li, W., Xu, G., Yang, N., 2017. Divergent proteome patterns of egg albumen from domestic chicken, duck, goose, turkey, quail and pigeon. Proteomics. 17 1700145.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution. Vol. 30. pp. 2725–2729.
Taylor, K., Rolfe, M., Reynolds, N., Kilanowski, F., Pathania, U., Clarke, D., Yang, D., Oppenheim, J., Samuel, K., Howie, S., 2009. Defensin-related peptide 1 (Defr1) is allelic to Defb8 and chemoattracts immature DC and CD4+ T cells independently of CCR6. Eur. J. Immunol. 39, 1353–1360.
Thomma, B.P., Cammue, B.P., Thevissen, K., 2002. Planta 216, 193–202 Plant defensins.
Tu, J., Li, D., Li, Q., Zhang, L., Zhu, Q., Gaur, U., Fan, X., Xu, H., Yao, Y., Zhao, X., 2015. Molecular evolutionary analysis of β-defensin peptides in vertebrates. Evol. Bioinforma. 11 (EBO. S25580).
Whenham, N., Lu, T.C., Maidin, M.B., Wilson, P.W., Bain, M.M., Stevenson, M.L., Stevens, M.P., Bedford, M.R., Dunn, I.C., 2015. Ovodefensins, an oviduct-specific antimicrobial gene family, have evolved in birds and reptiles to protect the egg by both sequence and intra-six-cysteine sequence motif spacing. Biol. Reprod. 92, 154.
Whittington, C.M., Papenfuss, A.T., Bansal, P., Torres, A.M., Wong, E.S., Deakin, J.E., Graves, T., Alsop, A., Schatzkamer, K., Kremitzki, C., 2008a. Defensins and the Convergent Evolution of Platypus and Reptile Venom Genes. (Genome research).
Whittington, C.M., Papenfuss, A.T., Kuchel, P.W., Belov, K., 2008b. Expression patterns of platypus defensin and related venom genes across a range of tissue types reveal the possibility of broader functions for OvDLPs than previously suspected. Toxicon 52, 559–565.
Xiao, Y., Hughes, A.L., Ando, J., Matsuda, Y., Cheng, J.-F., Skinner-Noble, D., Zhang, G., 2004. A genome-wide screen identifies a single β-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. BMC Genomics 5, 56.
Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.
Yu, L.-t., Xiao, Y.-P., Li, J.-J., Ran, J.-S., Yin, L.-Q., Liu, Y.-P., Zhang, L., 2018. Molecular characterization of a novel ovodefensin gene in chickens. Gene 678, 233–240.
Zasloff, M., 2002. Antimicrobial peptides of multicellular organisms. Nature 415, 389–395.
Zhang, J., 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18, 292–298.
Zharkikh, A., 1994. Estimation of evolutionary distances between nucleotide sequences. J. Mol. Evol. 39, 315–329.
Zhou, Y.S., Webb, S., Lettice, L., Tardif, S., Kilanowski, F., Tyrrell, C., MacPherson, H., Semple, F., Tennant, P., Baker, T., 2013. Partial deletion of chromosome 8 β-defensin cluster confers sperm dysfunction and infertility in male mice. PLoS Genet. 9, e1003826.
Zhu, S., 2008. Did cathelicidins, a family of multifunctional host-defense peptides, arise from a cysteine protease inhibitor? Trends Microbiol. 16, 353–360.
Zhu, S., Gao, B., 2013. Evolutionary origin of β-defensins. Dev. Comp. Immunol. 39, 79–84.